Functional Lower Bounds for Arithmetic Circuits and Connections to Boolean Circuit Complexity

Michael A. Forbes[*1], **Mrinal Kumar**[†2], **and Ramprasad Saptharishi**[‡3]

1 **Department of Computer Science, Princeton University, USA**
miforbes@csail.mit.edu
2 **Department of Computer Science, Rutgers University, USA**
mrinal.kumar@rutgers.edu
3 **Tel Aviv University, Israel**
ramprasad@cmi.ac.in

Abstract

We say that a circuit C over a field $\mathbb{F}$ *functionally* computes a polynomial $P \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ if for every $\mathbf{x} \in \{0,1\}^n$ we have that $C(x) = P(x)$. This is in contrast to *syntactically* computing P, when $C \equiv P$ as formal polynomials. In this paper, we study the question of proving lower bounds for homogeneous depth-3 and depth-4 arithmetic circuits for functional computation. We prove the following results :

- Exponential lower bounds for homogeneous depth-3 arithmetic circuits for a polynomial in VNP.
- Exponential lower bounds for homogeneous depth-4 arithmetic circuits with bounded individual degree for a polynomial in VNP.

Our main motivation for this line of research comes from our observation that strong enough functional lower bounds for even very special depth-4 arithmetic circuits for the Permanent imply a separation between #P and ACC^0. Thus, improving the second result to get rid of the *bounded individual degree* condition could lead to substantial progress in boolean circuit complexity. Besides, it is known from a recent result of Kumar and Saptharishi [9] that over constant sized finite fields, strong enough *average case* functional lower bounds for homogeneous depth-4 circuits imply superpolynomial lower bounds for homogeneous depth-5 circuits.

Our proofs are based on a family of new complexity measures called *shifted evaluation dimension*, and might be of independent interest.

1998 ACM Subject Classification F.2.1 Numerical Algorithms and Problems, I.1.1 Expressions and Their Representation

Keywords and phrases boolean circuits, arithmetic circuits, lower bounds, functional computation

Digital Object Identifier 10.4230/LIPIcs.CCC.2016.33

1 Introduction

Arithmetic circuits are one of the most natural models of computation for studying computation with multivariate polynomials. One of the most fundamental questions in this area

* Research supported by the Princeton Center for Theoretical Computer Science.
† Research supported in part by the Simons Graduate Fellowship.
‡ The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 257575.

31st Conference on Computational Complexity (CCC 2016).
Editor: Ran Raz; Article No. 33; pp. 33:1–33:19

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

of research is to show that there are low degree polynomials which cannot be efficiently computed by *small sized* arithmetic circuits. However, in spite of the significance of this question, progress on it has been sparse and our current state of understanding of lower bounds for arithmetic circuits continues to remain extremely modest.

Most of the research in algebraic complexity theory so far considers arithmetic circuits and multivariate polynomials as *formal* objects and studies the complexity of *syntactic* representation of polynomials over the underlying field. However, in this work, we aim to study the *semantic* or *functional* analogue of the complexity of computing multivariate polynomials. We formally define this notion below and then try to motivate the definition based on our potential applications.

▶ **Definition 1.1** (Functional equivalence)**.** Let $\mathbb{F}$ be any field and let D be a subset of $\mathbb{F}$. We say that two n-variate polynomials P_1 and P_2 in $\mathbb{F}[x_1, x_2, \ldots, x_n]$ are *functionally* equivalent over the domain D^n if

$$\forall \mathbf{x} \in D^n \quad , \quad P_1(\mathbf{x}) = P_2(\mathbf{x}) \, .$$

This definition of functional equivalence naturally extends to the case of arithmetic circuits functionally computing a family of polynomials, as defined below.

▶ **Definition 1.2** (Functional computation)**.** Let $\mathbb{F}$ be any field and let D be a subset of $\mathbb{F}$. A circuit family $\{C_n\}$ is said to functionally compute a family of polynomials $\{P_n\}$ over the domain D^n if

$$\forall n \in \mathbb{N}, \mathbf{x} \in D^n \quad , \quad C_n(\mathbf{x}) = P_n(\mathbf{x}) \, .$$

Having defined functional computation, we will now try to motivate the problem of proving functional lower bounds for arithmetic circuits.

1.1 Motivation

Improved boolean circuit lower bounds

In the late 80s there was some spectacular progress on the question of lower bounds for bounded depth boolean circuits. In particular, Razborov and Smolensky [16, 15] showed exponential lower bounds for constant depth boolean circuits with AND ($\wedge$), OR ($\vee$), Negations ($\neg$) and $\mod p$ gates for a prime p (i.e the class of $\mathsf{AC}^0[\mathsf{p}]$ circuits). However, the question of proving lower bounds for constant depth boolean circuits which also have $\mod q$ gates for a composite q (i.e the class of general ACC^0 circuits) remained wide open. In general, one major obstacle was that the techniques of Razborov and Smolensky failed for composite moduli, and we could not find alternative techniques which were effective for the problem. Although it is widely believed that the majority function should be hard for such circuits, till a few years ago, we did not even know to show that there is such a language in NEXP (the class of problems in nondeterministic exponential time). In a major breakthrough on this question, Williams [17] showed that there is a function in NEXP which requires ACC^0 circuits of superpolynomial size. Along with the result itself, the paper introduced a new proof strategy for showing such lower bounds. However, it still remains wide open to show that there is a function in deterministic exponential time, which requires ACC^0 circuits of superpolynomial size.

One of our main motivations for studying functional lower bounds for arithmetic circuits is the following lemma which shows that such lower bounds in fairly modest set up would imply a separation between $\#\mathsf{P}$ and ACC^0. A formal statement and a simple proof can be found in section 3.

▶ **Lemma 1.3** (Informal). *Let $\mathbb{F}$ be any field of characteristic zero or at least $\exp(\omega(\text{poly}(\log n)))$. Then, a functional lower bound of $\exp(\omega(\text{poly}(\log n)))$ for the permanent of an $n \times n$ matrix over $\{0,1\}^{n^2}$ for depth-4 arithmetic circuits with bottom fan-in $\text{poly}(\log n)$ imply that $\#\mathsf{P} \neq \mathsf{ACC}^0$.*

In fact, we show that something slightly stronger is true. It suffices to prove functional lower bounds for the model of sums of powers of low degree polynomials for the conclusion in Theorem 1.3 to hold.

At this point, there are two possible interpretations of the statement of Theorem 1.3. For an optimist, it provides another approach to proving new lower bounds for ACC^0, while for a pessimist it points to the fact that the functional lower bounds for depth-4 arithmetic circuits could be possibly very challenging. What makes us somewhat optimistic about this strategy is the fact that in the last few years, we seem to have made substantial progress on the question of proving lower bounds for homogeneous depth-4 circuits in the syntactic setting [6, 4, 7, 10]. In particular, even though the depth-4 circuits obtained in the proof of Theorem 1.3 are not homogeneous, an exponential lower bound for sums of powers of low degree polynomials is known in the syntactic set up. Therefore, it makes sense to try and understand if these bounds can be extended to the functional set up as well.

Lower bounds for homogeneous depth-5 circuits

In a recent work by Kumar and Saptharishi [9], it was shown that over constant size finite fields, *average case functional* lower bounds for homogeneous depth-4 circuits implies lower bounds for homogeneous depth-5 circuits[1]. More precisely, the following lemma was shown:

▶ **Lemma 1.4** ([9]). *Let $\mathbb{F}_q$ be a finite field such that $q = O(1)$. Let P be a homogeneous polynomial of degree d in n variables over $\mathbb{F}_q$, which can be computed by a homogeneous depth-5 circuit of size at most $O\left(\exp\left(d^{0.499}\right)\right)$. Then, there exists a homogeneous depth-4 circuit C' of bottom fan-in $O(\sqrt{d})$ and top fan-in at most $O\left(\exp\left(d^{0.499}\right)\right)$ such that*

$$\Pr_{x \in \mathbb{F}_q^n}[P(x) \neq C'(x)] \leq \exp(-\Omega(\sqrt{d}))\,.$$

Informally, the lemma shows that over small finite fields strong enough *average case* functional lower bounds for homogeneous depth-4 arithmetic circuit with bounded bottom fan-in are sufficient to show superpolynomial lower bounds for homogeneous depth-5 circuits. Even though in [9], the authors do not take this route to eventually prove their lower bounds, this connection seems like a strong motivation to study the question of proving functional lower bounds for bounded depth arithmetic circuits.

Functional lower bounds for bounded depth arithmetic circuits

It is immediately clear from the definition that *syntactic* computation implies *functional* computation, but vice-versa may not be necessarily true. In this sense, proving lower bounds for functional computation could be potentially harder than proving lower bounds for syntactic computation. From this point of view, once we have syntactic lower bounds for a certain class of circuits, it seems natural to ask if these bounds can be extended to the functional framework as well. The last few years have witnessed substantial progress on the

[1] In fact, such lower bounds for homogeneous depth-4 circuits with bounded bottom fan-in suffice for this application.

question of proving lower bounds for variants of depth-4 arithmetic circuits, and in this work we explore the question of whether these bounds can be extended to the functional setting.

Applications to proof complexity lower bounds

Functional lower bounds have recently found applications for obtaining lower bounds for algebraic proof systems. In particular, Forbes, Shpilka, Tzameret, and Wigderson [3] have given lower bounds in various algebraic circuit measures for any polynomial agreeing with certain functions of the form $\mathbf{x} \mapsto \frac{1}{p(\mathbf{x})}$, where p is a constant-degree polynomial (which is non-zero on the boolean cube). In particular, they used such lower bounds to obtain lower bounds for the various subclasses of the Ideal Proof System (IPS) of Grochow and Pitassi [5].

In the next section, we explore the connections between syntactic and functional computation in a bit more detail, and discuss why the techniques used in proving syntactic lower bounds do not seem conducive to prove lower bounds in the functional setting. Hence, the problem of proving functional lower bounds might lead us to more techniques for arithmetic circuit lower bounds.

1.2 Functional vs syntactic computation

We now discuss the differences and similarities between functional and syntactic computation in a bit more detail. The following observation is easy to see.

▶ Observation 1.5. The following properties follow from Theorem 1.2:

- Any two polynomials P_1 and P_2 which are syntactically equivalent are also functionally equivalent for every choice of D.
- If two polynomials of individual degrees bounded by d are functionally equivalent over any domain of size at least $d+1$, then they are also syntactically equivalent.
- In particular, any two multilinear polynomials which are functionally equivalent over the hypercube $\{0,1\}^n$ are also syntactically equivalent.

For the rest of the paper, our domain of interest will be $D = \{0,1\}$ and we will be interested in polynomials which are functionally the same over the hypercube $\{0,1\}^n$. For brevity, for the rest of the paper, when we say that two polynomials are functionally equivalent, we mean that the domain is the hypercube. As an additional abuse of notation, when we say that a circuit C is functionally equivalent to a polynomial P, we mean that for every $\mathbf{x} \in \{0,1\}^n$, $C(\mathbf{x}) = P(\mathbf{x})$. Observe that functional equivalence over the hypercube is precisely the same as syntactic equivalence when we work modulo the ideal generated by the polynomials $\{x_i^2 - x_i : i \in [n]\}$. However, we find the functional view easier and more convenient to work with.

At this point, one might ask why is the choice of D as $\{0,1\}$ a natural one? The motivation for studying a domain of size 2 stems from the fact that most of the polynomials for which we have syntactic arithmetic circuit lower bounds, are multilinear. For instance, the permanent (Perm), the determinant (Det), the Nisan-Wigderson polynomials (NW) and the iterated matrix multiplication polynomial (IMM) are known to be hard for many natural classes of arithmetic circuits, homogeneous depth three circuits being one such class. Since for any $D \subseteq \mathbb{F}$ such that $|D| \geq 2$, D^n is an interpolating set for multilinear polynomials, it seems natural to ask if there is a small homogeneous depth three arithmetic circuit which is functionally equivalent to any of these polynomials.

Another reason why $\{0,1\}^n$ seems a natural domain to study functional algebraic computation is due to potential connections to boolean circuit lower bounds. It seems natural to ask if the techniques discovered in the quest for arithmetic circuit lower bounds can be

adapted to say something interesting about questions in boolean circuit complexity. And, Theorem 1.3 seems like an encouraging step in this direction.

1.2.1 Functional lower bounds and partial derivatives

Almost all the bounded depth arithmetic circuit lower bounds so far have been proved using techniques based on the partial derivatives of a polynomial. This includes exponential lower bounds for homogeneous depth-3 circuits [11] and lower bounds for homogeneous depth-4 arithmetic circuits [6, 4, 7, 10]. At a high level, the proofs have the following structure:

- Define a function $\Gamma : \mathbb{F}[\mathbf{x}] \to \mathbb{N}$, called the complexity measure, which serves as an indicator of the hardness of a polynomial.
- For all *small* arithmetic circuits in the model of interest, show that Γ has a non-trivial upper bound.
- For the target hard polynomial, show that Γ is large. Comparing this with the upper bound in step 2 leads to a contradiction if the hard polynomial had a small arithmetic circuit.

The precise measure Γ used in these proofs varies, but they all build upon the the notion of partial derivatives of a polynomial. The idea is to define $\Gamma(P)$ to be the dimension of a linear space of polynomials defined in terms of the partial derivatives of P. In the syntactic set up, if a circuit C computes a polynomial P, then any partial derivative of C must be equivalent to the corresponding partial derivative of P. This observation along with bounds on the dimension of the partial derivative based linear spaces, led to circuit lower bounds.

However, this clearly breaks down in the case when our only guarantee is that the circuit C and the polynomial P agree as functions on all of $\{0,1\}^n$. Apriori, it is not clear if we can say anything meaningful about how the partial derivatives of C and those of P are related to each other. An extreme case of this is the following example. Let the polynomials P and Q be defined as follows:

$$P = \left(\sum_{i=1}^{n} x_i\right)^n$$

and

$$Q = P \mod I_0$$

Here I_0 is the ideal generated by the polynomials $\{x_i^2 - x_i : i \in [n]\}$. The following items follow easily from the definitions:

- $\forall \mathbf{x} \in \{0,1\}^n, P(\mathbf{x}) = Q(\mathbf{x})$.
- The dimension of the span of partial derivatives of P is at most $n+1$.
- The dimension of the span of partial derivatives of Q is at least 2^n. This follows from the fact that the leading monomial of Q is $x_1 \cdot x_2 \cdots x_n$.

So, clearly the dimension of the partial derivatives of two polynomials which are functionally the same over $\{0,1\}^n$ can be wildly different. Thus, it seems tricky to extend the proofs of syntactic lower bounds to the functional setup. Nevertheless, we do manage to get around this obstacle in certain cases as our results in the next section show. Moreover, we also show that a general solution to this question offers a possibility of proving new lower bounds for boolean circuits, that have so far been beyond our reach so far.

1.3 Our results

We now state our main results.

As our first result, we show functional lower bounds for homogeneous[2] depth-3 circuits. In the syntactic setting such lower bounds were first shown by Nisan and Wigderson [11] using the partial derivative of a polynomial as the complexity measure. However, as we discussed in subsubsection 1.2.1, partial derivative based proofs do not extend to the functional setting in a straightforward manner. We get around this obstacle by working with a different but related complexity measure. We now formally state the theorem:

▶ **Theorem 1.6.** *Let $\mathbb{F}$ be any field. There exists a family $\{P_d\}$ of polynomials of degree d in $n = \text{poly}(d)$ variables in VNP such that any $\Sigma\Pi\Sigma$ circuit of formal degree d which is functionally equivalent to P_d over $\{0,1\}^n$ has size at least $\exp\left(\Omega\left(d \log n\right)\right)$.*

As our second result, we show similar functional analogues of the homogeneous depth-4 lower bounds of [7, 10] but under the restriction that the depth-4 circuit computes a polynomial of *low individual degree*. As discussed in the introduction, such lower bounds for depth-4 circuits with bounded bottom fan-in but unbounded individual degree would imply that $\#\mathsf{P} \neq \mathsf{ACC}^0$, and would be a major progress on the question of boolean circuit lower bounds.

▶ **Theorem 1.7.** *Let $\mathbb{F}$ be any field. There exists a family $\{P_d\}$ of polynomials of degree d in $n = \text{poly}(d)$ variables in VNP such that any $\Sigma\Pi\Sigma\Pi$ circuit of formal degree d and individual degree[3] $O(1)$ which is functionally equivalent to P_d over $\{0,1\}^n$ has size at least $\exp\left(\Omega\left(\sqrt{d}\log n\right)\right)$.*

Our techniques for the proof of Theorem 1.7 are again different from the proofs of homogeneous depth-4 lower bounds in the syntactic setting. We introduce a family of new complexity measures, which are functional in their definition (as opposed to partial derivative based measures), and use them to capture functional computation. The family of measures, called *Shifted Evaluation dimension* is a shifted analogue of the well known notion of evaluation dimension, which has had many applications in algebraic complexity (for instance, in multilinear formula, circuit lower bounds [13, 12, 14]). We believe that the measure is of independent interest, and could have other potential applications.

Elementary symmetric polynomials

In their paper [11], Nisan and Wigderson showed an exponential lower bound on the size of homogeneous depth-3 circuits computing the elementary symmetric polynomials. A curious consequence of our proof, is that we are unable to show an analogue of Theorem 1.6 for the elementary symmetric polynomials. One of the reasons for this is the fact that the elementary symmetric polynomials have a *small* evaluation dimension complexity (the complexity measure used for this lower bound), hence our proof technique fails. However, it turns out the at least over fields of sufficiently large characteristic, there are polynomial sized depth-3 circuits of low

[2] Our lower bounds require that the formal degree of the circuit and the degree of the polynomial are *close* to each other. Homogeneity guarantees this condition, but is a much stronger condition than what we need for our proofs to work.

[3] The bounds do not extend to the case of individual degree $\omega(\log n)$, but may still hold for extremely slowly growing functions of n. However, for clarity of presentation, we work with constant individual degree throughout this paper.

formal degree which are functionally equivalent to the elementary symmetric polynomials over $\{0,1\}^n$. The upper bounds are based on the simple observation that for any d and $x \in \{0,1\}^n$, the value of $Sym_d(x)$ (elementary symmetric polynomial of degree d) is equal to $\binom{h(x)}{d}$, where $h(x) = \sum_i x_i$ is the hamming weight of x. In particular, for $d = 1$, the polynomial $\sum_i x_i$ is functionally equivalent to Sym_1, the polynomial $\frac{(\sum_i x_i)(\sum_i x_i - 1)}{2}$ is functionally equivalent to Sym_2 and so on. In particular, there is a polynomial which is a product of d affine forms which is equivalent to Sym_d. However, over fields of low characteristic, the complexity of the elementary symmetric polynomials for functional computation by depth-3 (or even depth-4) circuits is not clear to us and is an interesting open question.

Comparison to Kayal, Saha, Tavenas [8]

In a recent independent result, Kayal, Saha and Tavenas showed exponential lower bounds for depth-4 circuits of bounded individual degree computing an explicit polynomial in VP. Their proof uses a complexity measure called *skew shifted partials* which is very similar in spirit to the notion of *shifted evaluation dimension*, the complexity measure we use. Even though the results seem related, none of them subsumes the other. For our proof, we require that the formal degree of the depth-4 circuit is small (homogeneity), in addition to the individual degree being small, whereas in [8] the authors only require the individual degree of the circuit to be small. In this sense, their result is for a more general model than ours. However, for our lower bounds, we only require the circuit to agree with the target hard polynomial over $\{0,1\}^n$ while the proof in [8] is for syntactically computing the hard polynomial. Hence, the results are incomparable.

1.4 Organization of the paper

We set up some notations to be used in the rest of the paper in section 2. We prove the connections between functional lower bounds for depth-4 circuits and lower bounds for ACC^0 in section 3. We introduce our main complexity measure in section 4. We define and study the properties of the hard polynomials for our lower bounds in section 5. We present the proof of Theorem 1.6 in section 6 and the proof of Theorem 1.7 in section 7.

2 Notation

We now setup some notation to be used for the rest of the paper.

- Throughout the paper, we shall use bold-face letters such as $\mathbf{x}$ to denote a set $\{x_1, \ldots, x_n\}$. Most of the times, the size of this set would be clear from context. We shall also abuse this notation to use $\mathbf{x}^{\mathbf{e}}$ to refer to the monomial $x_1^{e_1} \cdots x_n^{e_n}$.
- The set of formal variables in this paper denoted by $\mathbf{x}$ of size n shall often be partitioned into sets $\mathbf{y}$ and $\mathbf{z}$. We shall use $\mathbf{x} = \mathbf{y} \sqcup \mathbf{z}$ to denote this and use n_y and n_z to denote the sizes of $\mathbf{y}$ and $\mathbf{z}$ respectively.
- For an integer $m > 0$, we shall use $[m]$ to denote the set $\{1, \ldots, m\}$.
- We shall use the short-hand $\partial_{\mathbf{x}^{\mathbf{e}}}(P)$ to denote

$$\frac{\partial^{e_1}}{\partial x_1^{e_1}} \left(\frac{\partial^{e_2}}{\partial x_2^{e_2}} \left(\cdots (P) \cdots \right) \right).$$

- For a set of polynomials $\mathcal{P}$ shall use $\partial_{\mathbf{y}}^{=k}\mathcal{P}$ to denote the set of all k-th order partial derivatives of polynomials in $\mathcal{P}$ with respect to y variables only, and $\partial_{\mathbf{y}}^{\leq k}\mathcal{P}$ similarly.

Also, $\mathbf{x}^{=\ell}\mathcal{P}$ shall refer to the set of polynomials of the form $\mathbf{x}^{\mathbf{e}} \cdot P$ where $\mathsf{Deg}(\mathbf{x}^{\mathbf{e}}) = \ell$ and $P \in \mathcal{P}$. Similarly $\mathbf{x}^{\leq\ell}\mathcal{P}$.

- For a polynomial $P \in \mathbb{F}[\mathbf{x}]$ and for a set $S \subseteq \mathbb{F}^n$, we shall denote by $\mathsf{Eval}_S(P)$ the vector of the evaluation of P on points in S (in some natural predefined order like say the lexicographic order). For a set of vectors V, their span over $\mathbb{F}$ will be denoted by $\mathsf{Span}(V)$ and the dimension of their span by $\mathsf{Dim}(V)$.
- We use $\{0,1\}^n_{\leq k}$ to denote the set of all boolean vectors of length n which have at most k ones.

3 Functional lower bounds for depth-4 circuits and ACC^0

In this section, we show that strong enough functional lower bounds for even very special depth-4 arithmetic circuits are sufficient to imply new lower bounds for ACC^0. The proof follows from a simple application of a well known characterization of ACC^0 by Yao [18] and Beigel and Tarui [2]. The following version of the theorem is from Arora-Barak [1].

▶ **Theorem 3.1** ([18, 2]). *If a function $f : \{0,1\}^n \to \{0,1\}$ is in ACC^0, then f can be computed by a depth 2 circuit with a symmetric gate with quasipolynomial $\left(\exp(\log^{O(1)} n)\right)$ fan-in at the output level and $\vee$ gates with polylogarithmic $\left(\log^{O(1)} n\right)$ fan-in at the bottom level.*

We now prove the following lemma which shows *functional* upper bound for ACC^0.

▶ **Lemma 3.2.** *Let $\mathbb{F}$ be any field of characteristic zero or at least $\exp(\omega(\mathrm{poly}(\log n)))$. If a function $f : \{0,1\}^n \to \{0,1\}$ is in ACC^0, then there exists a polynomial $P_f \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that the following are true:*

- *For every $\mathbf{x} \in \{0,1\}^n$, $f(\mathbf{x}) = P_f(\mathbf{x})$.*
- *P_f can be computed by a quasipolynomial sized $\Sigma\wedge\Sigma\Pi$ circuit with bottom fan-in at most $\mathrm{poly}(\log n)$, which are depth-4 circuits where the product gates in the second level[4] are powering gates.*

Proof. From Theorem 3.1, we know that there exists a symmetric function h and multilinear polynomials $g_1, g_2, \ldots, g_t$ such that

- $t = \exp(\mathrm{poly}(\log n))$.
- For every $\mathbf{x} \in \{0,1\}^n$, $f(\mathbf{x}) = h(g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_t(\mathbf{x}))$.
- Each g_i is a multilinear polynomial in at most $\mathrm{poly}(\log n)$ variables.
- For every $\mathbf{x} \in \{0,1\}^n$ and $j \in [t]$, $g_j(\mathbf{x}) \in \{0,1\}$.

From the last item above, we know that the g_is only take boolean values on inputs from $\{0,1\}^n$. Since h is symmetric, it follows that its value on boolean inputs only depends upon the hamming weight of its input. Hence, h is in fact a function of $\sum_{i\in[t]} g_i$. Therefore, over any field of characteristic zero or larger than t, there exists a univariate polynomial P_h of degree at most t over reals, such that

$$\forall \mathbf{x} \in \{0,1\}^n, h\left(g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_t(\mathbf{x})\right) = P_h\left(\sum_{i\in[t]} g_i(\mathbf{x})\right).$$

[4] Throughout this paper, we will assume that our circuits are levelled with alternating $+$ and $\times$ gates. The output gate is level 1 and its inputs are at level 2 and so on.

The lemma now follows from the fact that each g_i is a multilinear polynomial in $\text{poly}(\log n)$ variables. ◀

Theorem 3.2 now immediately implies the following lemma.

▶ **Lemma 3.3.** *Let $\mathbb{F}$ be any field of characteristic zero or at least $\exp(\omega(\text{poly}(\log n)))$. Then, an $\exp(\omega(\text{poly}(\log n)))$ functional lower bound for a function on n variables for $\Sigma \wedge \Sigma\Pi^{[\text{poly}(\log n)]}$ circuits over $\mathbb{F}$ would imply that f is not in ACC^0.*

4 The complexity measure

In the lower bounds for homogeneous depth four circuits [7, 10], the complexity measure used was the *dimension of projected shifted partial derivatives*. The following definition is not the same as used in [7, 10], but this slight variant would be easier to work with for our applications. We abuse notation to call it "projected shifted partial derivatives" as it continues to have the essence of the original definition. A discussion on the precise differences between the following definition and the original definition of [7, 10] is present in Appendix A

▶ **Definition 4.1** (Projected shifted partial derivatives)**.** Let $\mathbf{x} = \mathbf{y} \sqcup \mathbf{z}$ with $|\mathbf{y}| = n_y$ and $|\mathbf{z}| = n_z$, and let S be the set of all strings in $\{0,1\}^{n_y+n_z}$ that are zero on the first n_y coordinates. If k, ℓ are some parameters, the *dimension of projected shifted partial derivatives* for any polynomial $P(\mathbf{y}, \mathbf{z}) \in \mathbb{F}[\mathbf{y}, \mathbf{z}]$, denoted by $\Gamma^{\text{PSPD}}_{k,\ell}(P)$, is defined as

$$\Gamma^{\text{PSPD}}_{k,\ell}(P) \quad := \quad \text{Dim}\left\{\text{Eval}_S\left(\mathbf{z}^{=\ell}\partial^{=k}_{\mathbf{y}}(P)\right)\right\}.$$

The above measure is still syntactic as partial derivatives are not useful in the functional setting. For the functional setting, we shall use a different measure for our lower bound that we call *the shifted evaluation dimension*. We now define the complexity measure that we shall be using to prove the lower bound. For brevity, we shall assume that our set of variables $\mathbf{x}$ is partitioned into $\mathbf{y}$ and $\mathbf{z}$. For our proofs, we shall use a carefully chosen partition. We now formally define the notion of *shifted evaluation dimension* of a polynomial below.

▶ **Definition 4.2** (Shifted evaluation dimension)**.** Let ℓ and k be some parameters and let $\mathbf{x} = \mathbf{y} \sqcup \mathbf{z}$ such that $|\mathbf{y}| = n_y$ and $|\mathbf{z}| = n_z$. For any polynomial $P \in \mathbb{F}[\mathbf{y}, \mathbf{z}]$, define $\Gamma_{k,\ell}(P)$ as

$$\Gamma^{\text{SED}}_{k,\ell}(P) \quad := \quad \text{Dim}\left\{\text{Eval}_{\{0,1\}^{n_z}}\left(\mathbf{z}^{=\ell} \cdot \{P(\mathbf{a}, \mathbf{z}) : \mathbf{a} \in \{0,1\}^{n_y}_{\leq k}\}\right)\right\}.$$

Informally, for every polynomial P, we fix a partition of the input variables into $\mathbf{y}$ and $\mathbf{z}$ and generate a linear space by the following algorithm.

- We take the projections of P obtained by setting each of the y variables to $0, 1$ such that the number of y variables set to 1 is at most k.
- We shift the polynomials obtained in step 1 by all monomials in variables $\mathbf{z}$ of degree ℓ.
- Observe that the polynomials obtained at the end of step two are polynomials only in the $\mathbf{z}$ variables. We now look at the evaluation vectors of these polynomials over $\{0,1\}^{n_z}$.

The complexity measure of the polynomial P is defined as the dimension of the linear space generated by the vectors obtained at the end of step 3 in the algorithm above. For our proof, we will pick a careful partition of the variables $\mathbf{x}$ into $\mathbf{y}$ and $\mathbf{z}$ and look at $\Gamma^{\text{SED}}_{k,\ell}(P)$. The following lemma highlights the key reason of utility of the above measure to functional lower bounds.

▶ **Lemma 4.3** (Functional equivalence and shifted evaluation dimension)**.** *Let $P \in \mathbb{F}[\mathbf{x}]$ and $Q \in \mathbb{F}[\mathbf{x}]$ be any two polynomials which are functionally equivalent over $\{0,1\}^n$. Then, for every choice of k, ℓ and partition $\mathbf{x} = \mathbf{y} \sqcup \mathbf{z}$*

$$\Gamma^{\mathrm{SED}}_{k,\ell}(P) \quad = \quad \Gamma^{\mathrm{SED}}_{k,\ell}(Q)\,.$$

Proof. The proof easily follows from the fact that the measure $\Gamma^{\mathrm{SED}}_{k,\ell}(P)$ is the dimension of a linear space which is generated by vectors which correspond to evaluations of P over subcubes of $\{0,1\}^n$. Hence, it would be the same for any two polynomials which agree as functions over $\{0,1\}^n$. ◀

▶ Remark. Observe that a lemma analogous to Theorem 4.3 is not true in general for partial derivative based measures. And hence, the proofs for syntactic lower bounds which are based on such measures does not immediately carry over to the functional setting.

4.1 Evaluations vs. partial derivatives

In this section, we show that for polynomials of low individual degree, the notion of shifted evaluation dimension can be used as a proxy for the notion of shifted partial derivatives. This is the key observation that drives the proofs of Theorem 1.6 and Theorem 1.7. We first consider the case when the polynomial is *set-multilinear* in which case derivatives can be directly related to careful evaluations.

4.1.1 For set-multilinear polynomials

The explicit polynomials we shall be working with in this paper would be *set-multilinear*. An example to keep in mind is Det_n or Perm_n where the variables can be partitioned into rows and each monomial involves exactly one variable from each part.

▶ **Definition 4.4** (Set-multilinear polynomials)**.** A polynomial P is said to be *set-multilinear* with respect to the a partition $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_r$ if every monomial of P involves exactly[5] one variable from each $\mathbf{x}_i$.

We begin with the following simple observation.

▶ Observation 4.5. Let $P \in \mathbb{F}[\mathbf{x}]$ be a set-multilinear with respect to a partition $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_r$. Let $\mathbf{y} = \mathbf{x}_1 \cup \cdots \cup \mathbf{x}_k$ for some $k \leq r$ and let $\mathbf{z} = \mathbf{x} \setminus \mathbf{y}$. Then, for any degree k monomial $\mathbf{y}^{\mathbf{e}}$ that is set-multilinear with respect to $\mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_k$, we have

$$\frac{\partial P}{\partial \mathbf{y}^{\mathbf{e}}} \quad = \quad P(\mathbf{e}, \mathbf{z}).$$

Proof. We shall prove this by induction on k. Suppose $\mathbf{y} = \mathbf{x}_1$ and $y_1 \in \mathbf{x}_1$. Since P is set-multilinear, we can write P as

$$P(\mathbf{x}_1, \cdots, \mathbf{x}_r) \quad = \quad \sum_{y_i \in \mathbf{x}_1} y_i \cdot P_i(\mathbf{x}_2, \cdots, \mathbf{x}_r).$$

Hence it follows that $\partial_{y_1}(P)$ equals P_1, which is also the partial evaluation of P where y_1 is set to 1 and all other $y_i \in \mathbf{x}_1$ is set to zero. Hence, if $y_1 = \mathbf{y}^{\mathbf{e}}$, then $\partial_{y_1}(P) = P(\mathbf{e}, \mathbf{x}_2, \cdots, \mathbf{x}_r)$. The claim follows by repeating this argument on $P(\mathbf{e}, \mathbf{x}_2, \cdots, \mathbf{x}_r)$ which continues to be set-multilinear. ◀

[5] sometimes in the literature the word 'exactly' is replaced by 'at most' but in this paper we would be dealing with this definition.

Theorem 4.5 immediately implies the following corollary, which shows that for set-multilinear polynomials *shifted* evaluation dimension and *shifted* partial derivatives are the same quantity if we choose our set of derivatives carefully.

▶ **Corollary 4.6.** *Let $P(\mathbf{x})$ be a set-multilinear polynomial with respect to $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_r$. Suppose $\mathbf{y} = \mathbf{x}_1 \cup \cdots \cup \mathbf{x}_k$ and $\mathbf{z} = \mathbf{x} \setminus \mathbf{y}$. Then if we consider the dimension of projected shifted partials with respect to set-multilinear monomials in $\mathbf{y}$, we have*

$$\Gamma^{\mathrm{PSPD}}_{k,\ell}(P) \quad \leq \quad \Gamma^{\mathrm{SED}}_{k,\ell}(P).$$

4.1.2 For low individual degree polynomials

We now proceed to show that an *approximation* of the Theorem 4.6 also holds for polynomials of low individual degree.

▶ **Lemma 4.7.** *Let $P(\mathbf{y}, \mathbf{z})$ be a polynomial with individual degree at most r. Then, for every choice of parameters k and ℓ*

$$\left\{ P(\mathbf{a}, \mathbf{z}) : a \in \{0,1\}^{n_y}_{\leq k} \right\} \quad \subseteq \quad \mathsf{Span}\left(\left(\partial^{\leq rk} P\right)_{\mathbf{y}=\mathbf{0}} \right).$$

Proof. For the rest of this proof, we shall think of P as an element $P_{\mathbf{z}}(\mathbf{y}) \in \mathbb{F}[\mathbf{z}][\mathbf{y}]$. Let $\mathbf{a}$ be any point in $\{0,1\}^{n_y}$. Then by the Taylor's expansion, we know that

$$P_{\mathbf{z}}(\mathbf{y} + \mathbf{a}) \quad = \quad \sum_{\mathbf{e}} \mathbf{a}^{\mathbf{e}} \cdot \partial_{\mathbf{y}^{\mathbf{e}}}(P_{\mathbf{z}})(\mathbf{y}).$$

If the support of $\mathbf{a}$ is at most k, then for every $\mathbf{e}$ such that $\|\mathbf{e}\|_0 > k$, we would have $\mathbf{a}^{\mathbf{e}} = 0$. Moreover, since P is a polynomial of individual degree at most r, it follows that if any coordinate of $\mathbf{e}$ is more than r then

$$\partial_{\mathbf{y}^{\mathbf{e}}}(P_{\mathbf{z}}) = 0.$$

In summary, for any $\mathbf{a}$ such that $\|\mathbf{a}\|_0 \leq k$,

$$\begin{aligned} P_{\mathbf{z}}(\mathbf{y} + \mathbf{a}) \quad &= \quad \sum_{\substack{\mathbf{e}:\|\mathbf{e}\|_0 \leq k, \\ \|\mathbf{e}\|_1 \leq rk}} \mathbf{a}^{\mathbf{e}} \cdot \partial_{\mathbf{y}^{\mathbf{e}}}(P_{\mathbf{z}})(\mathbf{y}) \\ \implies P_{\mathbf{z}}(\mathbf{a}) \quad = \quad P(\mathbf{a}, \mathbf{z}) \quad &= \quad \sum_{\substack{\mathbf{e}:\|\mathbf{e}\|_0 \leq k, \\ \|\mathbf{e}\|_1 \leq rk}} \mathbf{a}^{\mathbf{e}} \cdot \left(\partial_{\mathbf{y}^{\mathbf{e}}}(P_{\mathbf{z}})\right)_{\mathbf{y}=\mathbf{0}} \quad \in \quad \mathsf{Span}\left(\left(\partial^{\leq rk} P\right)_{\mathbf{y}=\mathbf{0}} \right). \end{aligned}$$ ◀

We are now ready to prove our main technical claim of this section.

▶ **Lemma 4.8.** *Let $P(\mathbf{y}, \mathbf{z})$ be a polynomial with individual degree at most r. Then, for every choice of parameters k and ℓ,*

$$\Gamma^{\mathrm{SED}}_{k,\ell}(P) \quad \leq \quad \Gamma^{\mathrm{PSPD}}_{rk,\ell}(P)$$

Proof. From Theorem 4.7, we know that

$$\begin{aligned} \left\{ P(\mathbf{a}, \mathbf{z}) : \mathbf{a} \in \{0,1\}^{n_y}_{\leq k} \right\} \quad &\subseteq \quad \mathsf{Span}\left(\left(\partial^{\leq rk} P\right)_{\mathbf{y}=\mathbf{0}} \right) \\ \implies \left\{ \mathbf{z}^{=\ell} \cdot P(\mathbf{a}, \mathbf{z}) : \mathbf{a} \in \{0,1\}^{n_y}_{\leq k} \right\} \quad &\subseteq \quad \mathsf{Span}\left(\mathbf{z}^{=\ell} \cdot \left(\partial^{\leq rk} P\right)_{\mathbf{y}=\mathbf{0}} \right) \end{aligned}$$

By looking at the evaluation vectors on $\{0,1\}^{n_z}$,

$$\begin{aligned}\left\{\mathsf{Eval}_{\{0,1\}^{n_z}}\left(\mathbf{z}^{=\ell}\cdot P(\mathbf{a},\mathbf{z})\right) : \mathbf{a}\in\{0,1\}^{n_y}_{\leq k}\right\} &\subseteq \mathsf{Span}\left(\mathsf{Eval}_{\{0,1\}^{n_z}}\left(\mathbf{z}^{=\ell}\cdot\left(\partial^{\leq rk}P\right)_{\mathbf{y}=\mathbf{0}}\right)\right)\\ &= \mathsf{Span}\left(\mathsf{Eval}_{\{0\}^{n_y}\times\{0,1\}^{n_z}}\left(\mathbf{z}^{=\ell}\cdot\partial^{\leq rk}P\right)\right)\end{aligned}$$

Taking the dimension of the linear spans on both sides completes the proof. ◀

5 Nisan-Wigderson polynomial families

In this section, we formally define the family of Nisan-Wigderson polynomials and mention some known results about lower bounds on the their projected shifted partials complexity [7, 10, 9]. These bounds will be critically used in our proof.

▶ **Definition 5.1** (Nisan-Wigderson polynomial families)**.** Let d, m, e be arbitrary parameters with m being a power of a prime, and $d, e \leq m$. Since m is a power of a prime, let us identify the set $[m]$ with the field $\mathbb{F}_m$ of m elements. Note that since $d \leq m$, we have that $[d] \subseteq \mathbb{F}_m$. The Nisan-Wigderson polynomial with parameters d, m, e, denoted by $\mathsf{NW}_{d,m,e}$ is defined as

$$\mathsf{NW}_{d,m,e}(\mathbf{x}) \quad = \quad \sum_{\substack{p(t)\in\mathbb{F}_m[t]\\ \mathsf{Deg}(p)<e}} x_{1,p(1)}\cdots x_{d,p(d)}$$

That is, for every univariate polynomial $p(t) \in \mathbb{F}_m[t]$ of degree less that e, we add one monomial that encodes the 'graph' of p on the points $[d]$.

This is a homogeneous, multilinear polynomial of degree d over dm variables with exactly m^e monomials. Furthermore, the polynomial is *set-multilinear* with respect to $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_d$ where $\mathbf{x}_i = \{x_{i1}, \cdots, x_{im}\}$.

We now state the following lemma which shows a lower bound on the $\Gamma^{\mathrm{PSPD}}_{k,\ell}(\mathsf{NW}_{d,m,e})$ for an appropriate choice of parameters. We will then use this bound along with Theorem 4.6 to show a lower bound on $\Gamma^{\mathrm{SED}}_{k,\ell}(NW_{d,m,e})$. The lower bound on $\Gamma^{\mathrm{PSPD}}_{k,\ell}(\mathsf{NW}_{d,m,e})$ was shown in two independent proofs by Kayal et al. [7] and by Kumar and Saraf [10]. The version stated below is from a strengthening of these bounds by Kumar and Saptharishi [9].

▶ **Lemma 5.2.** *For every d and $k = O(\sqrt{d})$ there exists parameters m, e, ϵ such that $m = \Theta(d^2)$ and $\epsilon = \Theta\left(\frac{\log d}{\sqrt{d}}\right)$ with*

$$\begin{aligned} m^k &\geq (1+\epsilon)^{2(d-k)}\\ m^{e-k} &= \left(\frac{2}{1+\epsilon}\right)^{d-k}\cdot \mathrm{poly}(m).\end{aligned}$$

For such a choice of parameters, let $\mathbf{x} = \{x_{ij} : i \in [d], j \in [m]\} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_d$ where $\mathbf{x}_i = \{x_{i1}, \ldots, x_{im}\}$. Let $\mathbf{y} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_k$ and $\mathbf{z} = \mathbf{x} \setminus \mathbf{y}$. If ℓ is a parameter that satisfies $\ell = \frac{n_z}{2}(1-\epsilon)$, then over any field $\mathbb{F}$, we have[6]

$$\Gamma^{\mathrm{PSPD}}_{k,\ell}(\mathsf{NW}_{d,m,e}(\mathbf{y},\mathbf{z})) \quad \geq \quad \binom{n_z}{\ell+d-k}\cdot\exp(-O(\log^2 d)).$$

[6] We remark that in the calculations in [7, 10, 9], the shifted monomials consist of both the $\mathbf{y}$ and $\mathbf{z}$ variables, while here we only shift by $\mathbf{z}$ variables. But the calculations still go through since the parameters continue to satisfy the constraints needed for soundness of the calculation.

From Theorem 4.6, we immediately have the following crucial lemma.

▶ **Lemma 5.3.** *Let d, m, e, ℓ be parameters as defined in Theorem 5.2 and let $\mathbf{y}$ and $\mathbf{z}$ be the partition of variables $\mathbf{x}$ as in Theorem 5.2. Then,over any field $\mathbb{F}$, we have*

$$\Gamma^{\mathrm{SED}}_{k,\ell}(\mathsf{NW}_{d,m,e}(\mathbf{y},\mathbf{z})) \quad \geq \quad \binom{n_z}{\ell+d-k} \cdot \exp(-O(\log^2 d)).$$

6 Functional lower bounds for depth-3 circuits

In this section, we complete the proof of Theorem 1.6. We start by defining the exact hard polynomial for which our lower bound is shown.

Hard polynomials for the lower bound

We will prove Theorem 1.6 for the polynomial $\mathsf{NW}_{d,m,e}$ for an appropriate choice of the parameters.

▶ **Lemma 6.1.** *Let the parameters e and d be chosen so that $e = d/2 - 1$, and let $k = e + 1$. Let the variables $\mathbf{x}$ in $\mathsf{NW}_{d,m,e}$ be partitioned into $\mathbf{y} = \{x_{ij} \ : \ i \in [k], j \in [m]\}$ and $\mathbf{z} = \mathbf{x} \setminus \mathbf{y}$. Then*

$$\Gamma^{\mathrm{SED}}_{k,0}(\mathsf{NW}_{d,m,e}(\mathbf{y},\mathbf{z})) \quad \geq \quad m^{d/2}\,.$$

Proof. Let the set of monomials S be defined as

$$S = \left\{ \prod_{i=1}^{k} x_{i,j_i} : j_i \in [m] \right\}.$$

Observe that for every monomial $\mathbf{x}^\alpha$ in S, the partial derivative of $\mathsf{NW}_{d,m,e}$ with respect to $\mathbf{x}^\alpha$, is a monomial in $\mathbf{z}$. This is due to the fact that $e < d/2$ and no two distinct univariate polynomials of degree $d/2$ can agree at more than $d/2$ many points. Moreover for every two distinct monomials $\mathbf{x}^\alpha$ and $\mathbf{x}^\beta$ in S,

$$\frac{\partial \mathsf{NW}_{d,m,e}}{\partial \mathbf{x}^\alpha} \quad \neq \quad \frac{\partial \mathsf{NW}_{d,m,e}}{\partial \mathbf{x}^\beta}\,.$$

Hence,

$$\Gamma^{\mathrm{PSPD}}_{k,0}(\mathsf{NW}_{d,m,e}) = |S| = m^{d/2}\,.$$

Since $\mathsf{NW}_{d,m,e}$ is a set-multilinear with respect to the rows of variable matrix, by Theorem 4.5, it follows that

$$\Gamma^{\mathrm{SED}}_{k,0}(\mathsf{NW}_{d,m,e}) = m^{d/2}\,.$$ ◀

Complexity of the model

▶ **Lemma 6.2.** *The $C(\mathbf{x})$ be a $\Sigma\Pi\Sigma$ circuit of formal degree d and top fan-in s. Then, for all choices of k and any partition of $\mathbf{x}$ into $\mathbf{y}$ and $\mathbf{z}$,*

$$\Gamma^{\mathrm{SED}}_{k,0}(C) \quad \leq \quad s \cdot 2^d\,.$$

Proof. Observe that for any choice of k and ℓ, $\Gamma^{\mathrm{SED}}_{k,\ell}$ is a subadditive measure. Therefore, it is enough to upper bound the value of $\Gamma^{\mathrm{SED}}_{k,0}()$ for every product gate in C by 2^d. Let

$$Q(\mathbf{y},\mathbf{z}) = \prod_{i=1}^{d} L_i$$

be any product gate of formal degree at most d in C. Since each L_i is a linear form, we can express it as $L_i = L_{yi} + L_{zi}$, where L_{yi} and L_{zi} are the parts of L_i consisting entirely of $\mathbf{y}$ and $\mathbf{z}$ variables respectively. Therefore,

$$Q(\mathbf{y},\mathbf{z}) = \sum_{S\subseteq[d]} \prod_{i\in S} L_{yi} \cdot \prod_{j\notin S} L_{zj}\,.$$

Now observe that by

$$\{Q(\mathbf{a},\mathbf{z}) : \mathbf{a}\in\{0,1\}^{n_y}\} \quad\subseteq\quad \mathsf{Span}\left(\left\{\prod_{j\notin S} L_{zj} : S\subseteq[d]\right\}\right)$$

Therefore,

$$\Gamma^{\mathrm{SED}}_{k,0}(C) \quad\leq\quad 2^d\,.$$

The lemma now follows by subadditivity. ◀

Wrapping up the proof

We are now ready to complete the proof of Theorem 1.6.

▶ **Theorem 6.3.** *Let $\mathbb{F}$ be any field, and let d, m, e be parameters such that $e = d/2 - 1$ and $m = \mathrm{poly}(d)$. Let C be a $\Sigma\Pi\Sigma$ circuit of formal degree d which is functionally equivalent to the polynomial $\mathsf{NW}_{d,m,e}$. Then*

$$\mathit{Size}(C) \geq m^{d/2}/2^d\,.$$

Proof. Let $k = e+1$ and consider a partition of variables into $\mathbf{y}$ and $\mathbf{z}$ where all the variables in the first k rows of the variable matrix are labelled $\mathbf{y}$ and the remaining variables are labelled $\mathbf{z}$. Now, the theorem immediately follows from Theorem 6.1 and Theorem 6.2. ◀

7 Functional lower bounds for depth-4 circuits

In this section, we prove Theorem 1.7. We first define the family of polynomials for which our lower bounds apply.

Hard polynomials for the lower bound

For the proof of Theorem 1.7, we would have to show that a statement in the spirit of Theorem 5.3 is also true for a *random projection* of our hard polynomial. Even though we believe[7] that this is true for the polynomial defined in Theorem 5.1, for simplicity, we modify our hard polynomial and in turn prove a lower bound for the following variant of it.

[7] In fact, [7, 10] showed such statements to be true.

▶ **Definition 7.1** (Hard polynomials for the lower bound)**.** Let d, m, e be parameters as defined in Theorem 5.1. Let $p = p(m, d)$ be a parameter and let

$$t = \frac{dm}{p}\,.$$

The polynomial $\mathsf{NW} \circ \mathsf{Lin}$ is defined as

$$\mathsf{NW} \circ \mathsf{Lin}_{d,m,e,p} = \mathsf{NW}_{d,m,e}\left(L(x_{1,1}), L(x_{1,2}), \ldots, L(x_{d,m})\right)$$

where for each $i \in [d], j \in [m]$, $L(x_{i,j})$ is defined as

$$L(x_{i,j}) = \sum_{u=1}^{t} x_{i,j,u}\,.$$

For the rest of this proof, we set $p = (md)^{-0.1}$, and for brevity, we will indicate $\mathsf{NW} \circ \mathsf{Lin}_{d,m,e,(md)^{0.1}}$ by $\mathsf{NW} \circ \mathsf{Lin}_{d,m,e}$. Observe that setting p sets t to be equal to $(md)^{1.1}$. We conclude this section with the next lemma where we show that $\mathsf{NW} \circ \mathsf{Lin}_{d,m,e}$ is *robust* under random restrictions where every variable is kept alive with a probability p.

▶ **Lemma 7.2.** *Let p and t be as stated above and let $n = dm$. Let P be a random projection of $\mathsf{NW} \circ \mathsf{Lin}$ obtained by setting every variable in $\{x_{i,j,h} : i \in [d], j \in [m], h \in [t]\}$ to zero with a probability equal to $1 - p$. Then, with a probability at least $1 - o(1)$, $\mathsf{NW}_{d,m,e}$ is a projection of P.*

Proof. For every $i \in [d]$, $j \in [m]$, define the set $A_{i,j}$ as

$$A_{ij} = \{x_{i,j,h} : h \in [t]\}\,.$$

When every variable is being set to zero with a probability $1 - p$, the probability that there exists an $i \in [d]$ and $j \in [m]$ such that all the variables in the set $A_{i,j}$ are set to zero is at most $dm(1-p)^t$. For $p = n^{-0.1}$, the probability is at most $n(1 - n^{-0.1})^{n^{1.1}}$ which is $\exp(-\Omega(n))$.

Therefore, with a probability at least $1 - \exp(-\Omega(n))$, each of the set $A_{i,j}$ has at least one variable alive in P. Now, we set all but one of them to zero for each i, j. Observe that the resulting projection of P is precisely $\mathsf{NW}_{d,m,e}$ up to a relabelling of variables. This proves the lemma. ◀

It should be noted that the polynomial $\mathsf{NW} \circ \mathsf{Lin}$ continues to remain set-multilinear with respect to he rows of the variable matrix.

Upper bound on the complexity of the model

We now show the upper bound on $\Gamma^{\mathrm{SED}}_{k,\ell}(C)$ when C is a depth-4 circuit of individual degree at most r and bottom support s. We will use the following upper bound on $\Gamma^{\mathrm{PSPD}}_{k,\ell}(C)$ from [7, 10].

▶ **Lemma 7.3.** *Let $C(\mathbf{y}, \mathbf{z})$ be a depth-4 circuit, of formal degree at most d and bottom support at most s. Let k and ℓ be parameters satisfying $\ell + ks < n_z/2$. Then*

$$\Gamma^{\mathrm{PSPD}}_{k,\ell}(C) \quad \leq \quad \mathit{Size}(C) \cdot \binom{O\left(\frac{d}{s}\right) + k}{k} \cdot \binom{n_z}{\ell + ks} \cdot \mathrm{poly}(n)\,.$$

The following lemma now immediately follows from Theorem 7.3 and Theorem 4.8.

▶ **Lemma 7.4.** *Let $C(\mathbf{y}, \mathbf{z})$ be a depth-4 circuit, of formal degree at most d, individual degree at most r and bottom support at most s. Let k and ℓ be parameters satisfying $\ell + krs < n_z/2$. Then*

$$\Gamma^{\mathrm{SED}}_{k,\ell}(C) \quad \leq \quad \mathit{Size}(C) \cdot \binom{O\left(\frac{d}{s}\right) + kr}{kr} \cdot \binom{n_z}{\ell + krs} \cdot \mathrm{poly}(n_z).$$

Wrapping up the proof

▶ **Theorem 7.5.** *Let d, m, e be parameters as defined in Theorem 5.2. Let C be a $\Sigma\Pi\Sigma\Pi$ circuit C of formal degree d and individual degree at most $r = O(1)$ over any field $\mathbb{F}$ such that C is functionally equivalent to $\mathsf{NW} \circ \mathsf{Lin}_{d,m,e}$. Then,*

$$\mathit{Size}(C) \geq \exp\left(\Omega\left(\sqrt{d}\log dm\right)\right).$$

Proof. If the size of C is larger than $\exp\left(\frac{\sqrt{d}\log dm}{1000r}\right)$, then we are already done, else the size of C is at most $\exp\left(\frac{\sqrt{d}\log dm}{1000r}\right)$. Let us set every variable in C and $\mathsf{NW} \circ \mathsf{Lin}_{d,m,e}$ to zero independently with a probability $1 - (md)^{-0.1}$. The following claim easily follows via a standard application of the union bound.

▶ **Claim 7.6.** *With probability at least $1 - o(1)$ over the random restrictions as defined above, every product gate at the bottom level of C with support at least $\frac{\sqrt{d}}{100r}$ is set to zero.*

From the above claim and from Theorem 7.2, it follows that there is a $\Sigma\Pi\Sigma\Pi$ circuit C' of formal degree d over $\mathbb{F}$ which is functionally equivalent to $\mathsf{NW}_{d,m,e}$. Let us relabel the variables as $\mathbf{y}$ and $\mathbf{z}$ as described in Theorem 5.2. Let $k = \sqrt{d}$ and let $\ell = \frac{n_z}{2} \cdot (1 - \epsilon)$ where $\epsilon = O\left(\frac{\log d}{\sqrt{d}}\right)$ to be chosen shortly. By Theorem 5.3, we know that for this choice of k and ℓ

$$\begin{aligned}
\Gamma^{\mathrm{SED}}_{k,\ell}(\mathsf{NW}_{d,m,e}(\mathbf{y}, \mathbf{z})) \quad &\geq \quad \binom{n_z}{\ell + d - k} \cdot \exp(-O(\log^2 d)) \\
&\geq \quad \binom{n_z}{\ell} \cdot (1+\epsilon)^{2d-2k} \cdot \exp(-O(\log^2 d))
\end{aligned}$$

Moreover, by Theorem 7.4, we know that

$$\begin{aligned}
\Gamma^{\mathrm{SED}}_{k,\ell}(C') \quad &\leq \quad (dm)^{\sqrt{d}/1000r} \cdot \binom{O\left(\frac{\sqrt{d}}{r}\right) + kr}{kr} \cdot \binom{n_z}{\ell + k \cdot r \cdot \frac{\sqrt{d}}{100r}} \cdot \mathrm{poly}(n_z) \\
&\leq \quad (dm)^{\sqrt{d}/1000r} \cdot 2^{O(\sqrt{d})} \cdot \binom{n_z}{\ell} \cdot (1+\epsilon)^{\frac{d}{50}} \cdot \exp(O(\log^2 d)) \\
&\leq \quad \exp\left(\sqrt{d}\log d/100r\right) \cdot 2^{O(\sqrt{d})} \cdot \binom{n_z}{\ell} \cdot (1+\epsilon)^{\frac{d}{50}} \cdot \exp(O(\log^2 d))
\end{aligned}$$

Now, observe that there exists a constant c such that if ϵ is set to $\frac{c\log d}{\sqrt{d}}$, then

$$\Gamma^{\mathrm{SED}}_{k,\ell}(\mathsf{NW}_{d,m,e}) > \Gamma^{\mathrm{SED}}_{k,\ell}(C').$$

But this is a contradiction since C' computes $\mathsf{NW}_{d,m,e}$. This completes the proof. ◀

8 Open problems

We end with some open questions:

- The main challenge would be to improve Theorem 1.7, and prove it for the model of sums of powers of low degree polynomials. It is not clear to us if the complexity measure used in this paper would be useful.
- The functional lower bounds proved in this paper are for *exact* functional computation. We believe that some of these bounds should also hold in the average case, where the circuit and the polynomial agree on a random point on $\{0,1\}^n$ with a high probability. It is not clear to us if the proof techniques in this paper can be adapted to say something in the average case setting. The most natural attempt to generalize the proofs seem to hit a *matrix rigidity* like obstacle.

Acknowledgements. Part of this work was done while the third author was visiting Rutgers. We are grateful to Eric Allender and DIMACS for funding the visit. We are also grateful to Pravesh Kothari and Madhu Sudan for many helpful conversations.

References

1 Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.

2 Richard Beigel and Jun Tarui. On acc. *Computational Complexity*, 4(4):350–366, 1994. doi:10.1007/BF01263423.

3 Michael A. Forbes, Amir Shpilka, Iddo Tzameret, and Avi Wigderson. Proof complexity lower bounds from algebraic circuit complexity. Manuscript, 2015.

4 Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 128–135, 2014. doi:10.1145/2591796.2591824.

5 Joshua A. Grochow and Toniann Pitassi. Circuit complexity, proof complexity, and polynomial identity testing. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, pages 110–119, 2014. Full version at arXiv:abs/1404.3820. doi:10.1109/FOCS.2014.20.

6 Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *Journal of the ACM*, 61(6):33:1–33:16, 2014. Preliminary version in the *28th Annual IEEE Conference on Computational Complexity (CCC 2013)*. doi:10.1145/2629541.

7 Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Circuits. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, 2014. doi:10.1109/FOCS.2014.15.

8 Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. On the size of homogeneous and of depth four formulas with low individual degree. *Electronic Colloquium on Computational Complexity (ECCC)*, 2015. eccc:TR15-181. URL: http://eccc.hpi-web.de/report/2015/181/.

9 Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. *Electronic Colloquium on Computational Complexity (ECCC)*, 2015. eccc:TR15-109. URL: http://eccc.hpi-web.de/report/2015/109/.

10 Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, 2014. doi:10.1109/FOCS.2014.46.

11 Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1997. doi:10.1007/BF01294256.

12 Ran Raz. Separation of multilinear circuit and formula size. *Theory of Computing*, 2(1):121–135, 2006. Preliminary version in the *45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*. doi:10.4086/toc.2006.v002a006.

13 Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *Journal of the ACM*, 56(2), 2009. Preliminary version in the *36th Annual ACM Symposium on Theory of Computing (STOC 2004)*. doi:10.1145/1502793.1502797.

14 Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Computational Complexity*, 18(2):171–207, 2009. Preliminary version in the *23rd Annual IEEE Conference on Computational Complexity (CCC 2008)*. doi:10.1007/s00037-009-0270-8.

15 Alexander A. 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.

16 Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing (STOC 1987)*, pages 77–82, 1987. doi:10.1145/28395.28404.

17 Ryan Williams. Non-uniform ACC Circuit Lower Bounds. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity (CCC 2011)*, pages 115–125, 2011.

18 Andrew Chi-Chih Yao. Separating the polynomial-time hierarchy by oracles. In *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1985)*, pages 1–10, Oct 1985. doi:10.1109/SFCS.1985.49.

A The evaluation perspective on projected shifted partial derivatives

The notion of projected shifted partial derivatives was first introduced by Kayal, Limaye, Saha and Srinivasan [7] in proving lower bounds for homogeneous depth-4 circuits. The following is the precise definition they used.

▶ **Definition A.1** (Projected shifted partial derivatives of [7])**.** Let k and ℓ be some parameters. The projected shifted partial derivatives of a polynomial $P(\mathbf{y}, \mathbf{z})$, denoted by $\Gamma_{k,\ell}^{\mathrm{PSPD}_0}(P)$, is defined as

$$\Gamma_{k,\ell}^{\mathrm{PSPD}_0}(P) \quad := \quad \mathsf{Dim}\left\{\mathrm{mult}\left(\mathbf{z}^{=\ell}\partial_{\mathbf{y}}^{=k}(P)\right)\right\}$$

where $\mathrm{mult}(f)$ is just the vector of coefficients of all *multilinear* monomials in f in a fixed predefined order.

An alternate way to interpret the above definition is to consider the shifted partial derivatives of P, and *reduce* them under the relation $x_i^2 = 0$, and only then list the coefficients of the surviving monomials. The rationale for this in [7] was to ensure that non-multilinear terms do not interact with multilinear terms in the shifted partial derivatives of P. Hence,

$$\Gamma_{k,\ell}^{\mathrm{PSPD}_0}(P) \quad = \quad \mathsf{Dim}\left\{\mathbf{z}^{=\ell}\partial_{\mathbf{y}}^{=k}(P) \mod \left\{x_i^2 \ : \ i \in [n]\right\}\right\}.$$

Another equally useful definition, which was also employed by Kumar and Saptharishi [9], is to reduce the shifted partial derivatives of P with respect to $x_i^2 = x_i$ instead. This also in

essence ensures that non-multilinear terms do not interact with the relevant multilinear terms by reducing their degree. We shall denote this by $\Gamma^{\mathrm{PSPD}_1}_{k,\ell}(P)$, which is formally defined to be

$$\Gamma^{\mathrm{PSPD}_1}_{k,\ell}(P) \quad := \quad \mathsf{Dim}\left\{\mathbf{z}^{=\ell}\partial^{=k}_{\mathbf{y}}(P) \mod \left\{(x_i^2 - x_i) \ : \ i \in [n]\right\}\right\}.$$

Since any polynomial f has a unique multilinear representation modulo $\left\{x_i^2 - x_i \ : \ i \in [n]\right\}$, it follows that its evaluations on $\{0,1\}^n$ completely determine the coefficients of the reduced polynomial $f \mod \left\{x_i^2 - x_i \ : \ i \in [n]\right\}$. Therefore, if $\Gamma^{\mathrm{PSPD}}_{k,\ell}(P)$ is defined as

$$\Gamma^{\mathrm{PSPD}_2}_{k,\ell}(P) \quad := \quad \mathsf{Dim}\left\{\mathsf{Eval}_{\{0,1\}^n}\left(\mathbf{z}^{=\ell}\partial^{=k}_{\mathbf{y}}(P)\right)\right\},$$

then it follows that

$$\Gamma^{\mathrm{PSPD}_2}_{k,\ell}(P) \quad = \quad \Gamma^{\mathrm{PSPD}_1}_{k,\ell}(P).$$

Finally, if P was set-multilinear with respect to $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_r$ and $\mathbf{y} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_k$, then all partial derivatives of order k with respect to $\mathbf{y}$ would be result in polynomials only in $\mathbf{z}$. Therefore for such set-multilinear polynomials,

$$\begin{aligned}\Gamma^{\mathrm{PSPD}_2}_{k,\ell}(P) &= \mathsf{Dim}\left\{\mathsf{Eval}_{\{0,1\}^n}\left(\mathbf{z}^{=\ell}\partial^{=k}_{\mathbf{y}}(P)\right)\right\} \\ &= \mathsf{Dim}\left\{\mathsf{Eval}_{\{0\}^{n_y}\times\{0,1\}^{n_z}}\left(\mathbf{z}^{=\ell}\partial^{=k}_{\mathbf{y}}(P)\right)\right\} \\ &=: \Gamma^{\mathrm{PSPD}}_{k,\ell}(P) \quad \text{as defined in Theorem 4.1.}\end{aligned}$$

The explicit polynomials for which we shall be show the lower bounds would indeed be set-multilinear and hence there is no loss incurred in restricting to only evaluations on $\{0\}^{n_y} \times \{0,1\}^{n_z}$.

For polynomials that are not set-multilinear, clearly

$$\begin{aligned}\Gamma^{\mathrm{PSPD}_2}_{k,\ell}(P) &= \mathsf{Dim}\left\{\mathsf{Eval}_{\{0,1\}^n}\left(\mathbf{z}^{=\ell}\partial^{=k}_{\mathbf{y}}(P)\right)\right\} \\ &\geq \mathsf{Dim}\left\{\mathsf{Eval}_{\{0\}^{n_y}\times\{0,1\}^{n_z}}\left(\mathbf{z}^{=\ell}\partial^{=k}_{\mathbf{y}}(P)\right)\right\} \quad =: \quad \Gamma^{\mathrm{PSPD}}_{k,\ell}(P).\end{aligned}$$

Hence for the purposes of upper-bounding $\Gamma^{\mathrm{PSPD}}_{k,\ell}()$ for say a term in the circuit computing P, taking fewer evaluations only helps.