Electronic Colloquium on Computational Complexity, Report No. 104 (2026)

Lower Bounds for Depth-5 Algebraic Circuits with Bounded Fan-in of Top Product Gates

Jules Armand * Amik Raj Behera† Sébastien Tavenas ‡

June 24, 2026

Abstract

We study depth-5 algebraic circuits over small finite fields with restricted fan-in of the top product gates. We show that there exists an explicit degree-d polynomial $P(\mathbf{x})$ such that any $\Sigma\Pi^{[\mathrm{poly}(d)]}\Sigma\Pi\Sigma$ circuit, computing $P(\mathbf{x})$, over a small finite field, requires size $2^{\Omega(\sqrt{d})}$. Our work builds upon and strengthens the work of [KS19], who showed $2^{\Omega(\sqrt{d})}$ lower bounds against $\Sigma\Pi^{[\mathcal{O}(\sqrt{d})]}\Sigma\Pi\Sigma$ circuits over small finite fields. It is known that proving $2^{\omega(d^{1/3}\log n)}$ lower bound for $\Sigma\Pi^{[a]}\Sigma\Pi$ circuits with $a = 2^{\mathcal{O}(d^{1/3}\log d)}$, over fields of characteristic 0, implies $\mathsf{VP} \neq \mathsf{VNP}$. In pursuit of this, we also prove superpolynomial lower bounds over small finite fields for $\Sigma\Pi^{[a]}\Sigma\Pi$ circuits where $a = 2^{\mathcal{O}(d^{\lambda}\log d)}$, for any constant $\lambda < 1/3$.

We use evaluations of the shifted partial derivatives to prove our lower bounds. We follow the same outline as [KS19], but with a more delicate analysis of the complexity measure. We use a family of the Nisan-Wigderson polynomials as a hard polynomial. We show that over small finite fields, setting the parameters of our measure and the hard polynomial with care, the method of shifted partial derivatives can yield lower bounds well beyond the homogeneity restriction on depth-4 circuits.

We also show an exponential gap between depth-3 and homogeneous depth-4 circuits over small finite fields. Previously, only a superpolynomial gap was known using [CM17] and depth reduction of polynomials in VP until homogeneous depth-4. We use the complexity measure of [GK98], and we use the Product of the Inner Product polynomial to show the separation.

*CNRS, LAMA, Université Savoie Mont Blanc, Chambéry, France. Email: jules.armand@univ-smb.fr

†Department of Computer Science, University of Copenhagen, Denmark. Supported by Srikanth Srinivasan's start-up grant from the University of Copenhagen. Most of this work was done when the author was visiting Laboratoire de Mathématiques (LAMA), Université Savoie Mont Blanc. The author thanks LAMA for their hospitality and encouraging research environment. Email: ambe@di.ku.dk

‡CNRS, LAMA, Université Savoie Mont Blanc, Chambéry, France. Email: sebastien.tavenas@univ-smb.fr

ISSN 1433-8092

Contents

1 Introduction

Algebraic complexity is the study of the computation of polynomials as formal objects with addition and multiplication as primitive arithmetic operations. In his seminal work, Valiant [Val79] introduced two classes, VP and VNP, which are algebraic analogues of the classes P and NP problem. The central question in algebraic complexity is to resolve the VP vs VNP problem - prove superpolynomial size lower bounds for an explicit polynomial. We refer the reader to surveys [SY10; Sap21] for a more formal treatment of the subject and especially for lower bounds. For general algebraic circuits, the best known lower bound is $\Omega(n \log n)$ by [BS83], which is slightly better than the best known general boolean circuit lower bound of $3.1n - o(n)$ by [LY22]. Over the past fifty years, strong lower bounds have been proved for various restricted algebraic circuits, and one particular restriction that has gained huge traction is *constant depth* circuits. Exponential lower bounds were proved for constant-depth Boolean circuits AC^0 in the late 80's by [FSS84; Hås87; Raz87; Smo87]. On the contrary, exponential lower bounds for algebraic circuits have remained deceptive. In this work, we consider some further restrictions on constant-depth circuits and prove lower bounds for them. We discuss more background and context for our result now.

In the following discussion, we will always use n and d to denote the number of variables and degree, respectively.

Known Lower Bounds in Algebraic Complexity Most of the lower bounds in algebraic complexity go via the method of *partial derivatives*. [Nis91] used it to prove an $\exp(\Omega(n))$ lower bound against the model of non-commutative algebraic branching programs[1] (the term 'partial derivatives' was used formally later). Subsequently, [NW97] formally introduced the method of partial derivatives to prove an $\exp(\Omega(n))$ lower bound against *homogeneous*[2] depth-3 circuits, over any field. [GK98; GR00] proved an $\exp(\Omega(n))$ lower bound against (non-homogeneous) depth-3 circuits over small finite fields. Proving lower bounds against depth-3 circuits over fields of characteristic 0 was the next challenge in sight. [SW01] proved a $\Omega(n^2)$ lower bound against depth-3 circuits over fields of characteristic 0. Later, [KST16a] proved a $\widetilde{\Omega}(n^3)$ lower bound for depth-3 circuits over any field. A breakthrough result of [LST25] proved $\exp(\omega(\log n))$ lower bounds against any constant-depth circuits over fields of characteristic 0 (also see [BDS24] for improved parameters). [For24] extended this result to any field. Proving stronger lower bounds for constant-depth circuits still remains a big challenge, and for a good reason, which we will explain in a moment.

[1] Algebraic branching programs are a class of algebraic circuits which is known to be at least as powerful as algebraic formulas and at most as powerful as algebraic circuits. The determinant polynomial, $\mathsf{DET}(\mathbf{x})$, can be computed by a $\mathrm{poly}(n)$-sized algebraic branching program and is also complete for this class under suitable reductions.

[2] For any circuit class $\mathcal{C}$, we say a circuit C is in homogeneous $\mathcal{C}$ if every gate of C computes a homogeneous polynomial. For homogeneous circuits, one can assume without loss of generality that every gate computes a polynomial of degree at most the degree of the output polynomial.

Depth Reduction (Efficient Parallelization) In the pursuit of resolving VP vs VNP, one would like to find structure within circuits in VP that can be exploited to prove lower bounds. As we saw above, we have lower bounds against constant-depth circuits. This begs the following natural question: *Can we reduce the depth of an arbitrary circuit 'efficiently'?*
[Bre74] showed that one can reduce the depth of any algebraic formula[3] of $\text{poly}(n)$-size to depth $\mathcal{O}(\log n)$ with only $\text{poly}(n)$ blowup in the size. [VSBR83] showed that any circuit in VP can be reduced to depth $\mathcal{O}(\log^2 n)$ with only $\text{poly}(n)$ blowup in the size. In a series of works, [AV08; Koi12; Tav15] showed that any circuit in VP (over any field) computing a degree-d polynomial can be converted to a *homogeneous* depth[4] 4 circuit with size[5] $\exp(\mathcal{O}(\sqrt{d}\log n))$.

From a lower bound perspective, the depth reduction results of [AV08; Koi12; Tav15] says that a lower bound of $\exp(\omega(\sqrt{d}\cdot\log n))$ against homogeneous depth-4 circuits for a degree-d polynomial in VNP implies VP $\neq$ VNP. In the next paragraph, we discuss the known lower bounds against small depth circuits.

Previous Lower Bounds against Homogeneous Depth-4 Circuits The last decade saw a flurry of progress in proving lower bounds against homogeneous depth-4 circuits. [Kay12; GKKS14] introduced a new complexity measure - *shifted partial derivatives*, which paved the way for exponential lower bounds against homogeneous depth-4 circuits. In an exciting line of work, [GKKS14; KSS14; FLMS15; KS14; KLSS14; KLSS17; KS17b] showed a lower bound of $\exp(\Omega(\sqrt{d}\log n))$ against homogeneous depth-4 circuits over any field, for a polynomial in VP. It seems this is close to the dream lower bound of $\exp(\omega(\sqrt{d}\cdot\log n))$ against homogeneous depth-4 circuits (which would imply VP $\neq$ VNP), yet there is no approach currently to go beyond this barrier!
Also note that the $\exp(\Omega(\sqrt{d}\log n))$ holds for a polynomial in VP, and this says that the depth-reduction result of [AV08; Koi12; Tav15] cannot be improved further for homogeneous depth-4 circuits.

Better Depth Reduction for Depth-4 Circuits With no improvement on lower bounds for homogeneous depth-4 circuits, we reconsider the model of computation and ask the following natural question: *Can we do better depth-reduction until (non-homogeneous) depth-4 circuits?*
Combining the results of [AV08; Koi12; Tav15] and [GKKS16] (for example, see [KS19, Section 7.5] for a proof), one can show that over fields of characteristic 0, every polynomial in VP can be computed by a depth-4 circuit with size $\exp(\mathcal{O}(d^{1/3}\log n))$, which is $\exp(o(\sqrt{d}\log n))$. Therefore, over fields of characteristic 0, non-homogeneity is more powerful than homogeneity

[3] An algebraic formula is a restriction of an algebraic circuit where the underlying circuit is a tree.
[4] Their results also hold for depth-Δ with size depending on Δ, but we focus on depth-4 for now.
[5] To understand the parameters, think of s, n, and d to be polynomially related to each other. We know that every degree-d polynomial in n variables can always be computed by a depth-2 $\Sigma\Pi$ circuit of size $\mathcal{O}(n^d)$, which is $\exp(\mathcal{O}(d\log d))$. However, the above-mentioned results say that if we go to depth-4, then it can always be computed by size $\exp(\mathcal{O}(\sqrt{d}\log d))$.

in depth-4 circuits[6]. The depth reduction result for depth-4 circuits says that a lower bound of $\exp(\omega(d^{1/3} \cdot \log n))$ against depth-4 circuits for a polynomial in VNP, over fields of characteristic 0, would imply VP $\neq$ VNP. This depth reduction result until depth-4 circuits is not known for fields of positive characteristic[7].

Previous Lower Bounds against Depth-4 Circuits Proving lower bounds against depth-4 circuits has been more challenging than homogeneous depth-4 circuits, and has seen a gradual progress. [Sha17] proved a $\widetilde{\Omega}(n^{1.5})$ lower bound and then [GST20] proved a $\widetilde{\Omega}(n^{2.5})$ lower bound for depth-4 circuits over fields of characteristic 0 for a polynomial in VNP. [LST25; For24] show superpolynomial lower bounds against depth-4 circuits over any field, for low-degree polynomials. More precisely, for $d = \mathcal{O}(\log n)$, they show a lower bound of $\exp(\Omega(d^{1/4} \log n))$.

Before we proceed to explain our motivation and our results, it will be convenient to state the following notation. For a parameter a, we use $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ to denote depth-5 circuits where the top product gate has fan-in a, and similarly $\Sigma\Pi^{[a]}\Sigma\Pi$.

1.1 Motivation

From the above discussion, we know that a lower bound of $\exp(\omega(d^{1/3} \log n))$ against depth-4 circuits over fields of characteristic 0, implies VP $\neq$ VNP. More specifically, it suffices to prove a lower bound of $\exp(\omega(d^{1/3} \log n))$ against $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit where $a = \exp(\mathcal{O}(d^{1/3}))$ over fields of characteristic 0. In this work, we consider the setting of small finite fields instead of characteristic 0 because small finite fields seem to be more amicable towards our known lower bound techniques. [GK98; GR00] proved $\exp(\Omega(d))$ lower bounds for depth-3 circuits over small finite fields and [CM17] improved the lower bound to $\exp(\Omega(d \log n))$. [LST25; For24] achieve a lower bound of $\exp(\Omega(d^{1/4} \log n))$ against depth-4 circuits over every field. The difference between the lower bounds obtained for homogeneous circuits compared to general circuits seems to stem from the fact that in the former models, the formal degree is bounded by d, i.e., we can assume that all intermediate polynomials calculated have degree at most d. In this work, we seek an intermediate goal and investigate the following question:

Are there natural restrictions of depth-4 circuits allowing large formal degrees for which we can prove exponential lower bounds, over small finite fields?

To motivate our line of investigation, we look at few concrete examples:

[6] This can be formally captured as follows. The above-mentioned $\exp(\Omega(\sqrt{d} \log n))$ lower bounds against homogeneous depth-4 circuits hold for a polynomial in VP. In other words, there is a polynomial in VP that cannot be computed by a homogeneous depth-4 circuit of size $\exp(o(\sqrt{d} \log n))$, but can be computed by a depth-4 circuit of size $\exp(\mathcal{O}(d^{1/3} \log n))$, over fields of characteristic 0.

[7] Characteristic 0 shows up in the proof of [GKKS16]. There is a step to express any power of a linear form as a sum of product of univariate polynomials. This conversion is only known over fields of characteristic 0.

- Over small finite fields, [KS19, Theorem 1.2] went beyond the homogeneity condition and showed an $\exp(\Omega(\sqrt{d}))$ lower bound against $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ where $a = \mathcal{O}(\sqrt{d})$. Their proof (inspired from [GK98]) crucially uses small finite fields to "nullify" bottom product gates with large degree, and the circuit "behaves" like a homogeneous depth-5 circuit.

- Over fields of characteristic 0, [SW01] proved $\Omega(n^2)$ lower bounds against (non-homogeneous) depth-3 circuits. To handle non-homogeneity, they showed that one can "nullify" product gates of high rank[8]. Thus it suffices to prove a lower bound against $\Sigma\Pi^{[a]}\Sigma$ circuits where $a \leqslant n/100$.

- Over any field, another series of works follow this idea to extend lower bounds to models allowing larger formal degree. In [KS17a; KST16b], the authors generalized lower bounds for multilinear models [Raz09] (a model is multilinear if any gate computes a multilinear polynomial) to multi-r-ic ones (a model is multi-r-ic if the formal degree of any gate in each variable is at most r).

- Over fields of characteristic 0, [FLST24, Theorem 5] showed that a depth-4 circuit computing a polynomial of degree-d polynomial $P(\mathbf{x})$ can be efficiently converted to a depth-4 circuit of formal degree $\text{poly}(d)$, computing $P(\mathbf{x})$. Here, efficient means a blowup of $\text{poly}(d)$.

The above mentioned results suggest that to handle (non-homogeneous) $\Sigma\Pi\Sigma\Pi$ circuits, a possible approach is to reduce to $\Sigma\Pi^{[a]}\Sigma\Pi$ circuits where $a = \text{poly}(d)$. In this work, we prove $\exp(\Omega(\sqrt{d}))$ lower bounds against $\Sigma\Pi^{[\text{poly}(d)]}\Sigma\Pi$ circuits over small finite fields.

We also revisit the complexity measure used in [GK98] to prove lower bounds against depth-3 circuits over small finite fields. [GK98] showed an $\exp(\Omega(n))$ lower bounds against depth-3 circuits over small finite fields, for a polynomial in VP. Later, using the same complexity measure, [CM17] improved the lower bound to $\exp(\Omega(n \log n))$, and [Sap21, Theorem 10.2] showed that the same lower bound holds even for the elementary symmetric polynomial[9]. These results using depth reduction of [AV08; Koi12; Tav15] show that polynomials in VP cannot be reduced until depth-3, akin to [GKKS16]. Our motivation is to understand whether the complexity measure of [GK98] could be used to prove exponential lower bounds against depth-4 circuits with a bound on top-product fan-in. We answer this negatively by achieving an $\exp(\Omega(n \log n))$ lower bound against depth-3 circuits over small finite fields for a polynomial that can even be computed by a $\text{poly}(n)$-sized homogeneous depth-4 circuit. We state it formally in Theorem 1.5.

1.2 Our Results

In this subsection, we discuss our main results. We prove our lower bounds for Nisan-Wigderson polynomials $\text{NW}_{d,m,e}$ (see Definition 3.1 for a formal definition) and for Product

[8] Here, the rank of a product gate refers to the rank of the affine polynomials that are multiplied together.

[9] In a beautiful construction, Ben-Or showed that elementary symmetric polynomials have $\mathcal{O}(n^2)$-sized depth-3 circuits over fields with $\geqslant (n+1)$ elements. However, for small finite fields, we are not aware of any better construction than depth-reduction of VP until homogeneous depth-4 circuits.

of Inner Products $\mathsf{PIP}_{m,d}$ (see Definition 3.2 for a formal definition). For $\mathsf{NW}_{m,d,e}$, the number of variables n is md and for $\mathsf{PIP}_{m,d}$, the number of variables n is $2md$.

Our first result is an exponential lower bound against depth-4 circuits with bounded fan-in on top-product gates, over small finite fields. Our techniques also extend naturally to depth-5 circuits with bounded fan-in on top-product gates, and thus we state our main result for depth-5.

Theorem 1.1 (Lower Bounds for Restricted Depth-5). *Fix a field $\mathbb{F}_q$. Let $n, d \in \mathbb{N}$ be growing parameters, where $d = \mathcal{O}(\sqrt{n})$. Let $\alpha \in \mathbb{N}$ be a parameter verifying $0.01 \leqslant \alpha \leqslant d^{1/3}$. There exists $m, e \in \mathbb{N}$ with $m = \Theta(d^{\alpha+1})$, $n = md$ such that the following holds:*

Any $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit with $a = \lceil d^{\alpha+0.5} \rceil$, over $\mathbb{F}_q$, computing the degree-d polynomial $\mathsf{NW}_{d,m,e}(\mathbf{x})$ requires size $\exp(\Omega_q(\sqrt{d/\alpha}))$.

Remark 1.2. *Firstly, the constant 0.01 is chosen for simplicity. Our proof can easily be extended to $C \leqslant \alpha$ where $C > 0$ is any absolute constants.*
Second, we bound α by $d^{1/3}$, even if the proof does not require it, but form that point on, the lower bound becomes trivial. The reason is that the number of variables in $\mathsf{NW}_{d,m,e}$ is $md = \Theta(d^{\alpha+2})$ and for $\alpha \geqslant d^{1/3}$, n is more than the lower bound. However, if $\alpha = d^{\lambda}$ for some positive constant $\lambda < 1/3$, we get superpolynomial lower bounds. We state it in Corollary 1.4.

Theorem 1.1 achieves the same lower bound as [KS19], but for a broader class of depth-5 circuits[10], which we state in the following corollary.

Corollary 1.3 (Exponential Lower Bounds for Restricted Depth-5). *Fix a field $\mathbb{F}_q$ and let $C > 0$ be an absolute constant. Let $n, d \in \mathbb{N}$ be growing parameters, where $d = \mathcal{O}(n^{1/C})$. There exists $m, e \in \mathbb{N}$ such that the following holds:*

Any $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit with $a = \lceil d^{C+0.5} \rceil$, over $\mathbb{F}_q$, computing the degree-d polynomial $\mathsf{NW}_{d,m,e}(\mathbf{x})$ requires size $\exp(\Omega_q(\sqrt{d}))$.

In characteristic 0, the threshold for $a = \exp(\mathcal{O}(d^{1/3}))$ is crucial since depth reduction of VP until depth-4 yields $\Sigma\Pi^{[a]}\Sigma\Pi$ circuits of size $\exp(\mathcal{O}(d^{1/3} \log n))$. We get superpolynomial lower bounds in the setting when the top product gates have fan-in d^{λ} for some constant $\lambda < 1/3$.

[10] In a homogeneous depth-5 circuit, the top product gates have fan-in at most d.

Corollary 1.4 (Superpolynomial Lower Bounds for Restricted Depth-4). *Fix a field* $\mathbb{F}_q$ *and let* $n, d \in \mathbb{N}$ *be growing parameters, where* $d \leqslant \widetilde{\mathcal{O}}(\log^{3.1} n)$. *Let* $\alpha = \mathcal{O}(d^{0.32})$. *There exists* $m, e \in \mathbb{N}$ *such that the following holds:*

Any $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ *circuit with* $a = \lceil d^{\alpha} \rceil = \exp(\mathcal{O}(d^{0.32} \log d))$, *over* $\mathbb{F}_q$, *computing the degree-*d *polynomial* $\mathsf{NW}_{d,m,e}(\mathbf{x})$ *requires size* $\exp(\omega(\log^{1.01} n))$.

Note that lower bound of [LST25; For24] for depth-4 circuits is $\exp(\Omega(d^{1/4} \log n))$ but they require $d \leqslant \mathcal{O}(\log n)$. Corollary 1.4 is a lower bound for a restriction of depth-4 circuits, but it holds for a wider range of degree parameters, i.e., it continues to hold even for $d = \mathcal{O}(\log^3 n)$.

Our next result is an exponential separation between depth-3 and homogeneous depth-4 circuits over small finite fields. Note that $\mathsf{PIP}_{m,d}$ can be computed by a homogeneous depth-4 circuit of size $\mathcal{O}(md)$, which just follows from the definition.

Theorem 1.5 (Exponential Gap b/w Depth-3 and Homogeneous Depth-4). *Fix a field* $\mathbb{F}_q$ *and let* $m, d \in \mathbb{N}$ *be growing parameters. Any* $\Sigma\Pi\Sigma$ *circuit over* $\mathbb{F}_q$ *computing the degree-*d *polynomial* $\mathsf{PIP}_{m,d}$ *requires size* $\exp(\Omega_q(d \log m))$.

The previous best gap between depth-3 circuits and homogeneous depth-4 circuits over small finite fields was a superpolynomial gap (this follows immediately from lower bound of [CM17] and depth-reduction results of VP until homogeneous depth-4). Theorem 1.5 gives an exponential separation between depth-3 and homogeneous depth-4 over small finite fields.

2 Technical Overview

We start by stating the usual framework for proving algebraic circuit lower bounds. Let $\mathcal{C}$ denote the circuit class against which we want to prove a lower bound. The usual framework has the following key steps:

1. Choose a sub-additive complexity measure Γ, i.e., $\Gamma(T_1 + T_2) \leqslant \Gamma(T_1) + \Gamma(T_2)$.
2. For the class $\mathcal{C}$, identify the "building blocks". For example, in $\Sigma\Pi\Sigma\Pi$ circuits, a building block is a $\Pi\Sigma\Pi$ circuit, since a small depth-4 circuit is a small sum of such building blocks.
3. Show that any polynomial $T(\mathbf{x})$ that can be computed "efficiently" by a building block, has small measure under Γ. Say $\Gamma(T) \leqslant m$ for every polynomial $T(\mathbf{x})$ that can be computed by a building block of size s. Using sub-additivity of Γ, we get that if f has a size s circuit in the class $\mathcal{C}$, then $\Gamma(f) \leqslant s \cdot m$.
4. Find an explicit polynomial $P(\mathbf{x})$ such that it has large measure under Γ. Say $\Gamma(P) \geqslant M$. Then, using the previous item, any circuit from the class $\mathcal{C}$ computing the polynomial P

must have size $s \geqslant M/m$.

Now we use the above outline to discuss our proof idea and our technical contribution. For simplicity, we first discuss the proof idea for depth-4 circuits (see Theorem 4.1 for a formal statement), and then we will discuss how to prove Theorem 1.1 using a few technical modifications.

We consider the class $\mathcal{C}$ to be $\Sigma\Pi^{[a]}\Sigma\Pi$ where $a < d^{\beta}$ for some parameter β (we will use $\beta = \alpha+1$ in our proof where α is the parameter from Theorem 1.1). We use the same complexity measure Γ as used in [KS19], defined as,

$$\Gamma(f) \;:=\; \dim\left[\mathrm{Eval}_S\left(\mathbf{x}^{=\ell}\left(\partial^{=k}(f)\right)\right)\right], \qquad \text{for some choice of } k, \ell \text{ and } S \subseteq \mathbb{F}_q^n \;.$$

The above measure is a "functional" variant of the shifted partial measure, i.e., we interpret the polynomials after shifting the partial derivatives, as functions on the subset $S \subseteq \mathbb{F}_q^n$. It is easy to convince oneself that Γ is a sub-additive measure. Note that for a degree-d polynomial $f(\mathbf{x})$, $\Gamma(f)$ is at most the dimension of the space of $(\ell + d - k)$-degree polynomials. This is the first step according to the outline.

Before discussing the second and third steps of the framework, we jump ahead and discuss the fourth step, i.e., to find an explicit "hard" polynomial with large complexity measure. We use the Nisan-Wigderson polynomials $\mathsf{NW}_{d,m,e}$ as our hard polynomial. [KS19] showed that for a careful choice of k, ℓ, S, d, m, e, the measure $\Gamma(\mathsf{NW}_{d,m,e})$ is close to the maximum value. See Section 3.4 for the precise lower bound. Now we focus on the upper bound part of the outline - our main contribution. Before discussing how we prove an upper bound on the complexity measure, we take a brief detour to the prior works to get a slightly better understanding. We also need the following notation: For a parameter r, $\Sigma\Pi\Sigma\Pi^{\{r\}}$ denotes depth-4 circuits where the bottom product gates have support at most r, i.e., every bottom product computes a polynomial of support at most r (for a polynomial, its support is defined as the maximum number of variables in any of its monomials).

Upper Bound of Previous Works Building on a series of works [FLMS15; KS14; KLSS14; KLSS17; KS17b], we have exponential lower bounds against $\Sigma\Pi\Sigma\Pi$ circuits. We also have exponential lower bounds against homogeneous $\Sigma\Pi^{[\mathcal{O}(\sqrt{d})]}\Sigma\Pi\Sigma$ circuits over finite fields by [KS19]. We present a silhouette of upper bounds in the above-mentioned works.

Silhouette of Upper Bounds in Prior Works

Let $\mathcal{C}$ denote the class of circuits for which we would like to prove a lower bound, and let μ be our choice of complexity measure. The upper bound is a two-step process:

- Find a more *structured* class $\mathcal{C}'$ against whom we can prove strong lower bounds using the measure μ. This requires showing a "small" upper bound on the $\mu(f')$ where f' is computed by a small circuit in $\mathcal{C}'$.
- Show that the measure μ of $\mathcal{C}$ and $\mathcal{C}'$ are essentially the same, i.e., for every polynomial f that has a small circuit in $\mathcal{C}$, there exists another polynomial f' that has a small circuit in $\mathcal{C}'$ such that $\mu(f)$ and $\mu(f')$ are roughly the same.

The choices of the complexity measure μ and a more structured class $\mathcal{C}'$ are interrelated. We want a class $\mathcal{C}'$ against which we can prove a strong lower bound, but also $\mathcal{C}'$ should be strong enough such that μ cannot practically distinguish it from the class $\mathcal{C}$. Now we discuss the upper bounds of [FLMS15; KS14; KLSS14; KLSS17; KS17b] on homogeneous depth-4 in slightly more detail. To draw an analogy from the previous paragraph, let $\mathcal{C}$ denote the class of homogeneous depth-4 circuits and let μ be a close variant of the shifted partial derivative (the exact description of μ is not required here).

- The more structured class $\mathcal{C}'$ is the class of homogeneous depth-4 circuits with bounded fan-in on bottom product gates, i.e., homogeneous $\Sigma\Pi\Sigma\Pi^{[\mathcal{O}(\sqrt{d})]}$. Suppose f is computed by a small homogeneous $\Sigma\Pi\Sigma\Pi$ circuit. They used bounded support and homogeneity to prove small upper bounds on the measure. This lets them prove strong lower bounds against homogeneous $\Sigma\Pi\Sigma\Pi^{[\mathcal{O}(\sqrt{d})]}$ circuits.
- To show the second item from the previous paragraph, they use *random restriction*, where a subset of variables is set to 0 at random[11]. Since a random restriction only sets a subset of variables to 0, this does not affect the complexity measure μ. Finally, they showed that with high probability (over the random restriction), the resulting polynomial can be computed by a small homogeneous $\Sigma\Pi\Sigma\Pi^{[\mathcal{O}(\sqrt{d})]}$ circuit.

Our Upper Bound Now we return to discussing our upper bound on the measure of $\Sigma\Pi^{[a]}\Sigma\Pi$ circuits. We use the same complexity measure as in [KS19] (see Definition 3.11 for a formal definition) and we will denote it by $\Gamma_{k,\ell,S}$ for some choice of k, ℓ, and $S \subseteq \mathbb{F}_q^n$. The choice of the subset S will be crucial. Similar to the abstraction mentioned above, our upper bound also has the following two steps:

[11] The variables are set to 0 with a carefully designed random process. One has to be careful here since the random restriction might make the "hard" polynomial a very simple polynomial to compute. So the random process has to be chosen in a way such that even after the random process, the "hard" polynomial continues to have a large complexity measure with high probability.

1. The more structured class $\mathcal{C}'$ is the class of $\Sigma\Pi^{[a]}\Sigma\Pi^{\{R\sqrt{d}\}}$ circuits for some parameter R, i.e., the bottom product gates have bounded support. For this, we show that one could still prove a "small"[12] upper bound under the measure Γ, even though the circuit is no longer homogeneous. The parameter R is chosen carefully. To get a small upper bound, R will be upper-bounded by some function of α and also of ℓ (here ℓ denotes the degree of our shifts). See Lemma 4.2 for a formal statement.

2. We then show that the measure Γ does not distinguish between $\mathcal{C}$ and $\mathcal{C}'$ as stated in the second item of the silhouette. For this, the choice of S comes into the picture. The subset S is chosen so that every bottom product gate with large support vanishes on S, and all of its k^{th} order partial derivatives also vanish on S. This ensures that the bottom product gates with large support do not affect the measure Γ. Vaguely speaking, this choice of S can be thought of as a small finite field analogue of the random restriction in the previous works. For this step to also hold, we need the bottom support $R\sqrt{d}$ to be at least a constant factor of k. See Lemma 4.3 for a formal statement.

We show that there is a choice of the parameters such that the above idea works out and yields a lower bound of $\exp(\Omega(\sqrt{d}))$.

To extend our lower bounds for $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$, we only need to prove a similar upper bound on the complexity measure as for $\Sigma\Pi^{[a]}\Sigma\Pi$. For this, we again proceed as mentioned above, with a small modification in how we choose the subset S. In depth-4, we chose S such that for every bottom product gate with large support, all of its k^{th} order partial derivatives should vanish on S. Now, instead of bottom product gates with large support, we consider bottom product gates with "large rank", where rank is measured as the rank of the affine polynomials that are multiplied together. See Lemma 5.3. The rest of the argument remains the same.

Comparison with [KS19] In [KS19], the authors aim to obtain a lower bound of the form $\exp(\Omega(\sqrt{n}))$ for homogeneous depth-5 circuits over small finite fields. Under these conditions, the authors show that they can restrict attention to circuits in which the top gates have fan-in bounded by $\mathcal{O}(\sqrt{d})$. In this paper, we want to go beyond the homogeneity constraint and instead allow circuits of large formal degree. The fan-in bound on the top product gates then becomes a parameter α of our problem, and we aim to obtain lower bounds that remain superpolynomial as α grows. Our approach is a careful analysis and an adaptation of the proof of [KS19] in our settings.

Proof Idea for Theorem 1.5 To show that $\mathsf{PIP}_{m,d}$ requires $\exp(\Omega(n))$-sized depth-3 circuits over small finite fields, we use the complexity measure introduced in [GK98]. The measure is the dimension of the evaluation space of partial derivatives on $\mathbb{F}_q^m \setminus \mathcal{E}$, for an exponentially small subset $\mathcal{E}$. To show that polynomials computable by small depth-3 circuits have small measure, one can "nullify" product gates that involve affine polynomials with large rank. This

[12] Here, "small" is in comparison to the measure for the hard polynomial

is the reason we remove a small subset $\mathcal{E}$ from $\mathbb{F}_q^n$. Then it is simple to show that the partial derivative space is spanned by few affine polynomials and thus have small dimension. On the other hand, $\mathsf{PIP}_{m,d}$ polynomial has the property that its d^{th} order partial derivatives contain m^d distinct monomials, and thus they are all linearly independent. In other words, the dimension is at least m^d. Since the partial derivatives are small degree polynomials, removing a small subset $\mathcal{E}$ from the set of evaluations does not affect the dimension.

Organization

In Section 3, we give the necessary preliminaries. In Section 4, we prove Theorem 4.1, which is a special case of Theorem 1.1. Since the proof is slightly simpler in the case of depth-4 circuits than in the case of depth-5 circuits, we present the proof of Theorem 4.1 as a warm-up to the proof of Theorem 1.1. In fact, we do most of the heavy calculations in the proof of Theorem 4.1 (more specifically in Section 4.3). In Section 5, we prove Theorem 1.1. We recommend the reader to read Section 4 before going to Section 5. In Section 6, we prove Theorem 1.5 (it is independent of Section 4 and Section 5).

3 Preliminaries

Let $n, d \in \mathbb{N}$ be growing parameters, where n denotes the number of variables and d denotes the degree. We will work over the field $\mathbb{F}_q$, where q is independent of n and d, i.e. $q = \mathcal{O}(1)$. We also have parameters $a, k, \ell, s, r \in \mathbb{N}$ that are functions depending on q, n, and d. Throughout the article, we will use a to denote the fan-in of the top product gates, k to denote the order of partial derivatives we consider for the complexity measure, ℓ to denote the degree of the shift we consider for the complexity measure, r to denote the support/rank of the bottom product gates, and s to denote the size of the circuits. We will use $\mathbf{x}$ to denote the set of underlying variables $\{x_1, \ldots, x_n\}$. We will use $\mathcal{O}_q(\cdot)$ and $\Omega_q(\cdot)$ to hide the dependencies on q.

In the following list, let $P \in \mathbb{F}_q[x_1, \ldots, x_n]$ be an arbitrary polynomial.

- For a monomial $\mathbf{m} = x_1^{e_1} \cdots x_n^{e_n}$, we use the shorthand notation $\partial P/\partial \mathbf{m}$ to denote the following polynomial:

$$\frac{\partial P}{\partial \mathbf{m}} = \frac{\partial}{\partial x_n^{e_n}} \left(\frac{\partial}{\partial x_n^{e_n}} \left(\cdots \left(\frac{\partial P}{\partial x_1^{e_1}} \right) \right) \right).$$

- For a parameter $k \in \mathbb{N}$, we will use $\partial^{=k}(P)$ to denote the space (over $\mathbb{F}_q$) of all k^{th} order partial derivatives of $P(\mathbf{x})$. We will use $\partial^{\leqslant k}(P)$ to denote the space of all partial derivatives of $P(\mathbf{x})$ of order $\leqslant k$.

- For a parameter $\ell \in \mathbb{N}$, we will use $\mathbf{x}^{\ell} \cdot P$ to denote the space (over $\mathbb{F}_q$) spanned by $\mathbf{m} \cdot P$ over all degree ℓ monomials $\mathbf{m}$.

- For a subset $S \subseteq \mathbb{F}_q^n$, we will use $\mathrm{Eval}_S(P)$ to denote the evaluation vector of the polynomial $P(\mathbf{x})$ over the points of S (we will fix some arbitrary ordering on the points of S).
- The *support* of $P(\mathbf{x})$, denoted by $\mathrm{Supp}(P)$, is defined as the maximum support size over all monomials with a non-zero coefficient in $P(\mathbf{x})$. In other words, a polynomial $P(\mathbf{x})$ with $\mathrm{Supp}(P) \leqslant r$ implies that every monomial that has a non-zero coefficient in $P(\mathbf{x})$ depends on $\leqslant r$ variables (distinct monomials can depend on distinct sets of $\leqslant r$ variables).
- For a set of vectors $\{v^{(1)}, \ldots, v^{(t)}\}$, we will use $\mathrm{Span}\{v^{(1)}, \ldots, v^{(t)}\}$ to denote their span (over $\mathbb{F}_q$) and $\dim\{v^{(1)}, \ldots, v^{(t)}\}$ to denote the dimension of their span.

Depth-4 and Depth-5 Circuits In this article, we will be dealing with depth-4 circuits and depth-5 circuits with some restrictions. Both depth-4 and depth-5 have two layers of product gates and we will refer to them as *top product* and *bottom product* gates. More specifically, the product gates at the lower layer (closer to the input gates) will be referred to as bottom product gates, and the product gates at the upper layer (closer to the output gate) will be referred to as top product gates.

If $f(\mathbf{x})$ is computed by a depth-4 circuit $\Sigma\Pi\Sigma\Pi$ of size s, then we can express $f(\mathbf{x})$ as:

$$f(\mathbf{x}) \;=\; T_1(\mathbf{x}) + \ldots + T_s(\mathbf{x}),$$

where every $T_i(\mathbf{x}) = Q_1^{(i)}(\mathbf{x}) \cdots Q_s^{(i)}(\mathbf{x})$. We will refer to each $T_i(\mathbf{x})$ as a *term*, and every polynomial $Q_j^{(i)}(\mathbf{x})$ as an *inner polynomial*. By definition, each $Q_j^{(i)}(\mathbf{x})$ is computed by a $\Sigma\Pi$ circuit of size s. We extend this to depth-5 similarly, where every $Q_j^{(i)}(\mathbf{x})$ is computed by a $\Sigma\Pi\Sigma$ circuit of size s. For a parameter r, we will use $\Sigma\Pi\Sigma\Pi^{\{r\}}$ to denote the class of depth-4 circuits where the bottom product gates have support bounded by r, i.e., the support of the polynomial computed at every bottom product is $\leqslant r$.

Nisan-Wigderson Polynomials Now we define the family of Nisan-Wigderson polynomials, for which we will prove lower bounds against depth-5 circuits.

Definition 3.1 (Nisan-Wigderson Polynomials). *Let $d, m, e \in \mathbb{N}$ be parameters with $d, e \leqslant m$ and m is a prime power. We identity $[m]$ with $\mathbb{F}_m$. The* Nisan-Wigderson *polynomial with parameters d, m, e, denoted by $\mathsf{NW}_{d,m,e}$ is defined as follows:*

$$\mathsf{NW}_{d,m,e}(\mathbf{x}) \;:=\; \sum_{\substack{R(t) \in \mathbb{F}_m[t] \\ \deg(R) < e}} x_{1,R(1)} \cdots x_{d,R(d)}.$$

In other words, $\mathsf{NW}_{d,m,e}$ has one monomial for every univariate polynomial $R(t)$ of degree strictly less than e, where the monomial is the product of variables corresponding to d points on the curve of $R(t)$.

The family of Nisan-Wigderson polynomials for every choice of $\{d, m, e\}$ is *explicit* in the

sense that they belong to the class VNP. To see this, [Val79, Proposition 4] proved a sufficient condition for a polynomial to be in VNP if any of its coefficients can be computed efficiently (also see [Sap21, Fact 1.2]). For $\mathsf{NW}_{d,m,e}$, one can compute the coefficient (which is simply 0 or 1) by checking if the indices of a given monomial lie on a degree $< e$ univariate polynomial or not, and this can be efficiently done by univariate interpolation.

Product of Inner Products Polynomials We also define the Product of Inner Product polynomials, which we will use to prove lower bounds against depth-3 circuits.

Definition 3.2 (Product of Inner Products). *Let* $m, d \in \mathbb{N}$. *The* Product of Inner Products, $\mathsf{PIP}_{m,d}$ *is,*

$$\mathsf{PIP}_{m,d}(\mathbf{x}, \mathbf{y}) := \prod_{i=1}^{d} \sum_{j=1}^{m} x_{ij} \cdot y_{ij}.$$

By definition, for every choice of $\{m, d\}$, the polynomial $\mathsf{PIP}_{m,d}$ is computed by a homogeneous depth-4 circuit of size $O(md)$.

3.1 Rank and Multiplicity

We define the notion of the rank of a product gate.

Definition 3.3 (Rank of a Product Gate). *Let* $C(\mathbf{x})$ *be an algebraic circuit over* $\mathbb{F}_q$ *and let* g *be a product gate in* $C(\mathbf{x})$ *with children* $h_1(\mathbf{x}), \ldots, h_t(\mathbf{x}) \in \mathbb{F}_q[x_1, \ldots, x_m]$*, i.e., the polynomial computed at the node* g *is* $h_1(\mathbf{x}) \cdots h_t(\mathbf{x})$*. Then* rank of product gate g *is defined as the* $\mathbb{F}_q$*-rank of* $\{h_1(\mathbf{x}), \ldots, h_t(\mathbf{x})\}$.

According to Definition 3.3, in a depth-4 circuit, if a bottom product gate has rank r, then its support is of size r. In a depth-5 circuit, if a bottom product gate has rank r, then its children (which are affine polynomials) have rank r.

Definition 3.4 (Multiplicity of a point). *Let* $Q(x_1, \ldots, x_n) \in \mathbb{F}_q[\mathbf{x}]$ *be a polynomial and let* $\mathbf{a} \in \mathbb{F}_q^n$ *be any point. The point* $\mathbf{a}$ *is said to have* multiplicity *at least* t *with respect to the polynomial* $Q(\mathbf{x})$ *if for every polynomial* $h(\mathbf{x}) \in \partial^{<t}(Q(\mathbf{x}))$*, we have* $h(\mathbf{a}) = 0$.
We use $\mathrm{Mult}(Q, \mathbf{a})$ *to denote the largest number* t *such that* $\mathbf{a}$ *has multiplicity* t *with respect to the polynomial* $Q(\mathbf{x})$.

We prove that if a set of degree-1 polynomials has sufficiently large rank, then a dense subset of $\mathbb{F}_q^n$ vanishes on the product of those degree-1 polynomials and their derivatives up to some order.

In other words, a dense subset of $\mathbb{F}_q^n$ has significant multiplicity with respect to the product. We state it formally below.

Claim 3.5 (Product of Large Rank Degree-1 Polynomials). *Let* $\{L_1(\mathbf{x}), \ldots, L_t(\mathbf{x})\}$ *be a set of degree-1 polynomials with rank larger than* r*. Let* $H(\mathbf{x}) := L_1(\mathbf{x}) \cdots L_t(\mathbf{x})$*. Then for every* $\kappa > 0$*, we have,*

$$\Pr_{\mathbf{a} \sim \mathbb{F}_q^n} \left[\mathrm{Mult}(H, \mathbf{a}) < (1-\kappa) \cdot \frac{r}{q} \right] \leqslant \exp(-\kappa^2 r/2q).$$

Proof of Claim 3.5. Assume without loss of generality that $\{L_1(\mathbf{x}), \ldots, L_r(\mathbf{x})\}$ is a $\mathbb{F}_q$-linearly independent set of non-constant functions. Each non-constant degree-1 polynomial L_i is 0 with probability exactly $1/q$ at a random $\mathbf{a} \in \mathbb{F}_q^n$. For every $i \in [r]$, let $X_i : \mathbb{F}_q^n \to \{0,1\}$ be the random variables denoting whether the L_i vanishes or not: $X_i(\mathbf{a}) = 1$ if and only if $L_i(\mathbf{a}) = 0$. Let $X := X_1 + \ldots + X_r$. We have

$$\mathbb{E}_{\mathbf{a} \sim \mathbb{F}_q^n}[X_i] = \frac{1}{q} \quad \text{and so} \quad \mathbb{E}_{\mathbf{a} \sim \mathbb{F}_q^n}[X] = \frac{r}{q}.$$

Since $\{L_1, \ldots, L_r\}$ is $\mathbb{F}_q$-linearly independent, the random variables $X_1, \ldots, X_r$ are independent.

Observe that for any $\mathbf{a} \in \mathbb{F}_q^n$, if $\mathrm{Mult}(H, \mathbf{a}) < t$, then the number of factors of H vanishing at $\mathbf{a}$ is less than t. So it implies that the number of random variables amongst $X_1, \ldots, X_r$ which are set to 1 at $\mathbf{a}$ is also less than t. Using the Chernoff bound, we get

$$\Pr_{\mathbf{a} \sim \mathbb{F}_q^n} \left[\mathrm{Mult}(H, \mathbf{a}) < (1-\kappa) \cdot \frac{r}{q} \right] \leqslant \Pr_{\mathbf{a}} \left[X(\mathbf{a}) < (1-\kappa) \cdot \frac{r}{q} \right] \leqslant \exp(-\kappa^2 r/2q).$$

■

If the degree-1 polynomials are simply variables, then we get the following corollary immediately, which will be useful while studying depth-4 circuits.

Corollary 3.6 (High Multiplicity of Large Support Monomials). *Let* m *be a monomial with* $|\mathrm{Supp}(m)| \geqslant r$*. Then for every* $\kappa > 0$*, we have,*

$$\Pr_{\mathbf{a} \sim \mathbb{F}_q^n} \left[\mathrm{Mult}(m, \mathbf{a}) < (1-\kappa) \cdot \frac{r}{q} \right] \leqslant \exp(-\kappa^2 r/2q).$$

3.2 Multilinearization

The complexity measure will involve evaluations over a translation of the Boolean cube. In the next claim, we show that a monomial of low support but potentially high degree can be replaced by a low-degree monomial that agrees on a translation of the Boolean cube. The following claim

was sketched in [KS19, Lemma 4.7], but we give a complete proof for the sake of completeness.

Claim 3.7. *Let $n \in \mathbb{N}$ and fix $\mathbb{F}$ to be an arbitrary field. Fix an arbitrary $\mathbf{c} \in \mathbb{F}^n$. Let m be a monomial in $\mathbb{F}[x_1, \ldots, x_n]$ of arbitrary degree. There exists a multilinear polynomial $P_m(\mathbf{x})$ of degree $|\mathrm{Supp}(m)|$ and sparsity $2^{|\mathrm{Supp}(m)|}$ such that for every $\mathbf{a} \in \{0,1\}^n$,*

$$P_m(\mathbf{c} + \mathbf{a}) = m(\mathbf{c} + \mathbf{a}).$$

In other words, $P_m(\mathbf{x})$ agrees with the monomial m on $\mathbf{c} + \{0,1\}^n$, i.e. a translate of the Boolean cube.

Remark 3.8. *In Claim 3.7, suppose the underlying field $\mathbb{F} = \mathbb{F}_q$. In this case, a natural strategy to replace a monomial m with low-support but potentially of high degree is to use the identity $z^q = z$ for every $z \in \mathbb{F}_q$. In particular, every occurrence of x^q can be replaced by x and the evaluation remains the same. Thus there exists a degree $|\mathrm{Supp}(m)| \cdot (q-1)$ monomial that agrees with m on $\mathbb{F}_q^n$. Claim 3.7 gives us a polynomial of degree $|\mathrm{Supp}(m)|$, which is better than $|\mathrm{Supp}(m)| \cdot (q-1)$, but only agrees on a translate of the Boolean cube.*

Proof of Claim 3.7. For simplicity, assume that $\mathrm{Supp}(m) = \{1, \ldots, r\} \subseteq [n]$ and let $m = x_1^{\alpha_1} \cdots x_r^{\alpha_r}$. We are only interested in evaluations in $\mathbf{c} + \{0,1\}^n$, which means that we are only interested in the setting when for every $i \in [n]$, x_i takes a value from the set $\{c_i, c_i + 1\}$. So for every $x_i^{\alpha_i}$, we replace it with a degree-1 polynomial that agrees with it on $\{c_i, c_i + 1\}$. We explain it in more detail below.

Fix any $i \in [n]$. Consider the following degree-1 polynomial,

$$L_i(x_i) \;:=\; (x_i - c_i) \cdot (c_i + 1)^{\alpha_i} \;-\; (x_i - (c_i + 1)) \cdot c_i^{\alpha_i}.$$

Observe that[13] $L_i(x_i)$ and $x_i^{\alpha_i}$ agree on the set $\{c_i, c_i + 1\}$. Now define the polynomial $P_m(\mathbf{x})$ as follows:

$$P_m(\mathbf{x}) \;:=\; L_1(x_1) \cdots L_r(x_r).$$

Similar to above, P_m agrees with the monomial m on $\mathbf{c} + \{0,1\}^n$. Clearly $\deg(P_m) \leqslant r = |\mathrm{Supp}(m)|$ and the sparsity of $P_m(\mathbf{x})$ is at most $2^r = 2^{|\mathrm{Supp}(m)|}$.

■

Corollary 3.9 (Multilinearization). *[KS19, Lemma 4.7]. Let S be a translate of the Boolean hypercube, i.e., there exists $\mathbf{c} \in \mathbb{F}_q^n$ such that $S = \mathbf{c} + \{0,1\}^n$. For every polynomial $P \in \mathbb{F}_q[\mathbf{x}]$, there exists a unique multilinear polynomial $\widetilde{P} \in \mathbb{F}_q[\mathbf{x}]$ with $\deg(\widetilde{P}) \leqslant \deg(P)$ such that, for every $\mathbf{a} \in S$, $\widetilde{P}(\mathbf{a}) = P(\mathbf{a})$.*

[13] The polynomial L_i could be easily constructed either by Lagrange interpolation or by solving a system of two linear equations.

Proof. The existence of $\widetilde{P}$ follows from Claim 3.7 applied to every monomial in $P(\mathbf{x})$. Uniqueness follows from the fact that no non-zero multilinear polynomial can vanish on $\mathbf{c} + \{0,1\}^n$. ■

We state a binomial coefficient approximation that will be useful in our calculations.

Claim 3.10. *Let $k, n \in \mathbb{N}$ be growing parameters and let $k = o(n)$. Then,*

$$\ln \binom{n}{k} \;=\; (1 + o(1)) \cdot k \cdot \ln \left(\frac{n}{k}\right).$$

Proof of Claim 3.10. Using Stirling's approximation, we have,

$$\left(\frac{n}{k}\right)^k \leqslant \binom{n}{k} \leqslant \left(\frac{ne}{k}\right)^k \quad \Leftrightarrow \quad k \cdot \ln\left(\frac{n}{k}\right) \leqslant \ln \binom{n}{k} \leqslant k \cdot \left(\ln\left(\frac{n}{k}\right) + \ln(e)\right).$$

Using $k = o(n)$, we get the desired bound in Claim 3.10. ■

3.3 Complexity Measure

We define the complexity measure Γ that we will use to prove our lower bounds. It is the same as used in [KS19].

Definition 3.11 (Complexity Measure in [KS19]). *Fix a field $\mathbb{F}$ and let $n \in \mathbb{N}$. Let $k, \ell \in \mathbb{N}$ and $S \subseteq \mathbb{F}_q^n$. For any polynomial $P \in \mathbb{F}[x_1, \ldots, x_n]$, we define $\Gamma_{k,\ell,S}(P)$ as follows:*

$$\Gamma_{k,\ell,S}(P) \;:=\; \dim\left[\mathrm{Eval}_S\left(\mathbf{x}^{=\ell}\left(\partial^{=k}(P)\right)\right)\right].$$

The first and foremost thing to check is that the complexity measure $\Gamma_{k,\ell,S}$ is sub-additive, i.e.,

$$\Gamma_{k,\ell,S}(T_1 + T_2) \;\leqslant\; \Gamma_{k,\ell,S}(T_1) \;+\; \Gamma_{k,\ell,S}(T_2), \qquad \text{for any two polynomials } T_1, T_2.$$

This essentially follows because partial derivatives are linear operators. We would like to note some trivial upper bounds on the complexity measure. For every choice of k, ℓ, S where S is a translation of the Boolean cube, we have the following upper bound on the measure (we explain it below): For every degree-d polynomial $P(\mathbf{x})$,

$$\Gamma_{k,\ell,S}(P) \;\leqslant\; \min\left\{\binom{n}{\ell + d - k}, \quad \binom{n}{\ell} \cdot \binom{n+k}{k}, \quad |S|\right\}. \tag{1}$$

To see why the first binomial coefficient appears, note that after taking k^{th} order partial derivative of a degree-d polynomial, we are left with a degree-$(d-k)$ polynomial, and then we multiply it

by a degree ℓ polynomial. This gives us a polynomial of degree $(\ell + d - k)$, and since we are only evaluating it over a translation of the Boolean cube, we only need to consider multilinear polynomials. For the second quantity, this is simply the product of the number of degree-ℓ multilinear monomials and the number of k^{th} order partial derivatives. The third quantity is because any function on S is always in the span of indicator functions.

3.4 Lower Bound on Nisan-Wigderson Polynomials

[KLSS17] and [KS17b] gave a lower bound on the measure for these families under the measure of shifted partials (the lower bound of [KLSS17] holds only for fields of characteristic 0, however, the lower bound of [KS17b] holds for every field). Later, [KS19] gave a tighter analysis of the lower bound, which is the one we are also going to use. We now discuss the choice of parameters, which are chosen carefully so that they yield a meaningful lower bound on the measure of the Nisan-Wigderson polynomials.

Choice of Parameters Let n denote the number of variables and d denote the degree of the polynomial. Recall that our complexity measure is $\Gamma_{k,\ell,S}$ where $k \in \mathbb{N}$ represents the order of derivatives we consider, $\ell \in \mathbb{N}$ represents the degree of monomials by which we shift, and subset $S \subseteq \mathbb{F}_q^n$ denotes the set of evaluation points. Then,

- $\beta \leqslant C \cdot d^{\lambda}$ for some absolute constants $C > 0$ and $0 \leqslant \lambda < 0.5$. We will fix β later (β will be chosen in a way such that the top-product fan-in is less than d^{β}).
- $k = \eta\sqrt{d}$, where η is a parameter of our choice with $\eta\beta = \Theta(1)$,
- $\varepsilon = \dfrac{\beta\eta}{2} \cdot \dfrac{\ln d}{\sqrt{d}}$,
- $\ell = \dfrac{n}{2} \cdot (1 - \varepsilon)$,
- m and e are chosen such that $d^{\beta} < m \leqslant 2d^{\beta}$ and they satisfy the following two inequalities:

$$m^k \geqslant (1+\varepsilon)^{2(d-k)}$$

$$m^{e-k} = \left(\frac{2}{1+\varepsilon}\right)^{d-k} \cdot \mathrm{poly}(m).$$

Note that $\mathsf{NW}_{d,m,e}(\mathbf{x})$ has dm variables. For our choice of parameters, $m = \Theta(d^{\beta})$, we get $n = \Theta(d^{\beta+1})$ variables. In the first condition, m^k refers to the number of k^{th} order partial derivatives where $\mathsf{NW}_{d,m,e}(\mathbf{x})$ is non-zero, and $(1+\varepsilon)^{2(d-k)}$ refers to the number of partial derivatives that [KS19] use to get the lower bound, stated in Lemma 3.12. It turns out the lower bound on $\Gamma_{k,\ell,S}(\mathsf{NW}_{d,m,e})$ is roughly $\binom{n}{\ell} \cdot (1+\varepsilon)^{2(d-k)}$.

Existence of the above parameters We follow roughly the same argument as given in [KS19, Appendix A], but with a slightly more detailed argument to justify our choice of using several times the parameter η.

We start by fixing m to be the smallest power of 2 that is $> d^\beta$. Say $m = 2^M$ where $m > d^\beta$. For the first condition, since $\varepsilon > 0$, we have

$$(1+\varepsilon)^{2(d-k)} < (e^\varepsilon)^{2d} = e^{\beta\eta\sqrt{d}\ln d} = d^{\beta k} < m^k.$$

For the second condition, we note:

$$m^{e-k} < \left(\frac{2}{1+\varepsilon}\right)^{d-k} \cdot \mathrm{poly}(m), \qquad \text{if } e = k$$

$$m^{e-k} > \left(\frac{2}{1+\varepsilon}\right)^{d-k} \cdot \mathrm{poly}(m), \qquad \text{if } e = d.$$

Thus there exists some $k < e < d$ for which the second condition holds.

Now we are ready to state the lower bound on the complexity measure of Nisan-Wigderson polynomials. We use a tighter analysis from [KS19]. In [KS19], they analyzed the measure for $\Gamma_{k,\ell,S}(\mathsf{NW}_{m,d,e})$ for $m = \Theta(d^2)$. However, we note that their lower bound continues to hold as long as the two above-mentioned conditions hold. Since the analysis is exactly similar, we skip the proof of it and refer the interested reader to [KS19, Appendix A].

Lemma 3.12. *[KS19, Lemma 5.9]. Let β, k, ℓ, m, e be as chosen above and $n = md$. Let $\mathcal{E} \subset \mathbb{F}_q^n$ be a subset with $|\mathcal{E}| \leqslant \delta \cdot q^n$, where $\delta = \exp(-\omega(\log^2 d))$. Then, there exists a subset $S \subset \mathbb{F}_q^n \backslash \mathcal{E}$ contained inside some translate of the Boolean hypercube, i.e., there exists some $\mathbf{c} \in \mathbb{F}_q^n$ such that $S \subseteq \mathbf{c} + \{0,1\}^n$, for which the following holds:*

$$\Gamma_{k,\ell,S}(\mathsf{NW}_{d,m,e}) \geqslant \binom{n}{\ell + d - k} \cdot \exp(-\mathcal{O}(\log^2 d)).$$

Note that the lower bound in Lemma 3.12 is quite close to the best we can expect from our complexity measure, as we observed in Equation (1).

4 Depth-4 with Bounded Top Product Fan-in

In this section, we prove an exponential lower bound against depth-4 circuits with bounded fan-in on top product gates, Theorem 4.1. This will be a special case of our main theorem, Theorem 1.1. The proof of Theorem 4.1 is slightly simpler than that of Theorem 1.1, so we

present it here as a warm-up for the main theorem.

Theorem 4.1 (Lower Bounds for Restricted Depth-4). *Fix a field $\mathbb{F}_q$. Let $n, d \in \mathbb{N}$ be growing parameters, where $d = \mathcal{O}(\sqrt{n})$. Let $\alpha \in \mathbb{N}$ be a parameter verifying $0.01 \leqslant \alpha \leqslant d^{1/3}$. There exists $m, e \in \mathbb{N}$ with $m = \Theta(d^{\alpha+1})$ such that the following holds:*

Any $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit with $a = \lceil d^{\alpha+0.5} \rceil$, over $\mathbb{F}_q$, computing the degree-d polynomial $\mathsf{NW}_{d,m,e}(\mathbf{x})$ requires size $\exp(\Omega_q(\sqrt{d/\alpha}))$.

Throughout this section, we fix $q = \mathcal{O}(1)$, i.e., the field $\mathbb{F}_q$ is a finite field of constant size.

4.1 Upper Bound when Bounded Support on Bottom Products

In this subsection, we start with the case of depth-4 circuits whose bottom gates have bounded support. Recall that the notation $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$ refers to circuits where the top product gates have fan-in at most a and the bottom product gates have support at most r.

Lemma 4.2. *Let $k, \ell, r \in \mathbb{N}$ satisfying $\ell + rk \leqslant n/2$. Let $S \subseteq \mathbb{F}_q^n$ be a subset contained inside some translate of the Boolean hypercube, i.e., there exists some $\mathbf{c} \in \mathbb{F}_q^n$ such that $S \subseteq \mathbf{c} + \{0,1\}^n$. Let $f \in \mathbb{F}_q[\mathbf{x}]$ be a degree-d polynomial that can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$ circuit of size s. Then,*

$$\Gamma_{k,\ell,S}(f) \;\leqslant\; s \cdot \binom{n}{\ell + rk} \cdot n \cdot \binom{a}{k}.$$

Proof of Lemma 4.2. Let $f(\mathbf{x}) = T_1(\mathbf{x}) + \cdots + T_s(\mathbf{x})$, where every $T_i(\mathbf{x})$ can be expressed as follows:

$$T_i(\mathbf{x}) \;=\; Q_1^{(i)}(\mathbf{x}) \cdots Q_a^{(i)}(\mathbf{x}).$$

Since the measure is subadditive, we restrict our attention to a single term $T(\mathbf{x}) = Q_1(\mathbf{x}) \cdots Q_a(\mathbf{x})$. From the hypothesis, we know that for every $j \in [a]$, $|\mathrm{Supp}(Q_j)| \leqslant r$. We have,

$$\begin{aligned}
\partial^{=k}(T) &\subseteq \mathrm{span}\left\{ \prod_j \partial^{=\alpha_j}(Q_j(\mathbf{x})) \;\middle|\; (\alpha_1, \ldots, \alpha_a) \in \mathbb{N}^a,\ |\boldsymbol{\alpha}| = k \right\} \\
&\subseteq \mathrm{span}\left\{ \prod_{j \notin U} Q_j(\mathbf{x}) \cdot \prod_{j \in U} R_{j,U}(\mathbf{x}) \;\middle|\; U \in \binom{[a]}{k}, R_{j,U} \text{ such that } \forall j, \mathrm{Supp}(R_{j,U}) \subseteq \mathrm{Supp}(Q_j) \right\}.
\end{aligned}$$

The second inclusion comes from the fact that any product $\prod_j \partial^{=\alpha_j}(Q_j(\mathbf{x}))$ has at most k factors which are not of the form $Q_j(\mathbf{x})$.

Fix a $U \in \binom{[a]}{k}$. We have $|\mathrm{Supp}(R_{j,U})| \leqslant |\mathrm{Supp}(Q_j)| \leqslant r$, where for the final inequality we used

that the inner product gates compute a polynomial of support $\leqslant r$. In particular,

$$\Gamma_{k,\ell,S}(T) = \dim\left[\mathrm{Eval}_S\left(\mathbf{x}^{=\ell}\left(\partial^{=k}(T)\right)\right)\right] \leqslant \dim\left[\mathrm{Eval}_S\left(\mathrm{span}\left\{R'_{U,\ell}\cdot\prod_{j\notin U}Q_j(\mathbf{x}) \;\middle|\; U\in\binom{[a]}{k}\right\}\right)\right]$$

where $R'_{U,\ell}$ are polynomials of support bounded by $\ell + kr$.

As stated in the hypothesis, the subset S is contained inside some translate of the Boolean hypercube. Let $R'_{j,U}$ denote the polynomial we get after replacing every non-zero monomial $m \in R_{j,U}$ with the corresponding polynomial $P_m(\mathbf{x})$ of degree $|\mathrm{Supp}(m)|$ from Claim 3.7. Since $|\mathrm{Supp}(R_{j,U})| \leqslant r$, we have $|\mathrm{Supp}(m)| \leqslant r$, and thus $\deg(R'_{j,U}) \leqslant r$.

From Claim 3.7, we know that for every $\boldsymbol{\gamma} \in \mathbb{F}_q^n$, we have $R'_{j,U}(\boldsymbol{\gamma}) = R_{j,U}(\boldsymbol{\gamma})$. Thus both $R'_{j,U}$ and $R_{j,U}$ have the same evaluation vector on $\mathbb{F}_q^n$, and in particular, the same evaluation vector on $S \subset \mathbb{F}_q^n$. More formally,

$$\Gamma_{k,\ell,S}(T) = \dim\left[\mathrm{Eval}_S\left(\mathbf{x}^{=\ell}\left(\partial^{=k}(T)\right)\right)\right] \leqslant \dim\left[\mathrm{Eval}_S\left(\mathrm{span}\left\{\prod_{j\notin U}Q_j(\mathbf{x})\cdot\mathbf{x}^{<\ell+rk} \;\middle|\; U\in\binom{[a]}{k}\right\}\right)\right].$$

From Corollary 3.9, we know that any polynomial can be replaced by a unique multilinear polynomial such that the evaluations over S remain the same. As $\Gamma_{k,\ell,S}$ only evaluates over a translate of the Boolean hypercube $\{0,1\}^n$, we only need to consider at the place of $R'_{U,\ell}$, multilinear monomials of degree bounded by $\ell + rk$. Thus using sub-additivity of $\Gamma_{k,\ell,S}$ and $\ell + rk \leqslant n/2$, we get,

$$\Gamma_{k,\ell,S}(f) \;\leqslant\; s\cdot\binom{a}{k}\cdot\sum_{i=0}^{\ell+rk}\binom{n}{i} \;\leqslant\; s\cdot n\cdot\binom{n}{\ell+rk}\cdot\binom{a}{k}.$$

This finishes the proof of Lemma 4.2. ∎

4.2 Reduction to Bottom Products with Bounded Support

In the previous subsection, we considered $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$ circuits and computed an upper bound on the complexity measure of polynomials that have small $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$ circuits. Now we discuss how to handle bottom products with arbitrary fan-in. We will show that every polynomial with a small $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit can be *approximated* by a polynomial that has a small $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$, with the additional property that the measure remains the same.

Lemma 4.3 (Reducing the Support of Bottom Product Gates). *Let $n, d \in \mathbb{N}$ be growing parameters, where $d = \mathcal{O}(\sqrt{n})$. Let $\alpha \in \mathbb{N}$ be a parameter such that $0.01 \leqslant \alpha \leqslant d^{1/3}$. Let $\beta := \lceil \alpha + 1 \rceil$ and $\eta := \sqrt{\beta - \alpha}/(40\sqrt{\beta q})$. Let $a = \lceil d^{\alpha+0.5} \rceil$ and $s \leqslant \exp(20\eta \cdot \sqrt{d})$. Let $k = \eta\sqrt{d}, \ell, m, e$ be as in Lemma 3.12.*

Suppose a polynomial $f \in \mathbb{F}_q[x]$ can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit of size s. For support size parameter $r := 200qk$, there exists a polynomial $g \in \mathbb{F}_q[x]$ and a subset $\mathcal{E} \subseteq \mathbb{F}_q^n$ of size at most $\exp(-\sqrt{d/(\beta q)}) \cdot q^n$ satisfying the following properties:

1. *The polynomials f and g agree on $\mathbb{F}_q^n \backslash \mathcal{E}$.*
2. *The polynomial g can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$ circuit of size s.*
3. *Let $S \subseteq \mathbb{F}_q^n \backslash \mathcal{E}$. Then the measure $\Gamma_{k,\ell,S}(f)$ equals the measure $\Gamma_{k,\ell,S}(g)$.*

Proof of Lemma 4.3. From the hypothesis, we know that the polynomial $f(x)$ can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit of size s, which means we can express $f(x)$ as follows:

$$f(x) \;=\; Q_1^{(1)}(x) \cdots Q_a^{(1)}(x) \;+ \ldots + \; Q_1^{(s)}(x) \cdots Q_a^{(s)}(x).$$

Since the circuit has size s, we also know that each polynomial $Q_j^{(i)}(x)$ has sparsity at most s. We refer to the polynomials $Q_j^{(i)}(x)$'s as *inner polynomials*.

Existence of a dense subset with high multiplicity Fix an inner polynomial $Q_j^{(i)}(x)$. Let m be a monomial with non-zero coefficient in $Q_j^{(i)}(x)$ that has support size at least r, i.e., $\mathrm{Supp}(m) \geqslant r$. Since $r/100q > k$, Corollary 3.6 guarantees the existence of a subset $\mathcal{E}_m \subseteq \mathbb{F}_q^n$ with

$$|\mathcal{E}_m| \;\leqslant\; \exp(-r/3q) \cdot q^n, \qquad\qquad \text{(using } \kappa = 0.99\text{)}$$

such that for every point $a \in \mathbb{F}_q^n \backslash \mathcal{E}_m$, we have $\mathrm{Mult}(m, a) \geqslant r/100q > k$.

Let $\mathcal{E}$ be the union of $\mathcal{E}_m$ over all monomials m (computed at the bottom product gates) that have support size at least r. Since the size of the $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit is s, we can have at most s such monomials. Using $s \leqslant \exp(20\eta \cdot \sqrt{d})$, we get,

$$|\mathcal{E}| \leqslant s \cdot \exp(-r/3q) \cdot q^n \leqslant \exp(-40\eta\sqrt{d}) \cdot q^n = \exp\left(-\frac{\sqrt{(\beta-\alpha)d}}{\sqrt{\beta q}}\right) \cdot q^n \leqslant \exp\left(-\sqrt{\frac{d}{\beta q}}\right) \cdot q^n.$$

Defining the approximating polynomial $g(x)$ For every inner polynomial $Q_j^{(i)}(x)$, define the polynomial $\widetilde{Q}_j^{(i)}(x)$ as follows: Remove all monomials in $Q_j^{(i)}(x)$ that have support $\geqslant r$. Before proceeding to define $g(x)$, we make an observation that will be useful later in the proof.

Observation 4.4. *Let* $\mathbf{u}$ *be a monomial of degree* $\leqslant k$*. Then we have,*

$$\frac{\partial \widetilde{Q}_j^{(i)}(\mathbf{x})}{\partial \mathbf{u}}(\mathbf{a}) \;=\; \frac{\partial Q_j^{(i)}(\mathbf{x})}{\partial \mathbf{u}}(\mathbf{a}), \qquad \text{for every } \mathbf{a} \in \mathbb{F}_q^n \backslash \mathcal{E}.$$

This is because $Q_j^{(i)}$ *and* $\widetilde{Q}_j^{(i)}$ *only differ in monomials with support* $\geqslant r$ *and every point in* $\mathbb{F}_q^n \backslash \mathcal{E}$ *has multiplicity strictly larger than* k *with respect to those monomials.*

Now we define $g(\mathbf{x})$ as follows:

$$g(\mathbf{x}) \;:=\; \widetilde{Q}_1^{(1)}(\mathbf{x}) \cdots \widetilde{Q}_a^{(1)}(\mathbf{x}) \;+ \ldots + \; \widetilde{Q}_1^{(s)}(\mathbf{x}) \cdots \widetilde{Q}_a^{(s)}(\mathbf{x}).$$

We now check that $g(\mathbf{x})$ satisfies the three properties mentioned in Lemma 4.3:

1. From the definition of $\mathcal{E}$, we have that for every $\mathbf{a} \in \mathbb{F}_q^n \backslash \mathcal{E}$, $g(\mathbf{a})$ equals $f(\mathbf{a})$.

2. Note that the polynomial $g(\mathbf{x})$ can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit with bottom product gates of support at most r and size s. More precisely, removing bottom product gates with support at least r from the $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit computing $f(\mathbf{x})$ yields a circuit for $g(\mathbf{x})$.

3. Let $S \subseteq \mathbb{F}_q^n \backslash \mathcal{E}$. Using Observation 4.4, we have that for any monomial $\mathbf{u}$ of degree k and any monomial $\mathbf{v}$ of degree ℓ,

$$\left(\mathbf{v} \cdot \frac{\partial \widetilde{Q}_j^{(i)}(\mathbf{x})}{\partial \mathbf{u}}\right)(\mathbf{a}) \;=\; \left(\mathbf{v} \cdot \frac{\partial Q_j^{(i)}(\mathbf{x})}{\partial \mathbf{u}}\right)(\mathbf{a}), \qquad \text{for every } \mathbf{a} \in \mathbb{F}_q^n \backslash \mathcal{E}.$$

   This implies that $\Gamma_{k,\ell,S}(f) \;=\; \Gamma_{k,\ell,S}(g)$.

This finishes the proof of Lemma 4.3. ∎

4.3 Proof of Theorem 4.1

In this subsection, we use Lemma 4.3, Lemma 4.2, and Lemma 3.12 to prove Theorem 4.1. Lemma 4.3 states that every polynomial that can be efficiently computed by a $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit can be approximated by a polynomial with same complexity measure, that can be efficiently computed by a $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$ circuit. Lemma 4.2 gives us an upper bound on the complexity measure of every polynomial computed by a small $\Sigma\Pi^{[a]}\Sigma\Pi^{\{r\}}$ circuit. Lemma 4.2 together with Lemma 4.3 gives us an upper bound on the complexity measure of a polynomial computed by a small $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit. Combining it with the lower bound on the measure of Nisan-Wigderson polynomials (Lemma 3.12) will give us the lower bound we are seeking. We recall Theorem 4.1 below and then prove it.

Theorem 4.1 (Lower Bounds for Restricted Depth-4). *Fix a field $\mathbb{F}_q$. Let $n, d \in \mathbb{N}$ be growing parameters, where $d = \mathcal{O}(\sqrt{n})$. Let $\alpha \in \mathbb{N}$ be a parameter verifying $0.01 \leqslant \alpha \leqslant d^{1/3}$. There exists $m, e \in \mathbb{N}$ with $m = \Theta(d^{\alpha+1})$ such that the following holds:*

Any $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit with $a = \lceil d^{\alpha+0.5} \rceil$, over $\mathbb{F}_q$, computing the degree-d polynomial $\mathsf{NW}_{d,m,e}(\mathbf{x})$ requires size $\exp(\Omega_q(\sqrt{d/\alpha}))$.

Proof of Theorem 4.1. Let $\beta := \lceil \alpha + 1 \rceil$. Note that $(\beta - \alpha)$ is always positive and $(\beta - \alpha) < 2$. Assume that the polynomial $\mathsf{NW}_{m,d,e}(\mathbf{x})$ of degree-d is computed by a $\Sigma\Pi^{[a]}\Sigma\Pi$ circuit of size s. Suppose for the sake of argument that $s < \exp(20\eta \cdot \sqrt{d})$. Let η, k, r, ℓ, m, e be as in Lemma 4.3. We quickly verify that these parameters satisfy the conditions in Lemma 4.2.

$$\ell + rk \;=\; \frac{n}{2}(1-\varepsilon) \;+\; 200q\eta^2 \cdot d \leqslant \frac{n}{2}(1-\varepsilon) \;+\; d < \frac{n}{2},$$

where for the final equality we used $d = \mathcal{O}(\sqrt{n})$.

Let $g(\mathbf{x})$ be the polynomial and $\mathcal{E}$ be the subset guaranteed from Lemma 4.3. We know that $|\mathcal{E}| \leqslant \delta \cdot q^n$ for $\delta = \exp(-\sqrt{d/(\beta q)})) = \exp(-\omega(\log^2 d))$. Let $S \subseteq \mathbb{F}_q^n \backslash \mathcal{E}$ be the translate of Boolean hypercube from Lemma 3.12 such that

$$\Gamma_{k,\ell,S}(\mathsf{NW}_{m,d,e}(\mathbf{x})) \;\geqslant\; \binom{n}{\ell + d - k} \cdot \exp(-\mathcal{O}(\log^2 d)).$$

From the second item of Lemma 4.3 and Lemma 4.2, we have

$$\Gamma_{k,\ell,S}(g) \;\leqslant\; s \cdot \binom{n}{\ell + rk} \cdot n \cdot \binom{a}{k}.$$

From the third item of Lemma 4.3, we know

$$\Gamma_{k,\ell,S}(\mathsf{NW}_{m,d,e}(\mathbf{x})) \;\leqslant\; s \cdot \binom{n}{\ell + rk} \cdot n \cdot \binom{a}{k}.$$

Lower Bound We now give a lower bound on the size s. Let Z denote the following quantity:

$$Z \;:=\; \binom{n}{\ell + d - k} \cdot \exp(-\mathcal{O}(\log^2 d)) \Big/ \left(\binom{n}{\ell + rk} \cdot n \cdot \binom{a}{k} \right).$$

Since $rk = 200q\eta^2 \cdot d < (d - k)$, simplifying the binomial coefficients,

$$Z \;=\; \frac{(n-\ell) - rk}{\ell + (d-k)} \cdots \frac{(n-\ell) - (d-k) + 1}{\ell + rk + 1} \cdot \frac{1}{n \cdot \binom{a}{k}} \;\cdot\; \exp(-\mathcal{O}(\log^2 d))$$

$$\geqslant \left(\frac{(n-\ell)-rk}{\ell+(d-k)}\right)^{d-(r+1)k} \cdot \frac{1}{n\cdot\binom{a}{k}} \cdot \exp(-\mathcal{O}(\log^2 d)).$$

628 Using $\ell = \frac{n}{2}\cdot(1-\varepsilon)$, we get,

$$Z \geqslant \left(\frac{(1+\varepsilon)-\frac{rk}{n/2}}{1-(\varepsilon-\frac{d-k}{n/2})}\right)^{d-(r+1)k} \cdot \frac{1}{n\cdot\binom{a}{k}} \cdot \exp(-\mathcal{O}(\log^2 d)).$$

629 Using $n = \Omega(d^2)$ and the inequality $1/(1-u) > (1+u)$ for $0 < u < 1$,

$$Z \geqslant \left(\left(1+\varepsilon-\frac{rk}{n/2}\right)\cdot\left(1+\varepsilon-\frac{d-k}{n/2}\right)\right)^{d-(r+1)k} \cdot \frac{1}{n\cdot\binom{a}{k}} \cdot \exp(-\mathcal{O}(\log^2 d)).$$

630 Using also $\varepsilon = (\beta\eta/2)\cdot \ln d/\sqrt{d}$, we know that for n, d large enough,

$$\left(1+\varepsilon-\frac{rk}{n/2}\right)\cdot\left(1+\varepsilon-\frac{d-k}{n/2}\right) \geqslant 1+2\varepsilon.$$

631 Let $\theta = \dfrac{\alpha+\beta}{2\beta} < 1$, using the inequality $(1+u)^{1/u} \geqslant e^{\theta}$ for $u > 0$ small enough, we get,

$$Z \geqslant \exp\left(2\theta\varepsilon\cdot(d-(r+1)k)\right) \cdot \frac{1}{n\cdot\binom{a}{k}} \cdot \exp(-\mathcal{O}(\log^2 d)).$$

632 Since $a = \Omega(d^{0.51})$, by using Claim 3.10 on $\binom{a}{k}$, we get,

$$\frac{1}{\binom{a}{k}} = \exp(-(1+o(1))\cdot k\cdot\ln(a/k)).$$

633 We have $a = \lceil d^{\alpha+0.5}\rceil$, i.e., $\ln(a/k) \leqslant \alpha\ln d + \ln(1/\eta) + 1$, where $\ln(1/\eta)+1 = O(\alpha)$. Simplifying,
634 we get,

$$\begin{aligned} Z &\geqslant \exp\left(2\theta\varepsilon d - 2\theta\varepsilon rk - k\alpha\ln d - o(k\alpha\ln d)\right) \qquad (2)\\ &= \exp\left(\left(\frac{\alpha+\beta}{2}\eta - 100(\alpha+\beta)q\eta^3 - \eta\alpha\right)\cdot\sqrt{d}\ln d - o(\sqrt{\alpha d}\ln d)\right). \end{aligned}$$

635 We now simplify the function that is multiplied with $\sqrt{d}\ln d$ inside the above exp function

using our choice of parameters:

$$\frac{\alpha+\beta}{2}\eta - 100(\alpha+\beta)q\eta^3 - \eta\alpha \;=\; \frac{\eta}{2}\left(\beta-\alpha-\frac{(\alpha+\beta)(\beta-\alpha)}{8\beta}\right) \;=\; \frac{\eta}{16\beta}(7\beta-\alpha)(\beta-\alpha)$$

$$> \;\frac{3}{8}\eta(\beta-\alpha) \;=\; \frac{3}{320}\frac{(\beta-\alpha)^{1.5}}{\sqrt{\beta q}} \;=\; \Omega_q(1/\sqrt{\alpha}),$$

where we use $(\beta-\alpha)\geqslant 1$ for the final inequality. This implies

$$Z \;=\; \exp(\Omega_q((\sqrt{d/\alpha})\cdot\log d)).$$

Recall that at the beginning of the proof, we assumed $s < \exp(\sqrt{(\beta-\alpha)d/(4\beta q)})$ (we made this assumption so we could use Lemma 4.2). The above lower bound on Z says that a $\Sigma\Pi^{[a]}\Sigma\Pi$ of size $\exp(\Omega_q(\sqrt{d/\alpha}))$ cannot compute the polynomial $\mathsf{NW}_{d,m,e}$. This finishes the proof of Theorem 4.1. ■

Depth-4 Circuits with Outer Powering Gates The above proof also simply extends when we replace the top product gates with powering gates, and no bound on the fan-in of the powering gates. We state it formally below, which even follows from [KS19], but we note it here for the sake of completeness.

Corollary 4.5 (Lower Bound for Depth-4 with Outer Powering Gates). *Let $n\in\mathbb{N}$ be a growing parameter and let $d\in\mathbb{N}$ denote the degree parameter, where d is a function of n. There exists an explicit polynomial $f(x_1,\ldots,x_n)\in\mathbb{F}_q[\mathbf{x}]$ of degree at most d satisfying the following:*
Every $\Sigma\wedge\Sigma\Pi$ circuit over $\mathbb{F}_q$ computing the polynomial $f(\mathbf{x})$ requires size $\exp(\Omega(\sqrt{d}))$.

Proof Sketch of Corollary 4.5. The proof is quite similar (and a bit simpler) than the proof of Theorem 4.1. Fix any natural number $w\in\mathbb{N}$. Let $\Sigma\wedge^{[w]}\Sigma\Pi$ denote the depth-4 circuits where every $\wedge$ gate computes a $\leqslant w^{th}$ power of its child. We have that $\Sigma\wedge^{[w]}\Sigma\Pi\subseteq\Sigma\Pi^{[w]}\Sigma\Pi$, and thus Lemma 4.3 and Lemma 4.2 applies on the class $\Sigma\wedge^{[w]}\Sigma\Pi$, but with different parameters. In particular, Lemma 4.2 yields an upper bound of $s\cdot\binom{n}{\ell+rk}\cdot n$, and there is no $\binom{a}{k}$ factor. The reason is that when we take k^{th} partial derivatives, there is no choice on which of the k inner polynomials to derive on, and hence no factor of $\binom{a}{k}$. The lower bound calculation from the proof of Theorem 4.1 continues to hold as usual, without the $\binom{a}{k}$ factor. This only improves the lower bound on Z in the proof of Theorem 4.1 and thus $Z=\Omega(\exp(\sqrt{d}\log d))$. As this holds for every $w\in\mathbb{N}$, we have a lower bound on $s=\Omega(\exp(\sqrt{d}))$, irrespective of w. This finishes the proof sketch of Corollary 4.5. ■

5 Depth-5 with Bounded Top Product Fan-in

In this section, we discuss the proof of Theorem 1.1. As discussed in the technical overview, the proof of Theorem 1.1 is quite similar to the proof of Theorem 4.1.

5.1 Upper Bound when Bottom Product Gates have Bounded Rank

We start by giving an upper bound on the complexity measure of polynomials that have a small $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit whose bottom product gates have bounded rank. The following lemma is an analogue of Lemma 4.2.

Lemma 5.1. *Let $k, \ell, r \in \mathbb{N}$ satisfying $\ell + rqk \leqslant n/2$. Let $S \subseteq \mathbb{F}_q^n$ be a subset contained inside some translate of the Boolean hypercube, i.e., there exists some $\mathbf{c} \in \mathbb{F}_q^n$ such that $S \subseteq \mathbf{c} + \{0,1\}^n$.*
Let $f(x_1, \ldots, x_n) \in \mathbb{F}_q[\mathbf{x}]$ be a degree-d polynomial that can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit of size s. If every bottom product gate has rank $\leqslant r$, then,

$$\Gamma_{k,\ell,S}(f) \;\leqslant\; s \cdot \binom{n}{\ell + rqk} \cdot n \cdot \binom{a}{k}.$$

Before we prove Lemma 5.1, it will be convenient to have the following claim.

Claim 5.2. *Let $Q(\mathbf{x}) \in \mathbb{F}_q[x_1, \ldots, x_n]$ be a polynomial that can be computed by a $\Sigma\Pi\Sigma$ circuit with product gate of rank $\leqslant r$ and size s. Then for every monomial $\mathbf{u}$, there exists a polynomial $P_{Q,\mathbf{u}}$ of degree $\leqslant r(q-1)$ such that*

$$\frac{\partial Q}{\partial \mathbf{u}}(\mathbf{a}) \;=\; P_{Q,\mathbf{u}}(\mathbf{a}), \qquad \text{for every } \mathbf{a} \in \mathbb{F}_q^m.$$

Proof of Claim 5.2. We know that $Q(\mathbf{x})$ can be expressed as:

$$Q(\mathbf{x}) \;=\; L_{11}(\mathbf{x}) \cdots L_{1s}(\mathbf{x}) \;+\; \ldots \;+\; L_{s1}(\mathbf{x}) \cdots L_{ss}(\mathbf{x}),$$

where for every $j \in [s]$, the rank of the set of degree-1 polynomials $\{L_{j1}(\mathbf{x}), \ldots, L_{js}(\mathbf{x})\}$ is at most r. Assume without loss of generality that for every j, the set $\{L_{j1}(\mathbf{x}), \ldots, L_{jr}(\mathbf{x})\}$ is a spanning set for span $\{L_{j1}, \ldots, L_{js}\}$. We have,

$$\frac{\partial Q(\mathbf{x})}{\partial \mathbf{u}} \;=\; \frac{\partial\, L_{11}(\mathbf{x}) \cdots L_{1s}(\mathbf{x})}{\partial \mathbf{u}} \;+\; \ldots \;+\; \frac{\partial\, L_{s1}(\mathbf{x}) \cdots L_{ss}(\mathbf{x})}{\partial \mathbf{u}}.$$

Consider any summand on the right-hand side:

$$\frac{\partial L_{p1}(\mathbf{x}) \cdots L_{ps}(\mathbf{x})}{\partial \mathbf{u}} \;\subseteq\; \operatorname{span}\left\{ \prod_{j \notin V} L_{pj}(\mathbf{x}) \;\middle|\; V \in \binom{[s]}{|\mathbf{u}|} \right\}.$$

Since $\{L_{p1}(\mathbf{x}), \ldots, L_{pr}(\mathbf{x})\}$ is a spanning set, we know that for every $V \in \binom{[s]}{|\mathbf{u}|}$, there exists non-negative integers $t_1, \ldots, t_r$ such that

$$\prod_{j \notin V} L_{pj}(\mathbf{x}) \;=\; L_{p1}(\mathbf{x})^{t_1} \cdots L_{pr}(\mathbf{x})^{t_r}.$$

Now define the polynomial $L_{p,V}(\mathbf{x})$ as follows: Keep replacing every occurrence of $L_{pj}(\mathbf{x})^q$ with $L_{pj}(\mathbf{x})$ until every $L_{pj}(\mathbf{x})$ occurs with degree less than q. Since $z^q = z$ for every $z \in \mathbb{F}_q$, we have,

$$\prod_{j \notin V} L_{pj}(\mathbf{a}) \;=\; L_{p,V}(\mathbf{a}), \qquad \text{for every } \mathbf{a} \in \mathbb{F}_q^n.$$

Note that $L_{p,V}$ is a polynomial of degree at most $r(q-1)$. Thus, there exists a polynomial $L_p(\mathbf{x})$ of degree at most $r(q-1)$ such that,

$$\frac{\partial(L_{p1} \cdots L_{ps})}{\partial \mathbf{u}}(\mathbf{a}) \;=\; L_p(\mathbf{a}), \qquad \text{for every } \mathbf{a} \in \mathbb{F}_q^n.$$

Now define $R_{Q,\mathbf{u}} := L_1 + \ldots + L_s$, where L_p is as defined above. Then we get,

$$\frac{\partial Q}{\partial \mathbf{u}}(\mathbf{a}) \;=\; R_{Q,\mathbf{u}}(\mathbf{a}), \qquad \text{for every } \mathbf{a} \in \mathbb{F}_q^n.$$

This finishes the proof of Claim 5.2. ■

Proof of Lemma 5.1. The proof proceeds similarly to the proof of Lemma 4.2. Let $f(\mathbf{x}) = T_1(\mathbf{x}) + \ldots + T_s(\mathbf{x})$, where every $T_i(\mathbf{x})$ can be expressed as follows:

$$T_i(\mathbf{x}) \;=\; Q_1^{(i)}(\mathbf{x}) \cdots Q_a^{(i)}(\mathbf{x}).$$

Furthermore, every $Q_j^{(i)}(\mathbf{x})$ is a sum of product of affine forms with rank at most r.
Since the complexity measure $\Gamma_{k,\ell,S}$ is sub-additive, we will restrict our attention to a single term $T(\mathbf{x}) = Q_1(\mathbf{x}) \ldots Q_a(\mathbf{x})$. We have,

$$\begin{aligned}
\partial^{=k}(T) \;&\subseteq\; \operatorname{span}\left\{ \prod_{j=1}^{a} \partial^{=\alpha_j}(Q_j(\mathbf{x})) \;\middle|\; \alpha_1, \ldots, \alpha_a \in \mathbb{N}^a,\ |\boldsymbol{\alpha}| = k \right\} \\
&\subseteq\; \operatorname{span}\left\{ \prod_{j \notin U} Q_j(\mathbf{x}) \prod_{j \in U} R_{j,U}(\mathbf{x}) \;\middle|\; U \in \binom{[a]}{k}, R_{j,U} \text{ such that } \forall j,\ R_{j,U} \in \partial^{\leqslant k}(Q_j) \right\}.
\end{aligned}$$

Fix $U \in \binom{[a]}{k}$ and $j \in U$, and consider the polynomial $R_{j,U}(\mathbf{x})$. Since $Q_j(\mathbf{x})$ can be computed by a $\Sigma\Pi\Sigma$ circuit with product gate of rank at most r and $R_{j,U} \in \partial^{\leqslant k}(Q_j)$, we can use Claim 5.2 to get the following: There exists a polynomial $R'_{j,U}(\mathbf{x})$ of degree at most $r(q-1)$ such that for every $\boldsymbol{\gamma} \in \mathbb{F}_q^n$, we have $R_{j,U}(\boldsymbol{\gamma}) = R'_{j,U}(\boldsymbol{\gamma})$. This means that both $R_{j,U}$ and $R_{j',U}$ have the same evaluation vector over $\mathbb{F}_q^n$, and in particular have the same evaluation vector over the subset $S \subseteq \mathbb{F}_q^n$. This

implies:

$$\Gamma_{k,\ell,S}(T) = \dim\left[\mathrm{Eval}_S\left(\mathbf{x}^{=\ell}\left(\partial^{=k}(T)\right)\right)\right] \leqslant \dim\left[\mathrm{Eval}_S\left(\mathrm{span}\left\{\prod_{j\notin U} Q_j(\mathbf{x})\cdot\mathbf{x}^{<\ell+rqk}\;\middle|\; U\in\binom{[a]}{k}\right\}\right)\right].$$

From Corollary 3.9, we know that any polynomial can be replaced by a unique multilinear polynomial such that the evaluations over S remain the same. As $\Gamma_{k,\ell,S}$ only evaluates over a translate of the Boolean hypercube $\{0,1\}^n$, we only need to consider multilinear monomials of degree at most $\ell + rqk$. Thus using sub-additivity of $\Gamma_{k,\ell,S}$ and $\ell + rqk \leqslant n/2$, we get,

$$\Gamma_{k,\ell,S}(f) \;\leqslant\; s\cdot\binom{a}{k}\cdot\sum_{i=0}^{\ell+rqk}\binom{n}{i} \;\leqslant\; s\cdot\binom{a}{k}\cdot n\cdot\binom{n}{\ell+rqk}.$$

This finishes the proof of Lemma 5.1. ∎

5.2 Proof of Theorem 1.1

Now we show how to reduce a depth-5 circuit to a depth-5 circuit with bounded rank of the bottom product gates. The following lemma is an analogue of Lemma 4.3.

Lemma 5.3 (Reducing the Rank of Bottom Product Gates). *Let $n, d \in \mathbb{N}$ be growing parameters, where $d = \mathcal{O}(\sqrt{n})$. Let $\alpha \in \mathbb{N}$ be a parameter such that $0.01 \leqslant \alpha \leqslant d^{1/3}$. Let $\beta := \lceil \alpha + 1\rceil$ and $\eta := \sqrt{\beta-\alpha}/(40q\sqrt{\beta})$. Let $a = \lceil d^{\alpha+0.5}\rceil$ and $s \leqslant \exp(20\eta\cdot\sqrt{d})$. Let $k = \eta\sqrt{d}, \ell, m, e$ be as in Lemma 3.12.*

Suppose a polynomial $f \in \mathbb{F}_q[\mathbf{x}]$ can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit of size s. For support size parameter $r := 200qk$, there exists a polynomial $g \in \mathbb{F}_q[\mathbf{x}]$ and a subset $\mathcal{E} \subseteq \mathbb{F}_q^n$ of size at most $\exp(-\sqrt{d/(\beta q)})\cdot q^n$ satisfying the following properties:

1. *The polynomials f and g agree on $\mathbb{F}_q^n\backslash\mathcal{E}$.*
2. *The polynomial g can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit of size s with bottom product gates of rank at most r.*
3. *Let $S \subseteq \mathbb{F}_q^n\backslash\mathcal{E}$. Then the measure $\Gamma_{k,\ell,S}(f)$ equals the measure $\Gamma_{k,\ell,S}(g)$.*

Proof of Lemma 5.3. From the hypothesis, we know that the polynomial $f(\mathbf{x})$ can be expressed as

$$f(\mathbf{x}) \;=\; Q_1^{(1)}(\mathbf{x})\cdots Q_a^{(1)}(\mathbf{x}) \;+\; \ldots \;+\; Q_1^{(s)}(\mathbf{x})\cdots Q_a^{(s)}(\mathbf{x}),$$

where each $Q_j^{(i)}(\mathbf{x})$ can be computed by a $\Sigma\Pi\Sigma$ circuit of size at most s. Furthermore, every $Q_j^{(i)}$ is a sum of product of at most s affine polynomials.

Existence of a dense subset with high multiplicity Fix a polynomial $Q_j^{(i)}(\mathbf{x})$ and denote it by $Q(\mathbf{x})$ for convenience in writing. We have,

$$Q(\mathbf{x}) \;=\; \ell_{11}(\mathbf{x}) \cdots \ell_{1s}(\mathbf{x}) \;+\; \ldots \;+\; \ell_{s1}(\mathbf{x}) \cdots \ell_{ss}(\mathbf{x}).$$

Let $j \in [s]$ be such that (if it exists) $\operatorname{rank}\{\ell_{j1}(\mathbf{x}), \ldots, \ell_{js}(\mathbf{x})\} > r$ and $\{\ell_{j1}(\mathbf{x}), \ldots, \ell_{j,r+1}(\mathbf{x})\}$ are linearly independent (maybe after relabeling). Since $r/100q > k$, using Claim 3.5 on $\{\ell_{j1}(\mathbf{x}), \ldots, \ell_{j,r+1}(\mathbf{x})\}$, we know there exists a subset $\mathcal{E}_{Q,j} \subseteq \mathbb{F}_q^m$ with

$$|\mathcal{E}_{Q,j}| \;\leqslant\; \exp(-r/3q) \cdot q^n, \qquad\qquad (\text{using } \kappa = 0.99)$$

such that for every point $\mathbf{a} \in \mathbb{F}_q^n \backslash \mathcal{E}_{Q,j}$, we have $\operatorname{Mult}(\ell_{j1} \cdot \ldots \cdot \ell_{j,r+1}, \mathbf{a}) \geqslant r/100q > k$.

The number of bottom product gates whose set of children is of rank larger than r is at most s since the size of the circuit is at most s. Let $\mathcal{E}$ be the union of $\mathcal{E}_{Q,j}$ over all such inner product gates. Using $s \leqslant \exp(20\eta \cdot \sqrt{d})$, we get,

$$|\mathcal{E}| \leqslant s \cdot \exp(-r/3q) \cdot q^n \leqslant \exp(-40\eta\sqrt{d}) \cdot q^n = \exp\left(-\frac{\sqrt{(\beta-\alpha)d}}{q\sqrt{\beta}}\right) \cdot q^n \leqslant \exp\left(-\frac{\sqrt{d}}{q\sqrt{\beta}}\right) \cdot q^n.$$

Defining the approximating polynomial $g(\mathbf{x})$ For every inner polynomial $Q_j^{(i)}(\mathbf{x})$, define a new polynomial $\widetilde{Q}_j^{(i)}(\mathbf{x})$ by removing every summand of rank larger than r. Before proceeding to define $g(\mathbf{x})$, we will make an observation that will be useful later in the proof.

Observation 5.4. *Let $Q_j^{(i)} = H_1 + \ldots + H_s$ and suppose H_t are bottom product gates of the circuit above of rank larger than r. Let $\mathbf{u}$ be any monomial of degree at most k. Then for every H_t,*

$$\frac{\partial H_t(\mathbf{x})}{\partial \mathbf{u}}(\mathbf{a}) \;=\; 0, \qquad\qquad \textit{for every } \mathbf{a} \in \mathbb{F}_q^n \backslash \mathcal{E}.$$

This immediately implies:

$$\frac{\partial \widetilde{Q}_j^{(i)}(\mathbf{x})}{\partial \mathbf{u}}(\mathbf{a}) \;=\; \frac{\partial Q_j^{(i)}(\mathbf{x})}{\partial \mathbf{u}}(\mathbf{a}), \qquad\qquad \textit{for every } \mathbf{a} \in \mathbb{F}_q^n \backslash \mathcal{E}.$$

Now we define the new polynomial $g(\mathbf{x})$ as follows:

$$g(\mathbf{x}) \;:=\; \widetilde{Q}_1^{(1)}(\mathbf{x}) \cdots \widetilde{Q}_a^{(1)}(\mathbf{x}) \;+\; \ldots \;+\; \widetilde{Q}_1^{(s)}(\mathbf{x}) \cdots \widetilde{Q}_a^{(s)}(\mathbf{x}).$$

We now show that g satisfies the three conditions mentioned in Lemma 5.3:

1. From the definition of $\mathcal{E}$, we have that for every $\mathbf{a} \in \mathbb{F}_q^n \backslash \mathcal{E}$, $g(\mathbf{a})$ equals $f(\mathbf{a})$.

2. From the definition of $\widetilde{Q}_j^{(i)}$, it is easy to see that if $Q_j^{(i)}(\mathbf{x})$ can be computed by a $\Sigma\Pi\Sigma$ circuit of size s, then $\widetilde{Q}_j^{(i)}$ can also be computed by a $\Sigma\Pi\Sigma$ circuit of size s and that the rank of $\widetilde{Q}_j^{(i)}$ is at most the rank of $Q_j^{(i)}$. Since the polynomial $f(\mathbf{x})$ can be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit of size s, then the polynomial $g(\mathbf{x})$ can also be computed by a $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit with bottom product gates of rank $\leqslant r$ of size s.

3. Let $S \subseteq \mathbb{F}_q^n \backslash \mathcal{E}$. Using Observation 5.4, we have that for any monomial $\mathbf{u}$ of degree at most k and any monomial $\mathbf{v}$ of degree ℓ,

$$\left(\mathbf{v} \cdot \frac{\partial g(\mathbf{x})}{\partial \mathbf{u}}\right)(\mathbf{a}) \;=\; \left(\mathbf{v} \cdot \frac{\partial f(\mathbf{x})}{\partial \mathbf{u}}\right)(\mathbf{a}), \qquad \text{for all } \mathbf{a} \in \mathbb{F}_q^n.$$

Thus $\Gamma_{k,\ell,S}(f)$ equals $\Gamma_{k,\ell,S}(g)$.

This finishes the proof of Lemma 5.3. ∎

Now we are ready to finish the proof of Theorem 1.1.

Proof of Theorem 1.1. Consider a $\Sigma\Pi^{[a]}\Sigma\Pi\Sigma$ circuit of size s computing the polynomial $\mathsf{NW}_{d,m,e}(\mathbf{x})$ of degree d. For a sake of argument, assume that $s < \exp(20\eta\sqrt{d})$.

Let η, k, r, ℓ, m, e be as in Lemma 5.3. Let $r = \dfrac{(\beta-\alpha)}{8\beta^2\eta q} \cdot \sqrt{d}$. These parameters satisfy the condition in Lemma 5.1.

$$\ell + rqk \;=\; \frac{n}{2}(1-\varepsilon) \;+\; 200q^2\eta^2 \cdot d \leqslant \frac{n}{2}(1-\varepsilon) \;+\; d < \frac{n}{2},$$

where for the final equality we used $d = \mathcal{O}(\sqrt{n})$.

Let $g(\mathbf{x})$ be the polynomial and $\mathcal{E}$ be the subset guaranteed from Lemma 5.3. We have that $|\mathcal{E}| \leqslant \delta \cdot q^n$ for $\delta = \exp(-\sqrt{d}/(q\sqrt{\beta}))$. Let $S \subseteq \mathbb{F}_q^n \backslash \mathcal{E}$ be the translate of Boolean hypercube from Lemma 3.12 such that

$$\Gamma_{k,\ell,S}(\mathsf{NW}_{m,d,e}) \;\geqslant\; \binom{n}{\ell + d - k} \cdot \exp(-\mathcal{O}(\log^2 d)).$$

From the second item of Lemma 5.3 and Lemma 5.1, we have

$$\Gamma_{k,\ell,S}(g) \;\leqslant\; s \cdot \binom{n}{\ell + rqk} \cdot n \cdot \binom{a}{k}.$$

From the third item of Lemma 5.3, we know

$$\Gamma_{k,\ell,S}(\mathsf{NW}_{m,d,e}) \;\leqslant\; s \cdot \binom{n}{\ell + rqk} \cdot n \cdot \binom{a}{k}.$$

The lower bound now follows from the same calculation as done in the proof of Theorem 4.1.

Indeed, the only modification in the parameters between Lemma 5.3 and Lemma 4.3 is that the parameter η is reduced by a factor of $1/\sqrt{q}$ (Notice that η appears implicitly in the definitions of k, r, and ε). On the other side, the term rk in Lemma 4.2 is replaced by the term rqk in Lemma 5.1. Consequently, in Equation (2), each term inside the exponential function is divided by $1/\sqrt{q}$. Hence, the final bound is the same up to the factor $1/\sqrt{q}$ in the exponent. This finishes the proof of Theorem 1.1. ■

6 Exponential Gap between Depth-3 and Homogeneous Depth-4 Circuits

In this section, we prove Theorem 1.5. To prove Theorem 1.5, we use the complexity measure defined by [GK98], that we define now.

Definition 6.1 (Grigoriev-Karpinski Measure). *Fix a field $\mathbb{F}_q$ and let $k, n \in \mathbb{N}$ be arbitrary. For a subset $\mathcal{E} \subseteq \mathbb{F}_q^n$, let $\overline{\mathcal{E}} := \mathbb{F}_q^n \backslash \mathcal{E}$. Then for every polynomial $P(x_1, \ldots, x_n)$,*

$$\mathsf{GK}_{k,\mathcal{E}}(P) \;:=\; \dim\left[\mathrm{Eval}_{\overline{\mathcal{E}}}\left\{\partial^{=k}(P)\right\}\right].$$

We first state an upper bound on the GK measure for polynomials that can be computed by small depth-3 circuits. The following lemma was implicitly used in [GK98] and also stated explicitly in [CM17, Lemma 7]. We use the following formulation from [Sap21] for simplicity.

Lemma 6.2 (Upper Bound of GK measure for Depth-3). *[Sap21, Lemma 10.4] Fix a field $\mathbb{F}_q$ and let $k \in \mathbb{N}$. Let $P \in \mathbb{F}_q[x_1, \ldots, x_n]$ be a polynomial that can be computed by a $\Sigma\Pi\Sigma$ circuit over $\mathbb{F}_q$ of size s. Then, for any $\tau > 0$ and $k \leqslant \tau/10q$, there exists a subset $\mathcal{E} \subseteq \mathbb{F}_q^n$ with $|\mathcal{E}| \leqslant s \cdot \exp(-\tau/8) \cdot q^n$ such that,*

$$\mathsf{GK}_{k,\mathcal{E}}(P) \;\leqslant\; s \cdot q^{\tau}.$$

Now we show a lower bound on the Grigoriev Karpinski measure for the polynomial PIP.

Lemma 6.3 (Lower Bound of GK measure for PIP). *Fix a field $\mathbb{F}_q$ and let $m, d \in \mathbb{N}$ be growing parameters. Then for every subset $\mathcal{E} \subseteq \mathbb{F}_q^{2md}$ with $|\mathcal{E}| < \exp(-\omega(d \log q)) \cdot q^{2md}$, we have,*

$$\mathsf{GK}_{d,\mathcal{E}}(\mathsf{PIP}_{m,d}(\mathbf{x}, \mathbf{y})) \;\geqslant\; m^d.$$

Proof of Lemma 6.3. We first observe that if we take the partial derivative of $\mathsf{PIP}_{m,d}$ with respect to a set-multilinear monomial $\mathbf{u}$ in $\mathbf{x}$ variables, then the partial derivative is the "counterpart"

of $\mathbf{u}$ in the $\mathbf{y}$-variables. More formally,

$$\frac{\partial^{=d}\ \mathsf{PIP}_{m,d}}{\partial\ (x_{1u_1}\cdots x_{du_d})} \;=\; y_{1u_1}\cdots y_{du_d}, \qquad \text{for every d-tuple } (u_1,\ldots,u_d)\in[m]^d.$$

This implies,

$$\left\{ y_{1u_1}\cdots y_{du_d} \;\middle|\; (u_1,\ldots,u_d)\in[m]^d \right\} \;\subseteq\; \partial^{=d}\left(\mathsf{PIP}_{m,d}\right).$$

Let $B_{m,d}$ be the set on the left side, i.e., $B_{m,d} := \left\{ y_{1u_1}\cdots y_{du_d} \;\middle|\; (u_1,\ldots,u_d)\in[m]^d \right\}$. Note that for every two distinct tuples $\mathbf{u}\in[m]^d$ and $\mathbf{v}\in[m]^d$, the monomials $y_{\mathbf{u}}$ and $y_{\mathbf{v}}$ are *distinct*. This implies that $B_{m,d}$ is a set of linearly independent monomials.

In the following claim, we show that the evaluation vectors of linearly independent degree-d polynomials over $\overline{\mathcal{E}}$ continue to be a linearly independent set.

Claim 6.4. *Let $B\subset\mathbb{F}_q[z_1,\ldots,z_n]$ be a set of non-zero $\mathbb{F}_q$-linearly independent degree-d polynomials and let $\mathcal{E}\subseteq\mathbb{F}_q^n$ be any subset with $|\mathcal{E}|<\exp(-\omega(d\log q))\cdot q^n$. Then the set*

$$\left\{ \mathrm{Eval}_{\overline{\mathcal{E}}}(Q) \;\middle|\; Q(\mathbf{z})\in B \right\},$$

is a linearly independent set.

Proof of Claim 6.4. We know that every non-zero degree-d polynomial over $\mathbb{F}_q$ is non-zero on at least $q^{-d}\cdot q^n$ points in $\mathbb{F}_q^n$. Any non-zero polynomial in the space spanned by polynomials in B is a non-zero degree-d polynomial, and thus is non-zero on at least $q^{-d}\cdot q^n=\exp(-d\log q)\cdot q^n$ points of $\mathbb{F}_q^n$. Fix any polynomial $Q\in\mathrm{span}(B)$. Since $|\mathcal{E}|<\exp(-\omega(d\log q))\cdot q^n$, there exists a point $\mathbf{a}\in\mathbb{F}_q^n\backslash\mathcal{E}$ such that $Q(\mathbf{a})\neq 0$. In other words, $\mathrm{Eval}_{\overline{\mathcal{E}}}(Q)$ is a non-zero evaluation vector. This implies that any non-zero linear combination of vectors in $\left\{\mathrm{Eval}_{\overline{\mathcal{E}}}(Q)\;\middle|\;Q(\mathbf{z})\in B\right\}$ is a non-zero vector. Hence, $\left\{\mathrm{Eval}_{\overline{\mathcal{E}}}(Q)\;\middle|\;Q(\mathbf{z})\in B\right\}$ is a linearly independent set. ■

Applying Claim 6.4 on the set $B_{m,d}$ and $\mathcal{E}$, we get,

$$\mathsf{GK}_{d,\mathcal{E}}(\mathsf{PIP}_{m,d}(\mathbf{x},\mathbf{y})) \;\geqslant\; \dim(B_{m,d}) \;=\; |B_{m,d}| \;=\; m^d.$$

This finishes the proof of Lemma 6.3. ■

We are now ready to finish the proof of Theorem 1.5, which we recall below.

Theorem 1.5 (Exponential Gap b/w Depth-3 and Homogeneous Depth-4). *Fix a field $\mathbb{F}_q$ and let $m,d\in\mathbb{N}$ be growing parameters. Any $\Sigma\Pi\Sigma$ circuit over $\mathbb{F}_q$ computing the degree-d polynomial $\mathsf{PIP}_{m,d}$ requires size $\exp(\Omega_q(d\log m))$.*

Proof of Theorem 1.5. Let $\tau = (d \log m)/(2 \log q)$ and let $s \leqslant \exp((1/32 \log q) \cdot d \log m)$. Suppose there is a $\Sigma\Pi\Sigma$ circuit over $\mathbb{F}_q$ of size s that computes the polynomial $\mathsf{PIP}_{m,d}(\mathbf{x}, \mathbf{y})$. Since $d \leqslant \tau/10q$ when m is sufficiently large, Lemma 6.2 guarantees the existence of a subset $\mathcal{E} \subseteq \mathbb{F}_q^{2md}$ with $|\mathcal{E}| \leqslant s \cdot \exp(-\tau/8) \cdot q^{2md} = \exp(-1/(32 \log q) \cdot d \log m) \cdot q^{2md}$ such that

$$\mathsf{GK}_{d,\mathcal{E}}(\mathsf{PIP}_{m,d}) \leqslant s \cdot q^{\tau}.$$

On the other hand, we also have $|\mathcal{E}| \leqslant \exp(-\omega(d \log q)) \cdot q^{2md}$ and thus by Lemma 6.3, we get,

$$\mathsf{GK}_{d,\mathcal{E}}(\mathsf{PIP}_{m,d}) \geqslant m^d = \exp(d \log m).$$

Comparing the two bounds on $\mathsf{GK}_{d,\mathcal{E}}(\mathsf{PIP}_{m,d})$, we get,

$$s \geqslant \exp(d \log m - \tau \log q) = \exp(\Omega(d \log m)).$$

This finishes the proof of Theorem 1.5. ∎

References

[AV08] Manindra Agrawal and V. Vinay. "Arithmetic circuits: A chasm at depth four". In: Cited by: 175; All Open Access, Green Open Access. 2008, pp. 67–75. DOI: 10.1109/FOCS.2008.32. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-58049118510&doi=10.1109%2fFOCS.2008.32&partnerID=40&md5=29915d441a65af6e353bc3d5ed556662 (cit. on pp. 4, 6).

[BS83] Walter Baur and Volker Strassen. "The complexity of partial derivatives". In: *Theoretical Computer Science* 22.3 (1983), pp. 317–330. ISSN: 0304-3975. DOI: https://doi.org/10.1016/0304-3975(83)90110-X. URL: https://www.sciencedirect.com/science/article/pii/030439758390110X (cit. on p. 3).

[BDS24] C.S. Bhargav, Sagnik Dutta, and Nitin Saxena. "Improved Lower Bound, and Proof Barrier, for Constant Depth Algebraic Circuits". In: *ACM Trans. Comput. Theory* 16.4 (Nov. 2024). ISSN: 1942-3454. DOI: 10.1145/3689957. URL: https://doi.org/10.1145/3689957 (cit. on p. 3).

[Bre74] Richard P. Brent. "The Parallel Evaluation of General Arithmetic Expressions". In: *J. ACM* 21.2 (Apr. 1974), pp. 201–206. ISSN: 0004-5411. DOI: 10.1145/321812.321815. URL: https://doi.org/10.1145/321812.321815 (cit. on p. 4).

[CM17] Suryajith Chillara and Partha Mukhopadhyay. "On the limits of depth reduction at depth 3 over small finite fields". In: *Information and Computation* 256 (2017),

pp. 35–44. ISSN: 0890-5401. DOI: https://doi.org/10.1016/j.ic.2017.04.007. URL: https://www.sciencedirect.com/science/article/pii/S0890540117300573 (cit. on pp. 1, 5, 6, 8, 32).

[For24] Michael A. Forbes. "Low-Depth Algebraic Circuit Lower Bounds over Any Field". In: *39th Computational Complexity Conference, CCC 2024, Ann Arbor, MI, USA, July 22-25, 2024*. Ed. by Rahul Santhanam. Vol. 300. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024, 31:1–31:16. DOI: 10.4230/LIPICS.CCC.2024.31. URL: https://doi.org/10.4230/LIPIcs.CCC.2024.31 (cit. on pp. 3, 5, 8).

[FLMS15] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. "Lower Bounds for Depth-4 Formulas Computing Iterated Matrix Multiplication". In: *SIAM Journal on Computing* 44.5 (2015), pp. 1173–1201. DOI: 10.1137/140990280. eprint: https://doi.org/10.1137/140990280. URL: https://doi.org/10.1137/140990280 (cit. on pp. 4, 9, 10).

[FLST24] Hervé Fournier, Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. "On the Power of Homogeneous Algebraic Formulas". In: *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*. Ed. by Bojan Mohar, Igor Shinkar, and Ryan O'Donnell. ACM, 2024, pp. 141–151. DOI: 10.1145/3618260.3649760. URL: https://doi.org/10.1145/3618260.3649760 (cit. on p. 6).

[FSS84] M. Furst, J.B. Saxe, and M. Sipser. "Parity, circuits, and the polynomial-time hierarchy." In: *Mathematical Systems Theory* 17 (1984), pp. 13–27. URL: https://doi.org/10.1007/BF01744431 (cit. on p. 3).

[GK98] Dima Grigoriev and Marek Karpinski. "An Exponential Lower Bound for Depth 3 Arithmetic Circuits". In: *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*. Ed. by Jeffrey Scott Vitter. ACM, 1998, pp. 577–582. DOI: 10.1145/276698.276872. URL: https://doi.org/10.1145/276698.276872 (cit. on pp. 1, 3, 5, 6, 11, 32).

[GR00] Dima Grigoriev and Alexander A. Razborov. "Exponential Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields". In: *Appl. Algebra Eng. Commun. Comput.* 10.6 (2000), pp. 465–487. DOI: 10.1007/S002009900021. URL: https://doi.org/10.1007/s002009900021 (cit. on pp. 3, 5).

[GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. "Approaching the Chasm at Depth Four". In: *J. ACM* 61.6 (2014), 33:1–33:16. DOI: 10.1145/2629541. URL: https://doi.org/10.1145/2629541 (cit. on p. 4).

[GKKS16] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. "Arithmetic Circuits: A Chasm at Depth 3". In: *SIAM Journal on Computing* 45.3 (2016), pp. 1064–1079. DOI: 10.1137/140957123. eprint: https://doi.org/10.1137/140957123. URL: https://doi.org/10.1137/140957123 (cit. on pp. 4–6).

[GST20] Nikhil Gupta, Chandan Saha, and Bhargav Thankey. "A Super-Quadratic Lower Bound for Depth Four Arithmetic Circuits". In: *35th Computational Complexity Conference, CCC 2020, Saarbrücken, Germany (Virtual Conference), July 28-31, 2020*. Ed. by Shubhangi Saraf. Vol. 169. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020, 23:1–23:31. DOI: 10.4230/LIPICS.CCC.2020.23. URL: https://doi.org/10.4230/LIPIcs.CCC.2020.23 (cit. on p. 5).

[Hås87] Johan Torkel Håstad. *Computational limitations for small-depth circuits*. Cambridge, MA, USA: MIT Press, 1987. ISBN: 0262081679 (cit. on p. 3).

[Kay12] Neeraj Kayal. "An exponential lower bound for the sum of powers of bounded degree polynomials". In: *Electron. Colloquium Comput. Complex.* TR12-081 (2012). ECCC: TR12-081. URL: https://eccc.weizmann.ac.il/report/2012/081 (cit. on p. 4).

[KLSS14] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. "Super-polynomial lower bounds for depth-4 homogeneous arithmetic formulas". In: *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*. Ed. by David B. Shmoys. ACM, 2014, pp. 119–127. DOI: 10.1145/2591796.2591823. URL: https://doi.org/10.1145/2591796.2591823 (cit. on pp. 4, 9, 10).

[KLSS17] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. "An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas". In: *SIAM Journal on Computing* 46.1 (2017), pp. 307–335. DOI: 10.1137/151002423. eprint: https://doi.org/10.1137/151002423. URL: https://doi.org/10.1137/151002423 (cit. on pp. 4, 9, 10, 18).

[KS17a] Neeraj Kayal and Chandan Saha. "Multi-k-ic depth three circuit lower bound". In: *Theory of Computing Systems* 61.4 (2017), pp. 1237–1251 (cit. on p. 6).

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. "A super-polynomial lower bound for regular arithmetic formulas". In: *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*. Ed. by David B. Shmoys. ACM, 2014, pp. 146–153. DOI: 10.1145/2591796.2591847. URL: https://doi.org/10.1145/2591796.2591847 (cit. on p. 4).

[KST16a] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. "An Almost Cubic Lower Bound for Depth Three Arithmetic Circuits". In: *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*. Ed. by Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi. Vol. 55. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016, 33:1–33:15. DOI: 10.4230/LIPICS.ICALP.2016.33. URL: https://doi.org/10.4230/LIPIcs.ICALP.2016.33 (cit. on p. 3).

[KST16b] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. "On the size of homogeneous and of depth four formulas with low individual degree". In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016, pp. 626–632 (cit. on p. 6).

[Koi12] Pascal Koiran. "Arithmetic Circuits: The Chasm at Depth Four Gets Wider". In: *Theor. Comput. Sci.* 448 (Aug. 2012), pp. 56–65. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2012.03.041. URL: https://doi.org/10.1016/j.tcs.2012.03.041 (cit. on pp. 4, 6).

[KS19] Mrinal Kumar and Ramprasad Saptharishi. "The Computational Power of Depth Five Arithmetic Circuits". In: *SIAM J. Comput.* 48.1 (2019), pp. 144–180. DOI: 10.1137/18M1173319. URL: https://doi.org/10.1137/18M1173319 (cit. on pp. 1, 4, 6, 7, 9–11, 16–19, 26).

[KS14] Mrinal Kumar and Shubhangi Saraf. "Superpolynomial Lower Bounds for General Homogeneous Depth 4 Arithmetic Circuits". In: *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I*. Ed. by Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias. Vol. 8572. Lecture Notes in Computer Science. Springer, 2014, pp. 751–762. DOI: 10.1007/978-3-662-43948-7_62. URL: https://doi.org/10.1007/978-3-662-43948-7%5C_62 (cit. on pp. 4, 9, 10).

[KS17b] Mrinal Kumar and Shubhangi Saraf. "On the Power of Homogeneous Depth 4 Arithmetic Circuits". In: *SIAM J. Comput.* 46.1 (2017), pp. 336–387. DOI: 10.1137/140999335. URL: https://doi.org/10.1137/140999335 (cit. on pp. 4, 9, 10, 18).

[LY22] Jiatu Li and Tianqi Yang. "3.1n - o(n) circuit lower bounds for explicit functions". In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2022. Rome, Italy: Association for Computing Machinery, 2022, pp. 1180–1193. ISBN: 9781450392648. DOI: 10.1145/3519935.3519976. URL: https://doi.org/10.1145/3519935.3519976 (cit. on p. 3).

[LST25] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. "Superpolynomial Lower Bounds Against Low-Depth Algebraic Circuits". In: *J. ACM* 72.4 (July 2025). ISSN: 0004-5411. DOI: 10.1145/3734215. URL: https://doi.org/10.1145/3734215 (cit. on pp. 3, 5, 8).

[Nis91] Noam Nisan. "Lower Bounds for Non-Commutative Computation (Extended Abstract)". In: *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, May 5-8, 1991, New Orleans, Louisiana, USA*. Ed. by Cris Koutsougeras and Jeffrey Scott Vitter. ACM, 1991, pp. 410–418. DOI: 10.1145/103418.103462. URL: https://doi.org/10.1145/103418.103462 (cit. on p. 3).

[NW97] Noam Nisan and Avi Wigderson. "Lower Bounds on Arithmetic Circuits Via Partial Derivatives". In: *Comput. Complex.* 6.3 (1997), pp. 217–234. DOI: 10.1007/BF01294256. URL: https://doi.org/10.1007/BF01294256 (cit. on p. 3).

[Raz09] Ran Raz. "Multi-linear formulas for permanent and determinant are of super-polynomial size". In: *Journal of the ACM (JACM)* 56.2 (2009), pp. 1–17 (cit. on p. 6).

[Raz87] Alexander A. Razborov. "Lower bounds on the size of bounded depth circuits over a complete basis with logical addition". In: *Mathematical notes of the Academy of Sciences of the USSR* 41 (1987), pp. 333–338. URL: https://api.semanticscholar.org/CorpusID:121744639 (cit. on p. 3).

[Sap21] Ramprasad Saptharishi. "A Survey of Lower Bounds in Arithmetic Circuit Complexity". In: *Github survey* (2021). URL: https://github.com/dasarpmar/lowerbounds-survey/releases/tag/v9.0.3 (cit. on pp. 3, 6, 14, 32).

[Sha17] Abhijat Sharma. "An Improved Lower Bound for Depth Four Arithmetic Circuits". Available at https://www.csa.iisc.ac.in/~chandan/thesis_reports/AbhijatSharma_MScThesis.pdf. Master's thesis. Indian Institute of Science, Bangalore, July 2017 (cit. on p. 5).

[SW01] Amir Shpilka and Avi Wigderson. "Depth-3 arithmetic circuits over fields of characteristic zero". In: *Comput. Complex.* 10.1 (2001), pp. 1–27. DOI: 10.1007/PL00001609. URL: https://doi.org/10.1007/PL00001609 (cit. on pp. 3, 6).

[SY10] Amir Shpilka and Amir Yehudayoff. "Arithmetic Circuits: A Survey of Recent Results and Open Questions". In: *Found. Trends Theor. Comput. Sci.* 5.3–4 (Mar. 2010), pp. 207–388. ISSN: 1551-305X. DOI: 10.1561/0400000039. URL: https://doi.org/10.1561/0400000039 (cit. on p. 3).

[Smo87] R. Smolensky. “Algebraic methods in the theory of lower bounds for Boolean circuit complexity”. In: *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*. STOC ’87. New York, New York, USA: Association for Computing Machinery, 1987, pp. 77–82. ISBN: 0897912217. DOI: 10.1145/28395.28404. URL: https://doi.org/10.1145/28395.28404 (cit. on p. 3).

[Tav15] Sébastien Tavenas. “Improved bounds for reduction to depth 4 and depth 3”. In: *Information and Computation* 240 (2015). MFCS 2013, pp. 2–11. ISSN: 0890-5401. DOI: https://doi.org/10.1016/j.ic.2014.09.004. URL: https://www.sciencedirect.com/science/article/pii/S0890540114001138 (cit. on pp. 4, 6).

[Val79] L. G. Valiant. “Completeness Classes in Algebra”. In: *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*. STOC ’79. Atlanta, Georgia, USA: Association for Computing Machinery, 1979, pp. 249–261. ISBN: 9781450374385. DOI: 10.1145/800135.804419. URL: https://doi.org/10.1145/800135.804419 (cit. on pp. 3, 14).

[VSBR83] L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. “Fast Parallel Computation of Polynomials Using Few Processors”. In: *SIAM Journal on Computing* 12.4 (1983), pp. 641–644. DOI: 10.1137/0212043. eprint: https://doi.org/10.1137/0212043. URL: https://doi.org/10.1137/0212043 (cit. on p. 4).

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il