

An Almost Cubic Lower Bound for Depth Three Arithmetic Circuits*

Neeraj Kayal¹, Chandan Saha², and Sébastien Tavenas³

1 Microsoft Research India, Bangalore, India

neeraka@microsoft.com

2 Indian Institute of Science, Bangalore, India

chandan@csa.iisc.ernet.in

3 Microsoft Research India, Bangalore, India

sebastien.tavenas@ens-lyon.org

Abstract

We show an almost cubic lower bound on the size of any depth three arithmetic circuit computing an explicit multilinear polynomial in n variables over any field. This improves upon the previously known quadratic lower bound by Shpilka and Wigderson [CCC, 1999].

1998 ACM Subject Classification F.1.1 Models of Computation

Keywords and phrases arithmetic circuits, depth-3 circuits, shifted partials

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.33

1 Introduction

An arithmetic circuit is a directed acyclic graph with leaves (nodes with in-degree zero) labeled by formal variables and other nodes labeled by addition (+) or multiplication (\times) operations. Nodes with out-degree zero are the output nodes; for simplicity and without losing generality we will assume that there is only one output node in a circuit. Non-leaf nodes are also referred to as addition or multiplication *gates*. Such a circuit naturally represents a multivariate polynomial; we say this polynomial is *computed* at the output node of the circuit (or simply computed by the circuit). Two parameters that determine the complexity of a circuit are its *size* and *depth*, which are respectively the number of edges and the length of the longest path from any input node to the output node of the circuit. Computations involving arithmetic operations can be naturally modeled by arithmetic circuits and hence study of these objects forms a fundamental aspect of complexity theory.

Research on arithmetic circuits received a great impetus from the seminal paper by Valiant [42] who defined two non-uniform complexity classes that are algebraic analogues of classes P and NP. These algebraic complexity classes are known as VP and VNP in the literature. Class VP consists of families of polynomials $\{g_n\}_{n \geq 1}$ such that the number of variables and the degree of g_n are $n^{O(1)}$, and there is an arithmetic circuit of size $n^{O(1)}$ computing g_n . A family of polynomials $\{f_n\}_{n \geq 1}$ is in VNP if there is another family of polynomials $\{g_n(\mathbf{x}, \mathbf{y})\}_{n \geq 1}$ in VP such that $f_n = \sum_{\mathbf{y} \in \{0,1\}^{|\mathbf{y}|}} g_n(\mathbf{x}, \mathbf{y})$. Valiant defined a notion of completeness for the classes VNP and VP, and showed that the family of permanent polynomials is VNP-complete whereas the family of determinant polynomials is *almost* complete for VP. This gave rise

* A full version of the paper (which contains all the missing proofs) can be found at <http://eccc.hpi-web.de/report/2016/006/>



© Neeraj Kayal, Chandan Saha, and Sébastien Tavenas;

licensed under Creative Commons License CC-BY

43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016).

Editors: Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi;

Article No. 33; pp. 33:1–33:15



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



to the famous ‘determinantal complexity of the permanent’ problem, a suitable resolution of which would imply $VP \neq VNP$ or equivalently a super-polynomial size lower bound for arithmetic circuits. We refer the reader to the surveys [26, 39], the book [4] and the paper [27] for more on these and other related algebraic complexity classes, their inter-relationships and their associations with Boolean complexity classes. Throughout this article, whenever we use the term ‘circuit(s)’ we will mean ‘arithmetic circuit(s)’.

Starting with Valiant’s work there has been significant progress in proving lower bounds for several restricted models of arithmetic circuits. Multilinear [31, 30, 34], noncommutative [28, 25], monotone [13] and special low-depth circuits [29, 8, 36, 35, 32, 1, 15, 16, 21, 19] are examples of such interesting circuit classes. But still, our knowledge of general circuit lower bound is rather limited. The best known lower bound for general circuits is Baur and Strassen’s $\Omega(n \log d)$ bound [40, 3] for circuits computing the simple polynomial $\sum_{i \in [n]} x_i^d$. A recent line of work on depth reduction, starting with [2, 43] and culminating with [20, 10, 41], has shown that a moderately strong lower bound for circuits of *depth three*¹ implies a super-polynomial lower bound for general circuits. Also, Raz [33] showed that a strong enough lower bound for a special kind of (namely, set-multilinear) depth three circuits implies a super-polynomial lower bound for general arithmetic formulas². These depth reduction results have opened up the possibility of proving a super-polynomial lower bound for general circuits/formulas by first proving strong lower bounds for low-depth, in particular depth three, circuits. The hope is depth three circuits, which have an apparent simple structure, might be more amenable to lower bound proofs. But, unfortunately, even at depth three we do not know of any super-polynomial lower bound over fields of characteristic zero!

Depth three circuits. In this paper, whenever we mention a depth three circuit we will mean a $\Sigma\Pi\Sigma$ circuit that has an addition gate at the top, followed by a layer of multiplication gates and finally a bottom layer of sum gates. Such a circuit is a “sum of product of linear polynomials” representation of the computed polynomial. The fan-in of the top addition gate is called the top fan-in, and that of the bottom layer of addition gates the bottom fan-in of the circuit. Observe that bottom fan-in can be at most $n + 1$ where n is the number of variables. The *multiplicative complexity* of a depth three circuit C is the sum of the fan-ins of the multiplication gates of the circuit, i.e. if $C = \sum_{i=1}^s l_{i1} \cdots l_{id_i}$ where l_{ij} ’s are linear polynomials then multiplicative complexity of C is $\sum_i d_i$. It is easy to see that multiplicative complexity is less than the size of a depth three circuit. Circuit C is *homogeneous* if l_{ij} ’s are homogeneous linear polynomials (a.k.a. linear forms).

Previous works on depth three circuit lower bound. In [38, 36], Shpilka and Wigderson proved an $\Omega(n^2)$ lower bound on the multiplicative complexity of depth three circuits computing the elementary symmetric polynomial

$$\text{ESYM}_n^d(x_1, \dots, x_n) \stackrel{\text{def}}{=} \sum_{S \subseteq [n], |S|=d} \prod_{i \in S} x_i$$

on n -variables and degree $d = \Theta(n)$. This bound is essentially optimal for fields of size more than n , as n -variate elementary symmetric polynomials can be computed by depth three

¹ over fields of characteristic zero

² a *formula* is a circuit whose underlying directed acyclic graph is a tree

circuits with multiplicative complexity $O(n^2)$ ³. A similar tight quadratic lower bound but for the power symmetric polynomial $\sum_{i \in [n]} x_i^n$ was shown in [12]. Also, a near quadratic lower bound is known for the determinant polynomial [38, 36]. The situation is a lot better over small fields or under the restriction of homogeneity. An exponential lower bound was shown by [8] (and by [9]) for depth three circuits over any fixed finite field computing the determinant polynomial (even if the circuit and the determinant are treated in the algebra of functions over the finite field). It was shown in [29] that any homogeneous depth three circuit computing ESYM_n^{2d} has size $\Omega((n/4d)^d)$. Recently, [16] showed a lower bound of $n^{\Omega(\sqrt{d})}$ for depth three circuits, with bottom fan-in bounded by n^ε for any fixed $\varepsilon < 1$, computing an explicit n -variate polynomial of degree d .

1.1 Our results

► **Theorem 1** (Depth-3 circuit lower bound). *There is a family of homogeneous multilinear polynomials $\{f_n\}_{n \geq 1}$ in VNP, where f_n is a $\Theta(n)$ -variate polynomial of degree $\Theta(n)$ such that any depth-3 circuit computing f_n has multiplicative complexity (and hence size) $\Omega\left(\frac{n^3}{(\ln n)^2}\right)$.*

Theorem 1 can be seen as an improvement in the state of the art of the long-standing quadratic lower bound for depth three circuits [38, 36], although our target polynomial family is harder – it is in VNP and not known to be in VP. Also, from our analysis, we arrive at a near quadratic lower bound for the *symmetric model* defined in [37] thereby improving upon the linear bound therein (Theorem 2).

Let ESYM_m^d be an elementary symmetric polynomial in m variables and of degree d . Borrowing terminologies from [37], a *symmetric circuit* has a bottom layer of plus gates computing linear polynomials, and a top gate that computes some elementary symmetric polynomial on the linear polynomials computed at the bottom level gates. Thus, a symmetric circuit with m bottom level gates outputs a polynomial of the form $\text{ESYM}_m^d(l_1, \dots, l_m)$ for some d , where l_1, \dots, l_m are linear polynomials computed by the m bottom level gates. The parameter m is defined as the *size* of the symmetric circuit. This model was shown to be complete or universal in [37] (i.e. every polynomial can be computed in this model), and linear lower bounds were shown on the size of the smallest symmetric circuit computing the determinant polynomial and the polynomial $\prod_{i=1}^{n/2} x_i + \prod_{i=n/2+1}^n x_i$. The following theorem improves this lower bound but once again the target polynomial family is likely harder than the ones studied in [37].

► **Theorem 2** (Symmetric circuit lower bound). *Let $\{f_n\}_{n \geq 1}$ be the polynomial family of Theorem 1. The size of the smallest symmetric circuit computing f_n is $\Omega\left(\frac{n^2}{(\ln n)^2}\right)$ over any infinite field.*

In an attempt to make progress in understanding lower bounds for circuit models where formal degree of the circuit is much higher than the number of variables (as might be the case for a depth three circuit), [17] posed the problem of proving lower bounds for homogeneous depth three circuits with formal degree much larger than the number of variables. The following theorem gives a solution to this problem.

► **Theorem 3** (Homogeneous depth three circuits with high degree). *For any positive integer $d = d(n) \geq n$, there exists an explicit family $\{f_{n,d}\}$ of n -variate polynomials of degree d*

³ this follows from an interpolation trick attributed to Michael Ben-Or in [29]

such that any homogeneous depth three circuit computing $f_{n,d}$ must have size at least $2^{\Omega(n)}$. Moreover, one can even choose such a family $f_{n,d}$ so that it can in fact be computed by a $(nd)^{O(1)}$ -sized algebraic branching program⁴.

The above theorem can be viewed as a generalization of the lower bound by [29] for homogeneous depth three circuits. Since elementary symmetric polynomials in n -variables have degree at most n , the lower bound in [29] holds for homogeneous depth three circuits with degree less than the number of variables. To the best of our knowledge, a lower bound of $(nd)^{\omega(1)}$ for homogeneous depth three circuits with degree d much greater than the number of variables n was not known. Theorem 3 fills in this gap in our understanding as long as $d = 2^{o(n)}$. However, note that the lower bound in the above theorem is independent of d , ideally one should get $d^{\Omega(n)}$ instead of $2^{\Omega(n)}$.

1.2 Proof ideas

Like in many of the previous works, we use a measure $\mu : \mathbb{F}[\mathbf{x}] \rightarrow \mathbb{N}$ to capture some ‘weakness’ of a circuit family as opposed to a ‘hard’ family of polynomials which leads to a lower bound for the circuit family. In both Theorem 1 and 3, the improvements are achieved by applying the *dimension of the shifted partials* measure, introduced in [14], and used subsequently (at times with certain crucial alterations) in many other recent lower bound results [11, 18, 7, 15, 22, 16, 24, 21, 19, 23]. The shifted partials measure is a generalization of the dimension of the partial derivatives measure used previously in [29, 36]. It is quite effective in proving lower bounds for the model of depth four ($\Sigma\Pi\Sigma\Pi$) circuits with formal degree close to the actual degree of the computed polynomial, and somewhat low bottom fan-in [11, 18]. In fact, all the recent lower bounds (for restricted depth 3 and 4 circuits) obtained using shifted partials ‘reduce’ to this case of depth four circuits one way or the other. We take a similar route here, but make the crucial observation that a simple “grouping” step in the analysis with shifted partials gives some *leeway to the formal degree* of the circuit and allows it to grow over the actual degree of the computed polynomial. This observation and a careful construction of the target family of polynomials to take advantage of this leeway are the primary sources of improvement of the depth three lower bound.

An immediate hurdle in proving lower bounds for depth three circuits is that the formal degree of the circuit can be much larger than the degree and number of variables of the computed polynomial. The existing proof techniques and measures have had limited success in handling high formal degree circuits [16, 23]. To get around this first hurdle, we begin by following the same approach as in [36] of pruning the circuit of high degree product gates by going modulo some linear polynomials picked from among the factors of such ‘heavy’ product gates. This step is exactly (borrowing terminologies from [36]) satisfying some affine linear constraints and restricting the circuit to an affine subspace. However, the degree threshold used to define ‘heavy’ product gates can now be chosen higher than that in [36] because of the ‘leeway to formal degree’ provided by shifted partials. In the pruned circuit, a simple “grouping” of linear polynomials in every product term of a depth three circuit turns out to be surprisingly effective in handling the remaining product gates. The grouping step transforms a depth three circuit to a depth four circuit with bottom fan-in more than 1, but at the same time brings down the number of factors in every product term. The tradeoff between the bottom fan-in and the number of factors per product term is then analyzed to obtain a suitable upper bound on the shifted partials dimension of a depth three circuit.

⁴ The definition of an algebraic branching program can be found, for example, in [39].

Finally, in order to maximize the gain and obtain a near cubic bound we need an explicit multilinear polynomial with *degree linear in the number of variables*, and that has close to the maximum possible shifted partials dimension even when restricted to an affine subspace. The polynomial family $\{f_n\}_{n \geq 1}$ in Theorem 1 is a variant of the family of Nisan-Wigderson polynomials used in [18, 15]. A notable difference between the Nisan-Wigderson families used in earlier works and the one used here is that the degree of f_n is linearly related to its number of variables, unlike $d = n^{o(1)}$ in previous works. Although, a greedy construction of a Nisan-Wigderson family can make degree $\Theta(n)$, it is not clear if such a family is in VNP. To ensure both – a VNP family and linear degree – we construct a family by ‘composing’ two smaller families of Nisan-Wigderson polynomials, one is obtained by a greedy algorithm and the other explicitly defined in [18, 15]. A detailed description of the polynomial family is given in Section 6.

Few more details on the polynomial families. Polynomial f_n in Theorem 1 is homogeneous with three sets of variables $\mathbf{u}, \mathbf{y}, \mathbf{x}$ such that $|\mathbf{u}| = |\mathbf{y}| = |\mathbf{x}| = \frac{10n}{9}$. (To avoid a few ceil and floor notations in the analysis, we shall assume without any loss of generality that n is divisible by $1872 = 9 \cdot 13 \cdot 16$.) Let $\mathbf{u} = \{u_1, \dots, u_{\frac{10n}{9}}\}$, $\mathbf{y} = \{y_1, \dots, y_{\frac{10n}{9}}\}$ and $\mathbf{x} = \{x_1, \dots, x_{\frac{10n}{9}}\}$. Every monomial of f_n is a product of a \mathbf{u} -monomial of degree $d_{\mathbf{u}} = n$, a \mathbf{y} -monomial of degree $d_{\mathbf{y}} = \lfloor \ln n \rfloor$, and an \mathbf{x} -monomial of degree $d_{\mathbf{x}} \in [\frac{2n}{13}, \frac{n}{3}]$. Thus the number of variables and the degree of f_n are both $\Theta(n)$. The \mathbf{x} and the \mathbf{y} variables are the primary variables; derivatives of f_n of order $\lfloor \ln n \rfloor$ with respect to the \mathbf{y} -variables give rise to \mathbf{x} -monomials with large ‘pairwise distance’ that help estimate the shifted partials dimension of the target polynomial. The \mathbf{u} -variables are auxiliary variables which ensure that the measure remains high for the target polynomial even when restricted to an affine subspace.

The polynomial family $\{f_{n,d}\}$ used in Theorem 3 is a simple variant of the multi- r -ic iterated matrix multiplication polynomial family used in [19].

1.3 Organization

Sections 3 to 6 are devoted to the proofs of Theorem 1 and 2. We prove Theorem 3 in Section 7.

2 Preliminaries

2.1 Basic notations

For any $m \in \mathbb{N}$, the set of natural numbers, the set $\{1, \dots, m\}$ is denoted by $[m]$. We use upper-case letters (like A or S) to denote sets of numbers, calligraphic upper-case letters (like \mathcal{B}, \mathcal{D} or \mathcal{L}) to denote sets of polynomials, and bold lower-case letters (like \mathbf{x} or \mathbf{y}) to denote sets of variables. When the base ring of polynomials is clear from the context, the ideal generated by a set of polynomials of the ring, say \mathcal{L} , is denoted by $\langle \mathcal{L} \rangle$. A circuit is denoted using typewriter font, as in \mathbb{C} or \mathbb{D} . For a set of numbers $S \subseteq [m]$, \bar{S} denote the complement of S . Sometimes, we use the notation $\text{poly}(n)$ to mean $n^{O(1)}$.

2.2 The measure

Although, the results in this paper can be derived using the shifted partials measure as it is in [14], we choose to work with a variant of this measure for better clarity in the analysis. This variant is similar in outlook to the *shifted skewed partials* measure used recently in [19],

although for our application there is no difference (or skew) between the number of \mathbf{x} and \mathbf{y} variables. Such a skew between $|\mathbf{y}|$ and $|\mathbf{x}|$ was important for the results in [19].

Let $A \subseteq \left[\frac{10n}{9}\right]$ of size $|A| = n$. Let $\mathbf{x}_A = \{x_i : i \in A\}$ and $g(\mathbf{y}, \mathbf{x}_A) \in \mathbb{F}[\mathbf{y}, \mathbf{x}_A]$. For $k, \ell \in \mathbb{N}$, define the measure $\text{SP}_{k,\ell,A} : \mathbb{F}[\mathbf{y}, \mathbf{x}_A] \rightarrow \mathbb{N}$ as follows.

$$\text{SP}_{k,\ell,A}(g) \stackrel{\text{def}}{=} \dim(\mathbf{x}_A^{\leq \ell} \cdot \sigma_{\mathbf{y}}(\partial_{\mathbf{y}}^k g)),$$

where $\partial_{\mathbf{y}}^k g$ is the set of all k -th order partial derivatives of g with respect to the \mathbf{y} -variables, and $\sigma_{\mathbf{y}} : \mathbb{F}[\mathbf{y}, \mathbf{x}_A] \rightarrow \mathbb{F}[\mathbf{x}_A]$ is a map that sets all the \mathbf{y} -variables to zero. Naturally, $\sigma_{\mathbf{y}}$ is a homomorphism from $\mathbb{F}[\mathbf{y}, \mathbf{x}_A]$ to $\mathbb{F}[\mathbf{x}_A]$, and $\sigma_{\mathbf{y}}(\mathcal{D})$ is defined by $\{\sigma_{\mathbf{y}}(h) : h \in \mathcal{D}\}$ for any set of polynomials $\mathcal{D} \subseteq \mathbb{F}[\mathbf{y}, \mathbf{x}_A]$. $\mathbf{x}_A^{\leq \ell}$ is the set of all monomials in the \mathbf{x}_A -variables of degree ℓ or less. For two sets of polynomials \mathcal{B} and \mathcal{D} , $\mathcal{B} \cdot \mathcal{D} \stackrel{\text{def}}{=} \{h_1 \cdot h_2 : h_1 \in \mathcal{B} \text{ and } h_2 \in \mathcal{D}\}$, and the dimension of a set of polynomials \mathcal{D} (denoted by $\dim(\mathcal{D})$) is the dimension of the vector space spanned by the polynomials in \mathcal{D} over the field \mathbb{F} .

It is worth noting that the above measure (as in [19]) can be thought of as a hybrid of the *rank of the partial derivatives matrix* measure of [28] and the shifted partials measure of [14]. The former measure has been refined and used in several other subsequent work, most notably in [31, 35], and is also identified with the *evaluation dimension* measure in [6] over fields of characteristic zero. The following proposition is easy to verify.

► **Proposition 4** (Sub-additivity). *For any $k, \ell \in \mathbb{N}$, $\mathbf{x}_A \subseteq \mathbf{x}$ and $g_1, g_2 \in \mathbb{F}[\mathbf{y}, \mathbf{x}_A]$,*

$$\text{SP}_{k,\ell,A}(g_1 + g_2) \leq \text{SP}_{k,\ell,A}(g_1) + \text{SP}_{k,\ell,A}(g_2).$$

3 Lower bounding the measure for the target polynomial family

We will show that the measure SP (from Section 2.2) is considerably large when applied suitably to the polynomial family $\{f_n\}_{n \geq 1}$. The precise statement is given in the theorem below.

Polynomials restricted to an affine subspace. Let $S \subseteq \left[\frac{10n}{9}\right]$ be a set of size $\frac{n}{9}$ and

$$\mathcal{L}_S = \{x_i - h_i\}_{i \in S} \tag{1}$$

be a set of $|\mathcal{L}_S| = |S| = \frac{n}{9}$ linear polynomials in $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}]$ such that $h_i \in \mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}]$ for every $i \in S$, where $\bar{S} = \left[\frac{10n}{9}\right] \setminus S$.

Denote the ideal of $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}]$ generated by the linear polynomials of \mathcal{L}_S by $\langle \mathcal{L}_S \rangle$. For any polynomial $f \in \mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}]$, let

$$f_{\langle \mathcal{L}_S \rangle} \stackrel{\text{def}}{=} f \pmod{\langle \mathcal{L}_S \rangle}$$

be the image of the polynomial f in the ring $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}]/\langle \mathcal{L}_S \rangle$. Since $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}]/\langle \mathcal{L}_S \rangle$ is isomorphic to $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}]$, $f_{\langle \mathcal{L}_S \rangle}$ can be represented by a polynomial in the ring $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}]$; this polynomial is obtained from f by replacing x_i by h_i for all $i \in S$. So, we will treat $f_{\langle \mathcal{L}_S \rangle}$ as an element of $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}]$.

Finally, let $f_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0}$ be the polynomial obtained from $f_{\langle \mathcal{L}_S \rangle} \in \mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}]$ by setting the \mathbf{u} -variables to 0/1-values as follows: $u_i = 0$ if $i \in S$, else $u_i = 1$. We will describe the family $\{f_n\}_{n \geq 1}$ and prove the following theorem in Section 6.

► **Theorem 5.** *Let n be the parameter that defines the polynomial family $\{f_n\}_{n \geq 1}$. Let $k = \lfloor \ln n \rfloor$, q be the smallest prime greater or equal to $\lceil \frac{n}{1000 \cdot \ln n} \rceil$, $\ell = \lfloor \frac{n^2}{32 \cdot k \cdot \ln q} \rfloor$. Then for every set $S \subseteq [\frac{10n}{9}]$ of size $|S| = \frac{n}{9}$, and every set of linear polynomials \mathcal{L}_S as in Equation (1), and $f = f_n$,*

$$\text{SP}_{k,\ell,\bar{S}}(f_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0}) \geq \frac{1}{2} \cdot q^k \cdot \binom{n+\ell}{n}.$$

Next, we show an upperbound of the measure for a depth-3 circuit and prove Theorem 1.

4 Upper bounding the measure for a depth three circuit

Pruning ‘heavy’ product gates from a depth three circuit. Let $\mathcal{C} = \sum_{i=1}^s T_i$ be a depth three circuit computing $f = f_n$, where T_i is a product term⁵ of \mathcal{C} . Let c_0 be a constant to be fixed later in the analysis. Then either of the following two cases is obviously true.

Case 1: The number of product terms of \mathcal{C} , with \mathbf{x} -degree greater or equal to $\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \rfloor$, is greater than $\frac{n}{9}$.

Case 2: The number of product terms of \mathcal{C} , with \mathbf{x} -degree greater or equal to $\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \rfloor$, is less than or equal to $\frac{n}{9}$.

If Case 1 is true then the multiplicative complexity of \mathcal{C} is at least $\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \rfloor \cdot \frac{n}{9} = \Omega(\frac{n^3}{(\ln n)^2})$ as $d_{\mathbf{x}} \in [\frac{2n}{13}, \frac{n}{3}]$ and we have nothing to prove in this case. If Case 2 is true then we can find a ‘few’ linear polynomials such that modulo these the circuit is free of ‘heavy’ product terms. This is stated formally in the lemma below and the corollary thereafter, and is directly inspired by a similar argument in [38, 36]. However, the threshold chosen to define ‘heavy’ product gates in [38, 36] is linear in n , whereas the one here has an extra $\frac{d_{\mathbf{x}}}{(\ln n)^2}$ factor that finally accounts for the improvement in the lower bound. As mentioned in Section 1, this is the leeway to the formal degree of the circuit provided by the analysis with shifted partials.

► **Lemma 6.** *Suppose the number of product terms of \mathcal{C} , with \mathbf{x} -degree greater or equal to $\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \rfloor$, is bounded by $\frac{n}{9}$. Then, there is a set $S \subseteq [\frac{10n}{9}]$ of size $\frac{n}{9}$ and a set of linear polynomials,*

$$\mathcal{L}_S = \{x_i - h_i\}_{i \in S}, \text{ where } h_i \text{ is a linear polynomial in } \mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}] \text{ for every } i \in S,$$

such that $f_{\langle \mathcal{L}_S \rangle} \in \mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}]$ is computed by a depth three circuit, say $\mathcal{C}_{\langle \mathcal{L}_S \rangle}$, satisfying the following:

1. top fan-in of $\mathcal{C}_{\langle \mathcal{L}_S \rangle}$ is upper bounded by the top fan-in of \mathcal{C} ,
2. every product term of $\mathcal{C}_{\langle \mathcal{L}_S \rangle}$ has \mathbf{x} -degree upper bounded by $\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \rfloor$.

The proof of the lemma is relatively straightforward and can be found in the full version paper.

► **Corollary 7.** *Polynomial $f_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0} \in \mathbb{F}[\mathbf{y}, \mathbf{x}_{\bar{S}}]$ is computed by a depth three circuit, say $\mathcal{C}_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0}$, with top fan-in bounded by the top fan-in of \mathcal{C} and every product term of $\mathcal{C}_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0}$ has \mathbf{x} -degree bounded by $\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \rfloor$.*

⁵ a product term corresponds to a multiplication gate of \mathcal{C}

Let us denote the circuit $\mathcal{C}_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0}$ by \mathcal{D} . Let $\mathcal{D} = \sum_{i=1}^s P_i$, where a term P_i is a product of linear polynomials in $\mathbb{F}[\mathbf{y}, \mathbf{x}_{\bar{S}}]$. Note that the pruned circuit \mathcal{D} has only \mathbf{y} and $\mathbf{x}_{\bar{S}}$ variables. Now, as the degrees of the intermediate gates are bounded, we can upperbound the shifted partial measure for \mathcal{D} by using a grouping argument (as it was done in [19]). We get then the next lemma.

► **Lemma 8.** *Let $k, \ell \in \mathbb{N}$ be as in Theorem 5. Then*

$$\text{SP}_{k, \ell, \bar{S}}(\mathcal{D}) \leq s \cdot \binom{\lceil 32c_0 d_{\mathbf{x}} \rceil}{k} \cdot \binom{n + \ell + kt}{n}, \text{ where } t = \left\lceil \frac{n}{32 \cdot (\ln n)^2} \right\rceil.$$

5 Putting together: Proof of Theorem 1

Let \mathcal{C} be a depth three circuit computing f_n . Then, as explained in Section 4, we have two cases to handle. In Case 1, the multiplicative complexity of \mathcal{C} is already $\Omega(\frac{n^3}{(\ln n)^2})$ and we have nothing to prove. Whereas, in Case 2, the circuit can be pruned of heavy product gates so that the polynomial $f_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0} \in \mathbb{F}[\mathbf{y}, \mathbf{x}_{\bar{S}}]$ is computed by a depth three circuit, say \mathcal{D} , whose top fan-in is upper bounded by the top fan-in of \mathcal{C} (by Corollary 7). Moreover, every product term of \mathcal{D} has \mathbf{x} -degree bounded by $\left\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \right\rfloor$ so that Lemma 8 is applicable now. The computations of the next lemma can be found in the full version paper.

► **Lemma 9.** *In Case 2, the top fan-in of \mathcal{D} (hence also the top fan-in of \mathcal{C}) is $\omega(n^3)$.*

Proof. By Theorem 5 and Lemma 8, the top fan-in s of \mathcal{D} can be lower bounded as follows:

$$s \geq \frac{1}{2} \cdot \frac{q^k \cdot \binom{n+\ell}{n}}{\binom{\lceil 32c_0 d_{\mathbf{x}} \rceil}{k} \cdot \binom{n+\ell+kt}{n}},$$

which will imply after some computations that $s = \omega(n^3)$. ◀

Thus, in Case 2, the top fan-in of \mathcal{D} (and hence \mathcal{C}) must be $\omega(n^3)$ and therefore putting Case 1 and 2 together, the multiplicative complexity of \mathcal{C} is $\min\{\Omega(\frac{n^3}{(\ln n)^2}), \omega(n^3)\} = \Omega(\frac{n^3}{(\ln n)^2})$ for sufficiently large n .

5.1 Proof of Theorem 2

The proof follows from Lemma 9. Suppose $f = f_n$ is computed by a symmetric circuit where l_1, \dots, l_m are the bottom level linear polynomials. Naturally, $f = \text{ESYM}_m^d(l_1, \dots, l_m)$ for some d , and hence (by Ben-Or's interpolation trick over any field of size more than m) f is also computed by a depth three circuit \mathcal{C} with top fan-in $m+1$ and degree of every product term bounded by m . If $m \geq \left\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \right\rfloor$ then we have nothing to prove. Suppose $m < \left\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \right\rfloor$. Then the condition of Case 2 (in Section 4) is satisfied as *every* product term of \mathcal{C} has \mathbf{x} -degree (in fact, total degree) bounded by $m < \left\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \right\rfloor$. But then, Lemma 9 tells us that \mathcal{C} has top fan-in $\omega(n^3)$ which contradicts with the fact that the top fan-in is $m+1 = O(n^2)$. So, it must be that $m \geq \left\lfloor \frac{c_0 n d_{\mathbf{x}}}{(\ln n)^2} \right\rfloor = \Omega\left(\frac{n^2}{(\ln n)^2}\right)$.

6 The polynomial family and proof of Theorem 5

6.1 Construction of the Nisan-Wigderson polynomial family

Let $\mathbf{z} = \{z_1, \dots, z_n\}$ be a set of n formal variables. For any two multilinear monomials m_1 and m_2 in the \mathbf{z} -variables of degree $d_{\mathbf{z}}$ each, let $|m_1 \cap m_2|$ be the number of variables common

between m_1 and m_2 . Define *distance* between the monomials m_1, m_2 as,

$$\Delta(m_1, m_2) \stackrel{\text{def}}{=} d_{\mathbf{z}} - |m_1 \cap m_2|.$$

As in the statement of Theorem 5, let q be the smallest prime greater or equal to $\lceil \frac{n}{1000 \ln n} \rceil$ and $k = \lfloor \ln n \rfloor$. The following lemma plays a central role in the construction of the polynomial family $\{f_n\}_{n \geq 1}$.

► **Lemma 10.** *There is a family $\{g_n(\mathbf{z})\}_{n \geq 1}$ in VNP such that $g_n(\mathbf{z})$ is a homogeneous multilinear polynomial of degree $d_{\mathbf{z}} \in [\frac{2n}{13}, \frac{n}{3}]$ in n \mathbf{z} -variables, and $\Delta(m_1, m_2) \geq \frac{n}{16}$ for any pair of distinct monomials m_1 and m_2 of g_n . Further, g_n is a sum of q^k distinct monomials.*

In the previous lemma, we need to find a family whose monomials are pairwise distant, such that the degree is linear in the number of variables and the family is in VNP. The Nisan-Wigderson polynomial family in [18] is in VNP but its degree is not linear. On the other hand, one can greedily get a family such that the degree is linear but which is not known to be in VNP. We show in the full version paper that by ‘composing’ these two families one can get both the desired properties.

The family $\{f_n\}_{n \geq 1}$. Let (m_1, \dots, m_{q^k}) be an ordered sequence of monomials of the polynomial $g_n(\mathbf{z})$ from the above lemma under lexicographic monomial ordering $z_1 \succ \dots \succ z_n$. Let $\mathbf{w} = \{w_1, \dots, w_n\}$ be n formal variables different from \mathbf{z} . The number of multilinear monomials in \mathbf{w} -variables of degree k is $\binom{n}{k} \geq (\frac{n}{k})^k = (\frac{n}{\lfloor \ln n \rfloor})^k \geq q^k$. Under lexicographic monomial ordering $w_1 \succ \dots \succ w_n$, let $(\beta_1, \dots, \beta_{q^k})$ be the ordered sequence of the first q^k monomials among all multilinear monomials in the \mathbf{w} -variables of degree k . Define the polynomial $F_n(\mathbf{w}, \mathbf{z})$ as,

$$F_n(\mathbf{w}, \mathbf{z}) \stackrel{\text{def}}{=} \sum_{j=1}^{q^k} \beta_j m_j. \tag{2}$$

Now let $\mathbf{u} = \{u_1, \dots, u_{\frac{10n}{9}}\}$, $\mathbf{y} = \{y_1, \dots, y_{\frac{10n}{9}}\}$ and $\mathbf{x} = \{x_1, \dots, x_{\frac{10n}{9}}\}$ be the sets of variables on which $f_n(\mathbf{u}, \mathbf{y}, \mathbf{x})$ is defined as follows.

$$f_n(\mathbf{u}, \mathbf{y}, \mathbf{x}) \stackrel{\text{def}}{=} \sum_{\substack{A \subseteq [\frac{10n}{9}] \\ |A|=n}} \prod_{i \in A} u_i \cdot F_n(\mathbf{y}_A, \mathbf{x}_A). \tag{3}$$

We assume the lexicographic order $x_1 \succ \dots \succ x_{\frac{10n}{9}}$ and $y_1 \succ \dots \succ y_{\frac{10n}{9}}$. The polynomial $F_n(\mathbf{y}_A, \mathbf{x}_A)$ is obtained by substituting the \mathbf{y}_A -variables $\{y_i : i \in A\}$ in place of the \mathbf{w} -variables and \mathbf{x}_A -variables $\{x_i : i \in A\}$ in place of the \mathbf{z} -variables such that the underlying lexicographic orders, $z_1 \succ \dots \succ z_n$ and $w_1 \succ \dots \succ w_n$, are obeyed. Note that $d_{\mathbf{y}} = \deg_{\mathbf{y}} f_n = k$, $d_{\mathbf{u}} = \deg_{\mathbf{u}} f_n = n$ and $d_{\mathbf{x}} = \deg_{\mathbf{x}} f_n = \deg_{\mathbf{z}} g_n = d_{\mathbf{z}} \in [\frac{2n}{13}, \frac{n}{3}]$. Further, the polynomial family $\{f_n\}_{n \geq 1}$ is in VNP: It would be clear from the proof of Lemma 10 that the computational problem of finding the ‘index’ of a given monomial in g_n can be solved in $\text{poly}(n)$ time, which in turn implies the coefficient of a given monomial in f_n can be found in $\text{poly}(n)$ time. The *index* of a monomial m in g_n is the position of m in the lexicographically ordered list of q^k monomials of g_n .

6.2 Proof of Theorem 5: The measure on the polynomial family

In this section, we show that the relevant measure is high for the family of polynomials (defined in Equation 3) even when restricted to an affine subspace. As in Section 3 (Equation

33:10 An Almost Cubic Lower Bound for Depth Three Arithmetic Circuits

1), let $S \subseteq [\frac{10n}{9}]$ be a set of size $\frac{n}{9}$. Let $\mathcal{L}_S = \{x_i - h_i\}_{i \in S}$ be any set of $\frac{n}{9}$ linear polynomials in $\mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}]$ such that $h_i \in \mathbb{F}[\mathbf{u}, \mathbf{y}, \mathbf{x}_{\bar{S}}]$ for every $i \in S$, where $\bar{S} = [\frac{10n}{9}] \setminus S$. Let $f = f_n(\mathbf{u}, \mathbf{y}, \mathbf{x})$ (as defined in Equation 3).

► **Observation 11.** $f_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0} = F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}}) \in \mathbb{F}[\mathbf{y}, \mathbf{x}_{\bar{S}}]$.

Proof. The polynomial $f_{\langle \mathcal{L}_S \rangle, \mathbf{u}_S=0}$ is obtained from f by substituting every x_i by h_i for every $i \in S$, and then setting $u_j = 0$ for every $j \in S$ and $u_j = 1$ otherwise. Since the only \mathbf{x} -variables occurring in $F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}})$ are from $\mathbf{x}_{\bar{S}}$, it remains untouched by the above substitutions. Finally, the setting of the \mathbf{u} -variables retains only $F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}})$ from the sum in Equation 3. ◀

So, we need to show that

$$\text{SP}_{k,\ell,\bar{S}}(F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}})) \geq \frac{1}{2} \cdot q^k \cdot \binom{n+\ell}{n}.$$

This part of the argument bears close resemblance to and is inspired by similar arguments in [7, 5]. We begin with the following observation.

► **Observation 12.** *The set $\partial_{\mathbf{y}}^{\leq k} F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}})$ consists of exactly the monomials of $g_n(\mathbf{x}_{\bar{S}})$. Hence, $\sigma_{\mathbf{y}}(\partial_{\mathbf{y}}^{\leq k} F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}}))$ also consists of exactly the monomials of $g_n(\mathbf{x}_{\bar{S}})$.*

Proof. Follows easily from the definition of the polynomial F_n in Equation 2. ◀

Reusing notation, let the monomials of $g_n(\mathbf{x}_{\bar{S}})$ be $\{m_1, \dots, m_{q^k}\}$ – these are monomials in $\mathbf{x}_{\bar{S}}$ -variables. By Lemma 10, $\Delta(m_i, m_j) \geq \frac{n}{16}$ for every $i \neq j$ and $\frac{2n}{13} \leq \deg(m_i) \leq \frac{n}{3}$ for every $i \in [q^k]$. Let

$$B_i \stackrel{\text{def}}{=} \mathbf{x}_{\bar{S}}^{\leq \ell} \cdot m_i, \quad \text{for } i \in [q^k].$$

Then,

$$\begin{aligned} \dim(\mathbf{x}_{\bar{S}}^{\leq \ell} \cdot \sigma_{\mathbf{y}}(\partial_{\mathbf{y}}^{\leq k} F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}}))) &= |B_1 \cup \dots \cup B_{q^k}| \\ \Rightarrow \text{SP}_{k,\ell,\bar{S}}(F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}})) &\geq \sum_{i=1}^{q^k} |B_i| - \frac{1}{2} \cdot \sum_{\substack{i,j \\ i \neq j}} |B_i \cap B_j| \\ &= q^k \cdot \binom{n+\ell}{n} - \frac{1}{2} \cdot \sum_{\substack{i,j \\ i \neq j}} |B_i \cap B_j|, \end{aligned} \quad (4)$$

as $|\bar{S}| = n$ and $|B_i| = \binom{n+\ell}{n}$.

► **Proposition 13.** *For every $i, j \in [q^k]$ and $i \neq j$, $|B_i \cap B_j| \leq \binom{n+\ell-n/16}{n}$.*

Proof. If a monomial m belongs to both B_i and B_j then $m = s_1 \cdot m_i = s_2 \cdot m_j$ where $\deg(s_1), \deg(s_2) \leq \ell$. Since $\Delta(m_i, m_j) \geq n/16$,

$$m = s' \cdot \frac{m_j}{\gcd(m_i, m_j)} \cdot m_i, \quad \text{where } \deg(s') \leq \ell - \frac{n}{16}.$$

Hence, the number of such monomials m is bounded by $\binom{n+\ell-n/16}{n}$. ◀

Therefore, by Equation (4), after some computations.

► **Claim 14.** $\text{SP}_{k,\ell,\bar{S}}(F_n(\mathbf{y}_{\bar{S}}, \mathbf{x}_{\bar{S}})) \geq \frac{1}{2} \cdot q^k \cdot \binom{n+\ell}{n}$.

This completes the proof of Theorem 5.

7 Homogeneous depth three circuits with large degree

We prove Theorem 3 in this section. The measure remains the same as before, but the notation is simplified a little bit (as we do not need to include a subset of variables in the definition of the measure). For any $g \in \mathbb{F}[\mathbf{y}, \mathbf{x}]$, define the measure $\text{SP}_{k,\ell} : \mathbb{F}[\mathbf{y}, \mathbf{x}] \rightarrow \mathbb{N}$ as

$$\text{SP}_{k,\ell}(g) \stackrel{\text{def}}{=} \dim(\mathbf{x}^{\leq \ell} \cdot \sigma_{\mathbf{y}}(\partial_{\mathbf{y}}^{\leq k} g)).$$

Like before, the measure is sub-additive, i.e. for $g_1, g_2 \in \mathbb{F}[\mathbf{y}, \mathbf{x}]$ and $k, \ell \in \mathbb{N}$,

$$\text{SP}_{k,\ell}(g_1 + g_2) \leq \text{SP}_{k,\ell}(g_1) + \text{SP}_{k,\ell}(g_2).$$

Moreover, the measure is invariant under multiplication by any fixed polynomial from $\mathbb{F}[\mathbf{x}]$ (the proof of the following lemma is very simple and appears in the full version paper):

► **Lemma 15.** *For any $g \in \mathbb{F}[\mathbf{y}, \mathbf{x}]$, $h \in \mathbb{F}[\mathbf{x}]$ and $k, \ell \in \mathbb{N}$, $\text{SP}_{k,\ell}(h \cdot g) = \text{SP}_{k,\ell}(g)$.*

The outline of the proof of Theorem 3 also remains the same: we show a suitable upper bound on the measure for the circuit, and a lower bound for the target family of polynomials. The target family of polynomials is basically a *multi-r-ic* variant of the iterated matrix multiplication polynomial defined and analysed in [19] – we will recall some parts of the analysis from there to lower bound the measure for the family of polynomials. Furthermore, this polynomial can be computed by an algebraic branching program of size polynomial in the number of variables and degree of the polynomial.

7.1 Upper bound for the circuit

Let \mathcal{C} be any homogeneous depth three circuit computing a polynomial in n variables $\mathbf{y} \uplus \mathbf{x}$ and of degree d . More precisely, by identifying the circuit with the polynomial it computes, $\mathcal{C} = T_1 + T_2 + \dots + T_s$, where the T_i 's are products of d homogeneous linear polynomials i.e. $T_i = l_{i1} \cdot l_{i2} \cdot \dots \cdot l_{id}$, where every l_{ij} is a linear form. Let us consider any one product term, say T . By grouping t linear forms together and multiplying the linear forms within each group, we obtain $T = Q_1 \cdot \dots \cdot Q_{\lceil \frac{d}{t} \rceil}$, where $\deg(Q_j) \leq t$ for every $j \in [\lceil \frac{d}{t} \rceil]$. By sub-additivity of the measure and following a similar argument as in the proof of Lemma 8, we get the following lemma.

► **Lemma 16.** *For any $k, \ell \in \mathbb{N}$ and $t \leq d$,*

$$\text{SP}_{k,\ell}(\mathcal{C}) \leq s \cdot \binom{\lceil \frac{d}{t} \rceil}{k} \cdot \binom{|\mathbf{x}| + \ell + kt}{|\mathbf{x}|}. \quad (5)$$

7.2 Lower bound for the polynomial family

The polynomial family. We define a polynomial on n variables $\mathbf{y} \uplus \mathbf{x}$ and of degree d , where d is any integer greater or equal to n .

For $w, k, r, \alpha \in \mathbb{N}$, consider the following polynomial.

$$F_{w,k,r,\alpha}(\mathbf{y}, \mathbf{x}) \stackrel{\text{def}}{=} g_1(\mathbf{y}_1, \mathbf{x}_1) \cdot g_2(\mathbf{y}_2, \mathbf{x}_2) \cdot \dots \cdot g_k(\mathbf{y}_k, \mathbf{x}_k),$$

where the g_i 's are polynomials over the indicated (disjoint) subsets of variables $\mathbf{y} = \mathbf{y}_1 \uplus \dots \uplus \mathbf{y}_k$ and $\mathbf{x} = \mathbf{x}_1 \uplus \dots \uplus \mathbf{x}_k$, and defined as,

$$g_i(\mathbf{y}_i, \mathbf{x}_i) \stackrel{\text{def}}{=} \sum_{a,b \in [w]} y_{i,a,b} \cdot \prod_{c \in [\alpha]} x_{i,c,a}^r \cdot x_{i,c+\alpha,b}^r.$$

33:12 An Almost Cubic Lower Bound for Depth Three Arithmetic Circuits

The number of \mathbf{y} -variables is $|\mathbf{y}| = kw^2$ and the number of \mathbf{x} -variables is $|\mathbf{x}| = 2k\alpha w$. The total number of variables in $F_{w,k,r,\alpha}$ is $(w^2 + 2\alpha w) \cdot k$, and it has degree $\tilde{d} = (2\alpha r + 1) \cdot k$. Our target polynomial is almost $F_{w,k,r,\alpha}$, except that we multiply it with a suitable power of a variable just to match its degree with the given degree parameter d which is any number more than the number of variables.

Let $n = (w^2 + 2\alpha w) \cdot k$ and $d \geq n$ be a given degree parameter. In the analysis, we eventually fix α and w to integer constants so that $n = \Theta(k)$. Set $r = \lceil \frac{d}{3\alpha k} \rceil$ and x be any arbitrarily fixed variable in \mathbf{x} . Our polynomial family $\{f_{n,d}\}$ is defined by

$$f_{n,d} \stackrel{\text{def}}{=} x^{d-\tilde{d}} \cdot F_{w,k,r,\alpha}. \quad (6)$$

This polynomial is well defined, i.e. $d \geq \tilde{d}$, as soon as $w \geq 3$. Observe that $f_{n,d}$ has the same set of n variables as $F_{w,k,r,\alpha}$ and has degree d . Let us record the values for k and r for the analysis later $k = \frac{n}{w^2+2\alpha w}$ and $r = \lceil \frac{d}{3\alpha k} \rceil$. Also, note that $f_{n,d}$ can be computed by a $\text{poly}(n, d)$ -size ABP.

The measure on the polynomial family. The following lemma was essentially proved in [19] (see Section 7.5 in there) with slightly different notations.

► **Lemma 17.** *Let $0 < \delta \leq 1/5$ be a constant and $w \geq 3$.*

1. *Then*

$$\text{SP}_{k,\ell}(F_{w,k,r,\alpha}) \geq M \cdot \binom{|\mathbf{x}| + \ell}{|\mathbf{x}|} - \frac{M^2}{2} \cdot \binom{|\mathbf{x}| + \ell - \lceil \delta k \rceil \cdot \alpha r}{|\mathbf{x}|} \text{ with } M = \left(\left\lfloor \frac{w^{2-\delta}}{2} \right\rfloor \right)^k.$$

2. *Moreover, if $\ell \geq |\mathbf{x}|$ and $2 \cdot |\mathbf{x}| \cdot \alpha r \geq \ell \beta \cdot \ln w$ where $\beta \geq 4(2-\delta)/\delta$ is a constant then we can also conclude that $\text{SP}_{k,\ell}(F_{w,k,r,\alpha})$ is lower bounded by $M \cdot \binom{|\mathbf{x}| + \ell}{|\mathbf{x}|} / 2$.*

For the choice of parameters below, $\frac{w^{2-\delta}}{2}$ is an integer. Hence, $M = \left(\frac{w^{2-\delta}}{2} \right)^k$.

► **Corollary 18.** *If the conditions of Lemma 17 are satisfied then it follows from Lemma 15*

$$\text{SP}_{k,\ell}(f_{n,d}) \geq \frac{M}{2} \cdot \binom{|\mathbf{x}| + \ell}{|\mathbf{x}|}.$$

7.3 Putting together: Proof of Theorem 3

Let us choose $t = \lfloor 2\varepsilon\alpha r \rfloor$, and $\ell = \lfloor \frac{|\mathbf{x}| \cdot t}{\varepsilon \beta \cdot \ln w} \rfloor$ with the following parameters $\alpha = 18$, $\delta = \frac{1}{5}$, $\beta = 36$, $\varepsilon = \frac{1}{200}$, and $w = 2^{10}$.

We can notice that $t > 0$ and $\lceil d/t \rceil \leq (2k)/\varepsilon$. Furthermore, the conditions $\ell \geq |\mathbf{x}|$, $2 \cdot |\mathbf{x}| \cdot \alpha r \geq \ell \beta \cdot \ln w$, and $\beta \geq 4(2-\delta)/\delta$ are satisfied. Hence, if \mathbf{C} is a homogeneous depth three circuit computing $f_{n,d}$, then by Lemma 16 and Corollary 18,

$$s \cdot \binom{\lceil \frac{d}{t} \rceil}{k} \cdot \binom{|\mathbf{x}| + \ell + kt}{|\mathbf{x}|} \geq \text{SP}_{k,\ell}(f_{n,d}) \geq \frac{M}{2} \cdot \binom{|\mathbf{x}| + \ell}{|\mathbf{x}|}.$$

It implies, after some computations that $s \geq 2^{\Omega(n)}$.

References

- 1 Manindra Agrawal, Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. Jacobian hits circuits: hitting-sets, lower bounds for depth- d occur- k formulas & depth-3 transcendence degree- k circuits. In *STOC*, pages 599–614, 2012.
- 2 Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 67–75, 2008.
- 3 W. Baur and V. Strassen. The complexity of partial derivatives. *Theoretical Computer Science*, 22(3):317–330, 1983.
- 4 Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*. Springer, 2000.
- 5 Suryajith Chillara and Partha Mukhopadhyay. Depth-4 lower bounds, determinantal complexity: A unified approach. In *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5-8, 2014, Lyon, France*, pages 239–250, 2014. doi:10.4230/LIPIcs.STACS.2014.239.
- 6 Michael A. Forbes and Amir Shpilka. Quasipolynomial-Time Identity Testing of Non-commutative and Read-Once Oblivious Algebraic Branching Programs. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 243–252, 2013.
- 7 Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. In *STOC*, pages 128–135, 2014.
- 8 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*, pages 577–582, 1998.
- 9 Dima Grigoriev and Alexander A. Razborov. Exponential complexity lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. In *39th Annual Symposium on Foundations of Computer Science, FOCS'98, November 8-11, 1998, Palo Alto, California, USA*, pages 269–278, 1998.
- 10 Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *Foundations of Computer Science (FOCS)*, pages 578–587, 2013.
- 11 Ankit Gupta, Neeraj Kayal, Pritish Kamath, and Ramprasad Saptharishi. Approaching the chasm at depth four. In *Conference on Computational Complexity (CCC)*, pages 65–73, 2013.
- 12 Maurice J. Jansen and Kenneth W. Regan. "Resistant" Polynomials and Stronger Lower Bounds for Depth-Three Arithmetical Formulas. In *Computing and Combinatorics, 13th Annual International Conference, COCOON 2007, Banff, Canada, July 16-19, 2007, Proceedings*, pages 470–481, 2007.
- 13 Mark Jerrum and Marc Snir. Some exact complexity results for straight-line computations over semirings. *J. ACM*, 29(3):874–897, 1982.
- 14 Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012.
- 15 Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. In *Foundations of Computer Science (FOCS)*, pages 61–70, 2014.
- 16 Neeraj Kayal and Chandan Saha. Lower Bounds for Depth Three Arithmetic Circuits with small bottom fanin. In *Conference on Computational Complexity*, pages 158–208, 2015.

- 17 Neeraj Kayal and Chandan Saha. Multi-k-ic depth three circuit lower bound. In *32nd International Symposium on Theoretical Aspects of Computer Science (STACS 2015)*, volume 30, pages 527–539, 2015.
- 18 Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *STOC*, pages 146–153, 2014.
- 19 Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. On the size of homogeneous and of depth four formulas with low individual degree. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:181, 2015.
- 20 Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012.
- 21 Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. *CoRR*, abs/1507.00177, 2015. URL: <http://arxiv.org/abs/1507.00177>.
- 22 Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*, pages 364–373, 2014.
- 23 Mrinal Kumar and Shubhangi Saraf. Arithmetic circuits with locally low algebraic rank. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:194, 2015.
- 24 Mrinal Kumar and Shubhangi Saraf. Sums of products of polynomials in few variables : lower bounds and polynomial identity testing. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:71, 2015.
- 25 Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for non-commutative skew circuits. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:22, 2015.
- 26 Meena Mahajan. Algebraic complexity classes. *CoRR*, abs/1307.3863, 2013. URL: <http://arxiv.org/abs/1307.3863>.
- 27 Guillaume Malod and Natacha Portier. Characterizing Valiant’s algebraic complexity classes. *J. Complex.*, 24(1):16–38, 2008.
- 28 Noam Nisan. Lower bounds for non-commutative computation (extended abstract). In *STOC*, pages 410–418, 1991.
- 29 Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1996.
- 30 Ran Raz. Separation of multilinear circuit and formula size. *Theory of Computing*, 2(1):121–135, 2006.
- 31 Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2), 2009.
- 32 Ran Raz. Elusive functions and lower bounds for arithmetic circuits. *Theory of Computing*, 6(1):135–177, 2010.
- 33 Ran Raz. Tensor-rank and lower bounds for arithmetic formulas. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 659–666, 2010.
- 34 Ran Raz and Amir Yehudayoff. Balancing syntactically multilinear arithmetic circuits. *Computational Complexity*, 17(4):515–535, 2008.
- 35 Ran Raz and Amir Yehudayoff. Lower Bounds and Separations for Constant Depth Multilinear Circuits. *Computational Complexity*, 18(2):171–207, 2009.
- 36 A. Shpilka and A. Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001.
- 37 Amir Shpilka. Affine projections of symmetric polynomials. In *Proceedings of the 16th Annual IEEE Conference on Computational Complexity, Chicago, Illinois, USA, June 18-21, 2001*, pages 160–171, 2001.

- 38 Amir Shpilka and Avi Wigderson. Depth-3 arithmetic formulae over fields of characteristic zero. In *IEEE Conference on Computational Complexity*, pages 87–, 1999. Available at <http://eccc.hpi-web.de/report/1999/023/>.
- 39 Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.
- 40 V. Strassen. Die Berechnungskomplexität von elementarsymmetrischen Funktionen und von Interpolationskoeffizienten. *Numerische Mathematik*, 20:238–251, 1973.
- 41 Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013.
- 42 L. G. Valiant. Completeness Classes in Algebra. In *STOC'79: Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 249–261, New York, NY, USA, 1979. ACM Press.
- 43 L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. Fast parallel computation of polynomials using few processors. *SIAM Journal on Computing*, 12(4):641–644, 1983.