On $\Sigma\wedge\Sigma\wedge\Sigma$ Circuits: The Role of Middle Σ Fan-in, Homogeneity and Bottom Degree.

Christian Engels[1], B. V. Raghavendra Rao[2], and Karteek Sreenivasaiah[3]

1 **Tokyo Institute of Technology, Tokyo, Japan**
engels@is.titech.ac.jp*

2 **Indian Institute of Technology Madras, Chennai, India**
bvrr@cse.iitm.ac.in

3 **Max-Planck-Institute for Informatics and Saarland University, Saarbrücken, Germany**
karteek@mpi-inf.mpg.de†

Abstract

We study polynomials computed by depth five $\Sigma\wedge\Sigma\wedge\Sigma$ circuits, i.e., polynomials of the form $\sum_{i=1}^{t} Q_i$ where $Q_i = \sum_{j=1}^{r_i} \ell_{ij}^{d_{ij}}$, ℓ_{ij} are linear forms and r_i, $t \geq 0$. These circuits are a natural generalization of the well known class of $\Sigma\wedge\Sigma$ circuits and received significant attention recently. We prove exponential lower bound for the monomial $x_1 \cdots x_n$ against the following sub-classes of $\Sigma\wedge\Sigma\wedge\Sigma$ circuits:

- Depth four $\Sigma\wedge\Sigma\wedge$ arithmetic circuits.
- Depth five $\Sigma\wedge\Sigma^{[\leq n]}\wedge^{[\geq 21]}\Sigma$ and $\Sigma\wedge\Sigma^{[\leq 2^{\sqrt{n}}/1000]}\wedge^{[\geq\sqrt{n}]}\Sigma$ arithmetic circuits where the bottom Σ gate is homogeneous;

Our results show precisely how the fan-in of the middle Σ gates, the degree of the bottom powering gates and the homogeneity at the bottom Σ gates play a crucial role in the computational power of $\Sigma\wedge\Sigma\wedge\Sigma$ circuits.

1998 ACM Subject Classification F. Theory of Computation

Keywords and phrases Arithmetic Circuits, Polynomial Identity Testing, Computational Complexity

Digital Object Identifier 10.4230/LIPIcs...

1 Introduction

Arithmetic circuits were introduced by Valiant [23] as a natural model for algebraic computation where he defined Algebraic Complexity classes based on various complexity measures on arithmetic circuits. In the same article, Valiant conjectured that the permanent polynomial, perm_n, does not have polynomial size arithmetic circuits. Following Valiant's work, there have been intensive research efforts towards the resolution of Valiant's hypothesis. Further, obtaining super polynomial size lower bounds for arithmetic circuits computing explicit polynomials is a pivotal problem in Algebraic Complexity Theory. However, for general classes of arithmetic circuits, the best known lower bound is barely superlinear [2].

Lack of progress on lower bounds against general arithmetic circuits lead researchers to explore restricted classes of circuits. Grigoriev and Karpinski [7] proved an exponential

* This work is supported in part by ELC project (KAKENHI No. 24106608).
† This author is supported by the IMPECS post doctoral fellowship.

ISSN 1433-8092

size lower bound for depth three circuits computing the permanent over finite fields of fixed size. However, extending these results to infinite fields or depth four arithmetic circuits remains elusive. Agrawal and Vinay [1] (see also [22, 13]) explained this lack of progress by establishing that proving exponential lower bounds against depth four arithmetic circuits is enough to resolve Valiant's conjecture. This was strengthened further to depth three circuits over infinite fields by Gupta et al. [9].

In the first major attempt at crossing the chasm at depth four, Gupta et al. [10] obtained a $2^{\Omega(\sqrt{n})}$ size lower bound for depth four homogeneous circuits computing perm_n where the fan-in of the bottom product gate is bounded by $O(\sqrt{n})$. Following this, Fournier et al. [6] obtained a super polynomial lower bound against depth four homogeneous circuits computing a polynomial in VP. Further, the techniques in [10, 11] have been generalized and applied to prove lower bounds against various classes of constant depth arithmetic circuits for polynomials in VP as well as in VNP. (See e.g., [19] and references therein.)

Most of the lower bound proofs against arithmetic circuits follow a common framework: 1) define a measure for polynomials that is sub-additive and/or sub-multiplicative, 2) show that the circuit class of interest has small measure and 3) show that the target polynomial has high measure. The number of complexity measures that are known to be sub-additive and/or sub-multiplicative is very limited and are mostly based on the space of partial derivatives. The space of partial derivatives was first used by Nisan and Wigderson [18] to prove lower bounds against a class of set multilinear arithmetic circuits which was further applied to several other special classes of arithmetic circuits. (See [20] for a survey.) The lower bound in [10] was achieved by a relatively new complexity measure, the dimension of the shifted partial derivatives, i.e., derivatives multiplied with monomials of certain degree. The dimension of the space of shifted partial derivatives and their projections were used extensively in subsequent works (see for e.g., [19] and references therein). Further, the dimension of evaluation of the polynomials for a subset of variables, viz., evaluation dimension [4, 19] was studied and used as a measure for arithmetic formulas in several recent lower bound results.(See [19] for details). Recent attempts at arithmetic analogues of natural proofs [8, 5] indicate that the complexity measures discussed above might be insufficient to resolve Valiant's hypothesis.

Apart from the complexity measure based framework mentioned above, there have been two other prominent approaches towards a resolution of Valiant's hypothesis: A geometric approach by Mulmuley and Sohoni [17] and an approach based on the real τ conjecture proposed by Shub and Smale [21].

The geometric approach to complexity theory [17] involves the study of class of varieties associated with each of the complexity classes and studying their representations.

The real τ conjecture of Shub and Smale [21] states that the number of real roots of a univariate polynomial computed by an arithmetic circuit of size s is bounded by a polynomial in s. Koiran [12] showed that any resolution of the real τ-conjecture or an integer variant of it, would imply a positive resolution of Valiant's hypothesis. There has been several approaches towards the resolution of the real τ-conjecture and it's variants by Koiran et al. [15, 14].

Circuit Model

We consider the class of depth five powering circuits, i.e., $\Sigma \wedge \Sigma \wedge \Sigma$ circuits. It was shown in [9] that any homogeneous polynomial f of degree d over a sufficiently large field computed by a circuit of size s can also be computed by a homogeneous $\Sigma \wedge^{[a]} \Sigma \wedge^{[d/a]} \Sigma$ circuit of size $s^{\sqrt{d \log n \log(sd)}}$ for any $1 < a < d$ where the superscript $[a]$ for a gate denotes the fan-in (degree in the case of $\wedge$ gates) at that level. This was an intermediary step in [9] which went on to obtain a depth three $\Sigma\Pi\Sigma$ circuit of size $2^{\sqrt{d \log n \log(sd)}}$ for f.

Thus, combined with the results in [22], to prove Valiant's hypotheses over infinite fields, it is enough to prove a $2^{\omega(\sqrt{n}\log n)}$ size lower bound against any one of the following classes of circuits: (1) homogeneous depth four $\Sigma\Pi^{[\sqrt{n}]}\Sigma\Pi^{[O(\sqrt{n})]}$ circuits, (2) homogeneous depth five $\Sigma\wedge^{[\sqrt{n}]}\Sigma\wedge^{[O(\sqrt{n})]}\Sigma$ circuits or (3) depth three $\Sigma\Pi\Sigma$ circuits .

While models (1) and (3) have received extensive attention in the literature, there have hardly been any lower bound results on model (2). It follows that obtaining a $2^{\omega(\sqrt{n}\log n)}$ lower bound for any one of the models above would give a similar lower bound to the other. However, known lower bounds for model (1) so far do not even imply a super polynomial lower bound for model (2) which leaves obtaining super polynomial lower bounds against the model wide open.

In this article, we prove lower bounds against two restrictions of model (2) mentioned above: $\Sigma\wedge\Sigma^{[\leq n]}\wedge^{[\geq 21]}\Sigma$ circuits and $\Sigma\wedge\Sigma^{[\leq 2^{\sqrt{n}/1000}]}\wedge^{[\geq\sqrt{n}]}\Sigma$ circuits with bottom gates computing homogeneous linear forms. Since the transformation from depth four $\Sigma\Pi^{[\sqrt{n}]}\Sigma\Pi^{[O(\sqrt{n})]}$ to depth five $\Sigma\wedge^{[a]}\Sigma\wedge^{[d/a]}\Sigma$ in [9] works against any chosen parameter $a<d$, the restrictions on the degree of the bottom $\wedge$ gates in the models we consider are general enough.

Throughout, it helps to interpret the polynomials computed by $\Sigma\wedge\Sigma\wedge\Sigma$ as sums of powers of projections of power symmetric polynomials where the n variate power symmetric polynomial of degree d is given by $p_d(x_1,\ldots,x_n)=x_1^d+\cdots+x_n^d$.

Our Results

We prove lower bounds against the restrictions of depth five $\Sigma\wedge\Sigma\wedge\Sigma$ circuits with powering gates mentioned above and homogeneous depth four $\Sigma\wedge\Sigma\wedge$ circuits.

We begin with the study of depth four $\Sigma\wedge\Sigma\wedge$ circuits. We show that any $\Sigma\wedge\Sigma\wedge$ circuit, where every Σ gate at the lower layer either computes a linear polynomial or a sum of powers of variables with degree at least two, requires $2^{\Omega(n)}$ size to compute the polynomial $x_1\cdots x_n$ (See Corollary 3.7). Though the result follows from a simple argument involving the dimension of projections of multilinear derivatives, it is surprising, given that allowing the lower layer of Σ gates to compute arbitrary sum of powers gives a $2^{O(\sqrt{n})}$ size circuit for $x_1\cdots x_n$ (see [19, Corollary 17.16]).

Extending our approach to depth five circuits, we show that any $\Sigma\wedge\Sigma^{[\leq n]}\wedge^{[\geq 21]}\Sigma$ circuit requires exponential size to compute the monomial $x_1\cdots x_n$:

▶ **Theorem 1.1.** *Let $g=\sum_{i=1}^{s}f_i^{\alpha_i}$ where $f_i=p_{d_i}(\ell_{i_1},\ldots,\ell_{i_n})+\beta_i$ and for every i, either $d_i=1$ or $d_i\geq 21$ and $\ell_{i_1},\ldots,\ell_{i_n}$ are homogeneous linear forms. If $g=x_1\cdot x_2\cdots x_n$ then $s=2^{\Omega(n)}$.*

▶ Remark. It may be noted that the arguments in [11] immediately imply exponential lower bounds for the monomial $x_1\cdots x_n$ against $\Sigma\wedge\Sigma\wedge^{[\leq 21]}\Sigma$ circuits. However, due to the limitations on the choice of parameters used in [11] and those in the proof of Theorem 1.1, the two results though seemingly complementary, do not imply a lower bound against general $\Sigma\wedge\Sigma\wedge\Sigma$ circuits.

The proof of Theorem 1.1 involves the dimension of the space of projected multilinear derivatives as a complexity measure for a polynomial f. It is computed by first projecting the partial derivative space of f to its multilinear subspace and then setting a subset of variables to 0. The dimension of the resulting space of polynomials is our measure of complexity for polynomials. Further, the method of projected multilinear derivatives also gives our second important result of the paper: An exponential lower bound against depth five powering circuits where the middle Σ layers have fan-in at most $2^{\sqrt{n}/1000}$ with the degree of the bottom $\wedge$ gates at least $\sqrt{n}$:

▶ **Theorem 1.2.** *Let* $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ *where* $f_i = p_{d_i}(\ell_{i_1}, \ldots, \ell_{i_{N_i}}) + \beta_i$, $\sqrt{n} \leq d_i \leq n$, $N_i \leq 2^{\sqrt{n}/1000}$, *and* $\ell_{i_1}, \ldots, \ell_{i_{N_i}}$ *are homogeneous linear forms. If* $g = x_1 \cdot x_2 \cdots x_n$ *then* $s = 2^{\Omega(n)}$.

It is not difficult to see that the polynomial $x_1 \cdots x_n$ has a homogeneous $\Sigma\wedge^{[\sqrt{n}]}\Sigma^{[O(2^{\sqrt{n}})]}\wedge^{[\sqrt{n}]}\Sigma$ circuit of size $2^{O(\sqrt{n})}$ (see Lemma 3.11). Theorem 1.2 shows that reducing the middle Σ fan-in by a constant factor in the exponent leads to an exponential lower bound, and it establishes the fan-in of the middle Σ gate as an important parameter.

As mentioned earlier, the restrictions on the degree of the bottom $\wedge$ gates in Theorems 1.1 and 1.2 are not much of a concern. However, the homogeneity condition on the lower Σ and $\wedge$ gates seems to be necessary in our proofs of Theorem 1.1 and Theorem 1.2. In fact, Saptharishi [19] in a result attributed to Forbes, showed that $x_1 \cdots x_n$ can be computed by $\Sigma\wedge\Sigma\wedge$ circuits of size $2^{O(\sqrt{n})}$ where the lower Σ gates are not necessarily homogeneous. A closer look at our technique reveals that, if the bottom Σ gates either compute a sum of powers of degree at least two except for a constant term or compute a linear form, then any such $\Sigma\wedge\Sigma\wedge$ circuit computing $x_1 \cdots x_n$ must have size $2^{\Omega(n)}$. (See Corollary 3.7). Moreover, there has to be at least $2^{\Omega(n)}$ powering gates at the bottom with degree 1.

Thus, it is important to study depth five powering circuits where the bottom Σ gates are not necessarily homogeneous. Towards this, in Section 4, we consider the widely used measure of the dimension of the shifted partial derivatives of a polynomial. We show:

▶ **Theorem 1.3.** *Let* $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ *where* $f_i = p_{d_i}(x_{i_1}, \ldots, x_{i_{m_i}}, \ell_{i_1}, \ldots, \ell_{i_{r_i}})$, $m_i \leq \frac{1}{40}n$, $r_i \leq n^{\epsilon}$, $d \leq 2^{o(n)}$, $\alpha_i \leq 2^{o(n)}$ *for all* i *where* $0 < \epsilon < 1$. *If* $g = x_1 x_2 \ldots x_n$ *then* $s = 2^{\Omega(n)}$.

It should be noted that Theorem 1.3 is much weaker than Theorems 1.1 and 1.2, however, it allows non-homogeneous Σ gates at the bottom. It seems that the restrictions on r_i in the above theorem are necessary if the lower bound argument uses the method of shifted partial derivatives. In particular, we show:

▶ **Lemma 1.4.** *Let* $k \leq \min\{l, d\}$ *and* $\alpha > 0$ *be large enough. Then*

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}\left(p_d(x_1, \ldots, x_n)^{\alpha}\right)\right\}\right) = \Omega\left(\frac{\binom{n}{k}\binom{n+l}{l}}{l^{l/(d-1)}}\right).$$

In the cases where $l/(d-1) = O(1)$ and $l = n^{O(1)}$ the above bound is tight up to a polynomial factor since $\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}\left(p_d(x_1, \ldots, x_n)^{\alpha}\right)\right\}\right) \leq \binom{n}{k}\binom{n+l}{l}$ and hence indicating that the restrictions on the r_is in Theorem 1.3 would be necessary if the dimension of shifted partial derivatives is used as the measure of complexity.

Finally, adapting the ideas developed in [4] to the case of projected multilinear derivatives, we obtain black-box deterministic quasi-polynomial time identity testing algorithms for the above mentioned restrictions of $\Sigma\wedge\Sigma\wedge\Sigma$ circuits (Corollary 5.1).

2 Preliminaries

An *arithmetic circuit* is a labelled directed acyclic graph. Vertices of zero in-degree are called *input* gates and are labelled by elements in $\mathbb{F} \cup \{x_1, \ldots, x_n\}$. Vertices of in-degree two or more are called *internal* gates and have their labels from $\{\times, +\}$. An arithmetic circuit has at least one vertex of zero out-degree called an *output* gate. We assume that an arithmetic circuit has exactly one output gate. A polynomial p_g in $\mathbb{F}[x_1, \ldots, x_n]$ can be associated with every gate g of an arithmetic circuit defined in an inductive fashion. Input gates compute their label. Let g be an internal gate with children $f_1, \ldots, f_m$ then $p_g = p_{f_1}$ op $\cdots$ op p_{f_m} where op $\in \{+, \times\}$ is the label of g. The polynomial computed by the circuit is the polynomial at one of the output gates and denoted by p_C. The size of an arithmetic circuit is the number of gates in it and is denoted by size(C). We will denote a fan-in/degree bound on a layer as a superscript to the corresponding gate e.g., $\Sigma \wedge \Sigma^{[\leq n]} \wedge^{[\geq 21]} \Sigma$ denotes the class of families of polynomials computed by depth five circuits with powering and sum gates, where the middle layer of sum gates have fan-in bounded from above by n and the bottom most powering gates have degree at least 21.

The following bound on the binomial coefficient is useful throughout the paper:

▶ **Proposition 2.1** ([16]). *Let $r \leq n$. Then $\log_2 \binom{n}{r} \approx nH(r/n)$, where H is the binary entropy function, $H(p) = -p\log_2(p) - (1-p)\log_2(1-p)$, and $\approx$ is equality up to an additive $o(n)$ error.*

We denote by $[n]$ the set $\{1, \ldots, n\}$. For a set of polynomials S, let $\mathcal{M}_{\leq d}(S)$ ($\mathcal{M}_{=d}(S)$) denote the set of all products of at most (exactly) d not necessarily distinct elements from S. Note that when S is a set of variables, $|\mathcal{M}_{\leq d}(S)| = \binom{|S|+d}{d}$. When the set S is clear from the context, we use $\mathcal{M}_{\leq d}$ ($\mathcal{M}_{=d}$) instead of $\mathcal{M}_{\leq d}(S)$ ($\mathcal{M}_{=d}(S)$).

For a subset S of variables, let $\mathcal{X}_a^b(S)$ denote the set of all multilinear monomials of degree $a \leq d \leq b$ in variables from the set S, i.e.,

$$\mathcal{X}_a^b(S) = \{\prod_{x_i \in S} x_i^{\delta_i} \mid a \leq \sum_{i=1}^{n} \delta_i \leq b, \delta_i \in \{0,1\}\}.$$

For two sets A and B, define $A \odot B \triangleq \{a \cdot b \mid a \in A, b \in B\}$. Additionally, we define $A \cdot f$ for some polynomial f to be the set $\{a \cdot f \mid a \in A\}$.

The notion of *shifted partial derivatives* is given as follows: For $k \geq 0$ and $f \in \mathbb{F}[x_1, \ldots, x_n]$ let $\partial^{=k} f$ denote the set of all partial derivatives of f of order k.

For $l \geq 0$, the (k, l) shifted partial derivative space of f, denoted by $\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k} f\right\}$, is defined as:

$$\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k} f\right\} \triangleq \mathbb{F}\text{-Span}\left\{\mathbf{m} \cdot \partial^{=k} f \ \mid\ \mathbf{m} \in \mathcal{M}_{\leq \ell}(x_1, \ldots, x_n)\right\}$$

where $\mathbb{F}\text{-Span}\left\{S\right\} \triangleq \{\alpha_1 f_1 + \cdots + \alpha_m f_m \mid f_i \in S \text{ and } \alpha_i \in \mathbb{F} \text{ for all } i \in [m]\}$. We need the following bound for the dimension of the space of shifted partial derivatives $x_1 \cdots x_n$:

▶ **Proposition 2.2** ([11]).

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial_{\mathsf{ML}}^{=k} x_1 \cdots x_n\right\}\right) = \dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k} x_1 \cdots x_n\right\}\right) \geq \binom{n}{k} \cdot \binom{n-k+l}{l}.$$

In the above, $\partial_{\mathsf{ML}}^{=k} f$ denotes the set of kth order multilinear derivative space of f, i.e., $\partial_{\mathsf{ML}}^{=k} f \triangleq \{\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} \mid i_1 < \ldots < i_k \in \{1, \ldots, n\}\}$.

3 The proof of Theorems 1.1 and 1.2

This section is devoted to the proof of Theorems 1.1 and 1.2. Our proof follows the standard two step approach for proving arithmetic circuit lower bounds: First, define a sub-additive measure that is low for every polynomial computed in the model. Second, show that the measure is exponentially larger for a specific polynomial p. Hence allowing us to conclude that any circuit in the model that computes p requires exponential size.

We consider a variant of the space of partial derivatives, viz., the *projected multilinear derivatives* as the complexity measure for polynomials.

The Complexity Measure

Let $f \in \mathbb{F}[x_1,\ldots,x_n]$. For $S \in \{1,\ldots,n\}$, let $\pi_S : \mathbb{F}[x_1,\ldots,x_n] \to \mathbb{F}[x_1,\ldots,x_n]$ be the projection map that sets all variables in S to zero, i.e., for every $f \in \mathbb{F}[x_1,\ldots,x_n]$, $\pi_S(f) = f(x_i = 0 \mid i \in S)$. Let $\pi_{\mathsf{m}}(f)$ denote the projection of f onto its multilinear monomials, i.e., if $f = \sum_{\alpha\in\mathbb{N}^n} c_\alpha \prod_{i=1}^n x_i^{\alpha_i}$ then $\pi_{\mathsf{m}}(f) = \sum_{\alpha\in\{0,1\}^n} c_\alpha \prod_{i=1}^n x_i^{\alpha_i}$. For $S \subseteq \{1,\ldots,n\}$ and $0 < k \le n$, the dimension of Projected Multilinear Derivatives (PMD) of a polynomial f is defined as:

$$\mathrm{PMD}_S^k(f) \triangleq \dim(\mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k} f))\right\}).$$

We omit the subscript S when either S is clear from the context or when it refers to an unspecified set S. It can be seen that PMD_S^k is sub-additive. A proof is given in the appendix.

▶ **Lemma 3.1.** *For any $S \subseteq \{1\ldots,n\}$, $k \ge 1$, and polynomials f and g:* $\mathrm{PMD}_S^k(f+g) \le \mathrm{PMD}_S^k(f) + \mathrm{PMD}_S^k(g)$.

A Lower Bound for the Measure

We establish a lower bound on the dimension of projected multilinear derivatives of the polynomial $x_1\cdots x_n$.

▶ **Lemma 3.2.** *For any $S \subseteq \{1,\ldots,n\}$ with $|S| = n/2+1$ and $k = 3n/4$ we have:*

$$\mathrm{PMD}_S^k(x_1\cdots x_n) \ge \binom{n/2-1}{n/4} \ge 2^{n/2}/n^2.$$

Proof. Let $T \subseteq \{1,\ldots,n\}$ with $|T| = k$. Then $\frac{\partial^k}{\partial T}(x_1\cdots x_n) = \prod_{i\notin T} x_i$. Note that if $S\cap\overline{T} = \emptyset$ then we have $\pi_S(\pi_{\mathsf{m}}(\frac{\partial^k}{\partial T}(x_1\cdots x_n))) = \prod_{i\notin T} x_i$ since setting variables in S to zero does not affect the variables in $\overline{T}$. Otherwise, if $S\cap\overline{T} \ne \emptyset$ then $\pi_S(\pi_{\mathsf{m}}(\frac{\partial^k}{\partial T}(x_1\cdots x_n))) = 0$. Thus, we have:

$$\mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k}(x_1\cdots x_n)))\right\} \supseteq \mathbb{F}\text{-Span}\left\{\prod_{i\in T} x_i \;\mid\; T\subseteq\overline{S}, |T| \le n/4\right\}.$$

Hence, $\mathrm{PMD}_S^k(x_1\cdots x_n) \ge \binom{n/2-1}{n/4} \ge 2^{n/2}/n^2$ using Stirling's approximation. ◀

$\Sigma\wedge\Sigma\wedge$ Circuits: The Curse of Homogeneity

In this subsection, we show that the dimension of projected multilinear derivatives of powers of power symmetric polynomials is low. To begin with, we observe that in the case of powers of power symmetric polynomials, the dimension of projected multilinear derivatives is either 0 or 1 for suitable values of k. The proof of the following lemma can be found in the appendix.

▶ **Lemma 3.3.** *For any $\beta_0,\beta_1,\ldots,\beta_n \in \mathbb{F}$, $\alpha, d \in \mathbb{N}$ and for any $S \subseteq \{1,\ldots,n\}$ with $|S|+k > n$, we have* $\mathrm{PMD}_S^k((\sum_{i=1}^n \beta_i x_i^d + \beta_0)^\alpha) \le 1$.

This immediately leads us to the following lower bound against $\Sigma\wedge\Sigma\wedge$ circuits:

▶ **Corollary 3.4.** *Let $f = f_1^{\alpha_1} + \cdots + f_s^{\alpha_s}$ where $f_i = \sum_{j=1}^n \beta_{ij} x_j^{d_i} + \beta_{i0}$, $\beta_{ij} \in \mathbb{F}$. If $f = x_1\cdots x_n$ then $s = 2^{\Omega(n)}$.*

Proof. Let $S \subseteq \{1,\ldots,n\}$ with $|S| = n/2+1$ and $k = 3n/4$. From Lemmas 3.3 and 3.1 we have $\mathrm{PMD}_S^k(f) \le \sum_{i=1}^s \mathrm{PMD}_S^k(f_i) \le s$. Hence by Lemma 3.2, $s \ge 2^{n/2}/n^2$ as required. ◀

The homogeneity condition for the bottom power gates is necessary due to the following result in [19]. Let $\mathrm{Sym}_{n,d} = \sum_{S\subseteq[n],|S|=d}\prod_{i\in S} x_i$, the elementary symmetric polynomial of degree d.

▶ **Proposition 3.5.** *[19, Corollary 17.16] For any $d > 0$, $\mathrm{Sym}_{n,d}$ can be computed by a $\Sigma\wedge\Sigma\wedge$ circuit of size $2^{O(\sqrt{n})}$.*[1]

Is it all about homogeneity at the bottom Σ gates? The answer is no. In fact, Lemma 3.3 can be generalized to the case of powers of polynomials in the span of $\{x_{i_j}^{\alpha_{i_j}} \mid 1 \le i_j \le n,\ \alpha_{i_j} \ge 2\}$, a proof is deferred to the Appendix.

▶ **Lemma 3.6.** *For any $\beta_0, \beta_1, \ldots, \beta_r \in \mathbb{F}$, $\alpha, d \in \mathbb{N}$ and for any $S \subseteq \{1,\ldots,n\}$ with $|S| + k > n$, we have $\mathrm{PMD}_S^k((\sum_{j=1}^r \beta_j x_{i_j}^{d_j} + \beta_0)^\alpha) \le 1$ where $1 \le i_j \le n$ and either $\forall j\ d_j \ge 2$ or $\forall j\ d_j = 1$.*

We get the following generalization of Corollary 3.4, a proof can be found in the Appendix.

▶ **Corollary 3.7.** *Let $f = f_1^{\alpha_1} + \cdots + f_s^{\alpha_s}$ where for every i, either f_i is a linear form or $f_i = \sum_{j=1}^n \beta_{i,l_j} x_{l_j}^{d_{i_j}} + \beta_{i0}$ for $d_{i_j} \ge 2$ and $\beta_{i,l_j} \in \mathbb{F}$. If $f = x_1 \cdots x_n$ then $s = 2^{\Omega(n)}$. Moreover, $|\{i \mid f_i \text{ is linear}\}| = 2^{\Omega(n)}$.*

$\Sigma\wedge\Sigma\wedge\Sigma$ Circuits: Middle Σ Fan-in versus the Bottom Degree

It might seem that the argument in Lemma 3.3 would immediately generalize to the case when the variables are replaced with homogeneous linear forms. However, the argument above fails even when the degree of the power symmetric polynomial is two (i.e., $d = 2$). Let $f = \ell_1^2 + \cdots + \ell_n^2 + \beta$ where $\ell_1, \ldots, \ell_n$ are homogeneous linear functions such that each of the ℓ_i have all n variables with non-zero coefficients and $\beta \ne 0$. It is not hard to see that the space $\partial_{\mathsf{ML}}^k f$ of the kth order derivatives of f^α is contained in the span of $\{f^{\alpha-k}\prod_{i=1}^n \ell_i^{\gamma_i} \mid \sum_i \gamma_i \le k\}$. Even after applying the projections π_{m} and π_S for any $S \subseteq \{1,\ldots,n\}$, with $|S| = (n/2)+1$, obtaining a bound on PMD_S^k better than the lower bound in Lemma 3.2 seems to be difficult. For, every multilinear monomial of degree $|n/2 - 1 - k|$ appears in at least one of the projected multilinear derivatives of f^α.

A natural approach to overcome the above difficulty could be to obtain a basis for the projected multilinear derivatives of f^α consisting of a small set of monomials and a small set of products of powers of the linear forms multiplied by suitable powers of f, similar to the proof of Lemma 3.3. Surprisingly, as shown below in Lemma 3.8, the approach works when the degree $d \ge 21$, although it requires an involved combinatorial argument.

▶ **Lemma 3.8.** *Suppose that $f = (\ell_1^d + \ldots + \ell_n^d + \beta)$ for some scalar β, and ℓ_j homogeneous linear forms, $1 \le j \le n$. Let $Y = \{\ell_i^{d-j} \mid 1 \le i \le n, 1 \le j \le d\}$ and $\lambda = 1/4 + \varepsilon$ for some $0 < \varepsilon < 1/4$. Then, for $k = 3n/4$ and any $S \subseteq \{1,\ldots,n\}$ with $|S| = n/2+1$, we have:*

$$\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k} f^\alpha) \subseteq \mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\mathcal{F}\odot\left(\mathcal{X}_{\lambda n}^{n/2-1}(\overline{S}) \cup \mathcal{M}_{\le(1+\varepsilon)n/d}(Y)\right)))\right\}$$

where $\mathcal{F} = \cup_{i=1}^k f^{\alpha-i}$ and $\overline{S} = \{1,\ldots,n\}\setminus S$.

[1] In [19], Corollary 17.16, it is mentioned that the resulting $\Sigma\wedge\Sigma\wedge$ circuit is homogeneous. However, a closer look at the construction shows that that the application of Fischer's identity produces sum gates that are not homogeneous.

Proof. Let $T \subseteq \{x_1, \ldots, x_n\}$ with $|T| = k$, let $f_T^{(k)}$ denote kth order partial derivative of f with respect to T. Note that $f_T^{(k)} \in \mathbb{F}\text{-Span}\left\{\ell_j^{d-k} \mid 1 \le j \le n\right\}$. Let L_i denote $\{\ell_j^{d-i} \mid 1 \le j \le n\}$ so that $f_T^{(k)} \in \mathbb{F}\text{-Span}\{L_k\}$. Then

$$\frac{\partial^k f^\alpha}{\partial T} \in \mathbb{F}\text{-Span}\left\{f^{\alpha-i} \odot D_i^T(f) \mid 1 \le i \le k\right\} \tag{1}$$

where $D_i^T(f) = \left\{\prod_{r=1}^{i} f_{T_r}^{(t_r)} \;\mid\; T_1 \uplus \ldots \uplus T_i = T, t_r = |T_r| > 0 \text{ and } 1 \le r \le i\right\}$. Intuitively, the set D_i^T contains one polynomial for each possible partition of T into i many parts. The polynomial corresponding to a particular partition is the product of the derivatives of f with respect to each of the parts. Now, the following claim bounds the span of D_i^T:

▶ **Claim 1.** *For any* $1 \le i \le k$, $D_i^T \subseteq \mathbb{F}\text{-Span}\left\{\bigodot_{r=1}^{k} L_r^{\odot j_r} \;\mid\; 1 \cdot j_1 + \cdots + k \cdot j_k = k\right\}$.

Proof. Let $T_1 \uplus \cdots \uplus T_i = T$ be a partition and let j_r denote the number of parts with cardinality r, i.e., $j_r = |\{j \;\mid\; |T_j| = r\}|$. Then

$$\prod_{|T_j|=r} f_{T_j}^{(r)} \in \mathbb{F}\text{-Span}\left\{\bigodot_{|T_j|=r} L_r\right\} = \mathbb{F}\text{-Span}\left\{L_r^{\odot j_r}\right\}.$$

Thus, $\prod_{r=1}^{i} f_{T_r}^{(t_r)} \in \mathbb{F}\text{-Span}\left\{\bigodot_{r=1}^{k} L_r^{\odot j_r}\right\}$. Note that for any partition $T_1 \uplus \cdots \uplus T_i$ of T, $\sum_{r=1}^{k} r \cdot j_r = k$. The claim follows. ◀

Continuing from (1), we have:

$$\begin{aligned}\frac{\partial^k f^\alpha}{\partial T} &\in \mathbb{F}\text{-Span}\left\{f^{\alpha-i} \odot D_i^T(f) \;\mid\; 1 \le i \le k\right\} \subseteq \mathbb{F}\text{-Span}\left\{\mathcal{F} \odot \{D_i^T(f) \;\mid\; 1 \le i \le d\}\right\} \\ &\subseteq \mathbb{F}\text{-Span}\left\{\mathcal{F} \odot \left\{\bigodot_{r=1}^{d} L_r^{\odot j_r} \;\mid\; 1 \cdot j_1 + \cdots + d \cdot j_d = k\right\}\right\}. \end{aligned} \tag{2}$$

It remains to show that the right side of (2) is spanned by a set of polynomials that satisfy the properties stated in the lemma. Claim 2 completes the proof of Lemma 3.8.

▶ **Claim 2.**

$$\pi_S(\pi_{\mathsf{m}}\left(\left\{\bigodot_{r=1}^{d} L_r^{\odot j_r} \;\mid\; 1 \cdot j_1 + \cdots + d \cdot j_d = k\right\}\right) \subseteq \mathbb{F}\text{-Span}\left\{\mathcal{X}_{\lambda n}^{n/2-1}(\bar{S}) \cup \mathcal{M}_{\le(1+\varepsilon)n/d}(Y)\right\}.$$

Proof. Note that the polynomials in L_j are homogeneous non-constant polynomials of degree $d - j$, and hence the set $\bigodot_{r=1}^{d} L_r^{\odot j_r}$ consists of homogeneous polynomials of degree $\sum_{r=1}^{d} j_r(d-r)$.

Let $\deg(\bigodot_{r=1}^{d} L_r^{\odot j_r})$ denote the degree of polynomials in the set $\bigodot_{r=1}^{d} L_r^{\odot j_r}$. The remaining argument is split into three cases depending on the value of $\deg(\bigodot_{r=1}^{d} L_r^{\odot j_r})$.

Case 1: $\deg(\bigodot_{r=1}^{d} L_r^{\odot j_r}) \ge n/2$ then $\pi_S(\pi_{\mathsf{m}}(\bigodot_{r=1}^{d} L_r^{\odot j_r})) = \{0\}$. Note that here we have crucially used the fact that the ℓ_j are homogeneous.

Case 2: $\lambda n \le \deg(\bigodot_{r=1}^{d} L_r^{\odot j_r})) < n/2$. In this case $\pi_S(\pi_{\mathsf{m}}(\bigodot_{r=1}^{d} L_r^{\odot j_r}))$ is spanned by the set of all multilinear monomials in the set of variables $\{x_j \;\mid\; j \notin S\}$ of degree at least λn and at most $n/2 - 1$. Therefore we have, $\pi_S(\pi_{\mathsf{m}}(\bigodot_{r=1}^{d} L_r^{\odot j_r})) \subseteq \mathbb{F}\text{-Span}\left\{\mathcal{X}_{\lambda n}^{n/2-1}(\bar{S})\right\}$.

Case 3: $\deg(\bigodot_{r=1}^{d} L_r^{\odot j_r})) < \lambda n$. As $\deg(\bigodot_{r=1}^{d} L_r^{\odot j_r})) = \sum_{r=1}^{d} j_r(d-r) < \lambda n$, we have:

$$\begin{aligned}\sum_{r=1}^{d} j_r \cdot d &\leq \sum_{r=1}^{d} j_r \cdot r + \lambda = k + \lambda n \qquad (\text{since } \sum_{r=1}^{d} r \cdot j_r = k.)\\ &= (\lambda + 3/4)n = (1+\varepsilon)n.\end{aligned}$$

Hence, $\pi_S(\pi_{\mathsf{m}}(\bigodot_{r=1}^{d} L_r^{\odot j_r}))$ is spanned by the set of all product of at most $(1+\varepsilon)n/d$ polynomials of the form ℓ_i^{d-j}, i.e., $\pi_S(\pi_{\mathsf{m}}(\bigodot_{r=1}^{d} L_r^{\odot j_r})) \subseteq \mathbb{F}\text{-Span}\left\{\mathcal{M}_{\leq(1+\varepsilon)n/d}(Y)\right\}$. ◀

◀

Using Lemma 3.8 above and choosing suitable parameters k and S we obtain the following upper bound on the dimension of projected multilinear derivatives:

▶ **Theorem 3.9.** *Let $f = (\ell_1^d + \ldots + \ell_n^d + \beta)$ where ℓ_j are homogeneous linear forms. For $d \geq 21$ and any $S \subseteq \{1,\ldots,n\}$ where $|S| = n/2+1$. Then*

$$\mathrm{PMD}_S^k(f^\alpha) \leq 2^{(0.498+o(1))n}.$$

Proof. By Lemma 3.8,

$$\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k} f^\alpha)) \subseteq \mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\{f^{\alpha-i}\}_{i=1}^{k} \odot \left\{\mathcal{X}_{\lambda n}^{n/2-1}(\bar{S}) \cup \mathcal{M}_{\leq(1+\epsilon)n/d}(Y)\right\}))\right\}.$$

Recall that $\lambda = \frac{1}{4} + \varepsilon$. We choose $\varepsilon = 1/50$ and hence $\lambda = 0.27$. We have:

$$\mathrm{PMD}_S^k(f^\alpha) \leq k \cdot (|\mathcal{X}_{\lambda n}^{n/2-1}(\bar{S})| + |\mathcal{M}_{\leq(1+\epsilon)n/d}(Y)|).$$

Now, since $1/4 < \lambda < 1/2$, we have

$$\begin{aligned}|\mathcal{X}_{\lambda n}^{n/2-1}(\bar{S})| &\leq (n/2 - 1 - \lambda n) \cdot \binom{n/2-1}{\lambda n} \leq c(n/2) \cdot \binom{n/2}{\lambda n}\\ &\leq (cn/2) \cdot 2^{\frac{n}{2}\cdot\mathcal{H}(2\lambda)} \leq (cn/2) \cdot 2^{0.498n}.\end{aligned}$$

Where c is an absolute constant. We bound $|\mathcal{M}_{\leq(1+\epsilon)n/d}(Y)|$ as follows:

$$\begin{aligned}|\mathcal{M}_{\leq(1+\epsilon)n/d}(Y)| &= \binom{|Y| + (1+\varepsilon)n/d}{(1+\varepsilon)n/d} = \binom{dn + (1+\varepsilon)n/d}{(1+\varepsilon)n/d} \leq 2^{(dn+(1+\varepsilon)n/d)\mathcal{H}\left(\frac{(1+\varepsilon)n/d}{dn+(1+\varepsilon)n/d}\right)}\\ &= 2^{n(d+(1+\varepsilon)/d)\mathcal{H}\left((1+\varepsilon)/(d^2+(1+\epsilon))\right)} \leq 2^{0.4955n} \qquad \text{for } d \geq 21.\end{aligned}$$

For the last inequality, note that for fixed n and ε, $(d+(1+\varepsilon)/d)\mathcal{H}((1+\varepsilon)/(d^2+(1+\epsilon))$ is a monotonically decreasing function of d, with $\lim_{d\to\infty}(d+(1+\varepsilon)/d)\mathcal{H}((1+\varepsilon)/(d^2+(1+\epsilon)) = 0$. Therefore, the bound holds for $d \geq 21$. This completes the proof. ◀

▶ **Corollary 3.10.** *Let $f = (\ell_1^d + \ldots + \ell_N^d + \beta)$ where ℓ_j are homogeneous linear forms. If d is such that $N \leq 2^{(d/1000)}$, $d \leq n$, and $n/d = o(n)$ then for any $\alpha > 0$,*

$$\mathrm{PMD}_S^k(f^\alpha) \leq 2^{(0.498+o(1))n}.$$

Proof. Proof is exactly same as that of Theorem 3.9 except for the bound on $|\mathcal{M}_{\leq(1+\epsilon)n/d}(Y)|$. As in the proof of Theorem 3.9 set $\varepsilon = 1/50$ and $\lambda = 0.27$. By Lemma 3.8,

$$\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k} f^\alpha)) \subseteq \mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\{f^{\alpha-i}\}_{i=1}^{k} \odot \left\{\mathcal{X}_{\lambda n}^{n/2-1}(\bar{S}) \cup \mathcal{M}_{\leq(1+\epsilon)n/d}(Y)\right\}))\right\}$$

where $Y = \{\ell_i^{d-j} \mid 1 \leq i \leq N, 1 \leq j \leq k\}$. By the arguments in the proof of Theorem 3.9 we have $|\mathcal{X}_{\lambda n}^{n/2-1}(\bar{S})| \leq 2^{(0.498+o(1))n}$. Finally,

$$|\mathcal{M}_{\leq (1+\epsilon)n/d}(Y)| = \binom{dN + (1+\varepsilon)n/d}{(1+\varepsilon)n/d} \leq \left(\frac{(dN + (1+\varepsilon)n/d)e}{(1+\varepsilon)n/d}\right)^{(1+\varepsilon)n/d}$$
$$= ((d^2N/n + 1)e)^{(1+\varepsilon)n/d}.$$

Now, if $d \leq n$, $N \leq 2^{d/1000}$ and $n/d = o(n)$ then we have $|\mathcal{M}_{\leq(1+\epsilon)n/d}(Y)| \leq 2^{.002n+o(n)}$ for large enough n and hence $\mathrm{PMD}_S^k(f^\alpha) \leq 2^{(0.498+o(1))n}$. ◀

Proof of Theorem 1.1

▶ **Theorem 1.1.** *Let $g = \sum_{i=1}^s f_i^{\alpha_i}$ where $f_i = p_{d_i}(\ell_{i_1}, \ldots, \ell_{i_n}) + \beta_i$, either $d_i = 1$ or $d_i \geq 21$ and $\ell_{i_1}, \ldots, \ell_{i_n}$ are homogeneous linear forms for every i. If $g = x_1 \cdots x_n$ then $s = 2^{\Omega(n)}$.*

Proof. Let $S = \{1, \ldots, n/2+1\}$ and $k = 3n/4$. Then by Theorem 3.9 we have $\mathsf{PMD}_S^k(f_i) \leq 2^{0.498n+o(n)}$. By the sub-additivity of PMD_S^k (Lemma 3.1), we have $\mathsf{PMD}_S^k(\sum_{i=1}^s f_i^{\alpha_i}) \leq s \cdot 2^{0.498n+o(n)}$. Since $\mathsf{PMD}_S^k(x_1 \cdots x_n) \geq 2^{n/2}/n^2$, we conclude $s \geq 2^{0.001n}$, as required. ◀

Proof of Theorem 1.2

▶ **Theorem 1.2.** *Let $g = \sum_{i=1}^s f_i^{\alpha_i}$ where $f_i = p_{d_i}(\ell_{i_1}, \ldots, \ell_{i_{N_i}}) + \beta_i$, $\sqrt{n} \leq d_i \leq n$, $N_i \leq 2^{\sqrt{n}/1000}$, and $\ell_{i_1}, \ldots, \ell_{i_N}$ are homogeneous linear forms. If $g = x_1 \cdot x_2 \cdots x_n$ then $s = 2^{\Omega(n)}$.*

Proof. Let $S = \{1, \ldots, n/2+1\}$ and $k = 3n/4$. Since $d_i \geq \sqrt{n}$ so that $N_i \leq 2^{d/1000}$ then, by Corollary 3.10, we have $\mathsf{PMD}_S^k(f_i) \leq 2^{0.498n+o(n)}$. By the sub-additivity of PMD_S^k (Lemma 3.1), we have $\mathsf{PMD}_S^k(\sum_{i=1}^s f_i^{\alpha_i}) \leq s \cdot 2^{0.498n+o(n)}$. Since $\mathsf{PMD}_S^k(x_1 \cdots x_n) \geq 2^{n/2}/n^2$, we conclude $s \geq 2^{0.001n}$ for large enough n, as required. ◀

A Separation within $\Sigma\wedge\Sigma\wedge\Sigma$ Circuits

An alert reader might have wondered if the restriction on the fan-in of the middle layer of Σ gates in Theorem 1.2 is a limitation of the method of projected multilinear derivatives. By a simple application of Fischer's identity [3], we get:

▶ **Lemma 3.11.** *Over fields of unbounded size, the polynomial $x_1 \cdots x_n$ can be computed by a homogeneous $\Sigma\wedge^{[\sqrt{n}]}\Sigma^{[O(2^{\sqrt{n}})]}\wedge^{[\sqrt{n}]}\Sigma$ circuit of size $2^{O(\sqrt{n})}$.*

This immediately leads to the following separation of homogeneous $\Sigma\wedge^{[\sqrt{n}]}\Sigma\wedge^{[\sqrt{n}]}\Sigma$ circuits:

▶ **Corollary 3.12.** *The class of polynomials computed by $\Sigma\wedge^{[\sqrt{n}]}\Sigma^{[2^{\sqrt{n}/1000}]}\wedge^{[\sqrt{n}]}\Sigma$ of size $2^{O(\sqrt{n})}$ is strictly contained in the class computed by $\Sigma\wedge^{[\sqrt{n}]}\Sigma^{[2^{\sqrt{n}}]}\wedge^{[\sqrt{n}]}\Sigma$ of size $2^{O(\sqrt{n})}$.*

4 Dimension of Shifted Partial Derivatives

This section is devoted to the study of shifted partial derivatives of polynomials that are computed by restricted $\Sigma\wedge\Sigma\wedge\Sigma$ circuits and proofs of Theorem 1.3 and Lemma 1.4.

Proof of Theorem 1.3

We begin with a simple upper bound on the dimension of the derivatives of powers of projections of p_d onto low-dimensional sub-spaces:

▶ **Lemma 4.1.** *Let $f = p_d(\ell_1, \ldots, \ell_t)$ where $\ell_1, \ldots, \ell_t$ are linear forms. Then for any $k > 0$, we have* $\dim\left(\mathbb{F}\text{-Span}\left\{\partial^{\leq k}_{\mathsf{ML}} f^\alpha\right\}\right) \leq (k+1)(dk)^r$ *where* $r = \dim(\mathbb{F}\text{-Span}\{\ell_1, \ldots, \ell_t\})$.

Proof. Without loss of generality, assume that $\ell_1, \ldots, \ell_r$ is a basis for $\mathbb{F}\text{-Span}\{\ell_1, \ldots, \ell_t\}$, $r \leq t$. Observe that

$$\partial^{\leq k}_{\mathsf{ML}} f^\alpha \subseteq \mathbb{F}\text{-Span}\left\{ f^{\alpha - i} \cdot \ell_1^{\beta_1} \cdots \ell_r^{\beta_r} \mid \sum_{j=1}^{r} \beta_j \leq dk \right\}_{i \in \{1, \ldots, k\}}$$

and therefore, $\dim\left(\mathbb{F}\text{-Span}\left\{\partial^{\leq k}_{\mathsf{ML}} f^\alpha\right\}\right) \leq (k+1)(dk)^r$ as required. ◀

Now, we bound the dimension of shifted partial derivatives of powers of the power symmetric polynomial. proof of Lemma 4.2 can be found in the Appendix.

▶ **Lemma 4.2.** *Let $f = p_d(x_{j_1}, \ldots, x_{j_m})$ for some $j_1, \ldots, j_m \in \{1, \ldots, n\}$. Then for any $\alpha, l, k \geq 1$*

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l} \partial^{=k}_{\mathsf{ML}} f^\alpha\right\}\right) \leq (k+1)\binom{n+m+k+l}{k+l}.$$

Note that when m is small (say $m \leq n/40$), the bound shown above is better than the straightforward bound $\binom{m}{k}\binom{n}{l}$ for suitable values of k and l. Combining Lemmas 4.1 and 4.2 with the sum and product rules for partial derivatives, we get:

▶ **Lemma 4.3.** *Let $\ell_1, \ldots \ell_t$ be linear forms in $\mathbb{F}[x_1, \ldots, x_n]$ with* $\dim(\mathbb{F}\text{-Span}\{\ell_1, \ldots, \ell_t\}) =$ *r and $f = p_d(x_{j_1}, \ldots, x_{j_m}, \ell_1, \ldots, \ell_t)$. Then for any $d > k > 0$, we have*

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l} \partial^{=k}_{\mathsf{ML}} f^\alpha\right\}\right) \leq (\alpha+1)(k+1)^3(dk)^r\binom{m+n+k+l}{k+l}.$$

A proof of Lemma 4.3 can be found in the Appendix. Finally, using sub-additivity of shifted partial derivatives and Lemma 4.3 we get:

▶ **Theorem 4.4.** *Let $d > k > 0$ and $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ where $f_i = p_{d_i}(x_{i_1}, \ldots, x_{i_{m_i}}, \ell_{i_1}, \ldots, \ell_{i_{r_i}})$ and $\ell_{i_1}, \ldots, \ell_{i_{m_i}}$ are linear forms in $x_1, \ldots, x_n$. Then for any $l > 0$ with $k + l > n + m$:*

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l} \partial^{=k}_{\mathsf{ML}} g\right\}\right) \leq s(\alpha+1)(k+1)^3(dk)^r\binom{n+m+k+l}{k+l}$$

where $m = \max_i m_i$ and $r = \max_i\{\dim(\mathbb{F}\text{-Span}\{\ell_{i_1}, \ldots, \ell_{i_{r_i}}\})\}$.

Combining Theorem 4.4 with Proposition 2.2 complete the proof of Theorem 1.3:

▶ **Theorem 1.3.** *Let $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ where $f_i = p_{d_i}(x_{i_1}, \ldots, x_{i_{m_i}}, \ell_{i_1}, \ldots, \ell_{i_{r_i}})$, $m_i \leq \frac{1}{40}n$, $r_i \leq n^\epsilon$, $d \leq 2^{n^{1-\gamma}}$ and $\alpha_i \leq 2^{n^\delta}$ for all i, for some $0 < \delta, \epsilon, \gamma < 1$. If $g = x_1 x_2 \ldots x_n$ then $s = 2^{\Omega(n)}$.*

Proof. Let $d \geq 2$ and $m = \max_i m_i$. By Proposition 2.2 and Theorem 4.4:

$$s \geq \frac{\binom{n}{k}\binom{n-k+l}{l}}{(\alpha+1)(k+1)^3(dk)^r\binom{n+m+k+l}{k+l}} \quad \text{where } \alpha = \max_i \alpha_i.$$

Taking the logarithm and using that $3\log(k+1) \leq 3\log dk$ since $d \geq 2$ gives us

$$\log s \geq \log\binom{n}{k} + \log\binom{n-k+l}{l} - \left(\log(\alpha+1) + \log\binom{n+m+k+l}{k+l} + (r+3)\log dk\right).$$

Note that $(r+3)\log dk \in o(n)$ if $d \leq 2^{n^{1-\gamma}}$. Now, using Proposition 2.1 and setting $k = n/10$ and $l = 10n$ we get $\log s \geq 0.0165n$. This proves the required bound. ◀

Proof of Lemma 1.4

▶ **Lemma 1.4.** *For $k \leq \min\{l, d\}$ and $\alpha > 0$ be large enough. Then*

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}\left(p_d(x_1,\ldots,x_n)^\alpha\right)\right\}\right) = \Omega\left(\frac{\binom{n}{k}\binom{n+l}{l}}{l^{l/(d-1)}}\right).$$

Proof. We have $\forall S \subseteq [n]$ with $|S| = k$, $\frac{\partial^k f^\alpha}{\partial S} = \alpha(\alpha-1)\ldots(\alpha-k+1)d^k f^{\alpha-k} \cdot \prod_{i\in S} x_i^{d-1}$ where $f = p_d(x_1,\ldots,x_n)$. Now,

$$\begin{aligned}
\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}f\right\} &= \mathbb{F}\text{-Span}\left\{\mathbf{m}\cdot\partial^{=k}f \;\middle|\; \mathbf{m} \in \mathcal{M}_{\leq l}(x_1,\ldots,x_n)\right\} \\
&\supseteq \mathbb{F}\text{-Span}\left\{\mathbf{m}\cdot\partial^{=k}_{\mathsf{ML}}f \;\middle|\; \mathbf{m} \in \mathcal{M}_{\leq l}(x_1,\ldots,x_n)\right\} \\
&= \mathbb{F}\text{-Span}\left\{\mathbf{m}\cdot f^{\alpha-k}\cdot\prod_{i\in S} x_i^{d-1} \;\middle|\; \begin{array}{c}\mathbf{m} \in \mathcal{M}_{\leq l}(x_1,\ldots,x_n),\\ S \subseteq \{1,\ldots,n\} \text{ with } |S| = k\end{array}\right\}.
\end{aligned}$$

Thus $\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}\left(f^\alpha\right)\right\}\right)$ is at least the number of monomials in the set

$$\Gamma_{k,l,d} \triangleq \left\{\prod_{i\in S} x_i^{d-1} \mid S \subseteq \{1,\ldots,n\}, |S| = k\right\} \odot \mathcal{M}_{\leq l}(\{x_1,\ldots,x_n\}).$$

We now lower bound $|\Gamma_{k,l,d}|$. Let $\mathcal{N}_{k,d} = \{\prod_{i\in S} x_i^{d-1} \mid |S| = k\}$. Consider the map $\varphi : \mathcal{N}_{k,d} \times \mathcal{M}_{\leq l} \to \Gamma_{k,l,d}$ where $(m_1, m_2) \mapsto m_1 \cdot m_2$. Note that if $d-1 \geq l$, the map φ is injective and hence $|\Gamma_{k,l,d}| \geq |\mathcal{N}_{k,d}||\mathcal{M}_{\leq l}| = \binom{n}{k}\cdot\binom{n+l}{l}$. When $d-1 < l$, it is enough to argue that the number of pre-images under φ for any element gamma $\Gamma_{k,l,d}$ has at most $l^{l/(d-1)}$.

Let $\gamma = x_{i_1}^{d-1}\cdots x_{i_k}^{d-1}\cdot m \in \Gamma_{k,l,d}$ where $i_1 < i_2 < \cdots < i_k$ and $m \in \mathcal{M}_{\leq l}$. To bound the number of pre-images of γ, let $\varphi(x_{j_1}\cdots x_{j_k}, m') = \gamma$. Since degree of m' is at most l, by comparing the degrees, it must hold that $|\{i_1,\ldots i_k\}\triangle\{j_1,\ldots,j_k\}| \leq \max\{l,k\}/(d-1)$. Therefore $|\varphi^{-1}(x_{i_1}^{d-1}\cdots x_{i_k}^{d-1}\cdot m)| \leq 1 + \binom{\max\{l,k\}}{l/(d-1)} \leq l^{l/(d-1)}$. ◀

5 Black Box Polynomial Identity Testing

Forbes [4] showed that lower bound for $x_1\cdots x_n$ against any model using the method of shifted partial derivatives can be translated into quasi polynomial time black-box PIT algorithm. Using the ideas from [4], we obtain deterministic quasi-polynomial time identity testing algorithm from the lower bounds obtained in Sections 3 and 4.

Let $\mathcal{C}$ be union of the class of polynomials computed by circuits in Theorems 1.1, 1.2 and 1.3. Then

▶ **Corollary 5.1.** *There is a deterministic $n^{O(\log s)}$ time algorithm that given a multilinear polynomial $g \in \mathcal{C}$ tests if $g \equiv 0$.*

References

1 Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *FOCS*, pages 67–75, 2008.

2 Walter Baur and Volker Strassen. The complexity of partial derivatives. *Theor. Comput. Sci.*, 22:317–330, 1983. URL: http://dx.doi.org/10.1016/0304-3975(83)90110-X, doi:10.1016/0304-3975(83)90110-X.

3 Ismor Fischer. Sums of like powers of multivariate linear forms. *Mathematics Magazine*, 67(1):59–61, 1994. URL: http://www.jstor.org/stable/2690560.

4 Michael A. Forbes. Deterministic divisibility testing via shifted partial derivatives. In *FOCS*. IEEE Computer Society, 2015.

5 Michael A. Forbes, Amir Shpilka, and Ben Lee Volk. Succinct hitting sets and barriers to proving algebraic circuits lower bounds. *CoRR*, abs/1701.05328, 2017. URL: http://arxiv.org/abs/1701.05328.

6 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. URL: http://doi.acm.org/10.1145/2591796.2591824, doi:10.1145/2591796.2591824.

7 Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *STOC*, pages 577–582, 1998. URL: http://doi.acm.org/10.1145/276698.276872.

8 Joshua A. Grochow, Mrinal Kumar, Michael E. Saks, and Shubhangi Saraf. Towards an algebraic natural proofs barrier via polynomial identity testing. *CoRR*, abs/1701.01717, 2017. URL: http://arxiv.org/abs/1701.01717.

9 Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 578–587, 2013. URL: http://dx.doi.org/10.1109/FOCS.2013.68, doi:10.1109/FOCS.2013.68.

10 Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *J. ACM*, 61(6):33:1–33:16, 2014. doi:10.1145/2629541.

11 Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012.

12 Pascal Koiran. Shallow circuits with high-powered inputs. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 7-9, 2011. Proceedings*, pages 309–320, 2011. URL: http://conference.itcs.tsinghua.edu.cn/ICS2011/content/papers/5.html.

13 Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012.

14 Pascal Koiran, Natacha Portier, and Sébastien Tavenas. A wronskian approach to the real tau-conjecture. *J. Symb. Comput.*, 68:195–214, 2015. URL: http://dx.doi.org/10.1016/j.jsc.2014.09.036.

15 Pascal Koiran, Natacha Portier, Sébastien Tavenas, and Stéphan Thomassé. A tau - conjecture for newton polygons. *Foundations of Computational Mathematics*, 15(1):185–197, 2015. doi:10.1007/s10208-014-9216-x.

16 David J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.

17 Ketan Mulmuley and Milind A. Sohoni. Geometric complexity theory I: an approach to the P vs. NP and related problems. *SIAM J. Comput.*, 31(2):496–526, 2001. URL: http://dx.doi.org/10.1137/S009753970038715X.

18 Noam Nisan and Avi Wigderson. Lower bounds for arithmetic circuits via partial derivatives (preliminary version). In *FOCS*, pages 16–25, 1995. URL: `http://dx.doi.org/10.1109/SFCS.1995.492458`.

19 Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Version 3.1.0, `https://github.com/dasarpmar/lowerbounds-survey/releases`, 2016.

20 Amir Shpilka, Amir Yehudayoff, et al. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends® in Theoretical Computer Science*, 5(3–4):207–388, 2010.

21 Michael Shub and Steve Smale. On the intractability of hilbert's nullstellensatz and an algebraic version of" NP != P ?". *Duke Mathematical Journal*, 81(1):47–54, 1995.

22 Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013. `doi:10.1007/978-3-642-40313-2_71`.

23 Leslie G. Valiant. Completeness classes in algebra. In *STOC*, pages 249–261, 1979. URL: `http://doi.acm.org/10.1145/800135.804419`.

A Proof of Lemma 3.1

Lemma 3.1 *For any $S \subseteq \{1\ldots,n\}$, $k \geq 1$, and polynomials f and g:*

$$\mathrm{PMD}_S^k(f+g) \leq \mathrm{PMD}_S^k(f) + \mathrm{PMD}_S^k(g).$$

Proof. Note that $\partial_{\mathsf{ML}}^{=k}(f+g) = \partial_{\mathsf{ML}}^{=k}(f) + \partial_{\mathsf{ML}}^{=k}(g)$, where for any two sets A and B, $A+B = \{a+b \mid a \in A, b \in B\}$. Therefore, $\mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k}(f+g)))\right\} \subseteq \mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k}f))\right\} \oplus \mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k}g))\right\}$ where $\oplus$ denotes the direct sum of vector spaces. Hence $\mathrm{PMD}_S^k(f+g) \leq \mathrm{PMD}_S^k(f) + \mathrm{PMD}_S^k(g)$. ◀

B Proof of Lemma 3.2

Lemma 3.2 *For any $S \subseteq \{1,\ldots,n\}$ with $|S| = n/2+1$ and $k = 3n/4$ we have:*

$$\mathrm{PMD}_S^k(x_1\cdots x_n) \geq \binom{n/2-1}{n/4} \geq 2^{n/2}/n^2.$$

Proof. Let $T \subseteq \{1,\ldots,n\}$ with $|T| = k$. Then $\frac{\partial^k}{\partial T}(x_1\cdots x_n) = \prod_{i\notin T} x_i$. Note that if $S \cap \overline{T} = \emptyset$ then we have $\pi_S(\pi_{\mathsf{m}}(\frac{\partial^k}{\partial T}(x_1\cdots x_n))) = \prod_{i\notin T} x_i$ since setting variables in S to zero does not affect the variables in $\overline{T}$. Otherwise, if $S\cap\overline{T} \neq \emptyset$ then $\pi_S(\pi_{\mathsf{m}}(\frac{\partial^k}{\partial T}(x_1\cdots x_n))) = 0$. Thus, we have:

$$\mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial_{\mathsf{ML}}^{=k}(x_1\cdots x_n)))\right\} \supseteq \mathbb{F}\text{-Span}\left\{\prod_{i\in T} x_i \;\middle|\; T \subseteq \overline{S}, |T| \leq n/4\right\}.$$

Hence, $\mathrm{PMD}_S^k(x_1\cdots x_n) \geq \binom{n/2-1}{n/4} \geq 2^{n/2}/n^2$ using Stirling's approximation. ◀

C Proof of Lemma 3.3

Lemma 3.3 *For any* $\beta_0, \beta_1, \ldots, \beta_n \in \mathbb{F}, \alpha, d \in \mathbb{N}$ *and for any* $S \subseteq \{1, \ldots, n\}$ *with* $|S|+k > n$*:*

$$\mathrm{PMD}^k_S((\sum_{i=1}^{n} \beta_i x_i^d + \beta_0)^\alpha) \leq 1.$$

Proof. Let $f = \sum_{i=1}^n \beta_i x_i^d + \beta_0$. For any $T \subseteq \{x_1, \ldots, x_n\}$ with $|T| = k$, $\frac{\partial^k f^\alpha}{\partial T} = \left(\gamma \prod_{x_i \in T} \beta_i x_i^{d-1}\right) f^{\alpha-k}$ for some $\gamma \in \mathbb{F}$. Note that for any monomial M such that support$(M) \cap S \neq \emptyset$, we have $\pi_S(M) = 0$. The condition $k + |S| > n$ implies that $T \cap S \neq \emptyset$ for any $T \subseteq \{x_1, \ldots, x_n\}$ with $|T| = k$. This means that S has at least one variable index in common with every monomial of the derivative $\frac{\partial^k f^\alpha}{\partial T}$ provided $d \geq 2$. Therefore $\mathrm{PMD}^k_S(f^\alpha) = 0$ when $d \geq 2$. For the case when $d = 1$, we have $\mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial^{=k}_{\mathsf{ML}}(f^\alpha)))\right\} \subseteq \mathbb{F}\text{-Span}\left\{f^{\alpha-k}\right\}$ and hence $\mathrm{PMD}^k_S(f^\alpha) \leq 1$. ◀

D Proof of Lemma 3.6

Lemma 3.6 *For any* $\beta_0, \beta_1, \ldots, \beta_r \in \mathbb{F}$*,* $\alpha, d \in \mathbb{N}$ *and for any* $S \subseteq \{1, \ldots, n\}$ *with* $|S|+k > n$*, we have* $\mathrm{PMD}^k_S((\sum_{j=1}^r \beta_j x_{i_j}^{d_j} + \beta_0)^\alpha) \leq 1$ *where* $1 \leq i_j \leq n$ *and either* $\forall j\ d_j \geq 2$ *or* $\forall j\ d_j = 1$*.*

Proof. Let $f = (\sum_{j=1}^r \beta_{i_j} x_{i_j}^{d_j} + \beta_0)$. Firstly, consider the case when $\forall j\ d_j \geq 2$. For convenience, we write $f = (\sum_{i=1}^n p_i(x_i) + \beta_0)$ where for $1 \leq i \leq n$, $p_i(x_i)$ is a univariate polynomial without a linear or constant term. Then, for any $T \subseteq \{x_1, \ldots, x_n\}$ with $|T| = k$, $\frac{\partial^k f^\alpha}{\partial T} = \left(\gamma \prod_{x_i \in T} \frac{\partial p_i(x_i)}{\partial x_i}\right) f^{\alpha-k}$ for some $\gamma \in \mathbb{F}$. Note that for $1 \leq i \leq n$, $\frac{\partial p_i}{\partial x_i}(0) = 0$. Thus for any T such that $T \cap S \neq \emptyset$, we have $\pi_S(\prod_{i \in T} \frac{\partial p_i}{\partial x_i}) = 0$. The condition $k + |S| > n$ implies that $T \cap S \neq \emptyset$ for any $T \subseteq \{x_1, \ldots, x_n\}$ with $|T| = k$. This means that S has at least one variable index in common with every monomial of the derivative $\frac{\partial^k f^\alpha}{\partial T}$ provided $d_j \geq 2\ \forall j$. Therefore $\mathrm{PMD}^k_S(f^\alpha) = 0$ when $d \geq 2$. For the case when $\forall j\ d_j = 1$, we have $\mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial^{=k}_{\mathsf{ML}}(f^\alpha)))\right\} \subseteq \mathbb{F}\text{-Span}\left\{f^{\alpha-k}\right\}$ and hence $\mathrm{PMD}^k_S(f^\alpha) \leq 1$. ◀

E Proof of Corollary 3.7

Corollary 3.7 *Let* $f = f_1^{\alpha_1} + \cdots + f_s^{\alpha_s}$ *where for every* i*, either* f_i *is a linear form or* $f_i = \sum_{j=1}^n \beta_{i,l_j} x_{l_j}^{d_{i_j}} + \beta_{i0}$ *for* $d_{i_j} \geq 2$ *and* $\beta_{i,l_j} \in \mathbb{F}$*. If* $f = x_1 \cdots x_n$ *then* $s = 2^{\Omega(n)}$*. Moreover,* $|\{i \mid f_i \textit{ is linear}\}| = 2^{\Omega(n)}$*.*

Proof. Let $S \subset \{1, \ldots, n\}$ with $|S| = n/2 + 1$ and $k = 3n/4$. From Lemmas 3.6 and 3.1 we have $\mathrm{PMD}^k_S(f) \leq \sum_{i=1}^s \mathrm{PMD}^k_S(f_i^{\alpha_i}) \leq s$. Hence by Lemma 3.2 we have $s \geq 2^{n/2}/n^2$. Further,since $\mathrm{PMD}^k_S(f_i^{\alpha_i})$ is non-zero only if f_i is a linear form, $|\{i \mid f_i \text{ is linear}\}| = 2^{\Omega(n)}$. ◀

F Proof of Lemma 3.11

Lemma 3.11 *Over fields of unbounded size, the polynomial* $x_1 \cdots x_n$ *can be computed by a homogeneous* $\Sigma \wedge^{[\sqrt{n}]} \Sigma^{[O(2^{\sqrt{n}})]} \wedge^{[\sqrt{n}]} \Sigma$ *circuit of size* $2^{O(\sqrt{n})}$*.*

Proof. Write $x_1 \cdots x_n = (x_1 \cdots x_{\sqrt{n}}) \times (x_{\sqrt{n}+1} \cdots x_{2\sqrt{n}}) \times \cdots \times (x_{n-\sqrt{n}+1} \cdots x_n)$. By Fischer's identity, each of the products $x_{(i-1)\sqrt{n}+1} \cdots x_{i\sqrt{n}}$ for $1 \leq i \leq \sqrt{n}$ can be computed by a $\Sigma \wedge \Sigma$ circuit of size $2^{O(\sqrt{n})}$. Applying the Fischer's identity once again, we get the required circuit. ◀

G Proof of Lemma 4.2

Lemma 4.2 *Let $f = p_d(x_{j_1}, \ldots, x_{j_m})$ for some $j_1, \ldots, j_m \in \{1, \ldots, n\}$. Then for any $\alpha, l, k \geq 1$*

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}_{\mathsf{ML}} f^{\alpha}\right\}\right) \leq (k+1)\binom{n+m+k+l}{k+l}.$$

Proof. Let $i_1 < i_2 < \cdots < i_k \in \{1, \ldots, n\}$. Note that

$$\frac{\partial^k f^{\alpha}}{\partial x_{i_1} \cdots \partial x_{i_k}} = \begin{cases} \alpha(\alpha-1)\cdots(\alpha-k+1)d^k f^{\alpha-k} x_{i_1}^{d-1} \cdots x_{i_k}^{d-1} & \text{if } \{i_1 \ldots, i_k\} \subseteq \{j_1, \ldots, j_m\}, \\ 0 & \text{otherwise.} \end{cases}$$

Now, relabelling the powers $x_{j_1}^{d-1}, \ldots, x_{j_m}^{d-1}$ as new variables $y_1, \ldots, y_m$ and shifting the resulting polynomials by monomials of degree at most l we get:

$$\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}_{\mathsf{ML}} f^{\alpha}\right\} \subseteq \mathbb{F}\text{-Span}\left\{\bigcup_{0 \leq i \leq k} f^{\alpha-i} \cdot S|_{y_1 = x_{j_1}^{d-1}, \ldots, y_m = x_{j_m}^{d-1}}\right\}$$

where $S = \mathcal{M}_{\leq k+l}(\{x_1, \ldots, x_n, y_1, \ldots, y_m\})$, the set of all monomials of degree at most $k+l$ in the variables $\{x_1, \ldots, x_n, y_1, \ldots, y_m\}$. Therefore,

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}_{\mathsf{ML}} f^{\alpha}\right\}\right) \leq (k+1) \cdot \binom{n+m+k+l}{k+l}. \qquad \blacktriangleleft$$

H Proof of Lemma 4.3

Lemma 4.3 *Let $\ell_1, \ldots \ell_t$ be linear forms in $\mathbb{F}[x_1, \ldots, x_n]$ with $\dim(\mathbb{F}\text{-Span}\{\ell_1, \ldots, \ell_t\}) = r$ and $f = p_d(x_{j_1}, \ldots, x_{j_m}, \ell_1, \ldots, \ell_t)$. Then for any $d > k > 0$, we have*

$$\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}_{\mathsf{ML}} f^{\alpha}\right\}\right) \leq (\alpha+1)(k+1)^3(dk)^r\binom{m+n+k+l}{k+l}.$$

Proof. We have $f^{\alpha} = \sum_{i=0}^{\alpha}\binom{\alpha}{i} p_d(x_{j_1}, \ldots, x_{j_m})^i p_d(\ell_1, \ldots, \ell_t)^{\alpha-i}$ then by sub-additivity,

$$\begin{aligned}
&\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}_{\mathsf{ML}} f^{\alpha}\right\}\right) \leq \sum_{i=0}^{\alpha}\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}_{\mathsf{ML}}\left(p_d(x_{j_1}, \ldots, x_{j_m})^i p_d(\ell_1, \ldots, \ell_t)^{\alpha-i}\right)\right\}\right) \\
&\leq \sum_{i=0}^{\alpha}\sum_{j=0}^{k}\dim\left(\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=j}_{\mathsf{ML}} p_d(x_{j_1}, \ldots, x_{j_m})^i\right\}\right)\dim\left(\partial^{\leq k-j}_{\mathsf{ML}}\mathbb{F}\text{-Span}\left\{p_d(\ell_1, \ldots, \ell_t)^{\alpha-i}\right\}\right) \\
&\leq (\alpha+1)(k+1)(k+1)(dk)^r\,(k+1)\binom{n+m+k+l}{k+l} \text{ by Lemma 4.1 and Lemma 4.2.}
\end{aligned}$$

For the penultimate inequality, note that

$$\begin{aligned}
&\mathbb{F}\text{-Span}\left\{\boldsymbol{x}^{\leq l}\partial^{=k}_{\mathsf{ML}} p_d(x_{j_1}, \ldots, x_{j_m})^i p_d(\ell_1, \ldots, \ell_t)^{\alpha-i}\right\} \\
&\subseteq \mathbb{F}\text{-Span}\left\{\bigcup_{j=0}^{k}\boldsymbol{x}^{\leq l}\partial^{=j}_{\mathsf{ML}} p_d(x_1, \ldots, x_m)^i \odot \partial^{\leq k-j}_{\mathsf{ML}}(p_d(\ell_1, \ldots, \ell_t)^{\alpha-i})\right\}.
\end{aligned} \qquad \blacktriangleleft$$

I Proof sketch of Corollary 5.1

Corollary 5.1 *There is a deterministic $n^{O(\log s)}$ time algorithm that given a multilinear polynomial $g \in \mathcal{C}_1 \cup \mathcal{C}_2$ with $\deg(g)$ tests if $g \equiv 0$.*

Proof (Sketch): The proof is a generalization of the arguments in Forbes [4] to projected multilinear derivatives. (See Proposition 4.18 in [4]). The argument when $g \in \mathcal{C}_1$ is exactly the same as in [4]. For the case when $g \in \mathcal{C}_2$, we argue that if $g \not\equiv 0$, then the trailing monomial in g will have at most $O(\log s)$ variables. Recall that trailing monomial of g, denoted by $\mathrm{TM}(g)$ is the smallest monomial with non-zero coefficient in g with respect to the lexicographic ordering induced by $x_1 > x_2 > \cdots > x_n$. Suppose S is the set of variables in $\mathrm{TM}(g)$. Since g is multilinear, we have $g|_{\overline{S}\to 0} = \prod_{i\in S} x_i$. Then, by Theorem 1.1 we have $s \geq 2^{\Omega(|S|)}$, and hence $|S| \leq c \log s$ for some constant $c > 0$. Now, testing if $g \equiv 0$ can be done by the following algorithm:

1. For all $S \subseteq \{1, \ldots, n\}$ with $|S| \leq c \log s$ do steps 2 & 3.
2. Let $g' \triangleq g(x_j = 0 \mid j \notin S)$.
3. For $a_S \in \{0,1\}^{|S|}$, if $g'(a_S) \neq 0$ then reject and halt.
4. Accept and halt.

Now, testing if $g \equiv 0$ can be done in time $n^{O(\log s)}$ by enumerating all $S \subseteq \{1, \ldots, n\}$, with $|S| \leq c \log s$, and testing if $g|_{\overline{S}\to 0} \equiv 0$. ◀

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il