Exponential lower bounds for some depth five powering circuits

Suryajith Chillara*
IIT Bombay.
suryajith@cse.iitb.ac.in

Christian Engels
IIT Bombay.
christian@cse.iitb.ac.in

B. V. Raghavendra Rao
Indian Institute of Technology, Madras.
bvrr@cse.iitm.ac.in

Ramprasad Saptharishi
TIFR, Mumbai.
ramprasad@tifr.res.in

Karteek Sreenivasaiah†
IIT Hyderabad.
karteek@iith.ac.in

November 5, 2018

Abstract

Depth five powering circuits are arithmetic circuits of the form $\Sigma\wedge\Sigma\wedge\Sigma$ where 'Σ' and '$\wedge$' represent gates that compute sum and power of their inputs respectively. Such circuits compute polynomials of the form $\sum_{i=1}^{t} Q_i^{\alpha_i}$, where each Q_i is a sum of powers of linear polynomials. These circuits are a natural generalization of the well known class of depth three powering circuits ($\Sigma\wedge\Sigma$ circuits). In this paper, we study the complexity of some restricted classes of depth five powering circuits that compute the monomial $x_1x_2\cdots x_n$. The restrictions we study are on the fan-ins of the Σ and $\wedge$ gates (denoted by $\Sigma^{[a]}$ and $\wedge^{[d]}$) while the bottom Σ gates compute homogeneous linear polynomials.

We prove size lower bound of $2^{\Omega(n)}$ against the following classes of depth five powering circuits that compute the monomial $x_1x_2\cdots x_n$.

- The $\Sigma\wedge\Sigma^{[m]}\wedge^{[\geq d]}\Sigma_{\mathrm{h}}$ circuits where $d \geq 8$ and $m < \frac{0.29n}{d^2}2^{0.955d}$.
- The $\Sigma\wedge\Sigma^{\{r\}}\wedge^{[=d]}\Sigma_{\mathrm{h}}$ circuits where $r < \varepsilon\cdot d$ is the rank of the linear forms under that sum gate, $\varepsilon > 0$ is a constant and $d \geq 8$.

From this, we infer that the monomial can be computed by a $\Sigma\wedge\Sigma^{[=2^{\sqrt{n}}]}\wedge^{[=\sqrt{n}]}\Sigma_{\mathrm{h}}$ (or a $\Sigma\wedge\Sigma^{\{=\sqrt{n}\}}\wedge^{[=\sqrt{n}]}$ Σ_{h}) circuit of size $2^{O(\sqrt{n})}$ but any $\Sigma\wedge\Sigma^{[\leq 2^{0.955\sqrt{n}}]}\wedge^{[\geq\sqrt{n}]}\Sigma_{\mathrm{h}}$ (or a $\Sigma\wedge\Sigma^{\{\leq\varepsilon\sqrt{n}\}}\wedge^{[=\sqrt{n}]}\Sigma_{\mathrm{h}}$) circuit computing it must be of size $2^{\Omega(n)}$.

Our results show 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.

*Research supported by TCS PhD fellowship. This work was done while the author was a graduate student at Chennai Mathematical Institute. Part of this research was done while the author was visiting TIFR, Mumbai where he was hosted by Ramprasad Saptharishi.

†This work was also done while the author was working at Max Planck Institute for Informatics, Saarbrücken supported by IMPECS post doctoral fellowship.

1 Introduction

Arithmetic circuits were introduced by Valiant [Val79] as a natural model for algebraic computation. He conjectured that the permanent polynomial, Perm_n, does not have polynomial sized arithmetic circuits. Following Valiant's work, there has been several efforts towards a resolution of Valiant's conjecture[1]. Therefore, 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 super-linear in the number of variables [BS83].

Lack of progress on lower bounds against the general arithmetic circuits lead researchers to explore restricted classes of circuits. Grigoriev and Karpinski [GK98] proved an exponential 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 [AV08] (see also [Koi12, Tav15, CKSV16]) 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. [GKKS13].

Gupta et al. [GKKS14] 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, Kayal et al. [KSS14] and Fournier et al. [FLMS14] obtained a super polynomial lower bound against depth four bounded fan-in circuits computing polynomials in VNP and VP respectively. Further, the techniques in [GKKS14, Kay12] 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 (cf. [Sap15] for a detailed survey of these results and references therein).

Existing 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.

Circuit Model. In this paper we consider the class of depth five powering circuits, i.e., $\Sigma\wedge\Sigma\wedge\Sigma$ circuits. It was shown in [GKKS13] that any (homogeneous) polynomial f of degree d over a sufficiently large field, that is 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 $s^{O(\sqrt{d\log n\log(sd)})}$ for a suitably chosen a. Here the superscript $[a]$ for a gate denotes the fan-in (degree in the case of $\wedge$ gates) at that level and a subscript 'h' for a Σ gate denotes that the summation is over homogeneous polynomials. This was an intermediary step in [GKKS13].

This, combined with the result of Tavenas [Tav15], shows that to prove Valiant's hypothesis over *large* 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.

Models (1) and (3) have received extensive attention in the literature compared to 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 others. However, known lower bounds for model (1) ([KSS14, FLMS14]) so far do not even imply a super polynomial lower bound for model (2). In this article, we consider two restrictions of model (2) mentioned above and prove exponential lower bounds against those circuits for computing the monomial $x_1x_2\cdots x_n$.

Throughout, it helps to interpret the polynomials computed by $\Sigma\wedge\Sigma\wedge\Sigma$ as sums of powers of linear

[1] Also, referred to as Valiant's hypothesis.

projections[2] of power symmetric polynomials. A power symmetric polynomial of degree d over the variables $\{y_1,\ldots,y_m\}$ is the polynomial $P_m^d(y_1,\ldots,y_m)=y_1^d+\cdots+y_m^d$. For any integers $d,n>0$, Ellison [Ell69] showed that for any n variate polynomial $f(x_1,\ldots,x_n)$ of degree d, there is an integer $m>0$ such that f can be written as a projection of P_m^d. In other words, there exists an integer $m=m(n)$ such that every polynomial can be expressed as the sum of dth powers of m linear polynomials. For a polynomial f, the minimum such m is called the Waring rank of f and is denoted by $\mathrm{wrk}(f)$. Fischer [Fis94] showed that $\mathrm{wrk}(x_1x_2\cdots x_n)$ is at most 2^{n-1} by giving an explicit set of linear forms (cf. Proposition 2.7). Using the technique of partial derivatives, Saxena [Sax08] showed that $\mathrm{wrk}(x_1x_2\cdots x_n)\geq 2^{\Omega(n)}$ which is a linear factor away from the upper bound (cf. [CKW11, Chapter 10]). Ranestad and Schreyer [RS00] proved that the $\mathrm{wrk}(x_1x_2\cdots x_n)$ is exactly 2^{n-1} using algebraic geometry. Recently, Balaji et al. [BLSS17] gave an elegant proof of the same using elementary linear algebra.

Let us now consider arithmetic circuits that use only sum and powering gates[3]. The expression in the form of the sum of powers of linear polynomials is a depth three powering circuit, a restriction of the general depth three circuits. Since there exists a depth three powering circuit of size at most $n2^{n-1}$ to compute a monomial, the computational model is *universal* for polynomial computations. In fact, there is a powering circuit of depth $(d+1)$ and size $O\left(n^d\cdot 2^{d\cdot n^{\frac{1}{d}}}\right)$ that computes $x_1x_2\cdots x_n$. The aforementioned lower bounds on the Waring rank imply a size lower bound of $2^{\Omega(n)}$ for any depth three powering circuit computing the monomial $x_1x_2\cdots x_n$. In [CKW11] Chen et al. posed the following open question which is another motivation for the work presented in this article.

Question 1.1 ([CKW11])**.** *Can the monomial $x_1x_2\cdots x_n$ be efficiently computed by a constant depth powering circuit?*

Saptharishi[4] [Sap15] observed that the monomial $x_1x_2\cdots x_n$ has non-trivial $\Sigma\wedge\Sigma\wedge$ and $\Sigma\wedge\Sigma\wedge\Sigma$ circuits of size $2^{O(\sqrt{n})}$ (cf. Lemma 3.1). Ideally, we would like to prove matching lower bounds but the current state of affairs is far away from this. We make partial progress.

Kayal [Kay12], using the technique of shifted partial derivatives proved an exponential bound of $2^{\Omega\left(\frac{n}{d}\right)}$ against any $\Sigma\wedge\Sigma_{\mathrm{h}}\Pi^{[d]}$ circuit computing the monomial $x_1x_2\cdots x_n$. If there is a $\Sigma\wedge\Sigma\Pi^{[d]}$ circuit of size s then there is a $\Sigma\wedge\Sigma^{[m]}\wedge^{[d]}\Sigma_{\mathrm{h}}$ circuit of size $m\cdot s$ where $m=2^d\cdot\binom{n+d}{n}$. Thus, Kayal's bound implies an exponential size lower bound for the $\Sigma\wedge\Sigma_{\mathrm{h}}^{[m]}\wedge^{[d]}\Sigma_{\mathrm{h}}$ circuit computing the monomial when $d\leq\sqrt{\frac{n}{\log n}}$ and $m=2^d\cdot\binom{n+d}{n}$.

Our Results. We consider the depth five powering circuits which compute polynomials of the form $g=\sum_{i=1}^s f_i^{\alpha_i}$ where $f_i=(\ell_1^d+\cdots+\ell_n^d+c_i)$, ℓ_js are homogeneous linear polynomials, c_is are non-zero elements, and d is at least 8. We use the dimension of projections of *multilinear derivatives* as a complexity measure to prove such a result. When $d=1$, this reduces to the case of depth three powering circuits. We summarize this result in the following theorem.

Theorem 1.2. *Let $g=\sum_{i=1}^s f_i^{\alpha_i}$ where $f_i=(\ell_1^{d_{i1}}+\cdots+\ell_m^{d_{im}}+c_i)$, ℓ_js are homogeneous linear polynomials over $\mathbb{F}[X]$ and $c_i\in\mathbb{F}^*$. Let d be the minimum of $\left\{d_{ij}\mid(i,j)\in[s]\times[m]\right\}$. If $g\equiv x_1x_2\cdots x_n$ then for $m=n$ and $d\geq 8$, $s=2^{\Omega(n)}$.*

[2] A polynomial $f(x_1,\ldots,x_n)$ is said to be a linear projection of the polynomial $g(y_1,\ldots,y_m)$ if there exist linear polynomials $\{\ell_1,\ldots,\ell_m\}$ over $\mathbb{F}[x_1,\ldots,x_n]$ such that $f=g(\ell_1,\ldots,\ell_m)$.

[3] A powering gate takes in the tuple (f,d) as the input and output the polynomial f^d. It is denoted by $\wedge$.

[4] Saptharishi attributes the observation to Forbes.

Further, we also observe that such a bound also holds for larger values of m. That helps us arrive at the following corollary.

Corollary 1.3. *For any integer n, the monomial $x_1x_2\cdots x_n$ can be computed by a circuit of the form $\Sigma\wedge\Sigma_{\mathrm{h}}^{[2\sqrt{n}]}\wedge^{[=\sqrt{n}]}\Sigma_{\mathrm{h}}$ of size $2^{O(\sqrt{n})}$ but any $\Sigma\wedge\Sigma^{[2^{0.955\sqrt{n}}]}\wedge^{[\geq\sqrt{n}]}\Sigma_{\mathrm{h}}$ computing it must be of size at least $2^{\Omega(n)}$.*

We also consider the depth five powering circuits of the form $g=\sum_{i=1}^{s}f_i^{\alpha_i}$ where $f_i=(\ell_1^d+\cdots+\ell_m^d+c_i)$ where ℓ_is are homogeneous linear polynomials, c_is are non-zero field elements, and ℓ_is form a low rank subspace. Formally, we prove the following theorem.

Theorem 1.4. *Let $g=\sum_{i=1}^{s}f_i^{\alpha_i}$ be such that $f_i=(\ell_1^d+\cdots+\ell_m^d+c_i)$, ℓ_is are homogeneous linear polynomials over $\mathbb{F}[X]$ and $c_i\in\mathbb{F}^*$. Let $r<\varepsilon d$ be the rank of the linear forms $\{\ell_1,\ell_2,\ldots,\ell_m\}$ for a parameter $\varepsilon<1$. If $g\equiv x_1x_2\cdots x_n$ then there exists a suitable value for the parameter ε such that $s=2^{\Omega(n)}$.*

This gives us the following insight.

Corollary 1.5. *For any integer n, the monomial $x_1x_2\cdots x_n$ can be computed by a circuit of the form $\Sigma\wedge\Sigma_{\mathrm{h}}^{\{=\sqrt{n}\}}\wedge^{[=\sqrt{n}]}\Sigma_{\mathrm{h}}$ of size $2^{O(\sqrt{n})}$ but any $\Sigma\wedge\Sigma^{\{\leq\varepsilon\sqrt{n}\}}\wedge^{[=\sqrt{n}]}\Sigma_{\mathrm{h}}$ computing it must be of size at least $2^{\Omega(n)}$.*

2 Preliminaries

Arithmetic Circuits. An arithmetic circuit over a field $\mathbb{F}$ over the set of variables $\{x_1,x_2,\ldots,x_n\}$ is a directed acyclic graph such that the internal nodes are labelled by addition or multiplication gates and the leaf nodes are labelled by variables or field elements. Any node with fan-out zero is an output gate. An arithmetic circuit computes a polynomial in the polynomial ring $\mathbb{F}[x_1,x_2,\ldots,x_n]$. The size of an arithmetic circuit is the number of nodes and the depth is the length of a longest path from the root to a leaf node. A powering circuit is an arithmetic circuit where the internal nodes are either addition ($+$) or powering gates ($\wedge$).

In this work, we will study two specific restrictions of powering circuits of depth five.

Definition 2.1 (Depth five powering circuits)**.** *A depth five powering circuit computes the sum of powers of sums of powers of linear forms. Formally, a $\Sigma\wedge\Sigma\wedge\Sigma$ circuit computes the polynomials of the form*

$$\sum_{i=1}^{s}\left(\sum_{j=1}^{m}\ell_{ij}^{d_{ij}}+c_i\right)^{\alpha_i}$$

where ℓ_is are linear forms over $\mathbb{F}[x_1,x_2,\ldots,x_n]$. $\Diamond$

Here are the following two restrictions of the depth five powering circuits that we consider.

- A $\Sigma\wedge\Sigma^{[m]}\wedge^{[\geq d]}\Sigma_{\mathrm{h}}$ circuit computes the polynomials of the form

$$\sum_{i=1}^{s}\left(\sum_{j=1}^{m}\ell_{ij}^{d_{ij}}+c_i\right)^{\alpha_i}$$

  such that, for any $i\in[s]$, the linear forms $\left\{\ell_{ij}\right\}_{i,j}$ are homogeneous linear polynomials and $d=\min\left\{d_{ij}\mid(i,j)\in[s]\times[m]\right\}$.

- A $\Sigma\wedge\Sigma^{\{r\}}\wedge^{[=d]}\Sigma_{\mathrm{h}}$ circuit computes the polynomials of the form

$$\sum_{i=1}^{s}\left(\sum_{j=1}^{m}\ell_{ij}^{d}+c_i\right)^{\alpha_i}$$

such that, m is unbounded and for any $i\in[s]$, the linear forms $\left\{\ell_{ij}\right\}_{i,j}$ are homogeneous linear polynomials and for every i, the rank of the linear system $\left\{\ell_{ij}:j\in[m]\right\}$ is at most r.

Notation. The superscript a in $\Sigma^{[a]}$ for a sum gate denotes the fan-in at that level. Similarly, superscript d in $\wedge^{[d]}$ denotes the degree of $\wedge$ gates. Further, Σ_{h} denotes a summation over homogeneous polynomials.

Approximation of the Binomial Coefficients

The following estimates are key to the asymptotic estimates required for our analysis.

Proposition 2.2. *For all n, k such that $k\le 0.5n$, the following bounds for the binomial coefficient $\binom{n}{k}$ hold.*

1. $\left(\frac{n}{k}\right)^k\le\binom{n}{k}\le\left(\frac{e\cdot n}{k}\right)^k$.
2. $\binom{n}{k}\sim 2^{n\cdot H\left(\frac{k}{n}\right)}$ *where H is the binary entropy function.*

The following property is true about the binary entropy function.

Claim 2.3. *For positive real numbers $\left\{a_1,a_2,\ldots,a_p,u_1,u_2,\ldots,u_p\right\}$, each of which is in $[0,1]$, such that $u_1+u_2+\cdots+u_p=1$,*

$$\sum_{i=1}^{p}u_iH(a_i)\le H\left(\sum_{i=1}^{p}u_ia_i\right)$$

where $H(q)$ is the binary entropy function.

The proof of this claim follows easily from the fact that the binary entropy function is a concave function over $(0,1)$.

Complexity Measure: Projected Multilinear Derivatives

We begin with a definition of the complexity measure used in this article:

Definition 2.4. *Let $k>0$ be an integer and let $S\subset[n]$ be a set of indices. The dimension of the projected multilinear derivatives is defined as*

$$\mathrm{PMD}_S^k(f)=\dim(\mathbb{F}\text{-}span\left\{\sigma_S\left(\pi_m(\partial_{\mathrm{ML}}^{=k}f)\right)\right\})$$

where $\partial_{\mathrm{ML}}^{=k}f$ is the set of partial derivatives of order k with respect to the multilinear monomials, σ_S sets the variables indexed by the set S to zero and π_m projects the polynomial to its mulitilinear component. ◊

Let us also consider a stripped down variant of the projected multilinear derivatives.

Definition 2.5. *For an integer $k > 0$, we define the dimension of the multilinear derivatives of order k (denoted by MD_k) as follows.*

$$\mathrm{MD}_k(f) = \dim(\mathbb{F}\text{-}span\{\pi_m(\partial_{\mathrm{ML}}^{=k} f)\})$$

where $\partial_{\mathrm{ML}}^{=k} f$ is the set of partial derivatives of order k with respect to the multilinear monomials and π_m projects the polynomial to its mulitilinear component. ◊

Since the polynomial we consider is a multilinear polynomial, it is sufficient to consider the derivatives with respect to multilinear monomials and multilinear projections. This measure is a restriction of the measure introduced by Nisan and Wigderson [NW97] and is similar in flavour to the *skew partial derivatives* used by Kayal et al. [KNS16].

It is easy to see that the measures are sub-additive. We crucially use this property in our proof.

Proposition 2.6. *The measures, the dimensions of the multilinear derivatives and projected multilinear derivatives are sub-additive. Formally,*

$$\mathrm{MD}_k(f_1 + f_2) \leq \mathrm{MD}_k(f_1) + \mathrm{MD}_k(f_2),$$
$$\mathrm{PMD}_S^k(f_1 + f_2) \leq \mathrm{PMD}_S^k(f_1) + \mathrm{PMD}_S^k(f_2).$$

We shall now mention the result of Fischer [Fis94] where he showed that $\mathrm{wrk}(x_1 x_2 \cdots x_n)$ is at most 2^{n-1} by giving an explicit set of linear forms.

Proposition 2.7 ([Fis94]). *For any n, the monomial $x_1 \cdots x_n$ can be expressed as a linear combination of 2^{n-1} powers of linear forms as following.*

$$2^{n-1} \cdot n! \cdot x_1 \cdots x_n = \sum_{(r_2, r_3, \ldots, r_n) \in \{\pm 1\}^{n-1}} (-1)^{\mathrm{wt}(\mathbf{r})} \cdot (x_1 + \sum_{j=1}^{n} r_i x_i)^n$$

where $\mathrm{wt}(\mathbf{r}) = |\{i \mid r_i = -1\}|$.

For the sake of completeness, we will construct a depth five powering circuit of size $2^{\Omega(\sqrt{n})}$ that computes $x_1 x_2 \cdots x_n$.

Proposition 2.8. *There is a $\Sigma \wedge \Sigma_{\mathrm{h}}^{[2^{\sqrt{n}-1}]} \wedge^{[\sqrt{n}]} \Sigma_{\mathrm{h}}^{[\sqrt{n}]}$ formula of size $2^{\sqrt{n}}$ computing the monomial $x_1 x_2 \cdots x_n$.*

Proof. Let the monomial $x_1 x_2 \cdots x_n$ be expressed as $m_1 m_2 \cdots m_{\sqrt{n}}$ where

$$m_i = x_{((i-1)\sqrt{n}+1)} x_{((i-1)\sqrt{n}+2)} \cdots x_{(i\sqrt{n})}$$

for all $i \in [\sqrt{n}]$. Invoking Proposition 2.7 with $m_1, m_2, \ldots, m_{\sqrt{n}}$ as the variables, we get a depth three powering circuit of size $\sqrt{n} \cdot 2^{\sqrt{n}-1}$ over $\mathbb{F}\left[m_1, m_2, \ldots, m_{\sqrt{n}}\right]$. Furthermore, using Proposition 2.7 again, each of these m_is can be expressed as depth three powering circuits of size $\sqrt{n} \cdot 2^{\sqrt{n}-1}$ over $\mathbb{F}\left[x_{(i-1)\sqrt{n}+1}, x_{(i-1)\sqrt{n}+2}, \ldots, x_{i\sqrt{n}}\right]$. The $\Sigma \wedge \Sigma_{\mathrm{h}}^{[2^{\sqrt{n}-1}]} \wedge^{[\sqrt{n}]} \Sigma_{\mathrm{h}}^{[\sqrt{n}]}$ circuit thus obtained is of size at most $n \cdot 2^{2\sqrt{n}-2}$. □

Saptharishi [Sap15] gives an elegant construction of a depth four powering circuit of size $2^{O(\sqrt{n})}$ using Ellison's lemma/Newton identities in the first step and the Fisher's identity in the second step.

3 Hardness of the Monomial

Lemma 3.1. *If S is any arbitrary subset of X and $k \in [n]$ is an integer such that $k > |S|$ then*

$$\mathrm{PMD}_S^k(x_1 x_2 \cdots x_n) = \binom{n-|S|}{k-|S|}.$$

Proof. The derivative space of the monomial $M = x_1 x_2 \cdots x_n$, of order k is spanned by the multilinear monomials of degree exactly $n-k$ over $\mathbb{F}[X]$. Let us denote it by $\partial_{\mathrm{ML}}^{=k} M$. Upon applying the projection σ_S to $\partial_{\mathrm{ML}}^{=k} M$, we can see that only those monomials that are non zero are those over $\mathbb{F}[X \setminus S]$. The degree of the non zero monomials in $\sigma_S(\partial^k M)$ is at most $n-k$. The total number of monomials over $\mathbb{F}[X \setminus S]$ of degree equal to $n-k$ is equal to $\binom{n-|S|}{n-k}$. Thus,

$$\mathrm{PMD}_S^k(x_1 x_2 \cdots x_n) = \binom{n-|S|}{n-k} = \binom{n-|S|}{n-|S|-(n-k)} = \binom{n-|S|}{k-|S|}.$$

□

By setting the size of S to zero, we get that $\mathrm{MD}_k(x_1 x_2 \cdots x_n) = \binom{n}{k}$.

4 $\Sigma\wedge\Sigma\wedge$ Circuits: The Curse of Homogeneity

Firstly, we observe that homogeneous $\Sigma\wedge\Sigma\wedge$ circuits of polynomial size cannot compute the monomial $x_1 \cdots x_n$ by eliminating bottom $\wedge$ gates of degree at least 2:

Observation 4.1. *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 $s = 2^{\Omega(n)}$.*

The homogeneity condition for the bottom power gates is necessary due to the following result in [Sap15]. 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 4.2. *[Sap15, Corollary 17.16] For any $d > 0$, $\mathrm{Sym}_{n,d}$ can be computed by a $\Sigma\wedge\Sigma\wedge$ circuit of size $2^{O(\sqrt{d})}\mathrm{poly}(n)$.*[5]

Is it all about homogeneity at the bottom Σ gates? The answer is no. In fact, Observation 4.1 can also be generalized to the case of powers of polynomials in the span of the set $\{x_{i_j}^{\alpha_{i_j}} \mid 1 \le i_j \le n,\ \alpha_{i_j} \ge 2\}$:

Lemma 4.3. *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 $\mathrm{PMD}_S^k(\sum_{j=1}^n \beta_j x_{i_j}^{d_j} + \beta_0)^\alpha \le 1$ where $1 \le i_j \le n$ and either $\forall j\ d_j \ge 2$ or $\forall j\ d_j = 1$.*

Proof. Let $f = (\sum_{j=1}^r \beta_{i_j} x_{i_j}^{d_j} + \beta_0)$. Firstly, consider the case when $\forall j\ d_j \ge 2$. For convenience, we write $f = (\sum_{i=1}^n p_i(x_i) + \beta_0)$ where for $1 \le i \le n$, $p_i(x_i)$ 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

[5] In [Sap15], 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 the application of Fischer's identity produces sum gates that are not homogeneous.

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 $\mathrm{PMD}_S^k f^\alpha = 0$ when $d \geq 2$. For the case when $\forall j\ d_j = 1$, we have $\mathbb{F}\text{-span}\, \pi_S(\pi_{\mathsf{m}}(\partial_{\mathrm{ML}}^{=k}(f^\alpha))) \subseteq \mathbb{F}\text{-span}\, f^{\alpha-k}$ and hence $\mathrm{PMD}_S^k f^\alpha \leq 1$. □

We get the following generalization of Observation 4.1:

Corollary 4.4. *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, $|\{i \mid f_i \text{ is linear}\}| = 2^{\Omega(n)}$.*

Proof. Let $S \subset \{1, \ldots, n\}$ with $|S| = n/2 + 1$ and $k = 3n/4$. From Lemma 4.3 and Proposition 2.6 we have $\mathrm{PMD}_S^k f \leq \sum_{i=1}^s \mathrm{PMD}_S^k f_i^{\alpha_i} \leq s$. Hence by Lemma 3.1 we have $s \geq 2^{n/2}/n^2$ as required. Further, $\mathrm{PMD}_S^k f_i^{\alpha_i}$ is non-zero only if f_i is a linear form, and hence $|\{i \mid f_i \text{ is linear}\}| = 2^{\Omega(n)}$. □

5 Weakness of the $\Sigma\wedge\Sigma^{[m]}\wedge^{[\geq d]}\Sigma_{\mathrm{h}}$ Circuits under the Measure

In this section, we show that $\Sigma\wedge\Sigma^{[m]}\wedge^{[\geq d]}\Sigma_{\mathrm{h}}$ circuits of polynomial size cannot compute the monomial $x_1 x_2 \cdots x_n$. We start by showing that the dimension of the multilinear derivative space for any polynomial computed by this model is low.

Lemma 5.1. *Let k and $t > \frac{n}{2}$ be some parameters. Let $f = (\ell_1^{d_i} + \cdots + \ell_m^{d_m} + c)$ where ℓ_is are homogeneous linear polynomials over $\mathbb{F}[X]$ and c is a non-zero field element. Then for any positive integer α,*

$$\mathrm{MD}_k(f^\alpha) \leq k\left(\binom{m+p}{p}\binom{k}{p} + \frac{n}{2}\binom{n}{t}\right)$$

where $d = \min\{d_1, d_2, \ldots, d_m\}$ and $p < \frac{t+k}{d}$.

Proof. Note that the space of partial derivatives of order k, of f lies in $\mathbb{F}\text{-span}\{\ell_i^{d_i - k} \mid i \in [m]\}$. Extending this, we obtain the following.

$$\begin{aligned}
\partial^k f^\alpha &\subseteq \mathbb{F}\text{-span}\{f^{\alpha-p} \cdot \partial^{k_1} f \cdots \partial^{k_p} f \mid p \in [k] \text{ and } k_1 + \cdots + k_p = k\} \\
&\subseteq \mathbb{F}\text{-span}\{f^{\alpha-p} \mid p \in [k]\} \otimes \mathbb{F}\text{-span}\{\partial^{k_1} f \cdots \partial^{k_p} f \mid p \in [k] \text{ and } k_1 + \cdots + k_p = k\}.
\end{aligned}$$

The dimension of $\mathbb{F}\text{-span}\{f^{\alpha-p} \mid p \in [k]\}$ is trivially upper bounded by k. Let W be the vector space $\mathbb{F}\text{-span}\{\partial^{k_1} f \cdots \partial^{k_p} f \mid k_1 + \cdots + k_p = k\}$. Then, $\mathrm{MD}_k(f^\alpha) \leq k \cdot \mathrm{MD}_k(W)$.

$$\begin{aligned}
W &= \mathbb{F}\text{-span}\{\partial^{k_1} f \cdots \partial^{k_p} f \mid k_1 + \cdots + k_p = k\} \\
&\subseteq \mathbb{F}\text{-span}\{\ell_{i_1}^{d_{i_1} - k_1} \cdots \ell_{i_p}^{d_{i_p} - k_p} \mid \bar{i} = (i_1, \ldots, i_p) \in [m]^p \text{ and } k_1 + \cdots + k_p = k\}.
\end{aligned}$$

If the degree of any term $T_{\bar{i}} = \ell_{i_1}^{d_{i_1} - k_1} \cdots \ell_{i_p}^{d_{i_p} - k_p}$ is greater than n then its contribution to $\mathrm{MD}_k(W)$ is zero. Let us now consider all the other terms in W whose degree is at most n. Let t be a degree threshold such that $t > \frac{n}{2}$.

- Let us consider the terms $T_{\bar{i}}$ such that their degree is in $[t, n]$. There are at most $\sum_{j\in[t,n]}\binom{n}{j} \leq \binom{n}{t}\cdot(n-t+1)$ many multilinear monomials over $\mathbb{F}[X]$ of degree at least t and at most n. Their contribution to $\mathrm{MD}_k(W)$ is at most $\frac{n}{2}\binom{n}{t}$.
- Otherwise, the degree of any other term $T_{\bar{i}}$ is at most $t-1$.

$$\begin{aligned} d_{i_1}-k_1+d_{i_2}-k_2+\cdots+d_{i_p}-k_p &< t \\ p\cdot\min\left\{d_{i_1},d_{i_2},\ldots,d_{i_p}\right\}-(k_1+k_2+\cdots+k_p) &< t \\ \implies p &< \frac{t+k}{d} \end{aligned}$$

since $d \leq \min\left\{d_{i_1},d_{i_2},\ldots,d_{i_p}\right\}$. The number of terms $\{T_{\bar{i}}\}_{\bar{i}\in[m]^p}$ of degree at most $t-1$ can be counted as follows. We can choose the indices $(i_1,i_2,\ldots,i_p)$ in $\binom{m+p}{m}$ ways, and choose $k_1,k_2,\ldots,k_p$ in at most $\binom{k}{p}$ ways such that $k_1+k_2+\cdots+k_p=k$.

Thus,

$$\mathrm{MD}_k(f^\alpha) \leq k\cdot\mathrm{MD}_k(W) \leq k\left(\binom{k}{p}\binom{m+p}{p}+\frac{n}{2}\binom{n}{t}\right).$$

□

Putting it all Together

Theorem 5.2. *Let $g=\sum_{i=1}^{s} f_i^{\alpha_i}$ be such that $f_i=(\ell_1^{d_{i1}}+\cdots+\ell_m^{d_{im}}+c_i)$ such that ℓ_is are homogeneous linear polynomials over $\mathbb{F}[X]$ and $c_i\in\mathbb{F}^*$. Let d be the minimum of $\left\{d_{ij}\mid(i,j)\in[s]\times[m]\right\}$. If $g\equiv x_1x_2\cdots x_n$ then for $m=n$ and $d\geq 8$, $s=2^{\Omega(n)}$.*

Proof. Since $g=\sum_{i=1}^{s} f_i^{\alpha_i}$ and from Proposition 2.6 we can infer the following.

$$\mathrm{MD}_k(g) \leq s\cdot\max_{i\in[s]}(\mathrm{MD}_k(f_i^{\alpha_i})).$$

From Lemma 3.1 and the fact that $g\equiv x_1x_2\ldots x_n$, we can infer that $\mathrm{MD}_k(g)=\binom{n}{k}$. By invoking Lemma 5.1 we can get an upper bound on $\max_{i\in[s]}(\mathrm{MD}_k(f_i^{\alpha_i}))$. Thus,

$$s\cdot k\cdot\left(\binom{k}{p}\binom{m+p}{p}+\binom{n}{t}\right)\geq\binom{n}{k}$$

$$\implies s \geq \frac{\binom{n}{k}}{k\left(\binom{k}{p}\binom{m+p}{p}+\frac{n}{2}\binom{n}{t}\right)}.$$

Fix the value of k to $0.5n$ so as to maximize the numerator.

- $\binom{n}{k}\approx 2^n$.
- $\binom{k}{p}\binom{m+p}{p}\approx 2^{k\cdot H\left(\frac{p}{k}\right)+(m+p)\cdot H\left(\frac{p}{m+p}\right)}=2^{n\cdot\left(0.5H\left(\frac{p}{k}\right)+\frac{m+p}{n}H\left(\frac{p}{m+p}\right)\right)}$.

- $\binom{n}{t} \approx 2^{n \cdot H\left(\frac{t}{n}\right)}$.

$$\log(sk) \geq n \cdot \left(1 - \max\left\{0.5H\left(\frac{p}{k}\right) + \frac{m+p}{n} H\left(\frac{p}{m+p}\right), H\left(\frac{t}{n}\right)\right\}\right).$$

We will set the value of t to a value which is away from $0.5n$ on the greater side so that $\binom{n}{t}$ is the non-dominant term. Let us fix the parameters as follows: $k = 0.5n$, $t = 0.545n$ and $m = n$. This setting of parameters forces that d must be at least 8 and thus $p < 0.13n$. By substituting the values, we get $\binom{k}{p}\binom{m+p}{p} = 2^{0.99815n}$ and $\binom{n}{t} = 2^{0.99415n}$. Thus, $s \geq 2^{0.00184n}$. □

Theorem 5.3. *Any $\Sigma \wedge \Sigma^{[m]} \wedge^{[\geq d]} \Sigma_{\mathrm{h}}$ circuit computing the monomial $x_1 x_2 \cdots x_n$ must be of size $2^{\Omega(n)}$ where $m \leq \frac{0.29n}{d^2} \cdot 2^{0.955d}$.*

Proof. From the proof of Theorem 5.2, we get that

$$s \geq \frac{\binom{n}{k}}{k\left(\binom{k}{p}\binom{m+p}{p} + \binom{n}{t}\right)}$$

and the value of p is at most $\frac{t+k}{d}$. Let us fix the value of k to $0.5n$ again so as to maximize the numerator and we will set the value of t to a value of $0.545n$ which is away from $0.5n$ on the greater side so that $\binom{n}{t}$ is the non dominant term.

Then we want the other term in the denominator to be exponentially smaller than $2^{nH\left(\frac{k}{n}\right)}$ (say by $2^{\gamma n}$ where $\gamma = 0.998$) at its maximum.

$$\max\left(\binom{k}{p}\binom{m+p}{p}, \binom{n}{t}\right) \leq 2^{\gamma n} \implies \frac{e^2 k(m+p)}{p^2} \leq 2^{\frac{\gamma n}{p}}$$

$$(m+p) \leq \left(\frac{p^2}{e^2 k}\right) \cdot 2^{\frac{\gamma n}{p}} \implies m \leq \left(\frac{p^2}{e^2 k}\right) \cdot 2^{\frac{\gamma n}{p}} - p < \left(\frac{(t+k)^2}{e^2 d^2 k}\right) \cdot 2^{\frac{\gamma n}{p}}$$

$$m < \frac{(1.045)^2 n}{0.5 e^2 d^2} \cdot 2^{\frac{\gamma d}{1.045}} = \frac{0.29n}{d^2} \cdot 2^{0.955d}.$$

This completes the proof. □

In particular, we infer the following by setting d to $\sqrt{n}$.

Corollary 5.4. *For any integer n, the monomial $x_1 x_2 \cdots x_n$ can be computed by a circuit of the form $\Sigma \wedge \Sigma_{\mathrm{h}}^{[2^{\sqrt{n}}]} \wedge^{[=\sqrt{n}]} \Sigma_{\mathrm{h}}$ of size $2^{O(\sqrt{n})}$ but any $\Sigma \wedge \Sigma^{[2^{0.955\sqrt{n}}]} \wedge^{[\geq \sqrt{n}]} \Sigma_{\mathrm{h}}$ computing it must be of size at least $2^{\Omega(n)}$.*

6 Weakness of the $\Sigma \wedge \Sigma^{\{r\}} \wedge^{[=d]} \Sigma_{\mathrm{h}}$ Circuits under the Measure

Similar to the previous section, the dimension of the multilinear derivative space can be used to show strong lower bounds against the size of $\Sigma \wedge \Sigma^{\{r\}} \wedge^{[=d]} \Sigma_{\mathrm{h}}$ circuits computing the monomial. We follow a very similar approach as in the previous section.

Lemma 6.1. *Let k and $t > \frac{n}{2}$ be parameters. Let $f = (\ell_1^d + \cdots + \ell_m^d + c)$ where ℓ_is are homogeneous linear polynomials over $\mathbb{F}[X]$ such that the rank of $\{\ell_1, \ell_2, \ldots, \ell_m\}$ is at most r and c is a non-zero field element. Then for any positive integer α,*

$$\mathrm{MD}_k(f^\alpha) \le k\left(\binom{k}{p} \cdot 2^{(p(d+r)-k)\cdot H\left(\frac{rp}{p(d+r)-k}\right)} + \frac{n}{2}\binom{n}{t}\right)$$

where $p < \frac{t+k}{d}$.

Proof. Similar to the proof of Lemma 5.1, we have the following.

$$\begin{aligned}
\partial^k f^\alpha &\subseteq \mathbb{F}\text{-span}\{f^{\alpha-p} \cdot \partial^{k_1} f \cdots \partial^{k_p} f \mid p \in [k] \text{ and } k_1 + \cdots + k_p = k\} \\
&\subseteq \mathbb{F}\text{-span}\{f^{\alpha-p} \mid p \in [k]\} \otimes \mathbb{F}\text{-span}\{\partial^{k_1} f \cdots \partial^{k_p} f \mid p \in [k] \text{ and } k_1 + \cdots + k_p = k\}.
\end{aligned}$$

The dimension of $\mathbb{F}\text{-span}\{f^{\alpha-p} \mid p \in [k]\}$ is trivially upper bounded by k. Let W be the vector space $\mathbb{F}\text{-span}\{\partial^{k_1} f \cdots \partial^{k_p} f \mid k_1 + \cdots + k_p = k\}$. Thus, $\mathrm{MD}_k(f^\alpha) \le k \cdot \mathrm{MD}_k(W)$.

$$\begin{aligned}
W &= \mathbb{F}\text{-span}\{\partial^{k_1} f \cdots \partial^{k_p} f \mid k_1 + \cdots + k_p = k\} \\
&\subseteq \mathbb{F}\text{-span}\{\ell_{i_1}^{d-k_1} \cdots \ell_{i_p}^{d-k_p} \mid \bar{i} = (i_1, \ldots, i_p) \in [m]^p \text{ and } k_1 + \cdots + k_p = k\} \\
&\subseteq \mathbb{F}\text{-span}\left\{\ell_{i_1}^{d-k_1} \mid i_1 \in [m]\right\} \otimes \cdots \otimes \mathbb{F}\text{-span}\left\{\ell_{i_p}^{d-k_p} \mid i_p \in [m]\right\}; \quad k_1 + \cdots + k_p = k.
\end{aligned}$$

If the degree of any term $T_{\bar{i}} = \ell_{i_1}^{d-k_1} \cdots \ell_{i_p}^{d-k_p}$ is greater than n then its contribution to $\mathrm{MD}_k(W)$ is zero. Let us now consider all the other terms in W whose degree is at most n. Let $t > \frac{n}{2}$ be a degree threshold that we shall fix later. Let us consider the terms $T_{\bar{i}}$ whose degree lies in $[t, n]$. There are at most $\binom{n}{t} \cdot (n - t + 1)$ many multilinear monomials over $\mathbb{F}[X]$ of degree at least t and at most n. Otherwise, the degree of $T_{\bar{i}}$ is at most $t - 1$.

$$\begin{aligned}
d - k_1 + d - k_2 + \cdots + d - k_p &< t \\
p \cdot d - (k_1 + k_2 + \cdots + k_p) &< t \\
\implies p &< \frac{t+k}{d}
\end{aligned}$$

Without loss of generality, let us suppose that the set $\{\ell_1, \ell_2, \ldots, \ell_r\}$ forms the linear basis for $\{\ell_1, \ell_2, \ldots, \ell_m\}$. Thus, ℓ_i can be written as a linear combination of the basis elements. Also,

$$\begin{aligned}
\partial^{k_i} f &\subseteq \mathbb{F}\text{-span}\left\{\ell_j^{d-k_i} \mid j \in [m]\right\} \\
&\subseteq \mathbb{F}\text{-span}\left\{\ell_1^{a_1} \ell_2^{a_2} \ldots \ell_r^{a_r} \mid a_1 + a_2 + \cdots + a_r = d - k_i\right\}.
\end{aligned}$$

The number of integral solutions to the equation $a_1 + a_2 + \cdots + a_r = d - k_i$ is at most $\binom{d-k_i+r}{r}$. Thus for a fixed $(k_1, k_2, \ldots, k_p)$ such that $k_1 + k_2 + \cdots + k_p = k$,

$$\dim\left(\bigotimes_{j\in[p]} \mathbb{F}\text{-span}\left\{\ell_{i_j}^{d-k_j} \mid i_j \in [m]\right\}\right) \leq \prod_{i=1}^{p} \binom{d-k_i+r}{r}.$$

The number of ways of choosing $(k_1, k_2, \ldots, k_p)$ is at most $\binom{k}{p}$. Therefore,

$$\begin{aligned}\dim(W) &\leq \binom{k}{p} \cdot \prod_{i=1}^{p} \binom{d-k_i+r}{r} \\ &\leq \binom{k}{p} \cdot 2^{\sum_{i=1}^{p}(d-k_i+r)\cdot H\left(\frac{r}{d-k_i+r}\right)} \\ &= \binom{k}{p} \cdot 2^{(p(d+r)-k)\sum_{i=1}^{p} \frac{(d-k_i+r)}{p(d+r)-k}\cdot H\left(\frac{r}{d-k_i+r}\right)}\end{aligned}$$

and by using Claim 2.3 we get that

$$\begin{aligned}2^{(dm+rm-k)\sum_{i=1}^{p} \frac{(d-k_i+r)}{p(d+r)-k}\cdot H\left(\frac{r}{d-k_i+r}\right)} &\leq 2^{(p(d+r)-k)\cdot H\left(\sum_{i=1}^{p} \frac{(d-k_i+r)}{p(d+r)-k}\cdot \frac{r}{d-k_i+r}\right)} \\ &= 2^{(p(d+r)-k)\cdot H\left(\frac{pr}{p(d+r)-k}\right)}.\end{aligned}$$

Thus,

$$\mathrm{MD}_k(f^{\alpha}) \leq k \cdot \mathrm{MD}_k(W) \leq k\left(\binom{k}{p} 2^{(p(d+r)-k)\cdot H\left(\frac{pr}{p(d+r)-k}\right)} + \frac{n}{2}\binom{n}{t}\right).$$

□

Putting it all Together

Theorem 6.2. *Let $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ be such that $f_i = (\ell_1^d + \cdots + \ell_m^d + c_i)$, ℓ_is are homogeneous linear polynomials over $\mathbb{F}[X]$ and $c_i \in \mathbb{F}$. Let $r < \varepsilon d$ be the rank of the linear forms $\{\ell_i \mid i \in [m]\}$ for a parameter $\varepsilon < 1$. If $g \equiv x_1 x_2 \cdots x_n$ then there exists a setting for the parameter ε such that $s = 2^{\Omega(n)}$.*

Proof. From Proposition 2.6,

$$\mathrm{MD}_k(g) \leq s \cdot \max_{i\in[s]}(\mathrm{MD}_k(f_i^{\alpha_i})).$$

From Lemma 3.1 and the fact that $g \equiv x_1 x_2 \cdots x_n$, we can infer that $\mathrm{MD}_k(g) = \binom{n}{k}$. By invoking Lemma 6.1 we can get an upper bound on $\max_{i\in[s]}(\mathrm{MD}_k(f_i^{\alpha_i}))$. Thus,

$$s \cdot k \cdot \left(\binom{k}{p} \cdot 2^{(p(d+r)-k)\cdot H\left(\frac{rp}{p(d+r)-k}\right)} + \frac{n}{2}\binom{n}{t}\right) \geq \binom{n}{k}$$

$$\implies s \geq \frac{\binom{n}{k}}{k \cdot \left(\binom{k}{p} \cdot 2^{(p(d+r)-k)\cdot H\left(\frac{rp}{p(d+r)-k}\right)} + \frac{n}{2}\binom{n}{t}\right)}$$

As before, let us fix the value of k to $0.5n$ so as to maximize the numerator and we will set the value of t to $0.545n$, a value which is away from $0.5n$ on the greater side so that $\binom{n}{t}$ is the non dominant term.

$$\begin{aligned}
&\log\left(\binom{k}{p}\cdot 2^{(p(d+r)-k)\cdot H\left(\frac{rp}{p(d+r)-k}\right)}\right)\\
&= p\log\left(\frac{ek}{p}\right) + (p(d+r)-k)\cdot H\left(\frac{rp}{p(d+r)-k}\right)\\
&\leq \frac{(t+k)}{d}\log\left(\frac{ek}{p}\right) + \frac{(t+k)(d+r)-kd}{d}\cdot H\left(\frac{\varepsilon(t+k)}{(d+r)(t+k)-kd}\right)\\
&= \frac{1.045n}{d}\log\left(\frac{ed}{1.045}\right) + (0.545+1.045\varepsilon)n\cdot H\left(\frac{1.045\varepsilon}{(0.545+1.045\varepsilon)}\right).
\end{aligned}$$

For a constant value of $d \geq 8$, the value of ε can be set to 0.1. Also, the value of ε can be raised to 0.4 when $d \geq 200$. For a reasonably large value of $d = \omega(1)$, the second summand dominates the first. For a value of $\varepsilon = 0.44001$, the second summand would compute to a value of $0.9996n$. This gives us a size lower bound of at least $2^{0.0004n}$. □

Again, by setting the value of d to $\sqrt{n}$, we get the following corollary.

Corollary 6.3. *For any integer n, the monomial $x_1x_2\cdots x_n$ can be computed by a circuit of the form $\Sigma\wedge\Sigma_{\mathrm{h}}^{\{=\sqrt{n}\}}\wedge^{[=\sqrt{n}]}\Sigma_{\mathrm{h}}$ of size $2^{O(\sqrt{n})}$ but any $\Sigma\wedge\Sigma^{\{\leq\varepsilon\sqrt{n}\}}\wedge^{[=\sqrt{n}]}\Sigma_{\mathrm{h}}$ computing it must be of size at least $2^{\Omega(n)}$.*

Acknowledgement

The authors would like to thank the organizers of NMI workshop on arithmetic circuit complexity 2017 for facilitating interactions between the authors that later led to this collaboration.

References

[AV08] Manindra Agrawal and V. Vinay. Arithmetic Circuits: A Chasm at Depth Four. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2008)*, pages 67–75, 2008. Pre-print available at `eccc:TR08-062`.

[BLSS17] Nikhil Balaji, Nutan Limaye, Sai Sandeep, and Srikanth Srinivasan. Personal communication, 2017.

[BS83] Walter Baur and Volker Strassen. The Complexity of Partial Derivatives. *Theoretical Computer Science*, 22:317–330, 1983.

[CKSV16] Suryajith Chillara, Mrinal Kumar, Ramprasad Saptharishi, and V. Vinay. The Chasm at Depth Four, and Tensor Rank : Old results, new insights. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:96, 2016.

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial Derivatives in Arithmetic Complexity. *Foundations and Trends in Theoretical Computer Science*, 2011.

[Ell69] W.J. Ellison. A 'Waring's Problem' for homogeneous forms. *Proceedings of the Cambridge Philosophical Society*, 65:663–672, 1969.

[Fis94] Ismor Fischer. Sums of like powers of multivariate linear forms. *Mathematics Magazine*, 67(1):59–61, 1994.

[FLMS14] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 128–135, 2014. Pre-print available at `eccc:TR13-100`.

[GK98] Dima Grigoriev and Marek Karpinski. An Exponential Lower Bound for Depth 3 Arithmetic Circuits. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC 1998)*, pages 577–582, 1998.

[GKKS13] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic Circuits: A Chasm at Depth Three. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2013)*, pages 578–587, 2013. Pre-print available at `eccc:TR13-026`.

[GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the Chasm at Depth Four. *Journal of the ACM*, 61(6):33:1–33:16, 2014. Preliminary version in the *28th Annual IEEE Conference on Computational Complexity (CCC 2013)*. Pre-print available at `eccc:TR12-098`.

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. In *Electronic Colloquium on Computational Complexity (ECCC)TR12-081*, 2012.

[KNS16] Neeraj Kayal, Vineet Nair, and Chandan Saha. Separation Between Read-once Oblivious Algebraic Branching Programs (ROABPs) and Multilinear Depth Three Circuits. In *Proceedings of Symposium on Theoretical Aspects of Computer Science, STACS*, pages 46:1–46:15, 2016.

[Koi12] Pascal Koiran. Arithmetic Circuits: The Chasm at Depth Four Gets Wider. *Theoretical Computer Science*, 448:56–65, 2012. Pre-print available at `arXiv:1006.4700`.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 146–153, 2014. Pre-print available at `eccc:TR13-091`.

[NW97] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1997. Available on `citeseer:10.1.1.90.2644`.

[RS00] Kristian Ranestad and Frank-Olaf Schreyer. Varieties of sums of powers. *JOURNAL FUR DIE REINE UND ANGEWANDTE MATHEMATIK*, pages 147–182, 2000.

[Sap15] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Github survey, 2015.

[Sax08] Nitin Saxena. Diagonal Circuit Identity Testing and Lower Bounds. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP 2008)*, pages 60–71, 2008. Pre-print available at `eccc:TR07-124`.

[Tav15] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Inf. Comput.*, 240:2–11, 2015. Preliminary version in the *38th Internationl Symposium on the Mathematical Foundations of Computer Science (MFCS 2013)*.

[Val79] Leslie G. Valiant. Completeness Classes in Algebra. In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC 1979)*, pages 249–261, 1979.