Electronic Colloquium on Computational Complexity, Report No. 15 (2015)

Multi-k-ic depth three circuit lower bound

Neeraj Kayal
Microsoft Research India
neeraka@microsoft.com

Chandan Saha
Indian Institute of Science
chandan@csa.iisc.ernet.in

Abstract

In a multi-k-ic depth three circuit every variable appears in at most k of the linear polynomials in every product gate of the circuit. This model is a natural generalization of multilinear depth three circuits that allows the formal degree of the circuit to exceed the number of underlying variables (as the formal degree of a multi-k-ic depth three circuit can be kn where n is the number of variables). The problem of proving lower bounds for depth three circuits with high formal degree has gained in importance following a work by Gupta, Kamath, Kayal and Saptharishi [GKKS13a] on depth reduction to high formal degree depth three circuits. In this work, we show an exponential lower bound for multi-k-ic depth three circuits for any arbitrary constant k.

1 Introduction

The recent years have witnessed some promising progress in arithmetic circuit lower bounds. A line of research attempts to better understand the prospect of proving super-polynomial arithmetic circuit lower bound by proving strong lower bounds for small depth circuits - thanks to the beautiful depth reduction results in these works [VSBR83, AV08, Koi12, Tav13]. A work by Gupta, Kamath, Kayal and Saptharishi [GKKS13a] showed that in order to separate VP from VNP, it is sufficient to prove a *strong-enough* lower bound for depth three circuits. The formal degree [1] of a depth three circuit can be much larger than the degree of the polynomial that it computes. This fact is exhibited in [GKKS13a]: quite interestingly, there is a depth three circuit with formal degree $n^{O(\sqrt{n})}$ (and also size $n^{O(\sqrt{n})}$) that computes Det_n, the determinant of an $n \times n$ symbolic matrix. Note that in this case the formal degree $n^{O(\sqrt{n})}$ is also much higher than the number of variables n^2. It follows from [GKKS13a] that if we are able to show an $n^{\omega(\sqrt{n})}$ size lower bound for depth three circuits of formal degree $n^{O(\sqrt{n})}$ computing the Perm_n (the permanent of an $n \times n$ symbolic matrix) then we would end up separating the circuit complexity of the determinant and the permanent polynomials (also proving $\mathsf{VP} \neq \mathsf{VNP}$).

1.1 Motivation and our result

The issue of large formal degree of a circuit, compared to the actual degree and the number of variables of the polynomial being computed, poses a challenge to the existing lower bound techniques in particular the complexity measures that have been used successfully to prove lower bounds for

[1] formal degree of a circuit C is the formal degree of its output gate. Formal degree of a $+$ gate is the maximum of the formal degrees of its children, whereas formal degree of a $\times$ gate is the sum of the formal degrees of its children.

ISSN 1433-8092

certain interesting models of circuits having low formal degree. The partial derivatives measure, the shifted partials and the closely related projected shifted partials, and the evaluation dimension are examples of such effective measures.

The partial derivatives measure was introduced and used by Nisan and Wigderson in an influential work [NW97] to prove an exponential lower bound for homogeneous [2] depth three circuits with formal degree less than the number of variables. A lower bound for depth three circuits with large formal degree will trivially imply a lower bound for homogeneous depth three circuits with large formal degree. This prompts us to pose the following problem,

Problem 1. *Over fields of characteristic zero, prove a super polynomial lower bound for homogeneous depth three circuits with formal degree $D = k \cdot n$, where k is an arbitrary constant and n is the number of variables.*[3]

In other words, can we prove a lower bound even if we allow the degree of the polynomial (being computed) to equal the formal degree of the depth three circuit that is only modestly higher than the number of variables? We do not know if the partial derivatives measure, or in fact any of the known measures and techniques, can be used to solve this problem. But, doing so might offer some insight into depth three circuits with large formal degree. We note that solving Problem 1 would automatically take us to the realm of *non-multilinear* polynomials.

Building on the partial derivatives measure, Kayal [Kay12] has introduced the shifted partials measure which has been used subsequently to prove an exponential lower bound for homogeneous depth four circuits [4] [KLSS14, KS14b] (albeit, using a variant of the shifted partials measure called the projected shifted partials) [5]. A recent work by Kayal and Saha [KS14a] uses the projected shifted partials measure to prove an exponential lower bound for depth three circuits with arbitrarily large formal degree but with somewhat low bottom fanin. It is not clear to us if the projected shifted partials can be used to solve Problem 1.

The evaluation dimension measure (defined later) has been used by Raz and Yehudayoff [RY09] to prove an exponential lower bound for multilinear [6] depth three circuits [7]. More precisely, they have shown a size lower bound of $2^{\Omega(d)}$ for any multilinear depth three circuit computing Det_d. Note that the formal degree of a multilinear depth three circuit is less or equal to the number of variables of the polynomial it computes. In the context of studying depth three circuits with large formal degree, a natural generalization of multilinear depth three circuits is the model of multi-k-ic depth three circuits (defined below) that allows the formal degree of the circuit to be higher than the number of variables.

[2] a circuit is homogeneous if every gate of the circuit computes a homogeneous polynomial (meaning, all monomials have the same degree)

[3] Over any fixed finite field, a solution to this problem already follows from the works of [GR98] and [GK98].

[4] with formal degree less than the number of variables

[5] [KLSS14] builds upon the works of [GKKS13b] and [KSS14].

[6] every variable occurs in at most one of the linear polynomials in every product gate of a multilinear depth three circuit

[7] in fact, their result is more general and applies to constant depth multilinear circuits. Also, their result builds on an earlier work by Raz [Raz09] who showed a quasi-polynomial lower bound for general multilinear formulas. Both [RY09] and [Raz09] use the rank of a partial derivatives matrix as a measure which can be shown to be the same as the evaluation dimension - a concept used in [FS13].

Definition 1. *A depth three circuit is* multi-k-ic *if every variable appears in at most k of the linear polynomials in every product gate of the circuit.*

For example, the expression $(x_1 + 2x_2)(4x_1 - x_3) + x_2^2 + (x_3 - x_2)(x_1 + x_2)$ is a multi-2-ic [8] depth three circuit. The formal degree of a multi-k-ic depth three circuit can be as high as $k \cdot n$, where n is the number of variables. A question, related to Problem 1, is the following: even if we allow the degree of the polynomial computed to equal the formal degree of the multi-k-ic circuit that computes it, can we prove a lower bound for this model?

Problem 2. *Prove an exponential lower bound for multi-k-ic depth three circuits for any arbitrary constant k.*

Could the evaluation dimension be useful in solving this problem[9]? In this work, we answer this question in the affirmative.

Theorem 1. *Let k be any arbitrary constant. There is a family of n-variate, degree $k \cdot n$ polynomials $\{f_n\}$ in* VNP *such that any multi-k-ic depth three circuit computing f_n must have size $2^{\Omega(n/2^{25k})}$.*

We will prove the above theorem in the rest of this article, but leave Problem 1 open. (We have not tried to optimize the constant 2^{25} in the above theorem.)

2 The measure - evaluation dimension

Let $f(x_1, \ldots, x_n)$ be a polynomial in $\mathbb{F}[x_1, \ldots, x_n]$, and $S = \{x_{i_1}, \ldots, x_{i_m}\}$ be a subset of the variables $\mathbf{x} = \{x_1, \ldots, x_n\}$. For a point $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{F}^m$, let $f_{S=\mathbf{a}} \in \mathbb{F}[\mathbf{x} \backslash S]$ denote the polynomial f evaluated at $x_{i_j} = a_j$ for every $j \in [m]$. Let $\mathsf{eval}_S(f)$ be the $\mathbb{F}$-linear space spanned by the polynomials $\{f_{S=\mathbf{a}}\}_{\mathbf{a} \in \mathbb{F}^m}$, i.e.

$$\mathsf{eval}_S(f) = \mathbb{F}\text{-span}\left(\{f_{S=\mathbf{a}} : \mathbf{a} \in \mathbb{F}^m\}\right).$$

Definition 2. *Evaluation dimension of a polynomial f with respect to a subset of variables S is defined as the dimension of the vector space $\mathsf{eval}_S(f)$. It is denoted by $\mathsf{evalDim}_S(f)$.*

Let us state a couple of useful properties of the evaluation dimension.

Lemma 2. *Let f and g be two polynomials in $\mathbb{F}[\mathbf{x}]$ and $S \subseteq \mathbf{x}$. Then*

1. (subadditivity) $\mathsf{evalDim}_S(f + g) \leq \mathsf{evalDim}_S(f) + \mathsf{evalDim}_S(g)$
2. (submultiplicativity) $\mathsf{evalDim}_S(f \cdot g) \leq \mathsf{evalDim}_S(f) \cdot \mathsf{evalDim}_S(g)$

Proof. The subadditivity property follows from the observation that every polynomial in the space $\mathsf{eval}_S(f + g)$ is a sum of a polynomial in $\mathsf{eval}_S(f)$ and a polynomial in $\mathsf{eval}_S(g)$. Now suppose the polynomials $f_1, \ldots, f_p$ form a basis of the space $\mathsf{eval}_S(f)$ and similarly $g_1, \ldots, g_q$ form a basis of $\mathsf{eval}_S(g)$. Then every polynomial in the space $\mathsf{eval}_S(f \cdot g)$ can be expressed as an $\mathbb{F}$-linear combination of polynomials $f_i g_j$ with $i \in [p]$ and $j \in [q]$. This shows the submultiplicativity property. □

[8] 'multiquadratic' sounds better here

[9] The works of Grenet, Koiran, Portier, and Strozecki [GKPS11] and of Agrawal, Saha, Saptharishi and Saxena [ASSS12] proved lower bounds for certain models of depth four circuits with high formal degree, using properties of the real-τ-conjecture and the Jacobian respectively. The top fanin of such depth four circuits is essentially low or can be assumed to be low without loss of generality - a feature that is crucially used in their proofs. We do not know if their techniques can be used to solve Problem 2.

3 An explicit polynomial with high evaluation dimension

Let g be a polynomial in $6n$ variables $\mathbf{u} = \{u_1, \ldots, u_{4n}\}$ and $\mathbf{x} = \{x_1, \ldots, x_{2n}\}$, and $k \in \mathbb{Z}^+$ be an arbitrary positive integer. To every set $A \subseteq [2n]$, associate a set B_A in the following manner:

- If $|A| \geq n$ then B_A is a fixed subset of A of size exactly equal to $\bar{A} = [2n] \backslash A$.
- If $|A| < n$ then B_A is a fixed subset of $\bar{A}$ of size exactly equal to A.

One way of fixing B_A is to take lexicographically the smallest subset. For a set $A \subseteq [2n]$ and $\mathbf{e} = \{e_1, \ldots, e_{|A|}\} \in \mathbb{Z}^{|A|}$, let $x_A^{\mathbf{e}} \stackrel{\text{def}}{=} \prod_{i \in A} x_i^{e_i}$ and $u_A \stackrel{\text{def}}{=} \prod_{i \in A} u_i$. Let $\bar{A} + 2n$ denote the set $\{i + 2n : i \in \bar{A}\}$, and $\bar{u}_{\bar{A}+2n} \stackrel{\text{def}}{=} \prod_{i \in \bar{A}} (1 - u_{i+2n})$. Define the polynomial $f_A(\mathbf{x})$ as follows.

$$f_A(\mathbf{x}) = \begin{cases} \sum_{\mathbf{e} \in \{0,\ldots,k\}^{|\bar{A}|}} x_{B_A}^{\mathbf{e}} \cdot x_{\bar{A}}^{\mathbf{e}} & \text{if } |A| \geq n \\ \sum_{\mathbf{e} \in \{0,\ldots,k\}^{|A|}} x_A^{\mathbf{e}} \cdot x_{B_A}^{\mathbf{e}} & \text{if } |A| < n \end{cases}$$

Define g as,

$$g = \sum_{A \subseteq [2n]} u_A \cdot \bar{u}_{\bar{A}+2n} \cdot f_A(\mathbf{x}). \tag{1}$$

Polynomial g satisfies the following property.

Lemma 3. *For every $A \subseteq [2n]$, there is an assignment of the $\mathbf{u}$ variables to field constants such that $\mathsf{evalDim}_{\mathbf{x}_A}(g)$, where $\mathbf{x}_A = \{x_i : i \in A\}$, (after setting the $\mathbf{u}$ variables) is $(k+1)^{\min(|A|,|\bar{A}|)}$.*

Proof. Let $A \subseteq [2n]$. Consider this assignment of the $\mathbf{u}$ variables for every $i \in [2n]$: $u_i = u_{i+2n} = 1$ if $i \in A$ and $u_i = u_{i+2n} = 0$ if $i \in \bar{A}$. Denote the polynomial g under this assignment by $g_{\mathbf{u}_A = \mathbf{1}}$, which equals $f_A(\mathbf{x})$. Hence,

$$\mathsf{evalDim}_{\mathbf{x}_A}(g_{\mathbf{u}_A = \mathbf{1}}) = \mathsf{evalDim}_{\mathbf{x}_A}(f_A).$$

Now, it is not difficult to see that the evaluation dimension of f_A with respect to $\mathbf{x}_A$ equals $(k+1)^{|A|}$ (respectively, $(k+1)^{|\bar{A}|}$) if $|A| < n$ (respectively, $|A| \geq n$). □

We also note that g defines a polynomial family in VNP, as the coefficient of a given monomial can be computed efficiently. The construction of g is inspired by a similar construction in an earlier work of Raz [Raz10].

Picking a random $\mathbf{x}_A$. Suppose we form a set A by picking every $i \in [2n]$ independently at random with probability $\frac{1}{2}$. By Chernoff bound, $|A| \in [(1-\delta)n, (1+\delta)n]$ with probability at least $1 - e^{-n\delta^2/3}$ for any $\delta > 0$. We will study the evaluation dimension of g and the multi-k-ic depth three circuit that computes it with respect to such a random $\mathbf{x}_A = \{x_i : i \in A\}$ after assigning field values to the $\mathbf{u}$-variables. The parameter δ will be a fixed function of k (to be specified later in Section 6).

Corollary 4. *By Lemma 3, if A is chosen randomly (as described above) then $\mathsf{evalDim}_{\mathbf{x}_A}(g_{\mathbf{u}_A = \mathbf{1}})$ is at least $(k+1)^{(1-\delta)n}$ with probability higher than $1 - e^{-n\delta^2/3}$, for any $\delta > 0$.*

The above corollary provides a lower bound on the evaluation dimension of g. We will now show an upper bound on the evaluation dimension of a multi-k-ic depth three circuit with respect to a random $\mathbf{x}_A$. This, together with Corollary 4, will give us the relevant lower bound as outlined below. In the rest of this article whenever we write A is 'random' we mean A is formed by picking every $i \in [2n]$ independently at random with probability $\frac{1}{2}$.

4 Proof outline

Let $C = \sum_{i=1}^{s} T^{(i)}$ be a multi-k-ic depth three circuit computing g (as defined in Equation 1), where every $T^{(i)}$ is a product of linear polynomials. We will refer to $T^{(i)}$ as a *product term* (or simply a *term*) of C. Since C is multi-k-ic, every variable appears in at most k linear polynomials in every $T^{(i)}$. Let $A \subseteq [2n]$ be a random set and $\mathbf{x}_A = \{x_i : i \in A\}$ be the corresponding subset of $\mathbf{x}$. For any polynomial $h(\mathbf{x}, \mathbf{u})$, denote by $h_{\mathbf{u}_A=\mathbf{1}}$ the polynomial h with $u_i = 1$ if $i \in A$ and $u_i = 0$ if $i \notin A$. Note that $h_{\mathbf{u}_A=\mathbf{1}}$ is a polynomial in only the $\mathbf{x}$-variables.

$$\begin{aligned} g &= C = \sum_{i=1}^{s} T^{(i)} \\ \Rightarrow g_{\mathbf{u}_A=\mathbf{1}} &= C_{\mathbf{u}_A=\mathbf{1}} = \sum_{i=1}^{s} T^{(i)}_{\mathbf{u}_A=\mathbf{1}} \\ \Rightarrow \mathsf{evalDim}_{\mathbf{x}_A}(g_{\mathbf{u}_A=\mathbf{1}}) &\leq \sum_{i=1}^{s} \mathsf{evalDim}_{\mathbf{x}_A}(T^{(i)}_{\mathbf{u}_A=\mathbf{1}}), \end{aligned}$$

where the last inequality follows from the subadditive property of the evaluation dimension (Lemma 2). Now, suppose we are able to show that $\mathsf{evalDim}_{\mathbf{x}_A}(T^{(i)}_{\mathbf{u}_A=\mathbf{1}})$ is upper bounded by a quantity $U(k, n, \delta)$ for every $i \in [s]$ with high probability over the random choice of A. Then by applying union bound,

$$\mathsf{evalDim}_{\mathbf{x}_A}(g_{\mathbf{u}_A=\mathbf{1}}) \leq s \cdot U(k, n, \delta),$$

also with high probability. In other words, by the above observation and Corollary 4, there exists a choice of A such that

$$\begin{aligned} (k+1)^{(1-\delta)n} &\leq \mathsf{evalDim}_{\mathbf{x}_A}(g_{\mathbf{u}_A=\mathbf{1}}) \leq s \cdot U(k, n, \delta) \\ \Rightarrow s &\geq \frac{(k+1)^{(1-\delta)n}}{U(k, n, \delta)}. \end{aligned}$$

This will give us a lower bound on the top fanin of C. We are now left with the task of finding a suitable expression for $U(k, n, \delta)$, which we do in the following section.

5 Evaluation dimension of a term of a multi-k-ic depth-3 circuit

Notations. Let us focus on a product term $T^{(i)} = T$ (say). Let $T = \prod_{j=1}^{d} \ell_j$, where ℓ_j is a linear polynomial and c be a positive integer constant (to be fixed later in Section 6). Split the linear

polynomials in T into three parts:

$$\begin{aligned} P^{(1)} &:= \prod_{j\in[d]} \ell_j \text{ such that } \ell_j \text{ has exactly one or no } \mathbf{x}\text{-variables} \\ P^{(2)} &:= \prod_{j\in[d]} \ell_j \text{ such that the number of } \mathbf{x}\text{-variables in } \ell_j \text{ is between two and } ck \\ P^{(3)} &:= \prod_{j\in[d]} \ell_j \text{ such that } \ell_j \text{ has greater than } ck \text{ } \mathbf{x}\text{-variables.} \end{aligned}$$

Also let,

$$m_i := \text{ the number of linear polynomials in } T \text{ with exactly } i \text{ } \mathbf{x}\text{-variables.}$$

Naturally, $T = P^{(1)} \cdot P^{(2)} \cdot P^{(3)}$. Also, the number of linear polynomials in $P^{(1)}$ is $m_0 + m_1$, the number of linear polynomials in $P^{(2)}$ equals $\sum_{i=2}^{ck} m_i$, and the number of linear polynomials in $P^{(3)}$ equals $\sum_{i>ck} m_i$.

Claim 5. *For any $A \subseteq [2n]$, $\mathsf{evalDim}_{\mathbf{x}_A}(T_{\mathbf{u}_A=\mathbf{1}}) \leq \mathsf{evalDim}_{\mathbf{x}_A}(P^{(2)}_{\mathbf{u}_A=\mathbf{1}}) \cdot \mathsf{evalDim}_{\mathbf{x}_A}(P^{(3)}_{\mathbf{u}_A=\mathbf{1}})$*

Proof. By the submultiplicativity property of evaluation dimension (Lemma 2),

$$\mathsf{evalDim}_{\mathbf{x}_A}(T_{\mathbf{u}_A=\mathbf{1}}) \leq \mathsf{evalDim}_{\mathbf{x}_A}(P^{(1)}_{\mathbf{u}_A=\mathbf{1}}) \cdot \mathsf{evalDim}_{\mathbf{x}_A}(P^{(2)}_{\mathbf{u}_A=\mathbf{1}}) \cdot \mathsf{evalDim}_{\mathbf{x}_A}(P^{(3)}_{\mathbf{u}_A=\mathbf{1}}).$$

Moreover, it is easy to see that $\mathsf{evalDim}_{\mathbf{x}_A}(P^{(1)}_{\mathbf{u}_A=\mathbf{1}}) = 1$. □

We will upper bound the evaluation dimension of $P^{(3)}_{\mathbf{u}_A=\mathbf{1}}$ with respect to $\mathbf{x}_A$ for any A, and the evaluation dimension of $P^{(2)}_{\mathbf{u}_A=\mathbf{1}}$ with respect to $\mathbf{x}_A$ for a random A. Let r_2 be the number of occurrences of the $\mathbf{x}$-variables among the linear polynomials in $P^{(2)}$ and r_3 be the number of occurrences of the $\mathbf{x}$-variables in $P^{(3)}$. Since every variable occurs in at most k linear polynomials in T and there are $2n$ $\mathbf{x}$-variables,

$$r_2 + r_3 \leq 2kn. \tag{2}$$

5.1 Evaluation dimension of $P^{(3)}$

Lemma 6. *For any $A \subseteq [2n]$, $\mathsf{evalDim}_{\mathbf{x}_A}(P^{(3)}_{\mathbf{u}_A=\mathbf{1}}) \leq 2^{\frac{r_3}{ck}}$.*

Proof. The evaluation dimension of $P^{(3)}_{\mathbf{u}_A=\mathbf{1}}$ with respect to the $\mathbf{x}_A$-variables cannot exceed 2^b, where b is the number of linear polynomials in $P^{(3)}$. Observe that the degree of $P^{(3)}$ with respect to the $\mathbf{x}$-variables is less than $\frac{r_3}{ck}$, as every linear polynomial in the product $P^{(3)}$ has more than ck $\mathbf{x}$-variables. □

5.2 Evaluation dimension of $P^{(2)}$

Coloring of linear polynomials. Every linear polynomial in the product $P^{(2)} = P$ (say) has more than one and less than or equal to ck $\mathbf{x}$-variables. We *color* the linear polynomials in P in such a way that no two linear polynomials with the same color have a common $\mathbf{x}$-variable. This coloring can be done greedily using at most $(k-1)ck + 1 \leq ck^2$ colors. Let the number of colors

used be q; we will identify these colors with $\{1, \ldots, q\}$. Now we can split the product P into at most $q \leq ck^2$ parts (one per color), say $Q^{(1)}, \ldots, Q^{(q)}$, such that every $Q^{(j)}$ is a product of linear polynomials in P that are colored j. This also implies that $Q^{(j)}$ is *multilinear* in the $\mathbf{x}$-variables. Naturally,

$$P = \prod_{j=1}^{q} Q^{(j)}.$$

To understand the evaluation dimension of P, we will focus on the polynomials $Q^{(j)}$.

Some more notations and bounds. Let $m_{i,j}$ be the number of linear polynomials in $Q^{(j)}$ with exactly i many $\mathbf{x}$-variables. Hence, $m_i = \sum_{j \in [q]} m_{i,j}$ for every integer $i \in [2, ck]$. Let A be a random subset of $[2n]$ (in the sense described in Section 3). Let $r_{i,j}$ be the number of linear polynomials in $Q^{(j)}$ with strictly more than i $\mathbf{x}$-variables and exactly i $\mathbf{x}_A$-variables. Note that only such linear polynomials with at least one $\mathbf{x}_A$-variable, but not all $\mathbf{x}$-variables are $\mathbf{x}_A$-variables, contribute to the evaluation dimension of P with respect to $\mathbf{x}_A$. We will refer to such linear polynomials as *partially touched* (by A) linear polynomials. The expected value of $r_{i,j}$ over the random choice of A is

$$\begin{aligned}
\mathcal{E}[r_{i,j}] &= \sum_{\ell=i+1}^{ck} \binom{\ell}{i} \cdot \frac{1}{2^\ell} \cdot m_{\ell,j} \\
&\geq \frac{i+1}{2^{ck}} \cdot \sum_{\ell=i+1}^{ck} m_{\ell,j} \\
&\geq \frac{1}{2^{ck-1}} \cdot \sum_{\ell=i+1}^{ck} m_{\ell,j} \quad (\text{as } i \geq 1). \qquad (3)
\end{aligned}$$

The above expression for the expectation can be derived from the observation that a linear polynomial with ℓ $\mathbf{x}$-variables ($\ell > i$) has exactly i $\mathbf{x}_A$-variables with probability $\binom{\ell}{i} \cdot \frac{1}{2^\ell}$. We will see how $r_{i,j}$ contributes to the evaluation dimension of P later. But, first, in order to get a handle on the value of $r_{i,j}$ we would like to argue that it is close to its expected value with high probability. Since $Q^{(j)}$ is multilinear, if $\mathcal{E}[r_{i,j}]$ is sufficiently large, we can apply Chernoff bound on $r_{i,j}$ and show that $(1-\delta)\mathcal{E}[r_{i,j}] \leq r_{i,j} \leq (1+\delta)\mathcal{E}[r_{i,j}]$ with high probability. By Equation 3, expectation of $r_{i,j}$ is large if $\sum_{\ell=i+1}^{ck} m_{\ell,j}$ is large. This motivates us to split $Q^{(j)}$ further depending on the value of $\sum_{\ell=i+1}^{ck} m_{\ell,j}$.

Splitting $Q^{(j)}$ further. Let τ_j be the maximum number less than ck such that

$$\sum_{\ell=\tau_j+1}^{ck} m_{\ell,j} \geq \frac{n}{ck^2 \cdot \Delta}, \qquad (4)$$

where $\Delta = \Delta(k)$ is a sufficiently large constant, dependent on k, to be fixed later in Section 6. Let $Q'^{(j)}$ be the product of those linear polynomials in $Q^{(j)}$ that contribute to $r_{i,j}$ for $i > \tau_j$, and $\tilde{Q}^{(j)}$

the product of those linear polynomials in $Q^{(j)}$ that contribute to $r_{i,j}$ for $i \in [1, \tau_j]$. By Equation 4,

$$\sum_{i=\tau_j+1}^{ck-1} r_{i,j} \leq \sum_{i=\tau_j+2}^{ck} m_{i,j} < \frac{n}{ck^2 \cdot \Delta}. \tag{5}$$

Let $P' = \prod_{j=1}^{q} Q'^{(j)}$ and $\tilde{P} = \prod_{j=1}^{q} \tilde{Q}^{(j)}$. Then,

$$\mathsf{evalDim}_{\mathbf{x}_A}(P_{\mathbf{u}_A=\mathbf{1}}) \leq \mathsf{evalDim}_{\mathbf{x}_A}(\tilde{P}_{\mathbf{u}_A=\mathbf{1}}) \cdot \mathsf{evalDim}_{\mathbf{x}_A}(P'_{\mathbf{u}_A=\mathbf{1}}),$$

as a linear polynomial contributes to the evaluation dimension of P only if it is partially touched (by A). By Equation 5, the number of linear polynomials in P' is upper bounded by

$$\sum_{j=1}^{q} \sum_{i=\tau_j+1}^{ck-1} r_{i,j} \leq \frac{n}{ck^2 \cdot \Delta} \cdot q \leq \frac{n}{\Delta} \quad (\text{as } q \leq ck^2).$$

Hence,

$$\mathsf{evalDim}_{\mathbf{x}_A}(P'_{\mathbf{u}_A=\mathbf{1}}) \leq 2^{\frac{n}{\Delta}}, \tag{6}$$

as the evaluation dimension cannot exceed 2^b, where b is the number of linear polynomials in P'. By choosing a large enough Δ in the analysis later, we will ensure that $\mathsf{evalDim}_{\mathbf{x}_A}(P'_{\mathbf{u}_A=\mathbf{1}})$ is negligible compared to other relevant terms.

Computing evaluation dimension of $\tilde{P}$. Since $\tilde{Q}^{(j)}$ is a product of those linear polynomials that contribute to $r_{i,j}$ for $i \in [1, \tau_j]$, by Equations 3 and 4,

$$\mathcal{E}[r_{i,j}] \geq \frac{1}{2^{ck-1}} \cdot \frac{n}{ck^2 \cdot \Delta},$$

for every $i \in [1, \tau_j]$. For any fixed $j \in [q]$, $Q^{(j)}$ is multilinear. Hence, by applying Chernoff bound,

$$\Pr\{|r_{i,j} - \mathcal{E}[r_{i,j}]| > \delta \cdot \mathcal{E}[r_{i,j}]\} \leq e^{-\frac{\delta^2 \cdot \mathcal{E}[r_{i,j}]}{3}} \leq e^{-\frac{\delta^2 \cdot n}{3 \cdot 2^{ck-1} ck^2 \Delta}}.$$

By union bound, $\Pr\{|r_{i,j} - \mathcal{E}[r_{i,j}]| > \delta \cdot \mathcal{E}[r_{i,j}]\}$ for any $j \in [q]$ and $i \in [1, \tau_j]$, is bounded by,

$$\varepsilon_1 := ck^2 \cdot ck \cdot e^{-\frac{\delta^2 \cdot n}{3 \cdot 2^{ck-1} ck^2 \Delta}}. \tag{7}$$

As n is much larger compared to the constants k, c, δ, Δ, the above 'error probability' ε_1 is negligible. Hence, with probability at least $1 - \varepsilon_1$,

$$(1 - \delta) \cdot \mathcal{E}[r_{i,j}] \leq r_{i,j} \leq (1 + \delta) \cdot \mathcal{E}[r_{i,j}] \tag{8}$$

for every $j \in [q], i \in [1, \tau_j]$.

Let r_i be the number of linear polynomials in $\tilde{P}$ with more that i $\mathbf{x}$-variables and exactly i $\mathbf{x}_A$-variables. Then,

$$\begin{aligned} r_i &= \sum_{j \in [q]: i \in [1, \tau_j]} r_{i,j} \\ \mathcal{E}[r_i] &= \sum_{j \in [q]: i \in [1, \tau_j]} \mathcal{E}[r_{i,j}]. \end{aligned}$$

The notation $j \in [q] : i \in [1, \tau_j]$ means the sum is over those $j \in [q]$ for which $i \in [1, \tau_j]$. By Equation 8,

$$(1-\delta)\mathcal{E}[r_i] \leq r_i \leq (1+\delta)\mathcal{E}[r_i]$$

with probability at least $1 - \varepsilon_1$. This implies

$$\begin{aligned}
r_i &\leq (1+\delta) \cdot \sum_{j \in [q]: i \in [1,\tau_j]} \mathcal{E}[r_{i,j}] \\
&= (1+\delta) \cdot \sum_{j \in [q]: i \in [1,\tau_j]} \sum_{\ell=i+1}^{ck} \binom{\ell}{i} \cdot \frac{1}{2^\ell} \cdot m_{\ell,j} \quad \text{(by Equation 3)} \\
&= (1+\delta) \cdot \sum_{\ell=i+1}^{ck} \binom{\ell}{i} \cdot \frac{1}{2^\ell} \cdot \sum_{j \in [q]: i \in [1,\tau_j]} m_{\ell,j} \\
&\leq (1+\delta) \cdot \sum_{\ell=i+1}^{ck} \binom{\ell}{i} \cdot \frac{1}{2^\ell} \cdot \sum_{j \in [q]} m_{\ell,j} \\
&= (1+\delta) \cdot \sum_{\ell=i+1}^{ck} \binom{\ell}{i} \cdot \frac{1}{2^\ell} \cdot m_\ell.
\end{aligned}$$

Let e_x be the number of occurrences of a variable $x \in \mathbf{x}_A$ in the linear polynomials in $\tilde{P}$. Then, by the above equation, with probability at least $1 - \varepsilon_1$,

$$\begin{aligned}
\sum_{x \in \mathbf{x}_A} e_x &= \sum_{i=1}^{ck-1} i \cdot r_i \\
&\leq (1+\delta) \cdot \sum_{i=1}^{ck-1} i \cdot \sum_{\ell=i+1}^{ck} \binom{\ell}{i} \cdot \frac{1}{2^\ell} \cdot m_\ell \\
&= (1+\delta) \cdot \sum_{i=1}^{ck-1} \sum_{\ell=i+1}^{ck} \binom{\ell-1}{i-1} \cdot \frac{1}{2^\ell} \cdot \ell \cdot m_\ell \\
&\leq (1+\delta) \cdot \sum_{\ell=2}^{ck} \sum_{i=1}^{\ell-1} \binom{\ell-1}{i-1} \cdot \frac{1}{2^\ell} \cdot \ell \cdot m_\ell \\
&= (1+\delta) \cdot \sum_{\ell=2}^{ck} (2^{\ell-1} - 1) \cdot \frac{1}{2^\ell} \cdot \ell \cdot m_\ell \\
&= (1+\delta) \cdot \sum_{\ell=2}^{ck} \left(1 - \frac{1}{2^{\ell-1}}\right) \cdot \frac{1}{2} \cdot \ell \cdot m_\ell \\
&\leq (1+\delta) \cdot \left(1 - \frac{1}{2^{ck-1}}\right) \cdot \frac{1}{2} \cdot \sum_{\ell=2}^{ck} \ell \cdot m_\ell.
\end{aligned}$$

Observe that $\sum_{\ell=2}^{ck} \ell \cdot m_\ell$ is the number of occurrences of the $\mathbf{x}$-variables in P. Hence, $\sum_{\ell=2}^{ck} \ell \cdot m_\ell =$

r_2 and so with probability at least $1-\varepsilon_1$,

$$\sum_{x\in\mathbf{x}_A} e_x \leq (1+\delta)\cdot\left(1-\frac{1}{2^{ck-1}}\right)\cdot\frac{r_2}{2}. \tag{9}$$

Let $\varepsilon_0 = e^{-\frac{\delta^2 n}{3}}$.

Lemma 7. *With probability at least $1-(\varepsilon_0+\varepsilon_1)$ over the random choice of A,*

$$\mathsf{evalDim}_{\mathbf{x}_A}(\tilde{P}_{\mathbf{u}_A=\mathbf{1}}) \leq \left[\left(1-\frac{1}{2^{ck-1}}\right)\cdot\frac{r_2}{2n}+1\right]^{(1+\delta)\cdot n}.$$

Proof. Since A is chosen randomly by picking every $i\in[2n]$ independently at random with probability $\frac{1}{2}$, $|\mathbf{x}_A|\leq(1+\delta)\cdot n$ with probability at least $1-\varepsilon_0$. The evaluation dimension of $\tilde{P}$ with respect to $\mathbf{x}_A$ cannot exceed the number of distinct $\mathbf{x}_A$-monomials in $\tilde{P}$ with coefficients from $\mathbb{F}[\mathbf{x}\backslash\mathbf{x}_A]$. The number of such monomials is upper bounded by $\prod_{x\in\mathbf{x}_A}(e_x+1)$. By AM-GM inequality,

$$\begin{aligned}\prod_{x\in\mathbf{x}_A}(e_x+1) &\leq \left[\frac{\sum_{x\in\mathbf{x}_A}(e_x+1)}{|\mathbf{x}_A|}\right]^{|\mathbf{x}_A|}\\ &\leq \left[\frac{(1+\delta)(1-\frac{1}{2^{ck-1}})\cdot\frac{r_2}{2}}{|\mathbf{x}_A|}+1\right]^{|\mathbf{x}_A|}, \quad \text{(by Equation 9)}\end{aligned}$$

with probability at least $1-\varepsilon_1$. Hence, with probability at least $1-(\varepsilon_0+\varepsilon_1)$,

$$\prod_{x\in\mathbf{x}_A}(e_x+1) \leq \left[\left(1-\frac{1}{2^{ck-1}}\right)\cdot\frac{r_2}{2n}+1\right]^{(1+\delta)n},$$

as the above expression increases with $|\mathbf{x}_A|$ and $|\mathbf{x}_A|\in[(1-\delta)n,(1+\delta)n]$ with probability at least $1-\varepsilon_0$. □

Corollary 8. *With probability at least $1-(\varepsilon_0+\varepsilon_1)$ over the random choice of A,*

$$\mathsf{evalDim}_{\mathbf{x}_A}(P^{(2)}_{\mathbf{u}_A=\mathbf{1}}) \leq \left[\left(1-\frac{1}{2^{ck-1}}\right)\cdot\frac{r_2}{2n}+1\right]^{(1+\delta)n}\cdot 2^{\frac{n}{\Delta}}.$$

Proof. Follows from the above lemma and Equation 6. □

5.3 Evaluation dimension of a term

Let T be a product term in a multi-k-ic depth three circuit.

Lemma 9. *With probability at least $1-(\varepsilon_0+\varepsilon_1)$ over the random choice of A,*

$$\mathsf{evalDim}_{\mathbf{x}_A}(T_{\mathbf{u}_A=\mathbf{1}}) \leq \left[\left(1-\frac{1}{2^{2ck}}\right)(k+1)\right]^n\cdot(k+1)^{\delta n},$$

if $c\geq 3, k\geq 4$ and $\Delta = 2^{2ck}$.

Proof. By Claim 5, Lemma 6 and Corollary 8,

$$\mathsf{evalDim}_{\mathbf{x}_A}(T_{\mathbf{u}_A=\mathbf{1}}) \leq \left[\left(1-\frac{1}{2^{ck-1}}\right)\cdot\frac{r_2}{2n}+1\right]^{(1+\delta)n}\cdot 2^{\frac{n}{\Delta}}\cdot 2^{\frac{r_3}{ck}}.$$

Recall from Equation 2, $r_2+r_3 \leq 2kn$. Let $r_2 \leq \alpha\cdot 2kn$ and $r_3 \leq (1-\alpha)\cdot 2kn$ where $0\leq\alpha\leq 1$. Then,

$$\mathsf{evalDim}_{\mathbf{x}_A}(T_{\mathbf{u}_A=\mathbf{1}}) \leq \left[\left(1-\frac{1}{2^{ck-1}}\right)\cdot\alpha k+1\right]^{n}\cdot 2^{\frac{n}{\Delta}}\cdot 2^{\frac{2(1-\alpha)n}{c}}\cdot\left[\left(1-\frac{1}{2^{ck-1}}\right)k+1\right]^{\delta n}.$$

Since $2^{\frac{1}{y}} \leq 1+\frac{1}{y}$ for every $y\geq 1$,

$$\left[\left(1-\frac{1}{2^{ck-1}}\right)\cdot\alpha k+1\right]\cdot 2^{\frac{1}{\Delta}}\cdot 2^{\frac{2(1-\alpha)}{c}} \leq \left[\left(1-\frac{1}{2^{ck-1}}\right)\alpha k+1\right]\cdot\left(1+\frac{1}{\Delta}\right)\cdot\left(1+\frac{2(1-\alpha)}{c}\right),$$

as $\Delta\geq 1$ and $c\geq 3$. The quantity $\left[\left(1-\frac{1}{2^{ck-1}}\right)\alpha k+1\right]\cdot\left(1+\frac{2(1-\alpha)}{c}\right)$ when treated as a function of $\alpha\in[0,1]$ is maximized at $\alpha=1$, assuming $c\geq 3, k\geq 4$. Therefore,

$$\begin{aligned}\left[\left(1-\frac{1}{2^{ck-1}}\right)\cdot\alpha k+1\right]\cdot 2^{\frac{1}{\Delta}}\cdot 2^{\frac{2(1-\alpha)}{c}} &\leq \left[\left(1-\frac{1}{2^{ck-1}}\right)k+1\right]\cdot\left(1+\frac{1}{\Delta}\right)\\ &\leq \left(1-\frac{1}{2^{2ck}}\right)\cdot(k+1) \qquad (\text{as } \Delta=2^{2ck}).\end{aligned}$$

This proves the lemma as $[(1-\frac{1}{2^{ck-1}})k+1]^{\delta n}\leq(k+1)^{\delta n}$. □

6 Proof of Theorem 1

Following the setting of parameters in Lemma 9, let $c=3,\Delta=2^{6k}$ and without loss of generality $k\geq 4$. Also, let

$$\delta=\frac{\ln(1+\frac{1}{2^{2ck+1}})}{2\cdot\ln(k+1)}=\frac{\ln(1+\frac{1}{2^{6k+1}})}{2\cdot\ln(k+1)},$$

and denote the upper bound in Lemma 9 by $U(k,n,\delta)$.

Lemma 10. *If $g(\mathbf{x},\mathbf{u})$, as defined in Equation 1, is computed by a multi-k-ic depth three circuit C then the top fanin s of C is at least $2^{\Omega\left(\frac{n}{2^{25k}}\right)}$.*

Proof. By union bound, with probability at least $1-(\varepsilon_0+s\cdot\varepsilon_1)$ over the random choice of A, the evaluation dimension of every term in C is upper bounded by $U(k,n,\delta)$. By Equation 7,

$$\varepsilon_1 := ck^2\cdot ck\cdot e^{-\frac{\delta^2\cdot n}{3\cdot 2^{ck-1}ck^2\Delta}}.$$

So, if $s\leq e^{\frac{\delta^2 n}{9\cdot 2^{3k}\cdot k^2\cdot\Delta}}$ then there exists an A such that evaluation dimension of every term of C is upper bounded by $U(k,n,\delta)$ (assuming n is sufficiently larger than k). Otherwise,

$$s > e^{\frac{\delta^2 n}{9\cdot 2^{3k}\cdot k^2\cdot\Delta}}=2^{\Omega\left(\frac{n}{2^{25k}}\right)}$$

and we already have the lower bound. If evaluation dimension of every term is upper bounded by $U(k,n,\delta)$ then following the discussion in Section 4,

$$\begin{aligned} s &\geq \frac{(k+1)^{(1-\delta)n}}{U(k,n,\delta)} \\ &= \left(1-\frac{1}{2^{2ck}}\right)^{-n} \cdot (k+1)^{-2\delta n} = 2^{\Omega\left(\frac{n}{2^{6k}}\right)}, \end{aligned}$$

after plugging in the value of δ from above. □

The proof of Theorem 1 is immediate from the above lemma.

7 Discussion

In order to gain a better understanding of the strengths and limitations of the existing complexity measures, like partial derivatives, (projected) shifted partials, evaluation dimension etc., it is perhaps worth exploring some natural and interesting models of circuits for which we still do not know of any super-polynomial lower bound. Such a model of circuits emerging from our work is *multi-k-ic* formulas: Let x be a variable and g be a gate. The formal degree of x at g, denoted $\deg_x(g)$, is defined as follows. If g is a $\times$-gate with children g_1 and g_2 then $\deg_x(g) = \deg_x(g_1) + \deg_x(g_2)$. If g is a $+$-gate with children g_1 and g_2 then $\deg_x(g) = \max\{\deg_x(g_1), \deg_x(g_2)\}$. If g is an input gate labelled with x then $\deg_x(g) = 1$, otherwise $\deg_x(g) = 0$. A formula is multi-k-ic if for every variable x and every gate g, the formal degree of x at g is bounded by k.

- Can we prove super-polynomial lower bounds for constant depth multi-k-ic formulas?
- Can we prove super-polynomial lower bounds for multi-k-ic formulas?

Acknowledgement

We would like to thank the anonymous reviewers for their helpful comments that have improved the presentation of this paper.

References

[ASSS12] Manindra Agrawal, Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. Jacobian hits circuits: hitting-sets, lower bounds for depth-d occur-k formulas & depth-3 transcendence degree-k circuits. In *STOC*, pages 599–614, 2012.

[AV08] Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *FOCS*, pages 67–75, 2008.

[FS13] Michael A. Forbes and Amir Shpilka. Quasipolynomial-Time Identity Testing of Noncommutative and Read-Once Oblivious Algebraic Branching Programs. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 243–252, 2013.

[GK98] Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *STOC*, pages 577–582, 1998.

[GKKS13a] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *Foundations of Computer Science (FOCS)*, pages 578–587, 2013.

[GKKS13b] Ankit Gupta, Neeraj Kayal, Pritish Kamath, and Ramprasad Saptharishi. Approaching the chasm at depth four. In *Conference on Computational Complexity (CCC)*, 2013.

[GKPS11] Bruno Grenet, Pascal Koiran, Natacha Portier, and Yann Strozecki. The Limited Power of Powering: Polynomial Identity Testing and a Depth-four Lower Bound for the Permanent. In *Proceedings of the 30th Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 127–139, 2011.

[GR98] Dima Grigoriev and Alexander A. Razborov. Exponential complexity lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. In *FOCS*, pages 269–278, 1998.

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

[KLSS14] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. In *Foundations of Computer Science (FOCS)*, pages 61–70, 2014.

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012.

[KS14a] Neeraj Kayal and Chandan Saha. Lower Bounds for Depth Three Arithmetic Circuits with small bottom fanin. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:89, 2014.

[KS14b] Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *Foundations of Computer Science (FOCS)*, pages 363–373, 2014.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *STOC*, pages 146–153, 2014.

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

[Raz09] Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2), 2009.

[Raz10] Ran Raz. Elusive functions and lower bounds for arithmetic circuits. *Theory of Computing*, 6(1):135–177, 2010.

[RY09] Ran Raz and Amir Yehudayoff. Lower Bounds and Separations for Constant Depth Multilinear Circuits. *Computational Complexity*, 18(2):171–207, 2009.

[Tav13] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013.

[VSBR83] L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. Fast parallel computation of polynomials using few processors. *SIAM Journal on Computing*, 12(4):641–644, 1983.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de