Finer Separations Between Shallow Arithmetic Circuits

Mrinal Kumar[*1] **and Ramprasad Saptharishi**[†2]

1 **Rutgers University, New Brunswick, NJ, USA**
mrinal.kumar@rutgers.edu
2 **Tel Aviv University, Israel**
ramprasad@cmi.ac.in

Abstract

In this paper, we show that there is a family of polynomials $\{P_n\}$, where P_n is a polynomial in n variables of degree at most $d = O(\log^2 n)$, such that

- P_n can be computed by linear sized homogeneous depth-5 circuits.
- P_n can be computed by poly(n) sized non-homogeneous depth-3 circuits.
- Any homogeneous depth-4 circuit computing P_n must have size at least $n^{\Omega(\sqrt{d})}$.

This shows that the parameters for the depth reduction results of [1, 11, 20] are tight for extremely restricted classes of arithmetic circuits, for instance homogeneous depth-5 circuits and non-homogeneous depth-3 circuits, and over an appropriate range of parameters, qualitatively improve a result of Kumar and Saraf [14], which showed that the parameters of depth reductions are optimal for algebraic branching programs.

1998 ACM Subject Classification I.1.1 Expressions and Their Representation

Keywords and phrases arithmetic circuits, lower bounds, separations, depth reduction

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2016.38

1 Introduction

An arithmetic circuit over a field $\mathbb{F}$ and variables $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ is a directed acyclic graph with nodes labelled by $+$ and $\times$ operations over $\mathbb{F}$ and leaves (nodes of in-degree 0) labelled by elements of $\mathbb{F}$ and $\mathbf{x}$. The circuit computes an n variate polynomial in $\mathbb{F}[\mathbf{x}]$ in the natural way. Arithmetic circuits are natural and intuitive models of computation in the algebraic setting as they allow us to represent multivariate polynomials succinctly. For an introduction to the area of arithmetic circuit complexity, we refer the interested reader to the excellent survey of Shpilka and Yehudayoff [19].

Bounded depth arithmetic circuits

Most of the recent research in the area of arithmetic circuit complexity is centered around the question of proving strong lower bounds for structured bounded depth arithmetic circuits, in particular homogeneous depth-4 arithmetic circuits [5, 9, 14]. The focus on such circuits is due to a result of Agrawal and Vinay [1] and subsequent strengthening by Koiran [11] and Tavenas [20], which show that strong enough lower bounds for such structured bounded

* Research supported in part by Simons Graduate Fellowship.
† The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 257575

depth circuits suffice for general arithmetic circuit lower bounds. For an outline of most of the recent results related to lower bounds for homogeneous depth-4 arithmetic circuits, we refer the reader to a survey of Saptharishi [17].

These set of structural results, collectively referred to as *depth reductions*, show that any homogeneous polynomial in n variables of degree $d = \text{poly}(n)$ which can be computed by an arithmetic circuit of size $\text{poly}(n)$ can also be computed by a homogeneous depth-4 arithmetic circuit of size $n^{O(\sqrt{d})}$. A natural question here is to try and understand if the parameters in the above result are asymptotically tight. This direction has previously been explored, and Kumar and Saraf [14] showed a lower bound of $n^{\Omega(\sqrt{d})}$ for a polynomial that has a $\text{poly}(n)$-sized arithmetic circuit. This implies that, in general, the size bound of $n^{O(\sqrt{d})}$ can not be improved to $n^{o(\sqrt{d})}$ for $\text{poly}(n)$-sized arithmetic circuits. However, as far as we know, it was not known if such improved depth reductions are conceivable for slightly restricted classes of arithmetic circuits, for instance, arithmetic formulas or constant depth arithmetic circuits. In this paper, we study this problem and show that at least for the case when $d = O(\log^2 n)$, one cannot hope to prove such improved depth reduction results, for even extremely restricted classes of arithmetic circuits such as linear size homogeneous depth-5 arithmetic circuits, or polynomial sized non-homogeneous depth-3 arithmetic circuits.

We now state our results, and elaborate on how they compare to the known results.

1.1 Our results

We prove the following theorems.

▶ **Theorem 1.1.** *Let $\mathbb{F}$ be any field. There is a family of polynomials $\{P_n\}$ over $\mathbb{F}$, where P_n is of degree $d = O(\log^2 n)$ on n variables such that P_n can be computed by a homogeneous depth-5 circuit of size $O(n)$ whereas any homogeneous depth-4 circuit computing P_n requires size $n^{\Omega(\sqrt{d})}$.*

▶ **Theorem 1.2.** *Let $\mathbb{F}$ be any field of characteristic zero. There is a family of polynomials $\{P_n\}$ over $\mathbb{F}$, where P_n is of degree $d = O(\log^2 n)$ on n variables such that P_n can be computed by a (non-homogeneous) depth-3 circuit of size $\text{poly}(n)$ whereas any homogeneous depth-4 circuit computing P_n requires size $n^{\Omega(\sqrt{d})}$.*

1.2 Comparison to earlier results

An $n^{\Omega(\sqrt{d})}$ lower bound for homogeneous depth-4 circuits was proved for an explicit polynomial of degree d in n variables in VNP by Kayal, Limaye, Saha and Srinivasan [9] and for the iterated matrix product (IMM) by Kumar and Saraf [14]. Improvements on this can happen on three fronts – (1) by improving the bound from $n^{\Omega(\sqrt{d})}$ to $n^{\omega(\sqrt{d})}$, or (2) by making the lower bound work for a class more general than homogeneous depth-4 circuits, or (3) by proving the lower bound for a polynomial "simpler" than IMM. This work is of the last category where the polynomial is computed by linear sized homogeneous depth-5 circuits or polynomial sized depth-3 circuits.

We elaborate more on this now.

Depth reduction to depth-4 as a springboard for stronger lower bounds

Let $\mathcal{C}$ be a class of arithmetic circuits. If we had a depth reduction result that showed that all homogeneous polynomials of degree d in n variables that can be computed by an arithmetic circuit $C \in \mathcal{C}$ of size $s(n)$ can also be computed by a homogeneous depth-4 arithmetic circuit

of size $s^{o(\sqrt{d})}$, then it follows from the results in [9, 14] that there is an explicit polynomial in VP (or VNP) that cannot be computed by polynomial size arithmetic circuits in $\mathcal{C}$. In this sense, the *efficient* reductions to homogeneous depth-4 circuits is a *springboard* to prove lower bounds for many potentially stronger classes of circuits.

The lower bound for IMM in [14] rules out this strategy when $\mathcal{C}$ is the class of algebraic branching programs, since it shows polynomial families (namely IMM) that have linear size ABPs but require homogeneous depth-4 circuits of size $n^{\Omega(\sqrt{d})}$. However the strategy could still, in principle, work for other interesting classes of arithmetic circuits such as arithmetic formulas, constant depth arithmetic circuits or, possibly the simplest of them all, the class of homogeneous depth-5 arithmetic circuits. Another simple class of circuits for which this strategy could be tried is the class of non-homogeneous depth-3 circuits, where superpolynomial lower bounds are not known when the size of the underlying field is large. Theorem 1.1 and Theorem 1.2 show that the above mentioned classes of arithmetic circuits cannot be reduced to homogeneous depth-4 arithmetic circuits of size $n^{o(\sqrt{d})}$, albeit for an appropriate range of parameters. So, even though quantitatively we do not prove improved lower bounds, qualitatively, we show near optimal separations between complexity classes which are much closer to each other that was earlier known. Unfortunately, we are only able to show such separations when the degree $d = O(\log^2 n)$.

Non-homogeneous depth-3 circuits

Theorem 1.2 shows a separation between non-homogeneous depth-3 circuits and homogeneous depth-4 circuits, in a low degree regime. Intuitively, to prove such a separation, we need a candidate family of hard polynomials which have polynomial sized non-homogeneous depth-3 circuits and are believed to require homogeneous depth-4 circuits of size $n^{\Omega(\sqrt{d})}$. At first glance, it seems unclear what this polynomial should be. The elementary symmetric polynomial of degree d is *not* a good candidate[1] as it can indeed be computed by a homogeneous depth four circuit of size $2^{O(\sqrt{d})}$ [7]. However, *a generic affine projection* of the elementary symmetric polynomial, as studied by Shpilka [18], is a natural candidate and is almost complete for this model.

In this paper, however, we do not directly work with this polynomial but it can be easily inferred that the lower bound applies to a generic affine projection of the elementary symmetric polynomial as well.

Depth hierarchy theorems for arithmetic circuits

Depth hierarchy theorems, which show an exponential, (and near optimal) separation between depth h and depth $h+1$ circuits [6, 16] constitute some of the most celebrated results in the theory of lower bounds for bounded depth boolean circuits. It is natural to ask if such separations can be shown for arithmetic circuits. Unfortunately, superpolynomial lower bounds are not known in general when the depth of the arithmetic circuits is more than four [2]. So, at this point, we can only hope to show such separations between homogeneous depth-5 and homogeneous depth-4 arithmetic circuits. Due to the depth reduction results, the best such separation one can hope to prove for an n variate degree d polynomial would be $n^{\Omega(\sqrt{d})}$. We prove a matching lower bound, as long as the degree d is at most $O(\log^2 n)$. In the arithmetic circuit literature, the question of depth hierarchy theorems has previously been studied by Raz and Yehudayoff [15], where they show superpolynomial separation

[1] Indeed, for low enough degrees, they are known to have fairly high shifted partials complexity [3].
[2] For homogeneous depth-5 circuits, such lower bounds are known only over small finite fields.

separation between multilinear circuits of product depth d and product depth $d+1$, for $d = O(1)$. In the non-multilinear world, to the best of our knowledge this is the first such attempt. Even in the context of constant depth multilinear circuits, the separation in [15] is between depth-4 and depth-6 circuits, and not between depth-4 and depth-5 circuits.

The complexity measure

The proof of Kayal et al. [9] and Kumar and Saraf [14] rely on the notion of projected shifted partials of a polynomial as a measure of its complexity. This measure can be thought of as a variant of shifted partials which tries to take advantage of the fact that the hard polynomial is multilinear. The measure in this paper takes advantage of *set-multilinearity* instead of just multilinearity, and such a variant was essentially used in [10], where they showed an $n^{O(\log n)}$ lower bound for iterated matrix multiplication and the determinant. Our proofs rely on a slightly different interpretation of the measure, which makes the proofs much more transparent. Intuitively, this measure tries to take advantage of the fact that the hard polynomial (Nisan-Wigderson design polynomials or the IMM) is not just multilinear, but in fact set-multilinear. In the regime where $d \ll n$, set multilinearity is a much more rigid restriction on a polynomial when compared to multilinearity, and in some sense our gain comes from this observation. Our hard polynomial for Theorem 1.1 is also a simple generic balanced depth-5 circuit.

One might wonder if the results in this paper could have been shown by using the dimension of the projected shifted partial derivatives as the complexity measure. In particular, can we show that the projected shifted partials complexity of a generic depth-5 circuit is sufficiently close to the largest possible value? This would suffice for Theorem 1.1. Although we do not have enough evidence to conjecture one way or the other, intuitively this problem seems tricky since so far the known analyses of the projected shifted partials of a polynomial seems to rely on pairwise distance between the monomials of the hard polynomial, either in the worst case (Nisan-Wigderson polynomial [9, 14]), or in the average case (IMM [14]). Clearly, the monomials in a generic depth-5 circuit do not have good distance in the worst case, and to the best of our understanding, the guarantees about distance in the average case seem a bit weaker than what would suffice to simulate the proof in [14] for a generic depth-5 circuit. However, this problem of proving lower bounds on the dimension of projected shifted partials of homogeneous depth-5 circuits is of independent interest, since even if the answer is negative and homogeneous depth-5 circuits do not have large enough projected shifted partials complexity, then we could use this as a measure to prove lower bounds for such circuits. So far, such lower bounds are only known over small finite fields [12].

2 Preliminaries

2.1 Notations

- Throughout the paper, we use bold-face letters such as $\mathbf{x}$ to denote a sets of variables. Most of the times, the size of this set would be clear from context. We use $\mathbf{x}^{\mathbf{e}}$ to refer to the monomial $x_1^{e_1} \cdots x_n^{e_n}$.
- We use the short-hand $\partial_{\mathbf{x}^{\mathbf{e}}}(P)$ to denote
$$\frac{\partial^{e_1}}{\partial x_1^{e_1}} \left(\frac{\partial^{e_2}}{\partial x_2^{e_2}} \left(\cdots (P) \cdots \right) \right).$$
- For a set of polynomials $\mathcal{P}$ use $\partial^{=k}\mathcal{P}$ to denote the set of all k-th order partial derivatives of polynomials in $\mathcal{P}$, and $\partial^{\leq k}\mathcal{P}$ similarly.

Also, $\mathbf{x}^{=\ell}\mathcal{P}$ refer to the set of polynomials of the form $\mathbf{x}^{\mathbf{e}} \cdot P$ where $\mathsf{Deg}(\mathbf{x}^{\mathbf{e}}) = \ell$ and $P \in \mathcal{P}$. Similarly $\mathbf{x}^{\leq\ell}\mathcal{P}$.

- For an integer $m > 0$, we use $[m]$ to denote the set $\{1, \ldots, m\}$.
- For a set of vectors (or polynomials) V, their span over $\mathbb{F}$ will be denoted by $\mathsf{Span}(V)$ and their dimension by $\mathsf{Dim}(V)$.
- For a subset $\mathbf{y}$ of variables and a polynomial $P \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$, by $\mathsf{Mult}_{\mathbf{y}}[P]$, we denote the polynomial $P' \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$ which is obtained by projecting P only to its monomials which are multilinear in $\mathbf{y}$.
  Similarly, for a set S of polynomials, $\mathsf{Mult}_{\mathbf{y}}[S]$ denotes the set of polynomials obtained by projecting every polynomial in S to the monomials which are multilinear in $\mathbf{y}$.

2.2 The hard polynomial

The hard function for the lower bounds will be a generic balanced $\Pi\Sigma\Pi\Sigma$ circuit with appropriate parameters. We define the polynomial $P_{m,d}$ as

$$P_{m,d} \quad = \quad \prod_{i=1}^{\sqrt{d}} \sum_{j=1}^{m} \prod_{i'=1}^{\sqrt{d}} \sum_{j'=1}^{m} x_{iji'j'}.$$

The polynomial $P_{m,d}$ depends on $m^2 d$ variables. It would be useful to have $L_{iji'} = \sum_{j'} x_{iji'j'}$ so that $P_{m,d} = \prod_i \sum_j \prod_{i'} L_{iji'}$.

Observe that the polynomial $P_{m,d}$ is a set multilinear polynomial for the partition of variables into $\{\mathbf{x}_{i*i'*} : i, i' \in [\sqrt{d}]\}$, where $\mathbf{x}_{i*i'*} = \{x_{iji'j'} \ : \ j, j' \in [m]\}$. There are d such sets and each is of size m^2.

The range of parameters we will be working with in this paper when $d = \delta \log^2 n$ for a small enough constant δ. For such small d, it follows from observations in [4] that the polynomial $P_{m,d}$ is computable by a polynomial sized non-homogeneous depth-3 circuit. More formally, the proof relies on the following lemma which is implicit in [4].

▶ **Lemma 2.1** ([4]). *Let C be a homogeneous $\Sigma\Pi^{[a]}\Sigma\Pi^{[b]}\Sigma$ circuit of size s over $\mathbb{C}$, the field of complex numbers, which computes an n-variate polynomial P. Then there is an equivalent $\Sigma\Pi\Sigma$ circuit C' of size $s' = \mathrm{poly}(2^a, 2^b, n, s)$ which computes P.*

Using this observation, we have the following lemma which shows that there is a small depth-3 circuit for $P_{m,d}$.

▶ **Lemma 2.2.** *Let P be an n variate polynomial of degree $d = O(\log^2 n)$ which is computed by a homogeneous $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}\Sigma$ circuit C of size s. Then, P is computable by a $\Sigma\Pi\Sigma$ circuit of size $\mathrm{poly}(n)$.* ◀

Thus, to prove Theorem 1.1 and Theorem 1.2, it suffices to show an $n^{\Omega(\sqrt{d})}$ lower bound on the size of homogeneous $\Sigma\Pi\Sigma\Pi$ arithmetic circuits computing $P_{m,d}$.

2.3 Some useful approximations

▶ **Lemma 2.3** ([5]). *Let n, a, b satisfy $a + b = o(n)$. Then,*

$$\frac{(n+a)!}{(n-b)!} \quad = \quad n^{a+b} \cdot \exp(O((a+b)^2/n)).$$

In particular, if $a + b = o(\sqrt{n})$, then the right hand side is $(1 + o(1)) \cdot n^{a+b}$.

▶ **Lemma 2.4.** *For all $x, y > 0$,*

$$e^{xy} \quad \geq \quad (1+x)^y \quad \geq \quad e^{\frac{xy}{x+1}}.$$

3 Proof of Theorem 1.1

The first step in previous lower bounds for homogeneous depth-4 circuits is using a random restriction to set each variable independently to zero with a certain probability. We shall first analyze the random restriction process on a homogeneous depth-4 circuit and also on the polynomial $P_{m,d}$.

3.1 The effect of a random restriction

Our restrictions $\mathcal{R}_p$ will be defined by setting every variable to zero with a probability $1-p$ and keeping it alive with a probability p.

▶ **Lemma 3.1.** *Let $\epsilon > 0$ be any fixed constant and let $p = \frac{1}{n^\epsilon}$. Let C be a $\Sigma\Pi\Sigma\Pi$ circuit of size $n^{\frac{\epsilon^2}{2}\sqrt{d}}$. Then with a probability at least $1 - o(1)$ over $\pi \leftarrow \mathcal{R}_p$, every product gate at the lowest level of C (closest to the leaves) that depends on more than $\epsilon\sqrt{d}$ distinct variables is set to zero in $\pi(C)$.*

Proof. Consider any product gate of support at least $\epsilon\sqrt{d}$ present at the bottom level of C. The probability that this gate is not set to zero in $\pi(C)$ is at most $\frac{1}{n^{\epsilon^2\sqrt{d}}}$. So, by a union bound over all the product gates in C, the probability that some gate of support at least $\epsilon\sqrt{d}$ survives in $\pi(C)$ is at most $n^{\frac{\epsilon^2}{2}\sqrt{d}} \cdot \frac{1}{n^{\epsilon^2\sqrt{d}}}$ which is $o(1)$. ◀

We now analyse the effect of random restrictions on our candidate hard function.

▶ **Lemma 3.2.** *Let ϵ be a small enough constant and let $p = \frac{1}{n^\epsilon}$, and let $P_{m,d}$ be the polynomial as defined in subsection 2.2. Then, with probability at least $1 - o(1)$ over $\pi \leftarrow \mathcal{R}_p$, the polynomial $\pi(P_{m,d})$ is of the form*

$$\pi(P_{m,d}) = \prod_{i=1}^{\sqrt{d}} \sum_{j=1}^{m} \prod_{i'=1}^{\sqrt{d}} L'_{iji'}$$

where each $L'_{iji'}$ is a non-zero linear form.

Proof. From our choice of parameters, observe that $n = m^2 d$, and since $d = O(\log^2 n)$, $m > n^{1/4}$. Now, for any fixed linear form $L_{iji'}$, the probability that $\pi(L_{iji'})$ equals zero is equal to $(1-p)^m = (1 - 1/n^\epsilon)^m$ which is less than $(1 - 1/n^\epsilon)^{n^{2\epsilon}} = \frac{1}{\omega(n)}$. Therefore, the probability that there exists a linear form $L_{iji'}$ such that $\pi(L_{iji'}) \equiv 0$ is $o(1)$, and the lemma follows. ◀

At this point, we will deterministically set all but one alive variable in each $L'_{iji'}$ in the above lemma to zero, and obtain the following corollary upto a relabelling of variables.

▶ **Corollary 3.3.** *Let ϵ be a fixed constant and $p = \frac{1}{n^\epsilon}$, and let $P_{m,d}$ be the polynomial as defined in subsection 2.2. Then, with probability at least $1 - o(1)$ over $\pi \leftarrow \mathcal{R}_p$, there is a $0,1$ projection of $\pi(P_{m,d})$ which is of the form*

$$P'_{m,d} = \prod_{i=1}^{\sqrt{d}} \sum_{j=1}^{m} \prod_{i'=1}^{\sqrt{d}} x_{iji'},$$

where each $x_{iji'}$ is a distinct variable.

Observe that Theorem 3.1 continues to hold under this additional deterministic restriction, as the bottom support of a depth-4 circuit does not increase under $0,1$ projections. Clearly $P'_{m,d}$ is computable by a homogeneous depth-4 circuit of bottom fan-in $\sqrt{d}$.

In order to complete the proof, it suffices to show that any homogeneous depth-4 circuit of *bottom support* bounded by $\sqrt{d}/10$ that computes $P'_{m,d}$ must have size $n^{\Omega(\sqrt{d})}$. In fact, Kumar and Saraf [13] have shown that any homogeneous depth-4 circuit of bottom fan-in at most $\sqrt{d}/10$ computing $P'_{m,d}$ must require size $n^{\Omega(\sqrt{d})}$ using the measure of dimension of shifted partial derivatives. Thus we need to find a way to lift this lower bound to the class of homogeneous depth-4 circuit of *bottom support* bounded by $\sqrt{d}/10$. To do this, we modify the measure of dimension of shifted partials in order to address small bottom support instead of small bottom fan-in.

3.2 The complexity measure

The measure is again the dimension of an appropriate linear space of polynomials.

▶ **Definition 3.4** (The complexity measure)**.** Let $\mathbf{x} = \mathbf{x}_1 \sqcup \cdots \sqcup \mathbf{x}_d$ be a partition of the variables into d sets. For any polynomial $P \in \mathbb{F}[\mathbf{x}]$, define $P' \in \mathbb{F}[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d, y_1, y_2, \ldots, y_d]$ be the the polynomial derived from P by replacing every occurence of the variable $x_{ij} \in \mathbf{x}_i$ by $y_i \cdot x_{ij}$. Then, the complexity measure

$$\Gamma_{k,\ell}(P) \quad := \quad \mathsf{Dim}_{\mathbb{F}}\left\{\left(\mathbf{x}^{=\ell} \cdot \mathsf{Mult}_{\mathbf{y}}[\partial^{=k}(P')]\right)\right\}.$$

We remark that all the derivatives and shifts in the definition of $\Gamma_{k,\ell}$ are taken with respect to the variables in $\mathbf{x}$. However, the multilinearization is done with respect to the $\mathbf{y}$ variables. As mentioned earlier, this measure was used in [10] where it was called *dimension of shifted projected partial derivatives*.

As is clear from the definition, the measure is subadditive, i.e for every pair of polynomials P and Q and for every pair of field constants α and β, the inequality $\Gamma_{k,\ell}(\alpha P + \beta Q) \leq \Gamma_{k,\ell}(P) + \Gamma_{k,\ell}(Q)$ holds for every choice of k and ℓ.

Throughout this paper, we will be using very simple connections between the measure $\Gamma_{k,\ell}$ and the well known notion of shifted partial derivatives of polynomials (first defined in [8]), defined as

▶ **Definition 3.5** (Shifted partial derivatives)**.** Define $P \in \mathbb{F}[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d]$ be a polynomial. Then, the *dimension of shifted partial derivatives* is defined as

$$\mathsf{Dim}_{\mathbb{F}}\left\{\left(\mathbf{x}^{=\ell} \cdot \partial^{=k}(P)\right)\right\}.$$

Observe that if a polynomial P is set-multilinear with respect to the partition of the variables in $\mathbf{x}$ into $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d$, then multilinearization with respect to the $\mathbf{y}$ variables does not kill any of the monomials in the partial derivatives. In particular, for a set multilinear polynomial P, and for every choice of k, ℓ, the quantitity $\Gamma_{k,\ell}(P)$ is exactly equal to the dimension of shifted partial derivatives of the polynomial P where we take derivatives of order k and shifts are of degree ℓ. This observation will be useful for us in the proof and is summarised below.

▶ **Observation 3.6.** *Let P be a set multilinear polynomial of degree d. Then for every choice of parameters k and ℓ,*

$$\Gamma_{k,\ell}(P) \;=\; \mathsf{Dim}\left(\mathbf{x}^{=\ell} \cdot \partial^{=k}(P)\right).$$

Since $P_{m,d}$ is set multilinear with respect to the partition

$$\mathbf{x} = \bigsqcup_{i,i' \le \sqrt{d}} \mathbf{x}_{i*i'*}$$

we use this partition for in the definition of $\Gamma_{k,\ell}$. To complete the proof, we use this measure to show that $P'_{m,d}$ cannot be computed by small homogeneous depth-4 circuit of bottom support bounded by $\sqrt{d}/10$.

3.3 Upper bound for a small bottom-support depth-4 circuit

▶ **Lemma 3.7.** *Let C be a homogeneous $\Sigma\Pi\Sigma\Pi$ circuit with bottom support at most s which computes a degree d polynomial in $\mathbb{F}[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d]$. Then, for every k and ℓ,*

$$\Gamma_{k,\ell}(C) \;\le\; \mathit{Size}(C) \cdot 2^{2d} \cdot \binom{d}{k} \cdot \binom{n+\ell+ks}{n}.$$

Proof. Since the measure $\Gamma_{k,\ell}$ is subadditive, we will prove an upper bound on $\Gamma_{k,\ell}$ for one product term in C. So, let $T = Q_1 \cdot Q_2 \cdots Q_t$, where each Q_i has support at most s. Without loss of generality, we can assume that $t \le d$ since the circuit C is homogeneous to start with.

Recall that in the first step, we replace every variable x_{ij} by $y_i \cdot x_{ij}$. This transforms $T = Q_1 \cdots Q_t$ into $T = Q'_1 \cdot Q'_2 \cdots Q'_t$. Every monomial $\mathbf{x}^\alpha$ in the $\mathbf{x}$ variables will be transformed to a monomial $\mathbf{y}^{\alpha'} \cdot \mathbf{x}^\alpha$ by this transformation. The key points are that $\mathbf{y}^{\alpha'}$ is only over d variables, and if $\mathbf{x}^\alpha$ is non-multilinear then so is $\mathbf{y}^{\alpha'}$.

Let us now consider the derivative of T with respect to a monomial $\mathbf{x}^\alpha$ of order k.

$$\partial_{\mathbf{x}^\alpha}(T') \in \mathsf{Span}\left\{\partial_{\mathbf{x}^\alpha}(Q'_A) \cdot Q'_{\overline{A}} \;:\; A \subseteq [t], |A| \le k\right\},$$

where Q'_A is a shorthand for $\prod_{i \in A} Q'_i$.

$$\mathsf{Mult}_{\mathbf{y}}\left[\partial_{\mathbf{x}^\alpha}(T')\right] \in \mathsf{Span}\left\{\mathsf{Mult}_{\mathbf{y}}\left[\partial_{\mathbf{x}^\alpha}(Q'_A) \cdot Q'_{\overline{A}}\right] \;:\; A \subseteq [t], |A| \le k\right\}.$$

Since we are interested in the multilinear component, it suffices to only focus on multilinear (in $\mathbf{y}$) monomials in both $\partial_{\mathbf{x}^\alpha}(Q'_A)$ and $Q'_{\overline{A}}$. Since Q'_A is a product of at most k polynomials, each of support-size bounded by s, the only monomials $\mathbf{x}^\beta$ that can contribute a multilinear $\mathbf{y}$-part can have degree at most ks. Therefore,

$$\begin{aligned}
\mathsf{Mult}_{\mathbf{y}}\left[\partial_{\mathbf{x}^\alpha}(Q'_A)\right] &\in \mathsf{Span}\left\{\mathbf{y}^\beta \cdot \mathbf{x}^\gamma \;:\; \mathsf{Deg}(\mathbf{x}^\gamma) \le ks \,,\, \mathbf{y}^\beta \text{ multilinear}\right\} \\
\mathsf{Mult}_{\mathbf{y}} Q'_{\overline{A}} &= \sum_{\beta'} \mathbf{y}^{\beta'} \cdot Q'_{\overline{A},\beta'} \\
\implies \mathsf{Mult}_{\mathbf{y}}\left[\partial_{\mathbf{x}^\alpha}(Q'_A) \cdot Q'_{\overline{A}}\right] &\in \mathsf{Span}\left\{\mathbf{y}^\beta \mathbf{y}^{\beta'} \cdot \mathbf{x}^\gamma \cdot Q'_{\overline{A},\beta'} \;:\; \mathsf{Deg}(\mathbf{x}^\gamma) \le ks \,,\, \mathbf{y}^\beta \mathbf{y}^{\beta'} \text{ multilinear}\right\}.
\end{aligned}$$

Taking the union over all shifts and all derivatives, we get

$$\begin{aligned}
\mathbf{x}^{=\ell} \cdot \mathsf{Mult}_{\mathbf{y}}[\partial^{=k}(T')] \;\subseteq\; \mathsf{Span}\Big\{&\mathbf{y}^\beta \mathbf{y}^{\beta'} \cdot \mathbf{x}^\gamma \cdot Q'_{\overline{A},\beta'} \;:\; A \subseteq [t] \,,\, |A| \le k \,, \\
&\text{degree } (\mathbf{x}^\gamma) \le \ell + ks \,,\, \mathbf{y}^\beta \mathbf{y}^{\beta'} \text{ is multilinear } \Big\}.
\end{aligned}$$

For any k, ℓ, it follows that

$$\Gamma_{k,\ell}(T') \;\le\; 2^{2d} \cdot \binom{d}{k} \cdot \binom{n+\ell+ks}{n}.$$

Using subadditivity, we obtain the lemma. ◀

3.4 Lower bound for the measure on $P'_{m,d}$

The final technical ingredient of our proof will be a lower bound on the dimension of shifted partials of the polynomial $P'_{m,d}$. The bound follows from the calculations in [13], but we provide the calculation here for completeness.

▶ **Lemma 3.8.** *Recall the polynomial*

$$P'_{m,d} = \prod_{i=1}^{\sqrt{d}} \sum_{j=1}^{m} \prod_{i'=1}^{\sqrt{d}} x_{iji'}$$

where each $x_{iji'}$ is a distinct variable. For $k = \sqrt{d}$ and any ℓ, we have

$$\mathsf{Dim}\left(\mathbf{x}^{=\ell} \cdot \partial^{=k}(P)\right) \geq \frac{1}{4} \cdot \left(\frac{n+\ell}{\ell}\right)^{\frac{1}{2}\cdot(d-\sqrt{d})} \cdot \binom{n+\ell-1}{n}.$$

Proof. To show that the shifted partials complexity of P is large, we will follow the outline in [13]. We consider the following subset $\mathcal{S}$ of monomials of degree equal to $k = \sqrt{d}$:

$$\mathcal{S} = \{x_{1a_11} \cdot x_{2a_21} \cdots x_{ka_k1} : a_1, a_2, \ldots, a_k \in [m]\}.$$

Firstly, note that for any monomial $\mathbf{x}^\alpha = x_{1a_11} \cdots x_{ka_k1} \in \mathcal{S}$, the derivative $\partial_{\mathbf{x}^\alpha}(P)$ is just the monomial

$$(x_{1a_12} \cdots x_{1a_1k}) \cdots (x_{ka_k2} \cdots x_{ka_kk}).$$

Thus, it suffices to get a lower bound of distinct monomials obtained as shifts of such derivatives. To assist this calculation, we pick a subset $\mathcal{S}'$ of the set $\mathcal{S}$ such that the distance between any two monomials in $\mathcal{S}'$ is 'large', and the size of $\mathcal{S}'$ is also 'large'. This can be done by picking the monomials which correspond to a good code of length k over the alphabet $\Sigma = \{1, 2, \ldots, m\}$. To this end, we pick a Reed-Somolon code of relative distance $1/2$ and rate $1/2$. This can be done as long as m is a prime power and $\sqrt{d} \leq m$. Let $\mathcal{S}'$ be a such set of size $m^{k/2}$ where any pair of monomials in $\mathcal{S}'$ differ on at least $\sqrt{d}/2$ locations.

When we take derivatives of P with respect to monomials in the set $\mathcal{S}'$, two monomials obtained from distinct elements of $\mathcal{S}'$ have distance at least $\Delta = \sqrt{d}(\sqrt{d}-1)/2 = (d-\sqrt{d})/2$. So, each of the shifted partial derivatives obtained by shifting the derivatives of P by monomials of degree ℓ is just a monomial, and a lower bound on the number of distinct monomials obtained in this way gives us a lower bound on $\mathsf{Dim}\left(\mathbf{x}^{=\ell} \cdot \partial^{=k}(P)\right)$. In fact, we shall choose an even smaller set $\mathcal{S}''$ to ensure the following bounds work out.

By the inclusion-exclusion approach of Chillara and Mukhopadhyay [2], for any set $\mathcal{S}'' \subset \mathcal{S}'$ we get the following:

$$\mathsf{Dim}\left(\mathbf{x}^{=\ell} \cdot \partial^{=k}(P)\right) \geq |\mathcal{S}''| \cdot \binom{n+\ell-1}{n} - \frac{|\mathcal{S}''|^2}{2} \cdot \binom{n+\ell-\Delta-1}{n}.$$

If we pick our parameters, such that the first term above is at least twice the second term, then we would be done. For this, we need

$$|\mathcal{S}''| \leq \frac{\binom{n+\ell-1}{n}}{\binom{n+\ell-\Delta-1}{n}}.$$

For our choice of parameters, $\ell, n \gg d^2$, the ratio $\frac{\binom{n+\ell-1}{n}}{\binom{n+\ell-\Delta-1}{n}}$ can be approximated by $\left(\frac{n+\ell}{\ell}\right)^{\Delta}$ within a factor $1 \pm o(1)$ by Theorem 2.3. So, it suffices if our choice of parameters satisfies (omitting floors)

$$|\mathcal{S}''| \;=\; \frac{1}{2} \cdot \left(\frac{n+\ell}{\ell}\right)^{\Delta}.$$

Plugging in Δ and the size of $\mathcal{S}''$ in the inclusion-exclusion bound, we get

$$\mathsf{Dim}\left(\mathbf{x}^{=\ell} \cdot \partial^{=k}(P)\right) \;\geq\; \frac{1}{4} \cdot \left(\frac{n+\ell}{\ell}\right)^{(d-\sqrt{d})/2} \cdot \binom{n+\ell-1}{n}. \qquad \blacktriangleleft$$

3.5 Putting it together

▶ **Theorem 3.9** (Theorem 1.1 restated)**.** *Let C be a homogeneous depth-4 arithmetic circuit which computes the polynomial $P_{m,d}$ for $d = 0.0001 \log^2 n$. Then, the size of C is at least* $\exp(\Omega(\sqrt{d} \log n))$.

Proof. Assume on the contrary that the polynomial $P_{m,d}$ can be computed by C, a homogeneous depth-4 circuit of size at most $\exp(0.001\sqrt{d} \log n)$. If we apply a random restriction that sets every variable to zero independently with probability $1/n^{0.1}$, by Theorem 3.1 (with $\epsilon = 0.1$), the circuit reduces to C', a homogeneous depth-4 circuit with bottom support bounded by $\sqrt{d}/10$ with probability $1 - o(1)$.

On the other hand by Theorem 3.3, the polynomial $P_{m,d}$ under such a random restriction still retains $P'_{m,d}$ as a projection with high probability. Fix a restriction that satisfies both these properties and we now have a homogeneous depth-4 circuit C'' with bottom support bounded by $\sqrt{d}/10$ and size at most $\exp(0.001\sqrt{d} \log n)$ that computes $P'_{m,d}$.

Let $k = \sqrt{d}$ and $\ell = \frac{n\sqrt{d}}{\log n}$. By Theorem 3.7, we have

$$\Gamma_{k,\ell}(C'') \quad \leq \quad \mathrm{Size}(C'') \cdot 2^{2d} \cdot \binom{n+\ell+(0.1)d}{n}.$$

On the other hand, by Theorem 3.8 and Theorem 3.6,

$$\Gamma_{k,\ell}(P'_{m,d}) \quad \geq \quad \frac{1}{4} \cdot \left(\frac{n+\ell}{\ell}\right)^{(d-\sqrt{d})/2} \cdot \binom{n+\ell-1}{n}.$$

Together, this implies that

$$\mathrm{Size}(C'') \geq \frac{1}{4} \cdot \frac{\binom{n+\ell-1}{n} \cdot \left(\frac{n+\ell}{\ell}\right)^{(d-\sqrt{d})/2}}{2^{2d} \cdot \binom{n+\ell+(0.1)d}{n}}.$$

For our regime of parameters, $\sqrt{d} = 0.01 \log n$ and hence $2^{2d} = n^{0.02\sqrt{d}} = \exp(0.02\sqrt{d} \log n)$. Simplifying the ratio of binomial coefficients using (Theorem 2.3), and using $\frac{d-\sqrt{d}}{2} > \frac{d}{3}$, we get

$$\begin{aligned}
\mathrm{Size}(C'') \;&\geq\; \frac{1}{\exp(0.02\sqrt{d} \log n)} \cdot \left(1 + \frac{n}{\ell}\right)^{d/3} \\
&\geq\; \frac{1}{\exp(0.02\sqrt{d} \log n)} \cdot \exp\left(\frac{(nd/3\ell)}{(n/\ell)+1}\right) \quad \text{(By Theorem 2.4)} \\
&>\; \exp\left(0.1\sqrt{d} \log n\right),
\end{aligned}$$

which contradicts the assumption on the size of C. Hence $\mathrm{Size}(C) \geq \exp(0.001\sqrt{d} \log n)$. $\blacktriangleleft$

4 Open questions

We end with some open questions.

- One question of great interest to us would be to show the lower bounds in this paper when the degree is larger. The other proofs of lower bounds for homogeneous depth-4 circuits [9, 14] tolerate degrees as high as $n^{1/2}$. We conjecture that the results in this paper are true even when the degree d and the number of variables n are polynomially related.
- Is the dimension of projected shifted partials of a generic homogeneous depth-5 circuit close to the largest possible value? This could offer one approach to resolving the first open problem.
- If the answer to the second problem above is negative, then we might be able to use projected shifted partials as a complexity measure to prove new lower bounds for homogeneous depth-5 arithmetic circuits. Hence, even proving non-trivial *upper bounds* on the projected shifted partials complexity of homogeneous depth-5 circuits would be very interesting.

Acknowledgements. Part of this work was done while the authors were visiting Microsoft Research, Bangalore in Summer 2014. We are grateful to Neeraj Kayal for many insightful discussions.

References

1 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. doi:10.1109/FOCS.2008.32.

2 Suryajith Chillara and Partha Mukhopadhyay. Depth-4 Lower Bounds, Determinantal Complexity : A Unified Approach. *Proceedings of the 31st Symposium on Theoretical Aspects of Computer Science (STACS 2014)*, 2014. Preliminary version at arXiv:1308.1640. doi:10.4230/LIPIcs.STACS.2014.239.

3 Hervé Fournier, Nutan Limaye, Meena Mahajan, and Srikanth Srinivasan. *The Shifted Partial Derivative Complexity of Elementary Symmetric Polynomials*, pages 324–335. Springer Berlin Heidelberg, 2015. doi:10.1007/978-3-662-48054-0_27.

4 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. doi:10.1109/FOCS.2013.68.

5 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)*. doi:10.1145/2629541.

6 Johan Håstad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing (STOC 1986)*, pages 6–20, 1986. doi:10.1145/12130.12132.

7 Pavel Hrubes and Amir Yehudayoff. Homogeneous formulas and symmetric polynomials. *Computational Complexity*, 20(3):559–578, 2011.

8 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. URL: http://eccc.hpi-web.de/report/2012/081/.

9 Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Circuits. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, 2014. doi:10.1109/FOCS.2014.15.

10 Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. Super-polynomial lower bounds for depth-4 homogeneous arithmetic formulas. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, 2014.

11 Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012. arXiv:1006.4700, doi:10.1016/j.tcs.2012.03.041.

12 Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. *Electronic Colloquium on Computational Complexity (ECCC)*, 2015. eccc:TR15-109. URL: http://eccc.hpi-web.de/report/2015/109/.

13 Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: it's all about the top fan-in. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 136–145, 2014. doi:10.1145/2591796.2591827.

14 Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, 2014. doi:10.1109/FOCS.2014.46.

15 Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Computational Complexity*, 18(2):171–207, 2009. Preliminary version in the *23rd Annual IEEE Conference on Computational Complexity (CCC 2008)*. doi:10.1007/s00037-009-0270-8.

16 Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. An average-case depth hierarchy theorem for boolean circuits. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2015)*, 2015. Preliminary version at arXiv:1504.03398. doi:10.1109/FOCS.2015.67.

17 Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Github survey, 2015. URL: https://github.com/dasarpmar/lowerbounds-survey/releases/.

18 Amir Shpilka. Affine projections of symmetric polynomials. *Journal of Computer and System Sciences*, 65(4):639–659, 2002. Preliminary version in the *16th Annual IEEE Conference on Computational Complexity (CCC 2001)*, eccc:TR01-035. doi:10.1016/S0022-0000(02)00021-1.

19 Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5:207–388, March 2010. doi:10.1561/0400000039.

20 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)*. doi:10.1016/j.ic.2014.09.004.