Electronic Colloquium on Computational Complexity, Report No. 153 (2013)

The Limits of Depth Reduction for Arithmetic Formulas: It's all about the top fan-in*

Mrinal Kumar† Shubhangi Saraf‡

Abstract

In recent years, a very exciting and promising method for proving lower bounds for arithmetic circuits has been proposed. This method combines the method of *depth reduction* developed in the works of Agrawal-Vinay [AV08], Koiran [Koi12] and Tavenas [Tav13], and the use of the shifted partial derivative complexity measure developed in the works of Kayal [Kay12] and Gupta et al [GKKS13a]. These results inspired a flurry of other beautiful results and strong lower bounds for various classes of arithmetic circuits, in particular a recent work of Kayal et al [KSS13] showing superpolynomial lower bounds for *regular* arithmetic formulas via an *improved depth reduction* for these formulas. It was left as an intriguing question if these methods could prove superpolynomial lower bounds for general (homogeneous) arithmetic formulas, and if so this would indeed be a breakthrough in arithmetic circuit complexity.

In this paper we study the power and limitations of depth reduction and shifted partial derivatives for arithmetic formulas. We do it via studying the class of depth 4 homogeneous arithmetic circuits. We show: (1) the first *superpolynomial lower bounds* for the class of homogeneous depth 4 circuits with top fan-in $o(\log n)$. The core of our result is to show *improved depth reduction* for these circuits. This class of circuits has received much attention for the problem of polynomial identity testing. We give the first nontrivial lower bounds for these circuits for any top fan-in ≥ 2. (2) We show that improved depth reduction *is not possible* when the top fan-in is $\Omega(\log n)$. In particular this shows that the depth reduction procedure of Koiran and Tavenas [Koi12, Tav13] cannot be improved even for homogeneous formulas, thus strengthening the results of Fournier et al [FLMS13] who showed that depth reduction is tight for circuits, and answering some of the main open questions of [KSS13, FLMS13]. Our results in particular suggest that the method of depth reduction and shifted partial derivatives may not be powerful enough to prove superpolynomial lower bounds for (even homogeneous) arithmetic formulas.

*Some of the results in this paper appeared in an earlier paper [KS13].

†Department of Computer Science, Rutgers University. Email: mrinal.kumar@rutgers.edu.

‡Department of Computer Science and Department of Mathematics, Rutgers University. Email: shubhangi.saraf@gmail.com.

ISSN 1433-8092

1 Introduction

In a seminal paper in 1979, Valiant [Val79] laid out a neat theoretical framework for the study of resource bounded algebraic computation and defined the complexity classes VP and VNP as the algebraic analogs of P and NP respectively. Since then, the problem of understanding whether VNP is different from VP has been a problem of fundamental significance in Algebraic Complexity Theory. To show that VNP is different from VP, it would suffice to show that the Permanent polynomial, which is a complete problem for VNP [Val79] does not have polynomial sized arithmetic circuits. Unfortunately, not much progress has been made towards proving superpolynomial arithmetic circuit lower bounds for any explicit polynomial in spite of the intensive attention that the problem has received. In recent years much effort has been invested in proving lower bounds for restricted classes of arithmetic circuits. The hope is that understanding restricted classed might shed light on how to approach the much more general and seemingly harder problem. Small depth circuits are one such class which have been quite intensively studied from this perspective, and even for small depth circuits, we really only understand lower bounds for depth 2 circuits and some classes of depth 3 and depth 4 circuits [NW95, SW01, GK98, GKKS13a, KSS13].

Recently a very promising and exciting new framework for proving lower bounds for arithmetic circuits has emerged. The framework consists of two major components. Let $\mathcal{C}$ be the class of circuits one wants to prove lower bounds for. The first step is to show that any circuit in $\mathcal{C}$ can be efficiently *depth reduced* to a depth 4 circuit with bounded bottom fan-in ($\Sigma\Pi\Sigma\Pi^{[t]}$ circuit). This depth reduction procedure was introduced and developed in the works of Agrawal-Vinay [AV08], Koiran [Koi12] and Tavenas [Tav13], building upon the initial depth reduction procedure of Valiant et al [VSBR83]. The second step is to prove strong lower bounds for $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits using the *shifted partial derivative* complexity measure, which was developed in the works of Kayal [Kay12] and Gupta et al [GKKS13a]. Recently this framework was used successfully to prove the first superpolynomial lower bounds for *regular formulas* [KSS13], and it seemed promising that such techniques could be used to prove lower bounds for more general classes such as *general* arithmetic formulas.

In this paper, we successfully apply this framework to prove the first superpolynomial lower bounds for homogeneous depth 4 circuits with bounded top fan-in. We prove our results via an *improved depth reduction* [1]. We also show that if the bound on the top fan-in is relaxed (even by a small amount), then *efficient depth reduction is not possible*. In particular this suggests that the method of depth reduction + shifted partial derivatives seems to be not powerful enough to prove lower bounds for (even) homogeneous arithmetic formulas. This result strengthens the results in [KSS13, FLMS13], and answers some of the main open questions posed in them.

We now outline the major results and the sequence of events that build up to the results of this paper. In the discussion in the rest of this section, we will refer to the class of circuits of depth 4 ($\Sigma\Pi\Sigma\Pi$ circuits) with bottom (product) fan-in bounded by t as $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits.

Depth Reduction: In a surprising result in 2008, Agrawal and Vinay [AV08] showed that any homogeneous polynomial which can be computed by a polynomial sized circuit of *arbitrary depth* can also be computed by subexponential sized homogeneous *depth* 4 $\Sigma\Pi\Sigma\Pi$ circuit. In other words, in order to prove superpolynomial (or even exponential) lower bounds for general arithmetic circuits, it suffices to prove exponential ($\exp(\Omega(n))$) lower bounds for just depth 4 arithmetic circuits[2]! In a follow up paper Koiran [Koi12] improved the parameters of this depth reduction theorem and showed that in order to prove superpolynomial lower bounds for general arithmetic circuits, it suffices to prove a lower bound of the form $\exp(\omega(\sqrt{n}\log^2 n))$ for

[1] By depth reduction, we really mean a reduction to homogeneous depth 4 circuits with bounded bottom fan-in. So, it makes sense to talk of depth reduction for depth 4 circuits.

[2] This result came as a big surprise, and indeed nothing like this is true in the Boolean world.

homogeneous $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuits (for polynomials of degree n). He also showed that to prove superpolynomial arithmetic formula lower bounds, it suffices to prove a slightly weaker lower bound of the form $\exp(\omega(\sqrt{n}\log n))$ for homogeneous $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuits. Tavenas [Tav13] further refined the parameters of Koiran's result and showed that for circuits lower bounds also, a lower bound of $\exp(\omega(\sqrt{n}\log n))$ would suffice. This sequence of works laid out an approach towards separating VNP from VP by just proving strong enough lower bounds for homogeneous circuits of depth 4. In a recent work along this line, Gupta, Kamath, Kayal and Saptharishi [GKKS13b] prove that strong enough lower bounds for *depth* 3 circuits suffice to show superpolynomial lower bounds for circuits of *arbitrary* depth, although in this case, we lose the property of homogeneity that was true for the reduction to depth 4. This loss in homogeneity seems quite severe, at least with respect to proving lower bounds, and we know only weak lower bounds for non-homogeneous depth 3 circuits [SW01]. (For the rest of the paper, this depth reduction to non-homogeneous depth 3 circuits will not be relevant.) More precisely, the results of Tavenas [Tav13] and Koiran [Koi12] state the following.

Theorem 1.1 ([Koi12, Tav13])**.** *Every polynomial size circuit of degree n in N variables can be transformed into an equivalent homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit with top fan-in*[3] *at most $\exp(O(\frac{n}{t}\log N))$.*

Depth 4 Lower Bounds and VNP vs VP: In light of the results of Agrawal-Vinay [AV08], Koiran [Koi12] and Tavenas [Tav13], proving lower bounds for homogeneous *depth* 4 circuits seems like an extremely promising direction to pursue in order to separate VNP from VP. In a breakthrough result in this direction, Gupta, Kamath, Kayal and Saptharishi [GKKS13a] proved that any homogeneous $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuit computing the permanent must have size (and top fan-in) $\exp(\sqrt{n})$. This was strengthened in a more recent work of Kayal, Saha and Saptharishi [KSS13], where it was shown that there is an explicit family of polynomials in VNP such that any homogeneous $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuit computing it must have size (and top fan-in) at least $\exp(\Omega(\sqrt{n}\log n))$. More precisely,

Theorem 1.2 ([GKKS13a, KSS13])**.** *For every n, there is an explicit family of polynomials in VNP in $N = \theta(n^2)$ variables and with degree $\theta(n)$ such that any homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit computing it must have top fan-in at least $\exp(\Omega(\frac{n}{t}\log N))$.*

The depth reduction results combined with the lower bounds for homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits is indeed a remarkable collection of results. As it stands, in order to separate VP from VNP, any small asymptotic improvement in the exponent on either the lower bound front or on the depth reduction front would be sufficient! In fact for any class of circuits $\mathcal{C}$ for which we can improve the depth reduction parameters of Theorem 1.1, we would get superpolynomial lower bounds for that class using Theorem 1.2.

Unfortunately, it seems that in general, we cannot hope for a better depth reduction. In a recent work, Fournier, Limaye, Malod and Srinivasan [FLMS13] gave an example of an explicit polynomial in VP (of degree n and in $N = n^{O(1)}$ variables) such that any homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit computing it must have top fan-in at least $\exp(\Omega(\frac{n}{t}\log N))$. This immediately implies that the depth reduction parameters in the result of Tavenas [Tav13] are *tight* for circuits. This observation, along with the fact that the hard polynomial used by Kayal et al [KSS13] has a shifted partial derivative span only a polynomial factor away from the maximum possible value suggests that the technique of depth reduction and shifted partial derivatives may not be strong

[3] For $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits where $t = \sqrt{n}$, observe that an upper bound of $\exp(O(\sqrt{n}\log N))$ on the top fan-in the circuit implies the same upper bound on size, since each product gate at the second layer computes a polynomial with at most $\exp(O(\sqrt{n}\log N))$ monomials. However for other values of t, the top fan-in bound will be the more relevant parameter for depth reduction.

enough to separate VNP from VP[4]. In a recent result, Chillara and Mukhopadhyay [CM13] gave a clean unified way of way of lower bounding the shifted partial derivative complexities of the polynomials considered by [KSS13, FLMS13].

Formula Lower Bounds: Even though improved depth reduction does not seem to be powerful enough to separate VNP from VP, it is conceivable that it could lead to superpolynomial lower bounds for other interesting classes, for instance homogeneous arithmetic formulas, or even general arithmetic formulas. This hope was further strengthened when Kayal et al [KSS13] used these precise ideas to prove superpolynomial lower bounds for a restricted class of formulas which they called *regular* formulas. (Regular formulas are formulas which have alternating sum and product layers. Moreover, for every fixed layer, the fan-ins of the gates in that layer are the same and the formal degree of the formula is at most a constant times the formal degree of the polynomial being computed.) Kayal et al proved their result by showing that one can reduce any polynomial size regular formula to a $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit (for a carefully chosen choice of t) of size asymptotically better in the exponent than the $\exp(\frac{n}{t}\log N)$ bound (which as we just discussed is known to be tight for circuits). This improvement in depth reduction immediately leads to superpolynomial lower bounds for regular formulas by using Theorem 1.2.

Removing the restriction on regularity and proving superpolynomial lower bounds for general formulas or even general homogeneous formulas would be a huge step forward - it would be by far the strongest and most natural class of arithmetic circuits for which we would be able to prove lower bounds, and it would represent a real breakthrough. The authors of the two papers [KSS13, FLMS13] left as a tantalizing open question whether formulas (or even homogeneous formulas) can have better depth reduction than circuits (such as is true for regular formulas). If true, this would imply superpolynomial lower bounds for (homogeneous) formulas. Indeed it seemed quite likely to be true since at the face of it, regularity of formulas of formulas did not seem like such a severe restriction at all (and indeed this was argued to be the case). Perhaps it could be also be true that every formula could be reduced to a regular formula with only a polynomial blow up in size. If so, the improved depth reduction for formulas (and hence the lower bounds) would follow from the improved depth reduction of regular formulas.

Thus to summarize, the main challenge that remained was to understand the limits of the techniques of depth reduction and shifted partial derivatives. In particular, are there any other interesting classes of circuits for which improved depth reduction is possible? Is improved depth reduction possible for arithmetic formulas?

2 Our results

In this paper we study the power and limitations of depth reduction for arithmetic formulas. We do this via studying depth reduction for depth 4 arithmetic circuits[5]. Let homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuits be the class of homogeneous depth 4 circuits with *top fan-in* bounded by r, and with *no restriction on the bottom fan-in*. This is a very natural class of circuits and is quite different in nature from $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits.

Our results are divided into two parts. In the first part we show the first superpolynomial lower bounds for homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuits when $r = o(\log n)$. The core of our result is an improved depth reduction result for these circuits. (As we pointed out, when we refer to 'depth reduction', we really mean a reduction to homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits. Thus it

[4]The reason this statement is not completely formal is that we still do not know know if the upper bounds on the shifted partial derivative measure for $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits is tight for all choices of derivatives and shifts, though the results of [FLMS13] and this paper show that they are indeed tight for many of the choices.

[5]Since depth 4 arithmetic circuits are also equivalent to depth 4 arithmetic formulas upto a polynomial blow up in size, we will use the term circuits and formulas interchangeably when referring to depth 4 circuits.

makes sense to talk about a depth reduction for $\Sigma\Pi\Sigma\Pi$ circuits as well.) $\Sigma\Pi\Sigma\Pi(r)$ circuits have received significant attention for the problems of polynomial identity testing and polynomial reconstruction [KMSV10, SV11, GKL12], however prior to this work there were no nontrivial lower bounds for this class of circuits for any value of $r \geq 2$.

In the second part we show that efficient depth reduction *is not possible* for homogeneous arithmetic formulas. We show this result by studying the very simple class of formulas given by homogeneous $\Sigma\Pi\Sigma\Pi(\log n)$ circuits. We show that for this class of circuits, improved depth reduction is not possible. This suggests that improved depth reduction (combined with the method of shifted partial derivatives) may unfortunately not be powerful enough to prove lower bounds for homogeneous $\Sigma\Pi\Sigma\Pi(\log n)$ circuits, and in particular not strong enough to prove lower bounds for homogeneous arithmetic formulas, answering the main open questions of [KSS13, FLMS13].

Informally, our main results are the following:

Main Theorem 1 (Informal): *There is an explicit family of polynomials in* VNP *of degree* n *in* $N = n^{O(1)}$ *variables such that for* $r = o(\log n)$*, any polynomial size homogeneous* $\Sigma\Pi\Sigma\Pi(r)$ *circuit computing it must have superpolynomial size.*

At the core of the result is the following "depth reduction" result:

Improved Depth Reduction (Slightly wishful):[6] *For* $r = o(\log n)$*, any polynomial size homogeneous* $\Sigma\Pi\Sigma\Pi(r)$ *circuit computing a polynomial of degree* n *and in* N *variables is equivalent to a homogeneous* $\Sigma\Pi\Sigma\Pi^{[t]}$ *circuit of size* $\exp\left(o(\frac{n}{t}\log N)\right)$ *for some choice of* t *such that* $\log^2 n \leq t \leq \epsilon n$.

Observe that the parameters of the depth reduction we obtain above improve upon the parameters of depth reduction given by [Koi12, Tav13].

We also show that when $r = \Omega(\log(n))$, depth reduction as above is no longer true.

Main Theorem 2 (Informal) *For* $r = \Omega(\log n)$*, there exists an explicit family of polynomials* $\{\mathcal{Q}_n\}_n$ *computed by a* $\mathsf{poly}(n)$ *size homogeneous* $\Sigma\Pi\Sigma\Pi(r)$ *circuit (and hence homogeneous formula) of degree* n *and in* N *variables, such that for every* t *such that* $\omega(\log n) \leq t \leq \epsilon n$*, any homogeneous* $\Sigma\Pi\Sigma\Pi^{[t]}$ *circuit computing* $\mathcal{Q}_n$ *must have top fan-in at least* $\exp\left(\Omega(\frac{n}{t}\log N)\right)$.

An immediate consequence of this result is that the depth reduction procedure of Tavenas [Tav13] is tight for homogeneous arithmetic formulas (strengthening the results of [FLMS13]).

At the core of our result is a *hierarchy* theorem for homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits which shows that homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits are a much richer class than homogeneous $\Sigma\Pi\Sigma\Pi^{[t/20]}$ circuits. We state this result more formally in Theorem 2.4.

It was shown in [KSS13] that any ABP (even non homogeneous) can be converted to a regular formula with a quasipolynomial blow up in size. If one could improve this transformation even slightly for formulas or even for homogeneous formulas, this would imply superpolynomial lower bounds for formulas/homogeneous formulas. Another consequence of our results is that such an improvement is not possible. We build upon the results of [KSS13] and show that the

[6] Indeed the above statement is not quite true, and our reduction turns out to be much more subtle. We do not depth reduce to a $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit, but one in which the sum of degrees of any $\epsilon n/t$ product gates at the bottom is at most ϵn. This is a more refined notion and a slightly more general class of circuits than $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits. We observe that the shifted partial derivative technique does not distinguish between these two kinds of circuits, and thus we are still able to obtain our lower bounds. Thus in spirit we still get depth reduction. In fact everywhere in this paper we could replace $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits with this slightly more general class of circuits, and none of the results would be affected.

conversion of general formulas to regular formulas must incur a quasipolynomial blow up in size.

Theorem (Conversion to Regular Formulas is Tight) *For $r = \Omega(\log n)$, there exists an explicit family of polynomials $\{\mathcal{Q}_n\}_n$ computed by a $\mathsf{poly}(n)$ size homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuit (and hence also homogeneous formula) of degree n and in $N = n^{O(1)}$ variables, such that any regular formula computing $\mathcal{Q}_n$ must have size $N^{\Omega(\log n)}$.*

In the sections below we formally state our results and elaborate on them in greater detail, as well as highlight some of the interesting corollaries of our proof techniques.

2.1 Lower bounds for $\Sigma\Pi\Sigma\Pi(r)$ circuits, $r = o(\log n)$

In the first part of the paper, we explore the limits of computation of depth 4 homogeneous circuits when the restriction for the bottom fan-in is removed. For the general model of (even homogeneous) $\Sigma\Pi\Sigma\Pi$ circuits, only extremely weak lower bounds seem to be known. Even PIT for $\Sigma\Pi\Sigma\Pi$ circuits is known only when the top fan-in is constant and the circuit is multilinear (in the multilinear case, the degree of the polynomials computed must anyway be bounded by the number of variables, and hence, multilinearity is a much bigger restriction than homogeneity[7]). The problem of showing lower bounds for depth 4 circuits with bounded top fan-in is hence a problem that is simpler than derandomizing PIT for the same model (at least in the black box model), and it seems to be the first crucial step in that direction. Moreover, even when the top fan-in is 2, prior to this work there were no lower bounds known. Unlike the class of depth 3 circuits with bounded top fan-in which cannot even compute all polynomials irrespective of the size of the circuit, the class of $\Sigma\Pi\Sigma\Pi(r)$ circuits is complete (even for $r = 1$). For more discussion on the completeness of this class, see Appendix A.

We consider homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuits, which are depth 4 homogeneous circuits whose *top fan-in* is bounded by r. When r is a constant we prove exponential lower bounds[8] for the class of $\Sigma\Pi\Sigma\Pi(r)$ circuits, and for any $r = o(\log n)$ we show superpolynomial lower bounds for $\Sigma\Pi\Sigma\Pi(r)$ circuits[9]. In particular, we prove the following theorem:

Theorem 2.1. *There exists an explicit family of polynomials in VNP, $\{NW_n\}_n$, such that for each n, NW_n has degree $\theta(n)$, and number of variables $\theta(n^2)$ and such that the following holds: Let C be a homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuit that computes NW_n. Let s be the size of C. Then*

$$s \geq \exp\left(n^{\Omega(1/r)} \log n\right).$$

Prior to this result, we are not aware of any such lower bounds for depth 4 circuits even when the top fan-in r equals 2.

Lower bounds for homogeneous $\Sigma\Pi\Sigma\Pi^*$ circuits: Another class of circuits we are able to prove a lower bound for is the class of depth 4 circuits where each product at the second layer (from the top) has the same degree sequence of incoming polynomials, and there is no restriction on the top fan-in.

For any degree sequence $\mathcal{D} = D_1, D_2, \ldots, D_k$ of non-negative integers such that $\sum D_i = n$, we study the class of homogeneous $\Sigma\Pi^{\mathcal{D}}\Sigma\Pi$ circuits, which are homogeneous circuits where

[7] In all the results of this paper, the restriction of homogeneity can be replaced by the restriction that all gates in the circuit compute polynomials of degree at most n.

[8] In the rest of the paper, by exponential lower bound we will mean a lower bound of the form 2^{n^ϵ} for some constant ϵ.

[9] It is important to observe that the reduction of a polynomial sized homogeneous $\Sigma\Pi\Sigma\Pi$ circuit with arbitrary bottom fan-in to a homogeneous $\Sigma\Pi\Sigma\Pi$ circuit with bounded bottom fan-in as given by the results of [AV08, Koi12] can lead to circuits of size $\exp(\Omega(n/t)\log n)$ and so Theorem 1.2 does not imply any nontrivial lower bounds for it.

each Π gate at the second layer is restricted to having its inputs be polynomials whose sequence of degrees is precisely $\mathcal{D}$. We show that for *every* degree sequence $\mathcal{D}$, any $\Sigma\Pi^{\mathcal{D}}\Sigma\Pi$ circuit computing NW_n (an explicit family of polynomials in VNP) must have size at least $\exp(n^{\epsilon})$, for some fixed absolute constant ϵ independent of $\mathcal{D}$. In particular, let the class of $\Sigma\Pi\Sigma\Pi^*$ circuits be the union of the classes of $\Sigma\Pi^{\mathcal{D}}\Sigma\Pi$ for all $\mathcal{D}$. Then our lower bounds hold for homogeneous $\Sigma\Pi\Sigma\Pi^*$ circuits as well.

Theorem 2.2. *There exists an explicit family of polynomials in VNP, $\{NW_n\}_n$, such that for each n, NW_n has degree $\theta(n)$, and number of variables $\theta(n^2)$ and such that the following holds: Let C be a homogeneous $\Sigma\Pi\Sigma\Pi^*$ circuit that computes NW_n. Let s be the size of C. Then*

$$s \geq \exp(n^{\epsilon}),$$

for some fixed absolute constant $\epsilon > 0$.

2.2 Depth reduction is tight for $\Sigma\Pi\Sigma\Pi(r)$ circuits, $r = \Omega(\log n)$

The main question that was left open by both the works of [KSS13] and [FLMS13] was to understand whether an improved depth reduction was possible for general (homogeneous) arithmetic formulas.

In particular, the following tantalizing questions naturally emerge and were left as open questions by the works of [KSS13] and [FLMS13].

- Can the depth reduction by Koiran and Tavenas [Koi12, Tav13] be improved for formulas: In other words, can one show that for every polynomial of degree n and in $N = n^{O(1)}$ variables which has a polynomial sized (homogeneous) formula, it can be reduced to a $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuit of size $N^{o(\sqrt{n})}$?
- Can every homogeneous arithmetic formula be converted to a regular formula with only a polynomial blow up in its size?

A positive answer to any of the above questions would suffice in proving superpolynomial lower bounds for general homogeneous arithmetic formulas. We settle both the questions and show that unfortunately neither is true.

We settle these questions by constructing a an explicit family of polynomials $\{\mathcal{Q}_n\}_n$, where $\mathcal{Q}_n$ is a polynomial in $\theta(n^2)$ variables and is of degree $\theta(n)$, such that for each n, $\mathcal{Q}_n$ can be computed by a *polynomial sized homogeneous formula*, but any $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuit computing $\mathcal{Q}_n$ must have top fan-in at least $2^{\Omega(\sqrt{n}\log N)}$. Moreover $\mathcal{Q}_n$ is computed by a polynomial size homogeneous $\Sigma\Pi\Sigma\Pi(r)$ formula for $r = \theta(\log n)$. More formally, we prove the following theorem.

Theorem 2.3 (Depth reduction is tight for formulas)**.** *There exists an explicit family of polynomials $\{\mathcal{Q}_n\}_n$ and an absolute constant $\epsilon > 0$ such that $\mathcal{Q}_n$ is of degree $\theta(n)$, in $N = \theta(n^2)$ variables, and computed by a $\mathsf{poly}(n)$ size homogeneous $\Sigma\Pi\Sigma\Pi(\log n)$ circuit (in particular a homogeneous arithmetic formula); and for every t such that $\omega(\log n) \leq t \leq \epsilon n$, any $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit computing $\mathcal{Q}_n$ must have top fan-in at least $\exp\left(\Omega(\frac{n}{t}\log N)\right)$.*

The above theorem follows by an interpolation argument applied to a *hierarchy* theorem for $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits, which is the heart of our argument. The hierarchy theorem shows that by increasing the bound on the bottom fan-in of $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits even slightly, we get a much richer class of arithmetic circuits. We believe this is an interesting result in its own right.

Theorem 2.4 (Hierarchy theorem for $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits)**.** *There exists an absolute constant $\epsilon > 0$ such that for every t with $\omega(\log n) \leq t \leq \epsilon n$, there exists an explicit family of polynomials $\{\mathcal{P}_{t,n}\}_n$ such that $\mathcal{P}_{t,n}$ is of degree n, has $N = n^2$ variables, and is computed by a $\mathsf{poly}(n)$ size homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit, and for every t' s.t. $t' < t/20$, any homogeneous $\Sigma\Pi\Sigma\Pi^{[t']}$ circuit computing $\mathcal{P}_{t,n}$ must have top fan-in at least $\exp\left(\Omega(\frac{n}{t}\log N)\right)$.*

These results immediately imply that Koiran's and Tavenas' depth reduction [Koi12, Tav13] is tight for formulas, for all but a small number of choices of the bottom fan-in. In particular, it is tight for the case where the bottom fan-in is bounded by $\sqrt{n}$. Interestingly enough, the polynomial size formulas computing $\mathcal{Q}_n$ are of depth 4. In fact, they are a sum of $O(\log n)$ *regular* homogeneous formulas of depth 4.

A corollary of our results is that any conversion of a general (homogeneous) formula to a regular formula must incur a quasipolynomial blow up in size. It was shown in [KSS13] that any algebraic branching program can be converted to a regular formula with a quasipolynomial blow up in size. Since it is widely believed that formulas are much weaker than ABPs, it was conjectured that formulas, or homogeneous formulas might have a more efficient conversion (which would suffice in proving superpolynomial lower bounds for homogeneous formulas!). We show however that this is not true. Combining our results with the result of [KSS13], we obtain the following (tight) lower bound for converting homogeneous formulas to regular formulas.

Theorem 2.5 (Lower bounds for reduction to Regular Formulas)**.** *There exists an explicit family of polynomials $\{\mathcal{Q}_n\}_n$ and an absolute constant $\epsilon > 0$ such that $\mathcal{Q}_n$ is of degree $\theta(n)$, in $N = \theta(n^2)$ variables, and computed by a $\mathsf{poly}(n)$ size homogeneous $\Sigma\Pi\Sigma\Pi(\log n)$ circuit (in particular a homogeneous arithmetic formula); and any regular formula computing $\mathcal{Q}_n$ must have size at least $N^{\Omega(\log n)}$.*

Organization of the paper: The rest of the paper is organized as follows. In Section 3, we introduce some preliminary notions about circuits and introduce notation which we will use in the rest of the paper. In Section 4 we prove our lower bound for homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuits when $r = o(\log n)$. In Section 5, we show that depth reduction is tight for homogeneous arithmetic formulas by showing it is tight for homogeneous $\Sigma\Pi\Sigma\Pi(\Omega(\log n))$ circuits. We conclude with some discussion and open problems in Section 6.

3 Preliminaries

Arithmetic Circuits: An arithmetic circuit over a field $\mathbb{F}$ and a set of variables $\overline{x} = \{x_1, x_2, \ldots, x_n\}$ is a directed acyclic graph such that every node in the graph is labelled by either a field element or a variable in $\overline{x}$ or one of the field operations $+, \times$. There could be one or more nodes with fan-out zero, called the output gates of the circuit. The nodes with fan-in zero indexed by the variables or field elements are called the leaf nodes. In this paper, unless otherwise mentioned, we will assume that there is unique output gate. We will refer to the length of the longest path from an output node to a leaf node as the depth of the circuit. A circuit is said to be homogeneous if the polynomial computed at every node in the circuit is a homogeneous polynomial. A circuit is said to be a formula if the underlying undirected graph is a tree. By a circuit of depth 4, we will refer to a circuit of the form $\Sigma\Pi\Sigma\Pi$, where the output gate is a + gate and all the nodes at a distance 2 from it are also labelled by +, and the remaining gates are labelled by $\times$. Observe that a *depth 4 circuit* can be converted into a *depth 4 formula* with only a polynomial blow up in size. We will therefore, use the term formula or circuit for a depth 4 circuit interchangeably in this paper. A homogeneous polynomial $P(\overline{x})$ of degree d computed by a depth 4 circuit is of the form

$$P(\overline{x}) = \sum_{i=1}^{r} \prod_{j=1}^{d_i} Q_{i,j}(\overline{x}) \tag{1}$$

Based upon this definition, we will now define the specific restrictions of depth 4 circuits that we study in this paper.

Homogeneous $\Sigma\Pi\Sigma\Pi^{[a]}$ circuits and homogeneous $\Sigma\Pi^{[b]}\Sigma\Pi^{[a]}$ circuits: The depth 4 $\Sigma\Pi\Sigma\Pi$ circuit in Equation 1, is said to be a $\Sigma\Pi^{[b]}\Sigma\Pi^{[a]}$ circuit, if each $Q_{i,j}(\overline{x})$ is a polynomial of degree at most a and each d_i is at most b. The depth 4 $\Sigma\Pi\Sigma\Pi$ circuit in Equation 1, is said to be a $\Sigma\Pi\Sigma\Pi^{[a]}$ circuit, if each $Q_{i,j}(\overline{x})$ is a polynomial of degree at most a. In this case we say that the *botton fan-in is bounded by a*. If the circuit is homogeneous, then we can assume without loss of generality that for each i, $\prod_{j=1}^{d_i} Q_{i,j}(\overline{x})$ is a polynomial of degree exactly d.

Observe that, for each i, by grouping together and multiplying out some of the $Q_{i,j}$, we can transform a homogeneous $\Sigma\Pi\Sigma\Pi^{[a]}$ circuit into a homogeneous $\Sigma\Pi^{[b]}\Sigma\Pi^{[a]}$ circuit, where $b = O(\frac{d}{a})$. This operation of grouping together and multiplying would increase the size of the resulting circuit, but notice that it does not affect the top fan-in of the circuit. Thus lower bounds on the top fan-in for $\Sigma\Pi^{[O(b)]}\Sigma\Pi^{[a]}$ circuits impy the same lower bounds on the top fan-in of $\Sigma\Pi\Sigma\Pi^{[a]}$ circuits.

Homogeneous $\Sigma\Pi\Sigma\Pi(r)$ Circuits: The depth 4 $\Sigma\Pi\Sigma\Pi$ circuit in Equation 1, is said to be a $\Sigma\Pi\Sigma\Pi(r)$ circuit if the fan-in of the summation(top fan-in) is bounded by r. Observe that there is no restriction on the bottom fan-in except that implied by the restriction of homogeneity.

For each $i \in [r]$, the product $P_i = \prod_{j=1}^{d_i} Q_{ij}$ is said to be computed by the product gate i. Therefore, $P = \sum_{i=1}^{r} P_i$. Here for every i and j, Q_{ij} is an n variate homogeneous polynomial being computed by a $\Sigma\Pi$ circuit. The homogeneity restriction on C implies that for every product gate i,

$$\deg(P) = d = \sum_{j=1}^{d_i} \deg(Q_{ij}) \tag{2}$$

With every product gate $i \in [r]$, we can associate a multiset (D_i, m_i), where

$$D_i = \{\deg(Q_{ij}) : j \in [d_i]\} \tag{3}$$

and m_i is a map from D_i to $\mathbb{N}$, which assigns to every element l in D_i, the number of $j \in [d_i]$ such that Q_{ij} has degree equal to l. For a homogeneous depth 4 circuit, computing a degree d polynomial, Equation 2 can be rewritten as

$$\deg(P) = d = \sum_{l \in D_i} l \times m_i(l) \tag{4}$$

for each i in $[r]$. $\Sigma\Pi\Sigma\Pi(r)$ circuits for which the multiset (D_i, m_i) is the same for every product gate $i \in [r]$, are said to be $\Sigma\Pi\Sigma\Pi^*$ circuits.

Regular Formula: The notion of regular formulas was introduced in [KSS13], where super-polynomial lower bounds for this model were proved.

Definition 3.1. *A formula computing a degree d polynomial in n variables is said to be regular, if it satisfies the following conditions:*

1. *It has alternating layers of sum and product gates.*
2. *All gates in a single layer have the same fan-in.*
3. *The formal degree of the formula is at most some constant multiple of the degree of the polynomial being computed.*

Shifted Partial Derivatives: The complexity measure used in showing lower bounds in this paper is the dimension of the shifted partial derivatives introduced in [Kay12] and used in [FLMS13], [GKKS13a] and [KSS13]. For a field $\mathbb{F}$, an n variate polynomial $P \in \mathbb{F}[\overline{x}]$ and a positive integer k, we denote by $\partial^{=k}P$, the set of all partial derivatives of order equal to k of P. For a polynomial P and a monomial m, we denote by $\partial_m P$ the partial derivative of P with respect to m. Our proof uses the notion of shifted partial derivatives of a polynomial as defined below.

Definition 3.2 ([GKKS13a]). *For an n variate polynomial $P \in \mathbb{F}[\overline{x}]$ and integers $k, \ell \geq 0$, the space of ℓ shifted k^{th} order partial derivatives of P is defined as*

$$\langle \partial^{=k} P \rangle_{\leq \ell} \stackrel{def}{=} \mathbb{F}\text{-}span\{ \prod_{i \in [n]} {x_i}^{j_i} \cdot g : \sum_{i \in [n]} j_i \leq \ell, g \in \partial^{=k} P\} \tag{5}$$

Nisan-Wigderson Polynomials: [10] We will now define the family of polynomials introduced in [KSS13]. These polynomials were used to prove improved lower bounds for homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits. (In an earlier paper by Gupta et al [GKKS13a], a slightly weaker lower bound was shown for the permanent.) For a prime power n, let $\mathbb{F}_n$ be a field of size n. For the set of n^2 variables $\{x_{i,j} : i, j \in [n]\}$ and $t \in [n]$, we define the degree n homogeneous polynomial $NW_{t,n}$ as

$$NW_{t,n} = \sum_{\substack{f(z) \in \mathbb{F}_n[z] \\ deg(f) < \lfloor \frac{n}{2t} \rfloor}} \prod_{i \in [n]} x_{i,f(i)}$$

Clearly, for every n and t, $NW_{t,n}$ is in VNP. The Nisan-Wigderson polynomial family $\{NW_n\}_n$ is a family of polynomials in VNP such that NW_n is a polynomial of degree $n+1$ in n^2+n variables $\{x_{i,j} : i, j \in [n]\} \cup \{y_i : i \in [n]\}$ defined as follows

$$NW_n = \sum_{i=1}^{n} y_i \cdot NW_{i,n}$$

For more on arithmetic circuits, we refer the interested reader to the survey by Shpilka and Yehudayoff [SY10].

4 Lower bounds for $\Sigma\Pi\Sigma\Pi(r)$ circuits, $r = o(\log n)$

In earlier works by Gupta et al [GKKS13a] and Kayal et al [KSS13], exponential lower bounds were shown for the class of homogeneous $\Sigma\Pi\Sigma\Pi$ circuits with bounded bottom fan-in. Without the restriction on the bottom fan-in, basically no lower bounds for $\Sigma\Pi\Sigma\Pi$ circuits are known. In this section we prove the first super polynomial lower bounds for homogeneous$\Sigma\Pi\Sigma\Pi$ circuits with bounded top fan-in. The main technical core of our result is a *depth reduction* result, very similar in spirit to those by Koiran [Koi12] and Tavenas [Tav13]. By exploiting the structure of these circuits, we show how to get *improved depth reduction* for these circuits. The proof of our depth reduction is quite different from that of [Koi12, Tav13], and is somewhat subtle. We don't reduce to a $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit, but a related and slightly more general class of circuits where instead of requiring an absolute bound on the bottom fan-in, we just require that in some sort of average sense, the bottom fan-in is small. In particular we reduce to a $\Sigma\Pi\Sigma\Pi$ circuit in which the sum of degrees of any $\epsilon n/t$ product gates at the bottom is at most ϵn. This is a more refined notion and a slightly more general class of circuits than $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits. We observe that the shifted partial derivative technique does not distinguish between these two kinds of circuits, and thus we are still able to use a variant of Theorem 1.2 to obtain our lower bounds. Thus in spirit we still get depth reduction. In fact everywhere in this paper we could replace $\Sigma\Pi\Sigma\Pi^{[t]}$ circuits with this slightly more general class of circuits, and none of the results would be affected.

There seem to be two main obstacles in extending the lower bounds of [GKKS13a, KSS13] to lower bounds for general depth 4 homogeneous $\Sigma\Pi\Sigma\Pi$ circuits. The lower bounds in [GKKS13a, KSS13] work only when the degrees of *all* polynomials feeding into the product gate at the second

[10] This definition is a slight variant of the definition in [KSS13]. We modify the definition to ensure homogeneity. It is not hard to see that all the bounds proved for the original polynomial in [KSS13] hold for this variant also.

layer are small (in other words, the bottom fan in is small), say $\leq \sqrt{n}$. It is easy to see that if the degrees of all polynomials feeding into the product gate at the second layer is large (i.e. the bottom fan-in of all the gates is large), say $\geq \sqrt{n}$, then for sparsity reasons and simple monomial counting, it is easy to obtain exponential lower bounds. The first obstacle is to handle the case when the degrees of some of the polynomials is small and for some of them it is large. For instance fix any arbitrary sequence $\mathcal{D}$ of degrees summing to n, and assume that the polynomials feeding into each product gate at the second from top layer have their degrees coming from this sequence. Is it still possible to obtain exponential lower bounds? The second obstacle to extending the results from [GKKS13a] is to find a way to combine the lower bounds for all these various cases into a common lower bound for the case when the circuit is composed of product gates of different kinds. For instance we know lower bounds when all product gates at the second layer have small incoming degrees and when all product gates have large incoming degrees. However we do not know how to combine these lower bounds into a single lower bound when the circuit is the sum of two circuits, one of the low degree kind, and one of the high degree kind. In this paper we show how to resolve the first obstacle. Moreover when the top fan-in is $o(\log n)$, the second obstacle turns out to not be a problem either.

Proof Overview: Most lower bounds for arithmetic circuits proceed by identifying some kind of "progress measure", and show that for any given circuit in a circuit class, the measure is small if the size of the circuit is small, whereas for the polynomial one is trying to compute (for instance the permanent), the measure is large. In the results by Gupta et al. [GKKS13a] and Kayal et al [KSS13], the progress measure used is the dimension of the ℓ shifted k^{th} order partial derivative $\dim(\langle \partial^{=k} P \rangle_{\leq \ell})$, for a suitable choice of ℓ and k. It is shown that every small depth 4 circuit with bounded bottom fan-in has small $\dim(\langle \partial^{=k} P \rangle_{\leq \ell})$ compared to that of an explicit polynomial in VNP, the NW_n polynomial. Thus if a depth 4 circuit with bounded bottom fan-in must compute NW_n, then it must be large. More precisely it is shown that every product gate $Q_i = \prod_{j=1}^{d} Q_{ij}$ has $\dim(\langle \partial^{=k} P \rangle_{\leq \ell})$ much smaller than that of the permanent, provided the degrees of the Q_{ij} are small. This is the core of the argument. Combined with the sub-additivity of $\dim(\langle \partial^{=k} P \rangle_{\leq \ell})$, the result easily follows.

Our proof builds upon the results of [GKKS13a] and [KSS13], and combines the use of the progress measure $\dim(\langle \partial^{=k} P \rangle_{\leq \ell})$ with the notion of "sparsity" to prove our improved depth reduction and the lower bounds for the polynomial family $\{NW_n\}_n$. Suppose $C = \sum_{i=1}^{r} \prod_{j=1}^{d_i} Q_{ij}$ is a homogeneous $\Sigma\Pi\Sigma\Pi$ circuit computing NW_n. If all the Q_{ij} had low degree, then the results of [GKKS13a] and [KSS13] give exponential lower bounds for the top fan-in of C. Also in the extreme case where all the Q_{ij} have high degree, then since C is homogeneous, the number of Q_{ij} per product gate $Q_i = \prod_{j=1}^{d} Q_{ij}$ must be small, and hence their product cannot have too many monomials[11]. If the number of monomials is too few, we would not even be able to get all the monomials in NW_n. In general, of course there might be some high degree and some low degree polynomials, and we attempt to interpolate between the two settings to obtain our results.

For each product gate $Q_i = \prod_{j=1}^{d_i} Q_{ij}$, recall that each Q_{ij} is a homogeneous polynomial of degree d_{ij} (say), and $\sum_{j=1}^{d} d_{ij} = n$. If the size of the circuit is at most s, then each Q_{ij} has at most s monomials. We decompose each product gate into its inputs Q_{ij} of *high degree* (those of degree $\geq t$) and its inputs Q_{ij} of *low degree* (those of degree <t). Observe that there cannot be too many (greater than n/t) high degree polynomials Q_{ij} as otherwise their product would have degree exceeding n. Thus the product of all the high degree Q_{ij} cannot have more than $s^{n/t}$ monomials. Let H be the product of the the high degree Q_{ij}, and L be the product of the low degree Q_{ij}. Then, by writing out H as a sum of monomials ($H = \sum_k h_k$) and multiplying each

[11]The number of monomials in each Q_{ij} is a most the size of the circuit.

monomial h_k with L, we can expand out Q as $\sum_k h_k \cdot L$. Note that L is a product of low degree polynomials. Also, each h_k is a monomial and hence a product of degree 1 polynomials. Thus we have expressed Q as a $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit, where now all the product gates multiply polynomials of degree at most t.

The hope at this point would be to apply this transformation to all the product gates and then possibly apply the result in [KSS13] to obtain a lower bound. The trouble with this argument is that under the transformation described, the top fan-in of the original circuit might blow up by a factor equaling the number of monomials in H, which could be nearly as large as $s^{n/t}$. With this loss in parameters, the bound given by the [KSS13] result gives nothing nontrivial. Thus in general one cannot choose an absolute threshold t and for all product gates choose degrees greater than t to be the high degree polynomials and the ones below t to the the low degree polynomials.

What we show is that by examining the degrees of the polynomials feeding into the product gates, one can carefully choose a threshold t that works for each product gate individually, though it might not be the same threshold for all gates. It turns out that this threshold that we find is purely a function of the degree sequence $\mathcal{D}$ of the product gate. Thus if all product gates have the *same* degree sequence, i.e. we have a $\Sigma\Pi^{\mathcal{D}}\Sigma\Pi$ circuit, then we obtain exponential lower bounds. However, for general $\Sigma\Pi\Sigma\Pi$ circuits it can be a problem, since if the threshold is different for different gates, we do not have any one single progress measure that works for all gates and thus for the entire circuit. However we are still able to show that for each gate, only very few thresholds are "bad", and when the top fan-in is $o(\log n)$, then we show there is a single threshold that will work for all gates to give superpolynomial lower bounds.

4.1 Proof of Theorem 2.1

In this subsection, we will present the proof of Theorem 2.1. Let us consider a homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuit C of size s computing NW_n. From Equation 1, this implies that

$$NW_n = \sum_{i=1}^{r} \prod_{j=1}^{d_i} Q_{ij} \tag{6}$$

where for every value of i and j, Q_{ij} is a homogeneous polynomial being computed by a subcircuit of depth 2 of C. Observe that Q_{ij} is being computed by a $\Sigma\Pi$ circuit and hence, the number of monomials with nonzero coefficients in a sum of products expansion of Q_{ij} will be at most the size of C. In other words, Q_{ij} is s sparse for each $i \in [r]$ and $j \in [d_i]$. Without loss of generality, we will assume that for every $i \in [r]$, $d_i = n$, since if $d_i < n$ for any i, we can always make it equal to n adding dummy polynomials that are the constant 1.

Let us now consider the polynomial computed at a product gate near the top of C. It is of the form $Q = \prod_{i\in[n]} Q_i$. Let us also assume without loss of generality that the Q_i are arranged in non-increasing order of their degrees. The idea of the proof, as described in the overview, would be to decompose the Q_i into *high degree* and *low degree* parts and then multiply out all the *high degree* parts and count on their sparsity to show that the product does not blow up the dimension of the space of shifted partial derivatives by too much. We will then use the following lemma implicit in the work of [GKKS13a], to obtain our bounds.

Lemma 4.1 (Implicit in [GKKS13a]). *Let $P = \prod_{i=1}^{d} \tilde{P}_i$ be a polynomial in N variables such that the sum of the degrees of any k of these d polynomials $\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_d$ is at most D. Then, for every integer $\ell \geq 0$,*

$$\dim(\langle \partial^{=k} f \rangle_{\leq \ell}) \leq \binom{d+k-1}{k} \binom{N+D-k+\ell}{N}.$$

Proof. The proof of the lemma is exactly the same calculation as in [GKKS13a]. We replace their bound of tk (which for them was the sum of degrees of k polynomials of degree at most t), by our bound of D. □

The following lemma is the core of our argument.

Lemma 4.2. *Let $Q = \prod_{j\in[n]} Q_i$ be a depth 3 $\Pi\Sigma\Pi$ homogeneous circuit of degree n in N variables, where each Q_i has at most s monomials. Let $0 < \epsilon < 1$ be any small constant. Consider $k = n^{i/m}$, for $1 \le i \le m$ and any integer $\ell \ge 0$. Then for all but $1/\epsilon$ choices of i ,*

$$\dim(\langle \partial^{=k} Q\rangle_{\le \ell}) \le s^{k\cdot n^{-1/m}} \cdot \binom{k/\epsilon + k - 1}{k}\binom{N + 4\epsilon n - k + \ell}{N}.$$

Proof. Since the Q_i's are arranged in order of decreasing degree, Q_1 has highest degree and $\mathcal{Q}_n$ has the smallest degree.

For $1 \le i \le m$, let $S_i = \{Q_j | j \le n^{i/m}\}$ be the set of the first $n^{i/m}$ of the Q_j's. For each i, we will sum the degrees of the Q_j's in $S_i \setminus S_{i-1}$. Let

$$D_i = \sum_{j \text{ s.t.} Q_j \in S_i \setminus S_{i-1}} \deg(Q_j).$$

Then $\sum_{i=1}^{m} D_i = n$. Thus there are at most $1/\epsilon$ choices of i for which $D_i \ge \epsilon n$. We will show that for all other choices of i, for $k = n^{i/m}$ and any integer $\ell \ge 0$, $\dim(\langle \partial^{=k} Q\rangle_{\le \ell}) \le s^{k\cdot n^{-1/m}} \cdot \binom{k/\epsilon+k-1}{k}\binom{N+4\epsilon n-k+\ell}{N}$.

Let us fix i such that $D_i \le \epsilon n$. We will split up the various Q_j's into those that are in S_{i-1} and those that are not. For those Q_j in S_{i-1}, we will exploit the fact that there aren't too many of them and they each have at most s monomials, to show that they do not affect the dimension of shifted partial derivatives by too much. For the rest of the Q_j we will take advantage of the fact that their degrees are not too large, and hence the sum of degrees of any k of them is small, and thus we will be able to bound the span of shifted partial derivatives of their product using the argument presented in [GKKS13a].

Let $H = \prod_{Q_j \in S_{i-1}} Q_j$, and let $Q_{\bar{H}} = Q/H$. Since each Q_i has at most s monomials, thus H has at most $s^{n^{(i-1)/m}}$ monomials. Hence we can express the polynomial Q as the sum of at most $s^{n^{(i-1)/m}}$ polynomials $P_1, P_2, \ldots P_u$, where each of the polynomials is the product of some monomial (from H), and the product of all the Q_j that are not in S_{i-1} (i.e. those in $Q_{\bar{H}}$) .

We will show that for each P_j, $1 \le j \le u$, for $k = n^{i/m}$,

$$\dim(\langle \partial^{=k} P_j\rangle_{\le \ell}) \le \binom{k/\epsilon + k - 1}{k}\binom{N + 4\epsilon n - k + \ell}{N}.$$

Since u is at most the number of monomials in H, thus $u \le s^{n^{(i-1)/m}} = s^{k\cdot n^{-1/m}}$. Since $Q = \sum_{j\in[u]} P_j$, the sub-additivity of $\dim(\langle \partial^{=k}\rangle_{\le \ell})$ will imply that

$$\dim(\langle \partial^{=k} Q\rangle_{\le \ell}) \le s^{k\cdot n^{-1/m}} \cdot \binom{k/\epsilon + k - 1}{k}\binom{N + 4\epsilon n - k + \ell}{N}.$$

Let us focus our attention on any one of these polynomials P_j, and call it P.

Then $P = h\cdot Q_{\bar{H}} = h\cdot\prod_{j\ge n^{(i-1)/m}+1} Q_j$, where h is a monomial of H and can be thus written as a product of degree one homogeneous polynomials. Let us rename the degree 1 polynomials in h and the different Q_j dividing $Q_{\bar{H}}$, so that $P = \hat{P}_1\hat{P}_2\cdots\hat{P}_\ell$.

Consider all the polynomials $\hat{P}_i$ dividing P which have degree at most $\epsilon n/k$, and group them together and multiply them so that each of the grouped polynomials now has degree at least

$\epsilon n/k$ and at most $2\epsilon n/k$. Clearly this can be done. Call the new set of polynomials (the grouped ones and the ones that had degree at least $\epsilon n/k$ to start out with) $\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_d$. Since the sum of their degrees is at most n, thus the total number d of these polynomials is at most k/ϵ.

Proposition 4.3. *The sum of the degrees of any k of these d polynomials $\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_d$ is at most $4\epsilon n$.*

Proof. Out of the k polynomials, we see what fraction lie among the "grouped" polynomials, and what lie among the original ungrouped polynomials. Recall that by the choice of i, and setting $k = n^{\frac{i}{m}}$, the sum of degrees of any $k - kn^{\frac{-1}{m}}$ of the $\hat{P}_i$ dividing P was at most ϵn. Since m is $o(\log n)$, the sum of the degrees of any k of them will be at most $2\epsilon n$. Thus, the contribution from the original ungrouped polynomials is at most $2\epsilon n$. Also, the contribution from the grouped polynomials can be at most $2\epsilon n$ since there are at most k of them, and each has degree at most $2\epsilon n/k$. Thus the total sum of degrees is at most $4\epsilon n$. □

Thus, $P = \prod_{i=1}^{d} \tilde{P}_i$ is a polynomial in N variables such that the sum of the degrees of any k of the d polynomials $\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_d$ is at most $D = 4\epsilon n$. Recall also that $d \leq k/\epsilon$. Hence, by Lemma 4.1, for any integer $\ell \geq 0$,

$$\dim(\langle \partial^{=k} P \rangle_{\leq \ell}) \leq \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N}.$$

□

Theorem 4.4. *Let C be a homogeneous $\Sigma\Pi\Sigma\Pi(r)$ circuit in N variables, of size s and of degree at most n. Then for all constants ϵ, with $0 < \epsilon < 1$, there exists a choice of i, with $1 \leq i \leq 2r/\epsilon$, such that for $k = n^{\epsilon i/2r}$, and for all integers $\ell \geq 0$,*

$$\dim(\langle \partial^{=k} C \rangle_{\leq \ell}) \leq r \cdot s^{k \cdot n^{-\epsilon/2r}} \cdot \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N}.$$

Proof. Let $m = 2r/\epsilon$. Let $C = \sum_{j=1}^{r} Q_j$. Let $i \in [m]$.

Then for each Q_j, by Lemma 4.2, for all but $1/\epsilon$ choices of i, for $k = n^{i/m}$,

$$\dim(\langle \partial^{=k} Q_j \rangle_{\leq \ell}) \leq s^{k \cdot n^{-1/m}} \cdot \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N}.$$

Hence for each Q_j we get at most $1/\epsilon$ choices of i that may not work to get the bound above, and we call those choices "bad" for Q_j. We call the rest of the choices "good" for Q_j. Thus by the union bound there are at most r/ϵ choices of i that are bad for some Q_j. Since $m > r/\epsilon$, thus there is a choice of $i \in [m]$ that is good for every Q_j.

Thus for any integer $\ell \geq 0$ and $k = n^{i/m}$, for all $j \in [r]$,

$$\dim(\langle \partial^{=k} Q_j \rangle_{\leq \ell}) \leq s^{k \cdot n^{-1/m}} \cdot \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N}.$$

Hence

$$\dim(\langle \partial^{=k} C \rangle_{\leq \ell}) \leq r \cdot s^{k \cdot n^{-1/m}} \cdot \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N}.$$

□

We can observe that the choice of the threshold and k for every product gate just depends upon the multiset of the degrees associated with the input feeding into it. In particular, if we start with a $\Sigma\Pi\Sigma\Pi^*$ circuit, then the value of the threshold and k that works for one product gate also works for the circuit in general. Hence, we have the following theorem which gives us an upper bound on the dimension of the shifted partial derivative space of a $\Sigma\Pi\Sigma\Pi^*$ circuit.

Theorem 4.5. *Let C be a homogeneous $\Sigma\Pi\Sigma\Pi^*$ circuit in N variables, of size s, top fan-in r and of degree at most n. Then for all constants ϵ, with $0 < \epsilon < 1$, there exists a choice of i, with $1 \leq i \leq m$, where $m = 1/\epsilon + 1$ such that for $k = n^{i/m}$, and for all integers $\ell \geq 0$,*

$$\dim(\langle \partial^{=k} C \rangle_{\leq \ell}) \leq r \cdot s^{k \cdot n^{-1/m}} \cdot \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N}.$$

It is important to note the difference between the bounds in Theorem 4.4 and Theorem 4.5. In Theorem 4.5, the exponent of s is independent of the top fan-in r as m is a constant.

In order to complete the proof now, we will look at the shifted partial derivative complexity of the circuit as well as of the polynomial NW_n under restrictions where all the variables $\{y_1, y_2, \ldots, y_n\}$ are set to constants. The partial derivatives as well the final shifts are just taken with respect to monomials in the n^2 variables $\{x_{1,1}, x_{1,2}, \ldots, x_{n,n}\}$. The following theorem tells us that under some restrictions of this type, NW_n has large complexity. This happens because under the restriction where $y_t = 1$ and $y_j = 0$ for every $j \neq t$, we obtain $NW_{t,n}$ from NW_n.

Theorem 4.6 ([KSS13])**.** *For any integers t, k, ℓ such that $\log^2 n \leq t \leq \frac{n}{100}$, $k = \lfloor \frac{n}{2t} \rfloor$, and $\ell = \lceil \frac{5n^2 t}{\log n} \rceil$,*

$$\dim(\langle \partial^{=k} NW_{t,n} \rangle_{\leq \ell}) \geq \frac{1}{n^3} \binom{n^2 + \ell + n - k}{n^2}$$

We will also use the following result from [KSS13] in our calculations.

Theorem 4.7 ([KSS13])**.** *For any fixed constant α and t, k, ℓ such that $\log^2 n \leq t \leq \frac{n}{100}$, $k = \lfloor \frac{n}{2t} \rfloor$ and $\ell = \lceil \frac{5n^2 t}{\log n} \rceil$, if*

$$E = \frac{\frac{1}{n^3}\binom{n^2+\ell+n-k}{n^2}}{\binom{\frac{\alpha n}{t}}{k}\binom{n^2+\ell+k(t-1)}{n^2}}$$

Then, $E \geq \exp(\Omega(\frac{n}{t} \log n))$.

For the range of values of t stated above, the value of k lies in the range $200 \leq k \leq \frac{n}{2 \log^2 n}$. To complete the proof, we will argue that after setting the y variables to a constant, there is a value of k in this range and an ℓ such that the dimension of the shifted partial derivative span of the circuit is small . Based on this value of k, we will then invoke a particular projection $NW_{t,n}$ of NW_n and then use the bound from Theorem 4.6.

Proof of Theorem 2.1. Let us consider a $\Sigma\Pi\Sigma\Pi(r)$ circuit of size s which computes the polynomial NW_n. As discussed, we will analyze the shifted partial derivative complexity of the circuit and the polynomial under the restriction that the $\{y_1, y_2, \ldots, y_n\}$ variables are set to constants. So, the degree of the polynomial computed is n and the number of alive variables is $N = n^2$. Let $0 < \epsilon < 1$ be a constant. We will now show that we can choose a value of k such that the conditions in Theorem 4.4 and Theorem 4.6 hold. From the proof of Theorem 4.4, we know that there are at most $\frac{r}{\epsilon}$ many choices of integer $0 < i < \frac{2r}{\epsilon}$ that are bad i.e that $k = n^{\frac{\epsilon \cdot i}{2r}}$ does not give us the upper bound on the complexity of the shifted partial derivatives as stated in Theorem 4.4. Now, all we need to show is that there is one such "good" i such that $200 \leq k = n^{\frac{\epsilon \cdot i}{2r}} \leq \frac{n}{2 \log^2 n}$. For this to hold, we need to show a "good" i in the range $\frac{2r}{\epsilon \log n} \log 200 < i < \frac{2r}{\epsilon}(1 - \frac{1 + 2 \log \log n}{\log n})$. The number of integers in this range is at least $\frac{2r}{\epsilon}(1 - 3\frac{\log \log n}{\log n})$, while the number of bad i is at most $\frac{r}{\epsilon}$. Hence, for n large enough, there is an i such that for the resulting k, the bound in Theorem 4.4 holds and $t = \frac{n}{2k}$ satisfies $\log^2 n \leq t \leq \frac{n}{100}$. Let us fix such a good k. Let us now fix $t = \frac{n}{2k}$, $\ell = \frac{5n^2 t}{\log n}$ and $\epsilon = \frac{1}{8}$. Now, let us consider the restriction of C when just y_t is set

to 1 and y_j is set to zero for every $j \neq t$. In this case, the circuit just computes $NW_{t,n}$. From Theorem 4.4, we get

$$\dim(\langle \partial^{=k} C \rangle_{\leq \ell}) \leq r \cdot s^{k \cdot n^{-\epsilon/2r}} \cdot \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N}$$

and from Theorem 4.6, we get

$$\dim(\langle \partial^{=k} NW_{t,n} \rangle_{\leq \ell}) \geq \frac{1}{n^3} \binom{n^2 + \ell + n - k}{n^2}$$

So, if C computes NW_n, then $\dim(\langle \partial^{=k} C \rangle_{\leq \ell}) \geq \dim(\langle \partial^{=k} NW_{t,n} \rangle_{\leq \ell})$. Thus

$$r \cdot s^{k \cdot n^{-\epsilon/2r}} \cdot \binom{k/\epsilon + k - 1}{k} \binom{N + 4\epsilon n - k + \ell}{N} \geq \frac{1}{n^3} \binom{n^2 + \ell + n - k}{n^2}.$$

Substituting the parameters, we get

$$r \cdot s^{k \cdot n^{-1/16r}} \geq \frac{\frac{1}{n^3}\binom{n^2+\ell+n-k}{n^2}}{\binom{\frac{\alpha n}{t}}{k}\binom{n^2+\ell+k(t-1)}{n^2}}$$

for some appropriate constant α dependent on ϵ. From theorem 4.7, this implies that

$$r \cdot s^{k \cdot n^{-1/16r}} \geq \exp(\Omega(\frac{n}{t} \log n)) = \exp(\Omega(k \log n))$$

Using the fact that r is at most s (in fact it is much much smaller), we conclude that

$$k \cdot n^{-1/16r} \cdot \log s \geq \Omega(k \log n).$$

Thus

$$\log s \geq \Omega(n^{1/16r} \log n)$$

and hence

$$s \geq \exp\left(n^{\Omega(1/r)} \log n\right).$$

□

A very similar calculation lets us prove Theorem 2.2.

Proof of Theorem 2.2. For a $\Sigma\Pi\Sigma\Pi^*$ circuit, the calculation will proceed exactly the same as above, and in the end, we will get

$$s \geq \exp\left(n^{\Omega(1/m)}\right),$$

which on substituting $m = 1/\epsilon + 1$, completes the proof. Thus, we obtain exponential lower bounds for $\Sigma\Pi\Sigma\Pi^*$ circuits computing the polynomial NW_n regardless of their top fan-in. □

5 Depth reduction is tight for $\Sigma\Pi\Sigma\Pi(\Omega(\log n))$ circuits

In this section, we will show that the depth reduction procedure of Koiran and Tavenas [Koi12, Tav13] as given in Theorem 1.1 is tight. On the way to this result, we will prove a *Hierarchy* theorem(Theorem 2.4) for formulas of depth 4 with bounded bottom fan-in. We will then build up on this proof, and prove Theorem 2.3 and Theorem 2.5. We will first provide an overview of the proof.

Proof Overview: We will construct an infinite family of polynomials $\{\mathcal{Q}_n\}_n$ (here n is a prime power), such that $\mathcal{Q}_n$ is a homogeneous polynomial in $N = \theta(n^2)$ variables of degree $n+1$ which can be computed by a polynomial sized homogeneous $\Sigma\Pi\Sigma\Pi(O(\log n))$ circuit. We will show that for each $\omega(\log n) \leq a \leq \frac{n}{800}$, $\mathcal{Q}_n$ requires homogeneous $\Sigma\Pi\Sigma\Pi^{[a]}$ circuits of top fan-in $2^{\Omega(\frac{n}{a}\log n)}$. In order to construct this polynomial family, we will construct for each $\omega(\log n) \leq t \leq \frac{n}{40}$, a family of polynomials $\{\mathcal{P}_{t,n}(\overline{x})\}$, such that each $\mathcal{P}_{t,n}(\overline{x})$ is a homogeneous polynomial in n^2 variables and of degree n, and can be computed by a polynomial sized homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ circuit. Moreover, we will show that any homogeneous $\Sigma\Pi\Sigma\Pi^{[t/20]}$ circuit computing it must have top fan-in at least $2^{\Omega(\frac{n}{t}\log n)}$. We will then apply the interpolation trick of [KSS13] to $\mathcal{P}_{t,n}$ for various t to obtain the $\mathcal{Q}_n$. The construction is heavily inspired by the idea of constructing hard polynomials using Nisan-Wigderson designs used in [KSS13]. To show the lower bound for each t, we will use ideas from [CM13] and [FLMS13], and show that for suitable k, $\partial^{=k}(P(\overline{x}))$ has a large number of elements whose leading monomials are at a "large distance" from each other.

5.1 Proof of Theorem 2.4

For the rest of this section, we will assume that n is a prime power. For each such n, we will identify the elements of the field $\mathbb{F}_n$ with the elements of the set $[n] = \{1, 2, \ldots, n\}$. For a parameter t which is a positive integer less than n, let us now partition the set $[n]$ into $\lceil \frac{n}{t} \rceil$ parts which are roughly equal and each is of size about t. For brevity, we will indicate $\frac{n}{t}$ by $\tilde{t}$. We will let $C_i = \{t(i-1)+1, t(i-1)+2, \ldots, ti\}$ denote the i^{th} such partition. Also, for every $j \in [\tilde{t}]$ and $i \leq t$, let C_j^i be the set of the i smallest elements in C_j. Let us also consider a parameter p which we will later set to an appropriately chosen constant. Let $\mathcal{S}_p$ be the set of all univariate polynomials of degree p over $\mathbb{F}$ and let ${\mathcal{S}_p}^{\tilde{t}}$ be the set of ordered $\tilde{t}$ tuples over $\mathcal{S}_p$. Clearly, $|\mathcal{S}_p|$ is $\theta(n^{p+1})$, when p is a constant. In the rest of the paper, we will use $\overline{x}$ to denote the set of n^2 variables $\{x_{i,j} : i, j \in [n]\}$ and $\overline{y}$ to denote the set of variables $\{y_1, y_2, \ldots, y_n\}$. We will use the following notion of distance between two monomials as defined in [CM13].

Definition 5.1 ([CM13])**.** *Let m_1 and m_2 be two monomials over a set of variables. Let S_1 and S_2 be the multiset of variables in m_1 and m_2 respectively, then the distance $\Delta(m_1, m_2)$ between m_1 and m_2 is the min$\{|S_1| - |S_1 \cap S_2|, |S_2| - |S_1 \cap S_2|\}$ where the cardinalities are the order of the multisets.*

We will also use the following notion of distance between strings or ordered tuples. For any two strings s_1, s_2 of the same length, the distance between them $\Delta(s_1, s_2)$ is the number of coordinates at which s_1, s_2 disagree with each other. For brevity, we will use αm to refer to $\lfloor \alpha m \rfloor$ for any positive integer m and any real number α.

Based on the notations defined, we define the class of polynomials $\mathcal{P}_{p,t,n}$:

$$\mathcal{P}_{p,t,n}(\overline{x}) = \prod_{j \in [\tilde{t}]} \sum_{f \in \mathcal{S}_p} \prod_{i \in C_j} x_{i,f(i)}$$

From the expression above, it follows that for every n, t and a constant p, $\mathcal{P}_{p,t,n}$ can be computed by a polynomial sized $\Pi\Sigma\Pi$ formula. Observe that in fact it can be computed by a *regular* formula[12]. We summarize this observation below.

Observation 5.2. *For every n and a constant p, $\mathcal{P}_{p,t,n}$ can be computed by a $\Pi\Sigma\Pi$ regular formula of size polynomial in n.*

[12]When t divides n the formula will be exactly regular, and if not a simple modifcation could make it regular, but the details are simple and irrelevant.

We now intend to use the setup introduced in [CM13] by Chillara and Mukhopadhyay to show that this polynomial requires homogeneous $\Sigma\Pi\Sigma\Pi^{[\frac{t}{20}]}$ circuits of top fan-in at least $2^{\Omega(\frac{n}{t}\log n)}$. This forms the basis for our hierarchy theorem. In [CM13] the following theorem is proved, which gives a sufficient condition to show hardness for homogeneous $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuits.

Theorem 5.3 (Theorem 3 in [CM13]). *Let $f(X)$ be a polynomial of degree n in $n^{O(1)}$ variables such that for some constant δ there are $n^{\delta k}$ different polynomials in $\partial^{=k}f$ for $k = \gamma\sqrt{n}$(where $0 < \gamma < 1$ is a constant) such that any two of their leading monomials have distance at least $d = \frac{n}{c}$ for a constant $c > 1$. Then, any homogeneous $\Sigma\Pi\Sigma\Pi^{[\sqrt{n}]}$ circuit that computes $f(X)$ must have top fan-in at least $2^{\Omega(\sqrt{n}\log n)}$.*

Although the result is stated for k being around $\sqrt{n}$, the theorem is also true for k in a larger range of values of k. To show our lower bounds, we will argue that there are a "large" number of k^{th} order partial derivatives of $\mathcal{P}_{p,t,n}$ for some appropriate k, whose leading monomials have the distance property stated above.

To this end, we will now try and understand the monomial structure of the partial derivatives of an appropriately chosen order of $\mathcal{P}_{p,t,n}$. Now, from the definition of $\mathcal{P}_{p,t,n}$, every monomial in it can be identified by an ordered tuple of length $\tilde{t}$ over the set of polynomials in $\mathcal{S}_p$ and vice versa. So, for any $\overline{f} = (f_1, f_2, \ldots, f_{\tilde{t}}) \in {\mathcal{S}_p}^{\tilde{t}}$, let

$$m_{\overline{f}} = \prod_{j\in[\tilde{t}]}\prod_{i\in C_j} x_{i,f_j(i)}$$

From the definitions above and that of $\mathcal{P}_{p,t,n}(\overline{x})$, it follows that

$$\mathcal{P}_{p,t,n}(\overline{x}) = \sum_{\overline{f}\in{\mathcal{S}_p}^{\tilde{t}}} m_{\overline{f}}$$

Let

$$m'_{\overline{f}} = \prod_{j\in[\tilde{t}]}\prod_{i\in C_j^{2p}} x_{i,f_j(i)}$$

When we finally set parameters, we will always have p is a constant while t increases with n. So, for n large enough, $m'_{\overline{f}}$ divides $m_{\overline{f}}$ and

$$\frac{m_{\overline{f}}}{m'_{\overline{f}}} = \prod_{j\in[\tilde{t}]}\prod_{i\in C_j\setminus C_j^{2p}} x_{i,f_j(i)}$$

Now, we set $k = 2p\tilde{t}$ and look at the partial derivatives of $\mathcal{P}_{p,t,n}$ of order k. For each $\overline{f} \in {\mathcal{S}_p}^{\tilde{t}}$, the degree of $m'_{\overline{f}}$ equals k. Hence, $\partial^{=k}\mathcal{P}_{p,t,n}$ includes the set of partial derivatives of $\mathcal{P}_{p,t,n}$ with respect to $m'_{\overline{f}}$ for each $\overline{f} \in {\mathcal{S}_p}^{\tilde{t}}$. From the definition of $m_{\overline{f}}$ and $m'_{\overline{f}}$, and the fact that each polynomial in $\mathcal{S}_p$ has degree equal to p and two distinct polynomials in $\mathcal{S}_p$ cannot agree on more than p points, for any $\overline{f} \in {\mathcal{S}_p}^{\tilde{t}}$ and $\overline{g} \in {\mathcal{S}_p}^{\tilde{t}}$,

$$\partial_{m'_{\overline{f}}} m_{\overline{g}} = \begin{cases} 0 & \overline{f} \neq \overline{g} \\ \frac{m_{\overline{f}}}{m'_{\overline{f}}} & \overline{f} = \overline{g} \end{cases}$$

From this discussion, the following lemma follows.

Lemma 5.4. *For every $\overline{f} \in {\mathcal{S}_p}^{\tilde{t}}$, $\partial_{m'_{\overline{f}}}\mathcal{P}_{p,t,n}$ is a monomial and equals $\frac{m_{\overline{f}}}{m'_{\overline{f}}}$.*

At this point, we might hope to argue that for each $\overline{f} \in {\mathcal{S}_p}^{\tilde{t}}$ and $\overline{g} \in {\mathcal{S}_p}^{\tilde{t}}$ such that $\overline{f} \neq \overline{g}$, the distance between the monomials $\frac{m_{\overline{f}}}{m_{\overline{f}}}$ and $\frac{m_{\overline{g}}}{m_{\overline{g}}}$ is large. This statement in itself is not true, for if $\overline{f}$ and $\overline{g}$ differ in just one coordinate, then the distance between $\frac{m_{\overline{f}}}{m_{\overline{f}}}$ and $\frac{m_{\overline{g}}}{m_{\overline{g}}}$ could be as small as $t-3p$, which as it turns out is insufficient to achieve the desired bounds. Observe that if $\overline{f}$ and $\overline{g}$ differ in i coordinates, then the distance between $\frac{m_{\overline{f}}}{m_{\overline{f}}}$ and $\frac{m_{\overline{g}}}{m_{\overline{g}}}$ is at least $it-3pi$. (We prove this fact in Lemma 5.7). To prove the lower bound, we will show that there is a "large" nice subset $\mathcal{N} \subseteq {\mathcal{S}_p}^{\tilde{t}}$ such any $\overline{f}$ and $\overline{g}$ in $\mathcal{N}$ differ in a constant fraction of all coordinates. The following lemma, which just follows from the existence and properties of Reed-Solomon codes guarantees the existence of such an $\mathcal{N}$.

Lemma 5.5. *Let $0 < \alpha < 1$ be any absolute constant and let q be a prime power. For any alphabet Σ of size q and positive integer m such that $m < q$, there is a set $\mathcal{C}$ of strings of length m over Σ of size $q^{(1-\alpha)m}$ such that any two strings in $\mathcal{C}$ are at a distance at least αm apart.*

Proof. Let $\mathcal{C}$ be the set of codewords obtained when the set $\Sigma^{(1-\alpha)m}$ is encoded using Reed-Solomon codes of message length $(1-\alpha)m$ and code length m. The distance of the code is αm and the number of codewords is $q^{(1-\alpha)m}$. Hence the set satisfies the properties stated in the statement. □

Lemma 5.5 immediately implies the existence of a set $\mathcal{N}$, when invoked with parameters $\Sigma = \mathcal{S}_p$, $m = \tilde{t}$. So, we have the following corollary.

Corollary 5.6. *For all α such that $0 < \alpha < 1$, there exists $\mathcal{N} \subseteq {\mathcal{S}_p}^{\tilde{t}}$ of size equal to $n^{(1-\alpha)(p+1)\tilde{t}}$ such that for any distinct pair $\overline{f}$ and $\overline{g}$ in $\mathcal{N}$, $\overline{f}$ and $\overline{g}$ differ in at least $\alpha\tilde{t}$ coordinates.*

Informally, the set $\mathcal{N}$ now gives us a large number of partial derivatives which are at a large distance from each other. We formalize this claim in the lemma below.

Lemma 5.7. *For $k = 2p\tilde{t}$, the set $\partial^{=k}\mathcal{P}_{p,t,n}$ has a subset S of size at least $n^{(1-\alpha)(p+1)\tilde{t}}$ such that every element in this subset is a monomial and any two such monomials are at a distance of at least $\alpha\tilde{t}(t-3p)$ from each other.*

Proof. Let us pick any two $\overline{f}$ and $\overline{g}$ in $\mathcal{N}$. Let $i \in [\tilde{t}]$ be an index such that $f_i \neq g_i$. Then over the set C_i, f_i and g_i can agree at at most p points. Therefore, the monomials $m_{\overline{f}}$ and $m_{\overline{g}}$ differ in at least $t-p$ variables of the form $x_{h,j}$ for $h \in C_i$. Now, for each i and each $\overline{f} \in {\mathcal{S}_p}^{\tilde{t}}$, $m'_{\overline{f}}$ contains exactly $2p$ variables $x_{h,j}$ with $h \in C_i$. Hence, $\frac{m_{\overline{f}}}{m'_{\overline{f}}}$ and $\frac{m_{\overline{g}}}{m'_{\overline{g}}}$ differ in at least $t-3p$ variables of the form $x_{h,j}$ for $h \in C_i$. So, each coordinate i where $\overline{f}$ and $\overline{g}$ differ from each other contributes $t-3p$ to the distance between $\frac{m_{\overline{f}}}{m'_{\overline{f}}}$ and $\frac{m_{\overline{g}}}{m'_{\overline{g}}}$. Hence, for every $\overline{f}$ and $\overline{g} \in \mathcal{N}$, $\frac{m_{\overline{f}}}{m'_{\overline{f}}}$ and $\frac{m_{\overline{g}}}{m'_{\overline{g}}}$ are at a distance at least $\alpha\tilde{t}(t-3p)$ apart. The lemma now follows from the fact that the size of $\mathcal{N}$ is at least $n^{(1-\alpha)(p+1)\tilde{t}}$. □

We now essentially have all the ingredients we need for showing lower bounds for homogeneous $\Sigma\Pi\Sigma\Pi^{[a]}$ circuits computing $\mathcal{P}_{p,t,n}$. We will use the following lemma which is implicit in [CM13]. A similar calculation also appears in [FLMS13].

Lemma 5.8 (Implicit in [CM13])**.** *Let $Q = \sum_{i=1}^{s'} Q_{i1}Q_{i2}\dots Q_{iz}$ where each Q_{ij} is an N variate polynomial of degree bounded by u. Also, for some $r \leq z$, suppose there are s elements in $\partial^{=r}Q$ such that the distance between the leading monomial of any two of these is at least d. Then, for any positive integer ℓ such that $\ell \leq \frac{Nd}{2\ln(s.N^2)}$ and $(ru-r)^2 = o(\ell)$,*

$$s^{'} \geq \frac{s(1-\frac{1}{N^2})}{\binom{z+r}{r}e^{\frac{N(ru-r)}{\ell}}}$$

.

For $p = 1$, we will call the corresponding polynomial family $\{\mathcal{P}_{p,t,n}\}_n$ as $\{\mathcal{P}_{t,n}\}_n$. We will now prove the following lower bound for homogeneous $\Sigma\Pi\Sigma\Pi^{[\frac{t}{20}]}$ circuits computing $\mathcal{P}_{t,n}$.

Theorem 5.9. *For any $\omega(\log n) \leq t \leq \frac{n}{40}$, any homogeneous $\Sigma\Pi\Sigma\Pi^{[\frac{t}{20}]}$ circuit computing $\mathcal{P}_{t,n}$ has top fan-in at least $2^{\Omega(\frac{n}{t}\log n)}$.*

Proof. We will invoke Lemma 5.8 after setting the parameters used in it appropriately. Let

- $p = 1$
- $\alpha = 0.9$
- $N = n^2$
- $u = \frac{t}{20}$
- $z = O(\frac{n}{t})$
- $r = \frac{2n}{t}$

Now, Lemma 5.7 implies that there is a set S of size $s = n^{0.2\tilde{t}}$ such that any two monomials in S are at a distance at least $d = \alpha\tilde{t}(t-3p) = 0.9n - 2.7\frac{n}{t}$. Observe that for $t > 270$, we get $d \geq 0.89n$. We now need to set ℓ to a value which satisfies the constraints in the hypothesis of Lemma 5.8 and which also implies a non trivial lower bound on $s^{'}$. The hypothesis of Lemma 5.8 requires that $\ell \leq \frac{Nd}{2\ln(s.N^2)}$. Substituting the values of p, N, s, d, we require $\ell \leq \frac{n^2 \times 0.89n}{2(4+\frac{0.2n}{t})\ln n}$. For $t < \frac{n}{20}$, any $\ell < \frac{0.89n^2t}{0.8\ln t}$ will satisfy this constraint since, $\frac{n^2\times 0.89n}{2(4+\frac{0.2n}{t})\ln n} \geq \frac{0.89n^2t}{0.8\ln t}$. Observe that the term $\binom{z+r}{r}$ is of the order $2^{O(\frac{n}{t})}$ by Shannon's entropy estimation. The other high order term in the denominator is $e^{\frac{Nru}{\ell}} = e^{\frac{n^3}{10\ell}}$. On the other hand, the highest order term in the numerator is $s = n^{0.2\frac{n}{t}}$. So, for a lower bound of the order $n^{\Omega(\frac{n}{t})}$, we will ensure that $0.2\frac{n\ln n}{t} \geq 1.1\frac{n^3}{10\ell}$. This requires $\ell > \frac{0.55n^2t}{\ln n}$. So, we need $\frac{0.55n^2t}{\ln n} < \ell < \frac{0.89n^2t}{0.80\ln t}$ and $o(\ell) = n^2$. Let us set $\ell = \frac{n^2t}{\ln n}$. For t being $\omega(\log n)$, this satisfies $o(\ell) = n^2$. Substituting all these values into the expression in Lemma 5.8, we get $s^{'} \geq 2^{\Omega(\frac{n}{t}\log n)}$.

□

This completes the proof of Theorem 2.4. We will now build upon this proof to obtain Theorem 2.3.

5.2 Proof of Theorem 2.3

So far, we have constructed a polynomial family $\mathcal{P}_{t,n}$ such that $\mathcal{P}_{t,n}$ requires homogeneous $\Sigma\Pi\Sigma\Pi^{[\frac{t}{20}]}$ circuits with top fan-in at least $n^{\Omega(\frac{n}{t})}$. We can now build upon the construction of $\mathcal{P}_{t,n}$ described so far to construct a single polynomial family which is hard for any homogeneous $\Sigma\Pi\Sigma\Pi^{[a]}$ circuit for every $\omega(\log n) \leq a \leq \frac{n}{800}$. We will now use a variation of the interpolation trick described in Lemma 14 in [KSS13]. The idea now is just to take a linear combination of $\mathcal{P}_{a,n}$ for $O(\log n)$ many such values of a, with coefficients being the variables $\overline{y}$, such that for every a such that $\omega(\log n) \leq a \leq \frac{n}{800}$, there is a t such that $20a \leq t \leq 400a$ and such that $\mathcal{P}_{t,n}$ is in the linear combination.

In particular let us define the following family of polynomials $\mathcal{Q}_n$:

$$\mathcal{Q}_n(\overline{x},\overline{y}) = \sum_{i=0}^{O(\log n)} y_i \cdot \mathcal{P}_{20^i,n}(\overline{x})$$

Observe that $\mathcal{Q}_n$ can be computed by a polynomial size homogeneous $\Sigma\Pi\Sigma\Pi(\log n)$ circuit.

If $\mathcal{Q}_n$ could be computed *efficiently* by a homogeneous $\Sigma\Pi\Sigma\Pi^{[a]}$ for some a, then so could any projection of the sum (i.e. we set all but one of the y_i to 0), i.e. so could $\mathcal{P}_{t,n}$. This contradicts Theorem 2.4. In particular we get that every $\omega(\log n) \leq a \leq \frac{n}{800}$, any homogeneous $\Sigma\Pi\Sigma\Pi^{[a]}$ circuit computing $\mathcal{Q}_n$ must have top fan-in at least $2^{\Omega(\frac{n}{a}\log n)}$. This completes the proof of Theorem 2.3.

5.3 Proof of Theorem 2.5

The proof of Theorem 2.5 follows immediately from Theorem 2.3 along with Theorem 15 from [KSS13].

6 Discussion and future directions

One of the main questions left open by our lower bounds on $\Sigma\Pi\Sigma\Pi(r)$ circuits is to remove the restriction on top fan-in, and to prove super polynomial lower bounds for all homogeneous $\Sigma\Pi\Sigma\Pi$ circuits. Currently, we have no nontrivial lower bounds for homogeneous $\Sigma\Pi\Sigma\Pi(\log n)$ circuits, even in the further special case when the family of circuits is a sum of $\Pi\Sigma\Pi^{[t]}$ circuits for different values of t. We identify this as the simplest class of circuits/formulas for which we don't know how to prove lower bounds. While this would still not suffice in proving lower bounds for general arithmetic circuits, this seems to be an important step in that direction. Another very interesting direction would be to give nontrivial PIT results for $\Sigma\Pi\Sigma\Pi(r)$ circuits when r is a constant. So far, we only know how to derandomize PIT when the $\Sigma\Pi\Sigma\Pi(r)$ circuits are multilinear, and our lower bound for $\Sigma\Pi\Sigma\Pi(r)$ circuits could be viewed as a first step in this direction.

One corollary of our results is a hierarchy theorem for $\Sigma\Pi\Sigma\Pi^{[t]}$ formulas. A very interesting question that we don't know how to answer is if there is a tighter hierarchy theorem. We believe that for every t, polynomial sized $\Sigma\Pi\Sigma\Pi^{[t]}$ formulas should be able to compute a much richer class of polynomials than polynomial sized $\Sigma\Pi\Sigma\Pi^{[t-1]}$ formulas. A special case that we do not know how to answer is the relative complexity of $\Sigma\Pi\Sigma\Pi^{[2]}$ formulas versus $\Sigma\Pi\Sigma\Pi^{[1]}$ formulas (which are basically depth 3 formulas). Another kind of hierarchy question that we don't fully understand but which we think would be very interesting is to understand the relative complexity of depth d formulas versus depth $d+1$ formulas for constant d. Perhaps a refinement of the depth reduction techniques of Koiran and Tavenas would shed light on these questions.

Acknowledgements

We would like to thank Amir Yehudayoff and Amir Shpilka for helpful discussions at several stages of this work. We would also like to thank Swastik Kopparty and Avi Wigderson for many helpful comments on an earlier version of this paper.

References

[AV08] M. Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *Proceedings of the 49th Annual FOCS*, pages 67–75, 2008.

[CM13] Suryajith Chillara and Partha Mukhopadhyay. Depth-4 lower bounds, determinantal complexity: A unified approach. *CoRR*, abs/1308.1640v3, 2013.

[FLMS13] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:100, 2013.

[GK98] D. Grigoriev and M. Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *Proceedings of the 30th Annual STOC*, pages 577–582, 1998.

[GKKS13a] A. Gupta, P. Kamath, N. Kayal, and R. Saptharishi. Approaching the chasm at depth four. *In Proceedings of CCC*, 2013.

[GKKS13b] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:26, 2013.

[GKL12] Ankit Gupta, Neeraj Kayal, and Satyanarayana V. Lokam. Reconstruction of depth-4 multilinear circuits with top fan-in 2. In *Proceedings of the 44th Annual STOC*, pages 625–642, 2012.

[Kay] N. Kayal. Personal communication.

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012.

[KMSV10] Z. Karnin, P. Mukhopadhyay, A. Shpilka, and I. Volkovich. Deterministic identity testing of depth 4 multilinear circuits with bounded top fan-in. In *Proceedings of the 42nd Annual STOC*, pages 649–658, 2010.

[Koi12] P. Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012.

[KS13] Mrinal Kumar and Shubhangi Saraf. Lower bounds for depth 4 homogenous circuits with bounded top fanin. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:68, 2013.

[KSS13] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:91, 2013.

[NW95] N. Nisan and A. Wigderson. Lower bounds on arithmetic circuits via partial derivatives. In *Proceedings of the 36th Annual FOCS*, pages 16–25, 1995.

[SV11] S. Saraf and I. Volkovich. Black-box identity testing of depth-4 multilinear circuits. In *Proceedings of the 43rd Annual STOC*, pages 421–430, 2011.

[SW01] A. Shpilka and A. Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001.

[SY10] A. Shpilka and A. Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, March 2010.

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

[Val79] L. G. Valiant. Completeness classes in algebra. In *Proceedings of the 11th Annual STOC*, STOC '79, pages 249–261, New York, NY, USA, 1979. ACM.

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

A Completeness of the model of $\Sigma\Pi\Sigma\Pi(r)$ circuits

Depth 3 and depth 4 circuits with bounded top fan-in ($\Sigma\Pi\Sigma(r)$ and $\Sigma\Pi\Sigma\Pi(r)$ respectively) have been extensively studied in the past especially in the context of polynomial identity testing (PIT). The question of lower bound for $\Sigma\Pi\Sigma(r)$ circuits is almost uninteresting since it can be shown quite easily that for $r < n$, $\Sigma\Pi\Sigma(r)$ circuits cannot compute the $n \times n$ permanent or determinant, *no matter what the size* of the circuit. Thus the class of $\Sigma\Pi\Sigma(r)$ circuits is not complete, in the sense that the class of circuits cannot even compute all polynomials. In contrast, the class of depth 2 $\Sigma\Pi$ circuits (with no restriction on top fan-in) is complete, but lower bounds are trivial for this model since any polynomial with m monomials needs a $\Sigma\Pi$ circuit of size at least m to compute it.

It was observed by Kayal [Kay] that if one considers the class of depth 4 circuits and one imposes the additional requirement that each product gate has at least 2 nontrivial factors, then the class of $\Sigma\Pi\Sigma\Pi(r)$ circuits with $r < n/2$ circuits is not complete. This is because if α is a common root of at least two of the factors of each of the product gates, then it would be a zero of multiplicity 2 of the polynomial computed by the circuit. Also if $r < n/2$ then such an α always exists. Hence if one starts with a polynomial that does not vanish at any point with multiplicity 2, then it cannot be computed by such a circuit. Thus in this case we can prove lower bounds easily.

The general class of depth 4 $\Sigma\Pi\Sigma\Pi(r)$ circuits even when $r = 1$ is a complete class, since it contains the class of depth 2 $\Sigma\Pi$ circuits. However lower bounds for $\Sigma\Pi\Sigma\Pi(r)$ circuits for $r \geq 2$ did not seem to be known prior to this work.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de