2014 IEEE Annual Symposium on Foundations of Computer Science

An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas

Neeraj Kayal*, Nutan Limaye†, Chandan Saha‡ and Srikanth Srinivasan§
*Microsoft Research India, Bangalore, India
Email: neeraka@microsoft.com
†Indian Institute of Technology Bombay, Mumbai, India
Email: nutan@cse.iitb.ac.in
‡Indian Institute of Science, Bangalore, India
Email: chandan@csa.iisc.ernet.in
§ Indian Institute of Technology Bombay, Mumbai, India
Email: srikanth@math.iitb.ac.in

***Abstract*—We show here a $2^{\Omega(\sqrt{d}\cdot\log N)}$ size lower bound for homogeneous depth four arithmetic formulas. That is, we give an explicit family of polynomials of degree d on N variables (with $N = d^3$ in our case) with $0, 1$-coefficients such that for any representation of a polynomial f in this family of the form**

$$f = \sum_i \prod_j Q_{ij},$$

where the Q_{ij}'s are homogeneous polynomials (recall that a polynomial is said to be homogeneous if all its monomials have the same degree), it must hold that

$$\sum_{i,j} (\textbf{Number of monomials of } Q_{ij}) \quad \geq \quad 2^{\Omega\left(\sqrt{d}\cdot\log N\right)}.$$

The above mentioned family, which we refer to as the Nisan-Wigderson design-based family of polynomials, is in the complexity class VNP. Our work builds on the recent lower bound results [1], [2], [3], [4], [5] and yields an improved quantitative bound as compared to the quasi-polynomial lower bound of [6] and the $N^{\Omega(\log\log N)}$ lower bound in the independent work of [7].

***Keywords*-Arithmetic circuits, shifted partial derivatives, lower bounds.**

I. Introduction

Understanding efficient computation and the VP versus VNP problem. The model of arithmetic circuits is an algebraic analogue of the model of Boolean circuits: an arithmetic circuit contains addition $(+)$ and multiplication $(\times)$ gates and it naturally computes a polynomial in the input variables over some underlying field. We typically allow the input edges to a $+$ gate to be labelled with arbitrary constants from the underlying field $\mathbb{F}$ so that a $+$ gate can in fact compute an arbitrary $\mathbb{F}$-linear combination of its inputs. In the field of arithmetic complexity, we seek to understand the phenomenon of efficient computation of (multivariate) polynomials via arithmetic circuits. A specific fundamental question is the VP versus VNP problem. The complexity classes VP and VNP consist of families of polynomials and can be viewed as algebraic analogues of the classes P and NP respectively[1]. This outstanding open problem asks whether there are families of polynomials which admit an efficient description[2] but are hard to compute[3]. The hope is that it might be possible to use algebraic and geometric insights along with the structure of arithmetic circuits to make progress towards settling the VP vs VNP question. Till date, research on arithmetic circuits has produced several interesting results that have enriched our understanding of the lower bound problem and the related problems on polynomial identity testing & reconstruction (or learning) of arithmetic circuits. The survey [9] gives an account of some of the results and outstanding open questions in this area.

Can computation be efficiently parallelized? While the resolution of the VP vs VNP question would be a big landmark in our quest to understand efficient arithmetic computation, another fundamental pursuit might be to be understand efficient *parallel* computation. Circuits of low depth[4] correspond to computations which are highly parallel. A relevant question here is whether computation can be efficiently parallelized. Specifically, if an N-variate polynomial f of degree d can be computed by a circuit $\mathcal{C}$ of size s, what is the size of a minimal Δ-depth circuit $\mathcal{C}'$ computing the same polynomial? Following the landmark result [10], a series of generalizations and improvements [11], [12], [13] showed that this can be done with $\mathcal{C}'$ being

[1] It is known that if VNP can be computed by arithmetic circuits of polynomial size and degree and which have the additional property that the constants from the underlying field have polynomially bounded bitlengths then it must follow that P = NP (cf. [8]).

[2] A polynomial (family) is said to admit an efficient desciption if the coefficient of any given monomial can be computed efficiently.

[3] The VP versus VNP is perhaps closer in spirit to the #P versus NC problem in Boolean complexity.

[4] Recall that the depth of a circuit is the maximum length of any path from an input node to the output node.

0272-5428/14 $31.00 © 2014 IEEE
DOI 10.1109/FOCS.2014.15

a homogeneous[5] Δ-depth circuit (with unbounded fanin gates) of size $s^{O(d^{2/\Delta})}$. We do not know if this result is optimal. A recent result by [14], combined with observations by Tavenas [13] and Wigderson[6] shows that over fields of characteristic zero, the size of $\mathcal{C}'$ can be improved to $s^{O(d^{1/(\Delta-1)})}$ albeit at the loss of homogeneity of $\mathcal{C}'$. On the other hand, recent results by [3] and [4] together imply that if $\mathcal{C}'$ satisfies some additional regularity conditions then $s^{O(d^{2/\Delta})}$ is optimal. Without the regularity restrictions, we do not know if either of these depth reductions is optimal - the main bottleneck being the nonavailability of lower bounds for low depth (homogeneous) circuits.

VP versus VNP and homogeneous depth four lower bounds. Note also that the depth reduction results of [10], [11], [12], [13] imply in particular that if a degree-d, N-variate polynomial f is in VP then it can be computed by a homogeneous depth four circuit[7] of size $N^{O(\sqrt{d})}$. This also opens another potential avenue of attack on the VP versus VNP problem - it suffices to prove strong enough homogeneous depth four lower bounds for any polynomial (family) in VNP. The implicit hope here is that low depth circuits being easier to analyze, it might be more feasible to prove such strong lower bounds against them. Thus proving lower bounds against low depth circuits is relevant both from the viewpoint of making progress on the VP versus VNP question and for understanding the limits to which arithmetic computation can be efficiently parallelized. In this work, we prove a lower bound of $N^{\Omega(\sqrt{d})}$ on the size of a homogeneous depth four circuit computing a polynomial (family) in VNP.

Previous work on super-polynomial lower bounds. Lower bounds for homogeneous formulas were first proved by Nisan and Wigderson [15], who introduced the method of *partial derivatives* in this setting. They used this approach to show an exponential lower bound for homogeneous depth-3 formulas and also some other interesting lower bound results on circuit size and depth. [8]

The use of partial derivatives (alongside other important ideas) has since been a recurrent theme in arithmetic circuit lower bounds. For depth-3 (possibly inhomogeneous) formulas over constant-sized finite fields, this method was used to prove an exponential lower bound by [18], [19]. Further, Raz [20] showed that any *multilinear* formula computing the determinant Det_n (or the permanent Perm_n) polynomial has $n^{\Omega(\log n)}$ size with subsequent separations[9] and refinements[10] in [21] and in [22]. There are also other works such as [23], which are based upon studying partial derivatives or associated matrices involving partial derivatives like the Jacobian or the Hessian[11].

The situation for depth-4 homogeneous formulas has been substantially improved by the recent work of [1], [2], followed by the work of [3] and [4]. These works have led to a $2^{\Omega(\sqrt{d}\log N)}$ lower bound for depth-4 homogeneous formulas with bottom fan-in $O(\sqrt{d})$ (where d is the degree of the N-variate 'target' polynomial on which the lower bound is shown). Further, [3] and [4] together imply a super-polynomial separation between *algebraic branching programs* (ABPs) and *regular formulas* - two natural submodels of arithmetic circuits. Quite interestingly, the work of [5] in fact showed a super-polynomial separation between *homogeneous depth-4 formulas* and *regular formulas*! At a high level, these separation results are obtained by showing that a polynomial computed by a regular formula can also be computed by a bounded bottom fan-in homogeneous depth-4 formula having *low* top fan-in. Now it was shown in [5] that there is a polynomial (family) computed by polynomial size homogeneous depth-4 formulas such that any bounded bottom fan-in homogeneous depth-4 formula computing the polynomial must have *high* top fan-in. This implied the separation between homogeneous depth-4 formulas and regular formulas.

A seemingly tempting problem left open in these works is if the lower bound of $2^{\Omega(\sqrt{d}\log N)}$ in the above statement could be improved to $2^{\omega(\sqrt{d}\log N)}$, since a super-polynomial lower bound for general circuits would ensue immediately. At the heart of these results lies the study of the space of *shifted partial derivatives* of polynomials and an associated measure called the *dimension of the shifted partials* - a technique introduced in [1], [2]. Loosely speaking, the dimension of the shifted partials of a polynomial g refers to the dimension of the $\mathbb{F}$-linear vector space generated by the set of polynomials obtained by multiplying (shifting) the partial derivatives of g with monomials of suitable degrees.

Homogeneous Formulas and Shifted Partials. A more modest (compared to the resolution VP versus VNP), but still a highly interesting milestone in arithmetic complexity

[5] Recall that the formal degree of a node in a circuit is defined inductively in the natural manner - leaf nodes labelled with variables (respectively with field constants) have formal degree 1 (respectively zero) and every internal $+$ gate (resp. $\times$ gate) is said to have formal degree equal to the maximum of (resp. the sum of) the formal degrees of its children. An arithmetic circuit is said to be homogeneous if it is syntactically homogeneous, i.e. at every intermediate $+$ gate, the inputs all have the same formal degree.

[6] personal communication

[7] with bottom fanin bounded by $O(\sqrt{d})$.

[8] Prior to this work, Smolensky [16] used this measure to prove certain lower bounds for boolean circuits, and Nisan [17] showed an exponential lower bound for noncommutative arithmetic formulas.

[9] Building upon [20], a super-polynomial gap between multilinear circuits and formulas was obtained in [21].

[10] Also building upon [20], a significantly better bound was later shown for bounded (i.e. constant) depth multilinear circuits [22]: A depth-d multilinear circuit computing Det_n or Perm_n has size $2^{n^{\Omega(1/d)}}$.

[11] A recent survey by Chen, Kayal and Wigderson [24] gives more applications of partial derivatives.

might be to prove superpolynomial lower bounds for homogeneous formulas[12]. Could the shifted partials technique be used to achieve the same? The work [5] poses an apparent 'hurdle' for achieving even a homogeneous depth-4 formula lower bound: the strategy of *directly reducing* a homogeneous depth-4 formula to a bounded bottom fan-in homogeneous depth-4 formula of *low* top fan-in (followed by applying the top fan-in lower bound on the latter kind of formulas) will not work! At this point, proving a lower bound for homogeneous depth-4 formulas seems like a natural step forward to understand the strengths and limitations of the shifted partials method better - this is a recurring open problem stated in [3], [4], [25], [13]. Further, with the hope of proving a super-polynomial lower bound for general homogeneous formulas, it would be good to have an *exponential* lower bound for homogeneous depth-4 formulas first.

Our result. We show here that a slightly modified (or augmented) version of the shifted partial measure can be used to obtain an *exponential* lower bound for depth-4 homogeneous formulas. For the ease of reference in this paper, we will call this modified measure the *projected shifted partials*. Loosely speaking, the idea is to *shift the derivatives of a polynomial by a carefully chosen set of monomials and then view these after 'projecting' them to an appropriate set of monomials*. Our results are formally stated below.

Theorem 1. *Let $\mathbb{F}$ be any field of characteristic zero. There is an explicit family of homogeneous polynomials of degree d in $N = d^3$ variables with zero-one coefficients such that any homogeneous $\Sigma\Pi\Sigma\Pi$ formula over $\mathbb{F}$ computing this family must have size at least $2^{\Omega(\sqrt{d}\cdot\log N)}$. In other words, for any representation of the degree d polynomial f in the family, of the form*

$$f = \sum_i \prod_j Q_{ij},$$

where the Q_{ij}'s are homogeneous polynomials, it must hold that

$$\sum_{i,j} (\textit{Number of monomials of } Q_{ij}) \quad \geq \quad 2^{\Omega\left(\sqrt{d}\cdot\log N\right)}.$$

The explicit polynomial f in the theorem above is a variant of the Nisan-Wigderson design-based polynomial introduced in [3] and further studied in [5], [7]. While this family of polynomials is explicit (in VNP), it is not known to be efficiently computable. Thus, as it stands, our main theorem has two limitations - it is valid only over fields of characteristic zero and the explicit family of polynomials that we give is not known to be efficiently computable.

[12] Recall that, homogeneous formulas can be simulated by polynomial size ABPs which in turn can be simulated by polynomial size circuits.

Comparison with our earlier work [6]. The projected shifted partials measure is closely related to the measure we used earlier in [6] to obtain a quasi-polynomial lower bound for homogeneous depth-4 formulas. But there are also important differences between the two. The definition of the measure in [6] has an unconventional (perhaps also undesirable) feature - it depends on the target polynomial, Iterated Matrix Multiplication, on which the lower bound was shown. This is not the case for our present (somewhat cleaner) measure that can be applied on any target polynomial family and achieves a much stronger lower bound (exponential) as opposed to the quasi-polynomial lower bound in [6]. The primary source of this improvement is the design of a more suitable complexity measure (via a better ordering of the linear operators involved and a more careful shifting) and a refined analysis of rank estimation of a certain matrix. On the other hand, the lower bound in [6] holds for the families of Iterated Matrix Multiplication and Determinant polynomials that are in VP as compared to the family of Nisan-Wigderson design-based polynomials which is in VNP but not known to be in VP.

An independent result by [7]. Kumar and Saraf [7] independently proved a superpolynomial ($N^{\Omega(\log\ \log N)}$) lower bound for homogeneous depth four circuits using another nice augmentation of the shifted partial measure that they call *bounded support shifted partials*. We do not know if this measure can be used to prove an exponential lower bound. Indeed, they explicitly state the problem of proving exponential lower bounds for homogeneous depth four circuits as an open problem which we happen to achieve here.

The rest of the paper is devoted to proving Theorem 1. Many proofs and computations are omitted due to lack of space. They may be found in the full version of the paper [26].

II. Overview of our proof

We now give an outline of the proof of Theorem 1. Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a homogeneous polynomial of degree d on N variables over a field $\mathbb{F}$. Consider a representation of f of the form

$$f = \sum_{i=1}^{s} \prod_j Q_{ij}, \tag{1}$$

where the Q_{ij}'s are homogeneous polynomials. Note that any polynomial can be written in this way - the challenge is to prove a lower bound on the total number of monomials appearing in the Q_{ij}'s. For each $i \in [s]$, the i-th term in such a representation is defined to be $T_i = \prod_j Q_{ij}$. First observe that we can assume without loss of generality that the degree of each term T_i is at most d (as we can simply discard terms of degree larger than d without changing the output). So now assume that the total number of monomials in this representation is small, say $2^{o(\sqrt{d}\cdot\log N)}$ (else we have nothing to prove). In particular, our assumption means that every Q_{ij} has at most $2^{o(\sqrt{d}\cdot\log N)}$ monomials.

Using Random Restrictions to reduce the support size. In the first step, we consider the identity (1) and in that set each variable to zero independently at random with probability $(1-p)$ (a variable is left untouched with probability p.) Then any monomial m in any of the Q_{ij}'s which contains t distinct variables will now survive (i.e. remain nonzero under this substitution) with probability p^t. So if we choose $p = \frac{1}{N^{\Theta(1)}}$ then via an application of the union bound we deduce that all monomials of support at least $t = \Omega(\sqrt{d})$ will be 'killed' (i.e. set to zero) under this substitution[13]. For ease of subsequent exposition, let us introduce the following notation/terminology.

1) **Support.** Let $m = x_1^{e_1} \cdot x_2^{e_2} \cdot \ldots \cdot x_N^{e_N}$ in $\mathbb{F}[x_1, x_2, \ldots, x_N]$ be a monomial. The support of m, denoted $\mathrm{Supp}(m)$ is the subset of variables appearing in it, i.e.

$$\mathrm{Supp}(m) \stackrel{\text{def}}{=} \{i \ : \ e_i \geq 1\} \subseteq [N].$$

The support size of a polynomial f, denoted $|\mathrm{Supp}(f)|$ is the maximum support size of any monomial appearing in f.

2) **Substitution maps.** Let $R \subseteq [N]$ be a set. The substitution map $\sigma_R : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ is the map which sets all the variables in R to zero, i.e. $\sigma_R(f) \stackrel{\text{def}}{=} f|_{x_i=0\ \forall i\in R}$. Formally, $\sigma_R : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ is a homomorphism such that for any monomial $m \in \mathbb{F}[\mathbf{x}]$, $\sigma_R(m) = m$ if the monomial m is supported outside R and is zero otherwise.

So the above discussion can now be summarized as follows. Let $t = \Theta(\sqrt{d})$ be a suitable integer. By choosing a set R at random in the above manner and applying σ_R to the identity (1), we obtain (with high probability) another identity

$$\sigma_R(f) = \sum_{i=1}^{s} \prod_{j} \sigma_R(Q_{ij}), \qquad (2)$$

where $\forall i, j \ : \ \sigma_R(Q_{ij})$ is homogeneous and $|\mathrm{Supp}(\sigma_R(Q_{ij}))| \leq t$. In this manner our problem reduces to proving lower bounds for representations of the form (2) which we refer to as t-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuits.

Lower bounds for low support homogeneous $\Sigma\Pi\Sigma\Pi$ circuits. We first note that the degree of a polynomial is an upper bound on its support size. From prior work by [1], [2], [3], [4], we have lower bounds for similar looking representations but in which the *degree* of every Q_{ij}, rather than its support, was bounded by t. We build on these works to devise a complexity measure that we refer to as *dimension of projected shifted partials*. We define this measure as follows.

1) **The projection map.** Let $s, e \geq 1$ be integers. The linear map $\pi_{e,s} : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ maps a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ to the component of degree e and support s of $f(\mathbf{x})$. Formally, it is defined as follows. We need to only specify it for monomials and it then extends by linearity to all of $\mathbb{F}[\mathbf{x}]$. For a monomial $m \in \mathbb{F}[\mathbf{x}]$, $\pi_{e,s}(m)$ equals m if m has degree exactly e and support size exactly s and zero otherwise.

2) **The Complexity Measure.** Let k, ℓ, e be integer parameters and $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a multivariate polynomial. We denote by $\boldsymbol{\partial}^{=k} f$ the set of all k-th order partial derivatives of f. Let $\mathbf{x}^{(=\ell,=s)}$ denote the set of monomials of degree exactly ℓ and support exactly s over the variables in $\mathbf{x}$. Let $A, B \subseteq \mathbb{F}[\mathbf{x}]$ be any two sets of polynomials. $A \cdot B$ stands for the set

$$A \cdot B \ \stackrel{\text{def}}{=} \ \{f \cdot g \ : \ f \in A \text{ and } g \in B\}.$$

For a linear map $\pi : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$, $\pi(A)$ denotes the set

$$\pi(A) \stackrel{\text{def}}{=} \{\pi(f) \ : \ f \in A\}.$$

The *dimension of projected shifted partial derivatives* of f (DPSP for short) is defined as

$$\mathsf{DPSP}_{k,\ell,e}(f) \stackrel{\text{def}}{=} \dim\left(\pi_{\ell+e,\ell+e}\left(\mathbf{x}^{(=\ell,=\ell)} \cdot \boldsymbol{\partial}^{=k} f\right)\right).$$

Recap - lower bounds for low degree depth four. It was shown in [2] that if f can be expressed as a sum of a small number of products of low degree polynomials, i.e. when the Q_{ij}'s have low degree, then the dimension of shifted partial derivatives of f, namely $\dim\left(\mathbf{x}^{(=\ell)} \cdot \boldsymbol{\partial}^{=k} f\right)$. is small. This was done by observing that there exist a *relatively small* number of sets $S_1, S_2, \ldots, S_m \subseteq \mathbb{F}[\mathbf{x}]$ such that every polynomial in $\boldsymbol{\partial}^{=k} f$ is in the $\mathbb{F}$-span of the polynomials in $\bigcup_{i\in[m]} S_i$. Moreover for each set S_i, the polynomials within S_i share a *large* common factor. This implies that for each i, $\dim(\mathbf{x}^{(=\ell)} \cdot S_i)$ is small and thereby that $\dim\left(\mathbf{x}^{(=\ell)} \cdot \boldsymbol{\partial}^{=k} f\right)$ is small as well. Combining this with a lower bound estimate on $\dim\left(\mathbf{x}^{(=\ell)} \cdot \boldsymbol{\partial}^{=k} f\right)$, one could then obtain a lower bound for expressing f as a sum of products of low degree polynomials.

Lower bounds for low support depth four. We modify the complexity measure used previously so that it works even for a sum of product of low support factors. Intuitively, by shifting (i.e. multiplying) the partial derivatives by a carefully chosen set of monomials and then projecting them to another appropriate set of monomials, we are able to

[13] This reduction from homogeneous $\Sigma\Pi\Sigma\Pi$ formulas to low support $\Sigma\Pi\Sigma\Pi$ formulas was communicated to the first author by Avi Wigderson. It was recently exploited by Kumar and Saraf in [7] and also independently discovered by some of the other authors of the present work.

ignore high-degree factors while paying a relatively small cost (in terms of the dimension of the relevant spaces). Specifically, we show that this measure is relatively small for t-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuits (Corollary 9 in section IV). We then find an explicit polynomial f whose projected shifted partials has large dimension and thereby obtain a $2^{\Omega\left(\frac{d}{t}\cdot\log N\right)}$ lower bound for t-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuits computing f. We further show that the dimension of projected shifted partials of f remains quite large even under random restrictions (with high probability) thereby obtaining a $2^{\Omega\left(\sqrt{d}\cdot\log N\right)}$ lower bound overall for general homogeneous $\Sigma\Pi\Sigma\Pi$ circuits.

Remark 2. In a way, the random restriction together with the projection map help us carry out an 'indirect reduction' from homogeneous $\Sigma\Pi\Sigma\Pi$ formulas to homogeneous $\Sigma\Pi\Sigma\Pi^{[t]}$ formulas thereby bypassing the apparent hurdle pointed out in [5]. This comes at a price though - the projection map also severely restricts the monomials with which we can shift the partial derivatives. To handle this loss, we are required to do a tighter analysis to lower bound the dimension of the projected shifted partials of our explicit family of polynomials.

Lower bounding the dimension of projected shifted partials. A crucial component of this proof is to show that the dimension of projected shifted partials of our explicit family of polynomials is large[14]. From the definition, it follows that this quantity is equal to the rank of a certain matrix $M(f)$ whose rows correspond to the polynomials in $\pi_{\ell+e,\ell+e}\left(\mathbf{x}^{(=\ell,=\ell)}\cdot\boldsymbol{\partial}^{=k}f\right)$ in the natural way - each row is just the coefficient vector of the corresponding polynomial. In order to show that $\text{rank}(M(f))$ is large for our choice of f, we show that the columns of the matrix $M(f)$ are *almost orthogonal*[15], i.e. the dot product of any two distinct column vectors is small relatively to their lengths, and thereby deduce that it must have high rank[16]. The latter deduction goes as follows. Let $B(f) \stackrel{\text{def}}{=} M(f)^T\cdot M(f)$. Note that the (i,j)-th entry of $B(f)$ is the dot product of the the i-th and the j-th columns of $M(f)$ and the fact that the columns of $M(f)$ are almost orthogonal means that $B(f)$ is *diagonally dominant* - i.e, its diagonal entries are much larger than the off-diagonal entries. Note also that the rank of $B(f)$ is a lower bound on the rank of $M(f)$. Noga Alon [29] gave the following lower bound on the rank of diagonally dominant matrices (via an application of Cauchy-Schwarz on the vector of nonzero eigenvalues of $B(f)$):

$$\text{rank}(B(f)) \geq \frac{\text{Tr}(B(f))^2}{\text{Tr}(B(f)^2)}.$$

For our application, we then estimate $\text{Tr}(B(f))^2$ and $\text{Tr}(B(f)^2)$ and show that the ratio is large for our choice of f (even under random restrictions). This then yields the claimed lower bound on the size of homogeneous depth four formulas computing f.

Organization. The rest of the paper is devoted to fleshing out this outline into a full proof. For the sake of clarity of exposition, we first focus our attention on t-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuits. We first give an upper bound (in section IV) on the dimension of projected shifted partials of any homogeneous t-supported $\Sigma\Pi\Sigma\Pi$ circuit $\mathcal{C}$. In section V we then give the construction of our polynomial f and show that choosing the parameters appropriately yields a lower bound of $2^{\Omega\left(\frac{d}{t}\cdot\log N\right)}$ on the top fanin of homogeneous t-supported $\Sigma\Pi\Sigma\Pi$ circuits computing f - assuming that f has large projected shifted partials dimension. In section VI we show that our polynomial does indeed have a large projected shifted partials dimension. Finally, in section VII we analyze the effect of random restrictions and show that the dimension of shifted partials of f remains large under random restrictions thereby yielding a $2^{\Omega\left(\sqrt{d}\cdot\log N\right)}$ lower bound overall.

III. Preliminaries

Vector Spaces of Polynomials and linear maps. Let $U,V\subseteq\mathbb{F}[\mathbf{x}]$ be two vector spaces of polynomials and let $\pi:\mathbb{F}[\mathbf{x}]\mapsto\mathbb{F}[\mathbf{x}]$ be a linear map. Define

$$\pi(U)\stackrel{\text{def}}{=}\{\pi(f)\quad:\quad f\in U\}\subseteq\mathbb{F}[\mathbf{x}].$$

Note that $\pi(U)$ must be a subspace in $\mathbb{F}[\mathbf{x}]$. Also define

$$U+V\stackrel{\text{def}}{=}\mathbb{F}\text{-span}\left(\{f+g\quad:\quad f\in U,g\in V\}\right).$$

Let us record a basic fact from linear algebra as applicable to us.

Proposition 3. *Let $U,V\subseteq\mathbb{F}[\mathbf{x}]$ be two vector spaces of polynomials and let $\pi:\mathbb{F}[\mathbf{x}]\mapsto\mathbb{F}[\mathbf{x}]$ be any linear map. Then*

$$\pi(U+V)=\pi(U)+\pi(V)\quad\textit{and}\quad\dim(\pi(U))\leq\dim(U).$$

Numerical estimates.

Proposition 4 (Stirling's Formula, cf. [30])**.** $\ln(n!) = n\ln n - n + O(\ln n)$

Stirling's formula can be used to obtain the following estimates.

[14] In prior work one needed to estimate the dimension of shifted partials of a given f and it was shown that in many interesting cases this could be successfully accomplished simply by counting leading monomials. This corresponds to lower bounding the rank of $M(f)$ by finding a submatrix which is upper triangular. We do not know if the modified measure allows one to embed large triangular submatrices inside $M(f)$ but if this can be done then it could be one way to prove the same lower bound over arbitrary fields.

[15] Our inspiration for this method of lower bounding the rank comes from the beautiful recent work by Barak, Dvir, Wigderson and Yehudayoff [27] and a subsequent improvement by Dvir, Saraf and Wigderson [28],

[16] Note that if the columns of $M(f)$ were exactly orthogonal (i.e. the dot product is zero) then its rank would equal the number of columns.

Lemma 5. *Let $a(n),\ f(n),\ g(n) : \mathbb{Z}_{>0} \to \mathbb{Z}$ be integer valued function such that $(|f| + |g|) = o(a)$. Then,*

$$\ln\frac{(a+f)!}{(a-g)!} \quad = \quad (f+g)\ln a \ \pm\ O\left(\frac{f^2+g^2}{a}\right)$$

Depth-4 arithmetic formulas. We recall some basic definitions regarding arithmetic circuits and formulas; for a more thorough introduction, see the survey [9]. Let Y be a finite set of variables. An arithmetic formula C over $\mathbb{F}$ is a rooted tree the leaves of which are labelled by variables in Y and elements of the field $\mathbb{F}$, and internal nodes (called *gates*) by $+$ and $\times$. This computes a polynomial $f \in \mathbb{F}[Y]$ in a natural way. By the *size* of a formula, we mean the number of vertices in the tree, and by the *depth* of a formula, we mean the longest root-to-leaf path in the tree. Our focus here is on *depth-4 formulas* [17], which are formulas that can be written as sums of products of sums of products, otherwise known as $\Sigma\Pi\Sigma\Pi$ formulas. We will prove lower bounds for *homogeneous* $\Sigma\Pi\Sigma\Pi$ formulas which are $\Sigma\Pi\Sigma\Pi$ formulas such that each node computes a homogeneous polynomial (i.e. a polynomial whose every monomial has the same degree). Given a $\Sigma\Pi\Sigma\Pi$ formula, the layer 0 nodes will refer to the leaf nodes, the layer 1 nodes to the Π-gates just above the leaf nodes, etc. The *top fan-in* refers to the fan-in of the root node on layer 4. We also consider variants of $\Sigma\Pi\Sigma\Pi$ formulas with bounds on the fan-ins of the Π gates. By $\Sigma\Pi^{[D]}\Sigma\Pi^{[t]}$ formulas, we mean $\Sigma\Pi\Sigma\Pi$ formulas where the fan-ins of the layer 1 and layer 3 Π gates are *at most* t and D respectively.

IV. Upper bounding the measure for low support $\Sigma\Pi\Sigma\Pi$ circuits.

Consider a homogeneous $\Sigma\Pi\Sigma\Pi$ ciruit $\mathcal{C}$ of the form

$$\mathcal{C} = \sum_i \prod_j Q_{ij}, \quad \text{where } |\mathrm{Supp}(Q_{ij})| \le t \text{ for every } Q_{ij}.$$

We will see how the measure defined in Section II can be used to pinpoint a weakness of such a circuit. Let us first note two simple properties of our projection map π. The next two propositions are straightforward to verify and we omit the proof.

Proposition 6. *Let $Q(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a homogeneous polynomial of degree d and $m(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a monomial of degree a. Then $\pi_{d+a,d+a}\left(m(\mathbf{x}) \cdot Q(\mathbf{x})\right)$ equals 0 if* if $|\mathrm{Supp}(m)| <$ *a and it equals $m(\mathbf{x}) \cdot \sigma_A(\pi_{d,d}(Q)) = m(\mathbf{x}) \cdot \pi_{d,d}(\sigma_A(Q))$ if $A \stackrel{\text{def}}{=} \mathrm{Supp}(m)$ has size a.*

[17]we will interchangeably use the terms 'depth-4 circuits', as depth-4 circuits can be converted to depth-4 formulas with only a polynomial blow-up in size

Our measure, namely $\mathsf{DPSP}_{k,\ell,e}(f) \stackrel{\text{def}}{=} \pi_{\ell+e,\ell+e}(\mathbf{x}^{=(\ell,\ell)} \cdot \partial^{=k} f)$ has the following properties.

Proposition 7. *For any pair of polynomials $f, g \in \mathbb{F}[\mathbf{x}]$ and any 3-tuple of integers k, ℓ, e*

1) **[Subadditivity.]** $\mathsf{DPSP}_{k,\ell,e}(f+g) \le \mathsf{DPSP}_{k,\ell,e}(f) + \mathsf{DPSP}_{k,\ell,e}(g)$.
2) **[Subprojectivity.]** *If $g = \sigma_A(f)$ for some subset A, i.e. g is obtained from f by setting some subset A of variables to zero, then $\mathsf{DPSP}_{k,\ell,e}(g) \le \mathsf{DPSP}_{k,\ell,e}(f)$.*
3) **[Zeroness for low-support polynomials.]** *If all monomials of f have support strictly less than e then $\mathsf{DPSP}_{k,\ell,e}(f) = 0$.*

We will now upper bound how large the measure can be for any term T of a low support homogeneous $\Sigma\Pi\Sigma\Pi$-circuit $\mathcal{C} = T_1 + T_2 + \ldots + T_s$. Let us focus on a term T in our t-supported homogeneous $\Sigma\Pi\Sigma\Pi$-circuit $\mathcal{C}$ so that T is of the form $T = Q_1 \cdot Q_2 \cdot \ldots \cdot Q_m, \quad |\mathrm{Supp}(Q_i)| \le t \quad \text{for each } i \in [m]$, where the Q_i's are homogeneous polynomials and T is of degree d. We will now upper bound $\mathsf{DPSP}_{k,\ell,d-k}(T)$.

Preprocessing. First note that we can assume without loss of generality that every Q_i (except perhaps one) has degree at least $t/2$ for if not, then we can replace two such Q_i's by their product $(Q_i \cdot Q_j)$. The product $(Q_i \cdot Q_j)$ has degree at most t and therefore also support at most t. Continuing this process of combining factors of small degree, we end up in a situation where every factor (except perhaps one) has degree at least $t/2$. In such a situation, the number of factors m can at most be $m \le 1 + \frac{d}{t/2} = 1 + \frac{2d}{t}$.

Lemma 8. *Let T be as in the above paragraph. For any k and any $\ell < \frac{N}{2} - kt$ we have*

$$\mathsf{DPSP}_{k,\ell,d-k}(T) \le \binom{2d/t+1}{k} \cdot \binom{N}{\ell + k \cdot t}.$$

Combining the above upper bound for a term with the subadditivity of our measure we immediately get:

Corollary 9. *Let $\mathcal{C}$ be a t-supported degree d homogeneous $\Sigma\Pi\Sigma\Pi$ circuit with top fanin s, i.e $\mathcal{C}$ is a degree d homogeneous circuit of the form*

$$\mathcal{C} = \sum_{i=1}^{s} Q_{i1} \cdot Q_{i2} \cdot \ldots \cdot Q_{im_i}, \quad |\mathrm{Supp}(Q_{ij})| \le t.$$

Then for every k and every $\ell < \frac{N}{2} - kt$ we have

$$\mathsf{DPSP}_{k,\ell,d-k}(\mathcal{C}) \le s \cdot \binom{2d/t+1}{k} \cdot \binom{N}{\ell + k \cdot t}.$$

Consequently, for any N-variate homogeneous polynomial f of degree d, any homogeneous t-supported $\Sigma\Pi\Sigma\Pi$-circuit C computing f must have top fanin at least

$$s \ge \frac{\mathsf{DPSP}_{k,\ell,d-k}(f)}{\binom{2d/t+1}{k} \cdot \binom{N}{\ell+k\cdot t}}.$$

The proofs of Lemma 8 and Corollary 9 are omitted. In the next section we construct an explicit polynomial f for which $\mathsf{DPSP}_{k,\ell,d-k}(f)$ is large and then use the above to deduce a lower bound on the top fanin of any t-supported $\Sigma\Pi\Sigma\Pi$-circuit computing f.

V. The lower bound for low support homogeneous $\Sigma\Pi\Sigma\Pi$ circuits.

We will now construct an explicit homogeneous, multilinear polynomial f of degree d on $N = d^3$ variables for which our measure, namely $\mathsf{DPSP}_{k,\ell,d-k}(f)$ is large. We will then see that this implies that any t-supported $\Sigma\Pi\Sigma\Pi$-circuit computing f must have large top fanin.

A. The Construction of an Explicit Polynomial

Our explicit polynomial is parametrized by an integer parameter r that we call NW_r and it is a variant of the Nisan-Wigderson design polynomial from [3]. Let d be a prime power and $\mathbb{F}_d$ be the finite field of size d. Let $\mathbb{F}_{d^2} \supseteq \mathbb{F}_d$ be the quadratic extension field of $\mathbb{F}_d$. We refer to the elements of the finite field $\mathbb{F}_{d^2}$ simply as $\{1, 2, \cdots, d^2\}$ where the first d among these belong to the subfield $\mathbb{F}_d$. Fix an integer r. Our explicit polynomial is $\mathsf{NW}_r(x_{1,1}, x_{1,2}, \ldots, x_{d,d^2})$ which is defined to be

$$\sum_{h(z)\in\mathbb{F}_{d^2}[z],\deg(h)\le r} \quad \prod_{i\in[d]} x_{i,h(i)}.$$

From the definition above, it is clear that for all r, NW_r is an explicit homogeneous, multilinear polynomial of degree d on $N = d^3$ variables. our main technical lemma stated below is a lower bound on the dimension of projected shifted partials of the design polynomial NW_r.

Lemma 10. [Main Technical Lemma.] *Let NW_r be the Nisan-Wigderson design-based polynomial defined above. Over any field $\mathbb{F}$ of characteristic zero, for $r = \frac{d}{3}$ and $k = o(d)$ and $\ell = \frac{N}{2}\cdot\left(1 - \frac{k\ln d}{d}\right)$ we have* $\mathsf{DPSP}_{k,\ell,d-k}(\mathsf{NW}_r) \ge \frac{1}{d^{O(1)}} \cdot \min\left(\binom{N}{\ell+d-k}, \binom{d}{k}^2 \cdot d^k \cdot k! \cdot \binom{N}{\ell}\right)$.

The proof of the lemma is outlined in Section VI. We can use this lemma to deduce a lower bound on the top fanin of any t-supported homogeneous $\Sigma\Pi\Sigma\Pi$ circuit computing $\mathsf{NW}_{d/3}$.Consider a t-supported $\Sigma\Pi\Sigma\Pi$ circuit $\mathcal{C}$ of top fanin s computing $\mathsf{NW}_{d/3}$. We fix our choice of parameters as follows: $k = \delta \cdot \frac{d}{t}$ (for a small enough constant $\delta > 0$), $\ell = \frac{N}{2} \cdot \left(1 - \frac{k\ln d}{d}\right)$. By Lemma 10 and Corollary 9 and using some (omitted) computations we can prove

$$\begin{aligned} s &\ge \frac{\mathsf{DPSP}_{k,\ell,d-k}(\mathsf{NW}_{d/3})}{\binom{2d/t+1}{k} \cdot \binom{N}{\ell+k\cdot t}} \\ &\ge \frac{1}{d^{O(1)} \cdot \binom{2d/t+1}{k}} \min\left(\frac{\binom{d}{k}^2 d^k k! \binom{N}{\ell}}{\binom{N}{\ell+kt}}, \frac{\binom{N}{\ell+d-k}}{\binom{N}{\ell+kt}}\right) \\ &\ge 2^{\Omega\left(\frac{d}{t}\cdot\log N\right)} \end{aligned}$$

VI. Proof of the main technical lemma

In this section we give a proof sketch for Lemma 10, which shows that the dimension of projected shifted partial derivatives of the Nisan-Wigderson design based polynomial is large. Let $e \stackrel{\text{def}}{=} (d-k)$ throughout the rest of this section.

Preliminaries. Note that in the construction in section V of NW_r, there is a 1-1 correspondence between the variable indices in $[N]$ and points in $\mathbb{F}_d \times \mathbb{F}_{d^2}$, which we identify with $[d] \times [d^2]$. Being homogeneous and multilinear of degree d, the monomials of NW_r are in 1-1 correspondence with sets in $\binom{[N]}{d} \equiv \binom{[d]\times[d^2]}{d}$. Indeed, from the construction it is clear that the coefficient of any monomial in NW_r is either 0 or 1 and that there is a 1-1 correspondence between monomials in the support of NW_r and univariate polynomials of degree at most r in $\mathbb{F}_{d^2}[z]$. Now since two distinct polynomials of degree r over a field have at most r common roots we get:

Proposition 11. [A basic property of our construction.] *For any two distinct sets $D_1, D_2 \in \binom{[d]\times[d^2]}{d}$ in the support of NW_r, we have $|D_1 \cap D_2| \le r < \frac{e}{2}$ (for $r = d/3$ and $k = o(d)$.)*

Our goal for the remainder of this section is to lower bound $\mathsf{DPSP}_{k,\ell,d-k}(NW_r)$ which is defined to be $\dim\left(\pi_{\ell+d-k,\ell+d-k}\left(\mathbf{x}^{(=\ell,=\ell)} \cdot \boldsymbol{\partial}^{=k}\mathsf{NW}_r\right)\right)$.

Reformulating our goal in terms of the rank of an explicit matrix. Let f be any homogeneous multilinear polynomial of degree d on N variables. By multilinearity, the only derivatives of f that survive are those with respect to multilinear monomials. Thus we have

$$\boldsymbol{\partial}^{=k} f = \left\{\partial^C f \;:\; C \in \binom{[N]}{k}\right\}.$$

Note that every k-th order derivative of f is homogeneous and multilinear of degree $(d-k)$. Combining this with proposition 6 we get that $\pi_{\ell+d-k,\ell+d-k}(\mathbf{x}^{(=\ell,=\ell)} \cdot \boldsymbol{\partial}^{=k} f) = \left\{\mathbf{x}_A \cdot \sigma_A\left(\partial^C f\right) \;:\; A \in \binom{[N]}{\ell}, \quad C \in \binom{[N]}{k}\right\}$. Thus we have

Proposition 12. *For any homogeneous multilinear polynomial f of degree d on N variables and for all integers k and ℓ:* $\mathsf{DPSP}_{k,\ell,d-k}(f) = \dim\left(\left\{\mathbf{x}_A \cdot \sigma_A\left(\partial^C f\right) \;:\; A \in \binom{[N]}{\ell}, \quad C \in \binom{[N]}{k}\right\}\right)$.

Now the $\mathbb{F}$-linear dimension of any set of polynomials is the same as the rank of the matrix corresponding to our set of polynomials in the natural way. Specifically,

Proposition 13. *Let f be a homogeneous multilinear polynomial of degree d on N variables. Let k, ℓ be integers. Define a matrix $M(f)$ as follows. The rows of $M(f)$ are labelled by pairs of subsets $(A, C) \in \binom{[N]}{\ell} \times \binom{[N]}{k}$ and* *columns are indexed by subsets $S \in \binom{[N]}{\ell+e}$. Each row (A, C) corresponds to the polynomial*

$$f_{A,C} \stackrel{\text{def}}{=} \mathbf{x}_A \cdot \sigma_A\left(\partial^C f\right)$$

in the following way. The S-th entry of the row (A, C) is the coefficient of $\mathbf{x}_S$ in the polynomial $f_{A,C}$. Then,

$$\mathsf{DPSP}_{k,\ell,d-k}(f) = \textit{rank}(M(f)).$$

So our problem is equivalent to lower bounding the rank of the matrix $M(f)$ for our constructed polynomial f. Now note that the entries of $M(f)$ are coefficients of appropriate monomials of f and it will be helpful to us in what follows to keep track of this information. We will do it by assigning a label to each cell of $M(f)$ as follows. We will think of every location in the matrix $M(f)$ being labelled with either a set $D \in \binom{[N]}{d}$ or the label `InvalidSet` depending on whether that entry contains the coefficient of the monomial $\mathbf{x}_D$ of f or it would have been zero regardless of the actual coefficients of f. Specifically, let us introduce the following notation. For sets A, B define:

1)
$$A \setminus\!\!\setminus B = \begin{cases} A \setminus B & \text{if} B \subseteq A \\ \texttt{InvalidSet} & \text{otherwise} \end{cases}$$

2)
$$A \uplus B = \begin{cases} A \cup B & \text{if} B \cap A = \emptyset \\ \texttt{InvalidSet} & \text{otherwise} \end{cases}$$

Then the label of the $((A, C), S)$-th cell of $M(f)$ is defined to be the set $(S \setminus\!\!\setminus A) \uplus C$. Equivalently, if the label of a cell of the (A, C)-th row of M is a set D then the column must be the one corresponding to $S = (D \setminus\!\!\setminus C) \uplus A$ (if C is not a subset of D or if $(D \setminus\!\!\setminus C)$ and A are not disjoint then D cannot occur in the row indexed by (A, C)). For the rest of this section, we will refer to $M(\mathsf{NW}_r)$ simply as the matrix M. Our goal then is to show that the rank of this matrix M is reasonably close (within a $\mathsf{poly}(d)$-factor) of the trivial upper bound, viz. the minimum of the number of rows and the number of columns of M. It turns out that our matrix M is a relatively sparse matrix and we will exploit this fact by using a relevant lemma from real matrix analysis to obtain a lower bound on its rank.

The Surrogate Rank. Consider the matrix $B \stackrel{\text{def}}{=} M^T \cdot M$. Then B is a real symmetric, positive semidefinite matrix. From the definition of B it is easy to show that:

Proposition 14. *Over any field $\mathbb{F}$ we have $\textit{rank}(B) \leq \textit{rank}(M)$. Over the field $\mathbb{R}$ of real numbers we have $\textit{rank}(B) = \textit{rank}(M)$.*

So it suffices to lower bound the rank of B. By an application of Cauchy-Schwarz on the vector of nonzero eigenvalues of B, one obtains:

Lemma 15. *[29] Over the field of real numbers $\mathbb{R}$ we have:*

$$\textit{rank}(B) \geq \frac{\mathrm{Tr}(B)^2}{\mathrm{Tr}(B^2)}.$$

Let us call the quantity $\frac{\mathrm{Tr}(B)^2}{\mathrm{Tr}(B^2)}$ as the surrogate rank of M, denoted $\mathrm{SurRank}(M)$. It then suffices to show that this quantity is within a $\mathsf{poly}(d)$ factor of $U = \min(\binom{d^3}{\ell+e}, \binom{d^3}{\ell} \cdot \binom{d^3}{k})$.

A. *Bounding* $\mathrm{SurRank}(M)$.

We now bound $\mathrm{Tr}(B)$ and $\mathrm{Tr}(B^2)$ for $B = M^T \cdot M$ in order to bound $\mathrm{SurRank}(M)$.

Calculating $\mathrm{Tr}(B)$. Calculating $\mathrm{Tr}(B)$ is fairly straightforward. From the definition of the matrix B we have:

Proposition 16. *For any $0, \pm 1$ matrix M (i.e. a matrix all of whose entries are either 0, or $+1$ or -1) we have*

$$\mathrm{Tr}(B) = \mathrm{Tr}(M^T \cdot M) = \textit{number of nonzero entries in } \boldsymbol{M}.$$

Now we can calculate the number of nonzero entries in M by going over all sets $D \in \binom{[N]}{d} \cap \mathrm{Supp}(\mathsf{NW}_r)$, calculating the number of cells of M labelled with D and adding these up. This yields:

Proposition 17.

$$\mathrm{Tr}(B) = d^{2r+2} \cdot \binom{d}{k} \cdot \binom{N-e}{\ell}.$$

Calculating $\mathrm{Tr}(B^2)$. From the definition of $B = M^T \cdot M$ and expanding out the relevant summations we get:

Proposition 18. $\mathrm{Tr}(B^2) = \sum_{(A_1,C_1),(A_2,C_2)} \sum_{S_1,S_2} M_{(A_1,C_1),S_1} \cdot M_{(A_1,C_1),S_2} \cdot M_{(A_2,C_2),S_1} \cdot M_{(A_2,C_2),S_2}$, *where* $(A_1, C_1), (A_2, C_2) \in \left(\binom{[N]}{\ell} \times \binom{[N]}{k}\right)^2$ *and* $S_1, S_2 \in \binom{[N]}{\ell+e}^2$.

We will use the following notation in doing this calculation. For a pair of row indices $((A_1, C_1), (A_2, C_2)) \in \left(\binom{[N]}{\ell} \times \binom{[N]}{k}\right)^2$ and a pair of column indices $S_1, S_2 \in \left(\binom{[N]}{\ell+e}\right)^2$, the box $\mathbf{b}$ defined by them, denoted $\mathbf{b} = 2-\mathsf{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$ is the four-tuple of cells $(((A_1, C_1), S_1), ((A_1, C_1), S_2), ((A_2, C_2), S_1), ((A_2, C_2), S_2))$. Since all the entries of our matrix M are either 0 or 1 we have:

Proposition 19.

$$\mathrm{Tr}(B^2) = \textit{Number of boxes } \mathbf{b} \textit{ with all four entries nonzero}.$$

For a box $\mathbf{b} = 2-\mathsf{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$, its tuple of labels, denoted $\mathrm{labels}(\mathbf{b})$ is the tuple of labels of the cells $((A_1, C_1), S_1), ((A_1, C_1), S_2), ((A_2, C_2), S_1), ((A_2, C_2), S_2))$ in that order. In other words, labels($\mathbf{b}$) $= ((S_1 \setminus A_1) \uplus C_1, (S_2 \setminus A_1) \uplus C_1, (S_1 \setminus A_2) \uplus C_2, (S_2 \setminus A_2) \uplus C_2)$. We then have

Proposition 20. $\mathrm{Tr}(B^2)$ *equals the number of boxes*

$$\mathbf{b} = 2 - box((A_1, C_1), (A_2, C_2), S_1, S_2)$$

such that all the four labels in labels($\mathbf{b}$) *are valid sets in the support of our design polynomial* NW_r.

So our problem boils down to counting the number of boxes in which all the four labels are valid sets in the support of our polynomial NW_r.

A somewhat involved computation (which can be found in the full version) allows us to bound the number of such boxes and yields

$\mathrm{Tr}(B^2) \leq (2k^3 d) \cdot (d^{4r+4}) \cdot \max\left(\frac{1}{d^k \cdot k!} \cdot \binom{N-2e}{\ell}, \binom{N-2e}{\ell-e} \cdot \binom{d}{k}^2\right)$.

This means that $\mathrm{SurRank}(B)$ may be bounded as

$$\begin{aligned} \frac{\mathrm{Tr}(B)^2}{\mathrm{Tr}(B^2)} &\geq \frac{1}{2k^3 d} \min\left(\frac{\binom{d}{k}^2 d^k k! \binom{N-e}{\ell}^2}{\binom{N-2e}{\ell}}, \frac{\binom{N-e}{\ell}^2}{\binom{N-2e}{\ell-e}}\right) \\ &= \frac{1}{d^{O(1)}} \min\left(\binom{d}{k}^2 d^k k! \binom{N}{\ell}, \binom{N}{\ell+e}\right) \end{aligned}$$

where, the last equality follows from our choice of r, k and ℓ. This proves our main technical lemma, namely lemma 10.

VII. The lower bound for general homogeneous $\Sigma\Pi\Sigma\Pi$ circuits.

As hinted in the introduction, the problem of lower bounding the size of general homogeneous $\Sigma\Pi\Sigma\Pi$ circuits reduces to proving lower bounds for low support homogeneous $\Sigma\Pi\Sigma\Pi$ circuits. We now give some details of this reduction.

Definition 21. *For a real number $p \in (0,1]$, define the distribution $\mathcal{D}_p$ on subsets of $[N]$ obtained by choosing every element in $[N]$ independently at random with probability $(1-p)$. Thus, $\mathcal{D}_p : 2^{[N]} \mapsto (0,1]$ and for any $R \subseteq [N]$ we have $\mathcal{D}_p(R) = (1-p)^{|R|} \cdot p^{N-|R|}$.*

Let NW_r be the Nisan-Wigderson design polynomial as constructed in section V. Let us consider a homogeneous $\Sigma\Pi\Sigma\Pi$-circuit $\mathcal{C}$ computing it, i.e. consider any representation of NW_r of the form $\mathsf{NW}_r = \sum_i \prod_j Q_{ij}$, where the Q_{ij}'s are also homogeneous polynomials. Suppose that the total number of monomials in the polynomials Q_{ij}'s is bounded by $\mathfrak{s}$. Then the following may be verified.

Lemma 22. *For any homomorphism $\sigma_R : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ we have $\sigma_R(\mathsf{NW}_r) = \sum_i \prod_j \sigma_R(Q_{ij})$. For a set R chosen randomly according to $\mathcal{D}_p$, we have: $\Pr_{R\sim\mathcal{D}_p}[\exists i,j : \sigma_R(Q_{ij})$ contains a monomial of support more than $t] \leq \mathfrak{s} \cdot p^t$.*

Choosing the parameters t, p and $\mathfrak{s}$: Set $t = \sqrt{d}$, $p = d^{-\epsilon}$ (for an sufficiently small $\epsilon > 0$ to be fixed later), and suppose $\mathfrak{s} < 2^{\frac{\epsilon}{2}\sqrt{d}\log d}$. Then, $\Pr_{R\sim\mathcal{D}_p}[\exists i,j : \sigma_R(Q_{ij})$ contains a monomial of support more than $t] < 2^{-\frac{\epsilon}{2}\sqrt{d}\log d} \ll 1$.

This means, there are "plenty of" subsets R such that the circuit $\mathcal{C}$ restricted to the variables in R (i.e. $\sigma_R(\mathcal{C})$) is a t-supported homogeneous depth-4 circuit. If we can now show that there exists such an R that also keeps $\mathsf{DPSP}_{k,\ell,e}(\sigma_R(\mathsf{NW}_r))$ sufficiently close to $\min\left(\binom{N}{k} \cdot \binom{N}{\ell}, \binom{N}{\ell+e}\right)$ then we are done as before (by our discussion in Section V-A). The following lemma together with Lemma 22 show this. The proof is omitted.

Lemma 23. $\Pr_{R\sim\mathcal{D}_p}\left[\mathsf{DPSP}_{k,\ell,e}(\sigma_R(\mathsf{NW}_r)) < \frac{p^k}{d^{\Theta(1)}} \cdot \min\left(\binom{N}{k} \cdot \binom{N}{\ell}, \binom{N}{\ell+e}\right)\right] < \frac{1}{d^{\Theta(1)}}$.

By Lemma 22 and 23, and applying union bound, there exists a subset R such that $\sigma_R(\mathcal{C})$ is a t-supported homogeneous depth-4 circuit and $\mathsf{DPSP}_{k,\ell,e}(\sigma_R(\mathsf{NW}_r)) \geq \frac{p^k}{d^{\Theta(1)}} \cdot \min\left(\binom{N}{k} \cdot \binom{N}{\ell}, \binom{N}{\ell+e}\right)$.

If we choose a sufficiently small constant ϵ then $p^k = d^{-\epsilon k}$ is sufficiently large and the top fanin of $\sigma_R(\mathcal{C})$ (also the top fanin of $\mathcal{C}$) is $2^{\Omega(\sqrt{d}\cdot\log N)}$. Recall that we arrived at this conclusion assuming that the *total sparsity* of $\mathcal{C}$, which was denoted by $\mathfrak{s}$, is less than $2^{\epsilon/2\cdot\sqrt{d}\cdot\log d}$. Therefore, overall we get a lower bound of $2^{\Omega(\sqrt{d}\cdot\log N)}$ on the size of the homogeneous depth-4 circuit $\mathcal{C}$ computing NW_r.

VIII. Conclusion

As mentioned in the introduction, proving good enough lower bounds (specifically $2^{\omega(\sqrt{d}\cdot\log N)}$) for homogeneous depth four formulas yields superpolynomial lower bounds for general arithmetic circuits. Our lower bound of $2^{\Omega(\sqrt{d}\cdot\log N)}$ comes temptingly close to this threshold. So a very natural question would be to improve the exponent. A more modest aim might be to further understand the power and limitations of our techniques/complexity measure. With this intent we formulate a concrete conjecture that might serve as the goal of such an undertaking.

Conjecture 24. There exist a (family of) homogeneous polynomial(s) f of degree d in $N = d^{O(1)}$ variables which can be computed by $\mathsf{poly}(d)$-sized homogeneous circuits of depth six but for which any homogeneous circuit of depth four must have superpolynomial (in d) size.

Acknowledgements

NK would like to thank Avi Wigderson for many helpful discussions including pointing out the use of random restrictions to reduce a general homogeneous $\Sigma\Pi\Sigma\Pi$ circuit into one with low support. NL and SS would like to thank Hervé Fournier and Guillaume Malod for useful discussions. CS and SS would like to thank Arnab Bhattacharya and Ramprasad Saptharishi for their feedback and encouragement.

References

[1] N. Kayal, "An exponential lower bound for the sum of powers of bounded degree polynomials," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 19, p. 81, 2012.

[2] A. Gupta, P. Kamath, N. Kayal, and R. Saptharishi, "Approaching the chasm at depth four," in *Proceedings of the Conference on Computational Complexity (CCC)*, 2013.

[3] N. Kayal, C. Saha, and R. Saptharishi, "A super-polynomial lower bound for regular arithmetic formulas," *To appear in STOC*, 2014.

[4] H. Fournier, N. Limaye, G. Malod, and S. Srinivasan, "Lower bounds for depth 4 formulas computing iterated matrix multiplication," *To appear in STOC*, 2014.

[5] M. Kumar and S. Saraf, "The Limits of Depth Reduction for Arithmetic Formulas: Its all about the top fan-in," *To appear in STOC*, 2014.

[6] N. Kayal, N. Limaye, C. Saha, and S. Srinivasan, "Super-polynomial lower bounds for homogeneous depth four arithmetic formulas," *To appear in STOC*, 2014.

[7] M. Kumar and S. Saraf, "Superpolynomial lower bounds for general homogeneous depth 4 arithmetic circuits," Electronic Colloquium on Computational Complexity (ECCC), Tech. Rep. 181, 2013.

[8] S. Skyum and L. G. Valiant, "A Complexity Theory Based on Boolean Algebra," *J. ACM*, vol. 32, no. 2, pp. 484–502, 1985.

[9] A. Shpilka and A. Yehudayoff, "Arithmetic circuits: A survey of recent results and open questions," *Foundations and Trends in Theoretical Computer Science*, vol. 5, no. 3-4, pp. 207–388, 2010.

[10] L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff, "Fast Parallel Computation of Polynomials Using Few Processors," *SIAM J. Comput.*, vol. 12, no. 4, pp. 641–644, 1983.

[11] M. Agrawal and V. Vinay, "Arithmetic circuits: A chasm at depth four," in *FOCS*. IEEE Computer Society, 2008, pp. 67–75.

[12] P. Koiran, "Arithmetic circuits: The chasm at depth four gets wider," *Theor. Comput. Sci.*, vol. 448, pp. 56–65, 2012.

[13] S. Tavenas, "Improved bounds for reduction to depth 4 and depth 3," in *MFCS*, 2013, pp. 813–824.

[14] A. Gupta, N. Kayal, P. Kamath, and R. Saptharishi, "Arithmetic circuits: A chasm at depth three," in *Foundations of Computer Science (FOCS)*, 2013.

[15] N. Nisan and A. Wigderson, "Lower bounds on arithmetic circuits via partial derivatives," *Computational Complexity*, vol. 6, no. 3, pp. 217–234, 1997.

[16] R. Smolensky, "On Interpolation by Analytic Functions with Special Properties and Some Weak Lower Bounds on the Size of Circuits with Symmetric Gates," in *FOCS*, 1990, pp. 628–631.

[17] N. Nisan, "Lower bounds for non-commutative computation (extended abstract)," in *STOC*, 1991, pp. 410–418.

[18] D. Grigoriev and M. Karpinski, "An exponential lower bound for depth 3 arithmetic circuits," in *STOC*, 1998, pp. 577–582.

[19] D. Grigoriev and A. A. Razborov, "Exponential complexity lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields," in *FOCS*, 1998, pp. 269–278.

[20] R. Raz, "Multi-linear formulas for permanent and determinant are of super-polynomial size," *J. ACM*, vol. 56, no. 2, 2009.

[21] ——, "Separation of multilinear circuit and formula size," *Theory of Computing*, vol. 2, no. 1, pp. 121–135, 2006.

[22] R. Raz and A. Yehudayoff, "Lower Bounds and Separations for Constant Depth Multilinear Circuits," *Computational Complexity*, vol. 18, no. 2, pp. 171–207, 2009.

[23] M. Agrawal, C. Saha, R. Saptharishi, and N. Saxena, "Jacobian hits circuits: hitting-sets, lower bounds for depth-d occur-k formulas & depth-3 transcendence degree-k circuits," in *STOC*, 2012, pp. 599–614.

[24] X. Chen, N. Kayal, and A. Wigderson, "Partial Derivatives in Arithmetic Complexity and Beyond," *Foundations and Trends in Theoretical Computer Science*, vol. 6, no. 1-2, pp. 1–138, 2011.

[25] M. Kumar and S. Saraf, "Lower Bounds for Depth 4 Homogenous Circuits with Bounded Top Fanin," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 20, p. 68, 2013.

[26] N. Kayal, N. Limaye, C. Saha, and S. Srinivasan, "An exponential lower bound for homogeneous depth four arithmetic formulas," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 21, p. 5, 2014.

[27] B. Barak, Z. Dvir, A. Yehudayoff, and A. Wigderson, "Rank bounds for design matrices with applications toc ombinatorial geometry and locally correctable codes," in *STOC*, 2011, pp. 519–528.

[28] Z. Dvir, S. Saraf, and A. Wigderson, "Breaking the quadratic barrier for 3-LCCs over the Reals," *To appear in STOC*, 2014.

[29] N. Alon, "Perturbed Identity Matrices Have High Rank: Proof and Applications," *Combinatorics, Probability & Computing*, vol. 18, no. 1-2, pp. 3–15, 2009.

[30] D. Romik, "Stirlings approximation for n!: The ultimate short proof?" *The American Mathematical Monthly*, vol. 107, no. 6, p. 556557.