Electronic Colloquium on Computational Complexity, Report No. 91 (2013)

A super-polynomial lower bound for regular arithmetic formulas.

Neeraj Kayal
Microsoft Research India
neeraka@microsoft.com

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

Ramprasad Saptharishi
Microsoft Research India
ramprasad@cmi.ac.in

June 17, 2013

Abstract

We consider arithmetic formulas consisting of alternating layers of addition $(+)$ and multiplication $(\times)$ gates such that the fanin of all the gates in any fixed layer is the same. Such a formula Φ which additionally has the property that its formal/syntactic degree is at most twice the (total) degree of its output polynomial, we refer to as a *regular formula*. As usual, we allow arbitrary constants from the underlying field $\mathbb{F}$ on the incoming edges to a $+$ gate so that a $+$ gate can in fact compute an arbitrary $\mathbb{F}$-linear combination of its inputs. We show that there is an (n^2+1)-variate polynomial of degree $2n$ in VNP such that any regular formula computing it must be of size at least $n^{\Omega(\log n)}$.

Along the way, we examine depth four ($\Sigma\Pi\Sigma\Pi$) regular formulas wherein all multiplication gates in the layer adjacent to the inputs have fanin a and all multiplication gates in the layer adjacent to the output node have fanin b. We refer to such formulas as $\Sigma\Pi^{[b]}\Sigma\Pi^{[a]}$-formulas. We show that there exists an n^2-variate polynomial of degree n in VNP such that any $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$-formula computing it must have top fan-in at least $2^{\Omega(\sqrt{n}\cdot\log n)}$. In comparison, Tavenas [Tav13] has recently shown that every $n^{O(1)}$-variate polynomial of degree n in VP admits a $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$-formula of top fan-in $2^{O(\sqrt{n}\cdot\log n)}$. This means that any further asymptotic improvement either in our lower bound for such formulas (to say $2^{\omega(\sqrt{n}\log n)}$) or in Tavenas' upper bound (to say $2^{o(\sqrt{n}\log n)}$) will imply that VP is different from VNP.

ISSN 1433-8092

1 Introduction

Background. Arithmetic circuits and arithmetic formulas are the most natural and well-studied models for arithmetic computation. The objects of study here are families of multivariate polynomials

$$\{\ f_n(x_1, x_2, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots, x_n]\ :\ n \geq 1\ \}$$

and the goal is to estimate the complexity of computing them via arithmetic circuits and formulas. Such a family is said to be in the class VP if f_n has degree at most $\mathrm{poly}(n)$ and can be computed by an arithmetic circuit of size $\mathrm{poly}(n)$. It is said to be in VNP[1] if it can be expressed as

$$f_n(\mathbf{x}) = \sum_{\mathbf{y} \in \{0,1\}^m} g_{n+m}(\mathbf{x}, \mathbf{y})\ , \quad \text{where } m = |\mathbf{y}| = \mathrm{poly}(n) \tag{1}$$

and $g_{n+m}(\mathbf{x}, \mathbf{y}) \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$ is (a member of a family) in VP. A central open problem is to determine if VP equals VNP, i.e. to determine if every polynomial in VNP has a polynomial-sized circuit. Note, however, that superpolynomial lower bounds even for arithmetic formulas are not known for *any* explicit function and that questions of this type are considered to be among the most challenging open problems in theoretical computer science[2]. In this work, we consider arithmetic formulas satisfying some additional (but natural and seemingly mild) restrictions that we refer to as *regular* arithmetic formulas. Roughly speaking, a regular arithmetic formula is one which has low formal/syntactic degree[3] and in which the underlying tree is "well-balanced". We build and improve upon the work of [Kay12, GKKS13a] to show superpolynomial lower bounds for such formulas for an explicit family of polynomials in VNP.

Regular Formulas - definition. We now make precise the restrictions on arithmetic formulas imposed here and argue that these restrictions, though not without loss of generality, are nevertheless fairly natural and relatively mild. First note that in any formula, by collapsing two adjacent $+$ gates (resp. $\times$ gates) into a single gate (with slightly larger fanin), we can assume without loss of generality that the formula consists of alternating layers of addition and multiplication gates. Hence throughout this paper we will assume that all arithmetic formulas are in this normal form. We will denote an arithmetic formula with Δ layers by a sequence of Δ symbols wherein each symbol (either Σ or Π) denotes the nature of the gates at the corresponding layer and the leftmost

[1] As its name suggests, the class VNP is an algebraic analog of the Boolean class NP. The source of this analogy is that the definition of VNP in equation (1) is very similar to that of NP which can be defined as consisting of (families of) boolean functions $f_n(\mathbf{x})$ that can be expressed as

$$f_n(\mathbf{x}) = \bigvee_{\mathbf{y} \in \{0,1\}^m} g_{n+m}(\mathbf{x}, \mathbf{y}), \quad \text{where } m = |\mathbf{y}| = \mathrm{poly}(n)$$

and $g_{n+m}(\mathbf{x}, \mathbf{y})$ is (a member of a family) in P.

[2] Of course the same is also true for the study of Boolean functions and their complexity (with respect to Boolean circuits and formulae, or Turing machines), but in the Boolean case we have a better understanding of the difficulty (via results on relativation [BGS75], natural proofs [RR94] and algebrization [AW09]). The analogous problems in arithmetic complexity are however much more structured and there has always been more hope for progress in the arithmetic setting.

[3] The formal or syntactic degree of a circuit is the formal degree of its output node; the formal degree of a node being defined inductively in the natural manner - leaf nodes have formal degree 1 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.

symbol indicates the nature of the gate at the output layer. When all the gates in a particular layer have the same fanin, we will use an integer superscript on the corresponding Σ or Π symbol to denote the common fanin of the gates in that layer. For example a $\Sigma^{[s]}\Pi^{[a]}\Sigma\Pi^{[b]}$-formula computes a polynomial of the form

$$f \quad = \quad \sum_{i=1}^{s} (Q_{i1} \cdot Q_{i2} \cdot \ldots \cdot Q_{ia}) \qquad \text{where } \deg(Q_{ij}) \leq b \text{ for all } i, j.$$

Definition 1. Regular Formula. *Let Φ be a $\Sigma^{[a_1]}\Pi^{[p_1]}\Sigma^{[a_2]}\Pi^{[p_2]} \ldots \Sigma^{[a_\Delta]}\Pi^{[p_\Delta]}\Sigma^{[a_{\Delta+1}]}$-formula. Note that the size of such a formula is $(\prod_{i\in[\Delta+1]} a_i) \cdot (\prod_{i\in[\Delta]} p_i)$ and the formal degree is $(\prod_{i\in[\Delta]} p_i)$. We will say that such a formula is $(a_1, p_1, a_2, p_2, \ldots, a_\Delta, p_\Delta, a_{\Delta+1})$-regular, or just regular for short, if in addition the formal degree $(\prod_{i\in[\Delta]} p_i)$ is at most twice[4] the (total) degree of the output polynomial.*

We now make some remarks on the severity of each of the imposed restrictions. First the restriction that all gates at the same layer must have the same fanin is *in itself* not a restriction at all - for example, it follows from [GKQ13, Proposition 5] that any formula can be converted into a layered formula where all fanins are exactly two with only a polynomial increase in size (albeit at the loss of homogeneity). Much more substantial is the restriction that the formal degree is comparable to the degree of the output polynomial. Even this is not a serious restriction for circuits and branching programs as any branching program/circuit can be *homogenized*[5] with only a polynomial increase in size. For formulas, however, homogenization is not known to be efficient and Raz [Raz10] comes up with the tightest analysis of the loss incurred in the natural homogenization of an arithmetic formula.

Regular Formulas - motivation. It turns out that for many interesting families of polynomials such as the determinant and permanent, the best known formulas are regular or can be made so without *any asymptotic loss*. For example, the best known circuit for the $n \times n$ permanent known as Ryser's formula [Rys63] is $(2^n, n, n)$-regular while the best known formula for the $n \times n$ determinant is of depth $(2 \log n + 1)$ and is $(n^3, 2, n^3, 2, \ldots, n^3, 2, n^3)$-regular. A notable exception however is the family of elementary symmetric polynomials of degree say n in n^2 variables which admits a $\Sigma^{[n^2+1]}\Pi^{[n^2]}\Sigma^{[2]}$-formula which is not regular as per our definition because the formal degree (n^2) is much larger than the degree of the output polynomial (n). Besides the fact that the best known formulas for many natural polynomials are regular, it turns out that any formula - in fact any algebraic branching program - can be converted into a regular formula with a relatively small loss in size. Specifically, every $n^{O(1)}$-sized arithmetic branching program can be converted to a regular formula of size $n^{O(\log n)}$ (Proposition 5) which also implies that any $n^{O(1)}$-sized arithmetic circuit can be converted to a regular formula of size $n^{O(\log^2 n)}$. Hence the class of regular formulas seems to be natural and quite strong, and understanding its computational power is a worthwhile

[4] There is nothing special about two, and in fact formal degree being bounded by a constant multiple of the degree of the output polynomial would also suffice

[5] Recall that a multivariate polynomial is said to be homogeneous if all its monomials have the same total degree. An arithmetic circuit/formula is said to be *homogeneous* if the polynomial computed at every internal node of the circuit/formula is a homogeneous polynomial. It is a folklore result (cf. the survey by Shpilka and Yehudayoff [SY10]) that as far as computation by polynomial-sized arithmetic circuits of unbounded depth is concerned one can assume without loss of generality that the circuit is homogeneous. Specifically, if a homogeneous polynomial f of degree d can be computed by an (unbounded depth) arithmetic circuit of size s, then it can also be computed by a homogeneous circuit of size $O(d^2 \cdot s)$.

endeavour.

Prior work - Lower Bounds. The best known lower bounds for arithmetic circuits is $\Omega(n \cdot \log n)$ ([BS83]) and for formulas is $\Omega(n^3)$ ([Kal85]). Stronger lower bounds have been obtained for restricted subclasses of circuits/formulas. Nisan [Nis91] gave an exponential lower bound for *noncommutative* arithmetic formulas. In the setting of multilinear formulas, Raz [Raz09] showed an $n^{\Omega(\log n)}$ lower bound for the determinant family, Det_n. Subsequently, Raz and Yehudayoff [RY08] showed an $2^{n^{\Omega(1/d)}}$ lower bound for depth-d multilinear circuits. Nisan and Wigderson [NW97] gave an exponential lower bound for any homogeneous depth-3 formula computing Det_n or Perm_n. Grigoriev and Karpinski [GK98], and a follow-up by Grigoriev and Razborov [GR00] proved exponential lower bounds for arbitrary depth-3 formulas over any fixed finite field. Over infinite fields, we only have a $\Omega(n^2)$ lower bound [SW01] for depth-3 formulas. Gupta, Kamath, Kayal and Saptharishi [GKKS13a] showed an $\exp(\Omega(\sqrt{n}))$ top fan-in lower bound for $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuits computing Det_n or Perm_n. Using this, Kumar and Saraf [KS13] proved an exponential size lower bound for homogeneous depth-4 circuits of bounded top fan-in, with no restriction on the bottom fan-in.

Prior work - Depth reduction. Complementing these lower bounds is a parallel stream of research that studies the loss in size incurred in converting a general circuit/formula into a more restricted but highly structured (and hopefully easier to analyze) circuit/formula. Valiant, Skyum, Berkowitz and Rackoff [VSBR83] showed that every polynomial in VP can in fact be computed by bounded fanin homogeneous circuits of depth $O(\log^2 n)$. This theme has been pursued further in a recent series of results [AV08, Raz10, Koi12, GKKS13b, Tav13] that convert general arithmetic circuits/formulas of into very shallow and well-structured circuits (of depth four or three) and obtain nontrivial upper bounds on the size of the resulting circuit. Most relevant here is the work of Tavenas [Tav13] who built upon and improved [AV08, Koi12] and showed that any $n^{O(1)}$-variate polynomial of degree n in VP can also be computed by a $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$-formula of top fan-in $2^{O(\sqrt{n}\cdot\log n)}$.

Our results. We give here a superpolynomial lower bound for regular arithmetic formulas.

Theorem 1. *Let $\mathbb{F}$ be any field. There is an explicit family of (n^2+1)-variate polynomials of degree $2n$ which belongs to the class* VNP*, that requires regular arithmetic formulas of size $n^{\Omega(\log n)}$ to compute it.*

In an intermediate step, we build on [GKKS13a] and obtain an improved lower bound for depth four ($\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$) formulas.

Theorem 2. *Let $\mathbb{F}$ be any field. Let $\alpha \in \mathbb{Z}_{\geq 1}$ be any fixed positive integer. For any $t = t(n)$ satisfying $\log^2 n < t < \frac{n}{100}$, there is an explicit family $\mathcal{F}_t$ of n^2-variate polynomials of degree n over $\mathbb{F}$ in* VNP *such that any $\Sigma\Pi^{[\alpha\cdot(n/t)]}\Sigma\Pi^{[t]}$-circuit computing it must have top fan-in at least $2^{(\Omega((n/t)\cdot\log n))}$. In particular, there exists an explicit family of n^2-variate polynomials of degree n over $\mathbb{F}$ in* VNP *such that any $\Sigma\Pi^{[\alpha\sqrt{n}]}\Sigma\Pi^{[\sqrt{n}]}$-circuit computing it must have top fanin at least $2^{\left(\Omega\left(\sqrt{n}\cdot\log n\right)\right)}$.*

Remarks.

1. **Relevance to** VP **versus** VNP**.** Combined with the result of Tavenas, this means that any further asymptotic improvement either in our top fan-in lower bound for $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$

circuits (to say $2^{\omega(\sqrt{n}\log n)}$) or in Tavenas' upper bound (to say $2^{o(\sqrt{n}\log n)}$) will imply that VP is different from VNP.

2. **A generalization.** As in [GKKS13a], there is a slightly more general version of Theorem 2 which is as follows. Let $m = \alpha \cdot \sqrt{n}$. There exists an explicit family $\{F_n(\mathbf{x}) \in \mathbb{F}[\mathbf{x}] \ : \ n \geq 1\}$ of n^2-variate polynomials of degree n over $\mathbb{F}$ in VNP such that for any expression of the form

$$F_n(\mathbf{x}) = \sum_{i=1}^{s} G_i\left(Q_{i1}, \cdots, Q_{im}\right) \tag{2}$$

where each $G_i \in \mathbb{F}[u_1, u_2, \ldots, u_m]$ is an arbitrary m-variate polynomial and each $Q_{ij} \in \mathbb{F}[\mathbf{x}]$ is a polynomial of degree at most $\sqrt{n}$ over the n^2 variables of F_n. Then the number of summands s must be at least $\exp\left(\Omega\left(\frac{n}{t}\cdot\log n\right)\right)$. Note that Theorem 2 is a special case of this where each G_i is just the product of its inputs, i.e.

$$G_i(u_1, u_2, \ldots, u_m) = u_1 \cdot u_2 \cdot \ldots \cdot u_m.$$

3. **Tightness of the depth four lower bound.** Our depth four lower bound is tight in the sense that the family $\mathcal{F}_t$ in Theorem 2 can be computed by $\Sigma\Pi^{[(n/t)]}\Sigma\Pi^{[t]}$-circuits of size $2^{O((n/t)\cdot\log n)}$. Indeed, from our construction, it will be clear that the n-th degree polynomial F_n in $\mathcal{F}_t$ has $2^{O((n/t)\cdot\log n)}$ monomials so that F_n can in fact be computed by depth two circuits (and therefore also $\Sigma\Pi^{[(n/t)]}\Sigma\Pi^{[t]}$-circuits) of size $2^{O((n/t)\cdot\log n)}$.

4. **Allowing for larger formal degree.** For simplicity of exposition, we chose to state the above theorem with α being a constant but it turns out that our lower bound holds even if α is n^{δ} for a tiny $\delta > 0$.

The rest of this paper is devoted to a proof of the above two theorems.

2 Brief sketch of ideas

Lower bounds for depth four circuits. Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a multivariate polynomial of degree n and let $t = t(n)$ be an integer. Our first step is to prove Theorem 2, i.e. to give an improved top fan-in lower bound for depth four, $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuits. We proceed as in [GKKS13a] and use the same complexity measure called the dimension of shifted partial derivatives defined as follows. For integers $k, \ell \in \mathbb{Z}_{\geq 0}$ the dimension of ℓ-shifted k-th order derivatives of F is

$$\dim(\langle \boldsymbol{\partial}^{=k} F\rangle_{\leq \ell}) \quad \stackrel{\text{def}}{=} \quad \dim\left\{\mathbf{x}^{\mathbf{i}}\cdot\frac{\partial^{|\mathbf{j}|}F}{\partial\mathbf{x}^{\mathbf{j}}} \ : \ |\mathbf{j}| = k,\ |\mathbf{i}| \leq \ell\right\}$$

In other words, $\dim(\langle \boldsymbol{\partial}^{=k} F\rangle_{\leq \ell})$ is the dimension of the space spanned by all degree ℓ polynomial combinations of k-th order partial derivatives of F. [GKKS13a] gives an upper bound (Lemma 4) on the dimension of shifted partials of $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuits. We use the same upper bound so that the source of our improvement is the construction of a polynomial F whose shifted partials dimension is larger.

Estimating the dimension of shifted partials. [GKKS13a] gave an estimate of the shifted partial dimension of the determinant polynomial and conjectured that for suitable choice of k and ℓ, this dimension is much larger for the permanent. However, estimating $\dim(\langle \boldsymbol{\partial}^{=k} F \rangle_{\leq \ell})$ is a nontrivial task and it may well be that computing $\dim(\langle \boldsymbol{\partial}^{=k} F \rangle_{\leq \ell})$ cannot in general be done efficiently, even when the polynomial F is given very verbosely as a dense list of coefficients. In particular, we do not have a way to obtain a good estimate for the dimension of shifted partials of the permanent polynomial. To overcome this difficulty we construct an explicit F in a way that enables us to estimate $\dim(\langle \boldsymbol{\partial}^{=k} F \rangle_{\leq \ell})$. The idea of this construction is to design a homogeneous multilinear polynomial $F(\mathbf{x})$ of degree n with zero-one coefficients in such a way that any k-th order derivative of F is just a monomial (that could possibly be zero). This is achieved by interpreting the non-zero monomials of F as a collection of subsets of $[n^2]$ of size n each. We observe that if any two distinct subsets in our collection intersect in less than k points then any k-th order derivative is a single monomial (possibly zero). Such set-systems with small pairwise intersection are well-studied and are known as Nisan-Wigderson designs [NW94]. We use an explicit construction of such a set-system based on low-degree univariate polynomials to then reduce our problem to estimating the probability that a random collection of points in the plane has a large intersection with a low-degree curve. We do this via an application of Chebychev's inequality and using some elementary properties of univariate polynomials over finite fields. This leads to a proof of Theorem 2.

From depth four lower bounds to regular formula lower bounds. In the way indicated above we obtain for each $t = t(n)$ a family of polynomials $\mathcal{F}_t$ that requires $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuit of top fan-in at least $\exp\left(\Omega\left(\left(\frac{n}{t}\right)\log n\right)\right)$. Using a small interpolation trick (Lemma 14), we can combine the $\mathcal{F}_t$'s into a single family $\mathcal{F} = \{F_n(\mathbf{x})\}$ which requires $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuits of top fan-in at least $\exp\left(\Omega\left(\left(\frac{n}{t}\right)\log n\right)\right)$ *for all choices of* t[6]. We then argue that this combined polynomial $F_n(\mathbf{x})$ requires large regular formulas. The proof of this proceeds by starting with a regular formula Φ for $F_n(\mathbf{x})$, and obtaining $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuits computing $F_n(\mathbf{x})$ for several t's. Then, by a suitable *amortized analysis* (Theorem 15) over the $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ lower bounds for various t's, we obtain a super-polynomial lower bound for the size of the formula Φ.

2.1 Organisation of the paper

Section 3 shall describe the notations used in this paper and present some basic preliminaries. In Section 4 and describe the polynomial for which we shall prove $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ lower bounds. Section 5 shall be completely devoted to lower bounding the dimension of shifted partial derivatives of the polynomial constructed. Section 6 shall obtain the regular formula lower bound from the $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ lower bounds, and Section 7 concludes with some discussion and future directions.

[6] Actually, for a certain technical reason, the $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ lower bound is valid for a large range of values of t. This suffices for our purpose.

3 Notation and Preliminaries

Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a multivariate polynomial. We would use $\boldsymbol{\partial}^{=k} f$ to denote the set of all k-th order partial derivatives of f, i.e.

$$\left(\boldsymbol{\partial}^{=k} f\right) \stackrel{\text{def}}{=} \left\{ \frac{\partial^{|\mathbf{j}|} f}{\partial \mathbf{x}^{\mathbf{j}}} \ : \ |\mathbf{j}| = k \right\}.$$

If $S \subseteq \mathbb{F}[\mathbf{x}]$, then,

$$\langle S \rangle_{\leq \ell} \stackrel{\text{def}}{=} \left\{ \mathbf{x}^{\mathbf{i}} \cdot f \ : \ f \in S \text{ and } |\mathbf{i}| \leq \ell \right\}$$

In particular, $\langle \boldsymbol{\partial}^{=k}(f) \rangle_{\leq \ell}$ will denote the set of polynomials obtained by taking a k-th order partial derivative of f and multiplying by a monomial of degree at most ℓ.

Throughout this paper we shall use $\exp(x)$ to denote e^x (where e is the base of the natural logarithm). Also, all logarithms unless stated otherwise would refer to the natural logarithm.

3.1 Preliminaries

Estimates for ratios of factorials. The following useful estimate on ratios of factorials follows easily from Stirling's estimate.

Lemma 3. (cf. [GKKS13a]) *Let $a(n)$, $f(n)$, $g(n) : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be integer valued function such that $(f+g) = o(a)$. Then,*

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

Dimension of shifted partial derivatives. The complexity measure in this lower bound would be the *dimension of shifted partials* used in [GKKS13a].

$$\dim(\langle \boldsymbol{\partial}^{=k} F \rangle_{\leq \ell}) \quad \stackrel{\text{def}}{=} \quad \dim \left\{ \mathbf{x}^{\mathbf{i}} \cdot \frac{\partial^{|\mathbf{j}|} F}{\partial \mathbf{x}^{\mathbf{j}}} \ : \ |\mathbf{j}| = k, \ |\mathbf{i}| \leq \ell \right\}$$

We will use the following upper bound from [GKKS13a] on the dimension of shifted partial derivatives of a $\Sigma\Pi^{[m]}\Sigma\Pi^{[t]}$-circuit.

Lemma 4. Upper Bounding the Dimension of Shifted Partials of a Depth-4 circuits, [GKKS13a]. *Let $F(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be an N-variate polynomial that can be expressed as*

$$F(\mathbf{x}) = \sum_{i \in [s]} G_i(Q_{i1}, Q_{i2}, \ldots, Q_{im}), \quad \text{where } \deg(Q_{ij}) \leq t \quad \text{for all } i \in [s], j \in [m], \tag{3}$$

then

$$\dim(\langle \boldsymbol{\partial}^{=k} F \rangle_{\leq \ell}) \leq \sum_{i \in [s]} \dim(\langle \boldsymbol{\partial}^{=k} G_i \rangle_{\leq 0}) \cdot \binom{N + (t-1)k + \ell}{N}.$$

In particular, when each $G_i(u_1, u_2, \ldots, u_m)$ equals $(u_1 \cdot u_2 \cdot \ldots \cdot u_m)$ so that F can be computed by a $\Sigma\Pi^{[m]}\Sigma\Pi^{[t]}$-circuit of top fanin s then

$$\dim(\langle \boldsymbol{\partial}^{=k} F \rangle_{\leq \ell}) \leq s \cdot \binom{m}{k} \cdot \binom{N + (t-1)k + \ell}{N}.$$

3.2 ABPs to regular formulas

In this section we shall show that any algebraic branching program can be converted to a regular formula of quasi-polynomial size.

Proposition 5. *Let f be an N-variate degree d polynomial computed by an algebraic branching program of size s. Then, f can be computed by a regular formula of size $s^{O(\log d)}$.*

Proof. Since any branching program can be homogenized with just a polynomial blow up in size [Nis91], we can assume without loss of generality that the given branching program is homogeneous.

Let $Y^{(1)}, \ldots, Y^{(d)}$ be $s \times s$ matrices of disjoint set of indeterminates, that is $Y^{(i)} = ((y_{jk}^{(i)}))$. Then, the polynomial corresponding to the iterated matrix mutliplication can be defined as

$$\mathrm{IMM}_{s,d}(y_{11}^{(1)}, \ldots, y_{ss}^{(1)}, \ldots, y_{11}^{(d)}, \ldots, y_{ss}^{(d)}) \quad \stackrel{\text{def}}{=} \quad \left(Y^{(1)} \cdot \cdots \cdot Y^{(d)}\right)_{1,1}$$

Since f is computed by a homogeneous branching program of size s, it follows that f can be obtained by an appropriate projection of $\mathrm{IMM}_{s,d}(\mathbf{y})$. Hence, it suffices to show that $\mathrm{IMM}_{s,d}(\mathbf{y})$ can be computed by a regular formula of quasi-polynomial size.
Consider the polynomial $\mathrm{IMM}_{s,d'}$ where d' is the smallest power of two greater than d. Then by a straightforward divide-and-conquer approach, we can construct a circuit C of size $\mathrm{poly}(s, d)$ and depth $2 \log d'$ with alternating layer of addition and multiplication , where each multiplication gate has fan-in 2, and each addition gate has fan-in s. By unfolding this into a formula, we indeed get a formula of size $s^{O(\log d)}$ where the fan-in of all gates in a layer is the same. Furthermore, since the formal degree is at most $d' \leq 2d$, the resulting formula is indeed regular. Since affine projections of regular formulas stay regular, we have that f can be computed by a regular formula of size $s^{O(\log d)}$. □

4 Construction of a VNP-family

Let $\mathbb{F}$ be a field and $t = t(n)$ be an integer. In this section we present the construction of a family $\mathcal{F}_t$ of n^2-variate polynomials $F_{t,n}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ of degree n in VNP so that the top fanin of any $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuit computing $F_{t,n}(\mathbf{x})$ is very large. Given the upper bound on the dimension of shifted partials of $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuit (Lemma 4), our goal is to make sure that the dimension of shifted partial derivatives of $F_{t,n}(\mathbf{x})$, that is $\dim(\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell})$, is quite large for certain well-chosen values of the parameters k and ℓ. Indeed, we set our parameters k and ℓ as follows. We choose $k = \lfloor \frac{n}{2t} \rfloor$ and $\ell = \lceil \frac{cn^2 t}{\log n} \rceil$ for some sufficiently large constant $c > 0$ (to be specified later).

The Construction. Let n be a power of a prime and let $\mathbb{F}_n$ be the finite field of size n. We will think of the set of the first n integers as elements of $\mathbb{F}_n$ via an arbitrary correspondence $\phi : [n] \mapsto \mathbb{F}_n$. Let $a(z) \in \mathbb{F}_n[z]$ be a univariate polynomial. Abusing notation, $a(i)$ will denote the evaluation of a at the i-th field element via the above correspondence, i.e. $a(i) \stackrel{\text{def}}{=} \phi^{-1}(a(\phi(i)))$. The n-th member $F_{t,n}(\mathbf{x})$ of our family $\mathcal{F}_t$ will be a homogeneous multilinear polynomial of degree n on

the n^2 variables $x_{11}, x_{12}, \ldots, x_{nn}$ and is defined as follows.

$$F_{t,n}(x_{11}, \ldots, x_{nn}) \quad = \quad \sum_{\substack{a(z) \in \mathbb{F}_n[z] \\ \deg a(z) < k}} x_{1,a(1)} \cdot x_{2,a(2)} \cdot \ldots \cdot x_{n,a(n)}. \tag{4}$$

To simplify the exposition, we will omit the correspondence ϕ and will often identify a variable $x_{i,j}$ by the point $(\phi(i), \phi(j)) \in \mathbb{F}_n \times \mathbb{F}_n$ and a monomial $\left(x_{1,a(1)} \cdot x_{2,a(2)} \cdot \ldots \cdot x_{n,a(n)}\right)$ of $F_{t,n}(\mathbf{x})$ with the corresponding univariate polynomial $a(z) \in \mathbb{F}_n[z]$ or the corresponding subset $S_a \stackrel{\text{def}}{=} \{(1, a(1)), \ldots, (n, a(n))\} \subseteq \mathbb{F}_n^2$. We denote by $\mathcal{S}_k$ the set of all S_a's in which the polynomial $a(z) \in \mathbb{F}_n[z]$ has degree less than k, i.e.

$$\mathcal{S}_k \stackrel{\text{def}}{=} \left\{ S_a \subseteq \mathbb{F}_n^2 \ : \ a(z) \in \mathbb{F}_n[z] \quad \text{and } \deg(a) < k \right\}.$$

Upper bounds for the family $\mathcal{F}_t$. Our construction is clearly very explicit so that the family $\mathcal{F}_t$ is immediately seen to be in VNP.

Proposition 6. *For any $t = t(n) \leq n$, the family $\mathcal{F}_t$ is in the class* VNP. *Moreover, $F_{t,n}(\mathbf{x})$ can be computed by $\Sigma\Pi$ circuits of size $2^{O((n/t)\cdot \log n)}$ (and therefore also by $\Sigma\Pi^{[(n/t)]}\Sigma\Pi^{[t]}$-circuits) of size $2^{O((n/t)\cdot \log n)}$.*

Proof. Let $F_{t,n}(\mathbf{x})$ in $\mathcal{F}_t$ be the polynomial of degree n as defined in equation (4). It is trivial to construct a polynomial time algorithm that given a monomial $\mathbf{x}^{\mathbf{e}}$ checks if the coefficient of this monomial in $F_{t,n}(\mathbf{x})$ is zero or one. Hence, by Valiant's criterion [Val79, Proposition 4], $\mathcal{F}_t$ is in the class VNP. Moreover, by definition, the number of monomials in $F_{t,n}(\mathbf{x})$ equals the number of polynomials $a(z) \in \mathbb{F}_n[z]$ of degree $(k-1)$. Thus $F_{t,n}(\mathbf{x})$ has $n^k = 2^{(\lfloor n/2t \rfloor)\cdot \log n}$ monomials and hence can be computed by $\Sigma\Pi$-circuits of size $2^{O((n/t)\cdot \log n)}$. □

Estimating Shifted Partials via monomial counting. Now notice that since any two distinct univariate polynomials $a(z) \neq b(z) \in \mathbb{F}_n[z]$ of degree less than k each can agree on at most $(k-1)$ points therefore the size of the intersection of the corresponding image sets, $|S_a \cap S_b|$, is less than k. This in turn means that any two monomials in $F_{t,n}(\mathbf{x})$ have at most $(k-1)$ variables in common so that every k-th order partial derivative of $F_{t,n}(\mathbf{x})$ is either zero or just a monomial. We thus have:

Proposition 7. *Let $t = t(n)$ be an integer and let $F_{t,n}(\mathbf{x}) \in \mathcal{F}_t$ be the degree n polynomial defined in equation (4). Let $k = \lfloor \frac{n}{2t} \rfloor$. Then the set of k-th order partial derivatives of $F_{t,n}(\mathbf{x})$ are monomials of the form*

$$\left\{ \prod_{(i,j) \in S} x_{ij} \quad : \quad |S| = (n-k) \ \text{ and } \exists S_a \in \mathcal{S}_k \ \text{ such that } \ S \subset S_a \right\}.$$

This characterization of the polynomials in $(\boldsymbol{\partial}^{=k} F_{t,n})$ also allows us to characterize the polynomials in $\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}$. First note that since $(\boldsymbol{\partial}^{=k} F_{t,n})$ consists only of monomials therefore so does $\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}$. Note that every monomial $m \in \langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}$ of support $S \subseteq \mathbb{F}_n \times \mathbb{F}_n$ has degree at most $(\ell + n - k)$ and is divisible by some monomial in $(\boldsymbol{\partial}^{=k} F_{t,n})$. Hence, S contains a subset S'_a of

size $(n-k)$ of some $S_a \in \mathcal{S}_k$. Conversely, suppose that S contains a subset $S'_a \subseteq S_a$ of size $(n-k)$. Then, m can also be obtained by differentiating $F_{t,n}(\mathbf{x})$ with respect to the variables $S_a \setminus S'_a$ to obtain $\prod_{(i,j)\in S'_a} x_{ij}$, and then shifted appropriately to get m. Hence we have:

Proposition 8. *The set $\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell}$ consists precisely of monomials of degree at most $\ell+n-k$ that contain at least $(n-k)$ elements of some $S_a \in \mathcal{S}_k$.*

This then enables us to reduce the task of estimating $\dim(\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell})$ (which in turn is equivalent to the seemingly formidable task of estimating the rank of an exponentially-sized matrix) to the combinatorial problem of counting the number of distinct monomials satisfying certain properties. The main technical lemma in this work shows that for the above choice of parameters k and ℓ, the number of monomials in $\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell}$ (and hence also $\dim(\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell})$) is nearly equal (upto $\mathrm{poly}(n)$ factors) to the number of monomials of degree $(\ell+n-k)$. Specifically, we have:

Lemma 9. Main Technical Lemma - counting the number of distinct shifted partials. *Suppose that $\log^2 n < t < \frac{n}{100}$ and $c > 5$. Then*

$$\dim(\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell}) \quad \geq \quad \left(\frac{1-o(1)}{n^2}\right) \cdot \binom{n^2+\ell+n-k}{n^2}$$

We defer the proof of this lemma to the next section. We first see why this leads to lower bounds for $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuits.

Using the technical lemma to obtain depth four lower bounds . We can now put together the upper bound of Lemma 4 and the lower bound given by Lemma 9 to obtain lower bounds for depth four circuits as in Theorem 2.

Theorem 2 (restated). Let $t : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ be an increasing function. Let $\mathbb{F}$ be any field. Let $\alpha \in \mathbb{Z}_{\geq 1}$ be any fixed positive integer. For every t, the family $\mathcal{F}_t$ as defined above is in VNP. Moreover, for any $t = t(n)$ satisfying $\log^2 n < t < \frac{n}{100}$, the polynomial $F_{t,n}(\mathbf{x}) \in \mathcal{F}_t$ is an n^2-variate polynomial of degree n over $\mathbb{F}$ such that any $\Sigma\Pi^{[\alpha\cdot(n/t)]}\Sigma\Pi^{[t]}$-circuit computing it must have top fan-in at least $2^{((n/50t)\cdot\log n)}$ for all large enough n. In particular, there exists an explicit n^2-variate polynomial of degree n such that any $\Sigma\Pi^{[\alpha\sqrt{n}]}\Sigma\Pi^{[\sqrt{n}]}$-circuit computing it must have top fan-in at least $2^{\Omega(\sqrt{n}\log n)}$.

Proof. Since $t < \frac{n}{100}$, we have that $k = \lfloor \frac{n}{2t} \rfloor > 1$. Consider $\dim(\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell})$ for $\ell = \lceil \frac{5n^2 t}{\log n} \rceil \leq \frac{6n^2 t}{\log n}$. By our main technical lemma (Lemma 9) we have:

$$\dim(\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell}) \quad \geq \quad \left(\frac{1-o(1)}{n^2}\right) \cdot \binom{n^2+\ell+n-k}{n^2}$$

Further if $F_{t,n}(\mathbf{x})$ is computed by a $\Sigma\Pi^{[\alpha\cdot(n/t)]}\Sigma\Pi^{[t]}$ circuit of top fan-in s, then by Lemma 4 we have

$$\dim(\langle \partial^{=k} F_{t,n} \rangle_{\leq \ell}) \quad \leq \quad s \cdot \binom{\alpha\cdot(n/t)}{k}\binom{n^2+\ell+k(t-1)}{n^2}$$

Hence

$$\begin{aligned}
s &\geq \frac{(1-o(1))\cdot\binom{n^2+\ell+n-k}{n^2}}{n^2\cdot\binom{\alpha\cdot(n/t)}{k}\binom{n^2+\ell+k(t-1)}{n^2}} \\
&= \frac{\binom{n^2+\ell+n-k}{n^2}}{\binom{n^2+\ell+k(t-1)}{n^2}}\cdot\exp\left(-O\left(\frac{n}{t}\right)\right)\cdot\exp\left(-2\log n\right) \\
&\geq \left(1+\frac{n^2}{\ell}\right)^{(n-tk)}\cdot\exp\left(-O\left(\frac{n}{t}\right)\right)\cdot\exp\left(-2\log n\right) \quad \text{(Using Lemma 3)} \\
&\geq \exp\left(\frac{n^2(n-tk)}{2\ell}\right)\cdot\exp\left(-O\left(\frac{n}{t}\right)\right)\cdot\exp\left(-2\log n\right) \quad \text{(Using } 1+x>e^{x/2} \text{ for } 0\leq x\leq 1\text{)} \\
&\geq \exp\left(\left(\frac{n}{24t}\right)\log n-2\log n\right)\cdot\exp\left(-O\left(\frac{n}{t}\right)\right) \\
&> \exp\left(\left(\frac{n}{50t}\right)\log n\right) \quad \text{for any } \log^2 n < t < \frac{n}{100} \text{ and sufficiently large } n
\end{aligned}$$

□

5 Proof of the main technical lemma

In this section we use the characterization of $\langle \boldsymbol{\partial}^{=k} F_{t,n}\rangle_{\leq \ell}$ given in Proposition 8 to obtain a proof of Lemma 9 and thereby obtain a lower bound on $\dim(\langle \boldsymbol{\partial}^{=k} F_{t,n}\rangle_{\leq \ell})$. Throughout the rest of this section, the choice of the integer parameters k and ℓ will be as in Section 4, namely

$$k = \left\lfloor \frac{n}{2t} \right\rfloor \quad \text{and} \quad \ell = \left\lceil \frac{5n^2 t}{\log n} \right\rceil.$$

From Proposition 8 we have that $\langle \boldsymbol{\partial}^{=k} F_{t,n}\rangle_{\leq \ell}$ consists precisely of monomials m of degree at most $\ell+n-k$ that contain at least $(n-k)$ elements of some $S_a \in \mathcal{S}_k$. Since this condition is completely determined by the *support* of the monomial m, it would be useful to have an estimate of the support size of a random monomial of degree at most $\ell+n-k$.

Lemma 10. Most likely support size for monomials of a given degree. *At least a $\frac{1}{n^2}$ fraction of all monomials of degree at most $\ell' = (\ell+n-k)$ over n^2 variables has support size $N_{\max} \stackrel{\text{def}}{=} \left\lfloor \frac{n^2\ell'-1}{n^2+\ell'+2} \right\rfloor + 1$.*

Proof. The number of monomials of degree at most ℓ' of support size s is precisely

$$N_s \stackrel{\text{def}}{=} \binom{n^2}{s}\cdot\binom{s+(\ell'-s)}{s} = \binom{n^2}{s}\cdot\binom{\ell'}{s}$$

To find out where N_s is maximised, consider the ratio of N_{s+1} and N_s

$$\frac{N_{s+1}}{N_s} = \frac{\binom{n^2}{s+1}\cdot\binom{\ell'}{s+1}}{\binom{n^2}{s}\cdot\binom{\ell'}{s}} = \frac{(n^2-s)(\ell'-s)}{(s+1)^2}$$

Hence, $N_{s+1} > N_s$ if and only if $(n^2-s)(\ell'-s) > (s+1)^2$, which happens if and only if $s < \frac{n^2\ell'-1}{n^2+\ell'+2}$. Hence, N_s increases until $s = \left\lfloor \frac{n^2\ell'-1}{n^2+\ell'+2} \right\rfloor + 1$, and decreases from then on. Hence, N_s is maximised at $N_{\max} = \left\lfloor \frac{n^2\ell'-1}{n^2+\ell'+2} \right\rfloor + 1$. □

Hence, for $\ell = \lceil \frac{cn^2 t}{\log n} \rceil$, at least an $\frac{1}{n^2}$ fraction of all degree at most $(\ell + n - k)$ monomials over n^2 variables have support

$$N_{\max} \stackrel{\text{def}}{=} \left\lfloor \frac{n^2\ell' - 1}{n^2 + \ell' + 2} \right\rfloor + 1 = n^2 - O\left(\frac{n^2 \log n}{t}\right).$$

That is,

$$\binom{n^2}{N_{\max}} \cdot \binom{\ell + n - k}{N_{\max}} \quad \geq \quad \frac{1}{n^2} \cdot \binom{n^2 + \ell + n - k}{n^2}.$$

Reducing to a probability estimation. We will now focus our attention on monomials m of support exactly $N_{\max}$ and degree at most $(\ell + n - k)$. From the fact that the membership of m in $\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}$ is completely determined by its support and using the correspondence between supports of monomials and subsets of points in $\mathbb{F}_n^2$ (Section 4), we see that our problem reduces to estimating the fraction of subsets of points in $\mathbb{F}_n^2$ satisfying certain conditions.

Lemma 11. *Let $\ell' = \ell + n - k$ and let*

$$s^* = n^2 - N_{\max} = n^2 - \left\lfloor \frac{n^2\ell' - 1}{n^2 + \ell' + 2} \right\rfloor - 1$$

Let p be the probability that a random set $T \subset \mathbb{F}_n^2$ of size s^* intersects one of the S_a's in $\mathcal{S}_k$ in at most k places. Then

$$\dim(\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}) \quad \geq \quad \frac{1}{n^2} \cdot p \cdot \binom{n^2 + \ell + n - k}{n^2}$$

Proof. Let p be the fraction of monomials of degree at most $\ell' = (\ell + n - k)$ over n^2 variables *having support size exactly* $N_{\max}$ which are in $\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}$. Then from Lemma 10 we have

$$\dim(\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}) \quad \geq \quad \frac{1}{n^2} \cdot p \cdot \binom{n^2 + \ell + n - k}{n^2}.$$

Now let m be a monomial of degree at most ℓ' and support size $N_{\max}$. We interpret the support of m as a subset S of $\mathbb{F}_n^2$ via the correspondence in Section 4. Then by Proposition 8 we have that m is in $\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}$ if and only if there exists an $S_a \in \mathcal{S}_k$ such that $|S \cap S_a| \geq (n - k)$. Let $T = \mathbb{F}_n^2 \setminus S$. Since every $S_a \in \mathcal{S}_k$ is of size exactly n we see that m is in $\langle \boldsymbol{\partial}^{=k} F_{t,n} \rangle_{\leq \ell}$ if and only $|T \cap S_a| \leq k$. Finally since the number of monomials having support S depends only on the size of S we have

$$p \quad = \quad \Pr_{\substack{T \subset \mathbb{F}_n^2 \\ |T| = s^*}} \left[|T \cap S_a| \leq k \text{ for some } S_a \in \mathcal{S}_k \right],$$

as required. □

In this way, our task reduces to estimating p, the probability that a subset of $\mathbb{F}_n^2$ of size s^* intersects the image of a degree k polynomial in at most k places. The rest of this section is devoted to showing that for our choice of parameters k and ℓ, this probability p is close to one.

Picking points independently at random versus picking a random subset of points. We wish to lower bound p, the probability that a random set $T \subset \mathbb{F}_n^2$ of size s^* satisfies the above property. Suppose that instead of picking a random subset T of a certain size, we sample each point in $\mathbb{F}_n^2$ independently at random with probability $\approx \frac{s^*}{n^2}$ then the resulting set of sampled points will be a random set of size $\approx s^*$ (with high probability) and therefore ought to satisfy the above property with roughly the same probability. The following lemma makes this precise and shows that we do not lose much with this change. First note that $s^* = n^2 - N_{\max}$ is $o(n^2)$ if $t \geq \log^2 n$.

Lemma 12. *Let $\mathcal{D}_1$ be the uniform distribution on all sets $T \subset \mathbb{F}_n^2$ of size $s^* = o(n^2)$, and let $\mathcal{D}_2$ be the distribution on sets $T \subseteq \mathbb{F}_n^2$ obtained by sampling each point $(\alpha, \beta) \in \mathbb{F}_n^2$ independently at random with probability $\frac{11}{10} \cdot \frac{s^*}{n^2}$. Let p_1 and p_2 be the probability that the resulting random set T intersects one of the $S_a \in \mathcal{S}_k$ in at most k places, when T is drawn according to $\mathcal{D}_1$ and $\mathcal{D}_2$ respectively. Then,*

$$p_1 \quad \geq \quad p_2 - \exp\left(-O\left(s^*\right)\right)$$

Proof. The expected size of a random set T sampled according to $\mathcal{D}_2$ is $\frac{11}{10} \cdot s^*$. By the Chernoff bound, a set T randomly chosen according to $\mathcal{D}_2$ has size at least s^* with probability at least $1 - \exp\left(-O\left(s^*\right)\right)$. Hence,

$$\begin{aligned}
p_2 &\leq \Pr_{T \in \mathcal{D}_2}\left[\exists S_a \in \mathcal{S}_k \ : \ |T \cap S_a| \leq k \text{ and } |T| \geq s^*\right] + \exp\left(-O\left(s^*\right)\right) \\
&= \left(\sum_{r=s^*}^{n^2} \Pr_{T \in \mathcal{D}_2}\left[\exists S_a \in \mathcal{S}_k \ : \ |T \cap S_a| \leq k \quad \middle| \quad |T| = r\right] \cdot \Pr_{T \in \mathcal{D}_2}\left[|T| = r\right]\right) + \exp\left(-O\left(s^*\right)\right) \\
&\leq \Pr_{T \in \mathcal{D}_2}\left[\exists S_a \in \mathcal{S}_k \ : \ |T \cap S_a| \leq k \quad \middle| \quad |T| = s^*\right] + \exp\left(-O\left(s^*\right)\right)
\end{aligned}$$

The last inequality follows because increasing the size of the set T only decreases the probability of intersecting one of the S_a's in at most k places. Specifically, consider picking a random set T of size s^* by first picking a set T' of size $r \geq s^*$, and choosing the first s^* elements as T. Obviously, if T' intersects some S_a in at most k places, then so does T. Since the first term of the last equation is exactly p_1, we have

$$p_2 \quad \leq \quad p_1 \ + \ \exp\left(-O\left(s^*\right)\right)$$

□

In our setting, $\ell = \lceil \frac{cn^2 t}{\log n} \rceil$ and hence $s^* \leq \frac{n^2 \log n}{ct}$ for all large enough n, where c is a constant independent of n. Hence, it suffices to prove a good enough lower bound for p_2 when each element is sampled with probability $q = \frac{\log n}{c't}$ for a constant $c' = \frac{10c}{11}$. Increasing the sampling probability q can only diminish the chance that a set T intersects one of the $S_a \in \mathcal{S}_k$ in at most k places. Hence a lower bound for p_2 with sampling probability $q = \frac{\log n}{c't}$ implies a lower bound for p_2 when sampling probability is $\frac{11 s^*}{10 n^2}$ as in Lemma 12.

Estimating p_2 via Chebychev. So we now sample a set $T \subseteq \mathbb{F}_n^2$ by sampling each point in $\mathbb{F}_n^2$ independently at random with probability q as above and we wish to show that with high probability $|T \cap S_a| \leq k$ for some $S_a \in \mathcal{S}_k$. We proceed as follows. For each polynomial $a(z) \in \mathbb{F}_n[z]$ of degree k, let X_a be the indicator random variable which is one if $|T \cap S_a| \leq k$, and zero otherwise.

Define $Y = \sum X_a$. p_2 is then simply the probability that Y is larger than zero. We first compute the expectation μ of Y (using linearity of expectation) and see that this is quite large. Thereafter we estimate the variance of Y, $\mathrm{Var}(Y)$ and use Chebychev to deduce that $1 - p_2 = \Pr[Y = 0]$ is quite small.

Lemma 13. *Assume $\log^2 n < t < \frac{n}{100}$, and $k = \lfloor \frac{\varepsilon n}{t} \rfloor$ is a non-zero integer for a fixed constant $0 < \varepsilon < 1$. Let T be a random set obtained by selecting each element of $\mathbb{F}_n^2$ independently with probability $q = \frac{\log n}{c't}$, where c' is a sufficiently large constant ($c' > \frac{2}{\varepsilon}$ would suffice). Then we have*

$$\Pr[|T \cap S_a| \leq k \text{ for some } S_a \in \mathcal{S}_k] \quad \geq \quad 1 - o(1)$$

Proof. Let p_m be the probability that such a random set T intersects a fixed set S of size m in at most k places.

$$p_m \quad = \quad \sum_{i=0}^{k} \binom{m}{i} \left(\frac{\log n}{c't}\right)^i \left(1 - \frac{\log n}{c't}\right)^{m-i}$$

For every polynomial $a(z)$, let X_a be the indicator random variable which is one if $|T \cap S_a| \leq k$, and zero otherwise. Define $Y = \sum X_a$. Then, it follows that $\mu \stackrel{\text{def}}{=} \mathbb{E}[Y] = n^k p_n$.

$$\begin{aligned}
\mu \quad = \quad n^k p_n \quad &\geq \quad n^k \cdot \binom{n}{k} \left(\frac{\log n}{c't}\right)^k \left(1 - \frac{\log n}{c't}\right)^{n-k} \\
&\geq \quad n^k \cdot \left(\frac{n}{\varepsilon(n/t)}\right)^k \left(\frac{\log n}{c't}\right)^k \exp\left(\frac{-2n \log n}{c't}\right) \quad \left(\because e^{-x} \leq 1 - \frac{x}{2} \text{ for } 0 \leq x \leq 1\right) \\
&\geq \quad \left(\frac{n^{\varepsilon(n/t)}}{n^{(2n/c't)}}\right) \cdot \left(\frac{1}{n}\right) \cdot \left(\frac{\log n}{\varepsilon c'}\right)^k
\end{aligned}$$

Hence, if c' is large enough ($c' > \frac{2}{\varepsilon}$), μ is exponential in $\frac{n \log n}{t}$.

By the Chebychev inequality,

$$\Pr[Y = 0] \quad \leq \quad \Pr[|Y - \mu| > 0.9\mu] \quad \leq \quad \frac{2\mathrm{Var}(Y)}{\mu^2}$$

Thus, it suffices to prove a good upper-bound for $\mathrm{Var}(Y)$.

$$\begin{aligned}
\mathrm{Var}(Y) \quad &= \quad \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\
&= \quad \sum_a \left(\mathbb{E}[X_a^2] - (\mathbb{E}[X_a])^2\right) \quad + \quad \sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - \mathbb{E}[X_a]\mathbb{E}[X_b]\right) \\
&= \quad \sum_a \left(p_n - p_n^2\right) \quad + \quad \sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - p_n^2\right) \\
&= \quad \mu(1 - p_n) \quad + \quad \sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - p_n^2\right) \\
\implies \Pr[Y = 0] \quad &\leq \quad \frac{2\mu(1 - p_n)}{\mu^2} \quad + \quad \left(\frac{2\sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - p_n^2\right)}{\mu^2}\right)
\end{aligned}$$

The first term tends to zero (as $\mu \to \infty$ for large n) and hence it suffices to bound the second term.

Let us split the summation in the second term into groups according to the intersection sizes of S_a and S_b.

$$\sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - p_n^2 \right) \quad = \quad \sum_{r=0}^{k-1} \sum_{\substack{a,b \\ |S_a \cap S_b| = r}} \left(\mathbb{E}[X_a X_b] - p_n^2 \right)$$

If S_a and S_b intersect in exactly r places, then there must be distinct $\alpha_1, \ldots, \alpha_r \in \mathbb{F}_n$ such that $a(z) = b(z) + (z - \alpha_1) \ldots (z - \alpha_r) g(z)$ for some $g(z)$ of degree less than $k - r$. Hence, the number of pairs a, b such that S_a and S_b intersect in r places is at most $n^k \binom{n}{r} n^{k-r} \leq \frac{n^{2k}}{r!}$.
Now, if a random set T intersects S_a and S_b in at most k places, then certainly T intersects $S_a \setminus S_b$ and $S_b \setminus S_a$ in at most k places (and both these sets are disjoint, and of size $n - r$). Hence,

$$\begin{aligned} \mathbb{E}[X_a X_b] \quad = \quad & \Pr[X_a = 1 \text{ and } X_b = 1] \quad \leq \quad p_{n-r}^2 \\ \implies \sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - p_n^2 \right) \quad & \leq \quad \sum_{r=0}^{k-1} \left(\frac{n^{2k}}{r!} \right) \cdot \left(p_{n-r}^2 - p_n^2 \right) \end{aligned}$$

In order to understand how different p_{n-r} is from p_n, let us think of a fixed set S of size n and an $S' \subset S$ of size $(n - r)$. Then, p_n is the probability that a random T intersects S in at most k places, and p_{n-r} is the probability that T intersects S' in at most k places. Now notice that any set T that intersects S' in at most k places that *does not* contain any element of $S \setminus S'$ would obviously intersect with S in at most k places. Hence,

$$\begin{aligned} \Pr\left[|T \cap S| \leq k \right] \quad & \geq \quad \Pr\left[\left| T \cap S' \right| \leq k \right] \cdot \Pr\left[T \cap (S \setminus S') = \varnothing \right] \\ & \geq \quad \Pr\left[\left| T \cap S' \right| \leq k \right] \cdot \left(1 - \frac{\log n}{c' t} \right)^r \\ \text{That is,} \quad p_n \quad & \geq \quad p_{n-r} \cdot \left(1 - \frac{\log n}{c' t} \right)^r \\ \implies p_{n-r} \quad & \leq \quad p_n \cdot \left(1 + O\left(\frac{\log n}{t} \right) \right)^r \\ \implies p_{n-r}^2 \quad & \leq \quad p_n^2 \cdot \left(1 + O\left(\frac{\log n}{t} \right) \right)^r \end{aligned}$$

Hence,

$$\sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - p_n^2 \right) \quad \leq \quad \sum_{r=0}^{k} \left(n^{2k} p_n^2 \right) \cdot \frac{\left[\left(1 + O\left(\frac{\log n}{t} \right) \right)^r - 1 \right]}{r!} \tag{5}$$

If $t = \Omega(n)$, then $k = \lfloor \frac{\varepsilon n}{t} \rfloor = O(1)$. In that case,

$$\begin{aligned} \sum_{r=0}^{k} \left(n^{2k} p_n^2 \right) \cdot \frac{\left[\left(1 + O\left(\frac{\log n}{t} \right) \right)^r - 1 \right]}{r!} \quad & \leq \quad \mu^2 \cdot k \cdot \left[\left(1 + O\left(\frac{\log n}{t} \right) \right)^k - 1 \right] \\ & \leq \quad \mu^2 \cdot k \cdot O\left(\frac{k \log n}{t} \right) \quad = \quad \mu^2 \cdot o(1) \end{aligned}$$

On the other hand, if $t = o(n)$ then $k = \omega(1)$ so that we get from (5),

$$\begin{aligned}\sum_{a \neq b} \left(\mathbb{E}[X_a X_b] - p_n^2\right) &\leq \mu^2 \cdot \left(\sum_{r=0}^{\infty} \frac{\left(1 + O\left(\frac{\log n}{t}\right)\right)^r}{r!} - \sum_{r=0}^{k} \frac{1}{r!}\right) \\ &\leq \mu^2 \cdot \left(e^{1+O\left(\frac{\log n}{t}\right)} - e + O\left(\frac{1}{k!}\right)\right) \\ &= \mu^2 \cdot o(1) \quad \text{since } t > \log^2 n \text{ and } k = \omega(1).\end{aligned}$$

Therefore,

$$\begin{array}{rcccccccc} \mathrm{Var}(Y) & \leq & \mu(1-p_n) & + & \mu^2 \cdot o(1) & & & & \\ \Pr[Y=0] & \leq & \dfrac{2\mathrm{Var}(Y)}{\mu^2} & \leq & \dfrac{2}{\mu} & + & o(1) & = & o(1) \end{array}$$

which tends to zero as n tends to infinity. Therefore, p_2 the probability T would intersect some S_a in at most k places is quite high $(1 - o(1))$. □

Lemma 9 now follows immediately from the last two lemmas.

6 Lower bound for regular formulas

In this section we will show how the lower bound for $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuits leads to lower bounds for regular formulas introduced in Section 1. From Section 5, for each $\log^2 n < t < \frac{n}{100}$, we have a construction of a family of polynomials $\mathcal{F}_t = \{F_{t,n}(\mathbf{x})\}$ which is $\exp((n/t) \cdot \log n)$-hard for the class of $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuits. We first show that these different families can be combined using a simple interpolation trick to get a *single* family of polynomials $\mathcal{F} = \{F_n\}$ which is $\exp((n/t) \cdot \log n)$-hard for $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuits *for all* $\log^2 n < t < \frac{n}{100}$. We will then prove a regular formula lower bound for this family by doing an ammortized analysis over the different ways of converting a regular formula into a $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$-circuit.

6.1 Combining multiple families into a single family.

Lemma 14. *Let $F_n(x_{11}, \ldots, x_{nn}, u)$ be the (n^2+1)-variate degree $2n$ polynomial defined as*

$$F_n(x_{11}, \ldots, x_{nn}, u) = \sum_{k=1}^{n} u^k \cdot F_{t,n}(x_{11}, \ldots, x_{nn})$$

Then, $\mathcal{F} = \{F_n : n \geq 1\} \in \mathsf{VNP}$ and for every $\log^2 n < t < \frac{n}{100}$, any $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuit computing F_n must have top fan-in at least $\exp\left(0.01\left(\frac{n}{t}\right)\log n\right)$ for all large enough n.

Proof. It is clear that $\mathcal{F}$ is in VNP since each $\mathcal{F}_t = \{F_{t,n}\}$ is in VNP. As for the lower bound, assume on the contrary that there is some $\log^2 n < t < \frac{n}{100}$ such that F_n can be computed by $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuits of top fan-in $\exp\left(0.01\left(\frac{n}{t}\right)\log n\right)$.

For every distinct[7] scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$, and for every $t \in [n]$ there exists scalars $\beta_{t,1}, \ldots \beta_{t,n}$ such that

$$F_{t,n}(x_{11}, \ldots, x_{nn}) \quad = \quad \sum_{i=1}^{n} \beta_{t,i} \cdot F_n(x_{11}, \ldots, x_{nn}, \alpha_i)$$

In particular we then have that $F_{t,n}$ can be computed by a $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuit C of top fan-in $n \cdot \exp\left(0.01\left(\frac{n}{t}\right)\log n\right)$, which is at most $\exp\left(0.02\left(\frac{n}{t}\right)\log n\right)$ for any $\log^2 n < t < \frac{n}{100}$. Since substituting $u = \alpha_i$ can only decrease the formal degree (which was $2n$ originally), the resulting circuit continues to stay a $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuit computing $F_{t,n}$. But such a statement contradicts the lower bound of Theorem 2. □

Remark. *This trick can be applied more generally in situations where we have families of polynomials $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_r$ where the i-th family $\mathcal{F}_i$ is $s_i(n)$-hard for a subclass of circuits $\mathcal{C}_i$. If every circuit subclass is subprojective (i.e. computing $f(\alpha, x_2, \ldots, x_n)$ is no harder than computing $f(x_1, x_2, \ldots, x_n)$) and sub-additive (i.e. complexity of $f(\mathbf{x}) + g(\mathbf{x})$ is at most the sum of the complexities of f and g), then we can construct in the above manner a single family $\mathcal{F}$ which is simultaneously $\frac{s_i(n)}{r}$-hard for each circuit class $\mathcal{C}_i$.*

6.2 From regular depth four circuits to regular formulas.

Theorem 15. *Let $F(x_1, \ldots, x_N)$ be a polynomial of degree d with the property that there exists a $\delta > 0$ such that for every $\log^2 d < t < \frac{d}{100}$, any $\Sigma\Pi^{[O(d/t)]}\Sigma\Pi^{[t]}$ circuit computing the polynomial F has top fan-in at least $\exp\left(\delta\left(\frac{d}{t}\right)\log N\right)$. Then, any regular formula computing F must be of size $N^{\Omega(\log d)}$.*

Proof. Let Φ be a $\Sigma^{[a_1]}\Pi^{[p_1]}\Sigma^{[a_2]}\Pi^{[p_2]} \ldots \Sigma^{[a_\Delta]}\Pi^{[p_\Delta]}\Sigma^{[a_{\Delta+1}]}$-regular formula computing F, whose formal degree $\prod p_i \leq 2d$.

For any i, let C_i be the depth-4 circuit obtained by converting the top $2i$ layers into a $\Sigma\Pi$ circuit, and the remaining layers by $\Sigma\Pi$ circuits. Since the formal degree of all polynomials at the $(2i+1)$-th layer is $t_i = p_{i+1} \ldots p_\Delta$, we have that C_i is a $\Sigma\Pi^{[O(d/t_i)]}\Sigma\Pi^{[t_i]}$-circuit. Further, the top fan-in of the circuit C_i is upper-bounded by $a_1 \cdot a_2^{p_1} \ldots a_i^{p_1 \ldots p_{i-1}}$. Hence if $\log^2 d < t_i < \frac{d}{100}$, applying the lower bound for $\Sigma\Pi^{[O(d/t_i)]}\Sigma\Pi^{[t_i]}$ computing F, we have

$$a_1 \cdot a_2^{p_1} \ldots a_i^{p_1 \ldots p_{i-1}} \quad \geq \quad \exp\left(\delta\left(\frac{d}{t_i}\right)\log N\right)$$

Using $t_i = p_{i+1} \ldots p_\Delta$ and $p_1 \ldots p_\Delta \leq 2d$, we get

$$a_1 \cdot a_2^{p_1} \ldots a_i^{p_1 \ldots p_{i-1}} \quad \geq \quad \exp\left[\left(\frac{p_1 \ldots p_i}{2}\right) \cdot (\delta \log N)\right] \qquad (\text{Eqn}_i)$$

Let ℓ be the smallest index[8] with $p_{\ell+1} \ldots p_\Delta < \frac{d}{100}$, and let h be the largest index with $p_{h+1} \ldots p_\Delta > \log^2 d$.

[7] If $\mathbb{F}$ is too small, then we can choose these scalars from a large enough extension field. Of course, any lower bound for the extension field would continue to hold over the smaller base field.

[8] As a convention, let $p_i \ldots p_j = 1$ if $i > j$. Hence, if say $p_\Delta > \frac{d}{100}$, we shall set $\ell = \Delta$.

Case 0: $h < \ell$

Note that since $p_\ell \dots p_\Delta \geq \frac{d}{100} > \log^2 d$, we have that $h \geq \ell - 1$. However $h = \ell - 1$ if and only if $p_\ell \geq \frac{d}{100 \log^2 d}$. In this case, consider the $\Sigma\Pi^{[O(d/t_\ell)]}\Sigma\Pi^{[t_\ell]}$ circuit C_ℓ computing F as defined above. The top fan-in of this circuit is $a_1 a_2^{p_1} \dots a_\ell^{p_1 \dots p_{\ell-1}}$. Since $t_\ell \leq \log^2 d$, we cannot directly apply Lemma 14. Nevertheless, by grouping low degree terms together (cf. [GKKS13a, Remark 11]), C_ℓ can be converted to a $\Sigma\Pi^{[O(d/\log^2 d)]}\Sigma\Pi^{[\log^2 d]}$ circuit without any blow-up in top fan-in. Hence, applying Lemma 14, we have

$$\begin{aligned}
(a_1 \dots a_\ell)^{p_1 \dots p_{\ell-1}} &\geq a_1 a_2^{p_1} \dots a_\ell^{p_1 \dots p_{\ell-1}} \geq \exp\left(\delta \cdot \frac{d}{\log^2 d} \log N\right) \\
\implies |\Phi| &\geq \exp\left(\delta \cdot \frac{d}{p_1 \dots p_{\ell-1} \cdot \log^2 d} \log N\right) \\
&\geq \exp\left(\delta \cdot \frac{d}{200 \log^2 d} \log N\right)
\end{aligned}$$

Case 1: $\ell \leq h$ and $p_\ell \dots p_h \leq \sqrt{d}$

Consider the depth-4 circuit C_{h+1} obtained by converting the first $2(h+1)$ layers into a $\Sigma\Pi$ circuit, and all subsequent layers by $\Sigma\Pi$ circuits as well. The fan-in of the bottom level multiplication gates is $p_{h+2} \dots p_\Delta \leq \log^2 d$ by the choice of h. However, we can still apply the lower bound for $t = \log^2 d$ (by a similar grouping of lower degree terms) to get

$$\begin{aligned}
(a_1, \dots, a_{h+1})^{p_1 \dots p_h} &\geq a_1 a_2^{p_1} \dots a_{h+1}^{p_1 \dots p_h} \geq \exp\left(\left(\frac{d}{\log^2 d}\right)(\delta \log N)\right) \\
\implies |\Phi| &\geq (a_1 \dots a_{h+1}) \geq \exp\left(\left(\frac{d}{\log^2 d \cdot p_1 \dots p_h}\right)(\delta \log N)\right)
\end{aligned}$$

Since $p_1 \dots p_{\ell-1} \leq 200$ by the choice of ℓ, and as $p_\ell \dots p_h \leq \sqrt{d}$ by assumption, we have that $p_1 \dots p_h \leq 200\sqrt{d}$. Hence, $|\Phi| = \exp\left(\Omega\left(\frac{\sqrt{d} \log N}{\log^2 d}\right)\right)$.

Case 2: $\ell \leq h$ and $p_\ell \dots p_h > \sqrt{d}$

The equation obtained as $(\text{Eqn}_h) \cdot (\text{Eqn}_{h-1})^{(p_{h-1}-1)} (\text{Eqn}_{h-2})^{(p_{h-2}-1)p_{h-1}} \dots (\text{Eqn}_\ell)^{(p_\ell - 1)p_{\ell+1} \dots p_{h-1}}$ has on the LHS an expression where the exponent of $a_\ell, \dots, a_h$ is $p_1 \dots p_{h-1}$, and all other a_i's have a smaller exponent. Hence we get

$$\begin{aligned}
(a_1 \dots a_h)^{p_1 \dots p_{h-1}} &\geq \exp\left[\frac{p_1 \dots p_{h-1}}{2} \cdot ((p_\ell - 1) + \dots + (p_{h-1} - 1) + p_h) \cdot (\delta \log N)\right] \\
\implies a_1 \dots a_h &\geq \exp\left[((p_\ell - 1) + \dots + (p_h - 1)) \cdot \left(\frac{\delta \log N}{2}\right)\right]
\end{aligned}$$

Thus, if we can show that $((p_\ell - 1) + \dots + (p_h - 1)) = \Omega(\log d)$, then it follows that $|\Phi| \geq a_1 \dots a_h =$

$\exp(\Omega(\log N \log d))$. By the AM-GM inequality,

$$\begin{aligned}((p_\ell - 1) + \cdots + (p_h - 1)) &\geq (h - \ell + 1)\left((p_\ell \ldots p_h)^{1/(h-\ell+1)} - 1\right) \\ &\geq (h - \ell + 1)\left(d^{1/2(h-\ell+1)} - 1\right) \\ &= r \cdot \left(e^{\log d/2r} - 1\right) \quad (\text{where } r = h - \ell + 1 > 0) \\ &\geq \frac{\log d}{2} \quad (\text{since } e^x > 1 + x)\end{aligned}$$

Therefore,

$$|\Phi| \geq a_1 \ldots a_h \geq \exp\left(0.5 \log d \cdot \left(\frac{\delta \log N}{2}\right)\right) = N^{\Omega(\log d)}$$

□

Combining Theorem 15 and Lemma 14, we immediately obtain a our main theorem, namely:

Theorem 1 (restated). Let $\mathbb{F}$ be any field. The family of polynomials $\mathcal{F} = \{F_n(\mathbf{x})\}$ as constructed above belongs to the class VNP. The n-th polynomial $F_n(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ of this family is an (n^2+1)-variate polynomial of degree $2n$ which requires regular arithmetic formulas of size $n^{\Omega(\log n)}$ to compute it.

7 Discussion and open problems

Having obtained superpolynomial regular formula lower bounds, a natural question is to see if one can prove superpolynomial lower bounds for less restricted formulas. Note that our notion of regular formulas entailed two restrictions - one on the formal degree and the other that the fanins of the gates at any layer be the same. It is natural to ask if this technique can be extended to prove super-polynomial lower bounds when say the restriction on the fanins at a layer being equal is removed. It is quite conceivable that the technique of dimension of shifted partials (or a slight modification thereof) may be powerful enough to prove such lower bounds. It is less clear if the technique of shifted partials is powerful enough to prove general formula lower bounds.

As far as the model of $\Sigma\Pi^{[O(\sqrt{d})]}\Sigma\Pi^{[\sqrt{d}]}$ circuits are concerned, the lower bound obtained from considering the dimension of shifted partials is fairly tight (since we have a matching upper bound). However, if we keep in mind that we eventually want to prove circuit/ABP/formula lower bounds, it is conceivable that some of the asymptotics can be improved for polynomials computed by small circuits/ABPs/formulas. There are at least three asymptotes in this lower bound, any of which on improving could result in separating the complexities of the determinant and permanent families:

- If we can construct an explicit n-variate degree d polynomial F that requires $\Sigma\Pi^{[O(\sqrt{d})]}\Sigma\Pi^{[\sqrt{d}]}$ circuits of top fanin $n^{\omega(\sqrt{d})}$.
- If we can show that any n-variate degree d polynomial F that *can be computed by a polynomial sized circuit* satisfies

$$\dim(\langle \boldsymbol{\partial}^{=k} F \rangle_{\leq \ell}) \leq \binom{\sqrt{d}}{k}\binom{n + \ell + (\sqrt{d}-1)k}{n} \cdot \exp(-\omega(\sqrt{d}\log n))$$

for $k = \varepsilon\sqrt{d}$ and $\ell = \frac{cn\sqrt{d}}{\log n}$.

- If we can show that any n-variate degree d polynomial that *can be computed by a polynomial sized ABP* can be computed by a regular formula of size $n^{o(\log d)}$.

Needless to say, a new complexity measure other than the dimension of shifted partials would be very enlightening, and possibly is even necessary for stronger lower bounds. Nevertheless, our lower bounds having come so close and given the recent progress in depth reduction [GKKS13b, Tav13] and lower bounds, one cannot help but be optimistic that the resolution of VP vs VNP may not be too distant in the future.

Acknowledgements

The authors would like to thank Ankit Gupta and Pritish Kamath for many discussions, especially for discussions pertaining to the power and limitations of the method of shifted partials and the choice of a candidate hard polynomial.

References

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

[AW09] S. Aaronson and A. Wigderson. Algebrization: A new barrier in complexity theory. *ACM Transactions on Computation Theory*, 1(1):1–54, 2009.

[BGS75] T. Baker, J. Gill, and R. Solovay. Relativizations of the P =? NP question. *SIAM Journal on Computing*, 4(4):431–442, 1975.

[BS83] W. Baur and V. Strassen. The complexity of partial derivatives. *Theoretical Computer Science*, 22(3):317–330, 1983.

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

[GKKS13a] Ankit Gupta, Neeraj Kayal, Pritish Kamath, and Ramprasad Saptharishi. Approaching the chasm at depth four. In *Conference on Computational Complexity (CCC)*, 2013.

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

[GKQ13] Ankit Gupta, Neeraj Kayal, and Youming Qiao. Random arithmetic formulas can be reconstructed efficiently. In *Conference on Computational Complexity (CCC)*, 2013.

[GR00] Dima Grigoriev and Alexander A. Razborov. Exponential lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. *Appl. Algebra Eng. Commun. Comput.*, 10(6):465–487, 2000.

[Kal85] K. Kalorkoti. A lower bound for the formula size of rational functions. *SIAM journal of Computing*, 14(3):678–687, 1985.

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

[Koi12] Pascal 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 Homogeneous Circuits with Bounded Top Fanin. Technical report, Electronic Colloquium on Computational Complexity (ECCC), 2013.

[Nis91] Noam Nisan. Lower bounds for non-commutative computation (extended abstract). In *STOC*, pages 410–418, 1991.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994.

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

[Raz09] R. Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *Journal of the Association for Computing Machinery*, 56(2), 2009.

[Raz10] Ran Raz. Tensor-rank and lower bounds for arithmetic formulas. In *STOC*, pages 659–666, 2010.

[RR94] A.A. Razborov and S. Rudich. Natural proofs. *Journal of Computer and System Sciences*, 55:204–213, 1994.

[RY08] R. Raz and A. Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. In *Proceedings of the 23rd IEEE Annual Conference on Computational Complexity*, pages 128–139, 2008.

[Rys63] H. J. Ryser. Combinatorial mathematics. *Math. Assoc. of America*, 14, 1963.

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

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

[Tav13] Sébastien Tavenas. Improved bounds for reduction to depth 4 and 3. In *Mathematical Foundations of Computer Science (MFCS)*, 2013.

[Val79] Leslie G. Valiant. Completeness Classes in Algebra. In *STOC*, pages 249–261, 1979.

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

ECCC ISSN 1433-8092

http://eccc.hpi-web.de