Electronic Colloquium on Computational Complexity, Revision 3 of Report No. 98 (2012)

Approaching the chasm at depth four

Ankit Gupta
Microsoft Research India
ankitgupta.iitkanpur@gmail.com

Pritish Kamath
Microsoft Research India
pritish.kamath@gmail.com

Neeraj Kayal
Microsoft Research India
neeraka@microsoft.com

Ramprasad Saptharishi
Chennai Mathematical Institute
ramprasad@cmi.ac.in

September 2, 2013

Abstract

Agrawal-Vinay [AV08], Koiran [Koi12] and Tavenas [Tav13] have recently shown that an $\exp(\omega(\sqrt{n}\log n))$ lower bound for depth four homogeneous circuits computing the permanent with bottom layer of $\times$ gates having fanin bounded by $\sqrt{n}$ translates to super-polynomial lower bound for general arithmetic circuits computing the permanent. Motivated by this, we examine the complexity of computing the permanent and determinant via such homogeneous depth four circuits with bounded bottom fanin.

We show here that any homogeneous depth four arithmetic circuit with bottom fanin bounded by $\sqrt{n}$ computing the permanent (or the determinant) must be of size $\exp(\Omega(\sqrt{n}))$.

1 Introduction

Background. The most natural and intuitive way to compute a polynomial is via an arithmetic circuit. In this model the inputs are variables $x_1, x_2, \ldots, x_n$ and the computation is performed using the operations $+, \times$. We typically allow arbitrary constants from a field $\mathbb{F}$ on the incoming edges to a $+$ gate so that the output of a $+$ gate is an arbitrary $\mathbb{F}$-linear combination of its inputs. The complexity measures associated with circuits are size and depth, which capture the number of operations and the maximal distance between an input and the output respectively.
Recall that the permanent is an n^2-variate homogeneous[1] polynomial of degree n defined as:

$$\mathsf{Perm}_n \quad = \quad \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i\sigma(i)}$$

The permanent, by virtue of being complete for the class VNP (an algebraic analogue of the class NP, defined in [Val79]), occupies a central position in the study of the complexity of counting problems. The best known arithmetic circuit for the permanent is actually a depth three homogeneous circuit of size $O(n^2 \cdot 2^n)$ and is called the Ryser's formula . Its illustrious sibling, the determinant, is widely believed to be comparatively easy, being complete for a subclass of VP (an algebraic

[1] A multivariate polynomial is said to be homogeneous if all its monomials have the same total degree.

ISSN 1433-8092

analogue of P, also defined in [Val79]). It is conjectured (cf. [AV08]) that any arithmetic circuit computing the $n \times n$ permanent must be of $\exp(n)$ size. Meanwhile, the arithmetic complexity of computing the determinant equals $\tilde{O}(n^\omega)$, where ω is the exponent of matrix multiplication. Resolving the arithmetic complexity of computing the permanent and the determinant (i.e. determining the exponent of matrix multiplication) are two of the most fascinating open problems of our times.

Prior Work. Lower bounds have been obtained earlier for depth three arithmetic circuits (with some restrictions) and constant depth multilinear circuits. Specifically, Nisan and Wigderson [NW97] showed that any homogeneous depth three circuit computing the permanent (also the determinant) must be of exponential size. Following that, Grigoriev and Karpinski [GK98], and Grigoriev and Razborov [GR00] showed that any depth three arithmetic circuit over a finite field computing the permanent (also the determinant) requires exponential size but proving lower bounds for depth three circuits over fields of characteristic zero (or even over the algebraic closure of a finite field) remains an outstanding open problem. In this direction Shpilka and Wigderson [SW01] proved quadratic lower bounds for depth three circuits over arbitrary fields (without the homogeneity restriction). Meanwhile, Raz [Raz09] showed that any multilinear formula computing the permanent (also the determinant) must be of super-polynomial size. Following this, Raz and Yehudayoff [RY08] proved exponential lower bounds for constant depth multilinear circuits.

The model. In this work, we focus our attention on depth four arithmetic circuits with bottom fanin bounded by a parameter b, and fan-in of and all multiplication gates in the layer adjacent to the output node have fanin at most a, which we denote by $\Sigma\Pi^{[a]}\Sigma\Pi^{[b]}$ circuits. A $\Sigma\Pi^{[a]}\Sigma\Pi^{[b]}$ circuit computes a polynomial of the form

$$C \quad = \quad \sum_{i=1}^{s} (Q_{i1} \cdot Q_{i2} \cdot \ldots \cdot Q_{id_i}) \tag{1}$$

where each $d_i \leq a$ and each Q_{ij} is polynomial of degree bounded by b. The number of summands s is called the top fanin of the circuit. Our motivation for investigating representations of the form (1) stems from a recent result of Agrawal and Vinay [AV08], and a subsequent strengthening by Koiran [Koi12] and Tavenas [Tav13]

Theorem 1. *[AV08, Koi12, Tav13] If there is a polynomial sized arithmetic circuit computing* Perm_n*, then there is a* $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$*-circuit of top fan-in* $2^{O(\sqrt{n}\log n)}$ *computing* Perm_n*.*

The contra-positive of the above statement is that it suffices to show a $2^{\omega(\sqrt{n}\log n)}$ lower bound for the top fan-in of $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuits computing the Perm_n to prove a super-polynomial circuit lower bound. Thus, a good enough lower bound for $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuits would imply super-polynomial lower bounds for Perm_n. In this paper, we give a lower bound for the permanent (or determinant) that comes very close to the above threshold.

Theorem 2. *Over any field* $\mathbb{F}$*, any* $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ *circuit that computes* Perm_n *(or* Det_n*) must have top fanin at least* $\exp\left(\Omega\left(\frac{n}{t}\right)\right)$*. In particular, any* $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ *circuit that computes* Perm_n *(or* Det_n*) must have top fan-in* $\exp\left(\Omega\left(\sqrt{n}\right)\right)$*.*

More generally, we show the following:

Theorem 3. *Let $t : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ be any increasing function such that $t(n) = o(n)$. Let $m = (cn/t)$, where $c \geq 1$ is any fixed constant. Suppose that over some field $\mathbb{F}$, the polynomial $\mathsf{Perm}_n(\mathbf{x})$ can be written as*

$$\mathsf{Perm}_n(\mathbf{x}) = \sum_{i=1}^{s} F_i\left(Q_{i1}, \cdots, Q_{im}\right) \tag{2}$$

where each $F_i \in \mathbb{F}[u_1, u_2, \ldots, u_m]$ is an arbitrary m-variate polynomial and each $Q_{ij} \in \mathbb{F}[\mathbf{x}]$ is a polynomial of degree at most t over the n^2 variables $\mathbf{x} = (x_{11}, \ldots, x_{nn})$ of Perm_n. Then the number of summands s must be at least $\exp\left(\Omega\left(\frac{n}{t}\right)\right)$.

Theorem 2 can be seen to be the special case of the above theorem where each F_i is simply the product of its input variables, i.e.

$$F_i(\mathbf{u}) := u_1 \cdot u_2 \cdot \ldots \cdot u_m \quad \text{for each } i \in [s].$$

Note that the only restriction on the F_i's in Theorem 3 above is that each of them is a $O(n/t)$-variate polynomial. In particular, the F_i's can have arbitrarily large degree and complexity. The rest of the paper is devoted to the proof of this theorem. Our proof will be completely elementary and self-contained. Moreover, the above lower bound holds for both Perm_n as well as Det_n. However, it is quite possible that this lower bound can be improved for Perm_n by improving the estimate of the dimension of a certain explicit vector space[2] that comes up in our proof. We discuss this in Section 8 and make a specific conjecture in this regard.

1.1 Subsequent results

Following our lower bound, further progress has been made using similar techniques. Kumar and Saraf [KS13] showed exponential lower bounds for homogeneous depth-4 circuits of bounded top fan-in. For the class of $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuits studied in this paper, Kayal, Saha and Saptharishi [KSS13] showed a lower bound of $n^{\Omega(\sqrt{d})}$ for an explicit polynomial of degree d in the class VNP. Shortly after that, Fournier, Limaye, Malod and Srinivasan [FLMS13] showed a similar $n^{\Omega(\sqrt{d})}$ lower bound for the iterated matrix multiplication polynomial which is in the class VP.

All the above results use the same *complexity measure* used in this paper, and we begin by giving a quick overview of our proof technique.

2 Basic Idea and Outline

Our key idea is to exploit the *shifted derivatives of a polynomial* - a notion that we now define. Let $\mathbb{F}$ be a field and $\mathbb{F}[\mathbf{x}]$ be the set of polynomials over $\mathbb{F}$ in the set of variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. For an n-tuple $\mathbf{i} = (i_1, i_2, \ldots, i_n) \in \mathbb{Z}^n_{\geq 0}$, $\mathbf{x}^{\mathbf{i}}$ denotes the monomial $(x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n})$ which has degree $|\mathbf{i}| \stackrel{\text{def}}{=} (i_1 + i_2 + \ldots + i_n)$. We shall use $\partial^{\mathbf{i}} f$ to denote the partial derivative of f with respect to the monomial $\mathbf{x}^{\mathbf{i}}$,

$$\partial^{\mathbf{i}} f \quad \stackrel{\text{def}}{=} \quad \frac{\partial^{i_1}}{\partial x_1^{i_1}}\left(\frac{\partial^{i_2}}{\partial x_2^{i_2}}\left(\cdots\left(\frac{\partial^{i_n} f}{\partial x_n^{i_n}}\right)\cdots\right)\right)$$

[2] As we will see later, the vector space being referred to corresponds to a slice of the ideal generated by the derivatives of Perm_n.

For a finite subset of polynomials $S \subseteq \mathbb{F}[\mathbf{x}]$, we shall use $\mathbb{F}\text{-span}\,(S)$ to refer to the set of all possible $\mathbb{F}$-linear combinations of polynomials in S. i.e.

$$\mathbb{F}\text{-span}\,(S) \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{|S|} \alpha_i \cdot f_i \ : \ \alpha_i \in \mathbb{F}, \quad f_i \in S \right\}$$

With these notational preliminaries in hand, we are now ready to define our key concept.

Definition 1 (Shifted Derivatives). *Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a multivariate polynomial. The* span of the ℓ-shifted k-th order derivatives *of f, denoted $\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell}$, is defined as*

$$\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell} \stackrel{\text{def}}{=} \mathbb{F}\text{-}span \left\{ \mathbf{x}^{\mathbf{i}} \cdot (\partial^{\mathbf{j}} f) \quad : \quad \mathbf{i}, \mathbf{j} \in \mathbb{Z}^n_{\geq 0} \quad \text{with} \ |\mathbf{i}| \leq \ell \quad \text{and} \ |\mathbf{j}| = k \right\}$$

$\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell}$ forms an $\mathbb{F}$-vector space and we denote by $\dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell})$ the dimension of this space.

Recent work in arithmetic complexity has shown how $\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell}$ can give insights into the structure and complexity of f in ways that are sometimes surprising and unexpected. Kayal [Kay12a] showed that $\langle \boldsymbol{\partial}^{=1} f \rangle_{\leq 1}$ yields a lie algebra that can help efficiently determine if f is equivalent (via an affine change of variables) to the permanent (or determinant). For $\ell = \infty$, note that $\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell}$ is precisely the ideal generated by the k-th order derivatives of f. Gupta, Kayal and Qiao [GKQ12] recently exploited the structure of $\langle \boldsymbol{\partial}^{=1} f \rangle_{\leq \infty}$ to devise an efficient reconstruction algorithm for random arithmetic formulas. Note that the dimension of partial derivatives employed by Nisan and Wigderson [NW97] in their lower bound proofs corresponds to looking at $\dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq 0})$. Closer to the present application, Kayal [Kay12b] showed how $\dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell})$ (for suitably chosen ℓ and k) can be used to prove an exponential lower bound for representing a polynomial as a sum of powers of bounded degree polynomials. We show here that for suitably chosen values of ℓ and k, $\dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell})$ is comparatively small when f is computed by a $\Sigma\Pi^{[O(n/t)]}\Sigma\Pi^{[t]}$ circuit (Corollary 10). Meanwhile $\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Perm}_n \rangle_{\leq \ell})$ is relatively large (Corollary 16). This gives the lower bound.

2.1 Intuition from algebraic geometry

In order to prove lower bounds for $\Sigma\Pi^{[a]}\Sigma\Pi^{[b]}$ circuits, we would like to use a property that "distinguishes" a term of the form $T = Q_1 \dots Q_a$ (where $\deg(Q_i) \leq b$) from a random polynomial of degree ab. One of the properties that is useful in this regard is that any point $\mathbf{t}$ such that $Q_1(\mathbf{t}) = \cdots = Q_a(\mathbf{t}) = 0$ is a root of T of multiplicity a. Further, if $a \ll n$, we must have a large number of such $\mathbf{t}$'s. A random polynomial, however, does not have many roots of large multiplicity, and the dimension of shifted partial derivatives try to capture exactly this distinction.

Given an ideal $I = \langle f_1, \dots, f_k \rangle$, the variety of I (denoted by $\mathbb{V}(I)$) is the set of common zeros of $f_1, \dots, f_k$. The *dimension of the variety* is a well-defined notion in algebraic geometry and intuitively captures the "largeness" of the variety. Let $\mathbb{F}[x_1, \dots, x_n]_{\leq \ell}$ denote the set of polynomials of degree at most ℓ, and $\gamma_\ell(I) = \dim\left(I \cap \mathbb{F}[x_1, \dots, x_n]_{\leq \ell}\right)$ which in other words is the dimension of all polynomials in I of degree bounded by ℓ. Intuitively, if $\gamma_\ell(I)$ is large, then the ideal I imposes many constraints, and hence the dimension of the variety $\mathbb{V}(I)$ must be small. This intuition that the growth of $\gamma_\ell(I)$ is inversely related to the dimension of I is captured by the notion of the

Hilbert polynomial. The interested reader can learn more about dimension of varieties and Hilbert polynomials from any standard text on algebraic geometry (for example [CLO07]).
In our setting, we would like to study the roots of large multiplicity of a homogeneous polynomial f of degree d, and if $I = \langle \partial^{=k} f \rangle$ then $\mathbb{V}(I)$ is precisely the roots of multiplicity at least $k+1$. Thus, if $\gamma_\ell(I)$ is large, then there are few roots of large multiplicity. Notice that $\gamma_\ell(I) = \dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell-(d-k)})$, the dimension of shifted partial derivatives (of suitable parameters). Hence larger the dimension of shifted partials then fewer the roots of large multiplicity. Thus, one expects the shifted partials of a term $T = Q_1 \dots Q_a$ to be small, whereas the shifted partials of polynomials like Det_n or Perm_n ought to be large.

Outline of the rest of the paper. We execute this idea in the rest of the paper as follows. In Section 4 we give an upper bound on $\langle \boldsymbol{\partial}^{=k} C \rangle_{\leq \ell}$ when C is a polynomial of the form

$$C = \sum_{i=1}^{s} F_i\left(Q_{i1}, \cdots, Q_{im}\right) \quad \text{where each } Q_{ij} \text{ has degree at most } t.$$

In Section 5, we give a lower bound estimate for $\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Perm}_n \rangle_{\leq \ell})$. We then combine these bounds to obtain a proof of our main theorem in Section 6. Finally, in Section 8, we conclude by discussing the possibility of improving the estimates for $\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Perm}_n \rangle_{\leq \ell})$ obtained here.

3 Preliminaries

Notation. We will use $[n]$ to denote the set $\{1, \cdots, n\}$ for any $n \geq 1$. $\mathbf{x}_n$ denotes the set of variables $\{x_1, x_2, \cdots, x_n\}$. However, when the context is clear, we would use just $\mathbf{x}$ instead of $\mathbf{x}_n$. Similarly for $\mathbf{y}$, $\mathbf{z}$, etc. We use $\boldsymbol{\partial}^{=k} f$ to denote the set of all k-th order partial derivatives of f. If $S \subseteq \mathbb{F}[\mathbf{x}]$, then,

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

Useful asymptotic estimates and inequalities. We now collect together some useful estimates for binomial coefficients that follow from Stirling's formula.

Definition 2. *The binary entropy function H_2 is defined as*

$$H_2(x) \quad = \quad -x \cdot \log_2(x) - (1-x) \cdot \log_2(1-x)$$

The natural-log version of the entropy function, denoted by H_e is defined analogously as

$$H_e(x) \quad = \quad -x \cdot \ln(x) - (1-x) \ln(1-x)$$

Lemma 4. *For any $0 < x < 1$, we have $x \ln \frac{1}{x} \leq H_e(x) \leq x \ln \frac{1}{x} + x$.*

Proposition 5 (Stirling's Formula, cf. [Rom])**.** $\ln(n!) \quad = \quad n \ln n - n + O(\ln n)$

Stirling's formula can be used to obtain the following estimates (proofs of which are in section 6.1).

Lemma 6. *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,*

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

Lemma 7. *For any constants $\alpha \geq \beta > 0$,*

$$\ln \binom{\alpha n}{\beta n} \quad = \quad aH_e\left(\frac{\beta}{\alpha}\right) n - O(\ln n)$$

4 Upper bounding the dimension of shifted partials of $\Sigma\Pi^{[m]}\Sigma\Pi^{[t]}$ circuits

In this section we give an upper bound on $\dim(\langle \boldsymbol{\partial}^{=k} C\rangle_{\leq \ell})$ where C is an expression of the form

$$C = \sum_{i=1}^{s} F_i\left(Q_{i1}, \cdots, Q_{im}\right), \quad \text{where } \deg(Q_{ij}) \leq t \quad \text{for each } i \in [s], j \in [m] \tag{3}$$

where in the special case of $F_i(u_1, \cdots, u_m) = u_1 \cdots u_m$, C becomes a $\Sigma\Pi^{[m]}\Sigma\Pi^{[t]}$ circuit.

We begin by noting that $\dim(\langle \boldsymbol{\partial}^{=k} f\rangle_{\leq \ell})$ is sub-additive.

Proposition 8. Sub-additivity. *For all $k, \ell \geq 0$, we have* $\dim(\langle \boldsymbol{\partial}^{=k}(f+g)\rangle_{\leq \ell}) \ \leq \ \dim(\langle \boldsymbol{\partial}^{=k} f\rangle_{\leq \ell}) + \dim(\langle \boldsymbol{\partial}^{=k} g\rangle_{\leq \ell})$.

Proof. By linearity of partial derivatives, we have $\mathbf{x}^{\mathbf{i}} \cdot \partial^{\mathbf{j}}(f+g) = \mathbf{x}^{\mathbf{i}} \cdot \partial^{\mathbf{j}} f + \mathbf{x}^{\mathbf{i}} \cdot \partial^{\mathbf{j}} g$. Hence,

$$\langle \boldsymbol{\partial}^{=k}(f+g)\rangle_{\leq \ell} \quad \subseteq \quad \mathbb{F}\text{-span}\left(\langle \boldsymbol{\partial}^{=k} f\rangle_{\leq \ell} \cup \langle \boldsymbol{\partial}^{=k} g\rangle_{\leq \ell}\right)$$

The proposition follows. □

Let C be an expression of the form (3). By Proposition 8, it suffices to understand the growth of $\dim(\langle \boldsymbol{\partial}^{=k} T\rangle_{\leq \ell})$ of a single term T of the form

$$T = F\left(Q_1, Q_2, \ldots, Q_m\right)$$

where $F(u_1, u_2, \ldots, u_m) \in \mathbb{F}[u_1, u_2, \ldots, u_m]$ is an m-variate polynomial and each $Q_i \in \mathbb{F}[\mathbf{x}_N]$ is N-variate polynomial of degree at most t. Towards this end, let us first compute the derivatives of such a term. By the chain rule, we have

$$\frac{\partial T}{\partial x_1} = \sum_{i \in [m]} \frac{\partial F}{\partial u_i}(\mathbf{Q}) \cdot \frac{\partial Q_i}{\partial x_1}, \quad \text{where } \mathbf{Q} = (Q_1, Q_2, \ldots, Q_m) \tag{4}$$

Let $(\boldsymbol{\partial}^{\leq k} F)(\mathbf{Q})$ be a shorthand notation for the set

$$\left\{(\partial^{\mathbf{i}} F)(\mathbf{Q}) \quad : \quad i \in \mathbb{Z}^m_{\geq 0}, \ \ |\mathbf{i}| \leq k\right\} \quad \subseteq \quad \mathbb{F}[\mathbf{x}_N]$$

Now each derivative $\left(\frac{\partial Q_i}{\partial x_1}\right)$ has degree at most $(t-1)$. Thus equation (4) implies in particular that

$$\begin{aligned}\frac{\partial T}{\partial x_1} &\in \mathbb{F}\text{-span}\left(\left\{\frac{\partial F}{\partial u_i}(\mathbf{Q}) \quad : \quad i \in [m]\right\} \cdot \left\{\mathbf{x}^{\mathbf{i}} \quad : \quad \mathbf{i} \in \mathbb{Z}^N_{\geq 0}, \quad |\mathbf{i}| \leq (t-1)\right\}\right) \\ &\subseteq \left\langle (\boldsymbol{\partial}^{\leq 1} F)(\mathbf{Q}) \right\rangle_{\leq (t-1)}\end{aligned}$$

There is of course nothing special about the choice of the variable x_1 so that for every $i \in [N]$ we have

$$\frac{\partial T}{\partial x_i} \in \left\langle (\boldsymbol{\partial}^{\leq 1} F)(\mathbf{Q}) \right\rangle_{\leq (t-1)}$$

which can be stated succinctly as

$$\boldsymbol{\partial}^{=1} T \subseteq \left\langle (\boldsymbol{\partial}^{\leq 1} F)(\mathbf{Q}) \right\rangle_{\leq (t-1)}$$

Differentiating equation (4) again with respect to x_2 we have

$$\begin{aligned}\frac{\partial^2 T}{\partial x_1 \cdot \partial x_2} &= \sum_{i \in [m]} \frac{\partial F}{\partial u_i}(\mathbf{Q}) \cdot \frac{\partial^2 Q_i}{\partial x_1 \cdot \partial x_2} + \sum_{i,j \in [m]} \frac{\partial^2 F}{\partial u_i \cdot \partial u_j}(\mathbf{Q}) \cdot \frac{\partial Q_i}{\partial x_1} \cdot \frac{\partial Q_j}{\partial x_2} \\ &\in \left\langle (\boldsymbol{\partial}^{\leq 2} F)(\mathbf{Q}) \right\rangle_{\leq (2t-2)}\end{aligned}$$

As before there is nothing special about the pair of variables x_1 and x_2 so that we have

$$\boldsymbol{\partial}^{=2} T \subseteq \left\langle (\boldsymbol{\partial}^{\leq 2} F)(\mathbf{Q}) \right\rangle_{\leq (2t-2)}$$

Continuing in this manner we see that

$$\boldsymbol{\partial}^{=k} T \subseteq \left\langle (\boldsymbol{\partial}^{\leq k} F)(\mathbf{Q}) \right\rangle_{\leq k(t-1)}$$

Therefore

$$\begin{aligned}\langle \boldsymbol{\partial}^{=k} T \rangle_{\leq \ell} &\subseteq \left\langle (\boldsymbol{\partial}^{\leq k} F)(\mathbf{Q}) \right\rangle_{\leq (\ell + k(t-1))} \\ &= \mathbb{F}\text{-span}\left((\boldsymbol{\partial}^{\leq k} F)(\mathbf{Q}) \cdot \left\{\mathbf{x}^{\mathbf{i}} \ : \ \mathbf{i} \in \mathbb{Z}^N_{\geq 0}, \quad |\mathbf{i}| \leq (\ell + k(t-1))\right\}\right)\end{aligned}$$

In particular this means that

$$\begin{aligned}\dim(\langle \boldsymbol{\partial}^{=k} T \rangle_{\leq \ell}) &\leq \dim\left((\boldsymbol{\partial}^{\leq k} F)(\mathbf{Q}) \cdot \left\{\mathbf{x}^{\mathbf{i}} \ : \ \mathbf{i} \in \mathbb{Z}^N_{\geq 0}, \quad |\mathbf{i}| \leq (\ell + k(t-1))\right\}\right) \\ &\leq \dim\left((\boldsymbol{\partial}^{\leq k} F)(\mathbf{Q})\right) \cdot \dim\left(\left\{\mathbf{x}^{\mathbf{i}} \ : \ \mathbf{i} \in \mathbb{Z}^N_{\geq 0}, \quad |\mathbf{i}| \leq (\ell + k(t-1))\right\}\right) \\ &= \binom{m+k}{k} \cdot \binom{N + (t-1)k + \ell}{N}\end{aligned}$$

Let us record the above as a proposition.

Proposition 9. *Let $T = F(\mathbf{Q})$ where $\mathbf{Q} = (Q_1, Q_2, \ldots, Q_m) \in (\mathbb{F}[\mathbf{x}_N])^m$ is an m-tuple of N-variate polynomials with each $Q_i \in \mathbb{F}[\mathbf{x}_N]$ having degree bounded by t. Then*

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

The following corollary follows directly from the above observation via sub-additivity (Proposition 8).

Corollary 10. *If* $C = \sum_{i=1}^{s} F_i\left(Q_{i1}, Q_{i2}, \ldots Q_{im}\right)$ *where each* $Q_{ij} \in \mathbb{F}[\mathbf{x}_N]$ *is a polynomial of degree bounded by* t*, then for any* $k \leq m$

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

In the next section we give a reasonable lower bound for $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell})$ for suitable choice of parameters k and ℓ.

5 Lower bounding the dimension of shifted partials of the Permanent

Reducing dimension computation to counting leading monomials. In this section, we shall present a lower bound for $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell})$. Let $\succ$ be any admissible monomial ordering[3]. Recall that the leading monomial of a polynomial $f \in \mathbb{F}[\mathbf{x}]$, denoted $\mathsf{LM}(f)$, is the largest monomial $\mathbf{x}^{\mathbf{i}}$ under the ordering $\succ$.

Proposition 11. *Let* $S \subseteq \mathbb{F}[\mathbf{x}]$ *be any finite set of polynomials. Then*

$$\dim(\mathbb{F}\text{-span}\,(S)) \quad = \quad \#\left\{\mathsf{LM}(f) \;:\; f \in \mathbb{F}\text{-span}\,(S)\right\}$$

The proof is a simple application of Gaussian elimination. As a corollary we obtain

Corollary 12. *For any polynomial* $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *we have*

$$\dim(\langle \boldsymbol{\partial}^{=k} f\rangle_{\leq \ell}) \quad \geq \quad \#\{\mathbf{x}^{\mathbf{i}} \cdot \mathsf{LM}(\boldsymbol{\partial}^{\mathbf{j}} f) \;:\; \mathbf{i},\mathbf{j} \in \mathbb{Z}_{\geq 0}^{|\mathbf{x}|},\ |\mathbf{i}| \leq \ell \quad \text{and} \quad |\mathbf{j}| = k\}$$

The lower bound given by this corollary is usually a severe underestimate but fortunately even this will suffice for our purpose for the case when $f = \mathsf{Perm}_n$.

Reduction to counting monomials with increasing subsequences. Let us fix $\succ$ to be the lexicographic monomial ordering induced by the following ordering on the variables: $x_{11} \succ \cdots \succ x_{1n} \succ x_{21} \succ \cdots \succ x_{nn}$. Note that any partial derivative of Perm_n is just the corresponding permanental minor (or just 'P-minor'). Hence by the above corollary we have

$$\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell}) \quad \geq \quad \#\left\{\mathbf{x}^{\mathbf{i}} \cdot \mathsf{LM}(M) \;:\; \begin{array}{l}\mathbf{x}^{\mathbf{i}} \text{ is a monomial of degree at most } \ell \text{ and} \\ M \text{ is an } (n-k)\times(n-k) \text{ P-minor}\end{array}\right\}$$

Note that the leading monomial under $\succ$ of any $(n-k)\times(n-k)$ P-minor M is just the product of the variables along the principal diagonal of M. Now if the variables along the principal minor of M are $(x_{i_1 j_1}, \cdots, x_{i_{n-k} j_{n-k}})$ then the indices satisfy

$$i_1 < i_2 < \ldots < i_{n-k} \quad \text{and } j_1 < j_2 < \ldots < j_{n-k}$$

This naturally leads to the following definition.

[3]For more on monomial orderings and their applications in algebraic geometry, we refer the interested reader to Chapter 2 of the text by Cox, Little and O'Shea [CLO07]

Definition 3. *We shall refer to a sequence of variables* $(x_{i_1 j_1}, \cdots, x_{i_t j_t})$ *as a* t-increasing sequence *if the indices satisfy*

$$i_1 < i_2 < \ldots < i_t \quad \text{and } j_1 < j_2 < \ldots < j_t$$

We will say that a monomial $A = \mathbf{x}^{\mathbf{j}}$ contains a t-increasing sequence *if there exists an increasing sequence* $(x_{i_1 j_1}, \cdots, x_{i_t j_t})$ *wherein every variable* $x_{i_r j_r}$ *(*$r \in [t]$*) divides* A*.*

In this terminology we would then say that the leading monomial of any $(n-k)\times(n-k)$ P-minor is exactly the product of the variables in an $(n-k)$-increasing sequence. Consequently for any P-minor M of size $(n-k)$ we have that $\mathbf{x}^{\mathbf{i}} \cdot \mathsf{LM}(M)$ contains an $(n-k)$-increasing sequence. Conversely, every monomial of degree at most $(n-k+\ell)$ that contains an $(n-k)$-increasing sequence can be written as the leading monomial of $\mathbf{x}^{\mathbf{i}} \cdot M$ for some monomial $\mathbf{x}^{\mathbf{i}}$ of degree at most ℓ and an $(n-k)\times(n-k)$ P-minor M. Hence we have:

Corollary 13. $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell})$ *is lower bounded by the number of distinct monomials of degree at most* $(n-k+\ell)$ *over* n^2 *variables* $(\{x_{ij} \ : \ i,j \in [n]\})$ *that contain an* $(n-k)$*-increasing sequence.*

In order to count the number of monomials of degree bounded by $(n-k+\ell)$ that contain an $(n-k)$-increasing sequence, we shall restrict ourselves to a very *small set* of variables to contribute the increasing sequence, and "fill-up" the remaining degree using the other variables. The "small set" that we consider here is just two diagonals – the principal diagonal and the one above it.

5.1 Restricting to two diagonals

We shall focus on the variables $D_{2,n} = \{x_{ii} \ : \ 1 \leq i \leq n\} \cup \{x_{i(i+1)} \ : \ 1 \leq i \leq n-1\}$.

Lemma 14. *Fix parameters* $n, r \geq 0$*, and let* $S_2(n,r)$ *be the number of distinct* r*-increasing sequences in* $D_{2,n}$*. Then for any* $d > 0$*, the number of monomials of degree bounded by* d *containing an* r*-increasing sequence is lower bounded by* $S_2(n,r) \cdot \binom{n^2-2n+d+r}{d-r}$

Proof. For any monomial, define the *support in* $D_{2,n}$ as the set of variables in $D_{2,n}$ that divide it. Our strategy is to obtain a lower bound on the number of monomials of degree bounded by d which contain an r-increasing sequence entirely inside its support in $D_{2,n}$. We shall start with any r-increasing sequence contained in $D_{2,n}$ and multiply this with suitable monomials to obtain monomials containing an r-increasing sequence. To avoid double counting, we shall multiply by monomials involving only those variables that do not alter the *leading* r-increasing sequence among all r-increasing sequences it contains.[4] Consider any particular r-increasing sequence, call it Q, in $D_{2,n}$. We show that the total number of monomials (of degree $\leq d$) with Q as the leading r-increasing sequence is at least $\binom{n^2-2n+d+r}{d-r}$, which clearly suffices to prove the lemma.
For any variable $x_{ij} \in D_{2,n}$, define its companions to be the variables to its right in the same row, or below it in the same column, i.e. $\{x_{ij'} \ : \ j' > j\} \cup \{x_{i'j} \ : \ i' > i\}$. Let Q' be the set of all companions of variables in Q which are in $D_{2,n}$. The key observation is that adding elements of Q' to Q does not alter the leading increasing sequence. For any increasing sequence that uses elements of Q', replacing every $x_{i'j'} \in Q'$ by the corresponding $x_{ij} \in Q$ for which it is a companion for yields a "higher" increasing sequence. Hence adding any subset $T \subseteq Q'$ to Q does not alter the leading increasing sequence.

[4] The *leading* r-increasing sequence is the one which is largest in the monomial ordering defined earlier.

Note that every element of $D_{2,n}$ besides x_{nn} has exactly one companion in $D_{2,n}$. Hence, there are at least $(r-1)$ other variables in $D_{2,n}$ we can freely use to augment Q to a degree $\leq d$ monomial without changing the increasing sequence. The total number of variables that can be used is at least $n^2-(2n-1)+r+(r-1)$, and the degree to augment is at most $d-r$. Hence, there are $\binom{n^2-2n+d+r}{d-r}$ distinct monomial of degree $\leq d$ that contain Q as the leading r-increasing sequence. □

Now, all we need to do is to compute $S_2(n,r)$, which is the number of r-increasing sequences contained in $D_{2,n}$.

Lemma 15. *The number of r-increasing sequences contained in $D_{2,n}$ is equal to $\binom{2n-r}{r}$*

Proof. Consider the $(2n-1)$ variables in $D_{2,n}$ in the sequence $x_{11}, x_{12}, x_{22}, \ldots, x_{nn}$. Picking an r-increasing sequence is the same as picking r of the $(2n-1)$ variables such that no two adjacent variables (in the above order) are chosen. This can be thought as distributing the $(2n-r-1)$ variables that won't be picked such that there is at least one variable between any two variables that are picked, and this is exactly equal to

$$\binom{(2n-r-1-(r-1))+(r+1)-1}{(r+1)-1} = \binom{2n-r}{r}$$

□

By setting $d=\ell+n-k$ and $r=n-k$ in Lemma 14 and using Lemma 15 with these parameters, we get the following lower bound for $\dim(\langle \boldsymbol{\partial}^{=k}\mathsf{Perm}_n\rangle_{\leq \ell})$ via Corollary 13.

Corollary 16. *For every $n,k,\ell \geq 0$,*

$$\dim(\langle \boldsymbol{\partial}^{=k}\mathsf{Perm}_n\rangle_{\leq \ell}) \;\geq\; \binom{n+k}{2k}\cdot\binom{n^2+\ell-2k}{\ell}$$

6 Putting it all together

We are now ready to prove the main theorem which is the following.

Theorem 3 (restated). Let $t:\mathbb{Z}_{\geq 0}\to\mathbb{Z}_{\geq 0}$ be any increasing function such that $t(n)=o(n)$. Let $m=(cn/t)$, where $c\geq 1$ is any fixed constant. Suppose that over some field $\mathbb{F}$, the polynomial $\mathsf{Perm}_n(\mathbf{x})$ can be written as

$$\mathsf{Perm}_n(\mathbf{x}) = \sum_{i=1}^{s} F_i\left(Q_{i1},\cdots,Q_{im}\right) \tag{5}$$

where each $F_i\in\mathbb{F}[u_1,u_2,\ldots,u_m]$ is an arbitrary m-variate polynomial and each $Q_{ij}\in\mathbb{F}[\mathbf{x}]$ is a polynomial of degree at most t over the n^2 variables $\mathbf{x}=(x_{11},\ldots,x_{nn})$ of Perm_n. Then the number of summands s must be at least $\exp\left(\Omega\left(\frac{n}{t}\right)\right)$.

Proof. The proof involves comparing the dimension of shifted partials for the two sides of a polynomial identity of the form

$$\mathsf{Perm}_n = \sum_{i=1}^{s} F_i\left(Q_{i1},\cdots,Q_{im}\right)$$

Corollary 10 can be used to upper bound the dimension of shifted partials of the right-hand side of the equation 5 so that we have

$$\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell}) \quad \leq \quad s \cdot \binom{cn/t+k}{k}\binom{n^2+\ell+(t-1)k}{n^2} \tag{6}$$

On the other hand, Corollary 16 gives a lower bound for $\dim(\langle \boldsymbol{\partial}^{=k}\mathsf{Perm}_n\rangle_{\leq \ell})$:

$$\dim(\langle \boldsymbol{\partial}^{=k}\mathsf{Perm}_n\rangle_{\leq \ell}) \quad \geq \quad \binom{n+k}{2k}\binom{n^2+\ell-2k}{\ell} \tag{7}$$

Both these equations imply that

$$s \quad \geq \quad \frac{\binom{n+k}{2k}\binom{n^2+\ell-2k}{\ell}}{\binom{cn/t+k}{k}\binom{n^2+\ell+(t-1)k}{n^2}}$$

We shall set parameters as $\ell = n^2 t$ and $k = \varepsilon(n/t)$ (for an $\varepsilon > 0$ that shall be chosen shortly). The proofs of the following estimates for binomial coefficients are straightforward applications of Lemma 6 and Lemma 7, and we defer the proofs to Section 6.1.

Claim 17. *For the above choice of parameters:*

(a) $\ln \binom{n+k}{2k} = 2\varepsilon\left(\frac{n}{t}\right)\left(\ln\left(\frac{t}{2\varepsilon}\right)+1\right) \pm O\left(\frac{n}{t^2}\right)$

(b) $\ln \binom{cn/t+k}{k} = (c+\varepsilon)H_e\left(\frac{\varepsilon}{c+\varepsilon}\right)\cdot\left(\frac{n}{t}\right) - O(\ln n)$

(c) $\ln \frac{\binom{n^2+\ell-2k}{\ell}}{\binom{n^2+\ell+(t-1)k}{n^2}} = -2\varepsilon\left(\frac{n}{t}\right)\ln t - \varepsilon\left(\frac{n}{t}\right) \pm O(1)$

Using this, we get

$$\ln s \geq \left(2\varepsilon \ln\left(\frac{1}{2\varepsilon}\right) + \varepsilon - (c+\varepsilon)H_e\left(\frac{\varepsilon}{c+\varepsilon}\right)\right)\left(\frac{n}{t}\right) \;\pm\; O\left(\frac{n}{t^2}\right)$$

which after an application of Lemma 4 yields

$$\begin{aligned}
\ln s &\geq \left(2\varepsilon \ln\frac{1}{2\varepsilon} + \varepsilon - \varepsilon\ln\left(\frac{c+\varepsilon}{\varepsilon}\right) - \varepsilon\right)\left(\frac{n}{t}\right) \pm O\left(\frac{n}{t^2}\right) \\
&\geq \left(\varepsilon \ln\left(\frac{1}{4\varepsilon(c+\varepsilon)}\right)\right)\left(\frac{n}{t}\right) \pm O\left(\frac{n}{t^2}\right) \\
&= \Omega\left(\frac{n}{t}\right), \quad \text{by choosing } \varepsilon = \frac{1}{2e(c+\sqrt{c^2+e^{-1}})}, \text{ thus making } \frac{1}{4\varepsilon(c+\varepsilon)} = e
\end{aligned}$$

Hence, $s = \exp\left(\Omega\left(\frac{n}{t}\right)\right)$ as claimed. □

6.1 Proofs of binomial estimates

Lemma 18. *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,*

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

Proof.

$$\frac{(a+f)!}{(a-g)!} = (a+f)(a+f-1)\cdots(a-g+1)$$

$$\implies \quad a^{f+g}\left(1-\frac{g}{a}\right)^{f+g} \quad \leq \quad \frac{(a+f)!}{(a-g)!} \quad \leq \quad a^{f+g}\left(1+\frac{f}{a}\right)^{f+g}$$

$$\implies \quad (f+g)\ln\left(1-\frac{g}{a}\right) \quad \leq \quad \ln\frac{(a+f)!}{(a-g)!} - (f+g)\ln a \quad \leq \quad (f+g)\ln\left(1+\frac{f}{a}\right)$$

Using the fact that $\frac{x}{1+x} \leq \ln(1+x) \leq x$ for $x > -1$, it is easy to see that both the LHS and RHS are bounded by $O\left(\frac{(f+g)^2}{a}\right)$. □

Lemma 19. *For any constants $\alpha \geq \beta > 0$,*

$$\ln \binom{\alpha n}{\beta n} \quad = \quad aH_e\left(\frac{\beta}{\alpha}\right) n - O(\ln n)$$

Proof. By Stirling's approximation (Proposition 5),

$$\begin{aligned}
\ln \frac{(\alpha n)!}{(\beta n)!((\alpha-\beta)n)!} &= (\alpha n)\ln(\alpha n) - \alpha n - (\beta n)\ln(\beta n) + \beta n \\
&\qquad - (\alpha-\beta)n\ln((\alpha-\beta)n) + (\alpha-\beta)n - O(\ln n) \\
&= n(\alpha\ln\alpha - \beta\ln\beta - (\alpha-\beta)\ln(\alpha-\beta)) - O(\ln n) \\
&= \alpha n \cdot H_e\left(\frac{\beta}{\alpha}\right) - O(\ln n)
\end{aligned}$$

□

Claim 20. *Suppose $\ell = n^2 t$ and $k = \varepsilon\left(\frac{n}{t}\right)$ where t is an increasing function of n such that $t = o(n)$, and $\varepsilon > 0$ is a constant. Then,*

(a) $\ln \binom{n+k}{2k} = 2\varepsilon\left(\frac{n}{t}\right)\left(\ln\left(\frac{t}{2\varepsilon}\right)+1\right) \pm O\left(\frac{n}{t^2}\right)$

(b) $\ln \binom{cn/t+k}{k} = (c+\varepsilon)H_e\left(\frac{\varepsilon}{c+\varepsilon}\right)\cdot\left(\frac{n}{t}\right) - O\left(\ln n\right)$

(c) $\ln \frac{\binom{n^2+\ell-2k}{\ell}}{\binom{n^2+\ell+(t-1)k}{n^2}} = -2\varepsilon\left(\frac{n}{t}\right)\ln t - \varepsilon\left(\frac{n}{t}\right) \pm O\left(1\right)$

Proof.

(a) $\binom{n+k}{2k} = \frac{(n+k)!}{(n-k)!} \cdot \frac{1}{(2k)!}$. Since $k = o(n)$, using Lemma 18 and Lemma 19 gives

$$\begin{aligned} \ln \binom{n+k}{2k} &= 2k \ln n - (2k) \ln(2k) + 2k \pm O\left(\frac{k^2}{n}\right) \\ &= 2\varepsilon \left(\frac{n}{t}\right) \left(\ln\left(\frac{t}{2\varepsilon}\right) + 1\right) \pm O\left(\frac{n}{t^2}\right) \end{aligned}$$

(b) Follows directly from Lemma 19.

(c)

$$\frac{\binom{n^2+\ell-2k}{\ell}}{\binom{n^2+\ell+(t-1)k}{n^2}} = \frac{(n^2)!}{(n^2-2k)!} \cdot \frac{(\ell+(t-1)k)!}{\ell!} \cdot \frac{(n^2+\ell-2k)!}{(n^2+\ell+(t-1)k)!}$$

Using the fact that $tk + n = O(n)$, Lemma 18 can be applied on each of these ratios to give

$$\begin{aligned} \ln \frac{\binom{n^2+\ell-2k}{\ell}}{\binom{n^2+\ell+(t-1)k}{n^2}} &= 2\varepsilon \left(\frac{n}{t}\right) \ln(n^2) + (t-1)k \ln(\ell - (t+1)k) \ln(n^2+\ell) \pm O(1) \\ &= 2\varepsilon \left(\frac{n}{t}\right) \ln\left(\frac{n^2}{\ell}\right) + (t+1)k \ln\left(\frac{\ell}{n^2+\ell}\right) \pm O(1) \\ &= -2\varepsilon \left(\frac{n}{t}\right) \ln t - (t+1)k \ln\left(1 + \frac{1}{t}\right) \pm O(1) \\ &= -2\varepsilon \left(\frac{n}{t}\right) \ln t - \varepsilon \left(\frac{n}{t}\right) \pm O(1) \end{aligned}$$

□

7 Limitations of the measure of shifted partials

Theorem 1 shows that a lower bound of $\exp(\omega(\sqrt{n} \log n))$ on the size of any $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$-circuit computing Perm_n, will prove a $n^{\omega(1)}$ lower bound on the size of general arithmetic circuits computing Perm_n, thus proving $\mathsf{VP} \neq \mathsf{VNP}$. The measure of dimension of shifted partials was able to show a lower bound of $\exp(\Omega(\sqrt{n}))$, and we believe that the measure might yield a $\exp(\Omega(\sqrt{n} \log n))$ lower bound for the Perm_n. However, with the current upper bound for a $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit, the current technique cannot yield a bound of $\exp(\omega(\sqrt{n} \log n))$.

Proposition 21. *For any choice of $\ell, k \geq 0$, we have*

$$\frac{\min\left(\binom{N+k}{k} \cdot \binom{N+\ell}{N}, \binom{N+\ell+d-k}{N}\right)}{\binom{\sqrt{d}+k}{k} \cdot \binom{N+\ell+(\sqrt{d}-1)k}{N}} = 2^{O(\sqrt{d} \log N)}$$

Note that the first term in the numerator of the above expression are trivial upper bounds is the total number of derivatives of order k and shifts of degree ℓ. The second term in the numerator correspond to the number of monomials of degree $\ell + d - k$ over N variables. These are two trivial upper bounds for $\dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell})$ for any N-variate degree d polynomial f. Hence, in other words,

above proposition states that using the estimate Corollary 10, the best lower bound we can obtain for a N-variate degree d polynomial computed by a $\Sigma\Pi^{[O(\sqrt{d})]}\Sigma\Pi^{[\sqrt{d}]}$ circuit is $\exp(O(\sqrt{d}\log N))$.

Proof of Proposition 21. In order for $\binom{N+\ell+d-k}{N}$ to be greater than $\binom{N+\ell+(\sqrt{d}-1)k}{N}$, we must have that $(\sqrt{d}-1)k < d$ which forces $k < \sqrt{d}$.
Further, since $\binom{N+\ell}{N} \leq \binom{N+\ell+(\sqrt{d}-1)k}{N}$, the above ratio is at most $\binom{N+k}{k}/\binom{\sqrt{d}+k}{k} = 2^{O(\sqrt{d}\log N)}$. □

8 Upper bound for $\Gamma_{k,\ell}(\mathsf{Det}_n)$

The proof of Theorem 2 remains valid if we replace every occurrence of Perm_n by Det_n but there turns out to be a very interesting distinction between these two polynomials with respect to the dimension of their shifted partial derivatives. In the particular case of the determinant, Corollary 13 can be strengthened to say that the number of monomials of degree at most $(n-k+\ell)$ with an increasing sequence of length $(n-k)$ is not just a lower bound but is exactly equal to $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Det}_n)\rangle_{\leq \ell})$. This follows from the following powerful result on Gröbner bases of determinantal ideals which has been proved independently by Narasimhan [Nar86], Sturmfels[Stu90] and Caniglia, Guccione and Guccione [CGG90].

Theorem 22 ([Nar86], [Stu90], [CGG90])**.** *Let $\succ$ be the lexicographic ordering on monomials defined in Section 5. Then the set of all order $r \times r$ minors of Det_n is the reduced Gröbner basis for the ideal generated by them under the monomial ordering $\succ$.*

It is known that the set of 2×2 permanental minors *do not* form a Gröbner basis for the ideal they generate. Thus it is presumable that $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell})$ is much larger compared to the determinant. In this section we show that the lower bound on the dimension of the shifted partials of the determinant obtained in Section 5 is fairly tight, and we believe that the bound for the permanent is far from tight. In particular, we show the following,

Theorem 23. *For every large enough $n > 0$, and parameters $k, \ell \geq 0$ satisfying $k\ell < n^3$ and $k = o(n)$,*

$$\dim(\langle \boldsymbol{\partial}^{=k}\mathsf{Det}_n\rangle_{\leq \ell}) \quad \leq \quad (k+1)^2\binom{n-1}{k}^2\binom{n^2+\ell-2k}{\ell}$$

Proof. Consider an $(n-k)$-increasing sequence Q. Define $\mathcal{M}(Q)$ as the set of all monomials of degree at most $(n-k+\ell)$ having the *leading* $(n-k)$-increasing sequence as Q. From Theorem 22 we have,

$$\dim(\langle \boldsymbol{\partial}^{=k}\mathsf{Det}_n\rangle_{\leq \ell}) \quad = \quad \sum_Q |\mathcal{M}(Q)|$$

Thus, if we find an upper bound on $|\mathcal{M}(Q)|$ for each Q, then we have an upper bound on $\dim(\langle \boldsymbol{\partial}^{=k}\mathsf{Det}_n\rangle_{\leq \ell})$.

Similar to the proof of Lemma 14, we shall count the number of *forbidden variables* for the increasing sequence Q (denoted by $F(Q)$, i.e. variables that will change the leading increasing sequence.

Suppose $Q = (x_{r_1,c_1}, x_{r_2,c_2}, \ldots, x_{r_{n-k},c_{n-k}})$ (where r_i's and c_i's are increasing). Then clearly, the following variables must be forbidden:

$$\begin{aligned} F(Q) \supseteq & \left\{x_{r_1,1}, \cdots, x_{r_1,(c_1-1)}\right\} \cup \left\{x_{1,c_1}, \cdots, x_{(r_1-1),c_1}\right\} \\ & \cup \bigcup_{i=2}^{n-k} \left(\left\{x_{r_i,(c_{i-1}+1)}, \cdots, x_{r_i,(c_i-1)}\right\} \cup \left\{x_{(r_{i-1}+1),c_i}, \cdots, x_{(r_i-1),c_i}\right\}\right) \end{aligned}$$

That is, for every $x_{r_i,c_i} \in Q$, the variables to its left in the same row with column index greater than c_{i-1}, and the variables above it in the same column with row index greater than r_{i-1} must be forbidden. In other words, there are at least as many forbidden variables as there are "gaps" in the row/column indices. To formalize this, define vectors $\mathbf{a}, \mathbf{b} \in \mathbb{Z}_{\geq 0}^{n-k}$ as,

$$\begin{aligned} a_1 &= r_1 - 1 & b_1 &= c_1 - 1 \\ a_i &= r_i - r_{i-1} - 1 & b_i &= c_i - c_{i-1} - 1 \qquad (\text{for } 2 \leq i \leq n-k) \end{aligned}$$

Intuitively, $\mathbf{a}$ and $\mathbf{b}$ respectively capture the row and column gaps between the elements of Q. Thus clearly $|\mathbf{a}| = \sum_{i=1}^{n-k} a_i \leq k$ and $|\mathbf{b}| = \sum_{i=1}^{n-k} b_i \leq k$, and there is a natural bijection between the set of all non-negative solutions of these two inequalities and the set of all $(n-k)$-increasing sequences. Hence, we have

$$\begin{aligned} |F(Q)| &\geq |\mathbf{a}| + |\mathbf{b}| \\ \implies |\mathcal{M}(D)| &\leq \binom{n^2 + \ell - |\mathbf{a}| - |\mathbf{b}|}{\ell} \end{aligned}$$

Now that we have an upper bound on $|\mathcal{M}(D)|$, we can give the following upper bound,

$$\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Det}_n \rangle_{\leq \ell}) = \sum_D |\mathcal{M}(D)| \leq \sum_{\substack{\mathbf{a},\mathbf{b} \in \mathbb{Z}_{\geq 0}^{n-k} \\ |\mathbf{a}| \leq k, |\mathbf{b}| \leq k}} \binom{n^2 + \ell - |\mathbf{a}| - |\mathbf{b}|}{\ell}$$

which can be rewritten as,

$$\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Det}_n \rangle_{\leq \ell}) \leq \sum_{a=0}^{k} \sum_{b=0}^{k} \binom{a + (n-k) - 1}{a} \binom{b + (n-k) - 1}{b} \binom{n^2 + \ell - (a+b)}{\ell}$$

where a and b are used to denote $|\mathbf{a}|$ and $|\mathbf{b}|$ respectively. The terms $\binom{a+(n-k)-1}{a}$ (respectively $\binom{b+(n-k)-1}{b}$) are the number of solutions to $|\mathbf{a}| = a$ (respectively $|\mathbf{b}| = b$).

We wish to see where the summand is maximized. Let $f(a, b)$ to be the corresponding term in the above summation. Then,

$$\frac{f(a+1, b)}{f(a, b)} = \frac{a + n - k}{a + 1} \cdot \frac{n^2 - a - b}{n^2 + \ell - a - b}$$

If $n^3 > k\ell$ and $k = o(n)$, then this ratio is always greater than 1 for any choice of $a \leq k$ for large enough values of n (the most significant term in the numerator is n^3, whereas the most significant

term in the denominator is $a\ell$, which is at most $k\ell$ and less than n^3). Thus, for any value of b, the summand is maximized for $a = k$. A similar calculation for b reveals that the summand is maximized for $b = k$ for any choice of $a \leq k$. Hence, $f(a, b)$ is maximized at $a = b = k$ and we obtain

$$\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Det}_n \rangle_{\leq \ell}) \quad \leq \quad (k+1)^2 \binom{n-1}{k}^2 \binom{n^2+\ell-2k}{\ell}$$

□

Hence, the bound given in Corollary 16 for the determinant is not-too-far from the actual value. We believe that dimension of shifted partial derivatives of the permanent is significantly larger, and we conclude by stating this as a conjecture.

Conjecture 24. *Let $\mathbb{F}$ be any field of characteristic zero. There exists suitable choices for the parameters $\ell = \ell(n)$ and $k = k(n)$ such that over $\mathbb{F}$ we have*

$$\frac{\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n) \rangle_{\leq \ell})}{\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Det}_n) \rangle_{\leq \ell})} = n^{\omega(1)}.$$

Acknowledgments. We would like to thank Ravi Kannan and Satya Lokam for useful discussions and providing some relevant references. We also want to thank Avi Wigderson for many helpful remarks on an earlier draft. In particular, Avi pointed out how our proof technique is also applicable to some extent in the setting of bounded depth multilinear circuits and gave an elegant combinatorial argument to the effect that the lower bound estimate of Section 5 remains essentially valid under a random restriction.

References

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

[CGG90] L. Caniglia, J. A. Guccione, and J. J. Guccione. Ideals of generic minors. *Commutative Algebra*, 18:2633–2640, 1990.

[CLO07] D.A. Cox, J.B. Little, and D. O'Shea. *Ideals, Varieties and Algorithms*. Undergraduate texts in mathematics. Springer, 2007.

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

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

[GKQ12] Ankit Gupta, Neeraj Kayal, and Youming Qiao. Random arithmetic formulas can be reconstructed efficiently. Technical report, Electronic Colloquium on Computational Complexity (ECCC), 2012.

[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.

[Kay12a] Neeraj Kayal. Affine projections of polynomials. In *STOC*, pages 643–662, 2012.

[Kay12b] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19, 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 homogenous circuits with bounded top fanin. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:68, 2013.

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

[Nar86] H. Narasimhan. The irreducibility of ladder determinantal varieties. *Journal of Algebra*, 102:162–185, 1986.

[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.

[Rom] Dan Romik. Stirlings approximation for n!: The ultimate short proof? *The American Mathematical Monthly*, 107(6):556557.

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

[Stu90] Bernd Sturmfels. Gröbner bases and stanley decompositions of determinantal rings. *Mathematische Zeitschrift*, 209:137–144, 1990.

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

[Tav13] Sébastien Tavenas. Improved bounds for reduction to depth 4 and 3. *CoRR*, abs/1304.5777, 2013.

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

ECCC ISSN 1433-8092

http://eccc.hpi-web.de