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

An exponential lower bound for homogeneous depth four arithmetic circuits with bounded bottom fanin

Ankit Gupta
Microsoft Research India
t-ankitg@microsoft.com

Pritish Kamath
Microsoft Research India
t-pritk@microsoft.com

Neeraj Kayal
Microsoft Research India
neeraka@microsoft.com

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

August 3, 2012

Abstract

Agrawal and Vinay [AV08] have recently shown that an exponential lower bound for depth four homogeneous circuits with bottom layer of × gates having sublinear fanin translates to an exponential lower bound for a general arithmetic circuit computing the permanent. Motivated by this, we examine the complexity of computing the permanent and determinant via homogeneous depth four circuits with bounded bottom fanin.

We show here that any homogeneous depth four arithmetic circuit with bounded bottom fanin computing the permanent (or the determinant) must be of exponential size.

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.

Recall that the permanent is an n^2-variate homogeneous[1]polynomial of degree n defined as:

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

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

ISSN 1433-8092

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 circuit for the permanent is actually a depth three homogeneous circuit of size $O(n^2 \cdot 2^n)$ and is called Ryser's formula. Its illustrious sibling, the determinant, is widely believed to be comparatively easy, being complete for a subclass of VP (an algebraic analogue of P, also defined in [Val79]). It is conjectured (cf. [AV08]) that any arithmetic circuit computing the permanent must be of exponential size. Meanwhile, the arithmetic complexity of computing the determinant equals $\widetilde{O}(n^\omega)$, 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.

The model. Let $t \geq 1$ be any integer. In this work, we focus our attention on depth four homogeneous[2] arithmetic circuits with bottom fanin at most t which we denote by $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(t)$. A $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(t)$ circuit computes a polynomial of the form

$$C \quad = \quad \sum_{i=1}^{s} (Q_{i1} \cdot Q_{i2} \cdot \ldots \cdot Q_{id}) \quad \text{where each } Q_{ij} \text{ is homogeneous and } \deg(Q_{ij}) \leq t. \tag{1}$$

Our motivation for investigating representations of the form (1) stems from a recent result of Agrawal and Vinay [AV08]. They showed[3] that if Perm_n can be computed by arithmetic circuits of size $2^{o(n)}$ then Perm_n can in fact be computed by $2^{o(n)}$-sized $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(t)$ circuits with $t = o(n)$. Recently Koiran [Koi12], building upon [AV08], showed that if Perm_n has circuits of size $\mathsf{poly}(n)$ then it also has $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(o(n))$ circuits of size $2^{O(\sqrt{n}\log^2 n)}$. This means that as far as the problem of proving exponential (or even superpolynomial) arithmetic lower bounds for the permanent is concerned, one can assume without any loss of generality that the circuit is in fact a $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(o(n))$ circuit. We show here that any $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(t)$ circuit[4] computing the permanent (or the determinant) must be of size at least $2^{\Omega(n/(2^t))}$. In particular, if t is bounded, then we obtain a $2^{\Omega(n)}$ lower bound in this model.

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

[2] An arithmetic circuit is said to be *homogeneous* if the polynomial computed at every internal node of the circuit 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 $O(d^2 \cdot s)$.

[3] Agrawal and Vinay [AV08] do not explicitly mention that the resulting depth four circuit is homogeneous but this can be easily inferred from their proof.

[4] Our lower bound remains valid even if the homogeneity condition is relaxed to allow nonhomogeneous depth four circuits wherein the degree of any intermediate polynomial is at most $O(n)$. See the statement of Theorem 1 and the following remark for details.

Karpinski [GK98] 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 superpolynomial size. Following this, Raz and Yehudayoff [RY08] proved exponential lower bounds for constant depth multilinear circuits. To put our result in this larger context, we remark here that $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(1)$ is exactly the class of homogeneous depth three circuits. Thus for $t \geq 2$ the class $\Sigma\Pi\Sigma\Pi^{[\mathrm{hom}]}(t)$ that we consider here is a superset of homogeneous depth three circuits but of course a subclass of depth four circuits. We are now ready to give a formal statement of our main result.

Theorem 1. *Let $\mathbb{F}$ be any field. Let C be a circuit of the form $C = \sum_{i=1}^{s} Q_{i1} \cdots Q_{id}$ where each $Q_{ij} \in \mathbb{F}[\mathbf{x}]$ is a polynomial of degree at most t. Suppose further that $d = O(n)$. If C computes the polynomial Perm_n, then $s \geq 2^{\Omega(n/2^t)}$.*

Remark: In Theorem 1 we can allow d to be slightly superlinear at the expense of a slightly worse lower bound. Specifically, if for some constant $\delta > 0$ we have $d \leq n^{2-\delta}$, then the number of summands s must be at least $2^{\Omega(n^\delta/2^t)}$ (cf. Theorem 15). Our proof is completely elementary and self-contained but it is possible that using some more sophisticated theorems from algebraic geometry, the bounds that we obtain can be improved. We discuss this in Section 7.

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 $(x_1^{i_1} \cdot x_2^{i_2} \cdot \ldots \cdot x_n^{i_n})$ which has degree $|\mathbf{i}| \stackrel{\text{def}}{=} (i_1 + i_2 + \ldots + i_n)$. $\partial^\mathbf{i} f$ denotes the partial derivative of f with respect to the monomial $\mathbf{x}^\mathbf{i}$,

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

Recall that for a subset of polynomials $S \subseteq \mathbb{F}[\mathbf{x}]$, the $\mathbb{F}$-span of S, denoted $\mathbb{F}\text{-span}(S)$, is the set of all possible $\mathbb{F}$-linear combinations of polynomials in S. i.e.

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

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*

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

$\langle \boldsymbol{\partial}^{=k} f\rangle_{\leq \ell}$ 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 $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(t)$ circuit (Corollary 7). Meanwhile $\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Perm}_n\rangle_{\leq \ell})$ is relatively large (Lemma 14). This gives the lower bound.

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 $\langle \boldsymbol{\partial}^{=k} C\rangle_{\leq \ell}$ for C being a polynomial computed by a $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(t)$ circuit, i.e. when C is of the form given in equation (1). 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 7, 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,

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

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*

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

Claim 2. *For any $0 < x < 1$,*

$$x \cdot \log_2 \frac{1}{x} \quad \leq \quad H_2(x) \quad \leq \quad x \cdot \log_2 \frac{1}{x} \ + \ 2x$$

Proposition 3 (Stirling's Formula, cf. [Rom])**.** $\ln(n!) \quad = \quad n \ln n - n + O(\ln n)$

Using Stirling's Formula, it is straightforward to derive the following asymptotic estimates which we would need in our proof.

Claim 4. *For any $a, b, n \geq 0$,*

1. $\log_2 \binom{an}{bn} \quad = \quad an \cdot H_2\left(\frac{b}{a}\right) + O(\log_2(abn))$
2. *If $a, b, \alpha, \beta = O(1)$, then*

$$\log_2 \binom{an^2 + bn}{\alpha n^2 + \beta n} \quad = \quad an^2 \cdot H_2\left(\frac{\alpha}{a}\right) + bn \log_2 a - \beta n \log_2 \alpha \\ -(b-\beta) n \log_2(a-\alpha) + O(\log_2 n)$$

In the particular case when $a = 2\alpha$, the above expression simplifies to

$$\log_2 \binom{2\alpha n^2 + bn}{\alpha n^2 + \beta n} \quad = \quad 2\alpha n^2 + bn + O(\log_2 n)$$

4 Upper bounding the dimension of shifted partials of depth four circuits.

In this section we give an upper bound on $\dim(\langle \boldsymbol{\partial}^{=k} C \rangle_{\leq \ell})$ when C is computed by a depth four circuit, i.e. C is of the form given in equation (1). We begin by noting that $\dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell})$ is sub-additive.

Proposition 5. *For all $k, \ell \geq 0$, we have $\dim(\langle \boldsymbol{\partial}^{=k}(f+g) \rangle_{\leq \ell}) \quad \leq \quad \dim(\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell}) + \dim(\langle \boldsymbol{\partial}^{=k} g \rangle_{\leq \ell})$.*

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,

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

The proposition follows. □

Let C be a depth-4 circuit computing a polynomial of the form

$$C \quad = \quad \sum_{i=1}^{s} Q_{i1} \cdot Q_2 \cdot \ldots \cdot Q_{id} \quad \text{where } \deg(Q_{ij}) \leq t.$$

By Proposition 5, it suffices to understand the growth of $\dim(\langle \boldsymbol{\partial}^{=k} C\rangle_{\leq \ell})$ of a single term $(Q_1 \cdot Q_2 \ldots \cdot Q_d)$.

Proposition 6. *If $f = Q_1 \cdots Q_d$ where each $Q_i \in \mathbb{F}[\mathbf{x}_n]$ is a polynomial of degree bounded by t. Then, for any $k \leq d$ and $\ell \geq 0$,*

$$\dim(\langle \boldsymbol{\partial}^{=k} f\rangle_{\leq \ell}) \quad \leq \quad \binom{d}{k}\binom{n+(t-1)k+\ell}{n}$$

Proof. Let $\mathbf{j} \in \mathbb{Z}^n_{\geq 0}$ be any n-tuple satisfying $|\mathbf{j}| = k$. We observe by repeated use of the product rule that

$$\partial^{\mathbf{j}}\left(Q_1 \ldots Q_d\right) \quad = \quad \sum_{\mathbf{j}_1 + \cdots + \mathbf{j}_d = \mathbf{j}} \left(\partial^{\mathbf{j}_1} Q_1\right) \ldots \left(\partial^{\mathbf{j}_d} Q_d\right)$$

Since $|\mathbf{j}| = k$, at most k of the $\mathbf{j}_i$'s are non-zero. Hence, each term in the above sum can be written as $\left(\prod_{i \in P} Q_i\right) \cdot Q'$ for some set $P \subset [d]$ of size $(d-k)$ and Q' having degree at most $(tk - k)$. Thus, every element of $\mathbf{x}^{\leq \ell}\boldsymbol{\partial}^{=k}(Q_1 \cdots Q_d)$ can be written as a linear combination of $\left(\prod_{i\in P} Q_i\right)\mathbf{x}^{\mathbf{i}}$ where $|P| = (d-k)$ and $\mathbf{x}^{\mathbf{i}}$ is a monomial of degree at most $\ell + (t-1)k$. The total number of monomials of degree at most $\ell + (t-1)k$ over n variables is $\binom{n+(t-1)k+\ell}{n}$, and the total number of choices for P is $\binom{d}{k}$. Hence we obtain,

$$\dim(\langle \boldsymbol{\partial}^{=k}(Q_1 \cdots Q_d)\rangle_{\leq \ell}) \quad \leq \quad \binom{d}{k}\binom{n+(t-1)k+\ell}{n}$$

□

The following corollary follows directly from the above observation via sub-additivity.

Corollary 7. *If $C = \sum_{i=1}^{s}\prod_{j=1}^{d} Q_{ij}$ where each $Q_{ij} \in \mathbb{F}[\mathbf{x}_n]$ is a polynomial of degree bounded by t, then for any $k \leq d$*

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

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 monomial ordering[5]. 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 8. *Let $S \subseteq \mathbb{F}[\mathbf{x}]$ be any set of polynomials. Then*

$$\dim(\mathbb{F}\text{-}span(S)) = \#\{\mathsf{LM}(f) \ : \ f \in \mathbb{F}\text{-}span(S)\}.$$

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

Corollary 9. *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's fix $\succ$ to be the lexicographic monomial ordering induced by the following ordering on the variables: $x_{11} \succ \cdots \succ x_{1n} \succ x_{21} \succ \cdots \succ x_{nn}$. Now 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

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

Now 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.

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

[5] 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]

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 increasing sequence of length $(n-k)$. Consequently for any P-minor M of size $(n-k)$ we have that $\mathbf{x}^{\mathbf{i}}\cdot\mathsf{LM}(M)$ contains an increasing sequence of length $(n-k)$. Conversely, every monomial of degree at most $(n-k+\ell)$ that contain an increasing sequence of length $(n-k)$ 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 a $(n-k)\times(n-k)$ P-minor M. Hence we have:

Corollary 10. *$\dim(\langle\boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq\ell})$ is lower bounded by the number of distinct monomials of degree at most $(n-k+\ell)$ over n^2 variables that contain an increasing sequence of length $(n-k)$.*

In order to count the number of monomials of degree bounded by $n-k+\ell$ that contain an 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'll focus on the variables $X' = \{x_{ii} \ : \ 1\leq i\leq n\}\cup\{x_{i(i+1)} \ : \ 1\leq i\leq n-1\}$. Let $S_2(n,r)$ denote the number of subsets of X' that contain an increasing sequence of length r.

Lemma 11. *Let $S_2(n,r)$ be the number of subsets of X', that contain an increasing sequence of length at least r. Then $S_2(n,r)$ satisfies the following recurrence.*

$$S_2(n,r) \quad = \quad 2S_2(n-1,r-1) \quad + \quad \sum_{i=1}^{r-1} 2^{i-1}S_2(n-i,r-i+1) \quad + \quad 2^{r-1}(2^{2(n-r)}-1)$$

Proof. Either x_{11} is present in the set or it is not. If it is present in the set, then any subset of the last $(n-1)$ rows and columns with an increasing sequence of length $r-1$ can be extended by x_{11}. Further we are free to add x_{12} to the set or not as it does not change the presence of an increasing sequence. Therefore,

$$\begin{array}{rcll} S_2(n,r) & = & S_2(n-1,r-1) & (x_{11} \text{ and } x_{12} \text{ present}) \\ & & + \ \ S_2(n-1,r-1) & (x_{11} \text{ present but not } x_{12}) \\ & & + \ \ S_2'(n,r) & (x_{11} \text{ not present}) \end{array}$$

where $S_2'(n,r)$ denotes the number of subsets that contain an increasing sequence of length r and *do not* include x_{11}. Thus it suffices to show that

$$S_2'(n,r) \quad = \quad \sum_{i=1}^{r-1} 2^{i-1}S_2(n-i,r-i+1) \quad + \quad 2^{r-1}(2^{2(n-r)}-1)$$

and this shall be proved by induction on r. The base case of $r = 1$ is the number of non-empty subsets of the $2n-2$ variables (x_{11} excluded) and hence is $2^{2(n-1)} - 1$. For the inductive step, a similar argument as earlier gives

$$\begin{aligned} S_2'(n,r) &= S_2'(n-1,r-1) && (x_{12} \text{ and } x_{22} \text{ are present}) \\ &\quad + \; S_2'(n-1,r-1) && (x_{12} \text{ is present but not } x_{22}) \\ &\quad + \; S_2(n-1,r) && (x_{12} \text{ is not present}) \\ &= 2S_2'(n-1,r-1) \quad + \quad S_2(n-1,r) \end{aligned}$$

Using the inductive hypothesis,

$$\begin{aligned} S_2'(n,r) &= 2 \cdot \left(\sum_{i=1}^{r-2} 2^{i-1} S(n-1-i, r-i) + 2^{r-2}(2^{2(n-r)} - 1) \right) \\ &\quad + \; S_2(n-1,r) \\ &= \sum_{i=1}^{r-1} 2^{i-1} S_2(n-i, r-i+1) \quad + \quad 2^{r-1}(2^{2(n-r)} - 1) \end{aligned}$$

Combining with the earlier recurrence for $S(n,r)$, we have our desired recurrence. □

Lower bounding $S_2(n,r)$

Define a new function $g(d,r) \overset{\text{def}}{=} \frac{1}{2^r} S_2(d+r,r)$. Hence, using the recurrence for $S_2(n,r)$, we have that $g(d,r)$ satisfies

$$\begin{aligned} 2^r g(d,r) &= 2 \cdot 2^{r-1} g(d,r-1) + \sum_{i=1}^{r-1} 2^{i-1} \cdot 2^{r-i+1} g(d-1, r-i+1) + 2^{r-1}(2^{2d} - 1) \\ \implies \quad g(d,r) &= g(d,r-1) + \sum_{i=1}^{r-1} g(d-1, r-i+1) + \frac{2^{2d} - 1}{2} \end{aligned}$$

Lemma 12. $g(d,r) \quad \geq \quad \frac{1}{2}\binom{r+2d-2}{2d}$

Proof. Since the above recurrence for g consists of only non-negative terms, the contribution of $g(0,2)$ in the unfolding of the recursion is a lower bound for $g(d,r)$.

For this, we shall view the *computation graph* of g by the natural dynamic program on a two dimensional grid where (x,y) would hold the value of $g(x,y)$. Each point (x,y) has edges coming in from $(x, y-1)$ and $(x-1, 2), \cdots, (x-1, y)$, as the value of $g(x,y)$ depends on these values in the recurrence. Hence, the number of times $g(0,2)$ is added in the computation of $g(d,r)$ is precisely equal to the number of paths between $(0,2)$ and (d,r) in the graph.

For every sequence $2 \leq a_0' \leq a_1 \leq a_1' \leq \cdots \leq a_{d-1} \leq a_{d-1}' \leq a_d \leq r$ we can define unique path $(0,2) \rightsquigarrow (0, a_0') \to (1, a_1) \rightsquigarrow (1, a_1') \to \cdots \rightsquigarrow (d-1, a_{d-1}') \to (d, a_d) \rightsquigarrow (d, r)$, and every path corresponds to such a sequence. Hence, the number of paths between $(0,2)$ to (d,r) is precisely the number of such sequences which is $\binom{r+2d-2}{2d}$. Hence, $g(d,r) \geq \binom{r+2d-2}{2d} g(0,2) = \frac{1}{2} \cdot \binom{r+2d-2}{2d}$. □

From the definition of g, we get the following lower bound for $S_2(n,r)$.

Corollary 13. $S_2(n,r) \quad \geq \quad 2^{r-1} \cdot \binom{2n-r-2}{r-2}$ $\square$

To get a lower bound for $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell})$, we'll just pick an element contributing to $S_2(n,r)$ and fill up the remaining degree using the other variables.

Lemma 14. *There exists $m \leq 2n$ such that*

$$\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell}) \quad \geq \quad \frac{1}{2n} S_2(n,n-k) \cdot \binom{n^2-n+\ell-k+1}{n^2-2n+m+1}$$

Proof. Let $\tilde{S_2}(n,r;m)$ denote the number of subsets of X' *of size* m that contain an increasing sequence of length r. Since $S_2(n,r) = \sum_m \tilde{S_2}(n,r;m)$, there must exist some $m \leq 2n$ such that $\tilde{S_2}(n,r;m) \geq \frac{1}{2n} S_2(n,r)$.
Any set S of size m contributing to $\tilde{S_2}(n,n-k;m)$ can be thought of as a monomial of degree m over variables in its support. The remaining degree of at most $\ell + n - k - m$, can be filled using $n^2 - 2n + 1 + m$ variables (those outside the two diagonals, and the support of S) can be filled up in $\binom{n^2-n+\ell-k+1}{n^2-2n+m+1}$ ways. $\square$

6 Putting it all together

We are now ready to prove the main theorem, which is a stronger form of Theorem 1.

Theorem 15. *Let C be a circuit of the form $C = \sum_{i=1}^{s} P_{i1} \cdots P_{id}$ where each P_{ij} is a polynomial of degree bounded by t and $d \leq n^{2-\delta}$ for some constant $\delta > 0$. If C computes the polynomial Perm_n, then $s \geq \exp\left(\frac{n^\delta}{2^t}\right)$.*

Proof. Say $d = \alpha n$ for some $\alpha \leq n^{1-\delta}$. Let $\ell = n^2$ and $k = \varepsilon n$ (for an $\varepsilon > 0$ that shall be chosen shortly). Then, by Corollary 7, we have that

$$\begin{aligned}\dim(\langle \boldsymbol{\partial}^{=k}(C)\rangle_{\leq \ell}) \quad &\leq \quad s \cdot \binom{d}{k}\binom{n^2+\ell+(t-1)k}{n^2} \\ &\leq \quad s \cdot \binom{\alpha n}{\varepsilon n}\binom{2n^2+(t-1)\varepsilon n}{n^2}\end{aligned}$$

Using the estimates from Claim 4,

$$\log_2(\dim(\langle \boldsymbol{\partial}^{=k}(C)\rangle_{\leq \ell})) \quad \leq \quad \log_2 s + 2n^2 + \left(\alpha H_2\left(\frac{\varepsilon}{\alpha}\right) + (t-1)\varepsilon\right) n + O(\log n) \qquad (2)$$

As for Perm_n, we get from Corollary 13 and Lemma 14 that

$$\begin{aligned}\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell}) \quad &\geq \quad \frac{1}{2n} \cdot 2^{n-k-1}\binom{n+k-2}{2k} \cdot \binom{n^2-n+\ell-k+1}{n^2-2n+m+1} \\ &= \quad \frac{1}{4n} \cdot 2^{(1-\varepsilon)n}\binom{(1+\varepsilon)n-2}{2\varepsilon n} \cdot \binom{2n^2-(1+\varepsilon)n+1}{n^2-2n+m+1}\end{aligned}$$

Using the estimates of Claim 4 again,

$$\log_2(\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell})) \quad \geq \quad 2n^2 + \left((1+\varepsilon)H_2\left(\frac{2\varepsilon}{1+\varepsilon}\right) - 2\varepsilon\right) n - O(\log n) \tag{3}$$

From Equation (2) and (3), we get

$$\begin{aligned} \log_2 s &\geq \left((1+\varepsilon)H_2\left(\frac{2\varepsilon}{1+\varepsilon}\right) - (t+1)\varepsilon - \alpha H_2\left(\frac{\varepsilon}{\alpha}\right)\right) n - O(\log n) \\ &= f_{\alpha,t}(\varepsilon)\cdot n - O(\log n) \quad \text{where } f_{\alpha,t}(\varepsilon) \text{ denotes the coefficient of } n. \end{aligned}$$

Using Claim 2 to bound the entropy terms, we get

$$\begin{aligned} f_{\alpha,t}(\varepsilon) &\geq 2\varepsilon\cdot\log_2\left(\frac{1+\varepsilon}{2\varepsilon}\right) - (t+1)\varepsilon - \varepsilon\log_2\left(\frac{\alpha}{\varepsilon}\right) - 2\varepsilon \\ &\geq \varepsilon\log_2\left(\frac{1}{\varepsilon}\right) - \varepsilon\log_2\left(2^{t+5}\alpha\right) \\ &= \varepsilon\cdot\log_2\left(\frac{1}{\varepsilon\cdot 2^{t+5}\alpha}\right) \quad = \quad \frac{1}{2^{t+6}\alpha} \quad \text{if } \varepsilon = \frac{1}{2^{t+6}\alpha} \\ \Longrightarrow \quad \log_2 s &\geq O\left(\frac{n^\delta}{2^t}\right) \end{aligned}$$

which yields the claimed lower bound for Perm_n. □

7 Discussion

The proof of Theorem 1 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 10 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 by Sturmfels[Stu88].

Theorem 16 ([Stu88])**.** *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. We conclude with the following conjecture.

Conjecture 17. *There exists choices for $\ell, k \geq 0$ such that $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n)\rangle_{\leq \ell})$ is super-polynomially larger (in n) than $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Det}_n)\rangle_{\leq \ell})$.*

Acknowledgments. We would like to thank Satya Lokam and Ravi Kannan for useful discussions and providing some relevant references.

References

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

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

[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.

[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. 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.

[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.

[Stu88] Bernd Sturmfels. Gröbner bases of determinantal ideals. Technical report, Johannes Kepler Universit"at Linz, 1988.

[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.

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