Electronic Colloquium on Computational Complexity, Revision 2 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

March 20, 2013

Abstract

Agrawal-Vinay [AV08] and Koiran [Koi12] have recently shown that an $\exp(\omega(\sqrt{n}\log^2 n))$ 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.
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 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 analogue of

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

ISSN 1433-8092

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 superpolynomial 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 homogeneous[2] arithmetic circuits with bottom fanin bounded by a parameter t which we denote by $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(t)$. A $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(t)$ circuit computes a polynomial of the form

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

where each Q_{ij} is homogeneous polynomial of degree bounded by t, and every summand has the same degree. 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].

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

The contrapositive of the above statement is that it suffices to show a $2^{\omega(\sqrt{n}\log^2 n)}$ lower bound for $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(\sqrt{n})$ circuits computing the Perm_n to prove a super-polynomial circuit lower bound. Thus, a good enough lower bound for $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(\sqrt{n})$ 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* $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(t)$ *circuit that computes* Perm_n *(or* Det_n*) must have top fanin* s *at least* $\exp\left(\Omega\left(\frac{n}{t}\right)\right)$*. In particular, any* $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(\sqrt{n})$ *circuit that computes* Perm_n *(or* Det_n*) must have size at least* $\exp\left(\Omega\left(\sqrt{n}\right)\right)$*.*

[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 $O(d^2 \cdot s)$.

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[3], 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 arbitarily 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 lowerbound can be improved for Perm_n by improving the estimate of the dimension of a certain explicit vector space[4] that comes up in our proof. We discuss this in section 8 and make a specific conjecture in this regard. Towards the end, we also show how the above lower bound implies similar lower bounds for a certain structured subclass of formulas of larger depth which we call regular formulas. 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 $(x_1^{i_1} \cdot x_2^{i_2} \cdot \ldots \cdot x_n^{i_n})$ 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 $\mathbf{x}^{\mathbf{i}}$,

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

For a finite 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.

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

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

[3] See also remark 11 on how to (easily) handle the case where some of the Q_{ij}'s have degree strictly smaller than t.

[4] 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.

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*

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

$\langle \boldsymbol{\partial}^{=k} f \rangle_{\leq \ell}$ 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 10). Meanwhile $\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Perm}_n \rangle_{\leq \ell})$ is relatively large (Corollary 17). 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 $\langle \boldsymbol{\partial}^{=k} C \rangle_{\leq \ell}$ 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. As a corollary of Theorem 2, we present a lower bound on the size of *regular formulae* computing the Permanent (or the Determinant) in Section 7. 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 $\alpha \geq \beta > 0$,*

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

4 Upper bounding the dimension of shifted partials of $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(t)$ 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

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

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

Proposition 8. Subadditivity. *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

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

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

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

which can be stated succinctly as

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

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

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

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*

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

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

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

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

Remark 11. In the proof of Theorem 2, we would need a bound on m as well to achieve a good upper bound for the dimension of shifted partial derivatives of a $\Sigma\Pi\Sigma\Pi^{[\text{hom}]}(t)$ circuit. However, off hand we do not have any upper bound on m since several factors of a single term could have very low degree (much less than t). Fortunately, our proof would give a lower bound on the top fanin, irrespective of the sparsities of the Q_{ij}'s. Hence we can multiply the Q_{ij}'s of low degree to ensure that each of them (except perhaps one) has degree at least $t/2$ and at most t. By this, assuming $\sum_j \deg(Q_{ij}) = n$, without loss of generality, we can claim that $m \leq 2n/t + 1$.

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

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

[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]

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

Corollary 13. *For any polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ we have*

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

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: $x_{11} \succ \cdots \succ x_{1n} \succ x_{21} \succ \cdots \succ x_{nn}$. 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.

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 14. $\dim(\langle \boldsymbol{\partial}^{=k}(\mathsf{Perm}_n) \rangle_{\leq \ell})$ *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 $D_{2,n} = \{x_{ii} \ : \ 1 \le i \le n\} \cup \{x_{i(i+1)} \ : \ 1 \le i \le n-1\}$.

Lemma 15. *Fix parameters $n, r \ge 0$, 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.[6] Consider any particular r-increasing sequence, call it Q, in $D_{2,n}$. We show that the total number of monomials (of degree $\le d$) with Q as the leading r-increasing sequence is atleast $\binom{n^2-2n+d+r}{d-r}$, 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.
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 $\le d$ monomial without changing the increasing sequence. The total number of variables to use is $n^2-(2n-1)+r+(r-1)$, 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 $\le d$ 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 16. *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 15 and using Lemma 16 with these parameters, we get the following lowerbound for $\dim(\langle \boldsymbol{\partial}^{=k} \mathsf{Perm}_n \rangle_{\le \ell})$ via Corollary 14.

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

Corollary 17. *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 it should be recalled 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 17 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 subsection 6.1.

Claim 18. *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 19. *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 20. *For any constants $\alpha \geq \beta > 0$,*

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

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

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

Proof.

(a) $\binom{n+k}{2k} = \frac{(n+k)!}{(n-k)!} \cdot \frac{1}{(2k)!}$. Since $k = o(n)$, using Lemma 19 and Lemma 20 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 20.

(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 19 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 Lower bound on size of regular formulas

Similar lower bounds can also be obtained for certain structured subclasses of formulas of larger depth that we call *regular formulas*. Recall the notion of *formal degree* of a node in a circuit. The formal degree of a leaf node is 1 and the formal degree of internal nodes is defined inductively in the natural manner. For a $+$ gate, the formal degree is the maximum of the formal degrees of its children while for a $\times$ gate the formal degree is the sum of the formal degrees of its children. The formal degree of a circuit is the formal degree of its output node.

Definition 4 (Regular Formula)**.** *We will say that a formula ϕ is a* regular formula *if the following hold,*

(a) the formula consists of alternating layers of $+$ and $\times$ gates

(b) the formal degree of all nodes in any fixed layer is the same

The product depth *of a formula shall be defined to be the number of multiplication layers.*

Remark: If a regular formula computing Perm_n or Det_n is also homogeneous then its formal degree is n. The best formulas that we know of for computing Perm_n or Det_n are regular and homogeneous. We define the size of a formula ϕ as simply the number of leaves in ϕ (note that the number of leaf nodes in ϕ is within a constant factor of the total number of nodes in ϕ). We obtain the following as a simple corollary of Theorem 2.

Theorem 22. *Any regular formula of formal degree n and product depth Δ computing Perm_n (or Det_n) must have size $\exp(\Omega(n^{1/\Delta}))$*

We would need a slightly stronger version of Theorem 2, which we state here without proof. It can nevertheless be proved in a similar way.

Theorem 23. *There exist $c \in \mathbb{R}$ and $n_0 \in \mathbb{N}$ such that for any $n > n_0$ and $t \in \mathbb{N}$, any $\Sigma\Pi\Sigma\Pi(t)$ formula (of formal degree n) computing Perm_n (or Det_n) must have size at least $\exp(cn/t)$.*

We will also need the following upper bound on the number of monomials that may be generated by a formula of size s and formal degree d.

Proposition 24. *Let ϕ be a formula of size s and formal degree d. Then the number of monomials in its output can at most be s^d.*

The proof is via an easy induction. We now proceed to prove the lower bound on the size of any regular formula computing the Permanent (or the Determinant).

Proof of theorem 22. Let C be any regular formula of size s and formal degree n (with $n > n_0$, where n_0 is the one obtained in Theorem 23) computing Perm_n. Let the product depth of C be Δ. Observe that since the formal degrees of all nodes at a layer in C is the same, all $\times$ gates in any fixed layer must have the same fanin. Suppose that for the layer closest to the output, each $\times$ gate has fanin p_1, for the next layer of $\times$ gates, each gate has fanin p_2 and so on. Thus the formal degree of C is $n = p_1 \cdot p_2 \cdot \ldots \cdot p_\Delta$. Let p_t be the highest fan-in among all the p_is. Then we have $p_t \geq n^{1/\Delta}$. We now obtain a depth-4 formula (of formal degree n) by collapsing the part of the formula above (and including) the t-th layer of product gates into a depth-2 formula, and collapsing the remaining

lower part into another depth-2 formula, and putting them together to form one depth-4 formula $\hat{C}$. The bottom fan-in of the product gates in $\hat{C}$ is $b = n/(p_1 \dots p_t)$. Applying proposition 24 we get that the top fan-in $\hat{s}$ of $\hat{C}$ is at most $\hat{s} \leq s^{p_1 \cdot p_2 \cdot \dots \cdot p_{t-1}}$.
Applying Theorem 23 to our depth four circuit $\hat{C}$, we get that the top fan-in $\hat{s} \geq \exp(cn/b))$. Combining these two inequalities we have

$$\begin{aligned} s^{p_1 \cdot p_2 \cdot \dots \cdot p_{t-1}} &\geq 2^{\Omega(n/b)} \\ &= 2^{\Omega(p_1 \cdot p_2 \cdot \dots \cdot p_t)} \end{aligned}$$

so that

$$\begin{aligned} s &\geq 2^{\Omega(p_t)} \\ &\geq 2^{\Omega(n^{1/\Delta})}, \quad \text{as required.} \end{aligned}$$

□

8 Discussion

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 14 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 25 ([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. We conclude with the following conjecture.

Conjecture 26. *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.

[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. 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.

[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.

[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.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de