Depth-4 Lower Bounds, Determinantal Complexity: A Unified Approach

Suryajith Chillara and Partha Mukhopadhyay

Chennai Mathematical Institute, Siruseri, India
`{suryajith, partham}@cmi.ac.in`

Abstract

Tavenas has recently proved that any $n^{O(1)}$-variate and degree n polynomial in VP can be computed by a depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit of size $2^{O(\sqrt{n}\log n)}$ [14]. So, to prove VP $\neq$ VNP it is sufficient to show that an explicit polynomial in VNP of degree n requires $2^{\omega(\sqrt{n}\log n)}$ size depth-4 circuits. Soon after Tavenas' result, for two different explicit polynomials, depth-4 circuit size lower bounds of $2^{\Omega(\sqrt{n}\log n)}$ have been proved (see [7] and [4]). In particular, using combinatorial design Kayal et al. [7] construct an explicit polynomial in VNP that requires depth-4 circuits of size $2^{\Omega(\sqrt{n}\log n)}$ and Fournier et al. [4] show that the iterated matrix multiplication polynomial (which is in VP) also requires $2^{\Omega(\sqrt{n}\log n)}$ size depth-4 circuits.

In this paper, we identify a simple combinatorial property such that any polynomial f that satisfies this property would achieve a similar depth-4 circuit size lower bound. In particular, it does not matter whether f is in VP or in VNP. As a result, we get a simple unified lower bound analysis for the above mentioned polynomials.

Another goal of this paper is to compare our current knowledge of the depth-4 circuit size lower bounds and the determinantal complexity lower bounds. Currently the best known determinantal complexity lower bound is $\Omega(n^2)$ for Permanent of a $n \times n$ matrix (which is a n^2-variate and degree n polynomial) [3]. We prove that the determinantal complexity of the iterated matrix multiplication polynomial is $\Omega(dn)$ where d is the number of matrices and n is the dimension of the matrices. So for $d = n$, we get that the iterated matrix multiplication polynomial achieves the current best known lower bounds in both fronts: depth-4 circuit size and determinantal complexity. Our result also settles the determinantal complexity of the iterated matrix multiplication polynomial to $\Theta(dn)$.

To the best of our knowledge, a $\Theta(n)$ bound for the determinantal complexity for the iterated matrix multiplication polynomial was known only for any constant $d > 1$ [6].

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes

Keywords and phrases Arithmetic Circuits, Determinantal Complexity, Depth-4 Lower Bounds

Digital Object Identifier 10.4230/LIPIcs.STACS.2014.239

1 Introduction

One of the main challenges in algebraic complexity theory is to separate VP from VNP. This problem is well known as Valiant's hypothesis [15]. This is an algebraic analog of the problem P vs NP. Recall that a multivariate polynomial family $\{f_n(\mathrm{X}) \in \mathbb{F}[x_1, x_2, \ldots, x_n] : n \geq 1\}$ is in the class VP if f_n has degree of at most poly(n) and can be computed by an arithmetic circuit of size poly(n). It is in VNP if it can be expressed as

$$f_n(\mathrm{X}) = \sum_{\mathrm{Y}\in\{0,1\}^m} g_{n+m}(\mathrm{X}, \mathrm{Y})$$

31st Symposium on Theoretical Aspects of Computer Science (STACS'14).
Editors: Ernst W. Mayr and Natacha Portier; pp. 239–250
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

where $m = |\mathrm{Y}| = \mathrm{poly}(n)$ and g_{n+m} is a polynomial in VP. Permanent polynomial characterizes the class VNP over the fields of all characteristics except 2 and the determinant polynomial characterizes the class VP with respect to the quasi-polynomial projections.

▶ **Definition 1.** The determinantal complexity of a polynomial f, over n variables, is the minimum m such that there are affine linear functions $A_{k,\ell}$, $1 \leq k, \ell \leq m$ defined over the same set of variables and $f = \det((A_{k,\ell})_{1 \leq k,\ell \leq m})$. It is denoted by $\mathrm{dc}(f)$.

To resolve Valiant's hypothesis, proving $\mathrm{dc}(\mathrm{perm}_n) = n^{\omega(\log n)}$ is sufficient. Von zur Gathen [16] proved $\mathrm{dc}(\mathrm{perm}_n) \geq \sqrt{\frac{8}{7}}n$. Later Cai [2], Babai and Seress [17], and Meshulam [10] independently improved the lower bound to $\sqrt{2}n$. In 2004, Mignon and Ressayre [11] came up with a new idea of using second order derivatives and proved that $\mathrm{dc}(\mathrm{perm}_n) \geq \frac{n^2}{2}$ over the fields of characteristic zero. Subsequently, Cai et al. [3] extended the result of Mignon and Ressayre to all fields of characteristic $\neq 2$.

For any polynomial f, Valiant [15] proved that $\mathrm{dc}(f) \leq 2(F(f) + 1)$ where $F(f)$ is the arithmetic formula complexity of f. Later, Nisan [12] proved that $\mathrm{dc}(f) = O(B(f))$ where $B(f)$ is the arithmetic branching program complexity of f.

Another possible way to prove Valiant's hypothesis is to prove that the permanent polynomial can not be computed by any polynomial size arithmetic circuit. In 2008, Agrawal and Vinay proved that any arithmetic circuit of sub-exponential size can be depth reduced to a depth-4 circuit maintaining a nontrivial upper bound on the size [1]. Subsequently, Koiran [8] and Tavenas [14] have come up with improved depth reductions (in terms of parameters). In particular, Tavenas proved that any $n^{O(1)}$-variate polynomial of degree n in VP can also be computed by a $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$-circuit of top fan-in $2^{O(\sqrt{n}\log n)}$.

In a recent breakthrough, Gupta et al. [5] proved a $2^{\Omega(\sqrt{n})}$ lower bound for the size of the depth-4 circuits computing the determinant or the permanent polynomial using the method of shifted partial derivatives. Subsequently, Kayal et al. [7] improved the situation by proving a $2^{\Omega(\sqrt{n}\log n)}$ depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$-circuit size lower bound for an explicit polynomial in VNP.

More precisely, in [7] the following family of polynomials constructed from the combinatorial design of Nisan-Wigderson [13] was considered:

$$\mathrm{NW}_{n,\epsilon}(\mathrm{X}) = \sum_{a(z) \in \mathbb{F}[z]} x_{1a(1)} x_{2a(2)} \cdots x_{na(n)} \,.$$

where $a(z)$ runs over all univariate polynomials of degree $< k = \epsilon\sqrt{n}$ where $0 < \epsilon < 1$ is a suitably fixed parameter, and $\mathbb{F}$ is a finite field of size n. Here we consider the natural identification of $\mathbb{F}$ with the set $\{1, 2, \ldots, n\}$. Since the number of monomials in $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ is $n^{O(\sqrt{n})}$, the result from [7] gives a tight bound of $2^{\Theta(\sqrt{n}\log n)}$ for the depth-4 circuit complexity of $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$. From the explicitness of the polynomial, it is clear that the polynomial family $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ is in VNP for any $0 < \epsilon < 1$.

Although the combined implication of [7] and [14] looks very exciting from the perspective of lower bounds, a recent result by Fournier et al. [4] shows that such a lower bound is also obtained by the iterated matrix multiplication polynomial which is in VP. Similar to the works of Gupta et al. [5] and Kayal et al. [7], Fournier et al. also used the method of shifted partial derivatives as their main technical tool. The iterated matrix multiplication polynomial of d generic $n \times n$ matrices $\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \ldots, \mathrm{X}^{(d)}$ is the $(1,1)$th entry of the product of the matrices. More formally, let $\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \ldots, \mathrm{X}^{(d)}$ be d generic $n \times n$ matrices with disjoint set of variables and $x_{ij}^{(k)}$ be the variable in $\mathrm{X}^{(k)}$ indexed by $(i,j) \in [n] \times [n]$. Then the iterated matrix multiplication polynomial (denoted by $\mathrm{IMM}_{n,d}$) is defined as follows:

$$\mathrm{IMM}_{n,d}(\mathrm{X}) = \sum_{i_1,i_2,\ldots,i_{d-1}\in[n]} x^{(1)}_{1i_1} x^{(2)}_{i_1 i_2} \ldots x^{(d-1)}_{i_{(d-2)} i_{(d-1)}} x^{(d)}_{i_{(d-1)}1}\,.$$

Notice that $\mathrm{IMM}_{n,d}(\mathrm{X})$ is a $n^2(d-2)+2n$-variate polynomial of degree d. To see that $\mathrm{IMM}_{n,d}(\mathrm{X}) \in \mathrm{VP}$, it is sufficient to observe that it can be computed by a polynomial-size algebraic branching program. For the sake of completeness, we recall the definition of the algebraic branching programs.

▶ **Definition 2.** An algebraic branching program (ABP), over the set of variables X and field $\mathbb{F}$ is a layered (i.e. the edges are only between two consecutive layers) directed acyclic graph G with two special vertices s and t. The weight of an edge is a linear form in $\mathbb{F}[\mathrm{X}]$. The weight of a path is the product of the weights of its edges. The polynomial computed by G is the sum of the weights of all the paths from s to t in G.

To prove $\mathrm{IMM}_{n,d}(\mathrm{X}) \in \mathrm{VP}$, one just needs to observe that for all $1 \le i \le d$ the matrix $\mathrm{X}^{(i)}$ can be identified with the adjacency matrix of the subgraph between the layers i and $i+1$. Hence, the result from [4] is also tight and shows the optimality of the depth reduction of Tavenas [14]. Recent work of Kumar and Saraf [9] shows that the depth reduction as shown by [14] is optimal even for the homogenous formulas. This strengthens the result of [4] who proved the optimality of depth reduction for the circuits.

One of the main motivations of our study comes from this tantalizing fact that two seemingly different polynomials $\mathrm{NW}_{n,\epsilon}(\mathrm{X}) \in \mathrm{VNP}$ and $\mathrm{IMM}_{n,d}(\mathrm{X}) \in \mathrm{VP}$ behave very similarly as far as the $2^{\Omega(\sqrt{n}\log n)}$-size lower bound for depth-4 circuits are concerned. In this paper, we seek a conceptual reason for this behaviour. We identify a simple combinatorial property such that any polynomial that satisfies it would require $2^{\Omega(\sqrt{n}\log n)}$-size depth-4 arithmetic circuits. We call it *Leading Monomial Distance Property.* In particular, it does not matter whether the polynomial is easy (i.e. in VP) or hard (i.e. the polynomial is in VNP but not known to be in VP). As a result of this abstraction we present a simple *unified* analysis of the depth-4 circuit size lower bounds for $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ and $\mathrm{IMM}_{n,d}(\mathrm{X})$.

To define the Leading Monomial Distance Property, we first define the notion of distance between two monomials.

▶ **Definition 3.** Let m_1, m_2 be two monomials over a set of variables. Let S_1 and S_2 be the (multi)-sets of variables corresponding to the monomials m_1 and m_2 respectively. The distance $\mathrm{dist}(m_1, m_2)$ between the monomials m_1 and m_2 is the $\min\{|S_1| - |S_1 \cap S_2|, |S_2| - |S_1 \cap S_2|\}$ where the cardinalities are the order of the (multi)-sets.

For example, let $m_1 = x_1^2 x_2 x_3^2 x_4$ and $m_2 = x_1 x_2^2 x_3 x_5 x_6$. Then $S_1 = \{x_1, x_1, x_2, x_3, x_3, x_4\}$, $S_2 = \{x_1, x_2, x_2, x_3, x_5, x_6\}$, $|S_1| = 6$, $|S_2| = 6$ and $\mathrm{dist}(m_1, m_2) = 3$.

We say that a $n^{O(1)}$-variate and n-degree polynomial has the Leading Monomial Distance Property, if the leading monomials of a *large subset* ($\approx n^{\sqrt{n}}$) of its span of the derivatives (of order $\approx \sqrt{n}$) have *good pair-wise distance.* Leading monomials are defined by defining a suitable order on the set of variables. We denote the leading monomial of a polynomial $f(\mathrm{X})$ by $\mathsf{LM}(f)$. More formally, we prove the following theorem in Section 4.

▶ **Theorem 4.** *Let $f(\mathrm{X})$ be a $n^{O(1)}$-variate polynomial of degree n. Let there be $s \ge n^{\delta k}$ (δ is any constant > 0) different polynomials in $\langle \partial^{=k}(f) \rangle$ for $k = \epsilon\sqrt{n}$ such that any two of their leading monomials have pair-wise distance of at least $\Delta \ge \frac{n}{c}$ for any constant $c > 1$, and $0 < \epsilon < \frac{1}{40c}$. Then any depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit that computes $f(\mathrm{X})$ must be of size $e^{\Omega_{\delta,c}(\sqrt{n}\ln n)}$.*

In fact, from the proof it will be clear that the theorem remains valid for any constant ϵ arbitrarily close to $\frac{1}{4c}$. For technical simplicity, we prefer to the state the above theorem in its current form.

Another motivation of this work is to find a connection between our current knowledge of the determinantal complexity lower bounds and the depth-4 circuit size lower bounds. The best known determinantal complexity lower bound for a $n^{O(1)}$-variate and n degree (Permanent) polynomial is $\Omega(n^2)$. Here we ask the following question: can we give an example of an explicit $n^{O(1)}$-variate degree n polynomial in VNP for which the determinantal complexity is $\Omega(n^2)$ and the depth-4 complexity is $2^{\Omega(\sqrt{n}\log n)}$? We settle this problem by showing a $\Omega(n^2)$ lower bound for $\mathrm{dc}(\mathrm{IMM}_{n,n}(\mathrm{X}))$ which is a $O(n^3)$-variate and n-degree polynomial. In particular, we prove the following theorem.

▶ **Theorem 5.** *For any integers n and $d > 1$, the determinantal complexity of the iterated matrix multiplication polynomial $\mathrm{IMM}_{n,d}$ is $\Omega(dn)$.*

Since $\mathrm{IMM}_{n,d}(\mathrm{X})$ has an algebraic branching program of size $O(dn)$ [12], from the above theorem it follows that $\mathrm{dc}(\mathrm{IMM}_{n,d}(\mathrm{X})) = \Theta(dn)$. This improves upon the earlier bound of $\Theta(n)$ for the determinantal complexity of the iterated matrix multiplication polynomial for any constant $d > 1$ [6]. Similar to the approach of [3] and [11], we also use the the rank of Hessian matrix as our main technical tool.

2 Organization

In Section 3, we state a few results from [5], [7], and [14]. In Section 4, we do a unified analysis of the depth-4 lower bound results of [7] and [4]. We prove the determinantal complexity lower bound of $\mathrm{IMM}_{n,d}(\mathrm{X})$ in Section 5. We state a few open problems in Section 6.

3 Preliminaries

The following beautiful lemma (from [5]) is the key to the asymptotic estimates required for the lower bound analyses.

▶ **Lemma 6** (Lemma 6, [5]). *Let $a(n), f(n), g(n) : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ be the integer valued functions such that $(f+g) = o(a)$. Then,*

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

In this paper, whenever we apply this lemma, $(f+g)^2$ will be $o(a)$. So, we will not worry about the error term (which will be asymptotically zero) generated by this estimate.

The $\Sigma\Pi^{[D]}\Sigma\Pi^{[t]}$ circuits are depth-4 arithmetic circuits with alternating layers of addition and multiplication gates where the fan-in of the multiplication gates in the bottom layer is bounded by a parameter t and the fan-in of the multiplication gates in the layer adjacent to the output gate is bounded by the parameter D. These circuits compute polynomials of the form $C = \sum_{i=1}^{s}\prod_{j=1}^{D_i} Q_{ij}(\mathrm{X})$ where the degree of the polynomial Q_{ij} is bounded by t for all i and j.

Building on the results of [1] and [8], Tavenas [14] proved the following theorem.

▶ **Theorem 7** (Theorem 4, [14]). *Let f be an N-variate polynomial computed by a circuit of size s and of degree d. Then f is computed by a $\Sigma\Pi^{[D]}\Sigma\Pi^{[t]}$ circuit C of size $2^{O(\sqrt{d\log(ds)\log N})}$. Furthermore, if f is homogenous, it will also the case for C.*

Following Tavenas' proof, one can choose $D = 15\sqrt{d}$ and $t = \sqrt{d}$. As a consequence, we infer that any $n^{O(1)}$-variate polynomial of degree n in VP can be computed by a $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit of size $2^{O(\sqrt{n}\log n)}$.

For a monomial $\mathbf{x}^{\mathbf{i}} = x_1^{i_1}x_2^{i_2}\ldots x_n^{i_n}$, let $\partial^{\mathbf{i}} f$ be the partial derivative of f with respect to the monomial $\mathbf{x}^{\mathbf{i}}$. The degree of the monomial is denoted by $|\mathbf{i}|$ where $|\mathbf{i}| = (i_1 + i_2 + \cdots + i_n)$. We recall the following definition of shifted partial derivatives from [5].

▶ **Definition 8.** Let $f(\mathrm{X}) \in \mathbb{F}[\mathrm{X}]$ be a multivariate polynomial. The span of the ℓ-shifted k-th order derivatives of f, denoted by $\langle\partial^{=k} f\rangle_{\leq\ell}$, is defined as

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

We denote by $\dim(\langle\partial^{=k} f\rangle_{\leq\ell})$ the dimension of the vector space $\langle\partial^{=k} f\rangle_{\leq\ell}$.

Let $\succ$ be any admissible monomial ordering. The *leading monomial* of a polynomial $f(\mathrm{X}) \in \mathbb{F}[\mathrm{X}]$, denoted by $\mathsf{LM}(f)$ is the largest monomial $\mathbf{x}^{\mathbf{i}} \in f(\mathrm{X})$ under the order $\succ$. The next lemma follows directly from Proposition 11 and Corollary 12 of [5].

▶ **Lemma 9.** *For any multivariate polynomial $f(\mathrm{X}) \in \mathbb{F}[\mathrm{X}]$,*

$$\dim(\langle\partial^{=k} f\rangle_{\leq\ell}) \geq \#\{\mathbf{x}^{\mathbf{i}} \cdot \mathsf{LM}(g) : \mathbf{i},\mathbf{j} \in \mathbb{Z}^n_{\geq 0} \textit{ with } |\mathbf{i}| \leq \ell, |\mathbf{j}| = k, \textit{ and } g \in \mathbb{F}\textit{-span}\{\partial^{\mathbf{j}} f\}\}\,.$$

In [7], the following upper bound on the dimension of the shifted partial derivative space for polynomials computed by $\Sigma\Pi^{[D]}\Sigma\Pi^{[t]}$ circuits was shown. This bound was implicit in the work of Gupta et al. [5].

▶ **Lemma 10** (Lemma 4, [7]). *If $C = \sum_{i=1}^{s'} Q_{i1}Q_{i2}\ldots Q_{iD}$ where each $Q_{ij} \in \mathbb{F}[\mathrm{X}_N]$ is a polynomial of degree bounded by t. Then for any $k \leq D$,*

$$\dim(\langle\partial^{=k}(C)\rangle_{\leq\ell}) \leq s'\binom{D}{k}\binom{N+\ell+k(t-1)}{N}\,.$$

4 Unified analysis of depth-4 lower bounds

In this section, we first prove a simple combinatorial lemma which we believe is the crux of the best known depth-4 lower bound results. In fact, the lower bounds on the size of $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuits computing the polynomials $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ and $\mathrm{IMM}_{n,n}(\mathrm{X})$ follow easily from this lemma by suitable setting of the parameters.

▶ **Lemma 11.** *Let $m_1, m_2, \ldots, m_s$ be the monomials over N variables s.t. $\mathrm{dist}(m_i, m_j) \geq \Delta$ for all $i \neq j$. Let M be the set of monomials of the form $m_i m'$ where $1 \leq i \leq s$ and m' is a monomial of length at most ℓ over the same set of N variables. Then, the cardinality of M is at least $\left(sB - s^2\binom{N+\ell-\Delta}{N}\right)$ where $B = \binom{N+\ell}{N}$.*

Proof. Let B_i be the set of all monomials $m_i m'$ where m' is a monomial of length at most ℓ. It is easy to see that $|B_i| = \binom{N+\ell}{N}$. We would like to estimate $|\cup_i B_i|$. Using the principle of inclusion and exclusion, we get $|\cup_{i=1}^s B_i| \geq \sum_{i\in[s]} |B_i| - \sum_{i,j\in[s],i\neq j} |B_i \cap B_j|$.

Now we estimate the upper bound for $|B_i \cap B_j|$ such that $i \neq j$. Consider the monomials M_i and M_j in B_i and B_j respectively. For M_i and M_j to match, M_i should contain at least Δ variables from m_j and similarly M_j should contain at least Δ variables from m_i. The rest of the at most $(\ell - \Delta)$ degree monomials should be identical in M_i and M_j. The number of such monomials over N variables is at most $\binom{N+\ell-\Delta}{N}$. Thus, $|B_i \cap B_j| \leq \binom{N+\ell-\Delta}{N}$.

Then the total number of monomials of the form $m_i m'$ for all $i \in [s]$ where m' is a monomial of length at most ℓ is lower bounded as follows:

$$|\cup_{i=1}^{s} B_i| \geq sB - s^2 \binom{N+\ell-\Delta}{N} = sB\left(1 - \frac{s}{B}\binom{N+\ell-\Delta}{N}\right).$$

◀

We use the above lemma to prove the main theorem of this section (restated from Section 1).

▶ **Theorem 12.** *Let $f(\mathrm{X})$ be a $n^{O(1)}$-variate polynomial of degree n. Let there be at least $n^{\delta k}$ (δ is any constant > 0) different polynomials in $\langle \partial^{=k}(f) \rangle$ for $k = \epsilon\sqrt{n}$ such that any two of their leading monomials have a distance of at least $\Delta \geq \frac{n}{c}$ for any constant $c > 1$, and $0 < \epsilon < \frac{1}{40c}$. Then any depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit that computes $f(\mathrm{X})$ must be of size $e^{\Omega_{\delta,c}(\sqrt{n}\ln n)}$.*

Proof. Consider a set of $s = n^{\delta k}$ polynomials $f_1, f_2, \ldots, f_s \in \langle \partial^{=k}(f) \rangle$ such that $\mathrm{dist}(\mathsf{LM}(f_i), \mathsf{LM}(f_j)) \geq n/c$ for all $i \neq j$. We denote by m_i, the leading monomial $\mathsf{LM}(f_i)$.

We now invoke Lemma 11 with the parameters $s = n^{\delta k}, \Delta = n/c$. Let N be the number of variables in f. From Lemma 11, we know that $|\cup_{i=1}^{s} B_i| \geq sB\left(1 - \frac{s}{B}\binom{N+\ell-\Delta}{N}\right)$. To get a good lower bound for $|\cup_{i=1}^{s} B_i|$, we need to upper bound $\frac{s}{B}\binom{N+\ell-\Delta}{N}$. Let us bound it by an inverse polynomial in n by suitably choosing ℓ. We set $\frac{s\binom{N+\ell-d}{N}}{\binom{N+\ell}{N}} \leq \frac{1}{p(n)}$ where $p(n)$ is a polynomial in n.

After simplification, we get $s\frac{(N+\ell-\Delta)!}{(N+\ell)!}\frac{\ell!}{(\ell-\Delta)!} \leq \frac{1}{p(n)}$. Using Lemma 6 we tightly estimate the subsequent computations. In particular, we always choose the parameter ℓ such that $\Delta^2 = o(N+\ell)$. This also shows that the error term given by Lemma 6 is always asymptotically zero and we need not worry about it.

We now apply Lemma 6 to derive $s\left(\frac{\ell}{N+\ell}\right)^{\Delta} \leq \frac{1}{p(n)}$ or equivalently $s\left(\frac{1}{1+\frac{N}{\ell}}\right)^{\Delta} \leq \frac{1}{p(n)}$. We use the inequality $1 + x > e^{x/2}$ for $0 < x < 1$ to lower bound $\left(1 + \frac{N}{\ell}\right)^{\Delta}$ by $e^{\frac{N\Delta}{2l}}$. Thus, it is enough to choose ℓ in a way that $s \cdot p(n) \leq e^{\frac{N\Delta}{2\ell}}$ or equivalently $\ell \leq \frac{N\Delta}{2\ln(s \cdot p(n))}$. By fixing $p(n) = n^2$ and substituting for the parameters k and Δ , we get $\ell \leq \frac{N\sqrt{n}}{4c\delta\epsilon\ln n}$. From Lemma 9, we get that the dimension of $\langle \partial^{=k} f \rangle_{\leq \ell} \geq (1 - \frac{1}{n^2})\ s\binom{N+\ell}{N}$.

Combining this with Lemma 10, we get $s' \geq \frac{(1-\frac{1}{n^2})s\binom{N+l}{N}}{\binom{D}{k}\binom{N+l+k(t-1)}{N}}$. Suppose we choose ℓ such that $(kt - k)^2 = o(\ell)$. Then, by applying Lemma 6 we can easily show the following:

$$s' \geq \frac{s\left(1 - \frac{1}{n^2}\right)}{\binom{D}{k}(1 + \frac{N}{l})^{(kt-k)}} \geq \frac{n^{\delta k}\left(1 - \frac{1}{n^2}\right)}{\binom{D}{k} e^{\frac{N}{\ell}kt}}.$$

Since $D = O(\sqrt{n})$ and $k = \epsilon\sqrt{n}$, we can estimate $\binom{D}{k}$ to be $e^{O_\epsilon(\sqrt{n})}$ by Shannon's entropy estimate for binomial coefficients. To get the required lower bound it is sufficient to choose ℓ such that $\frac{Nkt}{\ell} < (0.1)\delta k \ln n$. Since $t \leq \sqrt{n}$, it is enough to choose $\ell > \frac{10N\sqrt{n}}{\delta\ln n}$. By comparing the lower and upper bounds of ℓ, we can fix ϵ such that $\epsilon < \frac{1}{40c}$. Since ϵ depends only on c, we can infer that $s' = e^{\Omega_{\delta,c}(\sqrt{n}\ln n)}$. ◀

The above proof clearly goes through even if we set $\frac{Nkt}{\ell} < \mu\delta k \ln n$ for any $0 < \mu < 1$, and choose $\epsilon < \frac{\mu}{4c}$. But for simplicity, we prefer to state Theorem 12 in its current form.

In the next section, we show that the lower bounds on the size of $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuits computing $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ and $\mathrm{IMM}_{n,n}(\mathrm{X})$ can be obtained by simply applying Theorem 12.

Moreover, it shows that the lower bound arguments of $\mathrm{IMM}_{n,n}(\mathrm{X})$ are essentially same as the lower bound arguments of $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$.

4.1 Lower bounds on the size of depth-4 circuits computing $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ and $\mathrm{IMM}_{n,n}(\mathrm{X})$

Now we derive the depth-4 circuit size lower bound for $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ polynomial by a simple application of Theorem 12.

▶ **Corollary 13.** *For $0 < \epsilon < 1/80$, any depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit computing the polynomial $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ must be of size $2^{\Omega(\sqrt{n}\log n)}$.*

Proof. Recall that $\mathrm{NW}_{n,\epsilon}(\mathrm{X}) = \sum_{a(z)\in\mathbb{F}[z]} x_{1a(1)}x_{2a(2)}\dots x_{na(n)}$ where $\mathbb{F}$ is a finite field of size n and $a(z)$ is a univariate polynomial of degree $\leq k-1$ where $k = \epsilon\sqrt{n}$. Notice that any two monomials can intersect in at most $k-1$ variables.

We differentiate the polynomial $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ with respect to the first $k = \epsilon\sqrt{n}$ variables of each monomial. After differentiation, we get n^k monomials of length $(n-k)$ each. Since they are constructed from the image of univariate polynomials of degree at most $(k-1)$, the distance Δ between any two monomials $\geq n - 2k > n/2$. So to get the required lower bound we invoke Theorem 12 with $\delta = 1$ and $c = 2$. ◀

Next we derive the lower bound on the size of the depth-4 circuit computing $\mathrm{IMM}_{n,n}(\mathrm{X})$.

▶ **Corollary 14.** *Any depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit computing the $\mathrm{IMM}_{n,n}(\mathrm{X})$ polynomial must be of size $2^{\Omega(\sqrt{n}\log n)}$.*

Proof. Recall that $\mathrm{IMM}_{n,n}(\mathrm{X}) = \sum_{i_1,i_2,\dots,i_{n-1}\in[n]} x^{(1)}_{1i_1}x^{(2)}_{i_1i_2}\dots x^{(n-1)}_{i_{(n-2)}i_{(n-1)}}x^{(n)}_{i_{(n-1)}1}$. It is a polynomial over $(n-2)n^2 + 2n$ variables. We fix the following lexicographic ordering on the variables of the set of matrices $\{\mathrm{X}^{(1)}, \mathrm{X}^{(2)}, \dots, \mathrm{X}^{(n)}\}$ as follows: $\mathrm{X}^{(1)} \succ \mathrm{X}^{(2)} \succ \mathrm{X}^{(3)} \succ \dots \succ \mathrm{X}^{(n)}$ and in any $\mathrm{X}^{(i)}$ the ordering is $x^{(i)}_{11} \succ x^{(i)}_{12} \succ \dots \succ x^{(i)}_{1n} \succ \dots \succ x^{(i)}_{n1} \dots \succ x^{(i)}_{nn}$.

Choose a prime p such that $\frac{n}{2} \leq p \leq n$. Consider the set of univariate polynomials $a(z) \in \mathbb{F}_p[z]$ of degree at most $(k-1)$ for $k = \epsilon\sqrt{n}$ where ϵ is a small constant to be fixed later in the analysis.

Consider a set of $2k$ of the matrices $\mathrm{X}^{(2)}, \mathrm{X}^{(3+\frac{n}{4k})}, \dots, \mathrm{X}^{(2k+1+\frac{(2k-1)n}{4k})}$ such that they are $n/4k$ distance apart. Clearly $2k + 1 + \frac{(2k-1)n}{4k} < n$. For each univariate polynomial a of degree at most $(k-1)$, define a set $S_a = \{x^{(2)}_{1,a(1)}, x^{(3+\frac{n}{4k})}_{2,a(2)}, \dots, x^{(2k+1+\frac{(2k-1)n}{4k})}_{2k,a(2k)}\}$. Number of such sets is at least $\left(\frac{n}{2}\right)^k$ and $|S_a \cap S_b| < k$ for $a \neq b$. Now we consider a polynomial $f(\mathrm{X})$ which is a restriction of the polynomial $\mathrm{IMM}_{n,n}(\mathrm{X})$. By restriction, we simply mean that a few variables of $\mathrm{IMM}_{n,n}(\mathrm{X})$ are fixed to some elements from the field and the rest of the variables are left untouched. We define the restriction as follows:

$$x^{(q)}_{ij} = 0 \text{ if } r + \frac{(r-2)n}{4k} < q < (r+1) + \frac{(r-1)n}{4k} - 1 \text{ for } 2 \leq r \leq 2k \text{ and } i \neq j.$$

The rest of the variables are left untouched.

Next we differentiate the polynomial $f(\mathrm{X})$ with respect to the sets of variables S_a indexed by the polynomials $a(z) \in \mathbb{F}[z]$. Consider the leading monomial of the derivatives with respect to the sets S_a for all $a(z) \in \mathbb{F}[z]$. Since $|S_a \cap S_b| < k$, it is straightforward to observe that the distance between any two leading monomials is at least $k \cdot \frac{n}{4k} = \frac{n}{4}$. The intuitive justification is that whenever there is a difference in S_a and S_b, that difference can be stretched to a distance $\frac{n}{4k}$ because of the restriction that eliminates the non diagonal entries.

Now we prove the lower bound for the polynomial $f(\mathrm{X})$ by applying Theorem 12. Notice that $f(\mathrm{X})$ is a $n^{O(1)}$-variate polynomial of degree n such that there are at least $(n/2)^k > n^{\frac{1}{4}(2k)}$ different polynomials in $\langle \partial^{=2k}(f)\rangle$ such that any two of their leading monomials have distance $\Delta \geq n/4$. So we set the parameters $\delta = 1/4$ and $c = 4$ in Theorem 12. A simple calculation shows that the parameter ϵ can be fixed to something $< 1/320$.

Since $f(\mathrm{X})$ is a restriction of $\mathrm{IMM}_{n,n}(\mathrm{X})$, any lower bound for $f(\mathrm{X})$ is a lower bound for $\mathrm{IMM}_{n,n}(\mathrm{X})$ too. Otherwise, if $\mathrm{IMM}_{n,n}(\mathrm{X})$ has a $2^{o(\sqrt{n}\log n)}$ sized $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit, then we get a $2^{o(\sqrt{n}\log n)}$ sized $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit for $f(\mathrm{X})$ by substituting for the variables according to the restriction. ◀

5 Determinantal complexity of $\mathrm{IMM}_{n,d}(\mathrm{X})$

We start by recalling a few facts from [3]. Let $A_{k,\ell}(\mathrm{X})$, $1 \leq k, \ell \leq m$ be the affine linear functions over $\mathbb{F}[\mathrm{X}]$ such that the following is true:

$$\mathrm{IMM}_{n,d}(\mathrm{X}) = \det((A_{k,\ell}(\mathrm{X}))_{1\leq k,\ell\leq m})\,.$$

Consider a point $\mathrm{X}_0 \in \mathbb{F}^{n^2 d}$ such that $\mathrm{IMM}_{n,d}(\mathrm{X}_0) = 0$. The affine linear functions $A_{k,\ell}(\mathrm{X})$ can be expressed as $L_{k,\ell}(\mathrm{X} - \mathrm{X}_0) + y_{k,\ell}$ where $L_{k,\ell}$ is a linear form and $y_{k,\ell}$ is a constant from the field. Thus, $(A_{k,\ell}(\mathrm{X}))_{1\leq k,\ell\leq m} = (L_{k,\ell}(\mathrm{X} - \mathrm{X}_0))_{1\leq k,\ell\leq m} + \mathrm{Y}_0$. If $\mathrm{IMM}_{n,d}(\mathrm{X}_0) = 0$ then $\det(\mathrm{Y}_0) = 0$. Let C and D be two non-singular matrices such that $\mathrm{CY}_0\mathrm{D}$ is a diagonal matrix:

$$\mathrm{CY}_0\mathrm{D} = \begin{pmatrix} 0 & 0 \\ 0 & \mathrm{I}_s \end{pmatrix}.$$

Since $\det(\mathrm{Y}_0) = 0$, $s < m$. From the previous works [17], [2], [11], and [3], it is enough to assume that $s = m - 1$. Since the first row and the first column of $\mathrm{CY}_0\mathrm{D}$ are zero, we may multiply $\mathrm{CY}_0\mathrm{D}$ by $\mathrm{diag}(\det(\mathrm{C})^{-1}, 1, \ldots, 1)$ and $\mathrm{diag}(\det(\mathrm{D})^{-1}, 1, \ldots, 1)$ on the left and the right side. Without loss of generality, we may assume that $\det(\mathrm{C}) = \det(\mathrm{D}) = 1$. By multiplying with C and D on the left and the right and suitably renaming $(L_{k,\ell}(\mathrm{X} - \mathrm{X}_0))_{1\leq k,\ell\leq m}$ and Y_0 we get

$$\mathrm{IMM}_{n,d}(\mathrm{X}) = \det((L_{k,\ell}(\mathrm{X} - \mathrm{X}_0)_{1\leq k,\ell\leq m} + \mathrm{Y}_0))$$

where $\mathrm{Y}_0 = \mathrm{diag}(0, 1, \ldots, 1)$.

We use $\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X})$ to denote the Hessian matrix of the iterated matrix multiplication and is defined as follows:

$$\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}) = (H_{s;ij,t;k\ell}(\mathrm{X}))_{1\leq i,j\leq n, 1\leq s,t\leq d}$$

$$H_{s;ij,t;k\ell}(\mathrm{X}) = \frac{\partial^2 \mathrm{IMM}_{n,d}(\mathrm{X})}{\partial x_{ij}^{(s)} \partial x_{k\ell}^{(t)}}$$

where $x_{ij}^{(s)}$ and $x_{k\ell}^{(t)}$ denote the (i, j)th and (k, ℓ)th entries of the variable sets $\mathrm{X}^{(s)}$ and $\mathrm{X}^{(t)}$ respectively.

By taking second order derivatives and evaluating the Hessian matrices of $\mathrm{IMM}_{n,d}(\mathrm{X})$ and $\det((A_{k,\ell}(\mathrm{X}))_{1\leq k,\ell\leq m})$ at X_0, we obtain $\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0) = \mathrm{LH}_{\det}(\mathrm{Y}_0)\mathrm{L}^T$ where L is a $n^2 d \times m^2$ matrix with entries from the field. It follows that $\mathrm{rank}(\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0)) \leq \mathrm{rank}(\mathrm{H}_{\det}(\mathrm{Y}_0))$. It was observed in the earlier work of [11] and [3] that it is relatively easy to get an upper bound for $\mathrm{rank}(\mathrm{H}_{\det}(\mathrm{Y}_0))$. The main task is to construct a point X_0 such that $\mathrm{IMM}_{n,d}(\mathrm{X}_0) = 0$, yet the rank of $\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0)$ is high. We give an explicit construction of a point $\mathrm{X}_0 \in \mathbb{F}^{n^2 d}$ such that $\mathrm{IMM}_{n,d}(\mathrm{X}_0) = 0$ and $\mathrm{rank}(\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0)) \geq d(n-1)$. First for the sake of completeness, we briefly recall the upper bound argument for the rank of $\mathrm{H}_{\det}(\mathrm{Y}_0)$ from Section 2.1 of [3].

5.1 Upper bound for the rank of $\mathrm{H}_{\det}(\mathrm{Y}_0)$

When we take a partial derivative $\frac{\partial}{\partial x_{ij}}$ of the determinant, we get the minor after striking out the row i and column j. The second order derivative of $\det(\mathrm{Y})$ with respect to the variables y_{ij} and $y_{k\ell}$ eliminates the rows $\{i, k\}$ and the columns $\{j, \ell\}$. Considering the form of Y_0, the non-zero entries in $\mathrm{H}_{\det}(\mathrm{Y}_0)$ are obtained only if $1 \in \{i, k\}$ and $1 \in \{j, \ell\}$ and thus $(ij, k\ell)$ are of the form $(11, tt)$ or $(t1, 1t)$ or $(1t, t1)$ for any $t > 1$. Thus, $\mathrm{rank}(\mathrm{H}_{\det}(\mathrm{Y}_0)) = O(m)$.

5.2 Lower bound for the rank of $\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0)$

In this section, we prove Theorem 5. In particular, we give a polynomial time algorithm to construct a point X_0 explicitly such that $\mathrm{IMM}_{n,d}(\mathrm{X}_0) = 0$ and $\mathrm{rank}(\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0)) \geq d(n-1)$. Since $\mathrm{rank}(\mathrm{H}_{\det}(\mathrm{Y}_0)) = O(m)$ and $\mathrm{rank}(\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0)) \leq \mathrm{rank}(\mathrm{H}_{\det}(\mathrm{Y}_0))$, we get that $m = \Omega(dn)$. As mentioned in the section 1, the determinantal complexity of $\mathrm{IMM}_{n,d}(\mathrm{X})$ is $O(dn)$. Together, it implies that $m = \Theta(dn)$.

▶ **Theorem 15.** *For any integers $n, d > 1$, there is a point $\mathrm{X}_0 \in \mathbb{F}^{n^2 d}$ such that $\mathrm{IMM}_{n,d}(\mathrm{X}_0) = 0$ and $\mathrm{rank}(\mathrm{H}_{\mathrm{IMM}_{n,d}}(\mathrm{X}_0)) \geq d(n-1)$. Moreover, the point X_0 can be constructed explicitly in polynomial time.*

Proof. We prove the theorem by induction on d. For the purpose of induction, we maintain that the entries indexed by the indices $(1, 2), (1, 3), \ldots, (1, n)$ of the matrix obtained after multiplying the first $(d-1)$ matrices are not all zero at X_0.

We first prove the base case for $d = 2$. The corresponding polynomial is $\mathrm{IMM}_{n,2}(\mathrm{X}) = \sum_{i=1}^{n} x_{1i}^{(1)} x_{i1}^{(2)}$. It is easy to observe that the rank of the Hessian matrix is $2n > 2(n-1)$ at any point since each non-zero entry of the Hessian matrix is 1 and the structure of the Hessian matrix is the following:

$$\mathrm{H}_{\mathrm{IMM}_{n,2}}(\mathrm{X}) = \begin{bmatrix} 0 & B_{12} \\ B_{21} & 0 \end{bmatrix}$$

where $B_{21} = B_{12}^T$. The matrix B_{12} is formally described as follows.

$$(B_{12})_{x_{ij}^{(1)} x_{kl}^{(2)}} = \begin{cases} 1 & \text{if } i = l = 1 \text{ and } j = k \\ 0 & \text{otherwise.} \end{cases}$$

We set the values of the variables as follows: $x_{11}^{(1)} = 0$, $x_{11}^{(2)} = 1$, $x_{21}^{(2)} = x_{31}^{(2)} = \cdots = x_{n1}^{(2)} = 0$ and $x_{12}^{(1)}, x_{13}^{(1)}, \ldots, x_{1n}^{(1)}$ arbitrarily but not all to zero. The point thus obtained (say X_0) is clearly a zero of the polynomial $\mathrm{IMM}_{n,2}(\mathrm{X})$.

For induction hypothesis, assume that the statement of the theorem is true for the case where the number of matrices being multiplied is $\leq d$. Consider the polynomial $\mathrm{IMM}_{n,(d+1)}(\mathrm{X})$:

$$\mathrm{IMM}_{n,(d+1)}(\mathrm{X}) = \sum_{i_1, i_2, \ldots, i_{d-1}, i_d \in [n]} x_{1i_1}^{(1)} x_{i_1 i_2}^{(2)} \ldots x_{i_{(d-2)} i_{(d-1)}}^{(d-1)} x_{i_{(d-1)} i_d}^{(d)} x_{i_d 1}^{(d+1)}.$$

Let the matrix obtained after multiplying the first d matrices be the following:

$$\begin{bmatrix} P_{11}(\mathrm{X}) & P_{12}(\mathrm{X}) & \cdots & P_{1n}(\mathrm{X}) \\ P_{21}(\mathrm{X}) & P_{22}(\mathrm{X}) & \cdots & P_{2n}(\mathrm{X}) \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1}(\mathrm{X}) & P_{n2}(\mathrm{X}) & \cdots & P_{nn}(\mathrm{X}) \end{bmatrix}$$

where

$$P_{k\ell}(\mathrm{X}) = \sum_{i_1,i_2,\ldots,i_{d-1}\in[n]} x^{(1)}_{ki_1} x^{(2)}_{i_1i_2} \ldots x^{(d-1)}_{i_{(d-2)}i_{(d-1)}} x^{(d)}_{i_{(d-1)}\ell} \text{ for } 1 \le k, l \le n\,.$$

Thus, we have the following expression:

$$\mathrm{IMM}_{n,(d+1)}(\mathrm{X}) = P_{11}(\mathrm{X})x^{(d+1)}_{11} + P_{12}(\mathrm{X})x^{(d+1)}_{21} + \cdots + P_{1n}(\mathrm{X})x^{(d+1)}_{n1}\,.$$

Now consider the Hessian matrix $\mathrm{H}_{\mathrm{IMM}_{n,d+1}}(\mathrm{X})$ which is a $(d+1)n^2 \times (d+1)n^2$ sized matrix:

$$\mathrm{H}_{\mathrm{IMM}_{n,d+1}}(\mathrm{X}) = \begin{bmatrix} 0 & B_{1,2} & B_{1,3} & B_{1,4} & \cdots & B_{1,(d+1)} \\ B_{2,1} & 0 & B_{2,3} & B_{2,4} & \cdots & B_{2,(d+1)} \\ B_{3,1} & B_{3,2} & 0 & B_{3,4} & \cdots & B_{3,(d+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{(d+1),1} & B_{(d+1),2} & \cdots & \cdots & B_{(d+1),d} & 0 \end{bmatrix}.$$

Each $B_{i,j}$ is a block of size $n^2 \times n^2$ which is indexed by the variables from the matrices $M^{(i)}$ and $M^{(j)}$ with the corresponding variable sets $\mathrm{X}^{(i)}$ and $\mathrm{X}^{(j)}$. Consider the block $B_{(d+1),d}$ which is indexed by the variable sets $\mathrm{X}^{(d+1)}$ and $\mathrm{X}^{(d)}$. The only non-zero rows in $B_{(d+1),d}$ are indexed by the variables $x^{(d+1)}_{11}, x^{(d+1)}_{21}, \ldots, x^{(d+1)}_{n1}$. The potential non-zero entries for the row $x^{(d+1)}_{11}$ are indexed by the columns $x^{(d)}_{11}, x^{(d)}_{21}, \ldots, x^{(d)}_{n1}$. Similarly the potential non-zero entries for the row $x^{(d+1)}_{21}$ are indexed by the columns $x^{(d)}_{12}, x^{(d)}_{22}, \ldots, x^{(d)}_{n2}$ and so on.

Consider the entries indexed by the indices $(x^{(d+1)}_{11}, x^{(d)}_{11}), (x^{(d+1)}_{11}, x^{(d)}_{21}), \ldots, (x^{(d+1)}_{11}, x^{(d)}_{n1})$. They are $s_1, s_2, \ldots, s_n$ respectively and they can be expressed as follows:

$$s_j = \sum_{i_1,i_2,\ldots,i_{d-2}\in[n]} x^{(1)}_{1i_1} x^{(2)}_{i_1i_2} \ldots x^{(d-1)}_{i_{(d-2)}j} \text{ for } 1 \le j \le n.$$

For the other rows indexed by the variables $x^{(d+1)}_{21}, x^{(d+1)}_{31}, \ldots, x^{(d+1)}_{n1}$, the sequence of potential non-zero entries is the same $(s_1, s_2, \ldots, s_n)$ but their positions are shifted by a column compared to the previous non-zero row. Formally, we have the following:

$$(B_{(d+1),d})_{x^{(d+1)}_{ij} x^{(d)}_{kl}} = \begin{cases} s_k & \text{if } j = 1,\ l = i, \text{ and } i, k \in [n] \\ 0 & \text{otherwise.} \end{cases}$$

$s_1, s_2, \ldots, s_n$ are also the entries indexed by the indices $(1,1), (1,2), \ldots, (1,n)$ of the matrix obtained after multiplying the first $(d-1)$ matrices. By induction hypothesis, we know that the entries indexed by the indices $(1,2), \ldots, (1,n)$ are not all zero at the point X_0 which is a zero of the polynomial $\mathrm{IMM}_{n,d}(\mathrm{X})$. This also makes the rows indexed by the variables $x^{(d+1)}_{11}, x^{(d+1)}_{21}, \ldots, x^{(d+1)}_{n1}$ linearly independent. It is important to note that $P_{11}(\mathrm{X}) = \mathrm{IMM}_{n,d}(\mathrm{X})$.

Now, let us define a point such that it is a zero of the polynomial $\mathrm{IMM}_{n,(d+1)}(\mathrm{X})$. Let X_0 be the zero of the polynomial $P_{11}(\mathrm{X}) = \mathrm{IMM}_{n,d}(\mathrm{X})$. Now to construct the new point, we inductively fix the variables appearing in $P_{11}(\mathrm{X})$ by the values assigned by X_0. We set $x^{(d+1)}_{11} = 1$ and $x^{(d+1)}_{21} = x^{(d+1)}_{31} = \cdots = x^{(d+1)}_{n1} = 0$. We will fix the rest of the variables later. We call the new point which is a zero of the polynomial $\mathrm{IMM}_{n,(d+1)}(\mathrm{X})$, as X_0 as well.

Now, consider the first $d \times d$ blocks of the Hessian matrix $\mathrm{H}_{\mathrm{IMM}_{n,(d+1)}}(\mathrm{X}_0)$. It precisely represents the Hessian matrix of $P_{11}(\mathrm{X})$ which is also the Hessian matrix of the polynomial $\mathrm{IMM}_{n,d}(\mathrm{X})$ at the point X_0[1]. By induction hypothesis, the rank of this minor of $\mathrm{H}_{\mathrm{IMM}_{n,(d+1)}}(\mathrm{X}_0)$ is at least $d(n-1)$. The only non-zero entries in the columns indexed by the variable set $\mathrm{X}^{(d)}$ are indexed by the variables $x_{11}^{(d)}, x_{21}^{(d)}, \ldots, x_{n1}^{(d)}$. This is because the other variables of $\mathrm{X}^{(d)}$ do not appear in $\mathrm{IMM}_{n,d}(\mathrm{X})$. The row in $B_{(d+1)d}$ indexed by $x_{11}^{(d+1)}$ is the only row that interferes with any of the rows of $B_{1d}, B_{2d}, \ldots, B_{dd}$. The rows indexed by the variables $x_{21}^{(d+1)}, x_{31}^{(d+1)}, \ldots, x_{n1}^{(d+1)}$ in $B_{(d+1)d}$ are linearly independent of the rows of $B_{1d}, B_{2d}, \ldots, B_{dd}$. Hence the rank of $\mathrm{H}_{\mathrm{IMM}_{n,(d+1)}}$ at the point described is $\geq (d+1)(n-1)$.

For the purpose of induction, we must verify that the entries indexed by the indices $(1,2), (1,3), \ldots, (1,n)$ of the matrix obtained after multiplying the first d matrices are not all zero at X_0. These entries are the polynomials $P_{12}, P_{13}, \ldots, P_{1n}$. We shall express each of the polynomials in terms of $s_1, s_2, \ldots, s_n$ as follows:

$$P_{1j} = s_1 x_{1j}^{(d)} + s_2 x_{2j}^{(d)} + \cdots + s_n x_{nj}^{(d)} \text{ for } 2 \leq j \leq n.$$

By induction hypothesis, we already know that $s_2, s_3, \ldots, s_n$ are not all zero at X_0. Notice that the variables in $\mathrm{X}^{(d)} \setminus \{x_{11}^{(d)}, x_{21}^{(d)}, \ldots, x_{n1}^{(d)}\}$ were never set in the previous steps of induction[2]. Therefore, we can fix these variables suitably such that $P_{12}, P_{13}, \ldots, P_{1n}$ are not all zero when evaluated at the point X_0 (in fact, we can make all of them non-zero). It is clear that we construct the point X_0 in polynomial time. This completes the proof. ◀

6 Open Problems

In [5] it was proved that any $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit for computing the determinant or the permanent polynomial of a $n \times n$ matrix must be of size $2^{\Omega(\sqrt{n})}$. A natural question is to ask whether one can improve the lower bound to $2^{\Omega(\sqrt{n}\log n)}$. It is unclear whether the leading monomial distance property can be applied directly to Determinant or Permanent to prove such a result. We suspect that it will require a new idea.

▶ **Problem 16.** *Prove that any $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit computing Determinant or Permanent of a $n \times n$ matrix must be of size $2^{\Omega(\sqrt{n}\log n)}$.*

We do not have a good understanding of the determinantal complexity of the $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ polynomial. In particular, we would like to pose the following problem.

▶ **Problem 17.** *Prove that the determinantal complexity of the $\mathrm{NW}_{n,\epsilon}(\mathrm{X})$ polynomial is $\Omega_\epsilon(n^2)$.*

Acknowledgement. We are grateful to the anonymous STACS 2014 referees for their comments and suggestions.

[1] This can be easily seen from the setting of the variables $x_{11}^{(d+1)} = 1$ and $x_{21}^{(d+1)} = x_{31}^{(d+1)} = \cdots = x_{n1}^{(d+1)} = 0$.

[2] Because they do not appear in the polynomial P_{11}.

References

1 Manindra Agrawal and V Vinay. Arithmetic circuits: A chasm at depth four. In *Proceedings-Annual Symposium on Foundations of Computer Science*, pages 67–75. IEEE, 2008.

2 Jin-Yi Cai. A note on the determinant and permanent problem. *Information and Computation*, 84(1):119–127, 1990.

3 Jin-Yi Cai, Xi Chen, and Dong Li. A quadratic lower bound for the permanent and determinant problem over any characteristic$\neq$ 2. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 491–498. ACM, 2008.

4 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)*, 20:100, 2013.

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

6 Maurice Jansen. Lower bounds for the determinantal complexity of explicit low degree polynomials. *Theory of Computing Systems*, 49(2):343–354, 2011.

7 Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:91, 2013.

8 Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012.

9 Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: It's all about the top fan-in. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:153, 2013.

10 Roy Meshulam. On two extremal matrix problems. *Linear Algebra and its Applications*, 114:261–271, 1989.

11 Thierry Mignon and Nicolas Ressayre. A quadratic bound for the determinant and permanent problem. *International Mathematics Research Notices*, 2004(79):4241–4253, 2004.

12 Noam Nisan. Lower bounds for non-commutative computation. In *Proceedings of the twenty-third annual ACM symposium on Theory of computing*, pages 410–418. ACM, 1991.

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

14 Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013.

15 Leslie G Valiant. Completeness classes in algebra. In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 249–261. ACM, 1979.

16 Joachim von zur Gathen. Permanent and determinant. In *FOCS*, pages 398–401. IEEE Computer Society, 1986.

17 Joachim von zur Gathen. Permanent and determinant. *Linear Algebra and its Applications*, 96:87–100, 1987.