

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

Suryajith Chillara Partha Mukhopadhyay
Chennai Mathematical Institute, 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 $2^{O(\sqrt{n}\log n)}$ [Tav13]. So to prove $\text{VP} \neq \text{VNP}$, it is sufficient to show that an explicit polynomial $\in \text{VNP}$ of degree n requires $2^{\omega(\sqrt{n}\log n)}$ size depth-4 circuits. Soon after Tavenas's result, for two different explicit polynomials, depth-4 circuit size lower bounds of $2^{\Omega(\sqrt{n}\log n)}$ have been proved ([KSS13] and [FLMS13]). In particular, using combinatorial design Kayal et al. [KSS13] construct an explicit polynomial in VNP that requires depth-4 circuits of size $2^{\Omega(\sqrt{n}\log n)}$ and Fournier et al. 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 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 $n \times n$ matrix (which is a n^2 -variate and degree n polynomial) [CCL08]. 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 a constant $d > 1$ [Jan11].

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 [Val79]. This is an algebraic analog of the problem P vs NP. Permanent polynomial characterizes the class VNP over fields of all characteristics except 2 and the determinant polynomial characterizes the class VP with respect to 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 $\text{dc}(f)$.*

To resolve Valiant's hypothesis, proving $\text{dc}(\text{perm}_n) = n^{\omega(\log n)}$ is sufficient. Von zur Gathen [vzG86] proved $\text{dc}(\text{perm}_n) \geq \sqrt{\frac{8}{7}}n$. Later Cai [Cai90], Babai and Seress [vzG87], and Meshulam [Mes89] independently improved the lower bound to $\sqrt{2}n$. In 2004, Mignon and Ressayre [MR04] came up with a new idea of using second order derivatives and proved that $\text{dc}(\text{perm}_n) \geq \frac{n^2}{2}$ over the fields of characteristic zero. Subsequently, Cai et al. [CCL08] extended the result of Mignon and Ressayre to all fields of characteristic $\neq 2$.

For any polynomial f , Valiant [Val79] proved that $\text{dc}(f) \leq 2(F(f) + 1)$ where $F(f)$ is the arithmetic formula complexity of f . Later, Nisan [Nis91] proved that $\text{dc}(f) = O(B(f))$ 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 [AV08]. Subsequently, Koiran [Koi12] and Tavenas [Tav13] 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. [GKKS13] 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. [KSS13] 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 [KSS13] the following family of polynomials constructed from the combinatorial design of Nisan-Wigderson [NW94] were considered:

$$\text{NW}(X) = \sum_{a(z) \in \mathbb{F}[z]} x_{1a(1)} x_{2a(2)} \cdots x_{na(n)}$$

where $a(z)$ runs over all univariate polynomials of degree $< k$ for a suitable parameter $k = O(\sqrt{n})$ and \mathbb{F} is a finite field of size n . Here we consider the natural identification of \mathbb{F} with the set $\{1, 2, \dots, n\}$. Since the number of monomials in $\text{NW}(X)$ is $n^{O(\sqrt{n})}$, the result from [KSS13] gives a tight bound of $2^{\Theta(\sqrt{n}\log n)}$ for the depth-4 circuit complexity of $\text{NW}(X)$.

Although the combined implication of [KSS13] and [Tav13] looks very exciting from the perspective of lower bounds, a recent result by Fournier et al. [FLMS13] shows that such a lower bound is also obtained by the iterated matrix multiplication polynomial which is in VP. The iterated matrix multiplication polynomial of d generic $n \times n$ matrices $X^{(1)}, X^{(2)}, \dots, X^{(d)}$ is the $(1, 1)$ th entry of the product of the matrices. More formally, let $X^{(1)}, X^{(2)}, \dots, X^{(d)}$ be d generic $n \times n$ matrices with disjoint set of variables and $x_{ij}^{(k)}$ be the variable in $X^{(k)}$ indexed by $(i, j) \in [n] \times [n]$. Then the iterated matrix multiplication polynomial (denoted by $\text{IMM}_{n,d}$) is defined as follows.

$$\text{IMM}_{n,d}(X) = \sum_{i_1, i_2, \dots, i_{d-1} \in [n]} x_{1i_1}^{(1)} x_{i_1 i_2}^{(2)} \cdots x_{i_{d-2} i_{d-1}}^{(d-1)} x_{i_{d-1} 1}^{(d)}$$

Notice that $\text{IMM}_{n,d}(X)$ is a $n^2(d-2) + 2n$ -variate polynomial of degree d . The result from [FLMS13] is also tight since $\text{IMM}_{n,d} \in \text{VP}$ and very importantly their result proves the optimality of the depth reduction of Tavenas [Tav13].

One of the main motivations of our study comes from this tantalizing fact that two seemingly different polynomials $\text{NW}(X) \in \text{VNP}$ and $\text{IMM}_{n,d}(X) \in \text{VP}$ 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 give a simple *unified* analysis of the depth-4 circuit size lower bounds for $\text{NW}(X)$ and $\text{IMM}_{n,d}(X)$.

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

Definition 2. *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 $\Delta(m_1, m_2)$ 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 $S_1 = \{x_1, x_1, x_2, x_3, x_3, x_4\}$, $S_2 = \{x_1, x_2, x_2, x_3, x_5, x_6\}$ and $\Delta(m_1, m_2) = 3$.

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 $f(X)$ by $\text{LM}(f)$. More formally, we prove the following theorem in Section 4.

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

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 the depth-4 complexity is $2^{\Omega(\sqrt{n} \log n)}$? We settle this problem by showing a $\Omega(n^2)$ lower bound for $\text{dc}(\text{IMM}_{n,n}(X))$ which is a $O(n^3)$ -variate and n -degree polynomial. In particular, we prove the following theorem.

Theorem 2. *For any integers n and $d > 1$, the determinantal complexity of the iterated matrix multiplication polynomial $\text{IMM}_{n,d}$ is $\Omega(dn)$.*

Since $\text{IMM}_{n,d}(X)$ has an algebraic branching program of size $O(dn)$ [Nis91], from the above theorem it follows that $\text{dc}(\text{IMM}_{n,d}(X)) = \Theta(dn)$. This improves upon the earlier bound of $\Theta(n)$ for the determinantal complexity of the iterated matrix multiplication polynomial for a constant $d > 1$ [Jan11]. Similar to the approach of [CCL08] and [MR04], we also use the the rank of Hessian matrix as our main technical tool.

2 Organization

In Section 3, we state some results from [GKKS13]. In Section 4, we do a unified analysis of the depth-4 lower bound results of [KSS13] and [FLMS13]. We prove the determinantal complexity lower bound of $\text{IMM}_{n,d}(X)$ in Section 5.

3 Preliminaries

The following beautiful lemma (from [GKKS13]) is the key to the asymptotic estimates required for the lower bound analyses.

Lemma 1. *Let $a(n), f(n), g(n) : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$ 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.

Building on the results of [AV08] and [Koi12], Tavenas [Tav13] proved 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 top fan-in $2^{O(\sqrt{n}\log n)}$.

In [GKKS13], the authors introduced the method of *shifted partial derivatives* which is the key tool in the recent depth-4 lower bound results. They also gave an elegant way of upper bounding the dimension of the shifted partial derivative space of a polynomial computed by a $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit.

For a monomial $\mathbf{x}^{\mathbf{i}} = x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$, let $\partial^{\mathbf{i}} f$ be the partial derivative of f with respect to the monomial $\mathbf{x}^{\mathbf{i}}$. We recall the following definition of shifted derivatives from [GKKS13].

Definition 3. *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 by $\langle \partial^{=k} f \rangle_{\leq \ell}$, is defined as*

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

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

The following proposition follows from the standard Gaussian elimination technique.

Proposition 1 (Corollary 13, [GKKS13]). *For any multivariate polynomial $f(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]$,*

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

In [GKKS13], 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.

Lemma 2 (Corollary 10, [GKKS13]). *If $C = \sum_{i=1}^{s'} Q_{i1} Q_{i2} \dots Q_{iD}$ where each $Q_{ij} \in \mathbb{F}[X_N]$ is a polynomial of degree bounded by t , then for any $k \leq D$*

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

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 $\text{NW}(\mathbf{X})$ and $\text{IMM}_{n,d}(\mathbf{X})$ follow easily from this lemma by suitable setting of parameters.

Lemma 3. Let m_1, m_2, \dots, m_s be the monomials over N variables such that $\Delta(m_i, m_j) \geq d$. Each monomial m_i is extended by padding monomials of length at most ℓ over N variables. Let $B = \binom{N+\ell}{N}$. Then the total number of monomials generated by the extension is at least $\left(sB - s^2 \binom{N+\ell-d}{N}\right)$.

Proof. Let B_i be the set of monomials constructed by extending the monomial m_i . It is easy to see that $|B_i| = \binom{N+\ell}{N}$. We would like to estimate $|\cup_i B_i|$. From 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 d variables from m_j and similarly M_j should contain at least d variables from m_i . The rest of the at most $(\ell - d)$ degree monomials should be identical in M_i and M_j . The number of such monomials over N variables is at most $\binom{N+\ell-d}{N}$. Thus,

$$|B_i \cap B_j| \leq \binom{N+\ell-d}{N}$$

Then the total number of monomials after extension is lower bounded as follows.

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

□

To get a good lower bound for $|\cup_{i=1}^s B_i|$, we need to upper bound $\frac{s}{B} \binom{N+\ell-d}{N}$. Let us bound it by an inverse polynomial in N by suitably choosing ℓ : $\frac{s \binom{N+\ell-d}{N}}{\binom{N+\ell}{N}} \leq \frac{1}{p(N)}$ where $p(N)$ is a polynomial in N .

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

We now apply Lemma 1 to derive the following: $s \left(\frac{\ell}{N+\ell}\right)^d \leq \frac{1}{p(N)}$ or equivalently $s \left(\frac{1}{1+\frac{\ell}{N}}\right)^d \leq \frac{1}{p(N)}$. We use the inequality $1+x > e^{x/2}$ for $0 < x < 1$ to lower bound $\left(1+\frac{\ell}{N}\right)^d$ by $e^{\frac{Nd}{2\ell}}$. Thus, it is enough to choose ℓ in a way that $s \cdot p(N) \leq e^{\frac{Nd}{2\ell}}$ or equivalently $\ell \leq \frac{Nd}{2 \ln(s \cdot p(N))}$.

Let f be a polynomial such that there are at least s polynomials in $\langle \partial^{=k}(f) \rangle$ (for a suitable value of k) and any two of their leading monomials have distance of at least d . Then for the above choice of ℓ and from Proposition 1, we know that the dimension of $\langle \partial^{=k} f \rangle_{\leq \ell} \geq \left(1 - \frac{1}{p(N)}\right) s \binom{N+\ell}{N}$.

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

We get the following Theorem from the above discussion.

Theorem 3. Let $f(X)$ be a $n^{O(1)}$ -variate polynomial of degree n . Let there be $s \geq n^{\delta k}$ (δ is some constant > 0) different polynomials in $\{\partial^{=k}(f)\}$ for $k = \epsilon\sqrt{n}$ (where $0 < \epsilon < 1$ is any constant) such that any two of their leading monomials have distance of at least $d \geq \frac{n}{c}$ for a constant $c > 1$. Then any depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit that computes $f(X)$ must be of size $s' \geq e^{\Omega(\sqrt{n}\ln n)}$.

Proof. All we need to do is to choose ℓ appropriately so that we get the desired lower bound. Let N be the number of variables in f . Then from the upper bound estimate of ℓ , we get that $\ell \leq \frac{N(n/c)}{2(\delta k \ln n + 2 \ln n)}$ by fixing the inverse polynomial to $\frac{1}{n^2}$. So it is enough to choose ℓ such that $\ell \leq \frac{Nn}{4c\delta k \ln n} = \frac{N\sqrt{n}}{4c\delta\epsilon \ln n}$.

From the lower bound arguments mentioned above, we know that $s' \geq \frac{(1-\frac{1}{2})n^{\delta k}}{\binom{D+k}{k}e^{\frac{Nkt}{\ell}}}$. Since $D = c'\sqrt{n}$ for some $c' > 0$ and $k = \epsilon\sqrt{n}$, $\binom{D+k}{k} = e^{O(\sqrt{n})}$ by Shannon's entropy estimate for binomial coefficient. To get the required lower bound it is enough to choose ℓ such that $\frac{Nkt}{\ell} < \mu\delta k \ln n$ for some $0 < \mu < 1$. Since $t \leq \sqrt{n}$, it is enough to take $\ell > \frac{N\sqrt{n}}{\mu\delta \ln n}$.

By comparing the lower and upper bounds of ℓ , we get that $\epsilon < \frac{\mu}{4c}$. Once we fix μ , we need to choose $\epsilon < \min\{c', \frac{\mu}{4c}\}$. \square

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 $\text{NW}(X)$ and $\text{IMM}_{n,n}(X)$ can be obtained by simply applying the Theorem 3. Moreover, it shows that the lower bound arguments of $\text{IMM}_{n,n}(X)$ are essentially same as the lower bound arguments of $\text{NW}(X)$.

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

Now we derive the depth-4 circuit size lower bound for $\text{NW}(X)$ polynomial by a simple application of Theorem 3.

Corollary 1. Any depth-4 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuit computing the $\text{NW}(X)$ polynomial over n^2 variables would be of size $2^{\Omega(\sqrt{n}\log n)}$.

Proof. Recall that $\text{NW}(X) = \sum_{a(z) \in \mathbb{F}[z]} x_{1a(1)}x_{2a(2)} \dots x_{na(n)}$ where \mathbb{F} is a finite field of size n and $a(z)$ are the polynomials of degree $\leq k-1$. Notice that any two monomials can intersect in at most $k-1$ variables. Here we fix an ordering on the variables: $x_{11} \succ x_{12} \succ \dots \succ x_{nn}$.

We differentiate the polynomial $\text{NW}(X)$ 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 d between any two monomials $\geq n-2k > n/2$. So to get the required lower bound we apply Theorem 3 with $\delta = 1$ and $c = 2$. The parameter ϵ will get fixed from Theorem 3. \square

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

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

Proof. Recall that $\text{IMM}_{n,n}(X) = \sum_{i_1, i_2, \dots, i_{n-1} \in [n]} x_{1i_1}^{(1)} x_{i_1 i_2}^{(2)} \dots x_{i_{n-2} i_{n-1}}^{(n-1)} x_{i_{n-1} 1}^{(n)}$. It is a polynomial over $(n-2)n^2 + 2n$ variables. Next we fix the following lexicographic ordering on the variables of the set of matrices

$\{X^{(1)}, X^{(2)}, \dots, X^{(n)}\}$ as follows: $X^{(1)} \succ X^{(2)} \succ X^{(3)} \succ \dots \succ X^{(n)}$ and in any $X^{(i)}$ the ordering is $x_{11}^{(i)} \succ x_{12}^{(i)} \succ \dots \succ x_{1n}^{(i)} \succ \dots \succ x_{n1}^{(i)} \succ \dots \succ x_{nn}^{(i)}$.

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

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

The rest of the variables are left untouched.

Next we differentiate the polynomial $f(X)$ 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 non diagonal entries.

Now we prove the lower bound for the polynomial $f(X)$ by applying Theorem 3. Clearly $s \geq (n/2)^k > n^{\frac{1}{4}(2k)}$. So we set the parameter $\delta = 1/4$ and $c = 4$. The parameter ϵ will get fixed from Theorem 3. Since $f(X)$ is a restriction of $\text{IMM}_{n,n}(X)$, any lower bound for $f(X)$ is a lower bound for $\text{IMM}_{n,n}(X)$ too : If $\text{IMM}_{n,n}(X)$ 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 $f(X)$ by fixing the variables according to the restriction. \square

5 Determinantal complexity of $\text{IMM}_{n,d}$

We start by recalling a few facts from [CCL08]. Let $A_{k,\ell}(X)$, $1 \leq k, \ell \leq m$ be the affine linear functions over $\mathbb{F}[X]$ such that the following is true.

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

Consider a point $X_0 \in \mathbb{F}^{n^2 d}$ such that $\text{IMM}_{n,d}(X_0) = 0$. The affine linear functions $A_{k,\ell}(X)$ can be expressed as $L_{k,\ell}(X - X_0) + y_{k,\ell}$ where $L_{k,\ell}$ is a linear form and $y_{k,\ell}$ is a constant from the field. Thus, $(A_{k,\ell}(X))_{1 \leq k, \ell \leq m} = (L_{k,\ell}(X - X_0))_{1 \leq k, \ell \leq m} + Y_0$. If $\text{IMM}_{n,d}(X_0) = 0$ then $\det(Y_0) = 0$. Let C and D be two non-singular matrices such that CY_0D is a diagonal matrix.

$$CY_0D = \begin{pmatrix} 0 & 0 \\ 0 & I_s \end{pmatrix}$$

Since $\det(Y_0) = 0$, $s < m$. From the previous works [vzG87], [Cai90], [MR04], and [CCL08], it is enough to assume that $s = m - 1$. Since the first row and the first column of CY_0D are zero, we may multiply CY_0D by $\text{diag}(\det(C)^{-1}, 1, \dots, 1)$ and $\text{diag}(\det(D)^{-1}, 1, \dots, 1)$ on the left and the right side. Without loss of generality, we may assume that $\det(C) = \det(D) = 1$. By multiplying with C and D on the left and the right and suitably renaming $(L_{k,\ell}(X - X_0))_{1 \leq k, \ell \leq m}$ and Y_0 we get

$$\text{IMM}_{n,d}(X) = \det((L_{k,\ell}(X - X_0))_{1 \leq k, \ell \leq m} + Y_0)$$

where $Y_0 = \text{diag}(0, 1, \dots, 1)$.

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

$$\begin{aligned} H_{\text{IMM}_{n,d}}(X) &= (H_{s;ij,t;k\ell}(X))_{1 \leq i,j \leq n, 1 \leq s,t \leq d} \\ H_{s;ij,t;k\ell}(X) &= \frac{\partial^2 \text{IMM}_{n,d}(X)}{\partial x_{ij}^{(s)} \partial x_{k\ell}^{(t)}} \end{aligned}$$

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

By taking second order derivatives and evaluating the Hessian matrices of $\text{IMM}_{n,d}(X)$ and $\det((A_{k,\ell}(X))_{1 \leq k, \ell \leq m})$ at X_0 , we obtain $H_{\text{IMM}_{n,d}}(X_0) = \text{LH}_{\det}(Y_0)\text{L}^T$ where L is a $n^2d \times m^2$ matrix with entries from the field. It follows that $\text{rank}(H_{\text{IMM}_{n,d}}(X_0)) \leq \text{rank}(H_{\det}(Y_0))$. We give an explicit construction of a point $X_0 \in \mathbb{F}^{n^2d}$ such that $\text{IMM}_{n,d}(X_0) = 0$ and $\text{rank}(H_{\text{IMM}_{n,d}}(X_0)) \geq d(n - 1)$.

5.1 Upper bound for the rank of $H_{\det}(Y_0)$

This analysis is the same as the one given in [CCL08] and [MR04] and we briefly recall it here for the sake of completeness. The second order derivative of $\det(Y)$ 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 $H_{\det}(Y_0)$ 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 $t > 1$. This gives $\text{rank}(H_{\det}(Y_0)) = O(m)$.

5.2 Lower bound for the rank of $H_{\text{IMM}_{n,d}}(X_0)$

In this section, we prove Theorem 2. In particular, we give a polynomial time algorithm to construct a point X_0 explicitly such that $\text{IMM}_{n,d}(X_0) = 0$ and $\text{rank}(H(\text{IMM}_{n,d}(X_0))) \geq d(n - 1)$. From the discussion in Section 5.1 and the upper bound for $\text{dc}(\text{IMM}_{n,d}(X))$ from [Nis91], it is clear that such a construction is sufficient to prove Theorem 2.

Theorem 4. *For any integers $n, d > 1$, there is a point $X_0 \in \mathbb{F}^{n^2d}$ such that $\text{IMM}_{n,d}(X_0) = 0$ and $\text{rank}(H(\text{IMM}_{n,d}(X_0))) \geq d(n - 1)$. 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 $(1, 2), (1, 3), \dots, (1, n)$ 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 $\text{IMM}_{n,2}(X) = \sum_{i=1}^n x_{1i}^{(1)} x_{i1}^{(2)}$. 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:

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

where the only non-zero rows of B_{12} are shown in the figure below and $B_{21} = B_{12}^T$.

$$\begin{matrix} & x_{11}^{(2)} & x_{12}^{(2)} & \cdots & x_{1n}^{(2)} & x_{21}^{(2)} & x_{22}^{(2)} & \cdots & x_{2n}^{(2)} & \cdots & x_{n1}^{(2)} & x_{n2}^{(2)} & \cdots & x_{nn}^{(2)} \\ \begin{matrix} x_{11}^{(1)} \\ x_{12}^{(1)} \\ \vdots \\ x_{1n}^{(1)} \end{matrix} & \left[\begin{array}{cccccccccccc} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \end{array} \right. \end{matrix}$$

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 $x_{12}^{(1)}, x_{13}^{(1)}, \dots, x_{1n}^{(1)}$ arbitrarily but not all zero. The point thus obtained (say X_0) is clearly a zero of the polynomial $\text{IMM}_{n,2}(X)$.

Now consider the case where the first d matrices are multiplied. The final polynomial could be represented as follows.

$$\text{IMM}_{n,d}(X) = \sum_{i_1, i_2, \dots, i_{d-1} \in [n]} x_{1i_1}^{(1)} x_{i_1 i_2}^{(2)} \cdots x_{i_{d-2} i_{d-1}}^{(d-1)} x_{i_{d-1} 1}^{(d)}$$

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 $\text{IMM}_{n,(d+1)}(X)$.

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

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

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

where

$$P_{k\ell}(X) = \sum_{i_1, i_2, \dots, i_{d-1} \in [n]} x_{ki_1}^{(1)} x_{i_1 i_2}^{(2)} \cdots x_{i_{d-2} i_{d-1}}^{(d-1)} x_{i_{d-1} \ell}^{(d)}$$

Thus, we have the following expression.

$$\text{IMM}_{n,(d+1)}(\mathbf{X}) = P_{11}(\mathbf{X})x_{11}^{(d+1)} + P_{12}(\mathbf{X})x_{21}^{(d+1)} + \cdots + P_{1n}(\mathbf{X})x_{n1}^{(d+1)}$$

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

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

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 $\mathbf{X}^{(i)}$ and $\mathbf{X}^{(j)}$. Consider the block $B_{(d+1),d}$ which is indexed by the variable sets $\mathbf{X}^{(d+1)}$ and $\mathbf{X}^{(d)}$. The only non-zero rows in $B_{(d+1),d}$ are indexed by the variables $x_{11}^{(d+1)}, x_{21}^{(d+1)}, \dots, x_{n1}^{(d+1)}$. The potential non-zero entries for the row $x_{11}^{(d+1)}$ are indexed by the columns $x_{11}^{(d)}, x_{21}^{(d)}, \dots, x_{n1}^{(d)}$. Similarly the potential non-zero entries for the row $x_{21}^{(d+1)}$ are indexed by the columns $x_{12}^{(d)}, x_{22}^{(d)}, \dots, x_{n2}^{(d)}$ and so on.

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

$$s_j = \sum_{i_1, i_2, \dots, i_{d-2} \in [n]} x_{1i_1}^{(1)} x_{i_1 i_2}^{(2)} \cdots x_{i_{d-2} j}^{(d-1)}$$

For the other rows indexed by the variables $x_{21}^{(d+1)}, x_{31}^{(d+1)}, \dots, x_{n1}^{(d+1)}$, the sequence of potential non-zero entries is the same (s_1, s_2, \dots, s_n) but their positions are shifted by a column compared to the previous non-zero row. See the figure for the block $B_{(d+1),d}$ below for a clearer picture.

$$\begin{array}{c} x_{11}^{(d+1)} \\ \vdots \\ x_{21}^{(d+1)} \\ \vdots \\ \vdots \\ x_{n1}^{(d+1)} \\ \vdots \end{array} \begin{bmatrix} x_{11}^{(d)} & x_{12}^{(d)} & \cdots & x_{1n}^{(d)} & x_{21}^{(d)} & x_{22}^{(d)} & \cdots & x_{2n}^{(d)} & \cdots & x_{n1}^{(d)} & x_{n2}^{(d)} & \cdots & x_{nn}^{(d)} \\ s_1 & 0 & \cdots & 0 & s_2 & 0 & 0 & \cdots & \cdots & s_n & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & s_1 & \cdots & 0 & 0 & s_2 & \cdots & \cdots & \cdots & 0 & s_n & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & s_1 & 0 & \cdots & \cdots & s_2 & 0 & \cdots & \cdots & \cdots & s_n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

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

Now, let us define a point such that it is a zero of the polynomial $\text{IMM}_{n,(d+1)}(X)$. We set $x_{11}^{(d+1)} = 1$ and $x_{21}^{(d+1)} = x_{31}^{(d+1)} = \dots = x_{n1}^{(d+1)} = 0$. Inductively fix the variables appearing in $P_{11}(X)$ by the values assigned by X_0 which is a zero of the polynomial $\text{IMM}_{n,d}(X)$. We will fix the other variables suitably later. We call the new point X_0 as well.

Now, consider the first $d \times d$ blocks of the Hessian matrix $H_{\text{IMM}_{n,(d+1)}}(X_0)$. It precisely represents the Hessian matrix of $P_{11}(X)$ which is also the Hessian matrix of the polynomial $\text{IMM}_{n,d}(X)$ at the point X_0 ¹. By induction hypothesis, the rank of this minor of $H_{\text{IMM}_{n,(d+1)}}(X_0)$ is at least $d(n - 1)$. The only non-zero entries in the columns indexed by the variable set $X^{(d)}$ are indexed by the variables $x_{11}^{(d)}, x_{21}^{(d)}, \dots, x_{n1}^{(d)}$. This is because the other variables of $X^{(d)}$ do not appear in $\text{IMM}_{n,d}(X)$. 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 $B_{1d}, B_{2d}, \dots, B_{dd}$. The rows indexed by the variables $x_{21}^{(d+1)}, x_{31}^{(d+1)}, \dots, x_{n1}^{(d+1)}$ in $B_{(d+1)d}$ are linearly independent of the rows of $B_{1d}, B_{2d}, \dots, B_{dd}$. Hence the rank of $H_{\text{IMM}_{n,(d+1)}}$ 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 $(1, 2), (1, 3), \dots, (1, n)$ of the matrix obtained after multiplying the first d matrices are not all zero at X_0 . These entries are the polynomials $P_{12}, P_{13}, \dots, P_{1n}$. We shall express each of the polynomials in terms of s_1, s_2, \dots, s_n as follows.

$$\begin{aligned} P_{12} &= s_1 x_{12}^{(d)} + s_2 x_{22}^{(d)} + \dots + s_n x_{n2}^{(d)} \\ P_{13} &= s_1 x_{13}^{(d)} + s_2 x_{23}^{(d)} + \dots + s_n x_{n3}^{(d)} \\ &\vdots \\ P_{1n} &= s_1 x_{1n}^{(d)} + s_2 x_{2n}^{(d)} + \dots + s_n x_{nn}^{(d)} \end{aligned}$$

By induction hypothesis, we already know that s_2, s_3, \dots, s_n are not all zero at X_0 . Notice that the variables in $X^{(d)} \setminus \{x_{11}^{(d)}, x_{21}^{(d)}, \dots, x_{n1}^{(d)}\}$ were never set in the previous steps of induction². Therefore, we can fix these variables suitably such that $P_{12}, P_{13}, \dots, P_{1n}$ 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. \square

References

- [AV08] 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.

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

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

- [Cai90] Jin-Yi Cai. A note on the determinant and permanent problem. *Information and Computation*, 84(1):119–127, 1990.
- [CCL08] 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.
- [FLMS13] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:100, 2013.
- [GKKS13] 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.
- [Jan11] Maurice Jansen. Lower bounds for the determinantal complexity of explicit low degree polynomials. *Theory of Computing Systems*, 49(2):343–354, 2011.
- [Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012.
- [KSS13] 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.
- [Mes89] Roy Meshulam. On two extremal matrix problems. *Linear Algebra and its Applications*, 114:261–271, 1989.
- [MR04] Thierry Mignon and Nicolas Ressayre. A quadratic bound for the determinant and permanent problem. *International Mathematics Research Notices*, 2004(79):4241–4253, 2004.
- [Nis91] 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.
- [NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994.
- [Tav13] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013.
- [Val79] Leslie G Valiant. Completeness classes in algebra. In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 249–261. ACM, 1979.
- [vzG86] Joachim von zur Gathen. Permanent and determinant. In *FOCS*, pages 398–401. IEEE Computer Society, 1986.
- [vzG87] Joachim von zur Gathen. Permanent and determinant. *Linear Algebra and its Applications*, 96:87–100, 1987.