Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 201 (2015)

Limitations of sum of products of Read-Once Polynomials

C. Ramya[1] and B. V. Raghavendra Rao[1]

1 **Department of Computer Science and Engineering**
IIT Madras, Chennai INDIA
{ramya,bvrr}@cse.iitm.ac.in

Abstract

We study limitations of polynomials computed by depth two circuits built over read-once polynomials (ROPs) and over depth three syntactically multi-linear formulas. We prove an exponential lower bound for the size of the $\Sigma\Pi^{[N^{1/30}]}$ arithmetic circuits built over syntactically multi-linear $\Sigma\Pi\Sigma^{[N^{1/4}]}$ arithmetic circuits computing a product of variable disjoint linear forms on N variables, where the superscripts on gates denote bound on the fan-in. We extend the result to the case of $\Sigma\Pi^{[N^{1/30}]}$ arithmetic circuits built over ROPs of unbounded depth, where the number of variables with $+$ gates as a parent in an proper sub formula is bounded by $N^{1/4}$. We show that the same lower bound holds for the permanent polynomial. Finally we obtain an exponential lower bound for the sum of ROPs computing a polynomial in VP defined by Raz and Yehudayoff [18].

Our results demonstrate a class of formulas of unbounded depth with exponential size lower bound against the permanent and can be seen as an exponential improvement over the multilinear formula size lower bounds given by Raz [17] for a sub-class of multi-linear and non-multi-linear formulas. Our proof techniques are built on the one developed by Kumar et. al. [13] and are based on non-trivial analysis of ROPs under random partitions. Further, our results exhibit strengths and limitations of the lower bound techniques introduced by Raz [17].

Keywords and phrases Arithmetic Circuits, Permanent, Computational Complexity

1 Introduction

More than three decades ago, Valiant [23] developed the theory of Algebraic Complexity classes based on arithmetic circuits as the model of algebraic computation. Valiant considered the permanent polynomial perm_n defined over an $n \times n$ matrix $X = (x_{i,j})_{1\le i,j\le n}$ of variables:

$$\mathsf{perm}_n(X) = \sum_{\pi\in S_n} \prod_{i=1}^{n} x_{i,\pi(i)}$$

where S_n is the set of all permutations on $[n]$. Valiant [23] showed that the polynomial family $(\mathsf{perm}_n)_{n\ge 0}$ is complete for the complexity class VNP. Further, Valiant [23] conjectured that perm_n does not have polynomial size arithmetic circuits. Since then, obtaining super polynomial size lower bounds for arithmetic circuits computing perm_n has been a pivotal problem in Algebraic Complexity Theory. However, for general classes of arithmetic circuits, the best known size bound is an $\Omega(n \log d)$ lower bound due to Baur and Strassen for an n-variate degree d polynomial [2]. In fact, this is the only super linear lower bound we know for general arithmetic circuits. While the challenge of proving lower bounds for general classes of circuits still seems to be at a distance, naturally the focus has been on proving lower bounds for restricted classes of circuits computing $perm_n$.

Nisan and Wigderson [16] used partial derivatives to obtain exponential lower bounds against Depth 3 $\Sigma\Pi\Sigma$ circuits and set multilinear formulas. Later, Grigoriev and Karpinski [6] proved an exponential size lower bound for depth three circuits of constant size over finite

Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

ISSN 1433-8092

fields. In 2001, Shpilka and Wigderson [21] proved a quadratic lower bound for $\Sigma\Pi\Sigma$ circuits over infinite fields computing det_n(or $perm_n$) which has been improved recently to an almost cubic lower bound in [11]. Explaining the lack of progress in proving lower bounds even for $\Sigma\Pi\Sigma$ circuits, Agrawal and Vinay [1] showed that proving exponential lower bounds against depth four arithmetic circuits is enough to resolve Valiant's conjecture. This was improved subsequently in [22, 12]. From then on, depth 4 circuits have been in the limelight. Recently, Gupta et. al. [7] obtained $2^{\Omega(\sqrt{n})}$ top fan-in lower bound for $\Sigma\Pi^{[\mathcal{O}(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ circuits computing $\det_n$ or perm_n. The techniques introduced in [7, 8] have been generalized and applied to prove lower bounds against several classes of constant depth arithmetic circuits, regular arithmetic formulas and homogeneous arithmetic formulas. (See e.g., [9, 14, 10].)

Motivation and our results: A seminal work of Raz [17] showed that multilinear formulas (i.e., every gate in the formula computes a multilinear polynomial) computing $\det_n$ or $perm_n$ must have size $n^{\Omega(\log n)}$. In [17] Raz used rank of the partial derivative matrix as a complexity measure. Using the same complexity measure as [17], Raz and Yehudayoff [19] proved exponential lower bounds against constant depth multilinear formulas [19]. Subsequently, several generalizations of Raz's measure were introduced. Kumar et. al [13] extended the techniques developed in [17] to prove lower bounds against non-multilinear circuits and formulas of constant size using the rank of the polynomial coefficient matrix as a measure. (See Definition 1). In [5], Forbes and Shpilka introduced evaluation dimension as a complexity measure to prove exponential lower bounds against Read-Once oblivious algebraic branching programs. In [10], Kayal and Saha used evaluation dimension to prove exponential lower bound against Depth three multi-ic-k circuits. Despite the fact that over large fields, evaluation dimension with respect to a partition of the set of variables in the polynomial and rank of the partial derivative matrix with respect to that partition are essentially the same (see Chapter 4 in [4]), the evaluation perspective sometimes comes handy in proving lower bounds against non-multilinear circuits.

In this work, we attempt to push Raz's measure to unexplored circuit models. A formula is said to be a *read-once formula* (ROF) if every variable labels atmost one leaf in the formula. A polynomial computed by an ROFs is called a *read-once polynomial* (ROP). Observe that ROFs are the simplest class of multilinear formulas. ROFs have gained much attention after Shpilka and Volkovich [20] obtained an efficient identity testing algorithm for sum of a constant number of ROPs . As an essential ingredient in their result, Shpilka and Volkovich [20] proved a linear lower bound for a special class of ROPs to sum-represent the polynomial $x_1 \cdots x_n$. We prove an exponential lower bound against the same model as in [20] against a polynomial in VP defined by Raz-Yehudayoff [18].

▶ **Theorem 1.** *There is an explicit polynomial* $g \in \mathsf{VP}$ *such that for any* ROP*s* $f_1, \ldots, f_s$*, if* $\sum_{i=1}^{s} f_i = g$*, then* $s = \exp(\Omega(n/\log n))$.

It should be noted that the result in Raz [17] immediately implies a lower bound of $n^{\Omega(\log n)}$ for the sum of ROPs and hence our result is an exponential improvement.

A natural next step is to extend Theorem 1 to $\Sigma\Pi$ circuits built over ROPs ($\Sigma\Pi$ROP for short). More formally, we study the model $\sum_i \prod_j Q_{ij}$ where each Q_{ij} is an ROP. Because of the non-trivial product gate at the second level the polynomials computed can potentially be non-multilinear. Apart from being a natural generalization of $\Sigma\Pi\Sigma$ circuits, the class $\Sigma\Pi$ROP can be seen as building non-multilinear polynomials using the simplest possible multilinear polynomials viz. ROPs.

However, it can easily be shown that rank of the partial derivative matrix under a random partition is only a constant factor away in the exponent from the maximum possible value even for a product of variable disjoint linear forms with high probability. (See Lemma 22.)

This necessitates further restrictions on ROPs that could lead to exponential lower bound against $\Sigma\Pi$ROP using the rank of the partial derivative matrix as the measure of complexity. We show:

▶ **Theorem 2.** *Let $\mathcal{C}$ be the class of N-variate polynomials computed by multilinear formulas $\sum^{[r]}\prod\sum^{[N^{1/4}]}$. Then there is an explicit family of polynomials p_{lin} such that if $p_{lin} = \sum_{i=1}^{s}\prod^{[N^{1/30}]}\mathcal{C}$ then $s \cdot r = \exp(\Omega(N^{\epsilon}))$, for some $\epsilon > 0$.*

Our arguments do not directly generalize to the case of unbounded depth ROPs with small bottom Σ fan-in. We obtain a generalization of Theorem 2, allowing ROPs of unbounded depth with a more stringent restriction than the bottom Σ-fan-in. Let F be an ROF and for a gate v in F, let sum-fan-in(v) be the number of variables in the sub-formula rooted at v whose parents are labelled as $+$. Then $s(F)$ is the maximum value of sum-fan-in(v), where the maximum is taken over all $+$ gates in F of product height at least 1. Note that, in the case of $\Sigma\Pi\Sigma$ ROPs, $s(F)$ is the same as the bottom fan-in. For an ROP f, define $s(f)$ as the smallest $s(F)$ among all ROFs F computing f.

▶ **Theorem 3.** *Let $\mathcal{C}$ be the class of N-variate ROPs f with $s(f) \leq N^{1/4}$. For $N = n^2$, if $p_{lin} = \sum_{i=1}^{s}\prod^{[N^{1/30}]}\mathcal{C}$ then $s = \exp(\Omega(N^{\epsilon}))$, for some $\epsilon > 0$.*

As far as we know, in the commutative setting, this is the first exponential lower bound for a sub-class of non-multilinear and non-homogeneous formulas of unbounded depth. It can be noted that our result above does not depend on the depth of the ROPs.

Even though a product of linear forms is a simple linear projection of perm_n, Theorem 3 does not imply a lower bound for perm_n due to restrictions on s_F, since linear projections might change the bottom fan-in of the resulting ROPs. With a more involved analysis of permanent under random partitions, we have:

▶ **Theorem 4.** *Let $\mathcal{C}$ be the class of N-variate ROPs f with $s(f) \leq N^{1/4}$. For $N = n^2$, if $perm_n = \sum_{i=1}^{s}\prod^{[N^{1/30}]}\mathcal{C}$ then $s = \exp(\Omega(N^{\epsilon}))$ for some $\epsilon > 0$.*

Related Results : In [15], Mahajan and Tawari obtain a tight linear lower bound for number of ROPs required to sum-represent elementary symmetric polynomials. Though the model in [15] is the same as the one in this paper, our lower bounds are incomparable with that of [15]. Kayal [8] showed that at least $2^{n/d}$ many polynomials of degree d are required to represent the polynomial $x_1 \dots x_n$ as sum of powers. Our model is significantly different from the one in [8] since our model includes high degree monomials, though the powers are restricted to be sub-linear, whereas Kayal's argument works against arbitrary powers.

Our Techniques: Our techniques are broadly based on the rank of polynomial coefficient matrix introduced by Kumar et. al [13] as an extension of the partial derivative matrix introduced in [17]. It can be noted that the lower bounds obtained in [17] are super polynomial and not exponential. Though Raz-Yehudayoff [19] proved exponential lower bounds, their argument works only against bounded depth multilinear circuits. Further, the arguments in [17, 19] do not work for the case of non-multilinear circuits, and fail even in the case of products of two multilinear formulas. This is because rank of the partial derivative matrix, a complexity measure used by [17, 19] (see Section 2 for a definition) is defined only for multi-linear polynomials. Even though this issue can be overcome by a generalization introduced by Kumar et. al [13], the limitation lies in the fact that the upper bound of $2^{n-n^{\epsilon}}$ for an n^2 or $2n$ variate polynomial, obtained in [17] or [19] on the measure for the underlying arithmetic formula model is insufficient to handle products of two ROPs.

Our approach to prove Theorems 3 and 4 lie in obtaining an exponentially stronger upper bounds (see Lemma 21) on the rank of the partial derivative matrix of an ROP F on N variables where $s(F) \leq N^{1/4}$. Our proof is a technically involved analysis of the structure of ROPs under random partitions of the variables. Even though the restriction on $s(F)$ might look un-natural, in Lemma 22, we show that a simple product of variable disjoint linear forms in N-variables, with $s(F) \geq N^{2/3}$ achieves exponential rank with probability $1 - 2^{-\Omega(N^{1/3})}$. Thus our results highlight the strength and limitations of the techniques developed in [19, 13] to the case of non-multi-linear formulas.

The rest of the paper is organized as follows. Section 2 provides essential definitions used in the paper. Section 3 proves Theorem 1. Sections 4 proves the remaining results. Proofs omitted due to space constraints can be found in the appendix.

2 Preliminaries

In this section we review the computational model we study and the complexity measure used to prove the lower bounds.

Let $\mathbb{F}$ be an arbitrary field and $X = \{x_1, \ldots, x_N\}$ be a set of variables. An *arithmetic circuit* $\mathcal{C}$ over $\mathbb{F}$ is a directed acyclic graph with vertices of in-degree 0,1 or 2 and exactly one vertex of out-degree 0 called the output gate. The vertices of in-degree 0 are called *input gates* and are labeled by elements from $X \cup \mathbb{F}$. The vertices of in-degree more than 1 are labeled by either $+$ or $\times$. Thus every gate of the circuit naturally computes a polynomial. The polynomial f computed by $\mathcal{C}$ is the polynomial computed by the output gate of the circuit. The *size* of an arithmetic circuit is the number of gates in $\mathcal{C}$. Depth of $\mathcal{C}$ is the length of the longest path from an input gate to the output gate in $\mathcal{C}$. The *product height* of a gate v in $\mathcal{C}$ is the maximum number of Π gates along any path from v the root gate in $\mathcal{C}$. For g any gate in a circuit C, $\mathsf{var}(g)$ denote the set of variables that appear as leaf labels in the sub-circuit rooted at g. Abusing the notation, if g is a polynomial, then $\mathsf{var}(g)$ denotes the set of variables that g is dependent on. An arithmetic circuit is called an *arithmetic formula* if the underlying undirected graph is a tree.

We now review the polynomial coefficient matrix introduced in [13] and take a look its properties.

▶ Definition 1. (*Polynomial Coefficient Matrix*). Let $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$. Let $f \in \mathbb{F}[Y, Z]$ be a polynomial. The *polynomial coefficient matrix* of f(denoted by M_f) is a $2^m \times 2^m$ matrix defined as follows : For monic multilinear monomials p and q in variables Y and Z respectively, the entry $M_f[p, q] = A$ if and only if f can be uniquely expressed as $f = pq \cdot A + B$ where $A, B \in \mathbb{F}[Y, Z]$ such that

- $\mathsf{var}(A) \subseteq \mathsf{var}(p) \cup \mathsf{var}(q)$.
- For every monomial $m \in B$, either $pq \nmid m$ or $\mathsf{var}(m) \subsetneq \mathsf{var}(p) \cup \mathsf{var}(q)$.

▶ Observation 1. For a multilinear polynomial $f \in \mathbb{F}[Y, Z]$, the polynomial coefficient matrix [13] and the partial derivative matrix [17] are the same.

The matrix M_f has entries in $\mathbb{F}[Y, Z]$. Therefore $\text{rank}(M_f)$ is defined only under a substitution function. For $\mathcal{S} : Y \cup Z \to \mathbb{F}$, let $M_f|_{\mathcal{S}}$ be the matrix obtained by substituting every variable $w \in Y \cup Z$ to $\mathcal{S}(w)$ at each entry of M_f.

$$\text{maxrank}(M_f) \triangleq \max_{\mathcal{S}: Y \cup Z \to \mathbb{F}} \{\text{rank}(M_f|_{\mathcal{S}})\}$$

It is known that $\text{maxrank}(M_f)$ satisfies sub-additivity and sub-multiplicativity:

▶ **Lemma 5.** [13](Sub-additivity.) *Let $f, g \in \mathbb{F}[Y, Z]$. Then,* $\text{maxrank}(M_{f+g}) \leq \text{maxrank}(M_f)+$ $\text{maxrank}(M_g)$.

▶ **Lemma 6.** [13](Sub-multiplicativity.) *Let $Y_1, Y_2 \subseteq Y$ and $Z_1, Z_2 \subseteq Z$ such that $Y_1 \cap Y_2 = \emptyset$ and $Z_1 \cap Z_2 = \emptyset$. Then for any polynomials $f \in \mathbb{F}[Y_1, Z_1]$, $g \in \mathbb{F}[Y_2, Z_2]$, we have* $\text{maxrank}(M_{fg}) = \text{maxrank}(M_f) \cdot \text{maxrank}(M_g)$.

The proofs of Lemma 5 and 6 follow directly from [13].

▶ Observation 2. For any multilinear polynomial $f \in \mathbb{F}[Y, Z]$, the entries of M_f are constants from $\mathbb{F}$. Therefore $\text{maxrank}(M_f) = \text{rank}(M_f)$.

▶ Definition 2. (*Partition function*). A partition of X is a function $\varphi : X \to Y \cup Z \cup \{0, 1\}$ such that φ is an injection when restricted to $Y \cup Z$, i.e., $\forall x \neq x' \in X$, if $\varphi(x) \in Y \cup Z$ and $\varphi(x') \in Y \cup Z$ then $\varphi(x) \neq \varphi(x')$.

Let F be a formula with leaves labelled by elements in $X \cup \mathbb{F}$ and $\varphi : X \to Y \cup Z \cup \{0, 1\}$ be a partition function as in Definition 2. Denote by F^{φ} to be the formula obtained by replacing every variable x that appears as a leaf in F by $\varphi(x)$. Denote by f^{φ} the polynomial computed by F^{φ}. Then $f^{\varphi} \triangleq f(\varphi(X)) \in \mathbb{F}[Y, Z]$.

▶ Definition 3. (*Constant-Minimal Formula*) An arithmetic formula F is said to be *constant-minimal* if no gate u in F has both its children as constants from $\mathbb{F}$. Observe that for any arithmetic formula F, if there exists a gate u in F such that $u = a \ \mathsf{op}\ b, a, b \in \mathbb{F}$ then we can replace u in F by the constant $a\ \mathsf{op}\ b$, where $\mathsf{op} \in \{+, \times\}$. Thus we assume without loss of generality that any arithmetic formula F is constant-minimal.

We need some observations on formulas that compute natural numbers. Recall that an arithmetic formula F is said to be monotone if F does not contain any negative constants.

Let G be a monotone arithmetic formula where the leaves are labelled numbers in $\mathbb{N}$. Then for any gate v in G, the value of v (denoted by $\mathsf{value}(v)$) is defined as : If u is a leaf then $\mathsf{value}(u) = a$ where $a \in \mathbb{N}$ is the label of u. If $u = u_1\ \mathsf{op}\ u_2$ then $\mathsf{value}(u) = \mathsf{value}(u_1)\ \mathsf{op}\ \mathsf{value}(u_2)$, where $\mathsf{op} \in \{+, \times\}$. Finally, $\mathsf{value}(G)$ is the value of the output gate of G.

The following is a simple upper bound on the value computed by a monotone formula. See appendix for a proof.

▶ **Lemma 7.** *Let G be a binary monotone arithmetic formula with t leaves. If every leaf in G takes a value at most $N > 1$, then $\mathsf{value}(G) \leq N^t$.*

▶ Definition 4. (*rank-$(1,2)$-separator*). Let G be a monotone arithmetic formula with leaves labelled by either 1 or 2. A node u in G at product height at least 1 is called a *rank*-$(1,2)$-*separator* if u is a leaf and $\mathsf{value}(u) = 2$ or u is a sum gate($u = u_1 + u_2$) with $\mathsf{value}(u) \geq 2$ and $\mathsf{value}(u_1), \mathsf{value}(u_2) < 2$.

Note that no gate labelled $\times$ can be a *rank*-$(1,2)$-*separator*. The following lemma shows that any formula computing a large value should have a large number of *rank*-$(1,2)$-*separator*s. Proof can be found in the appendix

▶ **Lemma 8.** *Let F be a binary monotone arithmetic formula with leaves labelled by either 1 or 2. Suppose $\mathsf{value}(F) > 2^r$ then there are at least $\lceil \frac{r}{\log N} \rceil$ gates that are rank-$(1,2)$-separators, where N is the sum of labels of leaves in F.*

Finally, we will use the following variants of Chernoff-Hoeffding bounds.

▶ **Theorem 9.** [3](Chernoff-Hoeffding bound) *Let $X_1, X_2, \ldots, X_n$ be independent random variables. Let $X = X_1 + X_2 + \cdots + X_n$ and $\mu = \mathbb{E}[X]$. Then for any $\delta > 0$,*

(1) $\Pr[X \geq (1+\delta)\mu] \leq \mathrm{e}^{\frac{-\delta^2\mu}{3}}$ *when* $0<\delta<1$*; and*
(2) $\Pr[X \leq (1-\delta)\mu] \leq \mathrm{e}^{\frac{-\delta^2\mu}{2}}$ *when* $0<\delta<1$*; and*
(3) $\Pr[X \geq (1+\delta)\mu] \leq \mathrm{e}^{\frac{-\delta\mu}{3}}$ *when* $\delta>1$

3 Hardness of representation for Sum of ROPs

Let $X=\{x_1,\ldots,x_{2n}\}, Y=\{y_1,\ldots,y_{2n}\}, Z=\{z_1,\ldots,z_{2n}\}$. Define $\mathcal{D}'$ as a distribution on the functions $\varphi: X \to Y \cup Z$ as follows : For $1 \leq i \leq 2n$,

$$\varphi(x_i) \in \begin{cases} Y & \text{with prob. } \frac{1}{2} \\ Z & \text{with prob. } \frac{1}{2} \end{cases}$$

Observe that $|\varphi(X) \cap Y| = |\varphi(X) \cap Z|$ is not necessarily true. Let F be a binary arithmetic formula computing a polynomial f on the variables $X=\{x_1,\ldots,x_{2n}\}$. Note that any gate with at least one variable as a child can be classified as:

(1) type-A gates : sum gates both of whose children are variables,
(2) type-B gates : product gates both of whose children are variables,
(3) type-C gates : sum gates exactly one child of which is a variable and the other an internal gate; and
(4) type- D gates: product gates exactly one child of which is a variable and the other an internal gate

Given any ROF F, let there be a type-A gates, b type-B, c type-C and d type-D gates in F. Note that $2a+2b+c+d \leq 2n$.

▶ Observation 3. Let F be a binary arithmetic formula computing a polynomial f. Then we can construct a formula F' computing f such that no root to leaf path in F' has two consecutive type-C gates. Therefore, for any binary arithmetic formula F, without the loss of generality we have $c \leq a+b+d$.

Let $\varphi \sim \mathcal{D}'$. Let there be a' gates of type-A that achieve rank-1 under φ and let a'' gates of type-A that achieve rank-2 under φ. Then, $a = a' + a''$.

The following lemma gives an upper bound on the rank of M_{f^φ}. The proof can be found in the appendix.

▶ **Lemma 10.** [1] *Let F be an* ROF *computing an* ROP *f and $\varphi: X \to Y \cup Z$. Then,* $\mathrm{rank}(M_{f^\varphi}) \leq 2^{a''+\frac{a'}{2}+\frac{2b}{3}+\frac{c}{2}}$.

▶ **Lemma 11.** *Let F be a* ROF*. Let there be a type-A gates in F and a' be the number type-A gates in F that achieve* rank-1 *under $\varphi \sim \mathcal{D}$. Then,* $\Pr_{\varphi\sim\mathcal{D}'}\left[\frac{2}{5}a \leq a' \leq \frac{3}{5}a\right] = 1 - 2^{-a/100}$.

Proof. Let v be a type-A gate in F. Then $f_v = x_i + x_j$ for some $i,j \in [N]$. Then $\Pr[\mathrm{rank}(M_{f_v^\varphi}) = 1] = \Pr[(\varphi(x_i),\varphi(x_j) \in Z) \vee (\varphi(x_i),\varphi(x_j) \in Y)] = \frac{1}{2}$. Therefore, $\mu = \mathbb{E}[a'] = a/2$. Applying Theorem 9 (2) and (3) with $\delta = 1/5$, we get the required bounds. ◀

▶ **Lemma 12.** *Let f be an* ROP *on $2n$ variables and $\varphi \sim \mathcal{D}'$. Then with probability at least* $1-2^{-\Omega\left(\frac{n}{\log n}\right)}$, $\mathrm{rank}(M_{f^\varphi}) \leq 2^{n-\frac{n}{15\log n}}$.

[1] A brief outline of the proof of Lemma 10 was suggested by an anonymous reviewer, the details included here for completeness and since the details were worked out completely by the authors.

Proof. Consider the following two cases:

Case 1 : $a + c \geq \frac{2n}{\log n}$. Then either $a \geq \frac{n}{\log n}$ or $c \geq \frac{n}{\log n}$.

(i) Suppose $a \geq \frac{n}{\log n}$, by Lemma 10, we have $\text{rank}(M_{f^\varphi}) \leq 2^{a''+a'/2+2b/3+c/2} \leq 2^{a''+a'/2+b+c/2}$. Since $2a'' + 2a' + 2b + c + d \leq 2n$, $a'' + a'/2 + b + c/2 \leq n - a'/2$. By Lemma 11, $a' \geq \frac{2a}{5} \geq \frac{2n}{5\log n}$ with probability $1 - 2^{-\Omega(\frac{n}{\log n})}$. Therefore, $\text{rank}(M_{f\varphi}) \leq 2^{a''+a'/2+b+c/2} \leq 2^{n-a'/2} \leq 2^{n-\frac{n}{5\log n}}$.

(ii) Suppose $c \geq \frac{n}{\log n}$. By Observation 3, $a + b + d \geq c \geq \frac{n}{\log n}$, then either $a \geq \frac{n}{3\log n}$ or $b \geq \frac{n}{3\log n}$ or $d \geq \frac{n}{3\log n}$.

- If $a \geq \frac{n}{3\log n}$, similar to (i) we have $\text{rank}(M_{f\varphi}) \leq 2^{n-\frac{n}{15\log n}}$ with probability $1 - 2^{-\Omega(\frac{n}{\log n})}$.
- If $b \geq \frac{n}{3\log n}$ by Lemma 10, $\text{rank}(M_{f^\varphi}) \leq 2^{a+2b/3+c/2}$. Since $2a + 2b + c + d \leq 2n$, we have $a + \frac{c}{2} \leq n - b$. Therefore $\text{rank}(M_{f^\varphi}) \leq 2^{n-\frac{b}{3}} \leq 2^{n-\frac{n}{9\log n}} \leq 2^{n-\frac{n}{15\log n}}$.
- If $d \geq \frac{n}{3\log n}$, since $2a + 2b + c + d \leq 2n$, $a + b + \frac{c}{2} \leq n - \frac{d}{2}$. Therefore by Lemma 10 $\text{rank}(M_{f^\varphi}) \leq 2^{a''+a'/2+2b/3+c/2} \leq 2^{a+b+c/2} \leq 2^{n-\frac{d}{2}} \leq 2^{n-\frac{n}{6\log n}} \leq 2^{n-\frac{n}{15\log n}}$.

Case 2 : $a + c < \frac{2n}{\log n}$. Observe that $b \leq n$. By Lemma 10, $\text{rank}(M_{f^\varphi}) \leq 2^{a+2b/3+c} \leq 2^{2n/3+2n/\log n} \leq 2^{n-n/15\log n}$ for large enough n. ◀

The following polynomial was introduced by Raz and Yehudayoff [18].

▶ Definition 5. Let $n \in \mathbb{N}$ be an integer. Let $X = \{x_1, \ldots, x_{2n}\}$ and $\mathcal{W} = \{w_{i,k,j}\}_{i,k,j\in[2n]}$. For any two integers $i, j \in \mathbb{N}$, we define an interval $[i,j] = \{k \in \mathbb{N}, i \leq k \leq j\}$. Let $|[i,j]|$ be the length of the interval $[i,j]$. Let $X_{i,j} = \{x_p \mid p \in [i,j]\}$ and $W_{i,j} = \{w_{i',k,j'} \mid i', k, j' \in [i,j]\}$. For every $[i,j]$ such that $|[i,j]|$ is even we define a polynomial $g_{i,j} \in \mathbb{F}[X, \mathcal{W}]$ as $g_{i,j} = 1$ when $|[i,j]| = 0$ and if $|[i,j]| > 0$ then, $g_{i,j} \triangleq (1 + x_i x_j) g_{i+1,j-1} + \sum_k w_{i,k,j} g_{i,k} g_{k+1,j}$. where x_k, $w_{i,k,j}$ are distinct variables, $1 \leq k \leq j$ and the summation is over $k \in [i+1, j-2]$ such that the interval $[i,k]$ is of even length. Let $g \triangleq g_{1,2n}$.

In the following, we view g as polynomial in $\{x_1, \ldots, x_{2n}\}$ with coefficients from the rational function field $\mathbb{G} \triangleq \mathbb{F}(\mathcal{W})$. The following lemma builds on Lemma 4.3 in [18] and a proof can be found in the appendix.

▶ **Lemma 13.** *Let Let $X = \{x_1, \ldots, x_{2n}\}$, $Y = \{y_1, \ldots, y_{2n}\}$, $Z = \{z_1, \ldots, z_{2n}\}$ and $\mathcal{W} = \{w_{i,k,j}\}_{i,k,j\in[2n]}$ be sets of variables. Suppose $\varphi \sim \mathcal{D}'$ such that $||\varphi(X) \cap Y| - |\varphi(X) \cap Z|| = \ell$. Then for the polynomial g as in Definition 5 we have,* $\text{rank}(M_{g^\varphi}) \geq 2^{n-\ell/2}$.

▶ **Lemma 14.** *For $\mathcal{Q} \in \{Y, Z\}$,* $\Pr_{\varphi\sim\mathcal{D}'}[n - n^{2/3} \leq |\varphi(X) \cap \mathcal{Q}| \leq n + n^{2/3}] \geq 1 - 2^{-\Omega(n^{1/3})}$.

Proof. Proof is a simple application of Chernoff's bound (Theorem 9) with $\delta = 1/n^{1/3}$. ◀

▶ **Corollary 15.** $\Pr_{\varphi\sim\mathcal{D}'}[\text{rank}(M_{g^\varphi}) \geq 2^{n-n^{2/3}}] \geq 1 - 2^{-\Omega(n^{1/3})}$.

Proof. Apply Lemma 13 with $\ell = 2n/n^{1/3} = 2n^{2/3}$ and the probability bound follows from Lemma 14. ◀

Proof of Theorem 1

Proof. Suppose $s < 2^{o(n/\log n)}$. Then by Lemma 12 and union bound, probability that there is an i such that $\text{rank}(M_{f_i^\varphi}) \geq 2^{n-n/15\log n}$ is $s2^{-\Omega(\frac{n}{\log n})} = 2^{-\Omega(\frac{n}{\log n})}$ and hence by Lemma 5, $\text{rank}(M_{g^\varphi}) \leq s2^{n-n/15\log n} \leq 2^{n-n/20\log n}$ with probability $1 - 2^{-\Omega(\frac{n}{\log n})}$ for large enough n. However, by Corollary 15, $\text{rank}(M_{g^\varphi}) \geq 2^{n-n^{2/3}} > 2^{n-n/20\log n}$ with probability at least $1 - 2^{-\Omega(n^{1/3})}$, a contradiction. Therefore, $s = 2^{\Omega(n/\log n)}$. ◀

4 Sum of Products of ROPs

4.1 ROPs under random partition

Throughout the section, let $m \triangleq N^{1/3}$, $N \triangleq n^2$ and $\kappa \triangleq 20 \log n$. Let $X = \{x_{11}, \ldots, x_{nn}\}$ be a set of n^2 variables and $\mathcal{D}$ denote the distribution on the functions $\varphi : X \to Y \cup Z \cup \{0, 1\}$ defined as follows

$$\varphi(x_{ij}) \in \begin{cases} Y & \text{with prob. } \frac{m}{N} \\ Z & \text{with prob. } \frac{m}{N} \\ 1 & \text{with prob. } \frac{\kappa n}{N} \\ 0 & \text{with prob. } 1 - \left(\frac{2m+\kappa n}{N}\right) \end{cases}$$

The following Lemmas show that bottom $\times$ gates do not contribute much to the rank. Proofs can be found in the appendix.

▶ **Lemma 16.** *Let F be a* ROF *and $\varphi \sim \mathcal{D}$. Let $\mathcal{X}$ be a random variable that denotes the number of non-zero multiplication gates at depth 1. Then* $\Pr_{\varphi \sim D}\left[\mathcal{X} > (N^{1/4})\right] \leq 2^{-\Omega(N^{1/4})}$.

▶ **Lemma 17.** *Let F be an* ROF *computing an* ROP *f and $\varphi \sim \mathcal{D}$. Then there exists an* ROF *F' such that every gate in F' at depth-1 is an addition gate, and* $\text{rank}(M_{F^\varphi}) \leq \text{rank}(M_{F'^\varphi}) \times 2^{\mathcal{O}(N^{1/4})}$ *with probability atleast* $1 - 2^{-\Omega(N^{1/4})}$.

Recall that an arithmetic formula F over $\mathbb{Z}$ is said to be monotone if it does not have any node labelled by a negative constant. We have:

▶ **Lemma 18.** *Let F be an* ROF, *and $\varphi \sim \mathcal{D}$. Then there exists a monotone formula G such that* $\text{rank}(M_{F^\varphi}) \leq \mathsf{value}(G)$.

▶ Observation 4. Let F be an ROF and $\varphi \sim \mathcal{D}$. By Lemma 18, we have, $\Pr[\text{rank}(M_{F^\varphi}) > 2^r] \leq \Pr[\mathsf{value}(G) > 2^r]$.

Let F be an ROF and $\varphi \sim \mathcal{D}$. Then by Lemma 8 we have the following corollary,

▶ **Corollary 19.** $\Pr[\text{rank}(M_{F^\varphi}) > 2^r] \leq \Pr[\exists\, u_1, \ldots, u_{\frac{r}{\log N}} \in F^\varphi$ *s.t.* $\forall\, i\ u_i$ *is a rank-*$(1,2)$*-separator*$]$

Now all we need to do is to estimate the probability that a given set of nodes $u_1, \ldots, u_t$ where $t > \frac{r}{\log N}$ are a set of *rank*-$(1,2)$-*separators*.

▶ **Lemma 20.** *F be an* ROF *and let $u_1, \ldots, u_t$ be a set of $+$ gates in F that have product height at least 1 and are not descendants of each other. Suppose $s(F) \leq N^{1/4}$. Then* $\Pr_\varphi[\bigwedge_{i=1}^t u_i$ *is a rank-*$(1,2)$*-separator*$] \leq c^t N^{-5t/6}$, *for some constant $c > 0$.*

Proof. Note that for $1 \leq i \leq t$ $\text{rank}(M_{u_i^\varphi}) = 2$ only if $|\mathsf{var}(u_i^\varphi) \cap Y| \geq 1$ and $|\mathsf{var}(u_i^\varphi) \cap Z| \geq 1$. Therefore $\Pr[u_i$ is a $(1,2)$ separator$] \leq \Pr[|\mathsf{var}(u_i^\varphi) \cap Y| \geq 1$ and $|\mathsf{var}(u_i^\varphi) \cap Z| \geq 1] \leq \Pr[|\mathsf{var}(u_i^\varphi) \cap (Y \cup Z)| \geq 2]$. Let $\ell_{i_1}, \ldots, \ell_{i_{r_i}}$ be the addition gates at depth-1 in the sub-formula rooted at u_i. For $0 \leq i \leq t$, we define $S_i \triangleq \mathsf{var}(\ell_{i_1}) \cup \cdots \cup \mathsf{var}(\ell_{i_{r_i}})$. Then for $0 \leq i \leq t$, $\Pr[\, u_i$ is a $(1,2)$ separator$] \leq \Pr[|S_i \cap (Y \cup Z)| \geq 2]$. Since $|\mathsf{var}(u_i)| \leq s(F)$, we have $|S_i| \leq s(F) \leq N^{1/4}$. Since $(1 - 2m/N)^{|S_i|-2} \leq 1$, $|S_i| \leq N^{1/4}$ and $m = N^{1/3}$, we have

$$\begin{aligned} \Pr[|S_i \cap (Y \cup Z)| = 2] &= \binom{|S_i|}{2} \left(\frac{2m}{N}\right)^2 (1 - 2m/N)^{|S_i|-2} \leq \binom{|S_i|}{2} \left(\frac{2m}{N}\right)^2 \\ &\leq 2^2 s(F)^2 N^{-4/3} = \mathcal{O}(N^{-5/6}). \end{aligned}$$

Similarly, $\Pr[|S_i \cap (Y \cup Z)| = 3] \leq \mathcal{O})(N^{-5/4})$. By union bound $\Pr[|S_i \cap (Y \cup Z)| \geq 3] \leq |Y \cup Z| \Pr[|S_i \cap (Y \cup Z)| = 3] \leq N^{-11/12} \leq \mathcal{O}(N^{-5/6})$. Then for some constant $c > 0$

$$\Pr_{\varphi}\left[\bigwedge_{i=1}^{t} u_i \text{ is a } (1,2) \text{ separator}\right] \leq \prod_{i=1}^{t} \Pr[|S_i \cap (Y \cup Z)| \geq 2] \leq \prod_{i=1}^{t} \mathcal{O}(N^{-5/6}) = c^t N^{-(5t/6)} \blacktriangleleft$$

▶ **Lemma 21.** *Let f be an* ROP*on N variables computed by an* ROF *F, with $s(F) \leq N^{1/4}$. Then,* $\Pr_{\varphi \sim \mathcal{D}}[\text{rank}(M_{f^\varphi}) \geq 2^{N^{4/15}}] \leq 2^{-\Omega(N^{1/4})}$.

Proof. By Lemma 17, note that $\times$ gates in F with at least two variables as their input contribute a multiplicative factor of $2^{N^{1/4}}$ to $\text{rank}(M_{f^\varphi})$ with probability at least $1-2^{-\Omega(N^{1/4})}$. Thus, without loss of generality we can assume that F has not $\times$ gate with at more than two variables as its input. By Corollary 19 we have

$$\begin{aligned}\Pr[\text{rank}(M_{f^\varphi}) \geq 2^{N^{4/15}}] &\leq \Pr[\exists \textit{ rank-}(1,2)\textit{-separators } u_1, \ldots, u_{\frac{N^{4/15}}{\log N}}] \\ &\leq \Pr[\exists \textit{ rank-}(1,2)\textit{-separators } u_1, \ldots, u_{N^{1/4}}] \leq \binom{N}{N^{1/4}} c^{N^{1/4}} N^{-\frac{5}{6}N^{1/4}} \\ &\leq c^{N^{1/4}} e^{N^{1/4}} N^{(3/4)N^{1/4} - (5/6)N^{1/4}} \leq N^{-\Omega(N^{1/4})}.\end{aligned}$$

The penultimate inequality follows by Lemma 20 and union bound. For the last inequality, we use the fact that $\binom{n}{k} \leq (ne/k)^k$, where e is the base of natural logarithm. ◀

4.2 Polynomials with High Rank

In this section, we prove rank lower bounds for two polynomials under a random partition $\varphi \sim \mathcal{D}$. The first one is in VP and the other one is in VNP.

▶ **Lemma 22.** *Let $p_{lin} = \ell_1 \cdots \ell_{m'}$ where $\ell_j = \left(\sum_{i=(j-1)(N/2m)+1}^{jN/2m} x_i\right) + 1$, where $m' = 2m$. Then,* $\text{rank}(M_{p_{lin}^\varphi}) = 2^{\Omega(m)}$ *with probability* $1 - 2^{-\Omega(m)}$.

Proof. Let $p_{lin} = \ell_1 \cdots \ell_{m'}$ where $\ell_j = \left(\sum_{(j-1)(N/2m)+1}^{jN/2m} x_i\right) + 1$ and $m' = 2m$.

Define indicator random variables $\rho_1, \rho_2, \ldots, \rho_{m'}$, where $\rho_i = 1$ if $\text{rank}(M_{\ell_i^\varphi}) = 2$ and 0 otherwise. Observe that for any $1 \leq i \leq m', \text{rank}(M_{\ell_i^\varphi}) = 2$ iff $\ell_i^\varphi \cap Y \neq \emptyset$ and $\ell_i^\varphi \cap Z \neq \emptyset$. Therefore, $\Pr[\text{rank}(M_{\ell_i^\varphi}) = 2] = \Pr[\ell_i^\varphi \cap Y \neq \emptyset \text{ and } \ell_i^\varphi \cap Z \neq \emptyset]$. For any $1 \leq j \leq m'$, $\Pr[\ell_j^\varphi \cap Y \neq \emptyset \text{ and } \ell_j^\varphi \cap Z \neq \emptyset] \geq \frac{N}{2m}\left(\frac{N}{2m} - 1\right)\left(\frac{m}{N}\right)^2\left(1 - \frac{m}{N}\right)^{\frac{N}{2m}-2} \geq 1/16$ for large enough N. Let $\rho = \sum_{i=1}^{m'} \rho_i$. Then by linearity of expectation, $\mu \triangleq \mathbb{E}[\rho] = \sum_{i=1}^{m'} \mathbb{E}[\rho_i] \geq \frac{m}{8}$. Since $\mu \geq m/8$, we have $\Pr[\rho < (1-\delta)m/8] \leq \Pr[\rho < (1-\delta)\mu] = 2^{-\Omega(m)}$ by Theorem 9 with $\delta = 1/4$, since $\text{rank}(M_{p_{lin}^\varphi}) = 2^\rho$. ◀

Throughout the section let φ denote a function of the form $\varphi : X \to Y \cup Z \cup \{0,1\}$. Let X_φ denote the matrix $(\varphi(x_{ij}))_{1 \leq i,j \leq n}$. If and when φ involved in a probability argument, we assume that φ is distributed according to $\mathcal{D}$.

▶ Definition 6. Let $1 \leq i, j \leq n$. (i, j) is said to be a *Y-special* (respectively *Z-special*) if $\varphi(x_{ij}) \in Y$ (respectively $\varphi(x_{ij}) \in Z$), $\forall i' \in [n], i' \neq i \;\; \varphi(x_{i'j}) \in \{0,1\}$ and $\forall j' \in [n], j' \neq j \;\; \varphi(x_{ij'}) \in \{0,1\}$.

The following lemma is an application of Chernoff's bound. Proof can be found in the appendix.

▶ **Lemma 23.** *Let $\mathcal{Q} \in \{Y, Z\}$, φ as above and $\chi = |\varphi(X) \cap \mathcal{Q}|$ where $\varphi(X) = \{\varphi(x_{ij})\}_{i,j \in [n]}$. Then,* $\Pr_{\varphi \sim \mathcal{D}}\left[\frac{3m}{4} < \chi < \frac{5m}{4}\right] = 1 - 2^{-\Omega(m)}$.

Let $C_1, \ldots, C_n$ denote the columns of X_φ and $R_1, \ldots, R_n$ denote the rows of X_φ.

▶ Definition 7. Let $\mathcal{Q} \in \{Y, Z\}$. A column C_j, $1 \le j \le n$ is said to be $\mathcal{Q}$-*good* if $\exists i \in [n], \ \ \varphi(x_{ij}) \in \mathcal{Q}$; and $\forall i' \in [n], i' \neq i \ \ \varphi(x_{i'j}) \in \{0, 1\}$. $\mathcal{Q}$-*good* rows are defined analogously.

▶ Observation 5. Let C_i be a Y-good column in X_φ. Let $i, i' \in [n]$, $\mathcal{R}$ be the event that $\varphi(x_{ij}) \in Y$ and $\mathcal{T}$ be the event that $\varphi(x_{i'j}) \in Y$. The events $\mathcal{R}$ and $\mathcal{T}$ are mutually exclusive.

By Observation 5 and union bound we have:

▶ **Lemma 24.** *For $1 \le i \le n$, let C_i be a column in X_φ. Then for any $\mathcal{Q} \in \{Y, Z\}$,* $\Pr_{\varphi \sim \mathcal{D}}[C_i \textit{ is } \mathcal{Q}\textit{-good}] = n \cdot \frac{m}{N}\left(1 - \frac{2m}{N}\right)^{n-1}$.

For $\mathcal{Q} \in \{Y, Z\}$ let $\eta_\mathcal{Q} \triangleq |\{C_i \mid C_i \text{ is } \mathcal{Q}\text{-good}\}|$ and $\zeta_\mathcal{Q} \triangleq |\{R_j \mid R_j \text{ is } \mathcal{Q}\text{-good}\}$. A proof of following lemma can be found in the appendix.

▶ **Lemma 25.** *With notations as above, $\forall \mathcal{Q} \in \{Y, Z\}$, $\Pr_{\varphi \sim \mathcal{D}}[\eta_\mathcal{Q} \ge \frac{2m}{3}] = 1 - 2^{-\Omega(m)}$; and* $\Pr_{\varphi \sim \mathcal{D}}[\zeta_\mathcal{Q} \ge \frac{2m}{3}] = 1 - 2^{-\Omega(m)}$.

▶ **Lemma 26.** *For $\mathcal{Q} \in \{Y, Z\}$, let $\gamma_\mathcal{Q}$ denote the number of $\mathcal{Q}$-special positions in X_φ. Then* $\forall \mathcal{Q} \in \{Y, Z\}$, $\Pr_{\varphi \sim \mathcal{D}}\left[\gamma_\mathcal{Q} \ge \frac{m}{12}\right] = 1 - 2^{-\Omega(m)}$.

Proof. We argue for $\mathcal{Q} = Y$, the proof is analogous when $\mathcal{Q} = Z$. Let φ be distributed according to $\mathcal{D}$. Consider the following events on X_φ. E1 : $2m/3 \le |X_\varphi \cap Y| \le 5m/4$;
E2 : The number of Y-good columns and Y-good rows is at least $r \triangleq 2m/3$.

By Lemmas 23 and 25, X_φ satisfies the events E1 and E2 with probability $1 - 2^{-\Omega(m)}$. Henceforth we assume that X_φ satisfies the events E1 and E2.

Without loss of generality, let $R_1, \ldots, R_r$ be the first r Y-good rows in X_φ. For every Y-good row R_i, $1 \le i \le r$ there exists a corresponding witness column $C_j, j \in [n]$ such that $\varphi(x_{ij}) \in Y$. Without loss of generality, assume $C_1, \ldots, C_r$ be columns that are witnesses for $R_1, \ldots, R_r$ being Y-good. Further, let $X_\varphi(C_j)$ denote the set of values along the column C_j.

Suppose among $C_1, \ldots, C_r$, $t \ge 0$ columns are not Y-good, without loss of generality let them be $C_1, C_2, \ldots, C_t$.

Each of the column C_j has at least one variable from Y and hence the columns $C_1, \ldots, C_t$ contain at least t distinct variables from Y. By event E2, there are at least $\frac{2m}{3}$ Y-good columns that are distinct from $C_1, \ldots, C_t$, each containing exactly one distinct variable from Y. Since the total number of variables from Y in X_φ is at most $5m/4$ (by E1) we have, $t \le \frac{5m}{4} - \frac{2m}{3} \le \frac{7m}{12}$. That is, at most $7m/12$ of the columns among $C_1, \ldots, C_r$ are not Y-good. Therefore, at least $r - t$ of the columns among $C_1, \ldots, C_r$ are Y good and hence the number of Y-special positions in X_φ is atleast $r - t \ge (2/3 - 7/12)m = \frac{m}{12}$. We conclude, $\Pr_{\varphi \sim \mathcal{D}}\left[\gamma_Y \ge \frac{m}{12}\right] = 1 - 2^{-\Omega(m)}$. ◀

A row R in the matrix $A \in (Y \cup Z \cup \{0, 1\})^{n \times n}$ said to be 1-*good* if there is at least one 1 in R in a column other than Y-special and Z-special positions. The following observation is immediate :

▶ Observation 6. Let φ be distributed according to $\mathcal{D}$. Then for any row (column) R: $\Pr_{\varphi \sim \mathcal{D}}[\text{R is 1-good}] \ge (1 - 1/n^3)$.

Finally, we are ready to show that perm has high rank under a random $\varphi \sim \mathcal{D}$.

▶ **Theorem 27.** $\Pr[\mathrm{rank}(M_{\mathsf{perm}_n^\varphi}) \geq 2^{m/12}] \geq (1 - O(1/n^2))/2$.

We need a few notations and Lemmas before proving Theorem 27. Consider a $\varphi : X \to Y \cup Z \cup \{0,1\}$ and let the number of Y-special positions and the number of Z-special positions in X_φ are both be at least γ. Let $(i_1, j_1), (i_2, j_2), \ldots, (i_\gamma, j_\gamma)$ be a set of distinct Y- special positions that do not share any row or column and $(k_1, \ell_1), (k_2, \ell_2), \ldots, (k_\gamma, \ell_\gamma)$ be a set of distinct Z - special positions in X_φ that do not share any row or column.

Without loss of generality, suppose $i_1 < i_2 < \cdots < i_\gamma$ and $k_1 < k_2 < \cdots < k_\gamma$. Let $\mathcal{M}$ be the perfect matching $((i_1, j_1), (k_1, \ell_1)), \ldots, ((i_\gamma, j_\gamma), (k_\gamma, \ell_\gamma))$.
For an edge $\{(i_p, j_p), (k_p, \ell_p)\} \in \mathcal{M}$, $1 \leq p \leq \gamma$ consider the 2×2 matrix :

$$B_p = \begin{pmatrix} X_\varphi[i_p, j_p] & X_\varphi[i_p, \ell_p] \\ X_\varphi[k_p, j_p] & X_\varphi[k_p, \ell_p] \end{pmatrix}.$$

There exists a partition $\varphi : X \to Y \cup Z \cup \{0,1\}$ such that $\mathrm{rank}(M_{B_p^\varphi}) = 2$. Let A be the matrix obtained by permuting the rows and columns in X_φ such that A can be written as in the Figure 1 below. Since (i_p, j_p) is a Y-special position, (k_p, ℓ_p) is a Z-special position we

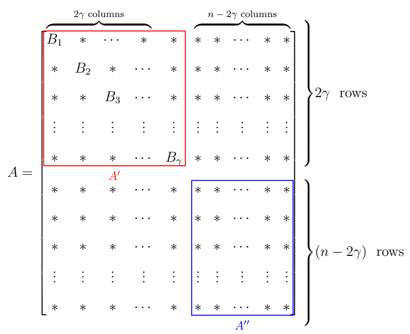

Figure 1 The matrix A after permuting the rows and columns. $*$ denotes unspecified entries.

have $X_\varphi[i_p, j_p] \in Y$, $X_\varphi[k_p, \ell_p] \in Z$. Also $X_\varphi[i_p, \ell_p] \in \{0,1\}$ and $X_\varphi[k_p, j_p] \in \{0,1\}$. Further, $\mathrm{rank}(M_{\mathsf{perm}(B_p)}) = 2$ if and only if $X_\varphi[k_p, j_p] = X_\varphi[i_p, \ell_p] = 1$. Consider the following events: F_1: $\gamma \geq m/12$; and F_2: Rows $i_1, \ldots, i_\gamma, k_1 \ldots, k_\gamma$ are 1-good. The following lemma estimates the probability of $\mathsf{perm}(A'') \neq 0$. Proof can be found in the appendix.

▶ **Lemma 28.** *Let A'' be matrix as in Figure 1. Then $\Pr_\varphi[\mathit{perm}(A'') \neq 0 \mid F_1, F_2] \geq 1 - \frac{1}{n^2}$.*

Let F_3 denote the event "$\mathsf{perm}(A'') \neq 0$". Define sets of matrices:

$$\mathcal{A} \triangleq \left\{ X_\varphi \mid \begin{matrix} X_\varphi \in F_1 \cap F_2 \cap F_3 \text{ and } \exists i \leq \\ \gamma, \mathrm{rank}(M_{\mathsf{perm}(B_i)}) = 1 \end{matrix} \right\}; \quad \mathcal{B} \triangleq \left\{ X_\varphi \mid \begin{matrix} X_\varphi \in F_1 \cap F_2 \cap F_3 \text{ and } \forall i \leq \\ \gamma, \mathrm{rank}(M_{\mathsf{perm}(B_i)}) = 2. \end{matrix} \right\}$$

▶ Observation 7. $\forall A \in \mathcal{A}$, $\mathrm{rank}(M_{\mathsf{perm}(A')}) < 2^\gamma$ and $\forall B \in \mathcal{B}, \mathrm{rank}(M_{\mathsf{perm}(B)}) \geq 2^\gamma$.

▶ **Lemma 29.** *Let $\mathcal{A}$ and $\mathcal{B}$ as defined above. Then (a) $\Pr_{\varphi \sim \mathcal{D}}[\mathrm{rank}(M_{\mathsf{perm}(X_\varphi)}) \geq 2^\gamma)] \geq \mathcal{D}(\mathcal{B})$; and (b) $\mathcal{D}(B) \geq \mathcal{D}(\mathcal{A})$, where $\mathcal{D}(S) = \Pr_{\varphi \sim \mathcal{D}}[X_\varphi \in S]$ for $S \in \{\mathcal{A}, \mathcal{B}\}$.*

Proof. (a) follows from Observation 7. For (b), we establish a one-one mapping $\pi : \mathcal{A} \to \mathcal{B}$ defined as follows. Let φ be such that $X_\varphi \in \mathcal{A}$. Consider $1 \leq p \leq \gamma$ such that $\mathrm{rank}(M_{\mathsf{perm}(B_p)}) = 1$. Then either $X_\varphi[k_p, j_p] = 0$ or $X_\varphi[i_p, \ell_p] = 0$ or both. If $X_\varphi[k_p, j_p] = 0$, then set $X_{\varphi'}[k_p, j_p] = 1$, and $X_{\varphi'}[k_p, \iota_p] = 0$ where $\iota_p \in [n] \setminus \{j_1 \ldots, j_\gamma, \ell_1 \ldots, \ell_\gamma\}$ is the first index from left such that $X_\varphi[k_p, \iota_p] = 1$. Similarly, if $X_\varphi[i_p, \ell_p] = 0$, then set $X_{\varphi'}[i_p, \ell_p] = 1$, and $X_{\varphi'}[i_p, \lambda_p] = 0$ where $\lambda_p \in [n] \setminus \{j_1 \ldots, j_\gamma, \ell_1 \ldots, \ell_\gamma\}$ is the first index from left such that $X_\varphi[k_p, \lambda_p] = 1$. Let φ' be the partition obtained from φ by applying the above mentioned swap operation for every $1 \le p \le \gamma$ with $\text{rank}(M_{\mathsf{perm}(B_p)}) = 1$, keeping other values of φ untouched. Clearly $X_{\varphi'} \in \mathcal{B}$. Set $\pi(X_\varphi) \mapsto X_{\varphi'}$. It can be seen that π is an one-one map. Further, for any fixed $A \in \mathcal{A}$, $\Pr_\varphi[X_\varphi = A] = \Pr_\varphi[X_\varphi = \pi(A)]$ since φ is independently and identically distributed for any position in the matrix. Thus we have $\mathcal{D}(\mathcal{A}) \le \mathcal{D}(\mathcal{B})$. ◀

Proof of Theorem 27. It is enough to argue that $\Pr_{\varphi\sim\mathcal{D}}[X\varphi \in \mathcal{A}\cup\mathcal{B}] = 1 - O(\frac{1}{n^2})$, as $\mathcal{A}\cap\mathcal{B} = \emptyset$. Now, $\Pr_{\varphi\sim\mathcal{D}}[X_\varphi \in \mathcal{A}\cup\mathcal{B}] = \Pr_{\varphi\sim\mathcal{D}}[F_1 \cap F_2 \cap F_3]$. By Lemma 26, $Pr_{\varphi\sim\mathcal{D}}[F_1] = 1 - 2^{-\Omega(m)}$. Form Observation 6 combined with union bound we have $Pr_{\varphi\sim\mathcal{D}}[F_2] \ge 1 - \gamma/n^3$ and by Lemma 28, $Pr_{\varphi\sim\mathcal{D}}[F_3 | F_1, F_2] \ge 1 - 2/n^2$. Thus we conclude $\Pr_{\varphi\sim\mathcal{D}}[F_1 \cap F_2 \cap F_3] = 1 - O(\frac{1}{n^2})$. As $\mathcal{D}(\mathcal{B} \cup \mathcal{A}) = \mathcal{D}(\mathcal{A}) + \mathcal{D}(\mathcal{B})$, by Lemma 29 we have $\Pr_{\varphi\sim\mathcal{D}}[\text{rank}(M_{\mathsf{perm}(X_\varphi)}) \ge 2^\gamma] \ge 1/2(1 - O(\frac{1}{n^2}))$. ◀

4.3 Putting them all together

Proof of Theorem 2

Proof. Suppose $p_{lin} = \sum_{i=1}^{s} \prod_{j=1}^{t} f_{i,j}$ where $f_{i,j}$ are syntactically multi-linear $\Sigma\Pi\Sigma$ formula, with $s < 2^{N^{1/4}}$, Let $f_{i,j} = \sum_{k=1}^{s'} T_{i,j,k}$, and $T_{i,j,k}$ are products of variable disjoint linear forms, and hence ROPs. Further, since the bottom fan-in of each $f_{i,j}$ is bounded by $N^{1/4}$, we have $s_{T_{i,j,k}} \le 2^{N^{1/4}}$. Then by Lemma 21 and union bound there is an i, j, k such that $\text{rank}(M_{T^\varphi_{i,j,k}}) \ge 2^{N^{4/15}}$ with probability at most $sts'2^{-\Omega(N^{1/4})}$. By Lemma 5 and 6, we have $\text{maxrank}(M_{p^\varphi_{lin}}) \le 2^{N^{4/15}}$ with probability $1 - o(1)$. However by Lemma 22, $\text{maxrank}(M_{p^\varphi_{lin}}) = \text{rank}(M_{p^\varphi_{lin}}) = 2^{\Omega(m)}$ with probability at least $1 - 2^{-\Omega(m)}$, a contradiction. Hence $ss' = 2^{\Omega(N^{1/4})}$. ◀

Proof of Theorem 3

Proof. Suppose $s = 2^{o(N^{1/4})}$. Then by Lemma 21, the probability that there is an $f_{i,j}$ with $\text{rank}(M_{f^\varphi_{i,j}}) \ge 2^{N^{4/15}}$ is at most $2^{-\Omega(N^{1/4})}s = o(1)$. By Lemma 5 and 6 and since $\text{maxrank}(M_{f^\varphi_{i,j}}) = \text{rank}(M_{f^\varphi_{i,j}})$, we have $\text{maxrank}(M_{p^\varphi_{lin}}) \le (s \cdot 2^{N^{4/15}})^{N^{1/30}} = 2^{o(N^{1/3})}$ with probability $1 - o(1)$. However by Lemma 22, $\text{maxrank}(M_{p^\varphi_{lin}}) = \text{rank}(M_{p^\varphi_{lin}}) = 2^{\Omega(m)}$ with probability $1 - 2^{-\Omega(m)}$, a contradiction. Hence $s = 2^{\Omega(N^{1/4})}$. ◀

Proof of Theorem 4

Proof. Suppose $s = 2^{o(N^{1/4})}$. Then by Lemma 21, Probability that there is an $f_{i,j}$ with $\text{rank}(M_{f^\varphi_{i,j}}) \ge 2^{N^{4/15}}$ is at most $2^{-\Omega(N^{1/4})}s = o(1)$. Then, by Lemma 5 and 6, we have $\text{maxrank}(M_{\mathsf{perm}^\varphi_n}) \le s \cdot (2^{N^{4/15}})^{N^{1/30}} = 2^{o(N^{1/3})}$ with probability $1 - o(1)$. However, by Theorem 27, $\text{maxrank}(M_{\mathsf{perm}^\varphi_n}) = \text{rank}(M_{\mathsf{perm}^\varphi_n}) = 2^{\Omega(m)}$ with probability $(1 - 1/n^2)/2$, a contradiction. Hence $s = 2^{\Omega(N^{1/4})}$. ◀

Acknowledgements: We thank anonymous reviewers of an earlier version of the paper for suggestions which improved the presentation. Further, we thank one of the anonymous reviewers for pointing an observation that lead to Lemma 10.

References

1. Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *FOCS*, pages 67–75, 2008.
2. Walter Baur and Volker Strassen. The complexity of partial derivatives. *Theor. Comput. Sci.*, 22:317–330, 1983.
3. Devdatt Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms.* Cambridge University Press, New York, NY, USA, 1st edition, 2009.
4. Michael Forbes. Polynomial identity testing of read-once oblivious algebraic branching programs. *PhD thesis, Massachusetts Institute of Technology*, 2014.
5. Michael A. Forbes and Amir Shpilka. Quasipolynomial-time identity testing of non-commutative and read-once oblivious algebraic branching programs. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 243–252, 2013.
6. Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *STOC*, pages 577–582, 1998.
7. Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *J. ACM*, 61(6):33:1–33:16, 2014.
8. Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012.
9. Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. Super-polynomial lower bounds for depth-4 homogeneous arithmetic formulas. In *STOC*, pages 119–127, 2014.
10. Neeraj Kayal and Chandan Saha. Multi-k-ic depth three circuit lower bound. In *STACS*, pages 527–539, 2015.
11. Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. An almost cubic lower bound for depth three arithmetic circuits. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:6, 2016.
12. Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012.
13. Mrinal Kumar, Gaurav Maheshwari, and Jayalal Sarma. Arithmetic circuit lower bounds via maxrank. In *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part I*, pages 661–672, 2013.
14. Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *FOCS*, pages 364–373, 2014.
15. Meena Mahajan and Anuj Tawari. Sums of read-once formulas: How many summands suffice? *Electronic Colloquium on Computational Complexity (ECCC)*, 22:204, 2015.
16. Noam Nisan and Avi Wigderson. Lower bounds for arithmetic circuits via partial derivatives (preliminary version). In *FOCS*, pages 16–25, 1995.
17. Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 633–641, 2004.
18. Ran Raz and Amir Yehudayoff. Balancing syntactically multilinear arithmetic circuits. *Computational Complexity*, 17(4):515–535, 2008.
19. Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Computational Complexity*, 18(2):171–207, 2009.
20. Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. In *STOC*, pages 507–516, 2008.
21. Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001.

22 Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013.
23 Leslie G. Valiant. Completeness classes in algebra. In *STOC*, pages 249–261, 1979.
24 J. H. van Lint and R. M. Wilson. *A Course In Combinatorics*. Cambridge University Press, 2nd edition, 2001.

A Proofs from Section 2

A.1 Proof of Lemma 7

Proof. The proof is by induction on the size of the formula. Base Case : $s = 1$

- If G has a single $+$ gate then $\mathsf{value}(G) \leq N + N \leq N^2$.
- If G has a single $\times$ gate then $\mathsf{value}(G) \leq N \cdot N = N^2$.

Induction Step : Let u be the output gate of G with children u_1 and u_2. Let the number of leaves in the sub formula rooted at u_1 and u_2 be t_1 and t_2 respectively.

- If u is a $+$ gate. Then, $\mathsf{value}(u) = \mathsf{value}(u_1) + \mathsf{value}(u_2)$. By induction hypothesis, $\mathsf{value}(u) \leq N^{t_1} + N^{t_2} \leq N^{t_1+t_2} \leq N^t$.
- If u is a $\times$ gate. Then, $\mathsf{value}(u) = \mathsf{value}(u_1) \times \mathsf{value}(u_2)$. By induction hypothesis, $\mathsf{value}(u) \leq N^{t_1} \times N^{t_2} \leq N^{t_1+t_2} \leq N^t$. ◀

A.2 Proof of Lemma 8

Proof. Let F be a binary monotone arithmetic formula with leaves labelled by either 1 or 2. The statement trivially holds when $\mathsf{value}(F) = 0$ or $\mathsf{value}(F) = 1$. Now suppose $\mathsf{value}(F) > 2$, First mark every gate u such that u is a *rank*-$(1,2)$-*separator* and remove sub-formula rooted at u except u. Consider any leaf v that remains unmarked. Then $\mathsf{value}(v) = 1$ and along the path from v to root there is no gate that is marked. Else v would have been removed. Consider the unique path from v to root in F. Let p be the first gate in the path such that $\mathsf{value}(p) > 2$. Since $\mathsf{value}(F) > 2$, such a gate p must exist. Let p_1 and p_2 be the children of p. Without loss of generality let p_1 be an ancestor of v. Since v was not removed and $\mathsf{value}(p) > 2$, we have $\mathsf{value}(p_2) \geq 2$. Therefore, there is atleast one marked node(say q) in the sub-formula rooted at p_2. Set $\mathsf{value}(q) = \mathsf{value}(q) + 1$ and remove v from the F, and make necessary short-circuiting of the parent of v. Repeat this process until every unmarked leaf in the formula is removed. Let $u_1, \ldots, u_t$ be the leaves of the resulting formula at the end of this process. For every $1 \leq i \leq t$, we have $2 \leq \mathsf{value}(u_i) \leq N$. By Lemma 7, $\mathsf{value}(F) \leq N^t$ and hence $2^r < N^t$. Therefore $t > \frac{r}{\log N}$ as required. ◀

B Proofs from Section 3

B.1 Proof of Lemma 10

Proof. Observe that for any type-D gate $g = h \times x$, $\text{rank}(M_{g^\varphi}) = \text{rank}(M_{(x \cdot h)^\varphi}) \leq \text{rank}(M_{h^\varphi})$, and hence type-$D$ gates do not contribute to the rank.

The proof is by induction on the structure of F. Let r be the root gate of F. Base case is when F has depth 1. Then,

- r is an type-A gate with children x_1, x_2 : $f = x_1 + x_2$. For any φ, $\text{rank}(M_{f^\varphi}) \leq 2$. Then $a = 1, b = 0, c = 0$. Therefore either $a' = 1$ or $a'' = 1$. In either case, $\text{rank}(M_{f^\varphi}) \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$.

- r is a type-B gate with children x_1, x_2 : $f = x_1 \cdot x_2$. For any φ, $\text{rank}(M_{f^\varphi}) \leq 1$. Then $a = 0, b = 1, c = 0$. Therefore $\text{rank}(M_{f^\varphi}) \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$.

For the induction step, we have the following cases based on the structure of F.

- r is a type-C gate with children x, h, i.e., $f = h + x$. For any φ, we have by sub-additivity $\text{rank}(M_{f^\varphi}) \leq \text{rank}(M_{h^\varphi}) + \text{rank}(M_{x^\varphi})$. Let a'_h, a''_h be the number of type-A gates in the sub-formula rooted at h that achieve rank-1 and rank-2 under φ respectively. Let b_h, c_h be the number of type-B and c type-C gates in the sub-formula rooted at h. We now have $a' = a'_h, a'' = a''_h, b = b_h, c = c_h + 1$, and $\text{rank}(M_{f^\varphi}) \leq \text{rank}(M_{h^\varphi}) + \text{rank}(M_{x^\varphi})$. By Induction hypothesis $\text{rank}(M_{h^\varphi}) \leq 2^{a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2}}$. First suppose the case when $a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2} \geq 1.5$, then, $\text{rank}(M_{f^\varphi}) \leq 2^{a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2}} + \text{rank}(M_{x^\varphi}) = 2^{a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2}} + 1 \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$. Now suppose $a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2} < 1.5$, observe that $a''_h \leq 1$ and $a'_h, b_h, c_h \leq 2$. Consider the following cases :
  - If $b_h = 2$, as $a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2} < 1.5$, we have $a'_h, a''_h, c_h = 0$. Therefore, when $b_h = 2$, $\text{rank}(M_f^\varphi) \leq 2 \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$.
  - If $a'_h = 2$ as $a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2} < 1.5$, we have $a''_h, b_h, c_h = 0$. In that case, $\text{rank}(M_f^\varphi) \leq 2 \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$.
  - If $c'_h = 2$ as $a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2} < 1.5$, we have $a''_h, a'_h, b_h = 0$. Such a formula cannot exist.
  - If $a''_h = 1$ then we have $a'_h = 0, b_h = 0, c_h = 0$ as $a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2} < 1.5$. In this case, $\text{rank}(M_f^\varphi) \leq 2 \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$.
  - Now the only remaining cases are $a''_h = 0$ and $a'_h, b_h, c_h \leq 1$. If $a''_h = 0$ then atmost two of a'_h, b_h, c_h can be non-zero as $a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2} < 1.5$. In any case, $\text{rank}(M_f^\varphi) \leq 2 \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$.
- $r = g * h$ be an internal gate with $* \in \{+, \times\}$. For $H \in \{g, h\}$, let a'_H, a''_H be the number of type-A gates that achieve rank-1 and rank-2 under φ respectively and b_H, c_H be the number of type-B and c type-C gates in the sub-formula rooted at H. Then, $\text{rank}(M_{f^\varphi}) \leq \text{rank}(M_{g^\varphi}) \cdot \text{rank}(M_{h^\varphi})$, and from Induction hypothesis $\text{rank}(M_{f^\varphi}) \leq 2^{a''_g + \frac{a'_g}{2} + \frac{2b_g}{3} + \frac{c_g}{2}} 2^{a''_h + \frac{a'_h}{2} + \frac{2b_h}{3} + \frac{c_h}{2}}$. Since $a' = a'_g + a'_h, a'' = a''_g + a''_h, b = b_h + b_g, c = c_g + c_h$ we have $\text{rank}(M_{f^\varphi}) \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$. ◀

B.2 Proof of Lemma 13

Proof. The proof builds on Lemma 4.3 in [18] as a base case and is by induction on $n + \ell$.
Base case: Either $\ell = 0$ or $\ell = 2n$. For $\ell = 0$, the statement follows by Lemma 4.3 in [18]. When $\ell = 2n$, then $\text{rank}(M_{g^\varphi}) = 1 = 2^{n - \ell/2}$.
Induction step: Without loss of generality, assume that $|\varphi(X) \cap Y| = |\varphi(X) \cap Z| + \ell$. There are three possibilities:
Case 1 : Let $\varphi(x_1) \in Y$ and $\varphi(x_{2n}) \in Z$ or vice versa. In this case

$$\begin{aligned} \text{rank}(M_{g^\varphi}) &\geq \text{rank}(M_{(1+x_1x_{2n})^\varphi})\,\text{rank}(M_{g^\varphi_{2,2n-1}}) = 2 \cdot \text{rank}(M_{g^\varphi_{2,2n-1}}) \\ &\geq 2 \cdot 2^{n-1-\ell/2} = 2^{n-\ell/2} \qquad \text{[By Induction Hypothesis.]} \end{aligned}$$

Case 2 : $\varphi(x_1) \in Y$ and $\varphi(x_{2n}) \in Y$. Then

$$\begin{aligned} \text{rank}(M_{g^\varphi}) &\geq \text{rank}(M_{(1+x_1x_{2n})^\varphi})\,\text{rank}(M_{g^\varphi_{2,2n-1}}) = 1 \cdot \text{rank}(M_{g^\varphi_{2,2n-1}}) \\ &\geq 2^{(2n-2)/2-(\ell-2)/2} = 2^{n-\ell/2}. \qquad \text{[By Induction Hypothesis.]} \end{aligned}$$

For the penultimate inequality above, note that $g_{2,2n-1}$ is defined on $X' = \{x_2, \dots, x_{2n-1}\}$ and $||\varphi(X') \cap Y| - |\varphi(X') \cap Z|| = \ell - 2$ and hence by Induction Hypothesis, $\text{rank}(M_{g^{\varphi}_{2,2n-1}}) \geq 2^{(2n-2)/2-(\ell-2)/2}$.

Case 3 $\varphi(x_1) \in Z$ and $\varphi(x_{2n}) \in Z$. Then there is an $i \in \{2, 2n-1\}$ such that $||\varphi(X_i) \cap Y| - |\varphi(X_i) \cap Z|| = 0$ and $||\varphi(X \setminus X_i) \cap Y| - |\varphi(X \setminus X_i) \cap Z|| = \ell$, where $X_i = \{x_1, \dots, x_i\}$. Then by the definition of g, over $\mathbb{G}$, $\text{rank}(M_{g^{\varphi}}) \geq \text{rank}(M_{g^{\varphi}_{1,i}}) \cdot \text{rank}(M_{g^{\varphi}_{i+1,2n}}) \geq 2^{i/2} \cdot 2^{(2n-i)/2-\ell/2} = 2^{n-\ell/2}$, since $\text{rank}(M_{g^{\varphi}_{1,i}}) = 2^{i/2}$ by Lemma 4.3 in [18], and $\text{rank}(M_{g^{\varphi}_{i+1,2n}}) \geq 2^{(2n-i)/2-\ell/2}$ by Induction Hypothesis. ◀

C Proofs from Section 4

C.1 Proof of Lemma 16

Proof. Consider a multiplication gate g in F at depth 1, with at least two variables as its input. Let m be the monomial (excluding the coefficient) computed by g, note that $d = \deg(m) \geq 2$. we have,

$$\Pr_{\varphi \sim D}[m^{\varphi} \neq 0] = \left(\frac{2m + \kappa n}{N}\right)^d \leq \left(\frac{2m + \kappa n}{N}\right)^2 \leq \left(\frac{2\kappa n}{N}\right)^2 \leq \left(\frac{2\kappa}{n}\right)^2 \leq \mathcal{O}\left(\frac{\kappa^2}{N}\right). \quad (1)$$

In the above, we have used the fact that $2m < \kappa n$ for large enough n. Corresponding to every product gate in F computing the monomial m_i, we define an indicator random variable $\mathcal{Y}_i$

$$\mathcal{Y}_i = \begin{cases} 1 & \text{if } m_i^{\varphi} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

By Equation 1, $\Pr[\mathcal{Y}_i = 1] \leq \frac{c\kappa^2}{N}$ where c is the constant hidden in the $\mathcal{O}$-notation. Let F have r product gates $\left(r \leq \frac{N}{2}\right)$ and $\mathcal{X} = \mathcal{Y}_1 + \mathcal{Y}_2 + \cdots + \mathcal{Y}_r$. Note that $\mathcal{Y}_i$ are independent random variables and $\mathbb{E}[\mathcal{X}] \leq r\frac{c\kappa^2}{N}$. Without loss of generality, assume $\mathsf{E}[\mathcal{X}] \neq 0$, else $r = 0$ and hence $\mathcal{X} = 0$. Choosing $\delta = (N^{1/4} - \mathsf{E}[\mathcal{X}])/\mathsf{E}[\mathcal{X}]$ and applying Theorem 9 (3)

$$\Pr_{\varphi \sim D}\left[\mathcal{X} > N^{1/4}\right] \leq 2^{-\delta \mathsf{E}[\mathcal{X}]/3} \leq 2^{\frac{-N^{1/4} + \mathsf{E}[\mathcal{X}]}{3}} = 2^{-\Omega(N^{1/4})}.$$

◀

C.2 Proof of Lemma 17

Proof. Given an arithmetic formula F we construct the formula F' by replacing every multiplication gate v at depth-1 in F by the constant 1. Let $\mathcal{X}$ the random variable as defined in the proof of Lemma 16. Then, by the construction of F',

$$\text{rank}(M_{F^{\varphi}}) \leq \text{rank}(M_{F'^{\varphi}}) \times 2^{\mathcal{X}}.$$

Now by Lemma 16, with probability atleast $1 - 2^{-\Omega(N^{1/4})}$ we have,

$$\text{rank}(M_{F^{\varphi}}) \leq \text{rank}(M_{F'^{\varphi}}) \times 2^{\mathcal{O}(N^{1/4})}.$$

◀

C.3 Proof of Lemma 18

Proof. Let F be an constant-minimal ROF, and $\varphi \sim \mathcal{D}$. Let G be a monotone formula obtained from F^{φ} as follows:
By short circuiting the gates if necessary, every leaf node v labelled by a constant is replaced by 1. For every gate v in F^{φ} with at least one leaf as a child,

- If $v = \prod_{j=1}^{k} v_j$, with $v_1, \ldots, v_i$, $i \geq 1$ are non-constant leaf gates, then replace the gates $v_1 \times v_2 \times \ldots \times v_i$ by the rank of the polynomial computed by $\varphi(v_1 \times v_2 \times \ldots \times v_i)$.
- Similarly, if $v = \sum_{j=1}^{k} v_j$, with $v_1, \ldots, v_i$, $i \geq 1$ are non-constant leaf gates, then replace the gates $v_1 + v_2 + \ldots + v_i$ by the rank of the polynomial computed by $\varphi(v_1 + v_2 + \ldots + v_i)$.

Clearly, the formula constructed above is monotone, since negative constants (if any) in F^{φ} have been replaced by 1. Then, by Lemmas 5 and 6, we have for any φ, $\text{rank}(M_{F^{\varphi}}) \leq \mathsf{value}(G)$. ◀

C.4 Proof of Lemma 23

Proof. Define indicator random variables χ_{ij} for $1 \leq i, j \leq n$:

$$\chi_{ij} = \begin{cases} 1 & \text{if } \varphi(x_{ij}) \in \mathcal{Q} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\chi = \sum_{i=1}^{n} \sum_{j=1}^{n} \chi_{ij}$ and $\mathbb{E}_{\varphi \sim \mathcal{D}}[\chi] = m$. Let $\delta = \frac{1}{4}$, then by Chernoff bounds in Theorem 9,

$$\Pr\left[\chi \geq \frac{5m}{4}\right] \leq \mathrm{e}^{-\frac{\delta^2 \mu}{3}} \leq \mathrm{e}^{-\frac{m}{48}} = 2^{-\Omega(m)};\ \text{and} \Pr\left[\chi \leq \frac{3m}{4}\right] \leq \mathrm{e}^{-\frac{\delta^2 \mu}{2}} \leq \mathrm{e}^{-\frac{m}{32}} = 2^{-\Omega(m)}.$$

Therefore, $\Pr_{\varphi \sim \mathcal{D}}\left[\frac{3m}{4} < \chi < \frac{5m}{4}\right] = 1 - 2^{-\Omega(m)}$. ◀

C.5 Proof of Lemma 25

Proof. Proof is a simple application for Chernoff's bound. We argue for the case of η_Y, the rest are analogous. For $1 \leq i \leq n$, let

$$\eta_i = \begin{cases} 1 & \text{if } C_i \text{ is Y-good column} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\eta_Y = \eta_1 + \cdots + \eta_n$ and by Observation 5 and Lemma 24 $\mathbb{E}[\eta_i] = \Pr[C_i \text{ is Y-good}] = n \cdot \frac{m}{N}\left(1 - \frac{2m}{N}\right)^{n-1}$. By linearity of expectation, $\mathbb{E}[\eta_Y] = n^2 \cdot \frac{m}{N}\left(1 - \frac{2m}{N}\right)^{n-1} = m\left(1 - \frac{2m}{N}\right)^{n-1}$ as $N = n^2$.

Set $\rho = \left(1 - \frac{2m}{N}\right)^{n-1}$ so that $\mathbb{E}[\eta_Y] = \rho m$. For $\delta = \frac{1}{4}$, we have by Theorem 9,

$$\Pr\left[\eta_Y \leq \left(1 - \frac{1}{4}\right)\rho m\right] \leq \mathrm{e}^{\frac{-(1/4)^2 \mu}{2}} \leq \mathrm{e}^{-\mu/32}.$$

As $m = o(n)$ and $N = n^2$, $\lim_{n \to \infty} \frac{2m}{N} = 0$. Thus for sufficiently large n, $\rho \geq 9/10$ and hence $\mu \geq 9m/10$. We conclude $Pr\left[\eta_Y \leq 27m/40\right] \leq 2^{-\Omega(m)}$. Since $27/40 > 2/3$ we have $\Pr[\eta_Y \geq \frac{2m}{3}] \geq 1 - 2^{-\Omega(m)}$ as required. ◀

C.6 Proof of Lemma 28

Proof. Permanent of any matrix M with entries from $Y \cup Z \cup \{0,1\}$ is zero if and only if M has an all zero $s \times t$ sub matrix such that $s+t = n+1$.(See Theorem 12.1 in [24].) We begin with a bound on the probability that there is at least one column/row with all zero entries. Note that under the event F_1 one can assume that the entries of the matrix A'' are in $\{0,1\}$, and the event F_2 is independent of the rows and columns of A''. Thus, for any position (i,j) in A'', we have $\Pr[\varphi(x_{i,j}) = 1 | F_1, F_2] = \kappa n/N(1-2m/N) \approx \kappa n/N$, for large enough n. Let U and V respectively denote the set of row and column indices of A''. Thus,

$$\Pr[\forall j \in V, \varphi(x_{ij}) = 0 | F_1, F_2] \leq \left(1 - \frac{\kappa n}{N}\right)^{n-2\gamma} \text{ and hence,}$$

$$\Pr[\exists i \in U\ \forall j \in V, \varphi(x_{ij}) = 0 | F_1, F_2] \leq n \cdot \left(1 - \frac{\kappa n}{N}\right)^{n-2\gamma} \quad \text{by union bound}$$

Since $\gamma = \mathcal{O}(m) = o(n)$ and $N = n^2$,

$$\Pr[\exists i \in U\ \forall j \in V, \varphi(x_{ij}) = 0 | F_1, F_2] \leq n \frac{\left(1 - \frac{\kappa}{n}\right)^n}{\left(1 - \frac{\kappa}{n}\right)^{2\gamma}}$$

As $n \to \infty$, the denominator $\left(1 - \frac{\kappa}{n}\right)^{2\gamma} \to 1$. Now, consider $1 < c < n'-1$, where $n' = n - 2\gamma$. We estimate the probability that there exists an $c \times (n'-c+1)$ all zero sub-matrix of A''. For any $c \times (n'-c+1)$ sub-matrix M of A'', $\Pr[M = 0 | F_1, F_2] = (1-\kappa/n)^{c(n'-c+1)}$.

As there are $\binom{n'}{c}^2$ many such sub-matrices M of A'', we get

$$\Pr[\exists M, M = 0 | F_1, F_2] \leq \binom{n'}{c}^2 (1-\kappa/n)^{c(n'-c+1)}$$

$$\leq (n'e/c)^c (1-\kappa/n)^{c(n'-c+1)} \approx e^{2c\log((n+1)/c) - \kappa c(n'-c+1)/n} \leq e^{-4\log n}$$

the last inequality follows since, $\kappa = 20\log n$, and hence $2c\log(n+1/c) - \kappa(c-1)(n'-c+1)/n \leq -2$ for large enough n.

$$\Pr[perm(\mathcal{A}'') = 0 \mid F_1, F_2] \leq n \cdot \left(1 - \frac{\kappa}{n}\right)^n + ne^{-4\log n} \leq n\left[\left(1 - \frac{\kappa}{n}\right)^{n/\kappa}\right]^\kappa + 1/n^3 \leq n \cdot \mathrm{e}^{-\kappa} \leq 1/n^2.$$

The penultimate inequality in the above is obtained by substituting $\kappa = 20\log n$. ◀

ECCC ISSN 1433-8092

http://eccc.hpi-web.de