Sum of Products of Read-Once Formulas

Ramya C.[1] **and B. V. Raghavendra Rao**[2]

1 **Dept. of Computer Science and Engineering, IIT Madras, Chennai, India**
ramya@cse.iitm.ac.in
2 **Dept. of Computer Science and Engineering, IIT Madras, Chennai, India**
bvrr@cse.iitm.ac.in

Abstract

We study limitations of polynomials computed by depth two circuits built over read-once formulas (ROFs). In particular,

1. We prove an exponential lower bound for the sum of ROFs computing the 2n-variate polynomial in VP defined by Raz and Yehudayoff [CC,2009].
2. We obtain an exponential lower bound on the size of arithmetic circuits computing sum of products of restricted ROFs of unbounded depth computing the permanent of an n by n matrix. The restriction is on the number of variables with + gates as a parent in a proper sub formula of the ROF to be bounded by sqrt(n). Additionally, we restrict the product fan in to be bounded by a sub linear function. This proves an exponential lower bound for a subclass of possibly non-multilinear formulas of unbounded depth computing the permanent polynomial.
3. We also show an exponential lower bound for the above model against a polynomial in VP.
4. Finally we observe that the techniques developed yield an exponential lower bound on the size of sums of products of syntactically multilinear arithmetic circuits computing a product of variable disjoint linear forms where the bottom sum gate and product gates at the second level have fan in bounded by a sub linear function.

Our proof techniques are built on the measure developed by Kumar et. al.[ICALP 2013] and are based on a non-trivial analysis of ROFs under random partitions. Further, our results exhibit strengths and provide more insight into the lower bound techniques introduced by Raz [STOC 2004].

1998 ACM Subject Classification F.1.1 Models of Computation, F.2.1 Numerical Algorithms and Problems

Keywords and phrases Arithmetic Circuits, Permanent, Computational Complexity, Algebraic Complexity Theory

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2016.39

1 Introduction

More than three decades ago, Valiant [26] 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:

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

where S_n is the set of all permutations on $[n]$.

Valiant [26] showed that the polynomial family $(perm_n)_{n\geq 0}$ is complete for the complexity class VNP. Further, Valiant [26] conjectured that $(perm_n)_{n\geq 0}$ does not have polynomial size arithmetic circuits (i.e. VP $\neq$ VNP). 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 seem to be afar, naturally the focus has been on proving lower bounds for restricted classes of circuits computing $perm_n$.

Recent research has focused on proving lower bounds for low depth circuits. Nisan and Wigderson [17] used partial derivatives to obtain exponential lower bounds against special cases of 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 over finite fields. In 2001, Shpilka and Wigderson [23] 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 [24, 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 $\Sigma\Pi^{[O(\sqrt{n})]}\Sigma\Pi^{[\sqrt{n}]}$ 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].)

Apart from constant depth circuits, there has been significant interest in proving lower bounds for unbounded depth circuits with additional structural restrictions such as multilinearity, restricted read etc. A seminal work of Raz [19] 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 [19] Raz used rank of the partial derivative matrix as a complexity measure. Using the same complexity measure as [19], Raz and Yehudayoff [21] proved exponential lower bounds against constant depth multilinear formulas. Subsequently, several generalizations of Raz's measure were introduced. Kumar et al. [13] extended the techniques developed in [19] to prove lower bounds against non-multilinear circuits and formulas of constant depth using the rank of the polynomial coefficient matrix as a measure. (See Definition 8). In [5], Forbes and Shpilka used evaluation dimension of polynomials as a complexity measure to prove exponential lower bounds against Read-Once oblivious algebraic branching programs. Further, in [10] Kayal and Saha used the evaluation dimension to obtain exponential lower bound against depth three multi-k-ic circuits. Over large fields, the evaluation dimension with respect to a partition of the set of variables in a polynomial and rank of the partial derivative matrix with respect to that partition are the same (see Chapter 4 in [4]). However, the evaluation perspective sometimes comes handy in proving lower bounds against non-multilinear circuits.

Motivation: While one direction of research proceeds in proving lower bounds for shallow arithmetic circuits (motivated by the depth reduction results in [1, 24]), the other direction has been on proving lower bounds for unbounded depth circuits with additional structural restrictions.

Despite a large number of lower bound results in the directions mentioned above, the techniques for proving lower bounds presently available to us are very limited, owing to

difficulty in coming up with complexity measures that are sub-additive and sub-multiplicative. In this context, it is important to understand the strength and limitations of existing complexity measures for arithmetic circuits to see their applicability to general classes of arithmetic formulae/circuits. We explore classes of arithmetic formulas where the techniques developed in [19, 13] can be extended and applied. In particular, we consider models that serve as a bridge between shallow arithmetic formulas (e.g., depth two and three formulas) and restricted class of unbounded depth formulas (e.g. multilinear formulas).

Models and Results: Focus of the paper will be on shallow formulas built over restricted formulas of unbounded depth, i.e., a hybrid between bounded depth formulas and restricted formulas of possibly unbounded depth. To begin with we consider the simplest possible restricted formulas of unbounded depth:

▶ **Definition 1** (Read-Once Formula)**.** A formula is said to be a *read-once formula* (ROF) if every variable labels at most one leaf in the formula. A polynomial computed by a read-once formula is called a *read-once polynomial* (ROP).

Observe that not all multilinear polynomials are read once. For instance, using the characterization of ROPs in [27] we can show that $\det_n$ and $perm_n$ are not read-once polynomial. Given that ROFs cannot compute all multilinear polynomials, it is natural to look for generalizations of ROFs that can compute all multilinear polynomials. As a first step, we consider the class $\Sigma \cdot \mathsf{ROF}$: a polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]$ is in $\Sigma \cdot \mathsf{ROF}$ if there exists ROFs $f_1, f_2, \ldots, f_s$ such that $g = \sum_{i=1}^{s} f_i$. Observe that $\Sigma \cdot \mathsf{ROF}$ is a subclass of multilinear unbounded depth formulas. Moreover, since each multilinear monomial is an ROP, any multilinear polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ is in $\Sigma \cdot \mathsf{ROF}$, thus making the model universal. It can be seen that the elementary symmetric polynomial in n variables of degree d denoted by $Sym_{n,d}$ can be computed by linear size $\Sigma \cdot \mathsf{ROF}$ [25]. While the model $\Sigma \cdot \mathsf{ROF}$ is powerful enough to compute elementary symmetric polynomials, we study its limitations. We show:

▶ **Theorem 2.** *There is an explicit $O(n)$ variate polynomial $g \in \mathsf{VP}$ such that for any* ROF*s $f_1, \ldots, f_s$, if $\sum_{i=1}^{s} f_i = g$, then $s = \exp(\Omega(n/\log n))$.*

Shpilka and Volkovich [22] obtained a deterministic quasi polynomial time identity testing algorithm for the sum of a constant number of ROPs. An essential ingredient in their result was a linear lower bound for a special class of ROPs computing $x_1 \cdots x_n$. We note that Theorem 2 is an exponential lower bound against the same model as in [22] against a polynomial in VP defined by Raz-Yehudayoff [20].

▶ Remark. It should be noted that the result in Raz [19] immediately implies a lower bound of $n^{\Omega(\log n)}$ for the sum of ROFs computing det_n or $perm_n$. We exhibit a polynomial in VP that requires a sum of exponential many ROFs to compute it.

Having looked at a subclass of multilinear unbounded depth formulas it is natural to look for non-multilinear unbounded depth formulas. We now introduce our main computational model: $\Sigma\Pi$ formulas built over ROFs ($\Sigma\Pi \cdot \mathsf{ROF}$ for short).

▶ **Definition 3** (Sum of Products of Read-Once Formula)**.** A polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]$ is in $\Sigma\Pi \cdot \mathsf{ROF}$ if there exists ROFs $Q_{ij}, i \in [s], j \in [t]$ such that $g = \sum_{i=1}^{s} \prod_{j=1}^{t} Q_{ij}$.

Since linear forms are computable by ROFs, $\Sigma\Pi \cdot \mathsf{ROF}$ is a natural generalization of $\Sigma\Pi\Sigma$ formulas. As every variable is trivially computed by ROF, any polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ can be computed by $\Sigma\Pi \cdot \mathsf{ROF}$. Also, $\Sigma\Pi \cdot \mathsf{ROF}$ is a subclass of non-multilinear unbounded

depth formulas and it contains possibly non-homogeneous and non-multilinear polynomials built using the simplest possible multilinear formulas viz. ROFs. We observe that there is a simple ROF which computes a product of variable disjoint linear forms such that rank of the partial derivative matrix under a random partition is close to the maximum possible value with high probability (see Lemma 34). This necessitates further restrictions on ROFs that could lead to exponential lower bound against $\Sigma\Pi \cdot \mathsf{ROF}$ using the rank of the partial derivative matrix as the measure of complexity.

Let F be an ROF and for a gate v in F, let $\mathsf{sum\text{-}fan\text{-}in}(v)$ 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 $\mathsf{sum\text{-}fan\text{-}in}(v)$, where the maximum is taken over all $+$ gates v in F of product height at least 1. For an ROP f, define $s(f)$ as the smallest $s(F)$ among all ROFs F computing f. Observe that the construction in [25] shows that $Sym_{n,d} \in \sum_i \prod_j Q_{ij}$ where each Q_{ij} is an ROF and $s(Q_{ij}) = 1$. Our main result is the following :

▶ **Theorem 4.** *Let $\mathcal{C}$ be the class of N-variate ROFs 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 $\epsilon > 0$.*

As far as we know, in the commutative setting, this is the first exponential lower bound for a sub-class of non-multilinear formulas of unbounded depth. In the non-commutative setting, Nisan [16] showed that $\det_n$ and $perm_n$ require $2^{\Omega(n)}$ size non-commutative arithmetic formula. It can be noted that our result above does not depend on the depth of the ROFs. Having proved an exponential lower bound against permanent which is in the class VNP, it is natural to ask if there are polynomials in VP that are hard to be computed by the model. We show the following :

▶ **Theorem 5.** *Let $\mathcal{C}$ be the class of N-variate ROFs F with $s(F) \leq N^{1/4}$. Let $N = n^2$. 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 $s = \exp(\Omega(N^{\epsilon}))$, for some $\epsilon > 0$.*

Since multilinear $\Sigma\Pi\Sigma$ circuits can be viewed as sum of depth two ROPs, we have the following corollary of Theorem 5,

▶ **Corollary 6.** *Let $\mathcal{C}$ be the class of N-variate polynomials computed by multilinear depth three $\sum^{[r]} \prod \sum^{[N^{1/4}]}$ formulas. 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 $\epsilon > 0$.*

Related Results: In [15], Mahajan and Tawari obtain a tight linear lower bound for number of ROPs required to sum-represent elementary symmetric polynomials. That is, they show that the elementary symmetric polynomial Sym_n^{n-1} can be written as a sum of $\lceil n/2 \rceil$ ROPs but cannot be written as a sum of k ROPs for any $k < \lceil n/2 \rceil$. Though the model in [15] is the same as the one in this paper, our lower bound shows that there is an explicit polynomial g that requires exponentially many ROPs to sum represent g. 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 [19]. It can be noted that the lower bounds obtained in [19] are super polynomial

and not exponential. Though Raz-Yehudayoff [21] proved exponential lower bounds, their argument works only against bounded depth multilinear circuits. Further, the arguments in [19, 21] 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 in [19, 21] (see Section 2 for a definition) is defined only for multilinear 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 [19] or [21] on the measure for the underlying arithmetic formula model is insufficient to handle products of two ROPs.

Our approach to prove Theorems 4 and 5 lie in obtaining exponentially stronger upper bounds (see Lemma 33) 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 34, 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 [21, 13] in the case of non-multilinear formulas.

The rest of the paper is organized as follows. Section 2 provides essential definitions used in the paper. Section 3 proves Theorem 2. Sections 4 proves the remaining results. Proofs omitted due to space constraints can be found in the full version of the paper [18].

2 Preliminaries

In this section we recall some basic definitions and introduce notations used in this article.

▶ **Definition 7** (Arithmetic Circuits)**.** Let $\mathbb{F}$ be a 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 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 2 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 $\mathcal{C}$. The *depth* of $\mathcal{C}$ is the length of the longest path from an input gate to the output gate in $\mathcal{C}$. An arithmetic circuit is called an *arithmetic formula* if the underlying undirected graph is a tree.

The *product height* of a gate v in $\mathcal{C}$ is the maximum number of $\times$ gates along any path from v to the root gate in $\mathcal{C}$. For g any gate in a circuit $\mathcal{C}$, $\mathsf{var}(g)$ 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. We now review the polynomial coefficient matrix introduced in [13]. Let $\mathbb{F}$ be a field and $X = \{x_1, \ldots, x_N\}$,$Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$ be disjoint sets of variables.

▶ **Definition 8** (Polynomial Coefficient Matrix)**.** 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 : 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 (1) $\mathsf{var}(A) \subseteq \mathsf{var}(p) \cup \mathsf{var}(q)$ and (2) for every monomial $m \in B$, either $pq \nmid m$ or $\mathsf{var}(m) \subsetneq \mathsf{var}(p) \cup \mathsf{var}(q)$.

▶ **Observation 9.** *For a multilinear polynomial $f \in \mathbb{F}[Y, Z]$, the polynomial coefficient matrix [13] and the partial derivative matrix [19] 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. The proofs of Lemma 10 and 11 follow directly from [13].

▶ **Lemma 10** (Sub-additivity, [13]). *Let $f, g \in \mathbb{F}[Y, Z]$. Then, we have that* $\text{maxrank}(M_{f+g}) \leq \text{maxrank}(M_f) + \text{maxrank}(M_g)$.

▶ **Lemma 11** (Sub-multiplicativity, [13]). *Let $Y_1, Y_2 \subseteq Y$ and $Z_1, Z_2 \subseteq Z$. Then for any polynomials $f \in \mathbb{F}[Y_1, Z_1]$, $g \in \mathbb{F}[Y_2, Z_2]$, we have* $\text{maxrank}(M_{fg}) \leq \text{maxrank}(M_f) \cdot \text{maxrank}(M_g)$. *Also, when $Y_1 \cap Y_2 = \emptyset$ and $Z_1 \cap Z_2 = \emptyset$ we have* $\text{maxrank}(M_{fg}) = \text{maxrank}(M_f) \cdot \text{maxrank}(M_g)$.

▶ **Observation 12.** *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 13** (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 13. 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 $f^{\varphi} \triangleq f(\varphi(X)) \in \mathbb{F}[Y, Z]$.

Consider a formula F all of whose leaves are labelled by constants. Then F computes a constant say α. Observe that in this case for any partition function $\varphi : X \to Y \cup Z \cup \{0, 1\}$, we have $\text{rank}(M_{\alpha^{\varphi}}) = 1$. However, Lemmas 10 and 11 we may get $\text{rank}(M_{\alpha^{\varphi}})$ as large as exponential in size of F. Hence we need a notion of formulas that use constants from $\mathbb{F}$ in a minimal fashion :

▶ **Definition 14** (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 state some observations on formulas that compute natural numbers. An arithmetic formula F is said to be monotone if no leaf in F is labelled by 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.

▶ **Lemma 15.** *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 16** (*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 $\mathsf{value}(u_1), \mathsf{value}(u_2) < 2$.

▶ **Lemma 17.** *Let F be a binary monotone arithmetic formula with leaves labelled by either* 1 *or* 2. *Suppose* 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 state the following variants of the well known Chernoff-Hoeffding bounds.

▶ **Theorem 18** (Chernoff-Hoeffding bound, [3])**.** *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; and
2. type-B gates: product gates both of whose children are variables; and
3. type-C gates: sum gates exactly one child of which is a variable; and
4. type-D gates: product gates exactly one child of which is a variable.

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 19.** *Let F be a binary arithmetic formula. Then there is a formula F' computing the same polynomial as F such that no root to leaf path in F' has two consecutive type-C gates. Therefore, for any binary formula F, without the loss of generality we have $c \leq a + b + d$.*

We say a gate G computing a polynomial g achieves rank-1 under φ if $\mathrm{rank}(M_{g^\varphi}) = 1$ and we say the gate G achieves rank-2 under φ if $\mathrm{rank}(M_{g^\varphi}) = 2$. 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 bounds the rank of M_{f^φ}.

▶ **Lemma 20.** *Let F be an* ROF *computing an* ROP *f and $\varphi : X \to Y \cup Z$. Then,* $\mathrm{rank}(M_{f^\varphi}) \leq 2^{a'' + \frac{a'}{2} + \frac{2b}{3} + \frac{c}{2}}$*, where a'', a', b and c are as defined above.*

▶ **Lemma 21.** *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 18 (2) and (3) with $\delta = 1/5$, we get the required bounds. ◀

▶ **Lemma 22.** *Let f be an* ROP *on $2n$ variables and $\varphi \sim \mathcal{D}'$. Then with probability at least* $1 - 2^{-\Omega\left(\frac{n}{\log n}\right)}$, $\mathrm{rank}(M_{f^\varphi}) \leq 2^{n - \frac{n}{15 \log n}}$.

Proof. Let F be an ROF computing f, and a, b, c, d, a' and a'' be as in the discussion preceding Lemma 20. We have 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 $a \geq \frac{n}{\log n}$, then by Lemma 20, 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 21, $a' \geq \frac{2a}{5} \geq \frac{2n}{5 \log n}$ with probability $1 - 2^{-\Omega\left(\frac{n}{\log n}\right)}$. 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 19, $a + b + d \geq c \geq \frac{n}{\log n}$, 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 $1 - 2^{-\Omega\left(\frac{n}{\log n}\right)}$.
- If $b \geq \frac{n}{3\log n}$ by Lemma 20, $\text{rank}(M_{f^\varphi}) \leq 2^{a+2b/3+c/2}$. 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 20 $\text{rank}(M_{f^\varphi}) \leq 2^{a''+a'/2+2b/3+c/2} \leq 2^{a+b+c/2} \leq 2^{n-\frac{d}{2}} \leq 2^{n-\frac{n}{6\log n}} \leq 2^{n-\frac{n}{15\log n}}$.

Case 2: $a + c < \frac{2n}{\log n}$. Observe that $b \leq n$. By Lemma 20, $\text{rank}(M_{f^\varphi}) \leq 2^{a+2b/3+c} \leq 2^{2n/3+2n/\log n} \leq 2^{n-n/15\log n}$ for large enough n. ◀

The following polynomial was introduced by Raz and Yehudayoff [20].

▶ **Definition 23.** 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 $W_{i,j} = \{w_{i',k,j'} \mid i', k, j' \in [i, j]\}$. Let $\mathbb{G} = \mathbb{F}(\mathcal{W})$, the rational function field. For every $[i, j]$ such that $|[i, j]|$ is even we define a polynomial $g_{i,j} \in \mathbb{G}[X]$ as $g_{i,j} = 1$ when $|[i, j]| = 0$ and if $|[i, j]| > 0$ then, $g_{i,j} \triangleq (1+x_ix_j)g_{i+1,j-1}+\sum_k w_{i,k,j}g_{i,k}g_{k+1,j}$. 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 $|[i, k]|$ is even. Let $g \triangleq g_{1,2n}$.

The following lemma builds on Lemma 4.3 in [20].

▶ **Lemma 24.** *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 23 we have,* $\text{rank}(M_{g^\varphi}) \geq 2^{n-\ell/2}$.

▶ **Lemma 25.** *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 18) with $\delta = 1/n^{1/3}$. ◀

▶ **Corollary 26.** $\Pr_{\varphi\sim\mathcal{D}'}[\text{rank}(M_{g^\varphi}) \geq 2^{n-n^{2/3}}] \geq 1 - 2^{-\Omega(n^{1/3})}$.

Proof. Apply Lemma 24 with $\ell = 2n/n^{1/3} = 2n^{2/3}$ and apply Lemma 25. ◀

Proof of Theorem 2

Proof. Suppose $s < \exp(o(n/\log n))$. Then by Lemma 22 and union bound, probability that there is an i such that $\text{rank}(M_{f_i^\varphi}) \geq \exp(n - n/15\log n)$ is $s\exp(-\Omega(n/\log n)) = \exp(-\Omega(n/\log n))$ and hence by Lemma 10, $\text{rank}(M_{g^\varphi}) \leq s\exp(n - n/15\log n) \leq \exp(n - n/20\log n)$ with probability $1 - \exp(-\Omega(n/\log n))$ for large enough n. However, by Corollary 26, $\text{rank}(M_{g^\varphi}) \geq \exp(n - n^{2/3}) > \exp(n - n/20\log n)$ with probability at least $1 - \exp(-\Omega(n^{1/3}))$, a contradiction. Therefore, $s = \exp(\Omega(n/\log n))$. ◀

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.

▶ **Lemma 27.** *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* $\Pr_{\varphi \sim D}\left[\mathcal{X} > (N^{1/4})\right] \leq \exp(-\Omega(N^{1/4}))$.

▶ **Lemma 28.** *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* $\text{rank}(M_{F^\varphi}) \leq \text{rank}(M_{F'^\varphi}) \cdot \exp(\mathcal{O}(N^{1/4}))$ *with probability atleast* $1 - \exp(-\Omega(N^{1/4}))$.

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 29.** *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 30.** *Let F be an* ROF *and $\varphi \sim \mathcal{D}$. By Lemma 29, 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 17 we have the following corollary,

▶ **Corollary 31.**

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

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-*$(1,2)$*-separator*s.

▶ **Lemma 32.** *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* $\Pr_\varphi[\bigwedge_{i=1}^t u_i \text{ is a rank-}(1,2)\text{-separator}] \leq c^t N^{-5t/6}$, *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 $\Pr[u_i \text{ is a } (1,2) \text{ separator}] \leq \Pr[|\mathsf{var}(u_i^\varphi) \cap Y| \geq 1 \text{ and } |\mathsf{var}(u_i^\varphi) \cap Z| \geq 1] \leq \Pr[|\mathsf{var}(u_i^\varphi) \cap (Y \cup Z)| \geq 2]$. Let $\ell_{i_1}, \ldots, \ell_{i_{r_i}}$ be the addition gates at depth-1 in the subformula 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 $0 \leq i \leq t$, $\Pr[\ u_i \text{ is a } (1,2) \text{ separator}] \leq \Pr[|S_i \cap (Y \cup Z)| \geq 2]$. 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 $c > 0$

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

▶ **Lemma 33.** *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 28, 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 no $\times$ gate with at more than two variables as its input. By Corollary 31 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 32 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 34.** *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, $\text{rank}(M_{p_{lin}{}^\varphi}) = \exp(\Omega(m))$ with probability $1 - \exp(-\Omega(m))$.*

Proof. Let $p_{lin} = \ell_1 \cdots \ell_{m'}$ where $\ell_j = \left(\sum_{i=(j-1)(N/2m)+1}^{jN/2m} x_i\right) + 1$ 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 $\Pr[\rho < (1-\delta)m/8] \leq \Pr[\rho < (1-\delta)\mu] = \exp(-\Omega(m))$ by Theorem 18 with $\delta = 1/4$, since $\text{rank}(M_{p_{lin}^\varphi}) = \exp(\rho)$. ◀

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 35.** 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 $\forall j' \in [n], j' \neq j \;\; \varphi(x_{ij'}) \in \{0,1\}$.

▶ **Lemma 36.** *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, $\Pr_{\varphi\sim\mathcal{D}}[3m/4 < \chi < 5m/4] = 1 - \exp(-\Omega(m))$.*

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 37.** Let $\mathcal{Q} \in \{Y, Z\}$. A column C_j, $1 \leq j \leq n$ 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 38.** *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 38 and union bound we have:

▶ **Lemma 39.** *For $1 \leq i \leq n$, let C_i be a column in X_φ. Then for any $\mathcal{Q} \in \{Y, Z\}$, $\Pr_{\varphi \sim \mathcal{D}}[C_i$ is $\mathcal{Q}$-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 $\zeta_\mathcal{Q} \triangleq |\{R_j \mid R_j \text{ is } \mathcal{Q}\text{-good}\}|$.

▶ **Lemma 40.** *With notations as above, $\forall \mathcal{Q} \in \{Y, Z\}$, $\Pr_{\varphi \sim \mathcal{D}}[\eta_\mathcal{Q} \geq \frac{2m}{3}] = 1 - \exp(-\Omega(m))$; and $\Pr_{\varphi \sim \mathcal{D}}[\zeta_\mathcal{Q} \geq \frac{2m}{3}] = 1 - \exp(-\Omega(m))$.*

▶ **Lemma 41.** *For $\mathcal{Q} \in \{Y, Z\}$, let $\gamma_\mathcal{Q}$ denote the number of $\mathcal{Q}$-special positions in X_φ. Then $\forall \mathcal{Q} \in \{Y, Z\}$, $\Pr_{\varphi \sim \mathcal{D}}\left[\gamma_\mathcal{Q} \geq \frac{m}{12}\right] = 1 - \exp(-\Omega(m))$.*

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 : $2m/3 \leq |X_\varphi \cap Y| \leq 5m/4$;
E2 : The number of Y-good columns and Y-good rows is at least $r \triangleq 2m/3$. By Lemmas 36 and 40, X_φ satisfies the events E1 and E2 with probability $1 - \exp(-\Omega(m))$. 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, $1 \leq i \leq r$ 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. 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, $t \leq 5m/4 - 2m/3 \leq 7m/12$. 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 $r - t \geq (2/3 - 7/12)m = m/12$. We conclude, $\Pr_{\varphi \sim \mathcal{D}}[\gamma_Y \geq m/12] = 1 - \exp(-\Omega(m))$. ◀

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 is immediate :

▶ **Observation 42.** *Let φ be distributed according to $\mathcal{D}$. Then for any row (column) R: $\Pr_{\varphi \sim \mathcal{D}}[R$ is 1-good$] \geq (1 - 1/n^3)$.*

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

▶ **Theorem 43.** $\Pr[\text{rank}(M_{\mathsf{perm}_n^\varphi}) \geq 2^{m/12}] \geq (1 - O(1/n^2))/2$.

We need a few notations and Lemmas before proving Theorem 43. 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

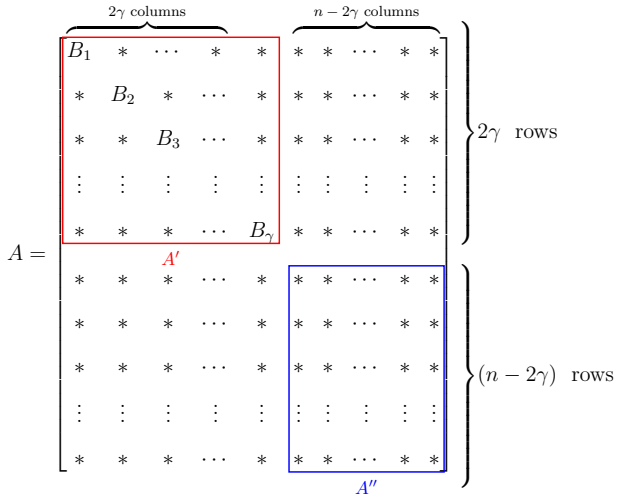

Figure 1 The matrix A after permuting the rows and columns. $*$ denotes unspecified entries.

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 $\text{rank}(M_{B_p^\varphi}) = 2$. 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.

Since (i_p, j_p) is a Y-special position, (k_p, ℓ_p) is a Z-special position we 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, $\text{rank}(M_{\mathsf{perm}(B_p)}) = 2$ 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 $\mathsf{perm}(A'') \neq 0$.

▶ **Lemma 44.** *Let A'' be matrix as in Figure 1. Then $\Pr_\varphi[perm(A'') \neq 0 \mid F_1, F_2] \geq 1 - \frac{1}{n^2}$.*

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

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

▶ **Observation 45.** $\forall A \in \mathcal{A}$, $\text{rank}(M_{\mathsf{perm}(A')}) < 2^\gamma$ *and* $\forall B \in \mathcal{B}, \text{rank}(M_{\mathsf{perm}(B)}) \geq 2^\gamma$.

▶ **Lemma 46.** *Let $\mathcal{A}$ and $\mathcal{B}$ as defined above. Then*
(a) $\Pr_{\varphi \sim \mathcal{D}}[\text{rank}(M_{\mathsf{perm}(X_\varphi)}) \geq 2^\gamma)] \geq \mathcal{D}(\mathcal{B})$*; 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 45. 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 $\text{rank}(M_{\mathsf{perm}(B_p)}) = 1$. 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 $1 \leq p \leq \gamma$ 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 $\mathcal{D}(\mathcal{A}) \leq \mathcal{D}(\mathcal{B})$. ◀

Proof of Theorem 43. 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 41, $Pr_{\varphi\sim\mathcal{D}}[F_1] = 1-2^{-\Omega(m)}$. From Observation 42 and the union bound we have $Pr_{\varphi\sim\mathcal{D}}[F_2] \geq 1-\gamma/n^3$. By Lemma 44, $Pr_{\varphi\sim\mathcal{D}}[F_3|F_1, F_2] \geq 1-2/n^2$. 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 46, $\Pr_{\varphi\sim\mathcal{D}}[\text{rank}(M_{\mathsf{perm}(X_\varphi)}) \geq 2^\gamma] \geq 1/2(1-O(\frac{1}{n^2}))$. ◀

4.3 Putting them all together

Proof of Corollary 6

Proof. Suppose $p_{lin} = \sum_{i=1}^{s}\prod_{j=1}^{t} f_{i,j}$ where $f_{i,j}$ are syntactically multilinear $\Sigma\Pi\Sigma$ formula, with $s < \exp(N^{1/4})$, 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 $s_{T_{i,j,k}} \leq \exp(N^{1/4})$. Then by Lemma 33 and union bound there is an i, j, k such that $\text{rank}(M_{T^\varphi_{i,j,k}}) \geq \exp(N^{4/15})$ with probability at most $sts'\exp(-\Omega(N^{1/4}))$. By Lemma 10 and 11, we have $\text{maxrank}(M_{p^\varphi_{lin}}) \leq 2^{N^{4/15}}$ with probability $1-o(1)$. However by Lemma 34, $\text{maxrank}(M_{p^\varphi_{lin}}) = \text{rank}(M_{p^\varphi_{lin}}) = \exp(\Omega(m))$ with probability at least $1-\exp(-\Omega(m))$, a contradiction. Hence $ss' = \exp(\Omega(N^{1/4}))$. ◀

Proof of Theorem 5

Proof. Suppose $s = \exp(o(N^{1/4}))$. Then by Lemma 33, the probability that there is an $f_{i,j}$ with $\text{rank}(M_{f^\varphi_{i,j}}) \geq \exp(N^{4/15})$ is at most $\exp(-\Omega(N^{1/4}))s = o(1)$. By Lemma 10 and 11 and since $\text{maxrank}(M_{f^\varphi_{i,j}}) = \text{rank}(M_{f^\varphi_{i,j}})$, we have $\text{maxrank}(M_{p^\varphi_{lin}}) \leq (s\cdot\exp(N^{4/15})^{N^{1/30}}) = \exp(o(N^{1/3}))$ with probability $1-o(1)$. However by Lemma 34, $\text{maxrank}(M_{p^\varphi_{lin}}) = \exp(\Omega(m))$ with probability $1-\exp(-\Omega(m))$, a contradiction. Hence $s = \exp(\Omega(N^{1/4}))$. ◀

Proof of Theorem 4

Proof. Suppose $s = \exp(o(N^{1/4}))$. Then by Lemma 33, Probability that there is an $f_{i,j}$ with $\text{rank}(M_{f^\varphi_{i,j}}) \geq \exp(N^{4/15})$ is at most $\exp(-\Omega(N^{1/4}))s = o(1)$. Then, by Lemma 10 and 11, we have $\text{maxrank}(M_{mathitperm^\varphi_n}) \leq s\cdot(\exp(N^{4/15}))^{N^{1/30}} = \exp(o(N^{1/3}))$ with probability $1-o(1)$. However, by Theorem 43, $\text{maxrank}(M_{mathitperm^\varphi_n}) = \text{rank}(M_{mathitperm^\varphi_n})\exp(\Omega(m))$ with probability $(1-1/n^2)/2$, a contradiction. Hence $s = \exp(\Omega(N^{1/4}))$. ◀

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 20.

References

1 Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *FOCS*, pages 67–75, 2008. doi:10.1109/FOCS.

2 Walter Baur and Volker Strassen. The complexity of partial derivatives. *Theor. Comput. Sci.*, 22:317–330, 1983. doi:10.1016/0304-3975(83)90110-X.

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. doi:10.1145/276698.276872.

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. doi:10.1145/2629541.

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. doi:10.1145/2591796.2591823.

10 Neeraj Kayal and Chandan Saha. Multi-k-ic depth three circuit lower bound. In *STACS*, pages 527–539, 2015. doi:10.4230/LIPIcs.STACS.2015.527.

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. Accepted at ICALP 2016. URL: http://eccc.hpi-web.de/report/2016/006.

12 Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012. doi:10.1016/j.tcs.2012.03.041.

13 Mrinal Kumar, Gaurav Maheshwari, and Jayalal Sarma. Arithmetic circuit lower bounds via maximum-rank of partial derivative matrices. *TOCT*, 8(3):8, 2016. doi:10.1145/2898437.

14 Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *FOCS*, pages 364–373, 2014. doi:10.1109/FOCS.2014.46.

15 Meena Mahajan and Anuj Tawari. Sums of read-once formulas: How many summands suffice? In *Computer Science – Theory and Applications – 11th International Computer Science Symposium in Russia, CSR 2016, St. Petersburg, Russia, June 9-13, 2016, Proceedings*, pages 266–279, 2016. doi:10.1007/978-3-319-34171-2_19.

16 Noam Nisan. Lower bounds for non-commutative computation (extended abstract). In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, May 5-8, 1991, New Orleans, Louisiana, USA*, pages 410–418, 1991. doi:10.1145/103418.103462.

17 Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1997. doi:10.1007/BF01294256.

18 C. Ramya and B. V. Raghavendra Rao. Limitations of sum of products of read-once polynomials. *CoRR*, abs/1512.03607, 2015. URL: https://arxiv.org/abs/1512.03607.

19 Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2), 2009. doi:10.1145/1502793.1502797.

20 Ran Raz and Amir Yehudayoff. Balancing syntactically multilinear arithmetic circuits. *Computational Complexity*, 17(4):515–535, 2008.

21 Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Computational Complexity*, 18(2):171–207, 2009. doi:10.1007/s00037-009-0270-8.

22 Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. *Computational Complexity*, 24(3):477–532, 2015. doi:10.1007/s00037-015-0105-8.

23 Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001. doi:10.1007/PL00001609.

24 Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Inf. Comput.*, 240:2–11, 2015. doi:10.1016/j.ic.2014.09.004.

25 Iddo Tzameret. Studies in algebraic and propositional proof complexity. *Ph.D Thesis*, page 33, 2008. URL: http://www.cs.rhul.ac.uk/home/tzameret/Iddo-PhD-thesis.pdf.

26 Leslie G. Valiant. Completeness classes in algebra. In *STOC*, pages 249–261, 1979. doi:10.1145/800135.804419.

27 Ilya Volkovich. Characterizing arithmetic read-once formulae. *TOCT*, 8(1):2, 2016. doi:10.1145/2858783.