Unbalancing Sets and an Almost Quadratic Lower Bound for Syntactically Multilinear Arithmetic Circuits

Noga Alon* Mrinal Kumar† Ben Lee Volk‡

Abstract

We prove a lower bound of $\Omega(n^2/\log^2 n)$ on the size of any syntactically multilinear arithmetic circuit computing some explicit multilinear polynomial $f(x_1, \ldots, x_n)$. Our approach expands and improves upon a result of Raz, Shpilka and Yehudayoff ([RSY08]), who proved a lower bound of $\Omega(n^{4/3}/\log^2 n)$ for the same polynomial. Our improvement follows from an asymptotically optimal lower bound for a generalized version of Galvin's problem in extremal set theory.

A special case of our combinatorial result implies, for every n, a tight $\Omega(n)$ lower bound on the minimum size of a family $\mathcal{F}$ of subsets of cardinality $2n$ of a set X of size $4n$, so that any subset of X of size $2n$ has intersection of size exactly n with some member of $\mathcal{F}$. This settles a problem of Galvin up to a constant factor, extending results of Frankl and Rödl [FR87] and Enomoto et al. [EFIN87], who proved in 1987 the above statement (with a tight constant) for odd values of n, leaving the even case open.

*Sackler School of Mathematics and Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv 6997801, Israel and CMSA, Harvard University, Cambridge, MA 02138, USA. Email: nogaa@tau.ac.il. Research supported in part by an ISF grant and by a GIF grant.

†Center for Mathematical Sciences and Applications, Harvard University, Cambridge, Massachusetts, USA. Email: mrinalkumar08@gmail.com. Part of this work was done while visiting Tel Aviv University.

‡Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel, Email: benleevolk@gmail.com. The research leading to these results has received funding from the Israel Science Foundation (grant number 552/16).

1 Introduction

An arithmetic circuit is one of the most natural and standard computational models for computing multivariate polynomials. Such circuits provide a succinct representation of multivariate polynomials, and in some sense, they can be thought of as algebraic analogs of boolean circuits. Formally, an arithmetic circuit over a field $\mathbb{F}$ and a set of variables $X = \{x_1, x_2, \ldots, x_n\}$ is a directed acyclic graph in which every vertex has in-degree either zero or two. The vertices of in-degree zero (called *leaves*) are labeled by variables in X or elements of $\mathbb{F}$, and the vertices of in-degree two are labeled by either $+$ (called *sum* gates) or $\times$ (called *product* gates). A circuit can have one or more vertices of out degree zero, known as the output gates. The polynomial computed by a vertex in any[1] given circuit is naturally defined in an inductive way: a leaf computes the polynomial which is equal to its label. A sum gate computes the polynomial which is the sum of the polynomials computed at its children and a product gate computes the polynomial which is the product of the polynomials at its children. The polynomials computed by a circuit are the polynomials computed by its output gates. The size of an arithmetic circuit is the number of vertices in it.

It is not hard to show (see, e.g., [CKW11]) that a random polynomial of degree $d = \mathsf{poly}(n)$ in n variables cannot be computed by an arithmetic circuit of size $\mathsf{poly}(n)$ with overwhelmingly high probability. A fundamental problem in this area of research is to prove a similar super-polynomial lower bound for an *explicit* polynomial family. Unfortunately, the problem continues to remain wide open and the current best lower bound known for general arithmetic circuits[2] is an $\Omega(n \log n)$ lower bound due to Strassen [Str73] and Baur and Strassen [BS83] from more than three decades ago. The absence of substantial progress on this general question has led to focus on the question of proving better lower bounds for restricted and more structured subclasses of arithmetic circuits. Arithmetic formulas [Kal85], non-commutative arithmetic circuits [Nis91], algebraic branching programs [Kum17], and low depth arithmetic circuits [NW97, GK98, GR00, Raz10, GKKS14, FLMS14, KLSS14, KS14, KS17] are some such subclasses which have been studied from this perspective. For an overview of the definition of these models and the state of art for lower bounds for them, we refer the reader to the surveys of Shpilka and Yehudayoff [SY10] and Saptharishi [Sap16].

Several of the most important polynomials in algebraic complexity and in mathematics in general are multilinear. Notable examples include the determinant, the permanent, and the elementary symmetric polynomials. Therefore, one subclass which has received a lot of attention in the last two decades and will be the focus of this paper is the class of *multilinear* arithmetic circuits.

1.1 Multilinear arithmetic circuits

For an arithmetic circuit Ψ and a vertex v in Ψ, we denote by X_v the set of variables x_i such that there is a directed path from a leaf labeled by x_i to v; in this case, we also say that v *depends* on x_i[3]. A polynomial P is said to be multilinear if the individual degree of every variable in P is at most one.

An arithmetic circuit Ψ is said to be *syntactically* multilinear if for every multiplication gate v in Ψ with children u and w, the sets of variables X_u and X_w are disjoint. We say that Ψ is *semantically* multilinear if the polynomial computed at every vertex is a multilinear polynomial. Observe that if Ψ is a syntactically multilinear circuit, then it is also semantically multilinear.

[1] Throughout this paper, we will use the terms gates and vertices interchangeably.

[2] In the rest of the paper, when we say a lower bound, we always mean it for an explicit polynomial family.

[3] We remark that this is a syntactic notion of dependency, since it is possible that every monomial with x_i might get canceled in the intermediate computation and might not eventually appear in the polynomial computed at v.

However, it is not clear if every semantically multilinear circuit can be efficiently simulated by a syntactically multilinear circuit.

A multilinear circuit is a natural model for computing multilinear polynomials, but it is not necessarily the most efficient one. Indeed, it is remarkable that all the constructions of polynomial size arithmetic circuits for the determinant [Csa76, Ber84, MV97], which are fundamentally different from one another, nevertheless share the property of being *non*-multilinear, namely, they involve non-multilinear intermediate computations which eventually cancel out. There are no subexponential-size multilinear circuits known for the determinant, and one may very well conjecture these do not exist at all.

Multilinear circuits were first studied by Nisan and Wigderson [NW97]. Subsequently, Raz [Raz09] defined the notion of multilinear formulas[4] and showed that any multilinear formula computing the determinant or the permanent of an $n \times n$ variable matrix must have super-polynomial size. In a follow up work [Raz06], Raz further strengthed the results in [Raz09] and showed that there is a family of multilinear polynomials in n variables which can be computed by a $\mathsf{poly}(n)$ size syntactically multilinear arithmetic circuits but require multilinear formulas of size $n^{\Omega(\log n)}$.

Building on the ideas and techniques developed in [Raz09], Raz and Yehudayoff [RY09] showed an exponential lower bound for syntactically multilinear circuits of constant depth. Interestingly, they also showed a super-polynomial separation between depth Δ and depth $\Delta + 1$ syntactically multilinear circuits for constant Δ.

In spite of the aforementioned progress on the question of lower bounds for multilinear formulas and bounded depth syntactically multilinear circuits, there was no $\Omega(n^{1+\varepsilon})$ lower bounds known for general syntactically multilinear circuits for any constant $\varepsilon > 0$. In fact, the results in [Raz06] show that the main technical idea underlying the results in [Raz09, Raz06, RY09] is unlikely to directly give a super-polynomial lower bound for general syntactically multilinear circuits. However, a weaker super-linear lower bound still seemed conceivable via similar techniques.

Raz, Shpilka and Yehudayoff [RSY08] showed that this is indeed the case. By a sophisticated and careful application of the techniques in [Raz09] along with several additional ideas, they established an $\Omega\left(\frac{n^{4/3}}{\log^2 n}\right)$ lower bound for an explicit n variate polynomial. Since then, this has remained the best lower bound known for syntactically multilinear circuits. In this paper, we improve this result by showing an almost quadratic lower bound for syntactically multilinear circuits for an explicit n variate polynomial. In fact, the family of hard polynomials in this paper is the same as the one used in [RSY08]. We now formally state our result.

Theorem 1.1. *There is an explicit family of polynomials $\{f_n\}$, where f_n is an n variate multilinear polynomial, such that any syntactically multilinear arithmetic circuit computing f_n must have size at least $\Omega(n^2/\log^2 n)$.*

For our proof, we follow the strategy in [RSY08]. Our improvement comes from an improvement in a key lemma in [RSY08] which addresses the following combinatorial problem.

Question 1.2. *What is the minimal integer $m = m(n)$ for which there is a family of subsets $S_1, S_2, \ldots, S_m \subseteq [n]$, each S_i satisfying $6\log n \leq |S_i| \leq n - 6\log n$ such that for every $T \subseteq [n], |T| = \lfloor n/2 \rfloor$, there exists an $i \in [m]$ with $|T \cap S_i| \in \{\lfloor |S_i|/2 \rfloor - 3\log n, \lfloor |S_i|/2 \rfloor - 3\log n + 1, \ldots, \lfloor |S_i|/2 \rfloor + 3\log n\}$?*

Raz, Shpilka and Yehudayoff [RSY08] showed that $m(n) \geq \Omega\left(n^{1/3}/\log n\right)$. For our proof, we show that $m(n) \geq \Omega\left(n/\log n\right)$.

[4] For formulas, it is known that syntactic multilinearity and semantically multilinearity are equivalent (See, e.g., [Raz09]).

In addition to its application to the proof of Theorem 1.1, Question 1.2 seems to be a natural problem in extremal combinatorics and might be of independent interest, and special cases thereof were studied in the combinatorics literature. In the next section, we briefly discuss the state of the art of this question and state our main technical result about it in Theorem 1.3.

1.2 Unbalancing Sets

The following question, which is of very similar nature to Question 1.2, is known as Galvin's problem (see [FR87, EFIN87]): What is the minimal integer $m = m(n)$, for which there exists a family of subsets $S_1, \ldots, S_m \subseteq [4n]$, each of size $2n$, such that for every subset $T \subseteq [4n]$ of size $2n$ there exists some $i \in [m]$ such that $|T \cap S_i| = n$?

It is not hard to show that $m(n) \leq 2n$. Indeed, let $S_i = \{i, i+1, \ldots, i+2n-1\}$, for $i \in \{1, 2, \ldots, 2n+1\}$, and let $\alpha_i(T) = |T \cap S_i| - |([4n] \setminus T) \cap S_i|$. Then $\alpha_i(T)$ is always an even integer, $\alpha_1(T) = -\alpha_{2n+1}(T)$, and $\alpha_i - \alpha_{i+1}(T) \in \{0, \pm 2\}$ if $i \leq 2n$. By a discrete version of the intermediate value theorem, it follows there exists $j \in [2n]$ such that $\alpha_j(T) = 0$, which implies that exactly n elements of S_j belong to T. Thus, the family $\{S_1, \ldots, S_{2n}\}$ satisfies this property.

As for lower bounds, a counting argument shows that $m(n) = \Omega(\sqrt{n})$, since for each fixed S of size $[2n]$ and random T of size $2n$,

$$\Pr[|T \cap S| = n] = \frac{\binom{2n}{n} \cdot \binom{2n}{n}}{\binom{4n}{2n}} = \Theta\left(\frac{1}{\sqrt{n}}\right).$$

Frankl and Rödl [FR87] were able to show that $m(n) \geq \varepsilon n$ for some $\varepsilon > 0$ if n is odd, and Enomoto, Frankl, Ito and Nomura [EFIN87] proved that $m(n) \geq 2n$ if n is odd, which implies that even the constant in the construction given above is optimal. Until this work, the question was still open for even values of n: in fact, Markert and West (unpublished, see [EFIN87]) showed that for $n \in \{2, 4\}$, $m(n) < 2n$.

For our purposes, we need to generalize Galvin's problem in two ways. The first is to lift the restriction on the set sizes. The second is to ask how small can the size of the family $\mathcal{F} = \{S_1, \ldots, S_m\} \subseteq 2^{[n]}$ be if we merely assume each balanced partition T is "τ-balanced" on some $S \in \mathcal{F}$, namely, if $||T \cap S| - |S|/2|| \leq \tau$ for some S (the main case of interest for us is $\tau = O(\log n)$). Of course, since T itself is balanced, very small or very large sets are always τ-balanced, and thus we impose the (tight) non-triviality condition $2\tau \leq |S| \leq n - 2\tau$ for every $S \in \mathcal{F}$.

Once again, by defining $S_i = \{i, i+1, \ldots, i+n/2-1\}$ (n is always assumed to be even), the family $\mathcal{F} = \left\{S_1, S_{1+\tau}, S_{1+2\tau}, ..., S_{1+\lfloor n/(2\tau)\rfloor \cdot \tau}\right\}$ gives a construction of size $O(n/\tau)$ such that every balanced partition T is τ-balanced on some $S \in \mathcal{F}$.

It is natural to conjecture that, perhaps up to a constant, this construction is optimal. Indeed, this is what we prove here.

Theorem 1.3. *Let n be any large enough even number, and let $\tau \geq 1$ be an integer. Let $S_1, \ldots, S_m \subseteq [n]$ be sets such that for all $i \in [m]$, $2\tau \leq |S_i| \leq n - 2\tau$. Further, assume that for every $Y \subseteq [n]$ of size $n/2$ there exists $i \in [m]$ such that $||Y \cap S_i| - |S_i|/2| < \tau$. Then, $m \geq \Omega(n/\tau)$.*

In particular, Theorem 1.3 proves a linear lower bound $m = \Omega(n)$ for the original problem of Galvin, even when the universe size is of the form $4k$ for even k.

We remark that the relevance of problems of this form to lower bounds in algebraic complexity was also observed by Jansen [Jan08] who considered the problem of obtaining a lower bound on homogenous syntactically multilinear algebraic branching program (which is a weaker model than syntactically multilinear circuits), and essentially proposed Theorem 1.3 as a conjecture. In fact,

a special case of this theorem (see Theorem 3.1), which has a simpler proof, is already enough to derive the improved lower bounds for syntactically multilinear circuits.

Alon, Bergmann, Coppersmith and Odlyzko [ABCO88] considered a very similar problem of balancing ± 1-vectors: they studied families of vectors $\mathcal{F} = \{v_1, \ldots, v_m\}$ such that $v_i \in \{\pm 1\}^n$ for $i \in [m]$, which satisfy the properties that for every $w \in \{\pm 1\}^n$ (not necessarily balanced), there exists $i \in [m]$ such that $|\langle v_i, w \rangle| \leq d$. They generalized a construction of Knuth [Knu86] and proved a matching lower bound which together showed that $m = \lceil n/(d+1) \rceil$ is both necessary and sufficient for such a set to exist. Galvin's problem seems like "the $\{0,1\}$ version" of the same problem, but, to quote from [ABCO88], there does not seem to be any simple dependence between the problems.

1.3 Proof overview

In this section, we discuss the main ideas and give a brief sketch of the proofs of Theorem 1.1 and Theorem 1.3. Since our proof heavily depends on the proof in [RSY08] and follows the same strategy, we start by revisiting the main steps in their proof and noting the key differences between the proof in [RSY08] and our proof. We also outline the reduction to the combinatorial problem of unbalancing set families in Question 1.2.

Proof sketch of [RSY08]

The proof in [RSY08] starts by proving a syntactically multilinear analog of a classical result of Baur and Strassen [BS83], where it was shown that if an n variate polynomial f is computable by an arithmetic circuit Ψ of size $s(n)$, then there is an arithmetic circuit Ψ' of size at most $5s(n)$ with n outputs such that the i-th output gate of Ψ' computes $f_i = \frac{\partial f}{\partial x_i}$. Raz, Shpilka and Yehudayoff show that if Ψ is syntactically multilinear, then the circuit Ψ' continues to be syntactically multilinear. Additionally, there is no directed path from a leaf labeled by x_i to the output gate computing f_i.[5]

Once we have this structural result, it would suffice to prove a lower bound on the size of Ψ'. For brevity, we denote the subcircuit of Ψ' rooted at the output gate computing f_i by Ψ'_i. As a key step of the proof in [RSY08], the authors identify certain sets of vertices $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_n$ in Ψ' with the following properties.

- For every $i \in [n]$, $\mathcal{U}_i$ is a subset of vertices in Ψ'_i.
- For every $i \in [n]$ and $v \in \mathcal{U}_i$, the number of $j \neq i$ such that $v \in \mathcal{U}_j$ is not too large (at most $O(\log n)$).

Observe that at this point, showing a lower bound of $s'(n)$ on the size of each $\mathcal{U}_i$ implies a lower bound of $\Omega(ns'(n)/\log n)$ on the size of Ψ' and hence Ψ. In [RSY08], the authors show that there is an explicit f such that each $\mathcal{U}_i$ must have size at least $\Omega(n^{1/3}/\log n)$, thereby getting a lower bound of $\Omega(n^{4/3}/\log^2 n)$ on the size of Ψ.

For our proof, we follow precisely this high level strategy. Our improvement in the lower bound comes from showing that each $\mathcal{U}_i$ must be of size at least $\Omega(n/\log n)$ and not just $\Omega(n^{1/3}/\log n)$ as shown in [RSY08]. We now elaborate further on the main ideas in this step in [RSY08] and the differences with the proofs in this paper.

We start with some intuition into the definition of the sets $\mathcal{U}_i$ in [RSY08]. Consider a vertex v in Ψ' which depends on at least k variables. Without loss of generality, let these variables be $\{x_1, x_2, \ldots, x_k\}$. From item 4 in Theorem 4.2, we know that the variable x_i does not appear in the

[5] See Theorem 4.2 for a formal statement.

subcircuit Ψ'_i. Therefore, the vertex v cannot appear in the subcircuits $\Psi'_1, \Psi'_2, \ldots, \Psi'_k$. So, if we define the set $\mathcal{U}_i$ as the set of vertices in Ψ'_i which depend on at least k variables, then $\mathcal{U}_i$ must be disjoint from vertices in at least k of the subcircuits $\Psi'_1, \Psi'_2, \ldots, \Psi'_n$. Picking $k \geq n - O(\log n)$ would give us the desired property. So, if we can prove a lower bound on the size of the set $\mathcal{U}_i$, we would be done. However, the definition of the set $\mathcal{U}_i$ so far turns out to be too general: indeed, it is not even a priori clear that the $\mathcal{U}_i$ has any other gates apart from the output gate of Ψ'_i.

As is often the case, the solution to this obstacle is to prove a stronger claim by imposing additional structure on the set $\mathcal{U}_i$. In [RSY08], the set $\mathcal{U}_i$ (called the *upper leveled* gates in Ψ'_i) is defined as the set of all vertices in Ψ'_i which depend on at least $n - 6\log n$ variables and have a child which depends on more than $6\log n$ variables and less than $n - 6\log n$ variables. This additional structure is helpful in proving a lower bound on the size of $\mathcal{U}_i$. We now discuss this in some more detail.

For every $i \in [n]$, let $\mathcal{L}_i$ be the set of vertices u in Ψ'_i, such that $6\log n < |X_u| < n - 6\log n$, and u has a parent in $\mathcal{U}_i$. These gates are referred to as *lower leveled* gates. Observe that $|\mathcal{U}_i| \geq \frac{|\mathcal{L}_i|}{2}$, since the in-degree of every vertex in ψ'_i is at most 2. The key structural property of the set $\mathcal{L}_i$ is the following (see Proposition 5.5 in [RSY08]).

Lemma 1.4 ([RSY08])**.** *Let $i \in [n]$, and let $h_1, h_2, \ldots, h_\ell$ be the polynomials computed by the gates in $\mathcal{L}_i$. Then, there exist multilinear polynomials $g_1, g_2, \ldots, g_\ell, g$ such that*

$$f_i = \sum_{j \in [\ell]} g_j \cdot h_j + g \tag{1.5}$$

where

- *For every $j \in [\ell]$, h_j and g_j are variable disjoint.*
- *The degree of g is at most $O(\log n)$.*

Observe that Equation 1.5 is basically a decomposition of a potentially-hard polynomial f_i in terms of the sum of products of multilinear polynomials in an intermediate number of variables. The goal is to show that for an appropriate explicit f_i, the number of summands on the right hand side of Equation 1.5 cannot be too small. A similar scenario also appears in the multilinear formula lower bounds and bounded depth multilinear formula lower bounds of [Raz09, Raz06, RY09] (albeit with some key differences). Hence, a natural approach at this point would be to use the tools in [Raz09, Raz06, RY09], namely the rank of the *partial derivative matrix*, to attempt to prove this lower bound. We refer the reader to Section 2.2 for the definitions and properties of the partial derivative matrix and proceed with the overview. For each $j \in [\ell]$, let the polynomial h_j in Lemma 1.4 depend on the variables $S_j \subseteq X$. The key technical step in the rest of the proof is to show that there is a partition of the set of variables $X = \{x_1, x_2, \ldots, x_n\}$ into Y and Z such that $|Y| = |Z|$ and for every $j \in [\ell]$, $||S_j \cap Y| - |S_j \cap Z|| \geq \Omega(\log n)$. In [RSY08], the authors show that there is an absolute constant $\varepsilon > 0$ such that if $\ell \leq \varepsilon n^{1/3}/\log n$, then there is an equipartition of X which *unbalances* all the sets $\{S_j : j \in [\ell]\}$ by at least $\Omega(\log n)$. Our key technical contribution (Theorem 1.3) in this paper is to show that as long as $\ell \leq \varepsilon n/\log n$, there is an equipartition which unbalances all the S_j's by at least $\Omega(\log n)$. This implies an $\Omega(n/\log n)$ on the size of each set $\mathcal{U}_i$, and thus an $\Omega(n^2/\log^2 n)$ lower bound on the circuit size.

Before we dive into a more detailed discussion on the overview and main ideas in the proof of Theorem 1.3 in the next section, we would like to remark that the lower bound question in Equation 1.5 seems to be a trickier question than what is encountered while proving multilinear formula lower bounds [Raz09, Raz06] or bounded depth syntactically multilinear circuit

lower bounds [RY09]. The main differences are that in the proofs in [Raz09, Raz06, RY09], the sets S_j have a stronger guarantee on their size (at least $n^{\Omega(1)}$ and at most $n - n^{\Omega(1)}$), and each of the summands on the right has *many* variable disjoint factors and not just two factors as in Equation 1.5. For instance, in the formula lower bound proofs the number of variable disjoint factors in each summand on the right is $\Omega(\log n)$, and for constant depth circuit lower bounds it is $n^{\Omega(1)}$. Together, these properties make it possible to show much stronger lower bounds on ℓ. In particular, it is known that a *random* equipartition works for these two applications, in the sense that it unbalances sufficiently many factors in each summand, thereby implying that the rank of the partial derivative matrix of the polynomial is small. Hence, for an appropriate[6] f_i, the number of summands must be large. However, since a set of size $O(\log n)$ is balanced under a random equipartition with probability $\Omega(1/\sqrt{\log n})$ and the identity in Equation 1.5 involves just two variable disjoint factors, taking a random equipartition would not enable us to prove any meaningful bounds.

Proof sketch of Theorem 1.3

Recall that our task is, given a small collection of subsets of $[n]$, to find a balanced partition which is unbalanced on each of the sets. Equivalently, we would like to prove that if $\mathcal{F}$ is a family of subsets such that every balanced partition balances at least one set in $\mathcal{F}$, then $|\mathcal{F}|$ must be large (of course, $\mathcal{F}$ must satisfy the conditions in Theorem 1.3).

We first sketch the proof of a special case (which suffices for the main application here), when $n = 4p$ and p is a prime. For the sake of simplicity, suppose also that all subsets $S \in \mathcal{F}$ are of even size, and assume further that for every subset $T \subseteq [n]$ of size $n/2$ there exists $S \in \mathcal{F}$ such that T completely balances S, namely, $|T \cap S| = |S|/2$. One possible approach to obtain lower bounds on $|\mathcal{F}|$ is via an application of the polynomial method as done, for example, in [ABCO88]. Define the following polynomial over, say, the rationals:

$$f(x_1, \ldots, x_n) = \prod_{S \in \mathcal{F}} (\langle x, \mathbb{1}_S \rangle - |S|/2).$$

By the assumption on $\mathcal{F}$, the polynomial f evaluates to 0 over all points in $\{0,1\}^n$ with Hamming weight exactly $n/2$. We can also argue, using the assumption on the set sizes in $\mathcal{F}$, that f is not identically zero, and clearly $\deg(f) \leq |\mathcal{F}|$. Thus, a lower bound on $\deg(f)$ translates to a lower bound on $|\mathcal{F}|$.

This idea, however, seems like a complete nonstarter, since there exists a degree 1 non-zero polynomial which evaluates to 0 over the middle layer of $\{0,1\}^n$, namely, $\sum_i x_i - n/2$.

A very clever solution to this potential obstacle was found by Hegedűs [Heg10]. Suppose $n = 4p$ for some prime p. The main insight in [Heg10] is to consider the polynomial f over $\mathbb{F}_p$, and to add the requirement that there exists some $z \in \{0,1\}^{4p}$, of Hamming weight *exactly* $3p$, such that $f(z) \neq 0$. This requirement rules out the trivial example $\sum_i x_i - n/2$, and Hegedűs was able to show that the degree of any polynomial with these properties must be at least $p = n/4$ (see Lemma 2.1 for the complete statement).

We are thus left with the task of proving that our polynomial evaluates to a non-zero value over some point $z \in \{0,1\}^{4p}$ of Hamming weight $3p$. This turns out to be not very hard to show, assuming each set is of size at least, say, $100 \log n$ and at most $n - 100 \log n$, by choosing a random such vector z. Indeed, it is not surprising that it is much easier to directly show that a highly

[6] f_i is chosen so that the the partial derivative matrix for f_i is of full rank for *every* equipartition.

unbalanced partition of $[n]$ (into $3n/4$ vs $n/4$) unbalances all the sets $\mathcal{F}$.[7]

As mentioned earlier, the case $n = 4p$ and $\tau \geq 100 \log n$ in Theorem 1.3 is considerably easier to prove and suffices for the application to circuit lower bounds. Proving this theorem for every even n and every $\tau \geq 1$ requires further technical ideas. We postpone this discussion to Section 3.2.

Organization of the paper

In the rest of the paper, we set up some notation and discuss some preliminary notions in Section 2, prove Theorem 1.3 in Section 3 and complete the proof of Theorem 1.1 in Section 4. Throughout the paper we assume, whenever this is needed, that n is sufficiently large, and make no attempts to optimize the absolute constants.

2 Preliminaries

For $n \in \mathbb{N}$, we denote $[n] = \{1, 2, \ldots, n\}$. For a prime p, we denote by $\mathbb{F}_p$ the finite field with p elements. For two integers i, j with $i \leq j$, we denote $[i, j] = \{a \in \mathbb{Z} : i \leq a \leq j\}$. The characteristic vector of a set $S \subseteq [n]$ is denoted by $\mathbb{1}_S \in \{0,1\}^n$.

As is standard, $\binom{[n]}{k}$ denotes the family $\{S \subseteq [n] : |S| = k\}$.

For an even $n \in \mathbb{N}$ and $Y \subseteq [n]$ such that $|Y| = n/2$, we call Y a *balanced partition* of $[n]$, with the implied meaning that Y partitions $[n]$ evenly into Y and $[n] \setminus Y$. The *imbalance* of a set $S \subseteq [n]$ under Y is $d_Y(S) := ||Y \cap S| - |S|/2|$. Observe the useful symmetry $d_Y(S) = d_Y([n] \setminus [S])$, which follows from the fact that $|Y| = n/2$. We say S is τ-unbalanced under Y if $d_Y(S) \geq \tau$.

We use the following lemma from [Heg10].

Lemma 2.1 ([Heg10])**.** *Let p be a prime, and let $f \in \mathbb{F}_p[x_1, \ldots, x_{4p}]$ be a polynomial. Suppose that for all $Y \in \binom{[4p]}{2p}$, it holds that $f(\mathbb{1}_Y) = 0$, and that there exists $T \subseteq [4p]$ such that $|T| = 3p$ and $f(\mathbb{1}_T) \neq 0$. Then $\deg(f) \geq p$.*

The proof of Lemma 2.1 in [Heg10] relies on the description of Gröbner basis for ideals of polynomials in $\mathbb{F}_p[x_1, \ldots, x_{4p}]$ which vanish on all points in $\{0,1\}^n$ of weight equal to $2p$. A complete description of the reduced Gröbner basis for such ideals was given by Hegedűs and Rónyai [HR03] and their proof builds up on a number of earlier partial results [ARS02, FG06] on this problem.

Here, we give an elementary proof for this lemma, which only relies on basic linear algebra, and was communicated to us by Srikanth Srinivasan. We begin with the following basic and well known fact, which we also prove for completeness.

Lemma 2.2. *Let V denote the subspace of functions $f : \{0,1\}^{4p} \to \mathbb{F}_p$ which can be represented as polynomials of degree at most $p-1$. Let $P : \{0,1\}^{4p} \to \mathbb{F}_p$ denote the parity function, i.e., $P(x_1, \ldots, x_{4p}) = (-1)^{\sum_i x_i}$. Then, the set*

$$B = \{P(\mathbf{x}) \cdot M(\mathbf{x}) : M \text{ is a multilinear monomial of degree at most } 3p\}$$

forms a basis for $V^\perp$ (with respect to the standard inner product $\langle f, g \rangle = \sum_{\mathbf{x} \in \{0,1\}^{4p}} f(\mathbf{x})g(\mathbf{x})$).

Proof. The elements of B are clearly linearly independent, as any non-zero linear combination of them is of the form $P(\mathbf{x}) \cdot Q(\mathbf{x})$ for a non-zero multilinear polynomial Q. To see that every element

[7] In our case, we need to argue that the imbalance is non-zero modulo p, which adds an extra layer of complication, although again, one which is not hard to solve.

$P(\mathbf{x})M(\mathbf{x})$ in B is perpendicular to V, let M' be a monomial of degree at most $p-1$. By definition,

$$\left\langle P(\mathbf{x})M(\mathbf{x}), M'(\mathbf{x})\right\rangle = \sum_{\mathbf{x}\in\{0,1\}^{4p}} P(\mathbf{x})\cdot M(\mathbf{x})M'(\mathbf{x}). \tag{2.3}$$

Since $M\cdot M'$ is a monomial of degree at most $4p-1$, there exists $i\in[4p]$ such that $M\cdot M'$ does not depend on x_i. Pair all elements of $\{0,1\}^{4p}$ according to the i-th axis, i.e., $\mathbf{x}$ is paired with $\mathbf{x}+e_i$, and observe that $P(\mathbf{x}) = -P(\mathbf{x}+e_i)$ while $M\cdot M'(\mathbf{x}) = M\cdot M'(\mathbf{x}+e_i)$, which implies that these two cancel each other out in the sum (2.3), so the inner product is 0.

The claim that B is a basis now follows from a dimension counting argument. □

Proof of Lemma 2.1. Let $f\in\mathbb{F}_p[x_1,\ldots,x_{4p}]$. Consider f as a function from $\{0,1\}^{4p}$ from $\mathbb{F}_p$, and suppose f is such that $f(\mathbf{y})=0$ for all $\mathbf{y}\in\{0,1\}^{4p}$ of weight $2p$, and $\deg f\le p-1$. Let $\mathbf{z}\in\{0,1\}^{4p}$ of weight $3p$. We will show that $f(\mathbf{z})=0$, thus implying the statement of the lemma.

Let V as in Lemma 2.2. We will argue that there is $g\in V^\perp$ such that $g(\mathbf{z})\neq 0$, and every other non-zero of g is obtained only on vectors $\mathbf{y}$ of weight $2p$. It then follows that

$$0 = \langle f, g\rangle = \sum_{\mathbf{x}\in\{0,1\}^{4p}} f(x)g(x) = f(\mathbf{z})g(\mathbf{z}),$$

as for every $\mathbf{x}\neq\mathbf{z}$, either $f(\mathbf{x})$ or $g(\mathbf{x})=0$, by the assumptions on f and g. This implies $f(\mathbf{z})=0$.

To show the existence of g, suppose without loss of generality $\mathbf{z}_i=1$ for $1\le i\le 3p$ and $\mathbf{z}_i=0$ for $3p+1\le i\le 4p$. Let

$$\tilde{g}(\mathbf{x}) = \left(1-\left(\sum_{i=1}^{4p} x_i\right)^{p-1}\right)\cdot(x_1\cdots x_{p+1})\cdot((1-x_{3p+1})\cdots(1-x_{4p}))\,. \tag{2.4}$$

The first term in (2.4) guarantees that $\tilde{g}$ is zero whenever the weight of $\mathbf{x}$ does not divide p. The second term in (2.4) guarantees that $\tilde{g}$ is zero on vectors of weight 0 and p. The third term in (2.4) guarantees that $\tilde{g}$ is zero on every vector of weight $4p$ or $3p$ other than $\mathbf{z}$. Furthermore, $\tilde{g}(\mathbf{z})=1$.

Since $\deg\tilde{g}\le 3p$, it follows from Lemma 2.2 that $g := P\cdot\tilde{g}\in V^\perp$ and has the required properties. □

2.1 Hypergeometric distribution

For parameters N, M, k, where $N\ge M$, by $\mathcal{H}(M,N,k)$, we denote the distribution of $|S\cap T|$, where S is any fixed subset of $[N]$ of size M, and T is a uniformly random subset of $[N]$ of size equal to k. Clearly,

$$\Pr[|S\cap T| = i] = \frac{\binom{M}{i}\binom{N-M}{k-i}}{\binom{N}{k}}\,.$$

The expected value of $|S\cap T|$ under this distribution is equal to kM/N. We need the following tail bound of hypergeometric distribution for our proof.

Lemma 2.5 ([Ska13])**.** *Let N, M, k, and $\mathcal{H}(M,N,k)$ be as defined above. Then, for every t*

$$\Pr[||S\cap T| - kM/N|\ge tk] \le \mathrm{e}^{-2t^2k}\,.$$

Lemma 2.6 (Hoeffding's inequality, [AS16])**.** *Let $X_1, X_2, \ldots, X_n$ be independent random variables taking values in $\{0,1\}$. Then,*

$$\Pr\left[\left|\sum_{i=1}^n X_i - \mathbb{E}[\sum_{i=1}^n X_i]\right|\ge t\right] \le 2\exp(-2t^2/n)\,.$$

2.2 Partial derivative matrix

For a circuit Ψ, we denote by $|\Psi|$ the size of Ψ, namely, the number of gates in it. For a gate v, we denote by X_v the set of variables that occur in the subcircuit rooted at v.

Let $X = \{x_1, \ldots, x_n\}$ be a set of variables, $Y \subseteq X$ (not necessarily of size $n/2$) and let $Z = X \setminus Y$. For a multilinear polynomial $f(X) \in \mathbb{F}[X]$, we define the *partial derivative matrix* of f with respect to Y, Z, denoted $M_{Y,Z}(f)$, as follows: the rows of M are indexed by multilinear monomials in Y. the columns of M are indexed by multilinear monomials in Z. The entry which corresponds to (m_1, m_2) is the coefficient of the monomial $m_1 \cdot m_2$ in f. We define $\mathrm{rank}_{Y,Z}(f) = \mathrm{rank}(M_{Y,Z}(f))$.

The following properties of the partial derivative matrix are easy to prove and well-documented (see, e.g., [RSY08]).

Proposition 2.7. *The following properties hold:*

1. *For every multilinear polynomial $f(X) \in \mathbb{F}[X]$, $Y \subseteq X$ and $Z = X \setminus Y$, $\mathrm{rank}_{Y,Z}(f) \leq \min\left\{2^{|Y|}, 2^{|Z|}\right\}$.*
2. *For every two multilinear polynomials $f_1(X), f_2(X) \in \mathbb{F}[X]$ and for every partition $X = Y \sqcup Z$, $\mathrm{rank}_{Y,Z}(f_1 + f_2) \leq \mathrm{rank}_{Y,Z}(f_1) + \mathrm{rank}_{Y,Z}(f_2)$.*
3. *Let $f_1 \in \mathbb{F}[X_1]$ and $f_2 \in \mathbb{F}[X_2]$ be multilinear polynomials such that $X_1 \cap X_2 = \emptyset$. Let $Y_i \subseteq X_i$ and $Z_i = X_i \setminus Y_i$ for $i \in \{1, 2\}$. Set $Y = Y_1 \cup Y_2, Z = Z_1 \cup Z_2$. Then $\mathrm{rank}_{Y,Z}(f_1 \cdot f_2) = \mathrm{rank}_{Y_1,Z_1}(f_1) \cdot \mathrm{rank}_{Y_2,Z_2}(f_2)$.*
4. *Let $f(X) \in \mathbb{F}[X]$ be a multilinear polynomial such that $X = Y \sqcup Z$ and $|Y| = |Z| = n/2$. Suppose $\mathrm{rank}_{Y,Z}(f) = 2^{n/2}$, and let $g = \partial f / \partial x$ for some $x \in X$. Then $\mathrm{rank}_{Y,Z}(g) = 2^{n/2-1}$.*
5. *Let $f(X) \in \mathbb{F}[X]$ be a multilinear polynomial of total degree d. Then for every partition $X = Y \sqcup Z$ such that $|Y| = |Z| = n/2$, $\mathrm{rank}_{Y,Z}(f) \leq 2^{(d+1)\log(n/2)}$.*

3 Unbalancing sets under a balanced partition

In this section, we prove Theorem 1.3. We start by proving a special case (see Theorem 3.1 below) when n equals $4p$ for some prime p, and $\tau \geq \Omega(\log n)$. This special case already suffices for the application to the proof of Theorem 1.1 (for infinitely many values of n), and has a somewhat simpler proof. We then move on to prove the case for general n and τ, which while being similar to the proof of Theorem 3.1, needs some additional ideas and care.

3.1 Special case : $n = 4p$ and $\tau \geq \Omega(\log n)$

Theorem 3.1. *Let p be a large enough prime, and let $\log p \leq \tau \leq p/1000$. Let $S_1, \ldots, S_m \subseteq [4p]$ be sets such that for all $i \in [m]$, $100\tau \leq |S_i| \leq 4p - 100\tau$. Further, assume that for every balanced partition Y of $[4p]$ there exists $i \in [m]$ such that $d_Y(S_i) < \tau$. Then, $m \geq \frac{1}{2} \cdot p/\tau$.*

We start with the following lemma, which shows that a small collection of sets can be unbalanced (modulo p) by a partition which is very unbalanced.

Lemma 3.2. *Let p be a large enough prime, and let $\log p \leq \tau \leq p/1000$. Let $S_1, \ldots, S_m \subseteq [4p]$ be sets such that for all $i \in [m]$, $100\tau \leq |S_i| \leq 2p$. Assume further $m \leq p$. Then, there exists $T \subseteq [4p]$, $|T| = 3p$ such that for all $i \in [m]$ and for all $-\tau + 1 \leq t \leq \tau$, $|S_i \cap T| \not\equiv \lfloor |S_i|/2 \rfloor + t \bmod p$.*

To prove Lemma 3.2, we use the following two technical claims. Let $\mu_{3/4}$ denote the probability distribution on subsets of $[4p]$ obtained by putting each $j \in [4p]$ in T with probability $3/4$, independently of all other elements.

Claim 3.3. *For a random set $T \sim \mu_{3/4}$, $\Pr[|T| = 3p] = \Theta(1/\sqrt{p})$.*

Proof. The probability that $|T| = 3p$ is given by $\binom{4p}{3p} \cdot (3/4)^{3p} \cdot (1/4)^{p}$, which is $\Theta(1/\sqrt{p})$, by Stirling's approximation. □

Claim 3.4. *Let $\log p \leq \tau \leq p/1000$ and let $S \subseteq [4p]$ such that $100\tau \leq |S| \leq 2p$. For a random set $T \sim \mu_{3/4}$, the probability that for some integer $-\tau + 1 \leq t \leq \tau$ it holds that $|T \cap S_i| = \lfloor |S_i|/2 \rfloor + t \bmod p$ is at most $1/p^5$.*

Proof. Denote $s = |S|$. Then $\mathbb{E}[|T \cap S|] = 3s/4$. We say T is bad for S if $|T \cap S| = \lfloor s/2 \rfloor + t + kp$ for some $-\tau \leq t \leq \tau + 1$ and $k \in \mathbb{Z}$. We claim this in particular implies that $||T \cap S_i| - 3s/4| \geq s/5$. Indeed, since $|T \cap S|$ is an integer in the interval $[0, 2p]$, and by the bounds on s, the only cases needed to be analyzed are $k = 0, \pm 1$.

If $|T \cap S| = \lfloor s/2 \rfloor + t - p$, then clearly $|T \cap S| \leq \lfloor s/2 \rfloor$ which implies the statement.

If $|T \cap S| = \lfloor s/2 \rfloor + t + p$, then, as $s \leq 2p$ and $\tau \leq s/100$,

$$|T \cap S| - 3s/4 \geq -s/4 - 1 + t + p \geq p/2 + t - 1 \geq s/4 + t - 1 \geq s/5$$

(The "-1" accounts for the fact that $s/2$ might not be an integer).

Finally, if $|T \cap S| = \lfloor s/2 \rfloor + t$, it holds that

$$|T \cap S| \leq s/2 + \tau \leq s/2 + 2s/100,$$

which again implies the statement.

By Chernoff Bound (see, e.g., [AS16]), $\Pr[||T \cap S_i| - 3s/4| \geq s/5] \leq 2^{-|S|/20} \leq 1/p^5$, hence T is bad for S with at most that probability. □

The proof of Lemma 3.2 is now fairly immediate.

Proof of Lemma 3.2. Pick $T \sim \mu_{3/4}$. By Claim 3.3, $|T| = 3p$ with probability $\Theta(1/\sqrt{p})$. Recall that T is bad for S_i if $|T \cap S_i| = \lfloor |S_i|/2 \rfloor + t \bmod p$ for $t \in \{-\tau + 1, \ldots, \tau\}$. By Claim 3.3, for each S_i, T is bad for S_i with probability at most $1/p^5$. Hence, the probability that there exists $i \in [m]$ such that T is bad for S_i is at most $m/p^5 \leq 1/p^4$.

It follows that with probability at most $1 - \Theta(1/\sqrt{p}) + 1/p^4 < 1$, either $|T| \neq 3p$ or T is bad for some S_i, and hence there exists a selection of T such that $|T| = 3p$ and T is good for all S_i's. □

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. Let $S_1, \ldots, S_m$ be a collection of sets as stated in the theorem. Since $d_Y(S_j) = d_Y([n] \setminus S_j)$, we can assume without loss of generality, by possibly replacing a set with its complement, that $|S_j| \leq 2p$ for all $j \in [m]$. We may further assume $m \leq p$ as otherwise the statement directly follows. For $j \in [m]$, define the following polynomials over $\mathbb{F}_p$:

$$B_j(x_1, \ldots, x_{4p}) = \prod_{t=-\tau+1}^{\tau} (\langle x, \mathbb{1}_{S_j} \rangle - \lfloor |S_j|/2 \rfloor - t),$$

where $x = (x_1, \ldots, x_{4p})$ and $\langle u, v \rangle = \sum u_i v_i$ is the usual inner product. Further, define

$$f(x_1, \ldots, x_{4p}) = \prod_{j=1}^{m} B_j(x_1, \ldots, x_{4p}),$$

as a polynomial over $\mathbb{F}_p$.

By assumption, for every $Y \in \binom{[4p]}{2p}$, $f(\mathbb{1}_Y) = 0$. This follows because $\langle \mathbb{1}_Y, \mathbb{1}_{S_j} \rangle = |Y \cap S_j|$, and by assumption, for some j is holds that $d_Y(S_j) < \tau$, so it must be that $|Y \cap S_j| - \lfloor |S_j|/2 \rfloor \in \{-\tau + 1, \ldots, 0, \ldots, \tau\}$, so that $B_j(\mathbb{1}_Y) = 0$.

Furthermore, Lemma 3.2 guarantees the existence of a set $T \in \binom{[4p]}{3p}$ such that $f(\mathbb{1}_T) \neq 0$, as the set T from Lemma 3.2 satisfies the property that $(\langle \mathbb{1}_T, \mathbb{1}_{S_j} \rangle - \lfloor |S_j|/2 \rfloor - t) \neq 0 \bmod p$ for all $-\tau + 1 \leq t \leq \tau$ and for all $j \in [m]$.

By Lemma 2.1, $\deg(f) \geq p$, and by construction, $\deg(f) \leq 2\tau \cdot m$, which implies the desired lower bound on m. □

3.2 General n and τ

In this section, we extend Theorem 3.1 for a more general range of parameters, by proving the following.

Theorem 3.5. *Let n be a large enough even natural number, and let $\tau \in \{1, 2, \ldots, n/10^6\}$ be a parameter. Let $S_1, S_2, \ldots, S_m \subseteq [n]$ be sets such that for each $i \in [m]$, $2\tau \leq |S_i| \leq n - 2\tau$. Furthermore, assume that for every balanced partition Y of $[n]$, there exists an i such that $d_Y(S_i) < \tau$. Then, $m \geq \frac{1}{10^5} \cdot n/\tau$.*

We remark that Theorem 3.1 suffices for the application to circuit lower bounds, and thus, a reader who is more interested in that aspect of this work may safely skip to Section 4.

Recall that in Theorem 3.1 we have required the universe size n to be of the form $4p$ for a prime p, and the sets $S_1, \ldots, S_m$ to be of size at least logarithmic in n (as commented earlier, we may assume $|S_i| \leq n/2$ for every i, by possibly replacing S_i with its complement).

Our strategy for general even[8] n and general τ will be very similar for the previous special case. In order to apply the useful Lemma 2.1, we start by "forcing" the universe size to be of the form $4p$. This is done by picking the largest number of the form $4p$ which is smaller than n (known results about the distribution of prime numbers guarantee the existence of such a prime such that $n - 4p \leq n^{0.6}$). We then randomly pick a subset of $A \subset [n]$ of size $n - 4p$ avoiding all the small sets and partition A in an arbitrary balanced manner. Such a subset is guaranteed, with high probability, to have a small intersection with every S_i, and thus for every such set the values of very few elements have been determined. Again, this intersection property is easier to show, by standard concentration bounds, when the sets S_i are somewhat large, whereas in our case they can be small. However, the fact that $|A|$ itself is sublinear in n enables us to handle all cases.

We now denote $\tilde{S}_i = S_i \setminus A$ and $\widetilde{[n]} = [n] \setminus A$, and, as before, we would like to find a set $T \subseteq \widetilde{[n]}$ of size exactly $3p$ that is unbalanced, modulo p, on every $\tilde{S}_i$ (and since $\tilde{S}_i$ is a very large subset of S_i, this property will extend to S_i itself). A naïve random choice, as is done in the proof of Theorem 3.1, will not work, since the probability of failure for very small sets will be too large to apply a union bound over all sets. Thus, we pick T using a different, and slightly more complicated, random procedure.

[8] In order to talk about balanced partitions of the universe, n clearly must be even. However, our techniques can be easily extended to odd integers, if one is willing to replace balanced partitions by almost-balanced partitions, that is, partitions $[n] = Y \sqcup Z$ such that $||Y| - |Z|| = 1$. We omit the straightforward details.

Given such T and A, the proof follows from a similar construction of a polynomial in a similar application of Lemma 2.1. We now provide the details.

We start by proving the existence of a set A as described above.

Lemma 3.6. *Let $\tau \geq 1$ be an integer and $S_1, S_2, \ldots, S_m$ be subsets of $[n]$, such that $m \leq 10^{-5}n/\tau$. Then, for every integer $a \leq n^{0.6}$, there exists an $A \subseteq [n]$ of size exactly a such that for every $i \in [m]$, $|A \cap S_i| \leq 0.01\,|S_i|$. Moreover, for each $i \in [m]$, if $|S_i| \leq 10^4\tau$, then $A \cap S_i = \emptyset$.*

Proof. Let $L = \bigcup_{i:|S_i|\leq 10^4\tau} S_i$ and let $\ell = |L|$. Since $m \leq 10^{-5}n$, we know that $\ell \leq m \cdot 10^4 \leq n/10$. Let A to be a uniformly random subset of $[n] \setminus L$ of size a.

We now show that with high probability A satisfies $|A \cap S_i| \leq 0.01\,|S_i|$ for every $i \in [m]$. We consider three cases.

- **Small sets: $|S_j| \leq 10^4\tau$.** By the choice of A, we know that A is disjoint from all subsets of size at most $10^4\tau$.
- **Large sets: $|S_j| \geq n^{0.31}$.** For any fixed set S_i of size at least $n^{0.31}$, by Lemma 2.5, we know that
$$\Pr\left[|A \cap S_i| - |A|\,|S_i|/(0.9n) \geq 0.009\,|S_i|\right] \leq \exp(-\Omega(|S_i|^2/|A|))\,.$$
Since $|A| \leq n^{0.6}$ and $|S_i| \geq n^{0.31}$, this probability is at most $\exp(-\Omega(n^{0.02}))$. Thus, by a union bound, we know that with probability at least $1 - \exp(-\Omega(n^{0.02}))$, for each S_i with $|S_i| \geq n^{0.31}$, $|A \cap S_i| \leq 0.01\,|S_i|$.
- **Sets of intermediate size: $10^4\tau \leq |S_j| \leq n^{0.31}$.** We now argue that for all such sets, $|A \cap S_i| \leq 100$, with high probability.

  To this end, we first upper bound the probability that the set A contains a fixed set S of size 100, and then take a union bound over all sets S of size $s = 100$ which are a subset of some S_i of intermediate size. Let S be a fixed set of size 100. Then,

$$\begin{aligned}
\Pr[S \subseteq A] &\leq \frac{\binom{n-\ell-s}{a-s}}{\binom{n-\ell}{a}} \\
&= \frac{(n-\ell-s)!}{(a-s)!(n-\ell-a)!} \cdot \frac{a!(n-\ell-a)!}{(n-\ell)!} \\
&= \frac{(n-\ell-s)!}{(a-s)!} \cdot \frac{a!}{(n-\ell)!} \\
&= \frac{(n-\ell-s)!}{(n-\ell)!} \cdot \frac{a!}{(a-s)!} \\
&\leq \left(\frac{a}{n-\ell-s}\right)^s \\
&\leq \left(\frac{n^{0.6}}{n-0.1n-n^{0.6}}\right)^s \qquad \text{(using bounds on } \ell \text{ and } a\text{)} \\
&\leq n^{-0.39s} \\
&\leq n^{-39} \qquad \text{(using s = 100)}
\end{aligned}$$

  For each S_i of size at most $n^{0.31}$ there are at most $(n^{0.31})^{100}$ subsets of size 100. Therefore, by a union bound, the probability that $|A \cap S_i| \geq 100$ for any subset S_i of size at most $n^{0.31}$ is at most $n^{-39} \cdot n \cdot n^{31} = n^{-7}$.

A union bound over all three cases completes the proof of the lemma. □

Having shown the existence of the set A as described in the proof outline, we turn to show the existence of a set T.

Lemma 3.7. *Let n be a natural number, p be a prime satisfying $n - n^{0.6} \leq 4p \leq n$ and let τ be an integer satisfying $1 \leq \tau \leq p/10^5$. Let $S_1, S_2, \ldots, S_m$ be subsets of $[n]$, such that $m \leq 10^{-5}n/\tau$ and for every $j \in [m]$, $2\tau \leq |S_j| \leq n/2$. Let $A \subseteq [n]$ be a set of size $n - 4p$ such that for every $j \in [m]$, $|A \cap S_j| \leq 0.01\,|S_j|$ and A is disjoint from all sets S_i of size at most $10^4\tau$. Let B be an arbitrary subset of A. Then, there exists a set $T \subseteq [n] \setminus A$ of size exactly $3p$, such that for every $j \in [m]$, if $|S_j| > 2\tau$ then for every integer t with $-\tau < t \leq \tau$, it holds that $|(T \cup B) \cap S_j| \neq \lfloor |S_j|/2 \rfloor + t \bmod p$. If $|S_j| = 2\tau$, the same holds for $-\tau < t < \tau$.*

Proof. Denote $\widetilde{[n]} = [n] \setminus A$, and $\tilde{S}_i = S_i \setminus A$ for all $i \in [m]$. We note that if $|S_i| \leq 10^4\tau$, then $\tilde{S}_i = S_i$. We construct the set T by a randomized algorithm, which consists of several steps. In the first step, we greedily select a small number of elements from each set $\tilde{S}_i$. The purpose of this step is to guarantee that $|T \cap S_i|$ is sufficiently far from 0, for every i. Next, we pick each of the remaining elements of $\widetilde{[n]}$ to T with probability 0.65. This constant is chosen so that with high probability (assuming $|\tilde{S}_i|$ is sufficiently large), the intersection $|T \cap \tilde{S}_i|$ is non-zero modulo p (and since $|\tilde{S}_i|$ and $|S_i|$ are very close, the same holds for $|T \cap S_i|$), and also with high probability the number of elements we have picked so far does not exceed $3p$.

The next step is again a deterministic, greedy step, which adds to T sufficiently many elements from each "bad" set S_i. Those are the sets of which too few elements were picked before. By standard concentration bounds, we do not expect to have many such large sets, and thus again we can control the number of elements added in this step.

Finally, assuming the number of elements that were picked so far is less than $3p$ (which happens with high probability), we add arbitrary elements to our set so that it will be of size exactly $3p$. Of course, we also have to argue that this step preserves the previous intersection requirements. This follows from the fact that we do not expect to add many elements in this step.

We now provide the more formal details. T is constructed using the following randomized algorithm.

- For every $j \in [m]$ such that $|S_j| \leq 6000\tau$, we add all elements of S_j to T_1. We then take 6000τ arbitrary elements from the remaining sets among $\tilde{S}_1, \tilde{S}_2, \ldots, \tilde{S}_m$. Since $m \leq 10^{-5}n/\tau$, the size of T_1 is at most $0.06n$. Without loss of generality, we take T_1 to be of size equal to $0.06n$.
- Let T_2 be the set obtained by picking every element in $\widetilde{[n]} \setminus T_1$ independently with probability 0.65.
- For every $j \in [m]$, such that $|S_j \cap (T_1 \cup T_2 \cup B)| \leq 0.52\,|S_j|$, include all elements in $\tilde{S}_j \setminus (T_1 \cup T_2)$ in the set T_3.
- If $|T_1| + |T_2| + |T_3| > 3p$, abort. Else, we add $3p - |T_1| - |T_2| - |T_3|$ arbitrary elements from $\widetilde{[n]} \setminus (T_1 \cup T_2 \cup T_3)$ into the set T_4.
- Let $T = T_1 \cup T_2 \cup T_3 \cup T_4$.

We will now argue that with a high probability, the algorithm above outputs a set T which satisfies the desired properties. To this end, we need the following claims, whose proofs we defer to the end of this section. The probabilities in these claims are all taken over the choice of T_2, which is the only randomized step in the algorithm.

Claim 3.8. *With probability at least $1 - n^{-5}$, all of the following events happen.*

- $0.64n \leq |T_2| \leq 0.66n$.
- $\forall j \in [m]$, *such that* $|S_j| \geq 1000 \log n$, $\left|\tilde{S}_j \cap T_2\right| \in [0.52\,|S_j|, 0.74\,|S_j|]$.
- *For every* $j \in [m]$, *if* $|S_j| \leq 6000\tau$, *then* $S_j \subseteq T$.
- *For every* $j \in [m]$, *if* $|S_j| \geq 6000\tau$, *then* $|S_j \cap T| \geq \max\{6000\tau, 0.52\,|S_j|\}$.

Claim 3.9 (T_3 is typically small)**.**

$$\Pr[|T_3| \leq 0.01n] \geq 0.99\,.$$

Claim 3.10 (T_4 is typically small)**.**

$$\Pr[|T_4| \leq 0.05n] \geq 1 - n^{-5}\,.$$

Probability of aborting and size of T. The algorithm aborts only in the case that $|T_1| + |T_2| + |T_3| > 3p$. We know that with probability 1, $|T_1| \leq 0.06n$. It follows from Claim 3.8 that with probability at least $1 - n^{-5}$, $|T_2| \leq 0.66n$ and from Claim 3.9 that with probability at least 0.99, $|T_3| \leq 0.01n$. Thus, with probability at least 0.98, $|T_1| + |T_2| + |T_3| \leq 0.73n$. Since $4p \leq n \leq 4p + O(p^{0.6})$, with probability at least 0.98, $|T_1| + |T_2| + |T_3| \leq 3p$. Also, whenever the algorithm does not abort, the set T_4 is picked so that T output by the algorithm satisfies $|T| = 3p$.

Intersection properties of T. For the rest of this argument, we assume that T_1, T_2, T_3, T_4 satisfy the properties in Claim 3.8, Claim 3.9 and Claim 3.10. We now argue that for every $j \in [m]$ it holds that $|(T \cup B) \cap S_j| \neq \lfloor |S_j|/2 \rfloor + t \bmod p$ for every integer t in the range specified in the statement of the Lemma.

We consider some cases based on the size of S_j.

- **Very small sets : $2\tau \leq |S_j| \leq 6000\tau$.** From Claim 3.8, all such sets are completely contained in T. Thus,
$$|(T \cup B) \cap S_j| - (\lfloor |S_j|/2 \rfloor + t) = \lceil |S_j|/2 \rceil - t\,.$$
Since $1 \leq \tau \leq p/10^5$, this remains non-zero modulo p for every $-\tau < t \leq \tau$ if $|S_j| > 2\tau$, and for every $-\tau < t < \tau$ if $|S_j| = 2\tau$.

- **Small sets : $6000\tau < |S_j| \leq 10^4\tau$.** From Claim 3.8, we know that for every $j \in [m]$, $|S_j \cap T| \geq 6000\tau$. We get that for every $-\tau < t \leq \tau$,
$$1 \leq |(T \cup B) \cap S_j| - (\lfloor |S_j|/2 \rfloor + t) \leq (10^4 + 1)\tau$$
Since $\tau \leq p/10^5$, $|(T \cup B) \cap S_j| - (\lfloor |S_j|/2 \rfloor + t)$ is non-zero modulo p for each $-\tau < t \leq \tau$.

- **Sets of intermediate size : $10^4\tau < |S_j| \leq 1000 \log n$.** Since by Claim 3.8, $|S_j \cap T| \geq 0.52\,|S_j|$, we get that for every $-\tau < t \leq \tau$,
$$198\tau \leq |(T \cup B) \cap S_j| - (\lfloor |S_j|/2 \rfloor + t) \leq 1000 \log n.$$
Thus, $|(T \cup B) \cap S_j| - (\lfloor |S_j|/2 \rfloor + t)$ remains non-zero modulo p.

- **Large sets :** $\max\{1000 \log n, 10^4\tau\} \leq |S_j| \leq n/2$**.** For such large sets, from Claim 3.8, Claim 3.9 and Claim 3.10, we know that

$$0.52\,|S_j| \leq |(T \cup B) \cap S_j| = \sum_{k=1}^{4} |T_k \cap S_j| + |B \cap S_j|$$
$$\leq 0.74\,|S_j| + |T_1| + |T_3| + |T_4| + 0.01\,|S_j| \leq 0.75\,|S_j| + 0.12n,$$

where we have also used the assumption that $|A \cap S_j| \leq 0.01\,|S_j|$, which in particular implies this upper bound for $|B \cap S_j|$, as $B \subseteq A$. Thus, as $|t| \leq \tau \leq 10^{-4}\,|S_j|$,

$$0.02\,|S_j| - \tau \leq |(T \cup B) \cap S_j| - (\lfloor |S_j|/2 \rfloor + t) \leq 0.251\,|S_j| + 0.12n\,.$$

Using $|S_j| \leq n/2$, $4p + n^{0.6} \geq n$ and $|S_j| \geq 10^4\tau$ we get that

$$0 < |(T \cup B) \cap S_j| - (\lfloor |S_j|/2 \rfloor + t) \leq 0.99p\,.$$

So, this quantity is also non-zero modulo p.

These three cases complete the proof of the lemma. □

We can now prove Theorem 3.5.

Proof of Theorem 3.5. We follow the outline discussed at the beginning of this section. Without loss of generality, we can assume that each set S_i has size at most $n/2$, else we work with the complement of S_i. Suppose, for the sake of contradiction, that $m \leq \frac{1}{10^5} \cdot n/\tau$. Let p be the largest prime such that $4p \leq n$. For large enough n, there is such a prime p such that $n - 4p \leq n^{0.6}$ (see [BHP01]).

Let $A \subseteq [n]$ be the set of size $n - 4p \leq n^{0.6}$ given by Lemma 3.6. Let B be an arbitrary subset of A of size $|A|/2$.

To every element $k \in [n] \setminus A$, we associate a formal variable x_k, and let $\mathbf{x} = \{x_k : k \in [n] \setminus A\}$ (note that $|\mathbf{x}| = 4p$). For each $j \in [m]$ such that $|S_j| > 2\tau$, define the following polynomials over $\mathbb{F}_p$:

$$B_j(\mathbf{x}) = \prod_{t=-\tau+1}^{\tau} \left(\sum_{k \in S_j \setminus A} x_k + |S_j \cap B| - \lfloor |S_j|/2 \rfloor - t \right).$$

If $|S_j| = 2\tau$, define a similar polynomial B_j where t ranges from $\tau + 1$ to $\tau - 1$. Further, let

$$f(\mathbf{x}) = \prod_{j=1}^{m} B_j(\mathbf{x})\,,$$

be a polynomial over $\mathbb{F}_p$. From the choice of the set A (see Lemma 3.6), we know that for every $j \in [m]$, B_j is a non-zero polynomial of degree smaller than 2τ.

There is a natural bijection between $[n] \setminus A$ and $[4p]$ (say, by ordering the elements of $[n] \setminus A$ by increasing order). Thus, we can naturally associate subsets Y' of $[4p]$ with subsets of $[n] \setminus A$, and indicator vector $\mathbb{1}_{Y'}$ with elements of $\{0,1\}^{\mathbf{x}}$.

We would like first to argue that f vanishes over all vectors of the form $\mathbb{1}_{Y'}$ for $Y' \in \binom{[4p]}{2p}$. Indeed, let Y' be such a set, and extend it to a balanced partition of $[n]$ by considering $Y = Y' \cup B$.

By the assumption, there is an index j such that $|Y \cap S_j| - \lfloor S_j \rfloor/2 \in \{-\tau + 1, \ldots, \tau\}$, and since $|Y \cap S_j| = |Y' \cap S_j| + |B \cap S_j|$, it follows that $B_j(\mathbb{1}_{Y'}) = 0$ and thus $f(\mathbb{1}_{Y'}) = 0$, as required.

Next, we want to show f does not vanish over a vector $\mathbb{1}_T$ for some $T \in \binom{[4p]}{3p}$.

Indeed, Lemma 3.7 precisely guarantees the existence of such a set $T \subseteq [n] \setminus A$, of size equal to $3p$, so that for all $j \in [m]$, $B_j(\mathbb{1}_T) \not\equiv 0 \bmod p$, and thus $f(\mathbb{1}_T) \neq 0$..

By Lemma 2.1, $\deg(f) \geq p$, and by construction, $\deg(f) \leq 2\tau \cdot m$, contradicting the assumed lower bound on m. □

Proofs of Claim 3.8, Claim 3.9 and Claim 3.10

We now prove the claims needed in the proof of Lemma 3.7. The arguments are based on standard concentration bounds.

Proof of Claim 3.8. The expected size of the set T_2 is equal to $0.65\,|[n] \setminus A|$. Using the fact that $|A| \leq n^{0.6}$ and by Lemma 2.6, we get that with probability at least $1 - \exp(-\Omega(n))$,

$$0.64n \leq |T_2| \leq 0.66n\,.$$

For the second item, observe that for any fixed $j \in [m]$, by Lemma 2.6, we have

$$\Pr\left[\left|\left|\tilde{S}_j \cap T_2\right| - 0.65\left|\tilde{S}_j\right|\right| \geq 0.09\left|\tilde{S}_j\right|\right] \leq 2\exp\left(-0.0162\left|\tilde{S}_j\right|\right)\,.$$

We know that $\tilde{S}_j \subseteq S_j$ and $\left|\tilde{S}_j\right| \geq 0.99\,|S|$. Thus,

$$\Pr\left[\left|\tilde{S}_j \cap T_2\right| \in [0.52\,|S_j|, 0.74\,|S_j|]\right] \geq 1 - 2\exp\left(-0.015\,|S_j|\right)\,.$$

For sets S_j of size at least $1000\log n$, this probability is high enough to take a union bound over all sets. So, we have the following.

$$\Pr\left[\forall j \in [m] \text{ such that } |S_j| \geq 1000\log n, \left|\tilde{S}_j \cap T_2\right| \in [0.52\,|S_j|, 0.74\,|S_j|]\right] \geq 1 - n^{-8}\,.$$

For the third and fourth items, observe that by construction, the set T_1 is a superset of all sets of size at most 6000τ and intersects every S_j on at least 6000τ elements. Moreover, since $\left|\tilde{S}_j\right| \geq 0.99\,|S_j|$, it follows that if $|S_j \cap (T_1 \cup T_2)| \leq 0.52\,|S_j|$, then sufficiently many elements will be included in the set T_3 so that $|S_j \cap (T_1 \cup T_2 \cup T_3)| \geq 0.52\,|S_j|$. □

Proof of Claim 3.9. For $j \in [m]$, we say that the set S is violated if $|S_j \cap (T_1 \cup T_2 \cup B)| \leq 0.52\,|S_j|$. Since T_1 intersects every set S_j on at least 6000τ elements, we know that any violated set S_j must satisfy $|S_j| \geq 10^4\tau$. So, from the proof of Claim 3.8, we get that the expected size of the set T_3 is given by

$$\mathbb{E}[|T_3|] \leq \sum_{j \in [m], |S_j| \geq 10^4\tau} \frac{2\,|S_i|}{\exp(0.015\,|S_i|)}\,.$$

From Claim 3.11 below, we know that this expectation can be upper bounded by

$$\mathbb{E}[|T_3|] \leq m \cdot \frac{2 \cdot \left|10^4\tau\right|}{\exp(0.015 \times 10^4\tau)}\,.$$

Since τ is at least 1 and $m \leq n/\tau$, we get

$$\mathbb{E}[|T_3|] \leq 10^{-10}n\,.$$

By Markov's inequality, we get the claim. □

Proof of Claim 3.10. This immediately follows from Claim 3.8. Observe that

$$|T_4| \leq 3p - |T_1| - |T_2| \ .$$

$|T_2| \geq 0.64n$ with probability at least $1 - n^{-5}$, and $|T_1| \geq 0.06n$ with probability 1. Thus, with probability at least $1 - n^{-5}$, $|T_4| \leq 0.05n$. □

Claim 3.11. *Let c be any positive constant. Then, for any $y \geq x \geq 1/c$, it holds that $x \cdot \mathrm{e}^{-cx} \geq y \cdot \mathrm{e}^{-cy}$.*

Proof. Let $f(x) = x \cdot \mathrm{e}^{-cx}$. The first derivative of $f(x)$ is

$$f'(x) = \mathrm{e}^{-cx} - cx\mathrm{e}^{-cx} \ .$$

It is easy to see that this is positive for $0 < x \leq 1/c$ and negative for $x > 1/c$. Therefore, $f(x)$, which vanishes at 0, increases as x increases from 0 to $1/c$, achieves its maximum at $x = 1/c$ and decreases thereafter. This implies the claim. □

4 Syntactically Multilinear Arithmetic Circuits

In this section, for the sake of completeness, we review the arguments of Raz, Shpilka and Yehudayoff [RSY08], and show how Theorem 3.1 implies a lower bound of $\Omega(n^2/\log^2 n)$. We mostly refer for [RSY08] for the proofs.

Specifically, we will show the following.

Theorem 4.1. *Let n be an even integer, and $X = \{x_1, \ldots, x_n\}$. Let $f(X) \in \mathbb{F}[X]$ be a multilinear polynomial such that for every balanced partition $X = Y \sqcup Z$, $\mathrm{rank}_{Y,Z}(f) = 2^{n/2}$. Let Ψ be a syntactically multilinear circuit computing f. Then $|\Psi| = \Omega(n^2/\log^2 n)$.*

The first step in proof of Theorem 4.1 is to show that if f is computed by a syntactically mutilinear circuit of size s, then there exists a syntactically multilinear circuit of size $O(s)$ that computes all the first-order partial derivatives of f, with the additional important property that for each i, the variable x_i does not appear in the subcircuit rooted at the output gate which computes $\partial f/\partial x_i$.

Theorem 4.2 ([RSY08], Theorem 3.1)**.** *Let Ψ be a syntactically multilinear circuit over a field $\mathbb{F}$ and the set of variables $X = \{x_1, \ldots, x_n\}$. Then, there exists a syntactically multilinear circuit Ψ', over $\mathbb{F}$ and X, such that:*

1. *Ψ' computes all n first-order partial derivatives $\partial f/\partial x_i$, $i \in [n]$.*
2. *$|\Psi'| \leq 5|\Psi|$.*
3. *Ψ' is syntactically multilinear.*
4. *For every $i \in [n]$, $x_i \notin X_{v_i}$, where v_i is the gate in Ψ' computing $\partial f/\partial x_i$.*

In particular, if v is a gate in Ψ', then it is connected by a directed path to at most $n - |X_v|$ output gates.

The proof of Theorem 4.2 appears in [RSY08], and mostly follows the classical proof of Baur and Strassen [BS83] of the analogous result for general circuits, with additional care in order to guarantee the last two properties.

Next we define two types of gates in a syntactically multilinear arithmetic circuits.

Definition 4.3. *Let Φ be a syntactically multilinear arithmetic circuit. Define $\mathcal{L}(\Phi, k)$, the set of lower-leveled gates in Φ, by*

$$\mathcal{L}(\Phi, k) = \{u : u \text{ is a gate in } \Phi,\ k < |X_u| < n - k, \text{ and } u \text{ has a parent } v \text{ with } |X_v| \geq n - k\}.$$

Define $\mathcal{U}(\Phi, k)$, the set of upper-leveled gates in Φ, by

$$\mathcal{U}(\Phi, k) = \{v : v \text{ is a gate in } \Phi,\ |X_v| \geq n - k, \text{ and } u \text{ has a child } v \in \mathcal{L}(\Phi, k)\}. \qquad \Diamond$$

The following lemma shows that if the set of lower-leveled gates is small, then there exists a partition $X = Y \sqcup Z$ under which the polynomial computed by the circuit is not of full rank.

Lemma 4.4. *Let Φ be a syntactically multilinear arithmetic circuit over $\mathbb{F}$ and $X = \{x_1, \ldots, x_n\}$, for an even integer n, computing f. Let $\tau = 3 \log n$ and $\mathcal{L} = \mathcal{L}(\Phi, 100\tau)$. If $|\mathcal{L}| < n/(10^5\tau)$, then there exists a partition $X = Y \sqcup Z$ such that $\operatorname{rank}_{Y,Z}(f) < 2^{n/2-1}$.*

We first sketch how Theorem 4.1 follows from Lemma 4.4. The proof is identical to the proof given in [RSY08] with slightly different parameters.

Proof of Theorem 4.1 assuming Lemma 4.4. Let Ψ' be the arithmetic circuit computing all n first-order partial derivatives of f, given by Theorem 4.2. Set $\tau = 3 \log n$ and let $\mathcal{L} = \mathcal{L}(\Psi', 100\tau)$ and $\mathcal{U} = \mathcal{U}(\Psi', 100\tau)$ as in Definition 4.3.

Denote $f_i = \partial f / \partial x_i$ and let v_i be the gate in Ψ' computing f_i, and Ψ'_i be the subcircuit of Ψ' rooted at v_i. Let $\mathcal{L}_i = \mathcal{L}(\Psi'_i, 100\tau)$. It is not hard to show (see [RSY08]) that $\mathcal{L}_i \subseteq \mathcal{L}$, and by Lemma 4.4 and item 4 in Proposition 2.7, it follows that $|\mathcal{L}_i| \geq n/(10^5\tau)$.

For every gate v in Ψ' define $C_v = \{i \in [n] : v \text{ is a gate in } \Psi_i\}$ to be the set of indices i such that there exists a directed path from v to the output gate computing f_i. For $i \in [n]$, let $\mathcal{U}_i = \{u \in \mathcal{U} : u \text{ is a gate in } \Psi'_i\}$, so that $\sum_{u \in \mathcal{U}} C_u = \sum_{i \in [n]} |\mathcal{U}_i|$.

Since the fan-in of each gate is at most two, $|\mathcal{L}_i| \leq 2|\mathcal{U}_i|$, and since every $u \in \mathcal{U}$ satisfies $|X_u| \geq n - 100\tau$, it follows by Theorem 4.2 that $|C_u| \leq 100\tau$. Thus, we get

$$n \cdot \frac{n}{10^5\tau} \leq \sum_{i \in [n]} |\mathcal{L}_i| \leq 2 \sum_{i \in [n]} |\mathcal{U}_i| = 2 \sum_{u \in U} C_u \leq 2|\mathcal{U}| \cdot 100\tau.$$

By item 2 in Theorem 4.2, and $\tau = 3 \log n$,

$$|\Psi| = \Omega(|\Psi'|) = \Omega(|\mathcal{U}|) = \Omega\left(\frac{n^2}{\log^2 n}\right). \qquad \square$$

It remains to prove Lemma 4.4. As the proof mostly appears in [RSY08], we only sketch the main steps.

Proof sketch of Lemma 4.4. Suppose $\mathcal{L} \leq n/(10^5\tau)$. By applying Theorem 3.5 to the family of sets $\{X_v : v \in \mathcal{L}\}$, it follows that there exists a balanced partition $Y \sqcup Z$ of X such that X_v is τ-unbalanced for every gate $v \in \mathcal{L}$ (one could get slightly improved constants in the case $n = 4p$ by applying Theorem 3.1).

The proof now proceeds in the exact same manner as the proof of Lemma 5.2 in [RSY08]. In Proposition 5.5 of [RSY08], it is shown that one can write

$$f = \sum_{i \in [\ell]} g_i h_i + g,$$

where $\mathcal{L} = \{v_1, \ldots, v_\ell\}$, h_i is the polynomial computed at v_i, and the set of variables appearing in g_i is disjoint from X_{v_i}.

In Claim 5.7 of [RSY08], it is shown that for every $i \in [\ell]$, $\mathrm{rank}_{Y,Z}(g_i h_i) \leq 2^{n/2-\tau}$. This uses the fact that X_{v_i} is τ-unbalanced, the upper bound in item 1 in Proposition 2.7, and item 3 in the same proposition.

In Proposition 5.8 of [RSY08], it is shown (with the necessary change of parameters) that the degree of g is at most 200τ.

Thus, by the fact that $\tau = 3 \log n$, item 5 and item 2 of Proposition 2.7, it follows that for large enough n,

$$\mathrm{rank}_{Y,Z}(f) \leq \ell \cdot 2^{n/2-\tau} + 2^{\tau^3} < 2^{n/2-1}. \qquad \square$$

4.1 An explicit full-rank polynomial

In this section, for the sake of completeness, we give a construction of a polynomial which is full-rank under any partition of the variables.

Construction 4.5 (Full rank polynomial, [RSY08])**.** *Let n be an even integer, and let $\mathcal{W} = \{\omega_1, \ldots, \omega_n\}$ and $X = \{x_1, \ldots, x_n\}$ be sets of variables. For a set $B \in \binom{[n]}{n/2}$, denote by $i_1 < \cdots < i_{n/2}$ the elements of B in increasing order, and by $j_1 < \cdots < j_{n/2}$ the elements of $[n] \setminus B$ in increasing order. Define $r_B = \prod_{\ell \in B} \omega_\ell$, and $g_B = \prod_{\ell \in [n/2]} (x_{i_\ell} + x_{j_\ell})$.*

Finally, define

$$f = \sum_{B \in \binom{[n]}{n/2}} r_B g_B. \qquad \diamondsuit$$

Claim 4.6 ([RSY08])**.** *For f from Construction 4.5, it holds that for every balanced partition of $X = Y \sqcup Z$, $\mathrm{rank}_{Y,Z}(f) = 2^{n/2}$, where the rank is taken over $\mathbb{F}(\mathcal{W})$.*

We give a proof which is shorter and simpler than the one given in [RSY08].

Proof of Claim 4.6. Fix a balanced partition $X = Y \sqcup Z$, and consider the matrix $M_{Y,Z}(f)$ where f is interpreted as a polynomial in $f \in (\mathcal{F}[\mathcal{W}])[X]$ (that is, the rows and columns of the matrix are indexed by X variables and its entries are polynomials in $\mathcal{W}$). We want to show that $\det(M_{Y,Z}(f)) \in \mathbb{F}[\mathcal{W}]$ is a non-zero polynomial. Fix $\omega_i = 1$ if $i \in Y$ and $\omega_i = 0$ otherwise. Under this restriction, $f = g_Y$. It is also not hard to see that $\det(M_{Y,Z}(g_Y)) \neq 0$, since this is a permutation matrix (this also follows from item 3 of Proposition 2.7). Thus, $\det(M_{Y,Z}(f))$ evaluates to a non-zero value under this setting of the variables $\mathcal{W}$, which implies it a non-zero polynomial. $\square$

Corollary 4.7. *Every syntactically multilinear circuit computing f has size at least $\Omega(n^2/\log^2 n)$.*

The polynomial f in Construction 4.5 is in the class VNP of explicit polynomials, but it is not known whether there exists a polynomial size multilinear circuit for f.

Raz and Yehudayoff [RY08] constructed a full-rank polynomial $g \in \mathbb{F}[X, \mathcal{W}']$ that has a syntactically multilinear circuit of size $O(n^3)$. Their construction also uses a set of auxiliary variables $\mathcal{W}'$ of size $O(n^3)$. Thus, if one measures the complexity as a function of $|X| \cup |\mathcal{W}'|$, the quadratic lower bound of Theorem 4.1 is meaningless, because a lower bound of $\Omega(n^3)$ holds trivially. However, we believe that since the rank is taken over $\mathbb{F}(\mathcal{W}')$, it is only fair to consider computations over $\mathbb{F}(\mathcal{W}')$, where any rational expression in the variables of $\mathcal{W}'$ is merely a field constant. Thus, in this setting, an input gate can be labeled by an arbitrarily complex rational function in the variables of $\mathcal{W}'$, and the complexity is measured as a function of $|X|$ alone. In this model the lower bound of Theorem 4.1 *is* meaningful, and furthermore, this example shows that the partial derivative matrix technique cannot prove an $\omega(n^3)$ lower bound.

Acknowledgments

Part of this work was done while the second author was visiting Tel Aviv University. We thank Amir Shpilka for the visit, for many insightful discussions, and for comments on an earlier version of this text. We also thank Srikanth Srivinasan for allowing us to include his proof of Lemma 2.1 in this paper, and to Andy Drucker for pointing out a correction in a previous version of this paper.

References

[ABCO88] Noga Alon, Ernest E. Bergmann, Don Coppersmith, and Andrew M. Odlyzko. Balancing sets of vectors. *IEEE Trans. Information Theory*, 34(1):128–130, 1988.

[ARS02] Richard P. Anstee, Lajos Rónyai, and Attila Sali. Shattering News. *Graphs and Combinatorics*, 18(1):59–73, 2002.

[AS16] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley Publishing, 4th edition, 2016.

[Ber84] Stuart J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Information Processing Letters*, 18(3):147 – 150, 1984.

[BHP01] R. C. Baker, G. Harman, and J. Pintz. The Difference Between Consecutive Primes, II. *Proceedings of the London Mathematical Society*, 83(3):532–562, 2001.

[BS83] Walter Baur and Volker Strassen. The Complexity of Partial Derivatives. *Theoretical Computer Science*, 22:317–330, 1983.

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial Derivatives in Arithmetic Complexity (and beyond). *Foundation and Trends in Theoretical Computer Science*, 2011.

[Csa76] László Csanky. Fast Parallel Matrix Inversion Algorithms. *SIAM J. Comput.*, 5(4):618–623, 1976.

[EFIN87] Hikoe Enomoto, Peter Frankl, Noboru Ito, and Kazumasa Nomura. Codes with given distances. *Graphs Combin.*, 3(1):25–38, 1987.

[FG06] Jeffrey B. Farr and Shuhong Gao. Computing Gröbner Bases for Vanishing Ideals of Finite Sets of Points. In *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 16th International Symposium, AAECC-16, Las Vegas, NV, USA, February 20-24, 2006, Proceedings*, pages 118–127, 2006.

[FLMS14] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 128–135, 2014. Pre-print available at `eccc:TR13-100`.

[FR87] Peter Frankl and Vojtěch Rödl. Forbidden intersections. *Trans. Amer. Math. Soc.*, 300(1):259–286, 1987.

[GK98] Dima Grigoriev and Marek Karpinski. An Exponential Lower Bound for Depth 3 Arithmetic Circuits. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC 1998)*, pages 577–582, 1998.

[GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the Chasm at Depth Four. *Journal of the ACM*, 61(6):33:1–33:16, 2014. Preliminary version in the *28th Annual IEEE Conference on Computational Complexity (CCC 2013)*. Pre-print available at eccc:TR12-098.

[GR00] Dima Grigoriev and Alexander A. Razborov. Exponential Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields. *Appl. Algebra Eng. Commun. Comput.*, 10(6):465–487, 2000. Preliminary version in the *39th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1998)*.

[Heg10] Gábor Hegedűs. Balancing sets of vectors. *Studia Sci. Math. Hungar.*, 47(3):333–349, 2010.

[HR03] Gábor Hegedűs and Lajos Rónyai. Gröbner bases for complete uniform families. *J. Algebraic Combin.*, 17(2):171–180, 2003.

[Jan08] Maurice J. Jansen. Lower Bounds for Syntactically Multilinear Algebraic Branching Programs. In *Proceedings of the 33rd Internationl Symposium on the Mathematical Foundations of Computer Science (MFCS 2008)*, volume 5162 of *Lecture Notes in Computer Science*, pages 407–418. Springer, 2008.

[Kal85] Kyriakos Kalorkoti. A Lower Bound for the Formula Size of Rational Functions. *SIAM J. Comput.*, 14(3):678–687, 1985.

[KLSS14] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Circuits. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, pages 61–70, 2014. Pre-print available at eccc:TR14-005.

[Knu86] Donald E. Knuth. Efficient balanced codes. *IEEE Trans. Information Theory*, 32(1):51–53, 1986.

[KS14] Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2014)*, pages 364–373, 2014. Pre-print available at eccc:TR14-045.

[KS17] Mrinal Kumar and Ramprasad Saptharishi. An Exponential Lower Bound for Homogeneous Depth-5 Circuits over Finite Fields. In *Proceedings of the 32nd Annual Computational Complexity Conference (CCC 2017)*, volume 79, pages 31:1–31:30, 2017. Pre-print available at eccc:TR15-109.

[Kum17] Mrinal Kumar. A Quadratic Lower Bound for Homogeneous Algebraic Branching Programs. In *Proceedings of the 32nd Annual Computational Complexity Conference (CCC 2017)*, volume 79, pages 19:1–19:16, 2017. Pre-print available at eccc:TR17-028.

[MV97] Meena Mahajan and V. Vinay. Determinant: Combinatorics, Algorithms, and Complexity. *Chicago J. Theor. Comput. Sci.*, 1997, 1997. Preliminary version in the *8th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 1997)*.

[Nis91] Noam Nisan. Lower bounds for non-commutative computation. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC 1991)*, pages 410–418, 1991. Available on citeseer:10.1.1.17.5067.

[NW97] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1997. Available on citeseer:10.1.1.90.2644.

[Raz06] Ran Raz. Separation of Multilinear Circuit and Formula Size. *Theory of Computing*, 2(1):121–135, 2006. Preliminary version in the *45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*. Pre-print available at eccc:TR04-042.

[Raz09] Ran Raz. Multi-Linear Formulas for Permanent and Determinant are of Super-Polynomial Size. *J. ACM*, 56(2):8:1–8:17, 2009. Preliminary version in the *36th Annual ACM Symposium on Theory of Computing (STOC 2004)*. Pre-print available at eccc:TR03-067.

[Raz10] Ran Raz. Elusive Functions and Lower Bounds for Arithmetic Circuits. *Theory of Computing*, 6(1):135–177, 2010.

[RSY08] Ran Raz, Amir Shpilka, and Amir Yehudayoff. A lower bound for the size of syntactically multilinear arithmetic circuits. *SIAM J. Comput.*, 38(4):1624–1647, 2008. Preliminary version in the *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007)*. Pre-print available at eccc:TR06-060.

[RY08] Ran Raz and Amir Yehudayoff. Balancing Syntactically Multilinear Arithmetic Circuits. *Computational Complexity*, 17(4):515–535, 2008.

[RY09] Ran Raz and Amir Yehudayoff. Lower Bounds and Separations for Constant Depth Multilinear Circuits. *Computational Complexity*, 18(2):171–207, 2009. Preliminary version in the *23rd Annual IEEE Conference on Computational Complexity (CCC 2008)*. Pre-print available at eccc:TR08-006.

[Sap16] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Github survey, https://github.com/dasarpmar/lowerbounds-survey/, 2016.

[Ska13] Matthew Skala. Hypergeometric tail inequalities: ending the insanity. *arXiv preprint arXiv:1311.5939*, 2013.

[Str73] Volker Strassen. Die Berechnungskomplexität Von Elementarsymmetrischen Funktionen Und Von Interpolationskoeffizienten. *Numerische Mathematik*, 20(3):238–251, June 1973.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic Circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5:207–388, March 2010.