Electronic Colloquium on Computational Complexity, Report No. 181 (2015)

On the size of homogeneous and of depth four formulas with low individual degree

Neeraj Kayal
Microsoft Research India
neeraka@microsoft.com

Chandan Saha
Indian Institute of Science
chandan@csa.iisc.ernet.in

Sébastien Tavenas
Microsoft Research India
sebastien.tavenas@free.fr

Abstract

Let $r \geq 1$ be an integer. Let us call a polynomial $f(x_1, x_2, \ldots, x_N) \in \mathbb{F}[\mathbf{x}]$ as a multi-r-ic polynomial if the degree of f with respect to any variable is at most r (this generalizes the notion of multilinear polynomials). We investigate arithmetic circuits in which the output is syntactically forced to be a multi-r-ic polynomial and refer to these as multi-r-ic circuits. We prove lower bounds for several subclasses of such circuits.

Specifically, first define the formal degree of a node α with respect to a variable x_i inductively as follows. For a leaf α it is 1 if α is labelled with x_i and zero otherwise; for an internal node α labelled with $\times$ (respectively $+$) it is the sum of (respectively the maximum of) the formal degrees of the children with respect to x_i. We call an arithmetic circuit as a multi-r-ic circuit if the formal degree of the output node with respect to any variable is at most r. We prove lower bounds for various subclasses of multi-r-ic circuits, including:

1. An $N^{\Omega(\log N)}$ lower bound against *homogeneous* multi-r-ic formulas (for an explicit multi-r-ic polynomial on N variables).

2. A $\left(\frac{n}{r^{1.1}}\right)^{\Omega\left(\sqrt{\frac{d}{r}}\right)}$ lower bound against *depth four* multi-r-ic circuits computing the polynomial $\mathrm{IMM}_{n,d}$ corresponding to the product of d matrices of size $n \times n$ each.

3. A $2^{\Omega\left(\sqrt{N}\right)}$ lower bound against depth four multi-r-ic circuits computing an explicit multi-r-ic polynomial on N variables.

ISSN 1433-8092

1 Introduction

Arithmetic Models of computation. An arithmetic circuit computes a polynomial function over some underlying field $\mathbb{F}$ via a sequence of operations involving $+$ and $\times$ starting from its inputs $x_1, x_2, \ldots, x_N$. We typically allow arbitrary constants from $\mathbb{F}$ on the incoming edges to a $+$ gate so that a $+$ gate can in fact compute an arbitrary $\mathbb{F}$-linear combination of its inputs. The complexity of a circuit is measured in terms of its size[1] and depth[2]. Being the most natural and intuitive way to compute polynomials, arithmetic circuits have been widely investigated[3]. A central open problem in this area is to prove arithmetic circuit lower bounds (for some explicit family of polynomials). Progress on it has been made in the form of lower bounds for some subclasses of circuits, one of the most significant ones being Raz's lower bound for multilinear[4] formulas[5] [Raz09]. We study formulas that are a natural generalization of the class of multilinear formulas. We now give some more motivation before giving a precise definition of the relevant circuit subclasses studied and the lower bounds obtained here.

Background. Motivated by the question of whether computation can be efficiently parallelized, one thread of work in this area [Hya79, VSBR83, AV08, Raz10, Koi12, Tav13, GKKS13a] gives the loss in size incurred in transforming a general circuit or formula into one of low-depth (sometimes with additional structural restrictions on the resulting low-depth circuits). In particular, these results say that proving sufficiently strong lower bounds for (subclasses of) low-depth circuits implies lower bounds for arbitrary circuits as well. Low-depth circuits being easier to analyze, this might be a potential pathway to general lower bounds. Somewhat promisingly, a lot of new lower bounds have recently been proved for various subclasses of arithmetic circuits, particularly for low-depth subclasses [Kay12, GKKS13b, KSS14, FLMS14, CM14, KS14a, KLSS14, KS14b, KS15a, KS15d, KS15c]. However, most of the present lower bounds can only handle subclasses of circuits having formal degree[6] which is rather low (equal to or sometimes slightly larger than the number of variables). To make

[1] The size of a circuit is the number of edges in the circuit. This corresponds to the number of binary operations in the computation.

[2] The depth of a circuit is the maximum length of a path from an input to the output node. This corresponds to the amount of parallelism afforded by the computation. The product-depth will correspond to the maximum number of product gates on a path from an input to the output.

[3] See for example the book by Bürgisser, Clausen and Shokrollahi [BCS97] or the more recent survey by Shpilka and Yehudayoff [SY10a] for an overview of the problems and results in this area.

[4] A polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is said to be multilinear if its degree with respect to any variable is at most 1. A circuit is said to be (syntactically) multilinear if the polynomial computed at every node is syntactically forced to be multilinear - for details see Definition 1 for the special case of $r = 1$.

[5] Recall that a formula is a circuit in which the underlying graph is in fact a tree. It is more convenient to work with the number of leaves of the tree as the size of the formula.

[6] The *formal degree* of a node α, denoted $\deg(\alpha)$, in a circuit is defined inductively as follows. For a leaf α, it is one if α is labelled with a variable and zero otherwise. For a $\times$ gate (respectively a $+$ gate), it is the sum of (respectively the maximum of) the formal degree of the children. The formal degree of a circuit is the formal degree of its output node. It represents what the degree of the output would have been if there were no cancellations and is always an upper bound on the degree of the output.

progress and remove the limitations of multilinearity and low formal degree, some recent work [KS15b, dO15] looks at a model that generalizes multilinear circuits/formulas. Such a generalization also appears in the hardness versus randomness trade-off result for bounded depth circuits [DSY09].

The multi-r-ic circuit model. Intuitively, a multi-r-ic circuit is an arithmetic circuit in which the output polynomial is syntactically constrained to have degree at most r with respect to any individual variable. We now make this precise.

Definition 1. *Define the formal degree of a node α with respect to a variable x_i in an arithmetic circuit inductively as follows. For a leaf α it is 1 if α is labelled with x_i and zero otherwise; for an internal node α labelled with $\times$ (respectively $+$) it is the sum of (respectively the maximum of) the formal degrees of the children with respect to x_i. We call an arithmetic circuit as a multi-r-ic circuit if the formal degree of the output node with respect to any variable is at most r.*

Note that the formal (total) degree of a N-variate multi-r-ic circuit can be $(r \cdot N)$ which is asymptotically larger than N when $r = \omega(1)$. In this work, we prove lower bounds for several subclasses of multi-r-ic circuits. Now, once one has a lower bound for some explicit polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ against a circuit subclass $\mathcal{C}$, it is desirable to try to make such an f come from as small a class $\mathcal{D}$ as possible[7]. Following this general theme, we have tried to minimize the complexity of our *target polynomial* f.

Our results. Our lower bounds hold over any field unless mentioned otherwise explicitly. Our first result is a superpolynomial lower bound for *homogeneous*[8] *multi-r-ic formulas.*

Theorem 1. Homogeneous multi-r-ic formulas. *Let $r = r(N)$ be any integer. There exists an explicit family of N-variate multi-r-ic polynomials $P_{N,r}$ such that any homogeneous multi-r-ic formula computing $P_{N,r}$ must have size at least $2^{\Omega(\log^2(N))}$. Moreover, if the field $\mathbb{F}$ contains r distinct r-th roots of unity and has size at least $(2Nr)$ then $P_{N,r}$ can be computed by a* $\mathsf{poly}(Nr)$*-sized (nonhomogeneous) depth three multi-r-ic formula.*

We probe a bit further in this direction and obtain improved lower bounds for homogeneous multi-r-ic formulas of low depth.

Theorem 2. Constant depth homogeneous multi-r-ic formulas. *Let $r = r(N)$ be an integer and let p be an integer. If $p = o\left(\frac{\log N}{\log r + \log\log N}\right)$ then there exists an explicit family of N-variate multi-r-ic polynomials $P_{N,r}$ such that any homogeneous multi-r-ic formula of product-depth p computing $P_{N,r}$ must have size at least $2^{\Omega\left(\frac{1}{r}\left(\frac{N}{4}\right)^{1/p}\right)}$. Moreover, if the field $\mathbb{F}$*

[7] This gives a separation between the classes $\mathcal{C}$ and $\mathcal{D}$ and enhances our understanding of the cost (in terms of size) that must be incurred in order to transform a circuit in $\mathcal{D}$ to an equivalent one in $\mathcal{C}$. It also enhances our understanding of the the proof techniques involved.

[8] Recall that a polynomial is called homogeneous if all its monomials have the same total degree. A circuit is called homogeneous if all its internal nodes compute a homogeneous polynomial.

contains r distinct r-th roots of unity and has size at least $(2Nr)$ then $P_{N,r}$ can be computed by a $\mathsf{poly}(Nr)$*-sized (nonhomogeneous) depth three multi-r-ic formula.*

The proofs of these two results follow the strategy of first facilitating the analysis by reducing the depth[9] and then proving lower bounds against the resulting low-depth formula. As one hopes to be able to implement some such strategy to obtain lower bounds against more powerful subclasses of circuits (maybe even general arithmetic circuits), it makes sense to prove lower bounds against low-depth multi-r-ic circuits (without the homogeneity restriction). Also, the "moreover" part of the above theorems shows that nonhomogeneous depth three multi-r-ic formulas are superpolynomially more powerful than homogeneous multi-r-ic formulas of arbitrary depth. This further motivates our next few results on lower bounds for low-depth multi-r-ic formulas without the homogeneity restriction. We represent a circuit of depth p with a sequence of p alternating symbols (either Σ or Π) wherein the leftmost symbol denotes the nature of the output gate. So for example, a $\Sigma\Pi\Sigma\Pi$ circuit is a depth four circuit where the output gate is an addition gate. We denote by $\mathrm{IMM}_{n,d}$ the $(1,1)$-th entry of the product of d matrices of size $n \times n$ each.

Theorem 3. Depth four multi-r-ic formulas computing IMM**.** *For any integer $r = r(n)$ such that $r^{1.1} \ll n$ and for any $d \gg r$, any multi-r-ic $\Sigma\Pi\Sigma\Pi$ circuit computing $\mathrm{IMM}_{n,d}$ where $d \geq \log^2 n$ must have size at least*[10] $\left(\frac{n}{r^{1.1}}\right)^{\Omega\left(\sqrt{\frac{d}{r}}\right)}$.

Moreover, one can notice that the reduction [Val79] of $\mathrm{IMM}_{n,d}$ as a projection of $\mathrm{Det}_{n\cdot d}$ maintains the multi-r-icity.

Corollary 4. Depth four multi-r-ic formulas computing determinant. *For any integer $r = r(n)$, any multi-r-ic $\Sigma\Pi\Sigma\Pi$ circuit computing Det_n must have size at least $2^{\Omega\left(\frac{\sqrt{n}}{r^{1.1}}\right)}$.*

Note that the target polynomials, namely $\mathrm{IMM}_{n,d}$ and Det_n, in the above theorems are multilinear polynomials. If we allow our target polynomial to be a multi-r-ic polynomial then we can obtain a lower bound that does not degrade at all as r increases.

Theorem 5. Depth four multi-r-ic formulas computing a multi-r-ic polynomial. *For any positive integer $r = r(N)$ there exists an explicit family $\{G_{N,r}\}$ of multi-r-ic N-variate polynomials such that any multi-r-ic $\Sigma\Pi\Sigma\Pi$-circuit computing $G_{N,r}$ must have size at least $2^{\Omega\left(\sqrt{N}\right)}$. Moreover, one can even choose such a family $G_{N,r}$ so that it can in fact be computed by a* $\mathsf{poly}(Nr)$*-sized multi-r-ic algebraic branching program.*

The three previous bounds hold for $\Sigma\Pi\Sigma\Pi$-circuits. Moreover, we can improve them in the case of $\Sigma\Pi\Sigma$-circuits.

Theorem 6. Depth three multi-r-ic formulas computing IMM**.** *For any integer $r = r(n)$ such that $r \ll n$, any multi-r-ic $\Sigma\Pi\Sigma$-circuit computing $\mathrm{IMM}_{n,d}$ has size at least $\left(\frac{n}{r}\right)^{\Omega(d)}$.*

[9] In this case, the depth reduction yields some sort of a depth four formula with a specific, well-chosen structure - see Lemma 11 for the details.

[10] Through all the paper, the exponents 1.1 and 1.2 can in fact be chosen as close as we want to 1.

Corollary 7. Depth three multi-r-ic formulas computing the determinant. *For any integer $r = r(n)$, any multi-r-ic $\Sigma\Pi\Sigma$-circuit computing Det_n has size at least $2^{\Omega(\frac{n}{r})}$.*

Theorem 8. Depth three multi-r-ic formulas computing a multi-r-ic polynomial. *For any positive integer $r = r(N)$ there exists an explicit family $\{G_{N,r}\}$ of multi-r-ic N-variate polynomials such that any multi-r-ic $\Sigma\Pi\Sigma$-circuit computing $G_{N,r}$ must have size at least $2^{\Omega(N)}$. Moreover, one can choose such a family $G_{N,r}$ so that it can in fact be computed by a $\mathsf{poly}(Nr)$-sized $\Pi\Sigma\Pi$-circuit.*

One can notice that for any constant r the lower bound $2^{\Omega(N)}$ (respectively $2^{\Omega(\sqrt{N})}$) in Theorem 8 (respectively Theorem 5) is optimal since there is a depth three circuit (respectively depth four circuit) of size $2^{O(N)}$ (respectively $2^{O(\sqrt{N})}$) computing the target polynomial.
Comparison to previous results. As mentioned earlier, most of the prior work failed to yield (superpolynomial) lower bounds when the degree of the polynomial being computed and/or the formal degree of the circuit was significantly larger than the number of variables. In this particular aspect, the above results represent significant progress. For example Theorems 1 and 5 yield (superpolynomial) lower bounds against natural subclasses of circuits wherein the formal degree is allowed to be arbitrarily larger than the number of variables. These results also yield improved lower bounds for some previously studied subclasses.

1. **Multilinear $\Sigma\Pi\Sigma\Pi$ Circuits.** While the focus of this work is on multi-r-ic formulas for $r > 1$ for which lower bounds were previously not known, our results have interesting consequences for the much more well-studied special case of $r = 1$ corresponding to multilinear formulas. Previously, the best known[11] lower bound against multilinear-$\Sigma\Pi\Sigma\Pi$ circuit computing any explicit N-variate polynomial of degree d was $2^{\Omega(\sqrt{d \cdot \log d})}$, where $d \ll N$. Note that this does not increase with N[12]. The special case of Theorem 3 for $r = 1$ yields a $n^{\Omega(\sqrt{d})} = \left(\frac{N}{d}\right)^{\Omega(\sqrt{d})}$ lower bound for multilinear-$\Sigma\Pi\Sigma\Pi$ circuits computing $\mathrm{IMM}_{n,d}$.

2. **Multi-r-ic $\Sigma\Pi\Sigma$-circuits.** Multi-r-ic depth three circuits were recently studied in [KS15a] and a lower bound of $2^{\Omega\left(\frac{N}{2^r}\right)}$ was obtained for an explicit multi-r-ic N-variate polynomial. In particular, no superpolynomial lower bound seems to have been known when $r = \omega(\log N)$. In comparison, Theorem 8 gives an exponential lower bound which is independent of r.

[11] Actually a lot of the work on multilinear formulas deal with polynomials such as the determinant and the permanent where the number of variables N is a fixed function of d, e.g. $N = d^2$ in the case of the permanent and the determinant. Therefore the statements of the results themselves do not reveal the structure of the lower bound as a function of both N and d. It seems that the proof technique employed in Raz [Raz09] and Raz-Yehudayoff [RY09] only yields a lower bound that is independent of N for when N is much larger than d (in a multilinear polynomial N cannot be smaller than d), a key initial step in their argument involving random restrictions *kills off* the extra variables so that the number of surviving variables is comparable to the degree and then works with this restricted polynomial.

[12] The work of [FLMS14] looks at multilinear-$\Sigma\Pi\Sigma\Pi$-circuits *with the additional restriction of homogeneity* and obtains a $n^{\Omega(\sqrt{d})}$ lower bound for $\mathrm{IMM}_{n,d}$.

2 Proof Overview

In this section we give an overview of the proof of some of these lower bounds with an emphasis on those ingredients of the proof that are new here as compared to prior work in the area.

Homogeneous multi-r-ic formulas. Our proof is a generalization (from $r = 1$ to arbitrary values of r) of the work of [HY11] and follows the same overall proof strategy of first doing a depth-reduction[13] in order to make the resulting expression easier to analyze and then proving lower bounds on the size of such expressions. Roughly, if f is any N-variate polynomial computed by a multi-r-ic formula Φ of size s then f can be written as

$$f = T_1 + T_2 + \ldots + T_s, \tag{1}$$

where each T_i is a multi-r-ic polynomial that can be expressed as a product of $(\log N)$-many homogeneous polynomials

$$T_i = Q_{i1} \cdot Q_{i2} \cdot \ldots \cdot Q_{i \log N} \quad \text{where each } Q_{ij} \text{ has at least } \sqrt{N}\text{-many } \textit{fresh variables} \tag{2}$$

(see Lemma 11 and the preceding discussion for the precise definitions and statements[14]). Moreover, if the formula Φ is also homogeneous then each Q_{ij} is homogeneous as well. We then carefully choose a subset of multi-r-ic monomials[15] which we refer to as extremal monomials (see Definition 5 for the precise statement) and employ a result from extremal combinatorics called *Sperner's theorem* to upper bound the number of extremal monomials in a homogeneous multi-r-ic term T of the form given by Equation (2). We then choose our target polynomial f to have the maximum possible number of extremal monomials (and also to be easily computed by a nonhomogeneous multi-r-ic depth three circuit). Finally looking at the *ratio* of the number of extremal monomials in f to that in T yields the stated lower bound.

Depth Four Circuits. The proof for multi-r-ic depth four circuits builds on some of the recent work [KLSS14, KS14b] on homogeneous depth four circuits and shares many ingredients with these. Let f be a polynomial computed by a small multi-r-ic depth four circuit. We first employ random restrictions to reduce the support size[16] of the monomials

[13] A similar depth reduction is also given in the exposition of Raz's proof in [SY10b].

[14] For comparison, we mention that in the special case of $r = 1$, the Q_{ij}'s can be ensured to have disjoint sets of variables. It seems unlikely that such a decomposition with Q_{ij}'s being variable disjoint can be obtained for arbitrary r.

[15] For comparison, we mention that [HY11] observe that *the ratio* of the maximum possible number of monomials in a homogeneous multilinear polynomial (a N-variate homogeneous multilinear polynomial contains at most $\binom{N}{\frac{N}{2}}$-many monomials) to the number of monomials in a term T of the form given by Equation (2) is superpolynomial and this essentially yields the lower bound. It seems quite implausible that this ratio of naive monomial counts will yield meaningful lower bounds when r is large.

[16] The support size of a monomial is the number of distinct variables appearing in it, i.e. if $m = x_1^{e_1} \cdot x_2^{e_2} \cdot \ldots \cdot x_N^{e_N}$ is a monomial then the support-size of m, denoted $|\mathrm{Supp}(m)|$ is the size of the set $\{i \;:\; e_i \geq 1\} \subseteq [N]$.

appearing in our depth four circuit. We then get a representation of the following form:

$$f(\mathbf{x}) = T_1 + T_2 + \ldots + T_s,$$

where each term T_i is a multi-r-ic polynomial of the form $T_i = Q_{i1} \cdot Q_{i2} \cdot \ldots \cdot Q_{iD}$, every monomial in each Q_{ij} has a relatively small number of variables. Since each T_i is multi-r-ic, the number of factors D in it can at most be $(N \cdot r)$ and in general this upper bound is tight. Now such representations did occur at the intermediate stages in some recent pieces of work [KLSS14, KS14b] and a complexity measure called *dimension of projected shifted partials* was devised to analyze these. It involves looking at all the low-order partial derivatives of f, multiplying these with monomials of a suitable degree, applying a carefully chosen linear operator $\pi : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ and looking at the dimension of the resulting set of polynomials. It turns out that when the number of factors D significantly exceeds the number of variables N the bounds on the dimension of shifted partials that were proved in earlier works do not seem to yield any nontrivial lower bound overall. The key observation here is that one can get around this difficulty using a complexity measure that is some sort of a hybrid of the shifted partials measure with what is used in the work of [Raz09, RY09]. Specifically, we partition our set of variables $\mathbf{x}$ into two sets $\mathbf{x} = \mathbf{y} \uplus \mathbf{z}$ where *the size of* $\mathbf{y}$ *is significantly larger compared to the size of* $\mathbf{z}$[17]. We observe that if instead of taking all (low order) derivatives, we take (low order) derivatives with respect to only the $\mathbf{y}$-variables *and subsequently set them to zero* then effectively the number of factors becomes more like $|\mathbf{z}| \cdot r$ which is much smaller than $(|\mathbf{x}| \cdot r) = (N \cdot r)$ while still giving us a large enough space of partial derivatives to work with. In order to highlight this idea and illustrate its power in a simpler situation, in Theorem 28 we first show how this can be used to obtain a $\left(\frac{N}{r^3}\right)^{\Omega(d)}$ lower bound against *multi-r-ic depth three* circuits computing an explicit *multilinear* polynomial of degree d on N variables. We then show how some of the other ingredients from some recent lower bound proofs [Kay12, GKKS13b], combined with some judicious bounds on the dimension of some relevant sets of polynomials (Lemma 30) can be used to obtain a $2^{\Omega(\sqrt{N})}$ lower bound for multi-r-ic depth four circuits computing an explicit polynomial (Theorem 5). This lower bound degrades with r in case the target polynomial is multilinear (Theorem 3).

3 Preliminaries

3.1 Notation.

$[n]$ shall denote the set of first n positive integers, i.e. $[n] = \{1, 2, \ldots, n\}$. For any finite set A and any integer $k \leq |A|$, $\binom{A}{k}$ shall denote the set of all subsets of A of size exactly k while 2^A will denote the set of all subsets of A. We will often use boldfaced letters for tuples of variables or numbers. For example $\mathbf{x}$ will usually denote an N-tuple of variables

[17] For comparison, [Raz09, RY09] also partition the variables into two sets but it is crucial to their argument that the two sets have nearly the same size and that the partition is chosen randomly. In contrast, we choose the partition deterministically and it is crucial to our argument that the two parts have rather unequal sizes.

$(x_1, x_2, \ldots, x_N)$ while $\mathbf{r} = (r_1, r_2, \ldots, r_N)$ will usually denote an N-tuple of non-negative integers (the dimensionality N will usually be clear from context).

Some numerical estimates. We will employ the following well-known numerical estimates:

Proposition 9. *1.* **Binomial Estimates.**

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k.$$

2. **Exponential Estimates.**

$$\begin{aligned} e^x &\geq 1 + x \quad \textit{for all } x \in \mathbb{R}, \quad \textit{and} \\ e^x &\leq 1 + 2x \quad \textit{for } \ x \in [0, 1]. \end{aligned}$$

Let $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ be an N-tuple of formal variables. We work with the ring of polynomials $\mathbb{F}[\mathbf{x}]$ over some underlying field $\mathbb{F}$. We will use the following notation related to (sets) of polynomials and maps between them.

Sets of Polynomials. For a subset of variables $\mathbf{z} \subset \mathbf{x}$, we will denote by $\mathbf{z}^{=\ell}$ the set of all monomials in the $\mathbf{z}$ variables of degree exactly ℓ. Thus if $\mathbf{z} = (z_1, z_2, \ldots, z_m)$ then

$$\mathbf{z}^{=\ell} \stackrel{\text{def}}{=} \left\{ z_1^{i_1} \cdot z_2^{i_2} \cdot \ldots z_m^{i_m} \ : \ (i_1, i_2, \ldots, i_m) \in \mathbb{Z}^m_{\geq 0}, \quad i_1 + i_2 + \ldots + i_m = \ell \right\}.$$

Similarly, $\mathbf{z}^{\leq \ell}$ shall denote the set of all monomials over the $\mathbf{z}$-variables of degree at most ℓ. Let $\mathbf{y} \subseteq \mathbf{x}$ be a subset of variables of $\mathbf{x}$. $\boldsymbol{\partial}_{\mathbf{y}}^{=k} f$ shall denote the set of all k-th order partial derivatives of f with respect to the $\mathbf{y}$ variables, i.e.

$$\boldsymbol{\partial}_{\mathbf{y}}^{=k} f \stackrel{\text{def}}{=} \left\{ \frac{\partial^k f}{\partial y_{i_1} \cdot \partial y_{i_2} \cdot \ldots \cdot \partial y_{i_k}} : i_1, i_2, \ldots i_k \in [|\mathbf{y}|] \right\}.$$

For two sets of polynomials $A, B \subseteq \mathbb{F}[\mathbf{x}]$, the set $A \cdot B$ shall be the set of pairwise products, i.e.

$$A \cdot B \stackrel{\text{def}}{=} \{ f(\mathbf{x}) \cdot g(\mathbf{x}) \ : \ f(\mathbf{x}) \in A, \ g(\mathbf{x}) \in B \}.$$

For a set of polynomials $A \subseteq \mathbb{F}[\mathbf{x}]$, the dimension of A will denote the dimension of the $\mathbb{F}$-vectorial space generated by A.

Degree with respect to a subset of variables. Let $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial and $\mathbf{z} \subseteq \mathbf{x}$ be a subset of variables. The $\mathbf{z}$-degree of f, denoted $\deg_{\mathbf{z}}(f)$, is the degree of f when viewed as a polynomial over the $\mathbf{z}$-variables with coefficients from the function field $\mathbb{F}(\mathbf{y})$, where $\mathbf{y} = \mathbf{x} \setminus \mathbf{z}$.

Support and z-support. Let $\mathbf{z} \subseteq \mathbf{x}$ be a subset of variables and $m = x_1^{e_1} \cdot x_2^{e_2} \cdot \ldots \cdot x_N^{e_N}$ in $\mathbb{F}[\mathbf{x}]$ be a monomial. The support of m (respectively the $\mathbf{z}$-support of m), denoted $\mathrm{Supp}(m)$

(respectively $\mathbf{z}$-support) is the subset of variables (respectively the subset of $\mathbf{z}$-variables) appearing in it, i.e.

$$\mathrm{Supp}(m) \stackrel{\text{def}}{=} \{i \ : \ e_i \geq 1\} \subseteq [N], \quad \mathrm{Supp}_{\mathbf{z}}(m) \stackrel{\text{def}}{=} \{i \ : \ x_i \in \mathbf{z} \quad \text{and } e_i \geq 1\} \subseteq [N].$$

Isomorphic polynomials. We will say that two N-variate polynomials $f(\mathbf{x}), g(\mathbf{x}) \in \mathbb{F}[x_1, x_2, \ldots, x_N]$ are isomorphic if one can be obtained from the other via a renaming of the variables, i.e. there exists a permutation $\pi \in S_N$ such that

$$f(x_1, x_2, \ldots, x_N) = g(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(N)}).$$

Linear Maps and homomorphisms. We will sometimes need to take a subset $\mathbf{y} \subseteq \mathbf{x}$ of variables from $\mathbf{x}$ and set them to zero in some relevant collection of polynomials. We employ the following notation in this regard. For a subset $\mathbf{y} \subseteq \mathbf{x}$, $\sigma_{\mathbf{y}} : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ shall denote the homomorphism corresponding to setting the variables in $\mathbf{y}$ to zero. It is formally defined as[18] :

$$\sigma_{\mathbf{y}}(x_i) = \begin{cases} 0 & \text{if } x_i \in \mathbf{y} \\ x_i & \text{otherwise.} \end{cases}$$

For any set of polynomials $S \subseteq \mathbb{F}[\mathbf{x}]$ and any (linear) map $\sigma : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$, $\sigma(S)$ shall denote the set of polynomials obtained by applying σ pointwise to every polynomial in S, i.e.

$$\sigma(S) \stackrel{\text{def}}{=} \{\sigma(f) \ : \ f \in S\}.$$

The Iterated Matrix Multiplication. Let $n, d \geq 2$ be two parameters. We consider polynomials defined on variable sets $\mathbf{x}_1, \ldots, \mathbf{x}_d$. For $i \in [d] \setminus \{1, d\}$, let $\mathbf{x}_i$ be the set of variables $x_{i,j,k}$ for $j, k \in [n]$; for $i = 1$, let $\mathbf{x}_1$ be the set of variables $x_{1,1,j}$ and for $i = d$, let $\mathbf{x}_d$ be the set of variables $x_{d,j,1}$ where $j \in [n]$. Let $\mathbf{x} = \bigcup_{i \in [d]} \mathbf{x}_i$.
The Iterated Matrix Multiplication polynomial on $\mathbf{x}$, denoted $\mathrm{IMM}_{n,d}$, is defined to be $\mathrm{IMM}_{n,d} = \sum_{j_1, \ldots, j_{d-1}} x_{1,1,j_1} \cdot x_{2,j_1,j_2} \cdot x_{3,j_2,j_3} \cdot \ldots \cdot x_{d-1,j_{d-2},j_{d-1}} \cdot x_{d,j_{d-1},1}$. Note that the polynomial $\mathrm{IMM}_{n,d}$ is the value of the product of d matrices $M_1, M_2, \ldots, M_d$ (of dimensions $1 \times n$, $n \times n$ ($d-2$ times), and $n \times 1$).

3.2 Multi-r-ic models

Let Φ be an arithmetic formula. $\widehat{\Phi} \in \mathbb{F}[\mathbf{x}]$ will denote the output polynomial computed by Φ while $\mathrm{Size}(\Phi)$ will denote the size (the number of leaves) in Φ. If α is a node of Φ, we will denote by Φ_α the sub-formula rooted at α. In analyzing formulas in which the formal degree of every variable is bounded, it is naturally convenient to keep track of the degree of each variable in the polynomials computed at intermediate nodes of the formula. We will employ the following notation for this purpose.

[18] This map then extends via $\mathbb{F}$-linearity and multiplicativity to all of $\mathbb{F}[\mathbf{x}]$.

Definition 2. *Let* $\mathbf{r} = (r_1, r_2, \ldots, r_N)$ *be an* N*-tuple of nonnegative integers.*

1. **Support of r.** *The support of* $\mathbf{r}$ *is the set of indices of the non-zero coordinates. More formally,*

$$\mathrm{Supp}(\mathbf{r}) \stackrel{\mathrm{def}}{=} \{i \in [N] \mid r_i \neq 0\}.$$

2. **Multi-r-ic polynomials.** *We say a polynomial* $f(\mathbf{x})$ *is* a multi-$\mathbf{r}$-ic polynomial *if* $\deg_{x_i}(f) \leq r_i$ *for all* $i \in [N]$.

3. **Multi-r-ic formulas.** *We say an arithmetic formula* Φ *is* a multi-$\mathbf{r}$-ic formula *if for all* $i \in [N]$*, the formal degree of* Φ *with respect to the variable* x_i *is at most* r_i.

For conciseness, a multi-$(r, r, \ldots, r)$-polynomial[19] (resp. formula) will be referred (as defined before) to simply as a multi-r-ic polynomial (resp. formula). In particular, multi-1-ic polynomials are exactly multilinear polynomials and multi-1-ic formulas are syntactically multilinear formulas. This notational device will aid us in analyzing formulas via the labelling of every node α with an N-tuple that upper bounds the syntactic degree of the polynomial computed at α with respect to the various formal variables. We refer to such labelled formulas as *certified multi-*$\mathbf{r}$*-ic formulas* and the precise properties that such labellings satisfy are captured in the definition below.

Definition 3. Certified multi-r-ic formulas. *Let* $\mathbf{r} = (r_1, r_2, \ldots, r_N)$ *be an* N*-tuple of nonnegative integers. A* certified multi-$\mathbf{r}$-ic formula *is an arithmetic formula such that each gate* α *is labelled by an* N*-tuple* $\mathbf{d}_\alpha = (d_1, \ldots, d_N)$ *of non-negative integers such that*

- *the output is labelled by* $\mathbf{r}$,
- *if* α *is the input variable* x_i*, then* $d_i \geq 1$,
- *if* α *is an addition gate with children* $\beta_1, \beta_2, \ldots, \beta_p$*, then* $\mathbf{d}_\alpha = \mathbf{d}_{\beta_1} = \mathbf{d}_{\beta_2} = \ldots = \mathbf{d}_{\beta_p}$,
- *and if* α *is a multiplication gate of children* $\beta_1, \beta_2, \ldots, \beta_p$ *then*[20]

$$\mathbf{d}_\alpha = \mathbf{d}_{\beta_1} + \mathbf{d}_{\beta_2} + \ldots + \mathbf{d}_{\beta_p}.$$

It is readily verified that a natural top-down labelling procedure provides such a labelling to any multi-$\mathbf{r}$-ic formula.

Proposition 10. *Let* $\mathbf{r} = (r_1, r_2, \ldots, r_N)$ *be an* N*-tuple of nonnegative integers. If* Φ *is a certified multi-*$\mathbf{r}$*-ic formula, then the formal degree of* Φ *with respect to the variable* x_i *is at most* r_i*. Conversely, if* Φ *is a multi-*$\mathbf{r}$*-ic arithmetic formula, then there exists a labelling of the vertices such that the labelled formula is certified multi-*$\mathbf{r}$*-ic.*

[19] Here, r is any positive integer.

[20] Here '+' naturally denotes component-wise addition of N-tuples.

Proof. The forward direction is immediate by definitions. Now consider the other direction. We will assign a labelling so that if a gate α is labelled by $\mathbf{d} = (d_1, d_2, \ldots, d_N)$, then the formal degree of α with respect to x_i is at most d_i. Let us show it by descending induction on Φ. If α is the output, we label it by $\mathbf{r}$. By assumption we have that the degree with respect to each variable x_i is bounded by r_i. If α is an addition gate labelled by $\mathbf{d}_\alpha$ such that its children are not labelled, we label all the children by $\mathbf{d}_\alpha$. As the formal degrees of the children are bounded by the degree of α, the condition is satisfied. Otherwise α is a multiplication gate labelled by $\mathbf{d}_\alpha = (d_1, d_2, \ldots, d_N)$ having non-labelled children $\beta_1, \ldots, \beta_p$. For $i \in [N]$, let the formal degree of β_j with respect to x_i be e_{ji}. Then the formal degree of α with respect to x_i equals $e_{1i} + \ldots + e_{pi}$ and is upper bounded by d_i. Then for all $j \geq 2$ we assign the label

$$\mathbf{d}_{\beta_j} \stackrel{\text{def}}{=} (e_{j1}, e_{j2}, \ldots, e_{jN})$$

to the child β_j and

$$\mathbf{d}_{\beta_1} \stackrel{\text{def}}{=} \mathbf{d}_\alpha - \mathbf{d}_{\beta_2} - \mathbf{d}_{\beta_3} - \ldots - \mathbf{d}_{\beta_p}$$

to the first child β_1. Thus the sum of the $\mathbf{d}_{\beta_j}$'s equals $\mathbf{d}_\alpha$ and it is easily checked that the i-th coordinate of $\mathbf{d}_{\beta_j}$ is an upper bound for the formal degree of β_j with respect to x_i (for all $j \in [p]$ and $i \in [N]$). $\square$

4 Homogeneous Multi-r-ic Formulas

In this section we implement the strategy outlined in section 2 to obtain superpolynomial lower bounds for homogeneous multi-r-ic formulas.

4.1 Log-Product Decomposition

In this section we show that if a multi-r-ic polynomial $f(\mathbf{x})$ is computed by a multi-r-ic formula Φ of size s then f can be written as a sum of s polynomials having a rather *special structure* that we will exploit. We first capture the structure of the summands in the following definition.

Definition 4. *Let $\mathbf{d} = (d_1, \ldots, d_N)$ be an N-tuple of nonnegative integers and $T(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial on N variables. Let $v, L \geq 0$ be integers. We will say that T has a $(\mathbf{d}, v, L)$-form if there exist v pairs*

$$(g_1(\mathbf{x}), \mathbf{d}_1), (g_2(\mathbf{x}), \mathbf{d}_2), \ldots, (g_v(\mathbf{x}), \mathbf{d}_v),$$

where each $\mathbf{d}_j \in \mathbb{Z}^N_{\geq 0}$ is an N-tuple and each $g_j(\mathbf{x})$ is a multi-$\mathbf{d}_j$-ic homogeneous polynomial such that:

1. *$T(\mathbf{x})$ is a multi-$\mathbf{d}$-ic polynomial,*
2. *$T(\mathbf{x}) = g_1(\mathbf{x}) \cdot g_2(\mathbf{x}) \cdot \ldots \cdot g_v(\mathbf{x})$,*

3. $\mathbf{d} = \mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_v$,

4. *for all* $i \in [v]$ *we have* $\left|\mathrm{Supp}(\mathbf{d}_i) \setminus \left(\bigcup_{1\le j<i} \mathrm{Supp}(\mathbf{d}_j)\right)\right| \ge L$.

Intuitively, the first three conditions specify that T is a multi-$\mathbf{d}$-ic polynomial that is a length-v product of multi-$\mathbf{d}_i$-ic polynomials. The fourth condition intuitively says that each factor g_i contains at least L *fresh variables*, i.e. variables which do not occur in the previous g_j's.

Lemma 11. *Let $v \ge 1$ be an integer, $\mathbf{d} = (d_1, \ldots, d_N)$ be an N-tuple of non-negative integers and f be a N-variate polynomial computed by a homogeneous multi-$\mathbf{d}$-ic formula of size s. If $|\mathrm{Supp}(\mathbf{d})| \ge 3^{v-1}\sqrt{N}$, then there exist s homogeneous $(\mathbf{d}, v, \sqrt{N})$-form polynomials $T_1, T_2, \ldots, T_s$ such that $f = T_1 + T_2 + \ldots + T_s$.*

We will need a suitable adaptation of an observation that is common to many depth reduction results for arithmetic (and even Boolean) formulas.

Proposition 12. *Let $f(\mathbf{x})$ be a polynomial computed by a certified multi-$\mathbf{d}$-ic formula Φ and let α be any node in Φ having label $\mathbf{d}_\alpha$. Then there exist certified formulas Ψ and Λ such that*

$$\widehat{\Phi} = \widehat{\Psi} \cdot \widehat{\Phi_\alpha} + \widehat{\Lambda}$$

where

1. *Λ is a certified multi-$\mathbf{d}$-ic formula, Ψ is a certified multi-$(\mathbf{d} - \mathbf{d}_\alpha)$-ic formula, and*

2. $\mathrm{Size}(\Phi_\alpha) + \mathrm{Size}(\Lambda) \le \mathrm{Size}(\Phi)$.

3. *If Φ is homogeneous then Ψ, Λ are also homogeneous formulas.*

4. *If Φ is of product-depth p then Ψ, Λ have product-depth at most p.*

See the appendix 10 for the proof. Here we see how to use it to get the desired depth reduction.

Proof of Lemma 11. By Proposition 10 we can assume that the formula Φ is in fact a certified multi-$\mathbf{d}$-ic homogeneous formula. We can also assume without loss of generality[21] that every node in the formulas has fanin[22] at most 2. We prove the result by induction on v and s.

Induction basis. If $v = 1$, then f is already in $(\mathbf{d}, 1, \sqrt{N})$-form.

If $s = 1$, the formula is just a leaf node and f depends on at most one variable - say x_i. Now since the support of $\mathbf{d}$ is large (at least $3^{v-1} \cdot \sqrt{N} \ge v \cdot \sqrt{N}$) we can decompose $\mathbf{d}$ as

$$\mathbf{d} = \mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_{v-1} + \mathbf{d}_v$$

such that:

[21] Since we use the number of leaves of a formula as a measure of its size.
[22] The fanin of a node is the number of inputs to a node.

1. $i \in \mathrm{Supp}(\mathbf{d}_1)$,
2. $|\mathrm{Supp}(\mathbf{d}_1)| = |\mathrm{Supp}(\mathbf{d}_2)| = \ldots = |\mathrm{Supp}(\mathbf{d}_{v-1})| = \sqrt{N}$,
3. $\mathrm{Supp}(\mathbf{d}_i) \cap \mathrm{Supp}(\mathbf{d}_j) = \emptyset$ for any $1 \leq i < j \leq v$.

Then, it is easy to check that f is in $(\mathbf{d}, v, \sqrt{N})$-form via the pairs

$$(f, \mathbf{d}_1), (1, \mathbf{d}_2), \ldots, (1, \mathbf{d}_v).$$

That concludes this case.

Induction step. Let us prove the lemma for a fixed value of (v, s) with $v \geq 2$. By hypothesis, f is computed by a homogeneous certified multi-$\mathbf{d}$-ic formula Φ. There are two cases.

- If one leaf α of Φ is such that $|\mathrm{Supp}(\mathbf{d}_\alpha)| \geq \frac{|\mathrm{Supp}(\mathbf{d})|}{3}$ where $\mathbf{d}_\alpha$ is the label of the node α. By Proposition 12 we have $\widehat{\Phi} = \widehat{\Psi} \cdot \widehat{\Phi_\alpha} + \widehat{\Lambda}$ where Λ is a multi-$\mathbf{d}$-ic formula of size at most $s-1$. Using the inductive hypothesis, it suffices to show that $(\widehat{\Psi} \cdot \widehat{\Phi_\alpha})$ is in $(\mathbf{d}, v, \sqrt{N})$-form. To see this first note that since α is a leaf node so $\widehat{\Phi_\alpha}$ depends on at most one variable - say x_a. Now since the support of $\mathbf{d}_\alpha$ is large (at least $3^{v-2} \cdot \sqrt{N} \geq (v-1) \cdot \sqrt{N}$) we can decompose $\mathbf{d}_\alpha$ as

  $$\mathbf{d}_\alpha = \mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_{v-1} + \mathbf{d}_v$$

  such that:

  1. $a \in \mathrm{Supp}(\mathbf{d}_1)$,
  2. $|\mathrm{Supp}(\mathbf{d}_1)| = |\mathrm{Supp}(\mathbf{d}_2)| = \ldots = |\mathrm{Supp}(\mathbf{d}_{v-1})| = \sqrt{N}$,
  3. $\mathrm{Supp}(\mathbf{d}_i) \cap \mathrm{Supp}(\mathbf{d}_j) = \emptyset$ for any $1 \leq i < j \leq (v-1)$.

  Also we have

  $$|\mathrm{Supp}(\mathbf{d} - (\mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_{v-1}))| \geq 3^{v-1} \cdot \sqrt{N} - (v-1) \cdot \sqrt{N} \geq \sqrt{N}.$$

  With the above facts in hand, it is easy to check that $(\widehat{\Psi} \cdot \widehat{\Phi_\alpha})$ is in $(\mathbf{d}, v, \sqrt{N})$-form via the pairs

  $$(\widehat{\Phi_\alpha}, \mathbf{d}_1), (1, \mathbf{d}_2), (1, \mathbf{d}_3), \ldots, (1, \mathbf{d}_{v-1}), (\widehat{\Psi}, \mathbf{d} - (\mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_{v-1})).$$

  That concludes this case.

- Otherwise all the leaves have support smaller than $\frac{|\mathrm{Supp}(\mathbf{d})|}{3}$. As the fan-in of every gate is bounded by 2, there exists a node α in Φ with label $\mathbf{d}_\alpha$ such that

  $$\frac{1}{3}\,|\mathrm{Supp}(\mathbf{d})| \leq |\mathrm{Supp}(\mathbf{d}_\alpha)| \leq \frac{2}{3} \cdot |\mathrm{Supp}(\mathbf{d})|\,.$$

By Proposition 12, there exist certified formulas Ψ and Λ such that

$$\widehat{\Phi} = \widehat{\Phi_\alpha} \cdot \widehat{\Psi} + \widehat{\Lambda},$$

where size of Φ_α and of Λ is at most $(s-1)$ each. So we apply the induction hypothesis on Φ_α and Λ. Now Λ is a certified multi-$\mathbf{d}$-ic formula so by induction hypothesis

$$\widehat{\Lambda} = T_1'(\mathbf{x}) + \ldots + T_{s_1}'(\mathbf{x})$$

where each T_i' has a $(\mathbf{d}, v, \sqrt{N})$-form and $s_1 \leq \mathrm{Size}(\Lambda)$. Now $\widehat{\Psi}$ is a homogeneous multi-$(\mathbf{d} - \mathbf{d}_\alpha)$-ic polynomial. Moreover,

$$|\mathrm{Supp}(\mathbf{d}_\alpha)| \geq \frac{1}{3}\,|\mathrm{Supp}(\mathbf{d})| \geq 3^{v-2} \cdot \sqrt{N}.$$

By the induction hypothesis,

$$\widehat{\Phi_\alpha} = T_1(\mathbf{x}) + \ldots + T_{s_2}(\mathbf{x}),$$

where each T_i has a $(\mathbf{d}_\alpha, v-1, \sqrt{N})$-form and $s_2 \leq \mathrm{Size}(\Phi_\alpha)$. Thus,

$$f = \left(\sum_{i=1}^{s_2} T_i \cdot \widehat{\Psi}\right) + \left(\sum_{i=1}^{s_1} T_i'\right).$$

Since

$$\begin{aligned} s_1 + s_2 &\leq \mathrm{Size}(\Lambda) + \mathrm{Size}(\Phi_\alpha) \\ &\leq \mathrm{Size}(\Phi) \end{aligned}$$

it suffices to show that for all i the polynomials $(T_i \cdot \widehat{\Psi})$ have a $(\mathbf{d}, v, \sqrt{N})$-form. Since each T_i is already in $(\mathbf{d}_\alpha, v-1, \sqrt{N})$ and $\widehat{\Psi}$ is a multi-$(\mathbf{d}-\mathbf{d}_\alpha)$-ic polynomial, it suffices to verify that $|\mathrm{Supp}(\mathbf{d}) \setminus \mathrm{Supp}(\mathbf{d}_\alpha)| \geq \sqrt{N}$. Now

$$\begin{aligned} |\mathrm{Supp}(\mathbf{d}) \setminus \mathrm{Supp}(\mathbf{d}_\alpha)| &\geq |\mathrm{Supp}(\mathbf{d})| - |\mathrm{Supp}(\mathbf{d}_\alpha)| \\ &\geq |\mathrm{Supp}(\mathbf{d})| - \frac{2}{3} \cdot |\mathrm{Supp}(\mathbf{d})| \\ &\geq \frac{1}{3} \cdot 3^{v-1} \cdot \sqrt{N} \\ &\geq \sqrt{N} \quad (\text{as } v \geq 2). \end{aligned}$$

This completes the inductive step and hence proves the lemma.

$\square$

Corollary 13. *Let $r \geq 1$ be an integer and let $\mathbf{r} = (r, r, \ldots, r)$. If a N-variate polynomial f is computed by a homogeneous multi-r-ic formula of size s then there exist s homogeneous multi-$\mathbf{r}$-ic polynomials $T_1, T_2, \ldots, T_s$ such that*

$$f = \sum_{i=1}^{s} T_i \quad \text{where each } T_i \text{ has a homogeneous } (\mathbf{r}, \log(N)/4, \sqrt{N})\text{-form.} \tag{3}$$

Proof. We can directly apply Lemma 11 for $v = \log(N)/4$. It is possible since $|\mathrm{Supp}(\mathbf{r})| = N \geq 3^{\frac{\log(N)}{4}-1} \cdot \sqrt{N}$. ∎

4.2 Counting extremal monomials

From the log-product decomposition described in the previous section, our problem boils down to understanding sums of $(\mathbf{d}, v, L)$-forms. Our next definition will help us describe the *weakness* of such summands that is being exploited here.

Definition 5. (Extremal monomials) *Let $\mathbf{d} = (d_1, d_2, \ldots, d_N) \in \mathbb{Z}_{\geq 0}^N$ be an N-tuple and $m \in \mathbb{F}[\mathbf{x}]$ be a multi-$\mathbf{d}$-ic monomial on N variables. We will call m as a* $\mathbf{d}$-extremal monomial *if the degree of m with respect to any variable x_j is either the minimum possible or the maximum possible amount, i.e.* $\deg_{x_j}(m) \in \{0, d_j\}$ *for all $j \in [N]$.*

In this subsection, we will show that a term of our log-product decomposition (Corollary 13) has a relatively small number of extremal monomials. Specifically,

Lemma 14. Upper bound on the number of extremal monomials in a term. *Let $T(\mathbf{x})$ be an N-variate polynomial and $\mathbf{d} \in \mathbb{Z}_{\geq 0}^N$ be an N-tuple of non-negative integers. If T is homogeneous and has a $(\mathbf{d}, v, L)$-form then the number of $\mathbf{d}$-extremal monomials in T is at most $\frac{2^N}{L^{v/2}}$.*

The rest of this subsection is devoted to a proof of this lemma. We will need the following result from extremal combinatorics due to Sperner [Spe28] (a proof of this can be found in the book [AS04]).

Theorem 15. *[Sperner's Theorem] Let N be an integer and $\mathcal{F} \subseteq 2^{[N]}$ be a set of subsets of $[N]$. Such an $\mathcal{F}$ is called an* antichain *if and only if for all distinct I and J in $\mathcal{F}$ we have $I \not\subseteq J$ and $J \not\subseteq I$. If $\mathcal{F} \subseteq 2^{[N]}$ is an antichain then*

$$|\mathcal{F}| \leq \binom{N}{N/2}.$$

We use it to first bound the number of extremal monomials in any homogeneous polynomial.

Lemma 16. *Let $\mathbf{d} \in \mathbb{Z}_{\geq 0}^N$ be an N-tuple and $g(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a multi-$\mathbf{d}$-ic polynomial. Let $N_1 \stackrel{\text{def}}{=} |\mathrm{Supp}(\mathbf{d})|$. If g is homogeneous then the number of $\mathbf{d}$-extremal monomials in g is at most*

$$\binom{N_1}{N_1/2} \leq \frac{2^{N_1}}{\sqrt{N_1}}.$$

Proof of Lemma 16. Let $\mathbf{d} = (d_1, d_2, \ldots, d_N)$. We can first assume that $d_i > 0$ and that $\deg_{x_i}(g) = d_i$ for all $i \in [N]$ (otherwise the corresponding variable x_i does not appear in any $\mathbf{d}$-extremal monomial in g and we can effectively remove this variable from consideration). Now looking at the support of the extremal monomials in g gives us a collection of subsets of $[N]$. Specifically, for any $\mathbf{d}$-extremal monomial m let:

$$S_m \stackrel{\text{def}}{=} \left\{ i \in [N] \ : \ \deg_{x_i}(m) = d_i \right\} \subseteq [N].$$

Let us consider the set

$$\mathcal{I}_g \stackrel{\text{def}}{=} \{S_m \ : \ m \text{ is } \mathbf{d}\text{-extremal}\} \quad \subseteq 2^{[N]}.$$

Via Sperner's theorem (Theorem 15), it is sufficient to prove that $\mathcal{I}_g$ is an antichain. If it is not the case, it means there exist $\mathbf{d}$-extremal monomials m_1 and m_2 in g such that S_{m_1} is a proper subset of S_{m_2}. In particular,

$$\deg(m_2) \stackrel{\text{def}}{=} \sum_{i \in S_{m_2}} d_i = \sum_{i \in S_{m_1}} d_i + \sum_{j \in (S_{m_2} \setminus S_{m_1})} d_j > \deg(m_1) \quad (\text{since every } d_j > 0).$$

This contradicts the premise that g is homogeneous. $\square$

Lemma 17. *Let $\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_v \in \mathbb{Z}^N_{\geq 0}$ be N-tuples and $g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_v(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be N-variate polynomials such that the i-th polynomial $g_i(\mathbf{x})$ is multi-$\mathbf{d}_i$-ic. Let*

$$\mathbf{d} \stackrel{\text{def}}{=} \mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_v \quad \text{and } T = g_1 \cdot g_2 \cdot \ldots \cdot g_v.$$

Let

$$N_i \stackrel{\text{def}}{=} \left| \mathrm{Supp}(\mathbf{d}_i) \setminus \left(\bigcup_{j<i} \mathrm{Supp}(\mathbf{d}_j) \right) \right|.$$

If all the g_i's are homogeneous polynomials then the number of $\mathbf{d}$-extremal monomials in T is at most

$$\frac{2^{N_1 + N_2 + \ldots + N_v}}{\sqrt{N_1 \cdot N_2 \cdot \ldots N_v}}.$$

Proof. We prove it by induction on v. We can first assume without loss of generality that $\mathrm{Supp}(\mathbf{d}) = [N]$ (otherwise for any $i \in [N] \setminus \mathrm{Supp}(\mathbf{d})$ the corresponding variable x_i does not appear in any monomial in T and we can effectively remove this variable from consideration). If $v = 1$, the result is directly given by Lemma 16. Now consider the case $T = g_1 \cdot g_2 \cdot \ldots \cdot g_v \cdot g_{v+1}$. Let

$$g = g_1 \cdot g_2 \cdot \ldots \cdot g_v \quad \text{and } \mathbf{e} = \mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_v = (e_1, \ldots, e_N)$$

so that

$$T = g \cdot g_{v+1} \quad \text{and } \mathbf{d} = \mathbf{e} + \mathbf{d}_{v+1}.$$

Now note that a $\mathbf{d}$-extremal monomial in T is a monomial from

$$\{\mathbf{e}\text{-extremal monomials in } g\} \cdot \{\mathbf{d}_{v+1}\text{-extremal monomials in } g_{v+1}\}.$$

Claim 18. *If m_1 is an $\mathbf{e}$-extremal monomial in g, the number of monomials m_2 from g_{v+1} such that $(m_1 \cdot m_2)$ is $\mathbf{d}$-extremal in T is at most*

$$\frac{2^{N_{v+1}}}{\sqrt{N_{v+1}}}.$$

Let us first assume that the claim is true. In this case, by induction hypothesis the number of $\mathbf{e}$-extremal monomials in g is at most

$$\frac{2^{N_1+\ldots+N_v}}{\sqrt{N_1 \cdot \ldots \cdot N_v}}$$

which implies the result. Let us prove the claim. Fix any $\mathbf{e}$-extremal monomial m_1. Now there is a natural partition induced on the set of variables as follows:

$$Y \stackrel{\text{def}}{=} \text{Supp}(\mathbf{e}) \quad \text{and} \quad Z \stackrel{\text{def}}{=} (\text{Supp}(\mathbf{d}_{v+1}) \setminus \text{Supp}(\mathbf{e})) = [N] \setminus Y.$$

Fix an $\mathbf{e}$-extremal monomial m_1 in g. We observe that there is a one-one correspondence between the set of all possible monomials m_2 such that $(m_1 \cdot m_2)$ is $\mathbf{d}$-extremal and subsets of Z. To see this observe that fixing m_1 completely determines the degree of m_2 with respect to any Y-variable[23] x_j - if $j \in Y$ then

$$\deg_{x_j}(m_2) = \begin{cases} 0 & \text{if } \deg_{x_j}(m_1) = 0 \\ d_j - e_j & \text{otherwise.} \end{cases}$$

For the remaining variables, i.e. the Z-variables, the degree of m_2 with respect to x_j must be either 0 or $(d_j - e_j) = d_j$. Define the subset of Z corresponding to m_2 as

$$S_{m_2} \stackrel{\text{def}}{=} \left\{ j \in Z \ : \ \deg_{x_j}(m_2) = d_j \right\}.$$

We now use the homogeneity of g_{v+1} and proceed as in the proof of Lemma 16 (by considering here the degree with respect to Z-variables) to deduce that the set

$$\mathcal{I}_{m_1} \stackrel{\text{def}}{=} \{S_{m_2} \ : \ m_2 \text{ is in } g_{v+1} \text{ and } (m_1 \cdot m_2) \text{ is a } \mathbf{d}\text{-extremal monomial}\} \subseteq 2^Z$$

forms an antichain in 2^Z. So by Sperner's theorem $\mathcal{I}_{m_1}$ can have size at most $\binom{|Z|}{|Z|/2} \le \frac{2^{N_{v+1}}}{\sqrt{N_{v+1}}}$. This proves the claim and hence the lemma as well. $\square$

The upper bound on the number of $\mathbf{d}$-extremal monomials in a $(\mathbf{d}, v, L)$-form follows immediately.

Proof of Lemma 14. Since T is in $(\mathbf{d}, v, L)$-form by definition there exist pairs $(\mathbf{d}_1, g_1), (\mathbf{d}_2, g_2), \ldots, (\mathbf{d}_v, g_v)$ such that each g_i is multi-$\mathbf{d}_i$-ic and

$$\mathbf{d} = \mathbf{d}_1 + \ldots + \mathbf{d}_v \quad \text{and } T = g_1 \cdot g_2 \cdot \ldots \cdot g_v,$$

[23] We say a variable x_j is a Y-variable iff $j \in Y$.

and

$$N_i \stackrel{\text{def}}{=} \left| \mathrm{Supp}(\mathbf{d}_i) \setminus \left(\bigcup_{j<i} \mathrm{Supp}(\mathbf{d}_j) \right) \right| \geq L.$$

Furthermore, since T is homogeneous all the g_i's are homogeneous as well. Applying Lemma 17 we get that the number of $\mathbf{d}$-extremal monomials in T is at most

$$\frac{2^{N_1+N_2+\ldots+N_v}}{\sqrt{N_1 \cdot N_2 \cdot \ldots \cdot N_v}} \leq \frac{2^N}{L^{v/2}},$$

as required. □

4.3 Putting things together.

Our *target polynomial* is a multi-r-ic adaptation of the elementary symmetric polynomial obtained by raising every variable to the r-th power. Specifically, let

$$P_{N,r}(x_1, \ldots, x_N) = \sum_{S \in \binom{[N]}{N/2}} \left(\prod_{j \in S} x_j^r \right). \tag{4}$$

Then $P_{N,r}$ is a homogeneous multi-r-ic polynomial of degree $\frac{rN}{2}$ which contains $\binom{N}{N/2}$ extremal monomials. A relatively straightforward adaptation of an observation attributed to Michael Ben-Or [NW96] implies that $P_{N,r}$ is easy to compute.

Proposition 19. *Let $\mathbb{F}$ be a field which contains r distinct rth roots of unity and size of $\mathbb{F}$ is at least $(2Nr)$. The polynomial $P_{N,r}$ defined above can be computed by a multi-r-ic (nonhomogeneous) $\Sigma\Pi\Sigma$ circuit of size $O(N^2 r)$ on $\mathbb{F}$.*

Proof of Theorem 1. Our target polynomial is the polynomial $P_{N,r}$ defined above. Suppose that it is computed by a homogeneous multi-r-ic formula Φ of size s. Let $\mathbf{r} = (r, r, \ldots, r)$. By Lemma 11, Φ can be written as a sum of size at most s of homogeneous $(\mathbf{r}, \log(N)/4, \sqrt{N})$-form polynomials. By Lemma 14 each one of these polynomials can compute at most $2^N/(2^{\log^2 N/16})$-many $\mathbf{r}$-extremal monomials. But $P_{N,r}$ has $\binom{N}{N/2}$-many $\mathbf{r}$-extremal monomials. Therefore we must have

$$\begin{aligned} s &\geq \frac{\binom{N}{N/2}}{2^N/(2^{\log^2 N/16})} \\ &= N^{\Omega(\log N)}. \end{aligned}$$

The moreover part follows from Proposition 19. □

5 Constant Depth Homogeneous Multi-r-ic Formulas

We follow the same overall proof strategy as in the previous section to obtain a lower bound for homogeneous multi-r-ic formulas of small product-depth. With definitions as in section 4, the depth reduction for low-depth formulas that we obtain is:

Lemma 20. *Let $r \geq 1$ be an integer and $\mathbf{r} = (r, r, \ldots, r)$. Let f be a degree-d polynomial computed by a homogeneous certified multi-$\mathbf{r}$-ic formula Φ of size s and product-depth bounded by p. For any positive integer $v \leq \left\lfloor \frac{1}{8r}\left(\frac{d}{2r}\right)^{1/p} \right\rfloor$, there exist s homogeneous $(\mathbf{r}, v, 2)$-form polynomials $T_1, T_2, \ldots, T_s$ such that $f = T_1 + T_2 + \ldots + T_s$.*

This depth reduction proceeds in two stages. The following definition will help us describe the structure of the output of the first stage.

Definition 6. *Let $\mathbf{d} = (d_1, \ldots, d_N)$ be an N-tuple of non-negative integers. We will say that a polynomial $T(\mathbf{x})$ is* $(\mathbf{d}, t, L)$-balanced *if T is multi-$\mathbf{d}$-ic and there exist t pairs $(\mathbf{d}_1, g_1), (\mathbf{d}_2, g_2), \ldots, (\mathbf{d}_t, g_t)$ where each g_i is a multi-$\mathbf{d}_i$-ic polynomial such that*

1. *$T = g_1 \cdot g_2 \cdot \ldots \cdot g_t$,*
2. *$\mathbf{d} = \mathbf{d}_1 + \mathbf{d}_2 + \ldots + \mathbf{d}_t$, and*
3. *$|\mathrm{Supp}(\mathbf{d}_i)| \geq L$ for all $i \in [t]$.*

Let us notice that if a polynomial is $(\mathbf{d}, a, L)$-balanced for a parameter $a > t$, then we can easily get a $(\mathbf{d}, t, L)$-balanced expression by grouping some factors.
The first stage flattens the formula into a sum of a small number of balanced polynomials and its proof idea comes from the paper [HY11]. Specifically, we have:

Lemma 21. *Let $r \geq 1$ be an integer and let $\mathbf{r} = (r, r, \ldots, r)$. Let Φ be a certified multi-$\mathbf{r}$-ic homogeneous formula of product-depth p computing a polynomial of degree $d \geq 2r$. For any positive integer $t \leq \frac{1}{4} \cdot \left(\frac{d}{2r}\right)^{1/p}$ there exist homogeneous $(\mathbf{r}, t, 2)$-balanced polynomials $T_1, \ldots, T_s$ such that $s \leq \mathrm{Size}(\Phi)$ and*

$$\widehat{\Phi} = T_1 + \ldots + T_s.$$

Proof of Lemma 21. First let us note the following:

Claim 22. *For any positive integers d and p with $d \geq 2r$, for any positive integer $t \leq \frac{1}{4} \cdot \left(\frac{d}{2r}\right)^{1/p}$, there exists a product node α in Φ such that $\deg(\alpha) \geq 2r$[24] and for every child β of α it holds that $\deg \beta < \frac{\deg(\alpha)}{4t}$. Moreover, $\widehat{\Phi_\alpha}$ is $(\mathbf{d}_\alpha, t, 2)$-balanced.*

[24] Recall that for a node α, the formal degree of α is denoted as $\deg(\alpha)$.

Proof. The proof is by induction on p. If $p = 1$ and $\beta = \beta_1 \times \ldots \times \beta_v$ is a product node in Φ, then $\deg(\beta) = d$ and $\deg(\beta_i) \leq 1 < d/4t$. So we can set $\alpha = \beta$. Assume that $p > 1$, and let $\beta = \beta_1 \times \ldots \times \beta_v$ be a product node in Φ with $\deg(\beta) = d$. If for every $i \in [v]$, $\deg(\beta_i) < d/(4t)$, then we can set $\alpha = \beta$. Otherwise there exists β_i such that $\deg(\beta_i) \geq d/(4t)$. In this case, Φ_{β_i} is of product-depth $p' < p$ and degree at least $d/(4t)$. We know $\deg(\beta_i) \geq d/4t \geq 2r$ and

$$(4t)^p \leq \frac{d}{2r} \leq 4t\frac{\deg \beta_i}{2r},$$
$$\text{so } (4t)^{p'} \leq (4t)^{p-1} \leq \frac{\deg \beta_i}{2r}.$$

In particular,

$$t \leq \frac{1}{4}\left(\frac{\deg \beta_i}{2r}\right)^{1/p'}.$$

By the inductive assumption, there exists a product node α in Φ_{β_i} such that $\deg(\alpha) \geq 2r$ and for every child β of α, $\deg \beta < \frac{\deg \alpha}{4t}$. We now show that $\widehat{\Phi_\alpha}$ is $(\mathbf{d}_\alpha, t, 2)$-balanced. Let the children of α be $\beta_1, \beta_2, \ldots, \beta_v$. We greedily merge pairs of polynomials $(\widehat{\Phi_{\beta_i}}, \widehat{\Phi_{\beta_j}})$ such that

$$|\mathrm{Supp}(\mathbf{d}_{\beta_i})| = \left|\mathrm{Supp}(\mathbf{d}_{\beta_j})\right| = 1 \quad \text{and } \mathrm{Supp}(\mathbf{d}_{\beta_i}) \neq \mathrm{Supp}(\mathbf{d}_{\beta_j})$$

and also add the corresponding $\mathbf{d}_{\beta_i}$ and $\mathbf{d}_{\beta_j}$. This means that $\widehat{\Phi_\alpha}$ can be written as

$$\widehat{\Phi_\alpha} = g_1 \cdot \ldots \cdot g_a \cdot h, \quad \text{where } \deg(h) \leq r$$

such that for each $i \in [a]$, we have $\deg(g_i) < 2 \cdot \frac{\deg(\alpha)}{4t}$ and the multi-degree label corresponding to g_i has support size at least 2. In particular,

$$2r \leq \deg(\alpha) \leq a\left(\frac{2\deg(\alpha)}{4t}\right) + r.$$

This implies that a is at least t. Finally, rewriting $\widehat{\Phi_\alpha}$ as $\widehat{\Phi_\alpha} = g_1 \cdot \ldots \cdot g_{a-1} \cdot (g_a \cdot h)$, we see that $\widehat{\Phi_\alpha}$ is $(\mathbf{d}_\alpha, a, 2)$-balanced where $a \geq t$. This proves the claim. □

Let α be a node given by Claim 22. By Proposition 12, there exist a $\mathbf{d}$-certified homogeneous formula Λ of product-depth at most p and a $(\mathbf{d} - \mathbf{d}_\alpha)$-certified homogeneous Ψ such that

$$\widehat{\Phi} = \widehat{\Psi} \cdot \widehat{\Phi_\alpha} + \widehat{\Lambda}.$$

By induction on size, $\widehat{\Lambda}$ can be expressed as a sum of at most $\mathrm{Size}(\Lambda)$-many $(\mathbf{d}, t, 2)$-balanced polynomials. Now $\widehat{\Phi_\alpha}$ is $(\mathbf{d}_\alpha, t, 2)$-balanced and so $(\widehat{\Psi} \cdot \widehat{\Phi_\alpha})$ is $(\mathbf{d}, t, 2)$-balanced (by grouping Ψ with another factor). Altogether, $\widehat{\Phi}$ can be written as a sum of $(\mathbf{d}, t, 2)$-balanced polynomials, the number of summands being at most $1 + \mathrm{Size}(\Lambda) \leq \mathrm{Size}(\Phi)$. This proves the lemma. □

Lemma 23. *Let $r, v \geq 1$ be integers and let $t = 2rv$. Let $\mathbf{r} = (r, r, \ldots, r)$. If $T(\mathbf{x})$ is a $(\mathbf{r}, t, 2)$-balanced polynomial then $T(\mathbf{x})$ has a $(\mathbf{r}, v, 2)$-form.*

Proof. By the premise of the lemma, there exist pairs $(\mathbf{d}_1, f_1), (\mathbf{d}_2, f_2), \ldots, (\mathbf{d}_t, f_t)$ such that

$$T = f_1 \cdot \ldots \cdot f_t, \quad \mathbf{r} = \mathbf{d}_1 + \ldots + \mathbf{d}_t, \quad \forall i \in [t] \;\; \mathrm{Supp}(\mathbf{d}_i) \geq 2, \quad f_i \text{ is multi-}\mathbf{d}_i\text{-ic.} \tag{5}$$

The proof is by reordering and regrouping the f_i's in a series of v steps. The following claim captures the invariants after the k-th step of this process.

Claim 24. *Let $k \in [0..v]$ be any integer. There exists a partition $[t] = A \uplus B \uplus C$ with $|A| = k$ and an order on the elements of $A = \{i_1, \ldots, i_k\}$ such that*

- *$f_{i_1} \cdot \ldots \cdot f_{i_k}$ is in $(\mathbf{e}, k, 2)$-form where $\mathbf{e} \stackrel{\mathrm{def}}{=} \mathbf{d}_{i_1} + \ldots + \mathbf{d}_{i_k}$,*
- *$|B| \geq t - 2rk$,*
- *and for all $j \in B$, $|\mathrm{Supp}(\mathbf{d}_j) \setminus \mathrm{Supp}(\mathbf{e})| \geq 2$*

Proof. Let us prove the claim by induction on k.
Base case $k = 0$. For the base case of $k = 0$, we can choose $A = C = \emptyset$ (so that $\mathbf{e} = (0, 0, \ldots, 0)$) and $B = [t]$ and the claim is easily verified using Equation (5).
Inductive step. Assume $k < v$ and let $[t] = A \uplus B \uplus C$ be the partition obtained after the k-th step. Let $\mathbf{e} = \sum_{i \in A} \mathbf{d}_i$ and let $Z = [N] \setminus \mathrm{Supp}(\mathbf{e})$ be the set of (indices of) variables which do not appear in A. We can associate to each f_i with $i \in B$ the support of $\mathbf{d}_i$ in Z:

$$\mathrm{Supp}_Z(\mathbf{d}_i) \stackrel{\mathrm{def}}{=} \mathrm{Supp}(\mathbf{d}_i) \cap Z.$$

We choose a factor f_u with $u \in B$ such that the size of its corresponding support ($\mathrm{Supp}(\mathbf{d}_u)$) in Z is minimum (as $k < v$, the set B is non-empty) and add it to A. We then remove from B *the neighbouring factors* $\mathcal{F}$ defined as

$$\mathcal{F} \stackrel{\mathrm{def}}{=} \{ j \in (B \setminus \{u\}) \;:\; |\mathrm{Supp}_Z(\mathbf{d}_j) \setminus \mathrm{Supp}_Z(\mathbf{d}_u)| \leq 1 \}.$$

The updated partition is given by

$$A' \stackrel{\mathrm{def}}{=} A \uplus \{u\}, \quad B' \stackrel{\mathrm{def}}{=} B \setminus (\{u\} \cup \mathcal{F}), \quad C' \stackrel{\mathrm{def}}{=} C \uplus \mathcal{F}.$$

To complete the induction, it suffices to verify that the size of B' is large enough (the other conditions follow from the choice of u and $\mathcal{F}$). In particular, by induction hypothesis $|\mathrm{Supp}_Z(\mathbf{d}_u)| \geq 2$. As $|\mathrm{Supp}_Z(\mathbf{d}_u)|$ is minimal, $|\mathrm{Supp}_Z(\mathbf{d}_u) \cap \mathrm{Supp}_Z(\mathbf{d}_j)| \geq |\mathrm{Supp}_Z(\mathbf{d}_u)| - 1$ for all $j \in \mathcal{F}$. Now counting the occurrences of the variables in $\mathrm{Supp}_Z(\mathbf{d}_u)$ in $\left(\mathbf{d}_u + \sum_{j \in \mathcal{F}} \mathbf{d}_j\right)$ and using the fact that T is multi-r-ic we get that

$$|\mathrm{Supp}_Z(\mathbf{d}_u)| \cdot r \geq |\mathrm{Supp}_Z(\mathbf{d}_u)| + (|\mathrm{Supp}_Z(\mathbf{d}_u)| - 1) \cdot |\mathcal{F}| \geq (|\mathrm{Supp}_Z(\mathbf{d}_u)| - 1) \cdot (|\mathcal{F}| + 1).$$

Hence

$$|\mathcal{F}| + 1 \leq r \cdot \frac{|\mathrm{Supp}_Z(\mathbf{d}_u)|}{|\mathrm{Supp}_Z(\mathbf{d}_u)| - 1} \leq 2r.$$

Thus

$$|B'| = |B| - (|\mathcal{F}| + 1) \geq t - 2r(k+1).$$

This completes the proof of the inductive step and hence of the claim.

□

Let us come back to the proof of the lemma. The previous claim when $k = v$ states that T can be written as $T = f_{i_1} \cdot \ldots \cdot f_{i_v} \cdot h$ where $f_{i_1} \cdot \ldots \cdot f_{i_v}$ is in $(\mathbf{e}, v, 2)$-form and $\mathbf{e} \stackrel{\text{def}}{=} \mathbf{d}_{i_1} + \ldots + \mathbf{d}_{i_v}$. Hence $T = f_{i_1} \cdot \ldots \cdot f_{i_{v-1}} \cdot (f_{i_v} \cdot h)$ is in $(\mathbf{r}, v, 2)$-form, as required. □

Proof of Lemma 20. Let $t = 2rv$. Then we have

$$t \leq \frac{1}{4} \cdot \left(\frac{d}{2r}\right)^{\frac{1}{p}}.$$

So we can apply Lemma 21. We get some $(\mathbf{r}, t, 2)$-balanced multi-r-ic polynomials $T_1, \ldots, T_s$ such that $s \leq \mathrm{Size}(\Phi)$ and $f = T_1 + \ldots + T_s$. By Lemma 23 each of these terms T_i has a $(\mathbf{r}, v, 2)$-form. □

Proof of Theorem 2. Our target polynomial is the polynomial $P_{N,r}$ defined in Equation (4) and used previously in the proof of Theorem 1. It has degree $d = \frac{rN}{2}$. Suppose that it is computed by a homogeneous multi-r-ic formula Φ of product-depth p and size s. Let $\mathbf{r} = (r, r, \ldots, r)$. By Proposition 10, there exists a labelling of the vertices of Φ such that the formula becomes certified multi-$\mathbf{r}$-ic. Let

$$v = \left\lfloor \frac{1}{8r} \left(\frac{d}{2r}\right)^{1/p} \right\rfloor = \left\lfloor \frac{1}{8r} \left(\frac{N}{4}\right)^{1/p} \right\rfloor.$$

By Lemma 20, Φ can be written as a sum of at most s homogeneous $(\mathbf{r}, v, 2)$-form polynomials. By Lemma 14 each one of these polynomials can compute at most $2^N/(2^{v/2})$-many $\mathbf{r}$-extremal monomials. But $P_{N,r}$ has $\binom{N}{N/2}$-many $\mathbf{r}$-extremal monomials. Therefore we must have

$$\begin{aligned} s &\geq \frac{\binom{N}{N/2}}{2^N/(2^{v/2})} \\ &= 2^{\Omega\left(\frac{1}{r} \cdot \left(\frac{N}{4}\right)^{1/p}\right)}. \end{aligned}$$

The moreover part follows from Proposition 19. □

6 A lower bound for depth three multi-r-ic circuits

6.1 The complexity measure.

In this section we prove a lower bound for multi-r-ic depth three circuits by employing some sort of a hybrid of the complexity measures used by Raz [Raz09] and by Nisan and Wigderson [NW96] and recently introduced in [KNS15] called *dimension of skewed partials*. We pick a suitable set of variables $\mathbf{y} \subset \mathbf{x}$, take k-th order derivatives with respect to these variables (for a suitably chosen value of k), then set the $\mathbf{y}$-variables to zero and finally count the dimension of the resulting set of polynomials. In the notation introduced in section 3 our measure is the dimension of the set $\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f\right)$, which we denote by $\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f\right)\right)$ and we sometimes refer to it as the *skewed partials complexity of f* ($\mathrm{SkP}_{\mathbf{y}}$-complexity of f for short).

6.2 Upper bounding the $\mathrm{SkP}_{\mathbf{y}}$-complexity of a depth three circuit.

We first observe that in order to upper bound the $\mathrm{SkP}_{\mathbf{y}}$-complexity of a $\Sigma\Pi\Sigma$-circuit $\mathtt{C}$, it suffices to upper bound the $\mathrm{SkP}_{\mathbf{y}}$-complexity of a single term.

Proposition 25. Sub-additivity of the complexity measure. *For any pair of polynomials $g(\mathbf{x}), h(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and any subset of variables $\mathbf{y} \subseteq \mathbf{x}$ and any integer $k \geq 0$ we have*

$$\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k}(g(\mathbf{x})+h(\mathbf{x}))\right)\right) \leq \dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} g(\mathbf{x})\right)\right) + \dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} h(\mathbf{x})\right)\right).$$

This is easily verified. We next derive an upper bound for the $\mathrm{SkP}_{\mathbf{y}}$-complexity of a term T of a multi-r-ic $\Sigma\Pi\Sigma$-circuit $\mathtt{C}$ by observing that the $\mathrm{SkP}_{\mathbf{y}}$-complexity of T depends only on a relatively small subset of the affine forms in T.

Lemma 26. Upperbound on $\mathrm{SkP}_{\mathbf{y}}$-complexity of a depth three circuit. *Let $\mathbf{x} = \mathbf{y} \uplus \mathbf{z}$ be any partition of the variable set with $|\mathbf{z}| = m$. If $\mathtt{C}$ is a multi-r-ic depth three circuit (i.e $\mathtt{C} = T_1 + T_2 + \ldots + T_s$, where each T_j is a product of affine forms and every variable occurs in at most r affine forms within a given T_j) then*

$$\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k}\mathtt{C}\right)\right) \quad \leq \quad s \cdot \left(\sum_{i=0}^{k} \binom{m \cdot r}{i}\right).$$

Proof. Let $\mathbf{x} = \mathbf{y} \biguplus \mathbf{z}$, where $|\mathbf{z}| = m$. Consider a term T of our circuit $\mathtt{C}$. T is a product of affine forms. Since the term T is assumed to be multi-r-ic, we have that each $\mathbf{z}$-variable appears in at most r affine forms inside T. In particular, there are at most $m \cdot r$ affine forms in T which contain a $\mathbf{z}$-variable. So let

$$T = \ell_1(\mathbf{y}, \mathbf{z}) \cdot \ell_2(\mathbf{y}, \mathbf{z}) \cdot \ldots \ell_t(\mathbf{y}, \mathbf{z}) \cdot Q(\mathbf{y}), \tag{6}$$

where $t \leq m \cdot r$, each $\ell_i(\mathbf{y}, \mathbf{z})$ is an affine form that depends on some $\mathbf{z}$-variable and $Q(\mathbf{y})$ is the product of all the affine forms in T which depend only on the $\mathbf{y}$-variables. To prove the lemma it suffices (via Proposition 25) to show that

$$\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} T\right)\right) \leq \sum_{i=0}^{k} \binom{t}{i}.$$

Suffices to show that

$$\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} T\right) \subseteq \mathbb{F}\text{-span}\left\{\prod_{i \in S} \sigma_{\mathbf{y}}(\ell_i(\mathbf{y}, \mathbf{z})) \;:\; S \subset [t], |S| \geq (t-k)\right\}.$$

Since $\sigma_{\mathbf{y}} : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ is a $\mathbb{F}$-linear map, this is in turn implied by

$$\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} T\right) \subseteq \mathbb{F}\text{-span}\left\{\left\{\prod_{i \in S} \ell_i(\mathbf{y}, \mathbf{z}) \;:\; S \subset [t], |S| \geq (t-k)\right\} \cdot \mathbb{F}[\mathbf{y}]\right\}.$$

This last statement is easily seen by starting from the definition of T given by Equation (6) and computing the appropriate derivatives. $\square$

6.3 A multilinear polynomial with high $\mathrm{SkP}_{\mathbf{y}}$-complexity.

So our problem boils down to finding an explicit polynomial $f(\mathbf{x})$ such that $\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f(\mathbf{x})\right)\right)$ is *large*. Additionally, it is desirable that f should be as easy to compute as possible. We show that there is a polynomial f computed by a small $\Pi\Sigma\Pi$-circuit[25] whose $\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f\right)\right)$-complexity is large.

Lemma 27. An explicit family with high $\mathrm{SkP}_{\mathbf{y}}$-complexity. *There is an explicit family of multilinear polynomials $\{f_{n,k}(\mathbf{x}) \;:\; n, k \geq 0\}$ of degree $d = 3k$ on $N = (n^2 \cdot k + 2nk)$ variables such that there exists a partition $\mathbf{x} = \mathbf{y} \uplus \mathbf{z}$ with $|\mathbf{z}| = m = 2nk$ and $|\mathbf{y}| = N - m = n^2 k$ such that $\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f_{n,k}\right)\right) = n^{2k}$. Moreover, $f_{n,k}$ can be obtained as a restriction of $\mathrm{IMM}_{n,d+2k}$ simply by substituting some subset of variables in $\mathrm{IMM}_{n,d+2k}$ to zero/one values.*

Proof. We first give the description of the family of polynomials $\{f_{n,k}(\mathbf{x}) \;:\; k \geq 0\}$ along with the associated partition of variables $\mathbf{x} = \mathbf{y} \uplus \mathbf{z}$. From the description itself, it will be clear that $f_{n,k}$ has the required degree and number of variables and is computed by a $\Pi^{[k]}\Sigma^{[n^2]}\Pi^{[3]}$-circuit. We first partition our set of $(n^2 k + 2nk)$ variables into two sets

$$\mathbf{x} = \mathbf{y} \uplus \mathbf{z}, \quad \text{where } |\mathbf{z}| = m = 2nk \quad \text{and } |\mathbf{y}| = N - m = n^2 k.$$

The $\mathbf{y}$ and $\mathbf{z}$ are further partitioned into k sets of equal size:

$$\mathbf{y} = \mathbf{y}_1 \uplus \mathbf{y}_2 \uplus \ldots \uplus \mathbf{y}_k, \quad \text{and } \mathbf{z} = \mathbf{z}_1 \uplus \mathbf{z}_2 \uplus \ldots \uplus \mathbf{z}_k$$

[25] This polynomial is a restriction of the iterated matrix multiplication polynomial and is implicitly there in the work of Fournier, Limaye, Malod and Srinivasan [FLMS14]. Vineet Nair pointed out to us that the relevant restriction of IMM as used here is in fact computed by a small $\Pi\Sigma\Pi$ circuit.

where for each $i \in [k]$ we have $|\mathbf{y}_i| = n^2$ and $|\mathbf{z}_i| = 2n$ and the overall structure of $f_{n,k}$ is:

$$f_{n,k}(\mathbf{y}, \mathbf{z}) = g_1(\mathbf{y}_1, \mathbf{z}_1) \cdot g_2(\mathbf{y}_2, \mathbf{z}_2) \cdot \ldots \cdot g_k(\mathbf{y}_k, \mathbf{z}_k)$$

where each $g_i(\mathbf{y}_i, \mathbf{z}_i)$ is a degree three polynomial over the indicated set of variables. For each $i \in [k]$, $\mathbf{y}_i$ will consist of the n^2 variables $\{y_{i,a,b} \ : \ a, b \in [n]\}$ and $\mathbf{z}_i$ consists of the $2n$ variables $\{z_{i,a,b} \ : \ a \in [n], b \in [2]\}$. Finally define

$$g_i(\mathbf{y}_i, \mathbf{z}_i) \stackrel{\text{def}}{=} \sum_{a,b \in [n]} y_{i,a,b} \cdot z_{i,a,1} \cdot z_{i,b,2}.$$

It remains to show that

$$\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f_{n,k}\right)\right) = n^{2k}.$$

To see this note that any k-th order derivative of f with respect to the $\mathbf{y}$-variables is either zero or a monomial in the $\mathbf{z}$-variables. There are $(n^2)^k$ nonzero derivatives of f with respect to the $\mathbf{y}$ variables (obtained by picking exactly one variable from each $\mathbf{y}_i$, $i \in [k]$) and moreover these give rise to distinct monomials in the $\mathbf{z}$-variables and hence are linearly independent as well. Therefore,

$$\dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f_{n,k}\right)\right) = (n^2)^k = n^{2k}, \quad \text{as required.}$$

For the moreover part, first note that from the definition of the family $f_{n,k}$ it follows that it is computed by a multilinear-$\Pi^{[k]}\Sigma^{[n^2]}\Pi^{[3]}$-circuit. So it is a restriction of $\mathrm{IMM}_{n,d+2k}$[26]. $\square$

6.4 Putting things together.

With the above upper and lower bounds in our hand, we are ready to prove a lower bound for multi-r-ic $\Sigma\Pi\Sigma$-circuits.

Theorem 28. *There exists an explicit family $f_{n,k}(\mathbf{x})$ of multilinear polynomials of degree $d = 3k$ on $N = \Theta(n^2 d)$ variables such that $f_{n,k}$ is computed by a $\mathsf{poly}(n,d)$-sized multilinear $\Pi\Sigma\Pi$-circuit*[27] *but any multi-r-ic $\Sigma\Pi\Sigma$-circuit computing $f_{n,k}$ must have top fanin at least*

$$\frac{3}{d}\left(\frac{n}{2e \cdot r}\right)^{\frac{d}{3}}.$$

Proof of Theorem 28. Let $f_{n,k}(\mathbf{y}, \mathbf{z})$ be the polynomial (family) along with the indicated partition of variables as described in Lemma 27. Indeed, from the definition of the family $f_{n,k}$ in the proof of Lemma 27, it is clear that it is computed by a multilinear-$\Pi^{[k]}\Sigma^{[n^2]}\Pi^{[3]}$-circuit.

[26] For more details refer to a subsequent lemma 34.

[27] More precisely, $f_{n,k}$ is computed by a multilinear-$\Pi^{[O(d)]}\Sigma^{[O(N)]}\Pi^{[O(1)]}$-circuit.

Suppose that a multi-r-ic $\Sigma\Pi\Sigma$ circuit $\mathtt{C}$ of top fanin s computes $f_{n,k}$. Then by Lemma 26 and Lemma 27 we have

$$n^{2k} = \dim\left(\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f_{n,k}\right)\right) \leq s \cdot \left(\sum_{i=0}^{k} \binom{(2nk)\cdot r}{i}\right) \leq s \cdot k \cdot \binom{(2nk)\cdot r}{k}.$$

Therefore

$$\begin{aligned} s &\geq \frac{n^{2k}}{k \cdot \binom{(2nk)\cdot r}{k}} \\ &\geq \frac{n^{2k}}{k \cdot \left(\frac{2nkre}{k}\right)^{k}} \\ &\geq \frac{1}{k} \cdot \left(\frac{n}{2re}\right)^{k}. \end{aligned}$$

as required. □

As it was noticed before, the polynomial $f_{n,k}$ is a restriction of the iterated matrix multiplication $\mathrm{IMM}_{n,d+2k}$. So, Theorem 6 directly follows from Theorem 28.

7 Depth four multi-r-ic circuits with low support

As mentioned in the overview, we prove a lower bound for multi-r-ic $\Sigma\Pi\Sigma\Pi$ circuits by first using random restrictions to reduce the number of variables appearing in any monomial and then proving lower bounds against such representations. Let us give names to such polynomials and circuits.

Definition 7. Support size of a polynomial and low-support $\Sigma\Pi\Sigma\Pi$ circuits. *Let $\mathbf{z} \subseteq \mathbf{x}$ be a subset of variables. The* support size *of a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ (resp. the $\mathbf{z}$-support size of f), denoted $|\mathrm{Supp}(f)|$ (resp. $|\mathrm{Supp}_{\mathbf{z}}(f)|$) is the maximum support size (resp. $\mathbf{z}$-support size) of any monomial appearing in f. We will call a depth four circuit $\mathtt{C}$ as a* τ-supported depth four circuit*, denoted as $\Sigma\Pi\Sigma\Pi^{\{\tau\}}$, if it is of the following form:*

$$\mathtt{C} = T_1 + T_2 + \ldots + T_s,$$

where each term T_i is of the form

$$T_i = Q_{i1} \cdot Q_{i2} \cdot \ldots \cdot Q_{it}, \quad \textit{where } \mathrm{Supp}_{\mathbf{z}}(Q_{ij}) \leq \tau \textit{ for all } i \in [s] \textit{ and } j \in [t].$$

7.1 The complexity measure.

Definition 8. Shifted Skewed Partials. *Let $\mathbf{x} = \mathbf{y} \uplus \mathbf{z}$ be a partition of our set of variables into two parts $\mathbf{y}$ and $\mathbf{z}$. For a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, define the dimension of shifted $\mathbf{y}$-partials, $\mathrm{SSP}_{\ell,\mathbf{y},k}(f)$ for short, as follows:*

$$\mathrm{SSP}_{\ell,\mathbf{y},k}(f) \stackrel{\text{def}}{=} \dim\left(\mathbf{z}^{\leq \ell} \cdot \sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} f\right)\right).$$

7.2 Upper bounding the SSP-complexity of a multi-r-ic depth four circuit.

Let $\mathtt{C}$ be a multi-r-ic depth four circuit with $\mathbf{z}$-bottom support bounded by τ. We now give an upper bound on $\mathrm{SSP}_{\ell,\mathbf{y},k}(\mathtt{C})$.

Proposition 29. Sub-additivity of the complexity measure. *For any pair of polynomials $g(\mathbf{x}), h(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and any partition of variables $\mathbf{x} = \mathbf{y} \uplus \mathbf{z}$ and any pair of integers $k, \ell \geq 0$ we have*

$$\mathrm{SSP}_{\ell,\mathbf{y},k}(g(\mathbf{y},\mathbf{z}) + h(\mathbf{y},\mathbf{z})) \leq \mathrm{SSP}_{\ell,\mathbf{y},k}(g(\mathbf{y},\mathbf{z})) + \mathrm{SSP}_{\ell,\mathbf{y},k}(h(\mathbf{y},\mathbf{z})).$$

We now provide the following upper bound on the SSP-complexity of a term of such a circuit.

Lemma 30. Upper bound on the $\mathrm{SSP}_{\ell,\mathbf{y},k}$-complexity of a low support circuit. *Let $\mathtt{C}$ be a τ-supported multi-r-ic $\Sigma\Pi\Sigma\Pi$-circuit and $\mathbf{y} \subseteq \mathbf{x}$ be any subset of variables. If $\mathtt{C}$ has top fanin s then*

$$\mathrm{SSP}_{\ell,\mathbf{y},k}(\mathtt{C}) \leq s \cdot \binom{\frac{2 \cdot m}{\tau}}{k} \cdot \binom{m + \ell + k \cdot r \cdot \tau}{m}, \quad \textit{where } m = |\mathbf{x} \setminus \mathbf{y}|\,.$$

Proof. By the subadditivity of our measure (Proposition 29), it suffices to prove a $\binom{\frac{2 \cdot m}{\tau}}{k} \cdot \binom{m+\ell+k \cdot r \cdot \tau}{m}$ upper bound for a term T which is of the form

$$T(\mathbf{y},\mathbf{z}) = Q_1(\mathbf{y},\mathbf{z}) \cdot Q_2(\mathbf{y},\mathbf{z}) \cdot \ldots \cdot Q_D(\mathbf{y},\mathbf{z}) \cdot R(\mathbf{y}), \tag{7}$$

where each $Q_i(\mathbf{y},\mathbf{z})$ is a polynomial of $\mathbf{z}$-support at most τ, $R(\mathbf{y})$ is an arbitrary polynomial over the indicated set of variables and T is multi-r-ic. Note that since each Q_i has $\mathbf{z}$-support at most τ and is multi-r-ic, the $\mathbf{z}$-degree of each Q_i is at most $r \cdot \tau$. We first do a preprocessing part in which we focus on the $\mathbf{z}$-degree of the Q_i's and combine those with low $\mathbf{z}$-degree. Specifically, if any two Q_i's have $\mathbf{z}$-degree less than $\frac{r \cdot \tau}{2}$ we multiply them together to obtain a new combined factor having $\mathbf{z}$-degree at most $r\tau$. After this preprocessing, we can assume without loss of generality that the $\mathbf{z}$-degree of every Q_i is between $\frac{r\tau}{2}$ and $r\tau$[28]. Now, since the polynomial T is multi-r-ic, its $\mathbf{z}$-degree is at most rm. Hence

$$\begin{aligned} m \cdot r &\geq \deg_{\mathbf{z}}(T) \\ &= \sum_{i \in [D]} \deg_{\mathbf{z}}(Q_i) \\ &\geq \sum_{i \in [D]} \frac{r\tau}{2} \\ &= D \cdot \frac{r\tau}{2}. \end{aligned} \tag{8}$$

[28]In fact, at the end, it could remain a factor of degree lower than $(r\tau/2)$. But either it can be merged with one of the previous factors (without exceeding the degree $r\tau$), or it implies that the average degree is at least $(r\tau/2)$. In all cases the inequality (8) still holds.

This yields the following upper bound on the number of factors D in T:

$$D \leq \frac{2m}{\tau}. \tag{9}$$

Now it suffices to show that

$$\mathbf{z}^{\leq \ell} \cdot \sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} T\right) \subseteq \mathbb{F}\text{-span}\left\{\bigcup_{S \in \binom{[D]}{D-k}} \left(\prod_{i \in S} \sigma_{\mathbf{y}}(Q_i)\right) \cdot \mathbf{z}^{\leq(\ell + kr\tau)}\right\}.$$

In turn, it suffices to show that

$$\sigma_{\mathbf{y}}\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} T\right) \subseteq \mathbb{F}\text{-span}\left\{\bigcup_{S \in \binom{[D]}{D-k}} \left(\prod_{i \in S} \sigma_{\mathbf{y}}(Q_i)\right) \cdot \mathbf{z}^{\leq(kr\tau)}\right\}$$

which in turn follows from

$$\left(\boldsymbol{\partial}_{\mathbf{y}}^{=k} T\right) \subseteq \mathbb{F}\text{-span}\left\{\bigcup_{S \in \binom{[D]}{D-k}} \left(\prod_{i \in S} Q_i(\mathbf{y}, \mathbf{z})\right) \cdot \mathbf{z}^{\leq(kr\tau)} \cdot \mathbb{F}[\mathbf{y}]\right\}. \tag{10}$$

We show this last containment via induction on k. For $k = 0$, the containment in (10) follows from the definition of T (Equation (7)) itself. For the inductive step, suppose that $g(\mathbf{y}, \mathbf{z}) \in \boldsymbol{\partial}_{\mathbf{y}}^{=k} T$ is a k-th order derivative of T. By the inductive assumption we have that $g(\mathbf{y}, \mathbf{z})$ can be expressed as a $\mathbb{F}$-linear combination of polynomials of the form

$$h = \left(\prod_{i \in S} Q_i(\mathbf{y}, \mathbf{z})\right) \cdot h_1(\mathbf{z}) \cdot h_2(\mathbf{y}), \quad \text{where } S \in \binom{[D]}{(D-k)} \text{ and } h_1(\mathbf{z}) \in \mathbf{z}^{\leq(k \cdot r \cdot \tau)}. \tag{11}$$

Let $R = \left(\prod_{i \in S} Q_i(\mathbf{y}, \mathbf{z})\right)$ and $y \in \mathbf{y}$ be a variable. Differentiating (11) with respect to y we have

$$\begin{aligned}
\frac{\partial h}{\partial y} &= \sum_{j\in S}\left(\frac{R}{Q_j}\right)\cdot\left(\frac{\partial Q_j}{\partial y}\right)\cdot h_1(\mathbf{z})\cdot h_2(\mathbf{y}) + \frac{R(\mathbf{y},\mathbf{z})}{Q_1(\mathbf{y},\mathbf{z})}\cdot Q_1(\mathbf{y},\mathbf{z})\cdot h_1(\mathbf{z})\cdot\frac{\partial h_2(\mathbf{y})}{\partial y}\\
&\in \mathbb{F}\text{-span}\left\{\left(\bigcup_{j\in S}\left(\frac{R}{Q_j}\right)\cdot\left(\frac{\partial Q_j}{\partial y}\right)\cdot\mathbf{z}^{\leq(kr\tau)}\cdot\mathbb{F}[\mathbf{y}]\right)\cup\frac{R(\mathbf{y},\mathbf{z})}{Q_1(\mathbf{y},\mathbf{z})}\cdot Q_1(\mathbf{y},\mathbf{z})\cdot\mathbf{z}^{\leq(kr\tau)}\cdot\mathbb{F}[\mathbf{y}]\right\}\\
&\subseteq \mathbb{F}\text{-span}\left\{\left(\bigcup_{j\in S}\left(\frac{R}{Q_j}\right)\cdot\mathbf{z}^{\leq(r\tau)}\cdot\mathbb{F}[\mathbf{y}]\cdot\mathbf{z}^{\leq(kr\tau)}\cdot\mathbb{F}[\mathbf{y}]\right)\cup\frac{R(\mathbf{y},\mathbf{z})}{Q_1(\mathbf{y},\mathbf{z})}\cdot\mathbf{z}^{\leq(r\tau)}\cdot\mathbb{F}[\mathbf{y}]\cdot\mathbf{z}^{\leq(kr\tau)}\cdot\mathbb{F}[\mathbf{y}]\right\}\\
&\subseteq \mathbb{F}\text{-span}\left\{\left(\bigcup_{j\in S}\left(\frac{R}{Q_j}\right)\cdot\mathbf{z}^{\leq((k+1)r\tau)}\cdot\mathbb{F}[\mathbf{y}]\right)\cup\frac{R(\mathbf{y},\mathbf{z})}{Q_1(\mathbf{y},\mathbf{z})}\cdot\mathbf{z}^{\leq((k+1)r\tau)}\cdot\mathbb{F}[\mathbf{y}]\right\}\\
&\subseteq \mathbb{F}\text{-span}\left\{\bigcup_{A\in\binom{S}{(D-k-1)}}\left(\prod_{j\in A}Q_j\right)\cdot\mathbf{z}^{\leq((k+1)r\tau)}\cdot\mathbb{F}[\mathbf{y}]\right\}\\
&\subseteq \mathbb{F}\text{-span}\left\{\bigcup_{A\in\binom{[D]}{(D-(k+1))}}\left(\prod_{j\in A}Q_j\right)\cdot\mathbf{z}^{\leq((k+1)r\tau)}\cdot\mathbb{F}[\mathbf{y}]\right\}.
\end{aligned}$$

In the third step above we use the fact that Q_j has $\mathbf{z}$-degree at most $r\tau$ and hence for all j we have that Q_j as well as $\frac{\partial Q_j}{\partial y}$ is in $\mathbb{F}$-span$\left\{\mathbf{z}^{\leq(r\tau)}\cdot\mathbb{F}[\mathbf{y}]\right\}$. This completes the induction step for the proof of (10) and hence proves the lemma. $\square$

As an immediate corollary, we obtain an upper bound on the SSP-complexity of a multi-r-ic circuit having bottom support bounded by τ.

7.3 A multilinear polynomial with high SSP-complexity.

The explicit polynomial. Our explicit polynomial is a generalization of the one in section 6 for multi-r-ic depth three circuits.

Lemma 31. *There exists an explicit family of polynomials $F_{n,d,k}(\mathbf{y},\mathbf{z})$ such that for any $\delta \leq 1/5, n \geq 3, d, k, \ell$ we have*

$$\mathrm{SSP}_{\ell,\mathbf{y},k}(F_{n,d,k}) \geq M\cdot\binom{|\mathbf{z}|+\ell}{|\mathbf{z}|} - \frac{M^2}{2}\cdot\binom{|\mathbf{z}|+\ell-(\delta\cdot\alpha\cdot k)}{|\mathbf{z}|},$$

where $M = (\frac{n^{2-\delta}}{2})^k$. Moreover, $F_{n,d,k}$ can be obtained as a restriction of $\mathrm{IMM}_{n,d+2k}$ simply by substituting some subset of variables in $\mathrm{IMM}_{n,d+2k}$ to zero/one values.

We now describe the family $F_{n,d,k}$ of the lemma above. Let $\alpha \stackrel{\text{def}}{=} \frac{d-k}{2k}$. The polynomial $F_{n,d,k}$ is of the form:

$$F_{n,d,k}(\mathbf{y},\mathbf{z}) \stackrel{\text{def}}{=} g_1(\mathbf{y}_1,\mathbf{z}_1)\cdot g_2(\mathbf{y}_2,\mathbf{z}_2)\cdot\ldots\cdot g_k(\mathbf{y}_k,\mathbf{z}_k),$$

where the g_i's are polynomials over the indicated (disjoint) subsets of variables defined as

$$g_i(\mathbf{y}_i, \mathbf{z}_i) \stackrel{\text{def}}{=} \sum_{a,b \in [n]} y_{i,a,b} \cdot \prod_{c \in [\alpha]} z_{i,c,a} \cdot z_{i,c+\alpha,b}$$

so that the overall number of $\mathbf{y}$-variables is $|\mathbf{y}| = (n^2) \cdot (k)$ and the overall number of $\mathbf{z}$-variables is $|\mathbf{z}| = (2\alpha n) \cdot (k)$. The degree of $F_{n,d,k}(\mathbf{y}, \mathbf{z})$ is $d = (2\alpha + 1) \cdot k$. The proof that this family has the properties claimed in lemma 31 is very similar to a corresponding one in the work of Fournier, Limaye, Malod and Srinivasan [FLMS14]. For the sake of completeness, we include a proof in section 10.

7.4 Putting things together

Armed with these upper and lower bounds on SSP-complexity, we are now ready to prove the lower bound for multi-r-ic $\Sigma\Pi\Sigma\Pi^{\{\tau\}}$-circuits.

Theorem 32. *Let $\tau = \tau(d)$ and any $r = r(d)$ be integers such that $r\tau = o(d)$, $r\tau \geq \log n$ and $r^{1.1} = o(n)$. For*[29] *$k = d/(1 + 5000r\tau)$, any multi-r-ic $\Sigma\Pi\Sigma\Pi^{\{\tau\}}$-circuit computing $F_{n,d,k}$ must have top fanin at least $(\frac{n}{r^{1.1}})^{\Omega\left(\frac{d}{r \cdot \tau}\right)}$.*

Proof. **Shorthands and choice of parameters.** Let $m = |\mathbf{z}|$ and $M = \left(\frac{n^{2-\delta}}{2}\right)^k$. Our main parameters are chosen as

$$\alpha = 2500 \cdot r \cdot \tau, \quad \ell = \left\lfloor \frac{m \cdot r \cdot \tau}{\varepsilon \cdot \beta \cdot \ln n} \right\rfloor$$

wherein the underlying constants are further chosen as

$$\varepsilon = 2 \cdot 10^{-4}, \quad \delta = \frac{1}{25}, \quad \beta = 200.$$

Our polynomial $F_{n,d,k}(\mathbf{x})$ is as described above. For the above choice of parameters we have[30]

$$M \cdot \binom{m + \ell}{m} - \frac{1}{2} \cdot M^2 \cdot \binom{m + \ell - (\delta \cdot \alpha \cdot k)}{m} \geq \frac{1}{2} \cdot M \cdot \binom{m + \ell}{m}. \tag{12}$$

From lemma 30 we get that any multi-r-ic $\Sigma\Pi\Sigma\Pi^{\{\tau\}}$-circuit computing $f_d = F_{n,d,k}$ must have

[29] By choosing a new d' (with $d \geq d' \geq d/2$) instead of d, we will always assume that $k = \varepsilon d/(\varepsilon + r\tau)$ is an integer.

[30] This estimate is easily verified via application of the numerical estimates in section 3. In general $\beta \geq \frac{4(2-\delta)}{\delta}$ suffices for ensuring (12). A proof is given in section 10.

top fanin at least

$$\begin{aligned}
s &\geq \frac{\mathrm{SSP}_{\ell,\mathbf{y},k}(f_d(\mathbf{y},\mathbf{z}))}{\binom{\frac{2\cdot m}{\tau}}{k}\cdot\binom{m+\ell+k\cdot r\cdot\tau}{m}} \\
&\geq \frac{M\cdot\binom{m+\ell}{m}-\frac{1}{2}\cdot M^2\cdot\binom{m+\ell-(\delta\cdot\alpha\cdot k)}{m}}{\binom{\frac{2\cdot m}{\tau}}{k}\cdot\binom{m+\ell+k\cdot r\cdot\tau}{m}} \quad \text{(using Lemma 31)} \\
&\geq \frac{\frac{1}{2}\cdot M\cdot\binom{m+\ell}{m}}{\binom{\frac{2\cdot m}{\tau}}{k}\cdot\binom{m+\ell+k\cdot r\cdot\tau}{m}} \quad \text{(using (12))} \\
&\geq \frac{1}{2}\cdot\frac{M}{\left(\frac{e\cdot\frac{2\cdot m}{\tau}}{k}\right)^k}\cdot\frac{(m+\ell)!\cdot m!\cdot(\ell+kr\tau)!}{m!\cdot\ell!\cdot(m+\ell+kr\tau)!} \quad \text{(via binomial estimates from section 3)} \\
&= \frac{1}{2}\cdot\left(\frac{n^{2-\delta}\cdot k\cdot\tau}{4e\cdot m}\right)^k\cdot\frac{(\ell+1)\cdot(\ell+2)\cdot\ldots\cdot(\ell+kr\tau)}{(m+\ell+1)\cdot(m+\ell+2)\cdot\ldots\cdot(m+\ell+kr\tau)} \\
&= \frac{1}{2}\cdot\left(\frac{n^{2-\delta}\cdot k\cdot\tau}{4e\cdot(2\alpha nk)}\right)^k\cdot\frac{1}{(1+\frac{m}{\ell+1})\cdot(1+\frac{m}{\ell+2})\cdot\ldots\cdot(1+\frac{m}{\ell+kr\tau})} \\
&= \frac{1}{2}\cdot\left(\frac{2\varepsilon\cdot n^{1-\delta}}{8e\cdot r}\right)^k\cdot\frac{1}{(1+\frac{m}{\ell+1})\cdot(1+\frac{m}{\ell+2})\cdot\ldots\cdot(1+\frac{m}{\ell+kr\tau})} \\
&\geq \frac{1}{2}\cdot\left(\frac{\varepsilon\cdot n^{1-\delta}}{4e\cdot r}\right)^k\cdot\frac{1}{(1+\frac{m}{\ell+1})^{kr\tau}} \quad \left(\text{as } 1+\frac{m}{\ell+i}\leq 1+\frac{m}{\ell+1}\ \forall i\geq 1\right) \\
&\geq \frac{1}{2}\cdot\left(\frac{\varepsilon\cdot n^{1-\delta}}{4e\cdot r}\right)^k\cdot e^{-\frac{m}{\ell+1}\cdot kr\tau} \quad \text{(via exponential estimates from section 3)} \\
&\geq \frac{1}{2}\cdot\left(\frac{\varepsilon}{4e\cdot r}\cdot n^{(1-\delta-\varepsilon\beta)}\right)^k \\
&= \frac{1}{2}\cdot\left(\frac{1}{2\cdot 10^4\cdot e}\cdot\frac{n^{0.92}}{r}\right)^{\frac{d}{2\alpha+1}} \\
&= \left(\frac{n}{r^{1.1}}\right)^{\Omega\left(\frac{d}{r\tau}\right)}.
\end{aligned}$$

□

7.5 A multi-r-ic polynomial with high SSP-complexity

We also observe here that if our target polynomial is also allowed to be multi-r-ic (so that the total degree of our target polynomial can be $\Theta(rN)$) then we can obtain a significantly improved lower bound - a lower bound that is independent of r (and exponential in $\left(\frac{N}{\tau}\right)$).

Theorem 33. *For any positive integers $r = r(N)$ and $\tau = \tau(N)$ with $\tau \geq 2$, $r\tau \geq \log n$ and $\tau = o(N)$, there exists a (explicit) family $\{F_{N,r}\}$ of multi-r-ic N-variate polynomials such* *that $F_{N,r}$ is computed by a* $\mathsf{poly}(Nr)$*-sized multi-r-ic* $\Pi\Sigma\Pi^{\{760\tau+1\}}$*-circuit but any multi-$r$-ic* $\Sigma\Pi\Sigma\Pi^{\{\tau\}}$*-circuit computing $F_{N,r}$ must have top fanin at least* $2^{\Omega(\frac{N}{\tau})}$.

Proof. Our target polynomial is just the same polynomial as in section 7.3 but wherein the $\mathbf{z}$-variables are raised to the r-th powers. In more detail, for positive integers n, k, r and α define the polynomial $F_{n,k,r,\alpha}$ as follows.

$$F_{n,k,r,\alpha}(\mathbf{y},\mathbf{z}) \stackrel{\text{def}}{=} g_1(\mathbf{y}_1,\mathbf{z}_1)\cdot g_2(\mathbf{y}_2,\mathbf{z}_2)\cdot\ldots\cdot g_k(\mathbf{y}_k,\mathbf{z}_k),$$

where the g_i's are polynomials over the indicated (disjoint) subsets of variables defined as

$$g_i(\mathbf{y}_i,\mathbf{z}_i) \stackrel{\text{def}}{=} \sum_{a,b\in[n]} y_{i,a,b}\cdot\prod_{c\in[\alpha]} z^r_{i,c,a}\cdot z^r_{i,c+\alpha,b}$$

so that the overall number of $\mathbf{y}$-variables is $|\mathbf{y}| = (n^2)\cdot(k)$ and the overall number of $\mathbf{z}$-variables is $|\mathbf{z}| = (2\alpha n)\cdot(k)$. The total number of variables in $F_{n,k,r,\alpha}$ is $N = (n^2+2\alpha n)\cdot k$. The degree of $F_{n,k,r,\alpha}(\mathbf{y},\mathbf{z})$ is $d = (2\alpha r+1)\cdot k$. The choice of parameters is

$$\alpha = \frac{\tau}{2\varepsilon},\quad \ell = \left\lfloor\frac{m\cdot\tau\cdot r}{\varepsilon\beta\ln n}\right\rfloor.$$

The analogous lower bound in this case is that for any $\delta\le 1/5, n\ge 3, d,\ell,k$ we have

$$\mathrm{SSP}_{\ell,\mathbf{y},k}(F_{n,k,r,\alpha}) \ge M\cdot\binom{|\mathbf{z}|+\ell}{|\mathbf{z}|} - \frac{M^2}{2}\cdot\binom{|\mathbf{z}|+\ell-(\delta\cdot\alpha\cdot r\cdot k)}{|\mathbf{z}|}, \tag{13}$$

where $M = (\frac{n^{2-\delta}}{2})^k$. Then for proving (13) is larger or equal to $M\binom{|\mathbf{z}|+\ell}{|\mathbf{z}|}/2$, we need $r\tau\ge\varepsilon\beta\ln n$. Thus taking $r\ge 1$, $\tau\ge 2$ is sufficient for the choice of parameters given below. Combining this with the upper bound for the circuit given by Lemma 30 and proceeding as in the proof of Theorem 32, we finally arrive at a lower bound of

$$s \ge \frac{1}{2}\left(\frac{\varepsilon\cdot n^{(1-\delta-\varepsilon\beta)}}{4e}\right)^k \text{ with } k = \left(\frac{\varepsilon N}{\varepsilon n^2+\tau n}\right).$$

If we now choose

$$\delta = \frac{1}{10},\quad \varepsilon = \frac{1}{760},\quad n = 2^{17}\quad\text{and}\quad \beta = 76,$$

we get

$$s \;=\; 2^{\Omega(\frac{N}{\tau})},\quad\text{as required.}$$

□

By choosing $\tau = 2$ in Theorem 33, it directly implies Theorem 8.

8 Depth four multi-r-ic circuits

In this section we give a lower bound for multi-r-ic $\Sigma\Pi\Sigma\Pi$ circuits computing the iterated multiplication polynomial, $\mathrm{IMM}_{n,d}$. The argument goes like this[31]. It turns out that there is a large set S of restrictions which when applied to IMM yields (an isomorphic copy of) the polynomial $F_{n,d,k}$ from section 7.3. At the same time, any small (multi-r-ic) $\Sigma\Pi\Sigma\Pi$-circuit $\mathtt{C}$ is converted to a small (multi-r-ic) low support $\Sigma\Pi\Sigma\Pi$ circuit under the action of a random restriction from S[32]. Since $F_{n,d,k}$ is hard for multi-r-ic low support depth four circuits (Theorem 32), we deduce that IMM must be hard for multi-r-ic $\Sigma\Pi\Sigma\Pi$ circuits. The next lemma is implicit in [FLMS14], however, to make this paper self-contained, we give a proof of it in section 10.

Lemma 34. (Implicit in [FLMS14]). *Let k, d, n be positive integers with $d = (2\alpha + 1)k$ where α is an integer. Consider $h(\mathbf{x}) := \mathrm{IMM}_{n,d+2k}(\mathbf{x})$[33]. There is a large set S consisting of restrictions $\sigma : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ of size $(n!)^{(d-k)}$ with the following properties:*

1. *For each $\sigma \in S$, $\sigma(h)$ is a polynomial isomorphic to the polynomial $F_{n,d,k}(\mathbf{y}, \mathbf{z})$ from section 7.3.*

2. *For any circuit $\mathtt{C}$, $\sigma(\mathtt{C})$ is a circuit of the same depth as $\mathtt{C}$ itself. Moreover, if $\mathtt{C}$ is multi-r-ic then so is $\sigma(\mathtt{C})$.*

3. *If $\mathtt{C}$ is a depth four circuit of size s then for any $\tau \geq 1$, with probability at least*
$$p = 1 - s \cdot \left(\frac{e}{n}\right)^{\tau} \text{ over the uniform, random choice of } \sigma,$$
$\sigma(\mathtt{C})$ is a depth four circuit with $\mathbf{z}$-bottom support at most τ.

Combined with Theorem 32, this immediately yields the lower bound for multi-r-ic depth four circuits (without any restriction on the bottom support) given in Theorem 3.

Proof of Theorem 3. Let
$$\tau = \left\lfloor \sqrt{\frac{d}{r}} \right\rfloor \quad \text{and still } k = \frac{d}{1 + 5000 r\tau}.$$

We give a proof by contradiction. Suppose if possible that $\mathrm{IMM}_{n,d}(\mathbf{x})$ has a multi-r-ic depth four circuit $\mathtt{C}$ of size
$$s = \left(\frac{n}{r^{1.1}}\right)^{o\left(\sqrt{\frac{d}{r}}\right)}. \tag{14}$$

This means that for k as chosen above, that $\mathrm{IMM}_{n,d+2k}(\mathbf{x})$ also has a multi-r-ic depth four circuit $\mathtt{C}$ of size $s = \left(\frac{n}{r^{1.1}}\right)^{o\left(\sqrt{\frac{d}{r}}\right)}$. So, for n large enough, $s \cdot \left(\frac{e}{n}\right)^{\tau} < 1$. Applying Lemma 34 we

[31] This part of the argument is essentially as in [FLMS14].

[32] with high probability

[33] In fact, $h(\mathbf{x}) := \mathrm{IMM}_{n,d+k+1}(\mathbf{x})$ is sufficient if one uses the original restriction from [FLMS14]. See the remark in the proof given in Appendix.

obtain that there exists[34] a restriction $\sigma : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ such that the following two properties hold:

1. $\sigma(\mathrm{IMM}_{n,d+2k})$ is isomorphic to the polynomial $F_{n,d,k}$ from section 7.3,
2. $\sigma(\mathtt{C})$ is a depth four circuit with $\mathbf{z}$-bottom support at most τ.

Thus $\sigma(\mathtt{C})$ computes (a polynomial isomorphic to) $F_{n,d,k}$. Moreover, as $d \geq \log^2 n$, we have $r\tau \geq \log n$. Hence by Theorem 32, $\sigma(\mathtt{C})$ must have top fanin at least

$$\left(\frac{n}{r^{1.1}}\right)^{\Omega\left(\frac{d}{r\tau}\right)} = \left(\frac{n}{r^{1.1}}\right)^{\Omega\left(\sqrt{\frac{d}{r}}\right)}.$$

But then the size s of $\mathtt{C}$ is at least as large as its top fanin which in turn is at least as large as the top fanin of $\sigma(\mathtt{C})$. Therefore

$$s \geq \left(\frac{n}{r^{1.1}}\right)^{\Omega\left(\sqrt{\frac{d}{r}}\right)},$$

which contradicts the assumption in Equation (14). This proves the theorem. □

8.1 A multi-r-ic polynomial requiring large multi-r-ic depth four circuits.

In this section we exhibit an explicit multi-r-ic polynomial in N variables such that any depth four multi-r-ic circuit computing it must have size at least $2^{\Omega(\sqrt{N})}$.

Proof of Theorem 5. Our explicit polynomial is just a variant of $\mathrm{IMM}_{n,d'+2k}$ obtained by raising some of the variables to the r-th powers. Specifically, we want to let the exponent 1 for the $\mathbf{x}$-variables which will be sent to the $\mathbf{y}$-variables during the restriction phase, and raise all others variables to the power r. More formally, if we define τ, k and α as follow:

$$\tau = \lfloor\sqrt{d'}\rfloor, \quad k = \frac{d}{1+760r\tau} \text{ and } \alpha = 380\tau$$

(we still have $d' + 2k = (2\alpha+3)\cdot k$ and $d = (2.\alpha.r+1).k$ as in Theorem 33),

then let $\mathrm{IMMP}_{n,d',r}(\mathbf{x})$ denote the following polynomial:

$$\mathrm{IMMP}_{n,d',r} \stackrel{\text{def}}{=} \mathrm{IMM}_{n,d'+2k}(x_{l,a,b}^{e_{l,a,b}})$$

where

$$e_{l,a,b} = \begin{cases} 1 & \text{if } l \text{ is of the form } i(2\alpha+3)+\alpha+2 \text{ for } i \text{ integer,} \\ r & \text{otherwise.} \end{cases}$$

[34] Indeed, every restriction from the set S of Lemma 34 satisfies the first property, and any random restriction from S satisfies the second one which high probability (larger than $1 - s\cdot(e/n)^{\tau}$).

From the above definition, it is clear that $\text{IMMP}_{n,d',r}$ is an explicit multi-r-ic polynomial computed by a $\mathsf{poly}(nd'r)$-sized algebraic branching program. Our explicit family of polynomials $G_{N,r}$ is simply $\text{IMMP}_{n,d',r}$ for n being a suitably large constant. Specifically, let $n = 2^{17}$. The number of variables $N = n^2 \cdot (d' + 2k - 2) + 2n$ and hence $d' = \Theta(N)$ as n is constant.

The rest of the proof is very similar to that of Theorem 3 and we sketch it below. Suppose if possible that there is a multi-r-ic depth four circuit $\mathtt{C}$ of size at most $2^{o\left(\sqrt{N}\right)}$ computing $\text{IMMP}_{n,d',r}$. Using the same set of restrictions as in the proof of Theorem 3, we obtain that there exists a restriction $\sigma : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ such that

1. $\sigma(\text{IMMP}_{n,d',r})$ is a polynomial isomorphic to the polynomial $F_{N,r}$ from Theorem 33.
2. $\sigma(\mathtt{C})$ is a multi-r-ic depth four circuit with bottom support at most $\tau = \Theta\left(\sqrt{N}\right)$.

But this means that $\sigma(\mathtt{C})$ is a multi-r-ic depth four circuit of bottom fanin $\tau = \Theta\left(\sqrt{N}\right)$ and top fanin $2^{o\left(\sqrt{N}\right)}$ that computes (a polynomial isomorphic to) $F_{N,r}$, contradicting Theorem 33. Hence any multi-r-ic depth four circuit computing $G_{N,r}$ must have size at least $2^{\Omega\left(\sqrt{N}\right)}$.

□

9 Discussion

One motivation behind this work was to generalize the results of [Raz09, RY09] from $r = 1$ which corresponds to multilinear formulas to arbitrary r. While we do make some progress towards this goal, the original problem(s) which motivated this work remain open:

Open Problem 35. *Prove super-polynomial lower bounds for constant depth multi-r-ic formulas (for some explicit family of polynomials).*

Open Problem 36. *Prove super-polynomial lower bounds for multi-r-ic formulas (for some explicit family of polynomials).*

Acknowledgments

We would like to thank Mitali Bafna and Ameya Velingker who pointed out Sperner's Theorem to us. We would also like to thank Sumant Hegde and Abhijat Sharma for useful discussions.

References

[AS04] Noga Alon and Joel H Spencer. *The probabilistic method.* John Wiley & Sons, 2004.

[AV08] Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *FOCS*, pages 67–75, 2008.

[BCS97] P. Bürgisser, M. Clausen, and A. Shokrollahi. *Algebraic Complexity Theory.* Springer, 1997.

[CM14] Suryajith Chillara and Partha Mukhopadhyay. Depth-4 lower bounds, determinantal complexity: A unified approach. In *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5-8, 2014, Lyon, France*, pages 239–250, 2014.

[dO15] Rafael Mendes de Oliveira. Factors of low individual degree polynomials. In *30th Conference on Computational Complexity, CCC 2015, June 17-19, 2015, Portland, Oregon, USA*, pages 198–216, 2015.

[DSY09] Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-randomness tradeoffs for bounded depth arithmetic circuits. *SIAM J. Comput.*, 39(4):1279–1293, 2009.

[FLMS14] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. In *STOC*, pages 128–135, 2014.

[GKKS13a] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *Foundations of Computer Science (FOCS)*, pages 578–587, 2013.

[GKKS13b] Ankit Gupta, Neeraj Kayal, Pritish Kamath, and Ramprasad Saptharishi. Approaching the chasm at depth four. In *Conference on Computational Complexity (CCC)*, pages 65–73, 2013.

[HY11] Pavel Hrubeš and Amir Yehudayoff. Homogeneous formulas and symmetric polynomials. *Computational Complexity*, 20(3):559–578, 2011.

[Hya79] Laurent Hyafil. On the parallel evaluation of multivariate polynomials. *SIAM J. Comput.*, 8(2):120–123, 1979.

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. Technical report, Electronic Colloquium on Computational Complexity (ECCC), 2012.

[KLSS14] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. In *Foundations of Computer Science (FOCS)*, pages 61–70, 2014.

[KNS15] Neeraj Kayal, Vineet Nair, and Chandan Saha. Separation between read-once oblivious algebraic branching programs (roabps) and multilinear depth three circuits. Technical Report 154, Electronic Colloquium on Computational Complexity (ECCC), 2015.

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012.

[KS14a] Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: it's all about the top fan-in. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 136–145, 2014.

[KS14b] Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*, pages 364–373, 2014.

[KS15a] Neeraj Kayal and Chandan Saha. Lower Bounds for Depth Three Arithmetic Circuits with small bottom fanin. In *Conference on Computational Complexity*, pages 158–208, 2015.

[KS15b] Neeraj Kayal and Chandan Saha. Multi-k-ic depth three circuit lower bound. In *32nd International Symposium on Theoretical Aspects of Computer Science (STACS 2015)*, volume 30, pages 527–539, 2015.

[KS15c] Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. *CoRR*, abs/1507.00177, 2015.

[KS15d] Mrinal Kumar and Shubhangi Saraf. Sums of products of polynomials in few variables : lower bounds and polynomial identity testing. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:71, 2015.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *STOC*, pages 146–153, 2014.

[NW96] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1996.

[Raz09] Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2), 2009.

[Raz10] Ran Raz. Tensor-rank and lower bounds for arithmetic formulas. In *STOC*, pages 659–666, 2010.

[RY09] Ran Raz and Amir Yehudayoff. Lower Bounds and Separations for Constant Depth Multilinear Circuits. *Computational Complexity*, 18(2):171–207, 2009.

[Spe28] Emanuel Sperner. Ein satz ber untermengen einer endlichen menge. *Mathematische Zeitschrift*, 27(1):544–548, 1928.

[SY10a] 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.

[SY10b] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.

[Tav13] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *MFCS*, pages 813–824, 2013.

[Val79] L. G. Valiant. Completeness Classes in Algebra. In *STOC '79: Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 249–261, New York, NY, USA, 1979. ACM Press.

[VSBR83] L.G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. Fast parallel computation of polynomials using few processors. *SIAM Journal on Computing*, 12(4):641–644, 1983.

10 Appendix

10.1 Proof of Proposition 12

Proposition 12 (restated). Let $f(\mathbf{x})$ be a polynomial computed by a certified multi-$\mathbf{d}$-ic formula Φ and let α be any node in Φ having label $\mathbf{d}_\alpha$. Then there exist certified formulas Ψ and Λ such that

$$\widehat{\Phi} = \widehat{\Psi} \cdot \widehat{\Phi_\alpha} + \widehat{\Lambda}$$

where

1. Λ is a certified multi-$\mathbf{d}$-ic formula, Ψ is a certified multi-$(\mathbf{d} - \mathbf{d}_\alpha)$-ic formula, and
2. $\mathrm{Size}(\Phi_\alpha) + \mathrm{Size}(\Lambda) \leq \mathrm{Size}(\Phi)$.
3. If Φ is homogeneous then Ψ, Λ are also homogeneous formulas.
4. If Φ is of product-depth p then Ψ, Λ have product-depth at most p.

Proof. Let $\beta_0 = \alpha, \beta_1, \beta_2, \ldots, \beta_t$ be the sequence of nodes in the (unique) path from α to the root node of Φ. So β_1 is the parent of α, β_2 is the grandparent of α and so on and β_t is the root node of Φ. It suffices to prove (via induction) that for every $i \geq 0$

$$\widehat{\Phi_{\beta_i}} = \widehat{\Psi_i} \cdot \widehat{\Phi_\alpha} + \widehat{\Lambda}_i$$

where

1. Λ_i is a certified multi-$\mathbf{d}_{\beta_i}$-ic formula and Ψ_i is a certified multi-$(\mathbf{d}_{\beta_i} - \mathbf{d}_\alpha)$-ic formula, and
2. $\text{Size}(\Phi_\alpha) + \text{Size}(\Lambda_i) \leq \text{Size}(\Phi_{\beta_i})$.

The base case $i = 0$. This is easy to verify.
The inductive step. Let β_{i+1} have children β_i and $\gamma_1, \ldots, \gamma_q$ with labels $\mathbf{d}_{\beta_i}$ and $\mathbf{d}_{\gamma_1}, \ldots, \mathbf{d}_{\gamma_q}$ respectively. There are two cases. The node β_{i+1} is a $+$ node. Then by definition we have $\mathbf{d}_{\beta_{i+1}} = \mathbf{d}_{\beta_i} = \mathbf{d}_{\gamma_1} = \ldots = \mathbf{d}_{\gamma_q}$. Now define

$$\Psi_{i+1} \stackrel{\text{def}}{=} \Psi_i, \quad \Lambda_{i+1} \stackrel{\text{def}}{=} \Lambda_i + \Phi_{\gamma_1} + \ldots + \Phi_{\gamma_q}.$$

Then Λ_{i+1} is a certified multi-$\mathbf{d}_{\beta_{i+1}}$-ic formula and Ψ_{i+1} is a certified multi-$(\mathbf{d}_{\beta_{i+1}} - \mathbf{d}_\alpha)$-ic formula. The other case of the inductive step is where β_{i+1} is a $\times$ gate. Then by definition we have $\mathbf{d}_{\beta_{i+1}} = \mathbf{d}_{\beta_i} + \mathbf{d}_{\gamma_1} + \ldots + \mathbf{d}_{\gamma_q}$. Now define

$$\Psi_{i+1} \stackrel{\text{def}}{=} \Psi_i \times \Phi_{\gamma_1} \times \ldots \times \Phi_{\gamma_q}, \quad \Lambda_{i+1} \stackrel{\text{def}}{=} \Lambda_i \times \Phi_{\gamma_1} \times \ldots \times \Phi_{\gamma_q},$$

with the natural labelling to the root nodes of Ψ_{i+1} and of Λ_{i+1} that this definition entails. Then Λ_{i+1} is computed by a certified multi-$\mathbf{d}_{\beta_{i+1}}$-ic formula while Ψ_{i+1} is computed by a certified multi-$\mathbf{e}$-ic formula, where

$$\mathbf{e} = (\mathbf{d}_{\beta_i} - \mathbf{d}_\alpha) + \mathbf{d}_{\gamma_1} + \ldots + \mathbf{d}_{\gamma_q} = \mathbf{d}_{\beta_{i+1}} - \mathbf{d}_\alpha.$$

Finally note that in the cases we have:

$$\begin{aligned}
\text{Size}(\Phi_{\beta_{i+1}}) &\geq \text{Size}(\Phi_{\beta_i}) + \text{Size}(\Phi_{\gamma_1}) + \ldots + \text{Size}(\Phi_{\gamma_q}) \\
&\geq (\text{Size}(\Phi_\alpha) + \text{Size}(\Lambda_i)) + \text{Size}(\Phi_{\gamma_1}) + \ldots + \text{Size}(\Phi_{\gamma_q}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{(by inductive hypothesis)} \\
&= \text{Size}(\Phi_\alpha) + (\text{Size}(\Lambda_i) + \text{Size}(\Phi_{\gamma_1}) + \ldots + \text{Size}(\Phi_{\gamma_q}) \\
&= \text{Size}(\Phi_\alpha) + \text{Size}(\Lambda_{i+1}).
\end{aligned}$$

This completes the proof of the inductive step. The last two properties of Ψ, Λ are also easily checked from their definitions. This completes the proof of the proposition.

$\square$

10.2 Proof of Lemma 31

Lemma 31 (restated). For any $\delta \leq 1/5$, $n \geq 3$, d, ℓ, k we have

$$\mathrm{SPP}_{\ell,\mathbf{y},k}(F_{n,d,k}) \geq M \cdot \binom{|\mathbf{z}|+\ell}{|\mathbf{z}|} - \frac{M^2}{2} \cdot \binom{|\mathbf{z}|+\ell-(\delta\cdot\alpha\cdot k)}{|\mathbf{z}|},$$

where $M = (\frac{n^{2-\delta}}{2})^k$.

Proof. For any couple of k-uplets $(\mathbf{a} = (a_1,\ldots,a_k), \mathbf{b} = (b_1,\ldots,b_k))$ in $\left([n]^k\right)^2$, let us define

$$\mathbf{y}_{\mathbf{a},\mathbf{b}} = (y_{1,a_1,b_1},\ldots,y_{k,a_k,b_k})$$

and

$$\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F) \stackrel{\text{def}}{=} \frac{\boldsymbol{\partial}^k F}{\boldsymbol{\partial} \mathbf{y}_{\mathbf{a},\mathbf{b}}} = \prod_{i=1}^{k} \prod_{c\in[\alpha]} z_{i,c,a_i} \cdot z_{i,c+\alpha,b_i}.$$

Notice that $\{\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F)\}$ is a subset of n^{2k} monomials belonging to the set $\sigma_{\mathbf{y}}(\boldsymbol{\partial}_{\mathbf{y}}^{=k}F)$. Hence,

$$\begin{aligned}
\mathrm{SPP}_{\ell,\mathbf{y},k}(F) &= \dim(\mathbf{z}^{\leq \ell}\cdot \sigma_{\mathbf{y}}(\boldsymbol{\partial}_{\mathbf{y}}^{=k}F)) \\
&\geq \dim(\mathbf{z}^{\leq \ell}\cdot \{\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F)\}) \\
&= \left|\mathbf{z}^{\leq \ell}\cdot \{\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F)\}\right|.
\end{aligned} \tag{15}$$

The third step is due to the fact that the dimension of the vectorial space generated by a set of monomials is exactly the cardinal of this set.
In the following, we will consider a subset of $\{\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F)\}$ made of monomials which are pairwise sufficiently far away. For that, let us define some distances. If $\mathbf{u}$ and $\mathbf{v}$ are two k-vectors,

$$\Delta(\mathbf{u},\mathbf{v}) \stackrel{\text{def}}{=} |\{i \mid u_i \neq v_i\}|.$$

And then

$$\Delta\left(\boldsymbol{\partial}_{\mathbf{a}_1,\mathbf{b}_1}(F), \boldsymbol{\partial}_{\mathbf{a}_2,\mathbf{b}_2}(F)\right) \stackrel{\text{def}}{=} \Delta(\mathbf{a}_1,\mathbf{a}_2) + \Delta(\mathbf{b}_1,\mathbf{b}_2).$$

Claim 37. *There exists $\mathcal{P}_{M,\delta}$ a subset of $\{\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F)\}$ of cardinal M such that if $\boldsymbol{\partial}_{\mathbf{a}_1,\mathbf{b}_1}(F)$ and $\boldsymbol{\partial}_{\mathbf{a}_2,\mathbf{b}_2}(F)$ are two distinct elements of $\mathcal{P}_{M,\delta}$, then*

$$\Delta\left(\boldsymbol{\partial}_{\mathbf{a}_1,\mathbf{b}_1}(F), \boldsymbol{\partial}_{\mathbf{a}_2,\mathbf{b}_2}(F)\right) \geq \delta \cdot k.$$

Proof. For any monomial m in $\{\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F)\}$, there are at most $\binom{2k}{\delta k}\cdot n^{\delta k}$ monomials from $\{\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F)\}$ which are at distance at most $\delta \cdot k$ (for the distance Δ). In particular such a $\mathcal{P}_{M,\delta}$ can be obtained by a greedy algorithm as soon as

$$M \cdot \binom{2k}{\delta k} \cdot n^{\delta k} \leq |(\boldsymbol{\partial}_{\mathbf{a},\mathbf{b}}(F))| = n^{2k}.$$

It implies we can choose M as large as

$$\begin{aligned}
\frac{n^{2k}}{\binom{2k}{\delta k} n^{\delta k}} &\geq \frac{n^{(2-\delta)k}}{(\frac{2ek}{\delta k})^{\delta k}} \\
&= \left[\left(\frac{\delta}{2e} \right)^{\delta} \cdot n^{2-\delta} \right]^k \\
&\geq \left(\frac{n^{2-\delta}}{2} \right)^k \quad \text{since } \delta \leq \frac{1}{5}.
\end{aligned}$$

□

Then, with Equation (15),

$$\begin{aligned}
\mathrm{SPP}_{\ell,\mathbf{y},k}(F) &\geq |\mathbf{z}^{\leq \ell} \cdot \mathcal{P}_{M,\delta}| \\
&= \left| \bigcup_{m \in \mathcal{P}_{M,\delta}} \left(\mathbf{z}^{\leq \ell} \cdot m \right) \right| \\
&\geq \sum_{m \in \mathcal{P}_{M,\delta}} |\mathbf{z}^{\leq \ell} \cdot m| - \frac{1}{2} \sum_{m_1 \neq m_2 \in \mathcal{P}_{M,\delta}} |(\mathbf{z}^{\leq \ell} \cdot m_1) \cap (\mathbf{z}^{\leq \ell} \cdot m_2)|. \qquad (16)
\end{aligned}$$

Let us upperbound the cardinal of $|(\mathbf{z}^{\leq \ell} \cdot m_1) \cap (\mathbf{z}^{\leq \ell} \cdot m_2)|$ for any $m_1 \neq m_2$. For any $\tilde{m}$ in $|(\mathbf{z}^{\leq \ell} \cdot m_1) \cap (\mathbf{z}^{\leq \ell} \cdot m_2)|$, we have $\tilde{m} = m_1 \cdot \tilde{m}_1$ where $\tilde{m}_1$ is a $\mathbf{z}$-monomial of degree at most ℓ. As $\Delta(m_1, m_2) \geq \delta \cdot k$, it implies there are at least $\delta k \alpha$ variables $\{t_1, \ldots, t_{\delta k \alpha}\}$ which appear in m_2 and not in m_1. So, these variables have to appear in $\tilde{m}_1$. In particular, $\tilde{m} = m_1 \cdot t_1 \cdot \ldots \cdot t_{\delta k \alpha} \cdot \tilde{m}_2$ where $\tilde{m}_2$ is a $\mathbf{z}$-monomial of degree at most $\ell - (\delta \alpha k)$. Consequently, for any pair of distinct monomials m_1, m_2 of $\mathcal{P}_{M,\delta}$,

$$|(\mathbf{z}^{\leq \ell} \cdot m_1) \cap (\mathbf{z}^{\leq \ell} \cdot m_2)| \leq \binom{|\mathbf{z}| + \ell - (\delta \alpha k)}{|\mathbf{z}|}.$$

Plugging this bound in Equation (16) directly implies the lemma. □

10.3 Equation (12) in the proof of Theorem 32

We want to constraint β to ensure

$$\begin{aligned}
&M \cdot \binom{m+\ell}{m} - \frac{1}{2} \cdot M^2 \cdot \binom{m + \ell - (\delta \cdot \alpha \cdot k)}{m} \geq \frac{1}{2} \cdot M \cdot \binom{m+\ell}{m} \\
\text{i.e.,} \quad &M \cdot \binom{m + \ell - \delta \alpha k}{m} \leq \binom{m+\ell}{m}.
\end{aligned}$$

Let us recall that

$$\frac{m\alpha}{\ell} \geq \frac{\beta}{2}\ln(n)$$
$$\text{and} \quad M = \left(\frac{n^{2-\delta}}{2}\right)^k.$$

We have

$$\begin{aligned}
M\frac{\binom{m+\ell-(\delta\alpha k)}{m}}{\binom{m+\ell}{m}} &= M\frac{(m+\ell-\delta\alpha k)!\ell!}{(\ell-\delta\alpha k)!(m+\ell)!} \\
&= M\frac{(\ell-\delta\alpha k+1)\cdot\ldots\cdot(\ell)}{(m+\ell-\delta\alpha k+1)\cdot\ldots\cdot(m+\ell)} \\
&\leq M\left(\frac{\ell}{m+\ell}\right)^{\delta\alpha k} \\
&= M\left(1-\frac{m}{m+\ell}\right)^{\delta\alpha k} \\
&\leq M\cdot e^{-\frac{m\delta\alpha k}{m+\ell}} \\
&\leq \left(\frac{n^{2-\delta}}{2}\right)^k \cdot e^{-\frac{m\delta\alpha k}{2\ell}} \qquad (17)\\
&\leq \left(2^{(2-\delta)\log(n)-\frac{\delta\beta\log(n)}{4}}\right)^k
\end{aligned}$$

which is lower or equal than 1 as soon as

$$2-\delta-\frac{\delta\beta}{4} \leq 0$$
$$\text{i.e.,} \quad \beta \geq \frac{4(2-\delta)}{\delta}.$$

For the inequaliy (17), we need $\ell \geq m$ which is true if $r\tau \geq \log n$, or more accurately if $r\tau \geq \varepsilon\beta\ln n$.

10.4 Proof of Lemma 34

We reprove here a result which is implicit in the paper [FLMS14].

Lemma 34 (restated). Let k, d, n be positive integers with $d = (2\alpha+1)k$ where α is an integer. Consider $h(\mathbf{x}) := \mathrm{IMM}_{n,d+2k}(\mathbf{x})$. There is a large set S consisting of restrictions $\sigma : \mathbb{F}[\mathbf{x}] \mapsto \mathbb{F}[\mathbf{x}]$ of size $(n!)^{(d-k)}$ with the following properties:

1. For each $\sigma \in S$, $\sigma(h)$ is a polynomial isomorphic to the polynomial $F_{n,d,k}$ from section 7.3.

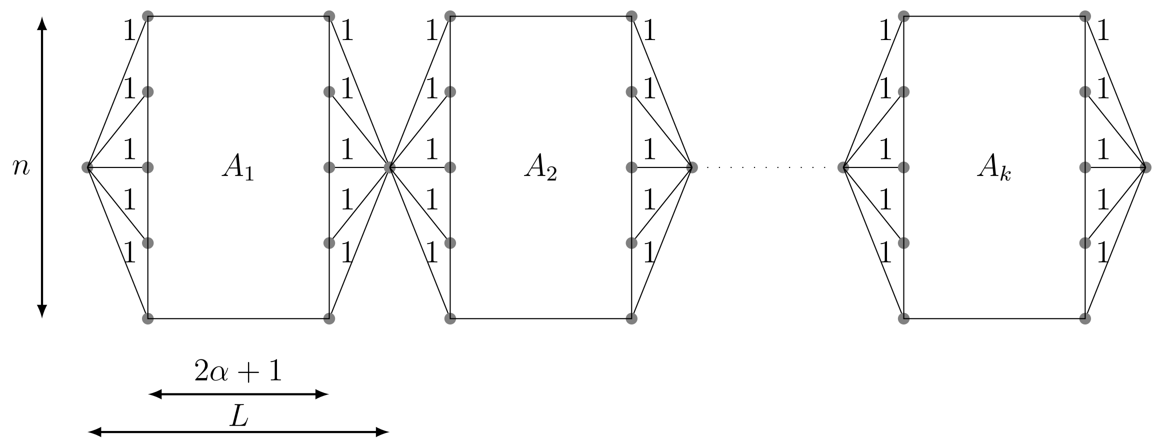
2. For any circuit $\mathtt{C}$, $\sigma(\mathtt{C})$ is a circuit of the same depth as $\mathtt{C}$ itself. Moreover, if $\mathtt{C}$ is multi-r-ic then so is $\sigma(\mathtt{C})$.

3. If $\mathtt{C}$ is a depth four circuit of size s then for any $\tau \geq 1$, with probability at least

$$p = 1 - s \cdot \left(\frac{e}{n}\right)^{\tau} \text{ over the random choice of } \sigma,$$

   $\sigma(\mathtt{C})$ is a depth four circuit with $\mathbf{z}$-bottom support at most τ.

Proof. Let $\mathfrak{S}_n$ be the set of the permutations of $[n]$ and let $L = 2 + d/k = 2\alpha + 3$.
Let first define the set of the restrictions. In fact, we will give in the same time a renaming of the variables which highlights the isomorphism to $F_{n,d,k}$. For any sequence of n-permutations $\boldsymbol{\pi} = (\pi_{1,1}, \ldots, \pi_{k,2\alpha}) \in (\mathfrak{S}_n)^{k \cdot 2\alpha}$ we will define the restriction $\sigma_{\boldsymbol{\pi}}$. We want each restriction $\sigma_{\boldsymbol{\pi}}(h)$ be a product of k polynomials.



More formally, for any matrix M_l with

- $l \equiv 1 \mod L$, we set

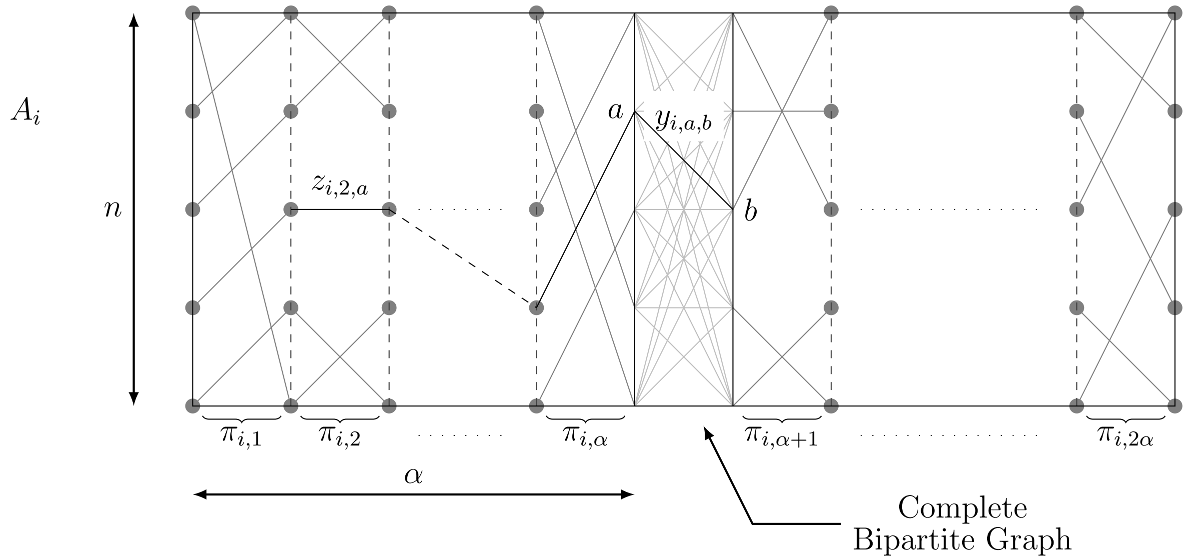
$$x_{l,a,b} = \begin{cases} 1 \text{ if } a = 1 \\ 0 \text{ otherwise,} \end{cases}$$

- $l \equiv 0 \mod L$, we set

$$x_{l,a,b} = \begin{cases} 1 \text{ if } b = 1 \\ 0 \text{ otherwise.} \end{cases}$$

Remark. *In the proof from [FLMS14], we can notice that the groups of two consecutive constant matrices (last matrix from a block and first matrix of the next block) are merged together. We choose this presentation since it highlights the fact the restriction is a product of length k.*

Now, we want to define the restriction for the other matrices.



More formally, for any matrix M_l with

- $l = (i-1)L + \alpha + 2$, we set $x_{l,a,b} = y_{i,a,b}$,
- $l = (i-1)L + q + 1$ with $1 \leq q \leq \alpha$, we set

$$x_{l,a,b} = \begin{cases} 0 \text{ if } \pi_{i,q}(a) \neq b \\ z_{i,q,c} \text{ otherwise where } c = \pi_{i,\alpha} \circ \pi_{i,\alpha-1} \circ \ldots \circ \pi_{i,q+1}(b), \end{cases}$$

- $l = (i-1)L + q + 2$ with $\alpha + 1 \leq q \leq 2\alpha$, we set

$$x_{l,a,b} = \begin{cases} 0 \text{ if } \pi_{i,q}(a) \neq b \\ z_{i,q,c} \text{ otherwise where } a = \pi_{i,q-1} \circ \pi_{i,q-2} \circ \ldots \circ \pi_{i,\alpha+1}(c). \end{cases}$$

The set of these restrictions will be denoted S. Then, for any $\sigma \in S$, the polynomial $\sigma(h)$ (where the renaming of the variables is done as above) equals $F_{n,d,k}$, which proves the point 1). The size of S is $(n!)^{2\alpha k} = (n!)^{d-k}$ as required.
The property 2) is immediate since

1. for any $x \in \mathbf{x}$, for any $\sigma \in S$, $\sigma(x)$ is a constant (in $\{0,1\}$) or a variable (in $\mathbf{y} \uplus \mathbf{z}$)
2. and if $\sigma(x_1) = \sigma(x_2) \in \mathbf{y} \uplus \mathbf{z}$ then $x_1 = x_2$.

Finally, let us prove the point 3). We uniformly randomly pick a sequence $\boldsymbol{\pi}$ of n-permutations. Let A and B be two subsets of $\mathfrak{S}_n$. So, for $i \neq j$ the events $\pi_i \in A$ and $\pi_j \in B$ are independent. Let $\mathbf{x}' \subseteq \mathbf{x}$ be the subset of the variables which can be transformed into a $\mathbf{z}$-variable. More formally,

$$\mathbf{x}' = \{x_{i,a,b} \mid \exists p, q \in \mathbb{N} \text{ s.t. } i = pL + q,\ p < k \text{ and } (2 \leq q \leq \alpha + 1 \text{ or } \alpha + 3 \leq q \leq 2\alpha + 2)\}.$$

For simplicity, we will also re-index the variables in $\mathbf{x}' = (x_{\underline{i,q},a,b})_{1\leq i\leq k, 1\leq q\leq 2\alpha}$ by skipping the matrices which do not contain $\mathbf{x}'$-variables.

$$x_{\underline{i,q},a,b} = \begin{cases} x_{(i-1)L+q+1,a,b} \text{ if } q \leq \alpha \\ x_{(i-1)L+q+2,a,b} \text{ otherwise.} \end{cases}$$

Let t be a monomial computed at the first level of the circuit $\mathtt{C}$ of $\mathbf{x}'$-support larger than τ. Let us define

$$C_{i,q} = \{(a,b) \mid x_{\underline{i,q},a,b} \text{ is a variable of } t\}.$$

In particular, we have

$$\sum_{i,q} |C_{i,q}| \geq \tau. \tag{18}$$

The restrictions $\sigma_{\boldsymbol{\pi}}(t)$ will be sent to 0 as soon as one of the $\mathbf{x}'$-variable $x_{\underline{i,q},a,b}$ of t is sent to 0, i.e. a soon as there exists a $\mathbf{x}'$-variable $x_{\underline{i,q},a,b}$ in t such that $\pi_{i,q}(a) \neq b$.
Hence,

$$\begin{aligned} \mathbb{P}_{\boldsymbol{\pi}}\left[\sigma_{\boldsymbol{\pi}}(t) \neq 0\right] &= \mathbb{P}_{\boldsymbol{\pi}}\left[\pi_{i,q}(a) = b \mid \forall x_{\underline{i,q},a,b} \in (\mathbf{x}' \cap t)\right] \\ &= \prod_{i=1}^{k} \prod_{q=1}^{2\alpha} \mathbb{P}_{\pi_{i,q}\in\mathfrak{S}_n}\left[\pi_{i,q}(a) = b \mid \forall (a,b) \in C_{i,q}\right]. \end{aligned} \tag{19}$$

The last equality holds since the events are independent.
Let us fix i and q, we will prove

Claim 38.

$$P_{i,q} \stackrel{\text{def}}{=} \mathbb{P}_{\pi_{i,q}\in\mathfrak{S}_n}\left[\pi_{i,q}(a) = b \mid \forall (a,b) \in C_{i,q}\right] \leq \left(\frac{e}{n}\right)^{|C_{i,q}|}. \tag{20}$$

Proof. In the proof, c will denote the cardinal $|C_{i,q}|$. The inequality is true if $c = 0$. Moreover, if (a, b_1) and (a, b_2) are in $C_{i,q}$ with $b_1 \neq b_2$, then $\pi_{i,q}(a) \neq b_1$ or $\pi_{i,q}(a) \neq b_2$ and so $P_{i,q} = 0$. Hence, we can assume that $1 \leq c \leq n$. Let $C_{i,q} = \{(a_1, b_1), \ldots, (a_c, b_c)\}$. The number of n-permutations π such that for all j, $\pi(a_j) = b_j$ is exactly the number of $(n-c)$-permutations, i.e. $(n-c)!$. Then, using Stirling's approximation[35]

$$\begin{aligned}
P_{i,q} &= \frac{(n-c)!}{(n)!} \\
&\leq \left(\frac{n-c}{e}\right)^{n-c} \left(\frac{e}{n}\right)^{n} \sqrt{\frac{n-c}{n}} \frac{e}{\sqrt{2\pi}} \\
&\leq \left(\frac{e}{n}\right)^{c} \left(1-\frac{c}{n}\right)^{n-c+\frac{1}{2}} \frac{e}{\sqrt{2\pi}} \\
&\leq \left(\frac{e}{n}\right)^{c} e^{-c+\frac{c^2}{n}-\frac{c}{2n}} \frac{e}{\sqrt{2\pi}} \\
&\leq \left(\frac{e}{n}\right)^{c} e^{-n+\frac{n^2}{n}-\frac{n}{2n}} \frac{e}{\sqrt{2\pi}} \\
&= \left(\frac{e}{n}\right)^{c} \sqrt{\frac{e}{2\pi}} \\
&\leq \left(\frac{e}{n}\right)^{c}
\end{aligned}$$

where the fifth step comes from the fact that the function $(-x+\frac{x^2}{n}-\frac{x}{2n})$ with $1 \leq x \leq n$ is maximized when $x = n$. $\square$

Finally, using Equations (18), (19) and (20)

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{\pi}}\left[\sigma_{\boldsymbol{\pi}}(t) \neq 0\right] &\leq \prod_{i=1}^{k} \prod_{q=1}^{2\alpha} \left(\frac{e}{n}\right)^{|C_{i,q}|} \\
&\leq \left(\frac{e}{n}\right)^{\tau}.
\end{aligned}$$

So, the probability a term with $\mathbf{x}'$-support larger than τ is still present in the restriction is at most $(e/n)^{\tau}$. By hypothesis, there are at most s terms, so the union bound implies that the probability that the restriction of the circuit does not contain any term with $\mathbf{z}$-support larger than τ is at least $1 - s \cdot (e/n)^{\tau}$. $\square$

[35] $\left(\frac{n}{e}\right)^{n} \sqrt{2\pi n} \leq n! \leq \left(\frac{n}{e}\right)^{n} e\sqrt{n}$

ECCC ISSN 1433-8092

http://eccc.hpi-web.de