Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 33 (2020)

New Exponential Size Lower Bounds against Depth Four Circuits of Bounded Individual Degree

Suryajith Chillara,
CRI, University of Haifa, Israel.
suryajith@cmi.ac.in

Abstract

Kayal, Saha and Tavenas [Theory of Computing, 2018] showed that for all large enough integers n and d such that $d \geqslant \omega(\log n)$, any *syntactic* depth four circuit of bounded individual degree $\delta = o(d)$ that computes the Iterated Matrix Multiplication polynomial ($\mathsf{IMM}_{n,d}$) must have size $n^{\Omega\left(\sqrt{d/\delta}\right)}$. Unfortunately, this bound deteriorates as the value of δ increases. Further, the bound is superpolynomial only when δ is $o(d)$. It is natural to ask if the dependence on δ in the bound could be weakened. Towards this, in an earlier result [STACS, 2020], we showed that for all large enough integers n and d such that $d = \Theta(\log^2 n)$, any *syntactic* depth four circuit of bounded individual degree $\delta \leqslant n^{0.2}$ that computes $\mathsf{IMM}_{n,d}$ must have size $n^{\Omega(\log n)}$.

In this paper, we make further progress by proving that for all large enough integers n and d, and absolute constants a and b such that $\omega(\log^2 n) \leqslant d \leqslant n^a$, any *syntactic* depth four circuit of bounded individual degree $\delta \leqslant n^b$ that computes $\mathsf{IMM}_{n,d}$ must have size $n^{\Omega(\sqrt{d})}$. Our bound is obtained by carefully adapting the proof of Kumar and Saraf [SIAM J. Computing, 2017] to the complexity measure introduced in our earlier work [STACS, 2020].

1 Introduction

Arithmetic circuits are directed acyclic graphs such that the leaf nodes are labeled by variables or constants from the underlying field, and every non-leaf node is labeled either by a $+$ or $\times$. Every node computes a polynomial by operating on its inputs with the operand given by its label. The flow of computation flows from the leaf to the output node. We refer the readers to the standard resources [SY10, Sap19] for more information on arithmetic formulas and arithmetic circuits.

One of the central questions in Algebraic Complexity Theory is to show super polynomial size lower bounds against *hard* polynomials. It is interesting to note that most polynomials of interest are multilinear. However, in the last few decades, we could not make substantial progress towards proving lower bounds for multilinear polynomials against general arithmetic circuits (cf. [SY10, Sap19]). It then is a natural strategy to prove lower bounds against restricted class of circuits. To begin with, we can restrict the polynomial computed at every gate to be multilinear. Under such a restriction, researchers made a lot of progress [NW97, Raz06, Raz04, RY08, RY09, HY11, RSY08, AKV18, CLS19, CELS18, CLS18]. Backed by this progress, we can now try to relax this restriction to circuits where the polynomial computed at every gate is of bounded individual degree.

Kayal and Saha [KS17a] first studied multi-δ-ic circuits of depth three and proved exponential lower bounds. Kayal, Saha and Tavenas [KST18] have extended this and proved exponential lower bounds at depth three and depth four, and superpolynomial lower bounds for homogeneous formulas. These circuits (formulas) that were considered were syntactically multi-δ-ic . That is, at any product node, any variable appears in the support of at most δ many operands, and the total of the individual formal degrees is also at most δ. Henceforth, all the multi-δ-ic depth four circuits that we talk about shall be syntactically multi-δ-ic .

ISSN 1433-8092

Recently, Kumar, Oliviera and Saptharishi [KdOS19] showed that proving lower bounds of the order $\exp(O(\sqrt{\delta \cdot N \log N}))$ for N variate polynomials against depth four multi-δ-ic circuits is sufficient to show superpolynomial lower bounds against general multi-δ-ic circuits. This provides us further motivation to study multi-δ-ic circuits of depth four.

Raz and Yehudayoff [RY09] showed a lower bound of $\exp(\Omega\left(\sqrt{d \log d}\right))$ against multilinear depth four circuits which compute a multilinear polynomial over N variables and degree $d \ll N$ (cf. [KST18, Footnote 9]). Kayal, Saha and Tavenas [KST18] have shown a lower bound of $\left(\frac{n}{\delta^{1.1}}\right)^{\Omega\left(\sqrt{\frac{d}{\delta}}\right)}$ for Iterated Matrix Multiplication polynomial over d many $n \times n$ matrices (denoted $\mathsf{IMM}_{n,d}$), that is computed by a multi-δ-ic depth four circuit. This lower bound deteriorates with the increasing value of δ and is superpolynomial only when δ is $o(d)$ and is strictly less than $n^{1.1}$. This raises a natural question if the dependence on the individual degree could be improved upon. Towards this, in the same paper, they showed a lower bound of $2^{\Omega(\sqrt{N})}$ for a N-variate polynomial that is not multilinear. Hegde and Saha [HS17] proved a bound of $2^{\Omega\left(\sqrt{\frac{N \log N}{\delta}}\right)}$ against a N-variate multilinear design polynomial which is in VNP, for a wider range of δ than [KST18].

In an earlier result [Chi20], we showed that for all $\delta < n^{0.2}$ and for a specific regime of degree $d = \Theta(\log^2 n)$, any multi-δ-ic depth four circuit computing $\mathsf{IMM}_{n,d}$ must have size $n^{\Omega(\log n)}$. Even though the bound holds for a very small range of degrees and degrees much smaller, this quasipolynomial bound holds true even for individual degree much larger than the syntactic degree of the polynomial.

In this paper, we extend [Chi20] and show that for all large enough integers n and d, and absolute constants a and b such that $b > a$ and $\omega(\log^2 n) \leqslant d \leqslant n^a$, any *syntactic* depth four circuit of bounded individual degree $\delta \leqslant n^b$ that computes $\mathsf{IMM}_{n,d}$ must have size $n^{\Omega(\sqrt{d})}$.

To prove our result, we use the dimension of Projected Shifted Skew Partial Derivatives as our complexity measure. This was introduced in [Chi20] and is an extension of the dimension of Shifted Skew Partial Derivatives, introduced by Kayal et al. [KST18]. This extension was inspired by the definition of Projected Shifted Partial Derivatives measure of Kayal, Limaye, Saha and Srinivasan [KLSS17]. The polynomial of interest is multilinear and it makes sense to project down the space of Shifted Skew Partial Derivatives to just those Shifted Skew Partial Derivatives which are multilinear.

For a polynomial P defined over the variable set $Y \sqcup Z$, the dimension of Projected Shifted Skew Partial Derivatives with respect to the parameters r' and m can informally be defined as follows.

$$\dim\left(\mathbb{F}\text{-span}\left\{\text{mult}\left(\mathbf{z}^{=m} \cdot \sigma_Y\left(\partial_Y^{=r'} P\right)\right)\right\}\right)$$

Here, $\mathbf{z}^{=m}$ denotes the set of monomials over Z-variables of degree m, $\sigma_Y(f)$ sets all the Y variable appearances in f to 0 and $\text{mult}(g)$ sets the coefficients of all the nonmultilinear monomials in g to 0. Using linear algebra and setting an order on the variables, counting the dimension of the aforementioned quantity can be reduced to counting leading monomials in the space $\left(\mathbb{F}\text{-span}\left\{\text{mult}\left(\mathbf{z}^{=m} \cdot \sigma_Y\left(\partial_Y^{=r'} P\right)\right)\right\}\right)$.

Comparison to [KST18, HS17]: Even though Kayal et al.[KST18] and us in this paper, have proved lower bounds for $\mathsf{IMM}_{n,d}$, when compared to Kayal et al. [KST18] our bound holds for a restricted range of d. In [KST18], they proved a lower bound for all $d \geqslant \log^2 n$. In contrast, our bound holds only for degrees smaller than n^a for an absolute constant a. On the other hand, for all $\omega(\log^2 n) \leqslant d \leqslant n^a$, we show better quantitative bounds than [KST18]. In this regime of d, we show a bound that does not deteriorate with the increase in individual degree of the depth four

circuit. Further our bound holds for a range of δ that is greater than d as well, which was not the case in either [KST18].

Hegde and Saha [HS17] proved an exponential lower bound for a N-variate polynomial of degree $\Theta(N)$ which is in VNP. Our result is mostly incomparable against this.

Comparison to [Chi20]: In [Chi20], we defined a polynomial Q_n which is a p-projection of $\mathsf{IMM}_{n,d}$ where $d = \Theta(\log^2 n)$. We then used random restrictions ρ on the variable set and reduced Q_n to a polynomial P_ρ. At the same time we showed that with a high probability, the depth four multi-δ-ic circuit under random restrictions would reduce to a depth four multi-δ-ic circuit such that every monomial at the bottom product gate is of low support. Then using the dimension of Projected Shifted Skew Partial Derivatives as our complexity measure, we showed that any multi-δ-ic depth circuit of low bottom support that computes P_ρ must have large size. Through union bound, we inferred that any multi-δ-ic depth four circuit computing Q_n must be large. Since Q_n was a p-projection of $\mathsf{IMM}_{n,\Theta(\log^2 n)}$, we get that any multi-$\delta$-ic depth four circuit computing $\mathsf{IMM}_{n,\Theta(\log^2 n)}$ must be large.

An important point to note in [Chi20], all the Skew Partial Derivatives were multilinear monomials. Thus, a leading monomial is given by the projected shift of the only multilinear monomial. This analysis fell short of giving bounds that were anything better than quasipolynomial. To overcome this hurdle, in this paper we ensure that the Skew Partial Derivatives are in fact multilinear polynomials over a large set of monomials. This puts us in a similar situation as in [KS17b]. The key observation in [KS17b] (cf. [Sap19, Section 20.3]) was that if β is a leading monomial in $\sigma_Y(\partial_\alpha(P))$, then for some γ, $\gamma \cdot \beta$ need not be a leading monomial in $\mathrm{mult}(\gamma \cdot \sigma_Y(\partial_\alpha(P)))$.

Similarity to [KS17b]: As discussed above, we extend the random restrictions and the careful analysis of Kumar and Saraf [KS17b] and adapt it to our complexity measure. We work with a different set of parameters than [KS17b] and thus, we build on their already *careful* analysis and refine it in certain places to suit our needs.

In this paper, we made a conscious decision to use a notation which is as close to that of [KS17b] as possible to help those readers who are already acquainted with the proof of [KS17b].

Proof overview:

The proof strategy is very similar to the previous work [FLMS15, KS17b, KST18, HS17, Chi20]. We first pick a random restriction V of the variables from a carefully crafted distribution D. We then show that with a high probability, a multi-δ-ic depth four circuit under such a restriction reduces to a multi-δ-ic depth four circuit of low bottom Z-support. Let $\Phi|_V$ be the circuit obtained after the restriction $V \leftarrow D$. Let $T_1, T_2, \ldots, T_s$ be the terms corresponding to the product gates feeding into the output sum gate of $\Phi|_V$. The output polynomial is obtained by adding the terms $T_1, T_2, \ldots, T_s$. Note that each of these T_i's is a product of low Z-support polynomials $Q_{i,j}$, that is, every monomial in these $Q_{i,j}$'s is supported on a small set of Z variables (say t many). One major observation at this point is to see that there can be at most $|Z| \cdot r$ many factors in any of the T_i's with non-zero Z-support.

From [Chi20], we get that the dimension of Projected Shifted Skew Partial Derivatives for *small* multi-δ-ic depth four circuits of low bottom Z-support is *not too large*. We then show that the Projected Skew Partial Derivative space of a polynomial $\mathsf{IMM}_{n,d}|_V$ is *large* using a Leading Monomial approach (cf. [Sap19, Section 20.3]). Thus, we infer that if $\Phi|_V$ were to compute $\mathsf{IMM}_{n,d}|_V$, then

$\Phi|_V$ cannot be *small*. Then we lift this argument to show that if Φ were to compute $\mathsf{IMM}_{n,d}$, then Φ cannot be *small*.

2 Preliminaries

Notation:

- For a polynomial f, we use $\partial^{=r'}_{Y}(f)$ to refer to the space of partial derivatives of order r' of f with respect to monomials of degree r' in Y.
- We use $\mathbf{z}^{=m}$ and $\mathbf{z}^{\leqslant m}$ to refer to the set of all the monomials of degree equal to m and at most m, respectively, in Z variables.
- We use $\mathbf{z}^{\leqslant m}_{\mathrm{ML}}$ to refer to the set of all the multilinear monomials of degree at most m in Z variables.
- We use $\mathbf{z}^{\leqslant m}_{\mathrm{NonML}}$ to refer to the set of all the non-multilinear monomials of degree at most m in Z variables.
- For a monomial m we use $|\mathrm{MonSupp}(m)|$ to refer to the size of the set of variables that appear in it.
- For a polynomial f, we use $|\mathrm{MonSupp}(f)|$ to refer to the maximum $|\mathrm{MonSupp}(m)|$ over all monomials in f.

Depth four circuits: A depth four circuit (denoted by $\Sigma\Pi\Sigma\Pi$) over a field $\mathbb{F}$ and variables $\{x_1, x_2, \ldots, x_n\}$ computes polynomials which can be expressed in the form of sums of products of polynomials. That is, $\sum_{i=1}^{s}\prod_{j=1}^{d_i} Q_{i,j}(x_1, \ldots, x_n)$ for some d_i's. A depth four circuit is said to have a bottom support of t (denoted by $\Sigma\Pi\Sigma\Pi^{\{t\}}$) if it is a depth four circuit and all the monomials in each polynomial $Q_{i,j}$ are supported on at most t variables.

Multi-δ-ic arithmetic circuits: Let $\overline{\delta} = (\delta_1, \delta_2, \cdots, \delta_N)$. An arithmetic circuit Φ is said to be a syntactically multi-$\overline{\delta}$-ic circuit if for all product gates $u \in \Phi$ and $u = u_1 \times u_2 \times \cdots \times u_t$, each variable x_i can appear in at most δ_i many of the u_i's ($i \in [t]$). Further the total formal degree with respect to every variable over all the polynomials computed at $u_1, u_2, \cdots, u_t$, is bounded by δ, i.e. $\sum_{j\in[t]} \deg_{x_i}(f_{u_j}) \leqslant \delta_i$ for all $i \in [N]$. If $\overline{\delta} = (\delta, \delta, \cdots, \delta)$, then we simply refer to them as multi-δ-ic circuits.

Complexity Measure: Let the variable set X be partitioned into two fixed, disjoint sets Y and Z such that $|Y|$ is a magnitude larger than $|Z|$. Let $\sigma_Y : \mathbb{F}[Y \sqcup Z] \mapsto \mathbb{F}[Z]$ be a map such that for any polynomial $f(Y, Z)$, $\sigma_Y(f) \in \mathbb{F}[Z]$ is obtained by setting every variable in Y to zero by it and leaving Z variables untouched. Let $\mathrm{mult} : \mathbb{F}[Z] \mapsto \mathbb{F}[Z]$ be a map such that for any polynomial $f(Y, Z)$, $\mathrm{mult}(f) \in \mathbb{F}[Z]$ is obtained by setting the coeficients of all the non-multilinear monomials in f to 0 and leaving the rest untouched. We use $\mathbf{z}^{\leqslant m} \cdot \sigma_Y(\partial^{=r'}_{Y} f)$ to refer to the linear span of polynomials obtained by multiplying each polynomial in $\sigma_Y(\partial^{=r'}_{Y} f)$ with monomials of degree at most m in Z

variables. We will now define our complexity measure, Dimension of Projected Shifted Skew Partial Derivatives (denoted by $\mathrm{PSSPD}_{r',m}$) as follows.

$$\mathrm{PSSPD}_{r',m}(f) = \dim\left(\mathbb{F}\text{-span}\left\{\mathrm{mult}\left(\mathbf{z}^{=m}\cdot\sigma_Y\left(\partial_Y^{=r'}f\right)\right)\right\}\right)$$

This is a natural generalization of Shifted Skew Partial Derivatives of [KST18]. The following proposition is easy to verify.

Proposition 1 (Sub-additivity). *Let r' and m be integers. Let the polynomials f, f_1, f_2 be such that $f = f_1 + f_2$. Then, $\mathrm{PSSPD}_{r',m}(f) \leqslant \mathrm{PSSPD}_{r',m}(f_1) + \mathrm{PSSPD}_{r',m}(f_2)$.*

Monomial Distance: We recall the following definition of distance between monomials from [CM19].

Definition 2 (Definition 2.7, [CM19]). *Let m_1, m_2 be two monomials over a set of variables. Let S_1 and S_2 be the multisets of variables corresponding to the monomials m_1 and m_2 respectively. The distance $\mathrm{dist}(m_1, m_2)$ between the monomials m_1 and m_2 is the $\min\{|S_1| - |S_1 \cap S_2|, |S_2| - |S_1 \cap S_2|\}$ where the cardinalities are the order of the multisets.*

For example, let $m_1 = x_1^2x_2x_3^2x_4$ and $m_2 = x_1x_2^2x_3x_5x_6$. Then $S_1 = \{x_1, x_1, x_2, x_3, x_3, x_4\}$, $S_2 = \{x_1, x_2, x_2, x_3, x_5, x_6\}$, $|S_1| = 6$, $|S_2| = 6$ and $\mathrm{dist}(m_1, m_2) = 3$. It is important to note that two distinct monomials could have distance 0 between them if one of them is a multiple of the other and hence the triangle inequality does not hold.

Polynomial Family: Let $X^{(1)}, X^{(2)}, \ldots, X^{(d)}$ be d generic $n \times n$ matrices defined over disjoint set of variables. For any $k \in [d]$, let $x_{i,j}^{(k)}$ be the variable in the matrix $X^{(k)}$ indexed by $(i, j) \in [n] \times [n]$. The Iterated Matrix Multiplication polynomial, denoted by the family $\{\mathsf{IMM}_{n,d}\}$, is defined as follows.

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

The following lemma (from [GKKS14]) is key to the asymptotic estimates required for the lower bound analyses.

Lemma 3 (Lemma 6, [GKKS14]). *Let $a(n), f(n), g(n) : \mathbb{Z}_{\geqslant 0} \to \mathbb{Z}_{\geqslant 0}$ be integer valued functions such that $(f + g) = o(a)$. Then,*

$$\ln\frac{(a+f)!}{(a-g)!} = (f+g)\ln a \pm O\left(\frac{(f+g)^2}{a}\right)$$

As in [KS17b], we also use the following strengthening of the Principle of Inclusion and Exclusion in our proof.

Lemma 4 (Strong Inclusion-Exclusion [KS17b]). *Let $W_1, W_2, \cdots, W_t$ be subsets of a finite set W. For a parameter $\lambda \geqslant 1$, let $\sum_{\substack{i,j\in[t]\\ i\neq j}} |W_i \cap W_j| \leqslant \lambda \sum_{i\in[t]} |W_i|$. Then, $\left|\cup_{i\in[t]} W_i\right| \geqslant \frac{1}{4\lambda}\sum_{i\in[t]}|W_i|$.*

We also need the following form of generalized Hamming bound [GRS19, Section 1.7].

Lemma 5. *For every $\Delta_0 < 2r'$, there exists a subset $\mathscr{P}_{\Delta_0} \subset [n]^{2r'}$ of size $\frac{n^{2r'-\Delta_0/2}}{\frac{\Delta_0}{2}\cdot\binom{2r'}{0.5\Delta_0}}$ such that for all $\mathbf{a}, \mathbf{a}' \in \mathscr{P}_{\Delta_0}$, $\mathrm{dist}(\mathbf{a}, \mathbf{a}') \geqslant \Delta_0$.*

Leading Monomial Approach of [KS17b]

Let P be any polynomial defined over the sets of variables $Y \sqcup Z$, of degree d'. Let $\mathscr{M}$ be a set of suitably chosen monomials over Y variables. Let σ_Y applied to a polynomial set all its Y variables to zero. For any monomial $\alpha \in \mathscr{M}$ let $M(\alpha)$ be the set of monomials in the support of the polynomial $\sigma_Y(\partial_\alpha(P))$. Note that all the monomials in $M(\alpha)$ are supported only on Z variables. For any $\beta \in M(\alpha)$, let

$$A_m(\alpha,\beta) = \left\{\gamma' \mid \mathrm{Supp}(\gamma') = \deg(\gamma') = d' - r' + m \text{ and } \exists \gamma \,:\, \gamma' = \mathsf{LM}\left(\gamma \cdot \sigma_Y(\partial_\alpha(P))\right) = \gamma \cdot \beta\right\}.$$

Proposition 6 (Equation 20.13, [Sap19]). *Let m, r' be integers. Then,*

$$\mathrm{PSSPD}_{r',m}(P) \geqslant \left| \bigcup_{\substack{\alpha \in \mathscr{M} \\ \beta \in M(\alpha)}} A_m(\alpha,\beta) \right|$$

Using the Inclusion-Exclusion principle we get the following.

$$\left| \bigcup_{\substack{\alpha \in \mathscr{M} \\ \beta \in M(\alpha)}} A_m(\alpha,\beta) \right| \geqslant \sum_{\substack{\alpha \in \mathscr{M} \\ \beta \in M(\alpha)}} |A_m(\alpha,\beta)| - \sum_{\substack{\alpha,\alpha' \in \mathscr{M} \\ \beta \in M(\alpha) \\ \gamma \in M(\alpha') \\ (\alpha,\beta) \neq (\alpha',\gamma)}} \left|A_m(\alpha,\beta) \cap A_m(\alpha',\gamma)\right|.$$

For any $\beta \in M(\alpha)$, let

$$S_m(\alpha,\beta) = \{\gamma \mid \deg(\gamma) = \mathrm{Supp}(\gamma) = m \text{ and } |\mathrm{Supp}(\gamma) \cap \mathrm{Supp}(\beta)| = \emptyset\}.$$

Lemma 7 (Lemma 20.17, [Sap19]). *For all $\alpha \in \mathscr{M}$,*

$$\sum_{\beta \in M(\alpha)} |A_m(\alpha,\beta)| \geqslant \left| \bigcup_{\beta \in M(\alpha)} S_m(\alpha,\beta) \right|.$$

Again, by using the Inclusion-Exclusion principle, we get that

$$\left| \bigcup_{\beta \in M(\alpha)} S_m(\alpha,\beta) \right| \geqslant \sum_{\beta \in M(\alpha)} |S_m(\alpha,\beta)| - \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} |S_m(\alpha,\beta) \cap S_m(\alpha,\gamma)|.$$

For a polynomial P and monomials $\alpha, \alpha' \in \mathscr{M}$, let $T_1(P,\alpha)$, $T_2(P,\alpha)$ and $T_3(P,\alpha,\alpha')$ be defined as follows.

$$T_1(P,\alpha) = \sum_{\beta \in M(\alpha)} |S_m(\alpha,\beta)|$$

$$T_2(P,\alpha) = \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} |S_m(\alpha,\beta) \cap S_m(\alpha,\gamma)|$$

$$\text{and} \quad T_3(P,\alpha,\alpha') = \sum_{\substack{\alpha,\alpha' \in \mathscr{M} \\ \beta \in M(\alpha) \\ \gamma \in M(\alpha') \\ (\alpha,\beta) \neq (\alpha',\gamma)}} \left|A_m(\alpha,\beta) \cap A_m(\alpha',\gamma)\right|.$$

Further, let $T_1(P), T_2(P)$ and $T_3(P)$ be defined as follows.

$$T_1(P) = \sum_{\alpha \in \mathscr{M}} T_1(P, \alpha)$$
$$T_2(P) = \sum_{\alpha \in \mathscr{M}} T_2(P, \alpha)$$
$$\text{and} \quad T_3(P) = \sum_{\alpha, \alpha' \in \mathscr{M}} T_3(P, \alpha, \alpha').$$

Thus we get that $\mathrm{PSSPD}_{r',m}(P) \geqslant T_1(P) - T_2(P) - T_3(P)$.

Choice of parameters:

- Set $\varepsilon' = 0.340$.
- Set $\eta = 0.05$ and thus $\varepsilon = \varepsilon' - \eta = 0.29$.
- Let $r = 2\tau r'$ and we set $\tau = 0.08$.
- Let c_0 be a constant that is at most 0.016.
- Let us set t to $2 \times 10^{-5} k'$.
- Since $(2k' + 3)r' = d$, we can set k' and r' to be of the same order and thus are $\approx O(\sqrt{d})$. Let $2c_0 r' \leqslant t$.
- Let $k = d - 3r' = 2k'r'$.
- Let $d \leqslant n^{0.01}$.
- We need to ensure that $\nu(1 - \tau) \geqslant 2\varepsilon$ and $1 - \nu - \varepsilon' > 0$. Thus, we set $\nu = 0.631$.
- We want $n^{\eta} \cdot 2^{k' - \varepsilon' \log n - 1} = \frac{2^{k'}}{2 \cdot n^{\varepsilon}} = \left(\frac{N}{N-m}\right)^{k'}$. Let us set m to $\frac{N}{2}(1 - \Gamma)$ such that $\Gamma = O_{\varepsilon}\left(\frac{\ln n}{k'}\right)$ and the constant hidden in the order is determined by the equation $(1 + \Gamma)^{k'} = 2n^{\varepsilon}$.

3 Deterministic and Random Restrictions

Let the d matrices be divided into r' contiguous blocks of matrices $B_1, B_2, \ldots, B_{r'}$ such that each block B_i contains $2k' + 3$ matrices and $d = (2k' + 3) \cdot r'$. By suitable renaming, let us assume that each block B_i contains the following matrices.

$$X^{(i,L,k'+1)}, \cdots, X^{(i,L,2)}, X^{(i,L,1)}, X^{(i)}, X^{(i,R,1)}, X^{(i,R,2)}, \cdots, X^{(i,R,k'+1)}.$$

Let us first consider the following set of restrictions, first deterministic and then randomized.

Deterministic Restrictions

Let $V_0 : X \mapsto Y_0 \sqcup Z_0 \sqcup \{0, 1\}$ be a deterministic restriction of the variables X in to disjoint variable sets Y_0, Z_0, and $\{0, 1\}$ as follows. For all $i \in [r']$,

- The variables in matrix in $X^{(i)}$ are each set to a distinct Y_0 variable. Henceforth, we shall refer to this as $Y^{(i)}$ matrix.
- The entries of the first row of matrix $X^{(i,L,k'+1)}$ are all set to 1 and the rest of the matrix to 0.
- The entries of the first column of matrix $X^{(i,R,k'+1)}$ are all set to 1 and the rest of the matrix to 0.
- The rest of the variables are all set to distinct Z_0 variables. Henceforth, for all $b \in \{L, R\}$ and $j \in [k']$, we shall refer to the matrix $X^{(i,b,j)}$ as $Z^{(i,b,j)}$ matrix.

Random Restrictions

Let $V_1 : Y_0 \sqcup Z_0 \mapsto Y \sqcup Z \sqcup \{0, 1\}$ be a random restriction of the variables $Y_0 \sqcup Z_0$ as follows.

- Matrix $Z^{(i,L,1)}$: For every column, pick n^η distinct elements uniformly at random and keep these elements alive. Set the other entries in this matrix to zero.
- Matrix $Z^{(i,R,1)}$: For every row, pick n^η distinct elements uniformly at random and keep these elements alive. Set the other entries in this matrix to zero.
- Matrices $Z^{(i,L,j)}$ for all $j \in [2, k' - \varepsilon' \log n]$: For every column, pick 2 distinct elements uniformly at random and set all the other entries to zero.
- Matrices $Z^{(i,R,j)}$ for all $j \in [2, k' - \varepsilon' \log n]$: For every row, pick 2 distinct elements uniformly at random and set all the other entries to zero.
- Matrices $Z^{(i,L,j)}$ for all $j > k' - \varepsilon' \log n$: For every column, pick 1 element uniformly at random and set the other elements in that row to zero.
- Matrices $Z^{(i,R,j)}$ for all $j > k' - \varepsilon' \log n$: For every row, pick 1 element uniformly at random and set the other elements in that row to zero.

Let D be the distribution of all the restrictions $V : X \mapsto Y \sqcup Z \sqcup \{0, 1\}$ such that $V = V_1 \circ V_0$ where V_0 and V_1 are deterministic and random restrictions respectively, as described above. Let N be used to denote the number of Z variables left after the restriction and $N = 2r'n(n^\eta + 2(k - \varepsilon' \log n - 1) + \varepsilon' \log n)$.

Effect of restrictions on Depth Four Multi-δ-ic Circuits

Lemma 8. *Let Φ be a multi-δ-ic depth four circuit of size at most $s \leqslant n^{\frac{t}{2}}$ that computes $\mathsf{IMM}_{n,d}$. Then with a probability of at least $1 - o(1)$, over $V \leftarrow D$, $\Phi|_V$ is a multi-δ-ic depth four circuit of bottom support at most t in Z variables.*

Proof. Let m be any monomial that is computed at the bottom layer of Φ. It is easy to see that $V \leftarrow D$ sets variables across matrices independently and also variables across the rows (columns) in each right (left) matrix independently. That is, the probability of setting two variables is independent unless they are from the same row in a matrix from the left side of the block. Let us first compute the probability that a monomial survives if all the variables in its support come from a single matrix. A monomial m survives of Z-support t survives if and only if the random restriction keeps these variables alive.

Matrices $Z^{(i,R,1)}$: Let these $t_1, \cdots, t_n$ be the distribution of variables across n rows such that $t = t_1 + \cdots + t_n$.

$$\mathbb{P}_{V \leftarrow D}[m|_V \neq 0] \leqslant \prod_{i \in [n]} \frac{\binom{n-t_i}{n^\eta - t_i}}{\binom{n}{n^\eta}} = \prod_{i \in [n]} \left(\frac{n^\eta}{n}\right)^{t_i} = n^{-(1-\eta)t}.$$

Matrices $Z^{(i,L,1)}$: Let these $t_1, \cdots, t_n$ be the distribution of variables across n rows such that $t = t_1 + \cdots + t_n$.

$$\mathbb{P}_{V \leftarrow D}[m|_V \neq 0] \leqslant \prod_{i \in [n]} \frac{\binom{n-t_i}{n^\eta - t_i}}{\binom{n}{n^\eta}} = \prod_{i \in [n]} \left(\frac{n^\eta}{n}\right)^{t_i} = n^{-(1-\eta)t}.$$

Matrices $Z^{(i,R,j)}$ for all $j \in [2, k' - \varepsilon' \log n]$: Let these $t_1, \cdots, t_n$ be the distribution of variables across n rows such that $t = t_1 + \cdots + t_n$.

$$\mathbb{P}_{V \leftarrow D}[m|_V \neq 0] \leqslant \prod_{i \in [n]} \left(\frac{2}{n}\right)^{t_i} = \left(\frac{2}{n}\right)^t.$$

Matrices $Z^{(i,L,j)}$ for all $j \in [2, k' - \varepsilon' \log n]$: Let these $t_1, \cdots, t_n$ be the distribution of variables across n columns such that $t = t_1 + \cdots + t_n$.

$$\mathbb{P}_{V \leftarrow D}[m|_V \neq 0] \leqslant \prod_{i \in [n]} \left(\frac{2}{n}\right)^{t_i} = \left(\frac{2}{n}\right)^t.$$

Matrix $Z^{(i,R,j)}$ for all $j > k' - \varepsilon' \log n$: Let these $t_1, \cdots, t_n$ be the distribution of variables across n rows such that $t = t_1 + \cdots + t_n$.

$$\mathbb{P}_{V \leftarrow D}[m|_V \neq 0] \leqslant \prod_{i \in [n]} \left(\frac{1}{n}\right)^{t_i} = \left(\frac{1}{n}\right)^t.$$

Matrix $Z^{(i,L,j)}$ for all $j > k' - \varepsilon' \log n$: Let these $t_1, \cdots, t_n$ be the distribution of variables across n columns such that $t = t_1 + \cdots + t_n$.

$$\mathbb{P}_{V \leftarrow D}[m|_V \neq 0] \leqslant \prod_{i \in [n]} \left(\frac{1}{n}\right)^{t_i} = \left(\frac{1}{n}\right)^t.$$

Thus, we can now infer that any monomial of support t stays alive with a probability of at most p where

$$p = \max\left\{n^{-(1-\eta)t}, n^{-t}, \left(\frac{2}{n}\right)^t\right\}.$$

By taking a union bound, we get that with a probability of at least $1 - s \cdot p = 1 - o(1)$, $\Phi|_V$ is a circuit of bottom Z-support at most t. □

We recall the following lemma from [Chi20]. We provide the proof of this lemma in Appendix A for the sake of completeness.

Lemma 9 (Section 3.1, [Chi20]). *Let N, r', m and t be positive integers such that $m + r't < \frac{N}{2}$. Let Ψ be a processed syntactic multi-δ-ic depth four circuit of size s and every monomial computed at the bottom product gate has Z-support of at most t. Then, $\mathrm{PSSPD}_{r',m}(\Psi)$ is at most $s \cdot \binom{\frac{2N\delta}{t}}{r'} \cdot \binom{N}{m+r't} \cdot (m + r't)$.*

Effect of Restrictions on $\mathsf{IMM}_{n,d}$

Let $g_{1,a}^{(i,L)}$ be the $(1,a)$th entry in product of matrices $\prod_{j=0}^{k'} X^{(i,L,k'+1-j)}|_V$. Let $g_{b,1}^{(i,R)}$ be the $(b,1)$th entry in product of matrices $\prod_{j=1}^{k'+1} X^{(i,R,j)}|_V$. Let $g^{(i)}$ the $(1,1)$th entry in the product of all the matrices in the block B_i. Then we can express $g^{(i)}$ as follows.

$$g^{(i)} = \sum_{a,b \in [n]} g_{1,a}^{(i,L)} \cdot y_{a,b}^{(i)} \cdot g_{b,1}^{(i,R)}.$$

Let $P|_V$ obtained by restricting $\mathsf{IMM}_{n,d}$ with the restriction $V \leftarrow D$. Thus,

$$P|_V = \prod_{i=1}^{r'} g^{(i)}.$$

Thus, we can say that $g_{1,a}^{(i,L)}$ is the $(a,1)$th entry in the product $\left(\prod_{j=0}^{k'} X^{(i,L,k'+1-j)}|_V\right)^T = \prod_{j=1}^{k'+1} \left(X^{(i,L,j)}|_V\right)^T$. Putting this together with the structure of random restrictions, we get the following observation.

Observation 10. *Structures of polynomials $g_{1,a}^{(i,L)}$ and $g_{a,1}^{(i,R)}$ are similar.*

Henceforth, without loss of generality, we will now consider that our polynomial $P|_V$ is composed of $2r'$ many Z-blocks, with suitable renaming.

$$P|_V = \prod_{i=1}^{2r'} h_i$$

4 Lower Bound Against Multi-δ-ic Depth Four Circuit

Let $P|_V$ be the polynomial obtained by restricting $\mathsf{IMM}_{n,d}$ with the restriction $V \leftarrow D$. For all $\alpha \in \mathscr{M}$, let $M(\alpha)$ be the set of monomials in the support of the polynomial $\sigma_Y(\partial_\alpha(P|_V))$. From the construction, it is easy to see that the cardinality of the set $M(\alpha)$ is the same for all $\alpha \in \mathscr{M}$. Let L_2 denote the cardinality of $M(\alpha)$. Recall that $L_2 = (n^\eta \cdot 2^{k'-\varepsilon' \log n - 1})^{r'}$.

Observation 11. *It is important to note that $\frac{\partial^{=r'}(P|_V)}{y^{(1)}_{(a_1,a_2)} \cdot y^{(2)}_{(a_3,a_4)} \cdot \ldots \cdot y^{(r')}_{(a_{2r'-1},a_{2r'})}} = \prod_{i \in [r']} g_{1,a_{2i-1}}^{(i,L)} \cdot g_{a_{2i},1}^{(i,R)}$, for any choice of $\mathbf{a} \in [n]^{2r'}$ is a multilinear polynomial over just the Z variables.*

Let $\Delta_0 \leqslant 2r' - r$ where $r = 2\tau r'$. From Lemma 5, we get that there is a set $\mathscr{P}_{\Delta_0} \subseteq [n]^{2r'}$ of size $\frac{n^{2r'-0.5\Delta_0}}{2^{O(r')}}$ such that for any pair of vectors $\alpha, \alpha' \in \mathscr{P}_{\Delta_0}$, $\left|\{i \mid \alpha_i = \alpha'_i\}\right|$ is at most r.

Let the carefully chosen set of monomials $\mathscr{M}$ over Y variables be defined as follows.

$$\mathscr{M} = \left\{ \prod_{i\in[r']} y^{(i)}_{(a_{2i-1},a_{2i})} : \alpha = (a_1, a_2, \ldots, a_{2r'}) \in \mathscr{P}_{\Delta_0} \right\}.$$

From the definition, we can infer the following.

Observation 12. *For any $\alpha \neq \alpha' \in \mathscr{M}$, $\partial_\alpha^{=r'}(P|_V)$ and $\partial_{\alpha'}^{=r'}(P|_V)$ are distinct polynomials.*

We shall henceforth use L_1 to denote the cardinality of $\mathscr{M}$ and $L_1 \geqslant \frac{n^{(1+\tau)r'}}{2^{O(r')}}$.

With respect to the polynomial $P|_V$, let $T_1(P|_V), T_2(P|_V)$ and $T_3(P|_V)$ be as defined in Section 2. The following lemma is the crux of all our arguments.

Lemma 13. *For all $\alpha, \alpha' \in \mathscr{M}$ such that $\alpha \neq \alpha'$,*

1. $T_1(P|_V, \alpha) = L_2 \cdot \binom{N-k}{m}$.
2. $\mathbb{E}_{V\leftarrow D}[T_2(P|_V, \alpha)] \leqslant L_2 \cdot \binom{N-k}{m} \cdot 2^{O(r')}$.
3. *(a)* $\mathbb{E}_{V\leftarrow D}[T_3(P|_V, \alpha, \alpha)] \leqslant L_2 \cdot \binom{N-k}{m} \cdot 2^{O(r')}$
   (b) $\mathbb{E}_{V\leftarrow D}[T_3(P|_V, \alpha, \alpha')] \leqslant L_2 \cdot \binom{N-k}{m} \cdot 2^{O(r')} \cdot \left(\left(\frac{m}{N-m}\right)^{k'} + \frac{1}{n^{(1-\tau)\nu}} \right)^{(2r'-r)}$.

Proof of this lemma can be found in Section 5. The following follows from a straightforward application of Markov inequality.

Lemma 14.

$$\mathbb{P}_{V\leftarrow D}\left[(T_2(P|_V) < 20 \cdot \mathbb{E}_{V'\leftarrow D}[T_2(P|_{V'})]) \wedge (T_3(P|_V) < 20 \cdot \mathbb{E}_{V'\leftarrow D}[T_3(P|_{V'})]) \right] \geqslant 0.9.$$

Lemma 15 (Lemma 8.8, [KS17b])**.** *With a probability of at least 0.9, there is a set $\mathscr{M}' \subseteq \mathscr{M}$ of size at least $4L_1/5$ such that for all $\alpha \in \mathscr{M}'$, $T_2(P|_V, \alpha) \leqslant 100 \cdot \mathbb{E}_{V\leftarrow D}[T_2(P|_V, \alpha)]/L_1$.*

From Lemma 4, Lemma 13 and Lemma 15, we can infer the following.

$$\sum_{\substack{\alpha\in\mathscr{M}'\\ \beta\in\mathscr{M}(\alpha)}} |S_m(\alpha,\beta)| = \sum_{\alpha\in\mathscr{M}'} T_1(P|_V, \alpha) \geqslant \frac{4L_1}{5} \cdot L_2 \cdot \binom{N-k}{m}$$

$$\sum_{\substack{\alpha\in\mathscr{M}'\\ \beta\neq\gamma\in\mathscr{M}(\alpha)}} |S_m(\alpha,\beta) \cap S_m(\alpha,\gamma)| = \sum_{\alpha\in\mathscr{M}'} T_1(P|_V, \alpha) \leqslant \frac{4L_1}{5} \cdot L_2 \cdot \binom{N-k}{m} \cdot 2^{O(r')}$$

$$\text{and thus } \left| \bigcup_{\substack{\alpha\in\mathscr{M}\\ \beta\in\mathscr{M}(\alpha)}} S_m(\alpha,\beta) \right| \geqslant \left| \bigcup_{\substack{\alpha\in\mathscr{M}'\\ \beta\in\mathscr{M}(\alpha)}} S_m(\alpha,\beta) \right| \geqslant \frac{L_1 \cdot L_2 \cdot \binom{N-k}{m}}{2^{O(r')}}.$$

Lemma 16. *If $\left(\frac{m}{N-m}\right)^{k'} \geqslant n^{-\nu(1-\tau)}$, then*

$$\mathbb{E}_{V\leftarrow D}[T_3(P|_V)] \leqslant L_1^2 \cdot L_2 \cdot \binom{N-k}{m} \cdot 2^{O(r')} \cdot \left(\frac{m}{N-m}\right)^{k'(2r'-r)}.$$

Proof. The proof follows from Item 3 of Lemma 13 and simple application of linearity of expectation.

$$
\begin{aligned}
&\mathbb{E}_{V\leftarrow D}\left[T_3(P|_V)\right]\\
&\leqslant \mathbb{E}_{V\leftarrow D}\left[\sum_{\alpha,\alpha'\in\mathscr{M}} T_3(P|_V,\alpha,\alpha')\right]\\
&= \mathbb{E}_{V\leftarrow D}\left[\sum_{\alpha\in\mathscr{M}} T_3(P|_V,\alpha,\alpha)\right] + \mathbb{E}_{V\leftarrow D}\left[\sum_{\substack{\alpha,\alpha'\in\mathscr{M}\\ \alpha\neq\alpha'}} T_3(P|_V,\alpha,\alpha')\right]\\
&\leqslant \sum_{\alpha\in\mathscr{M}} L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')} + \sum_{\substack{\alpha,\alpha'\in\mathscr{M}\\ \alpha\neq\alpha'}} L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(\left(\frac{m}{N-m}\right)^{k'}+\frac{1}{n^{(1-\tau)\nu}}\right)^{(2r'-r)}\\
&\leqslant L_1\cdot L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')} + L_1^2\cdot L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(\left(\frac{m}{N-m}\right)^{k'}+\frac{1}{n^{(1-\tau)\nu}}\right)^{(2r'-r)}\\
&\leqslant L_1^2\cdot L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(\left(\frac{m}{N-m}\right)^{k'}+\frac{1}{n^{(1-\tau)\nu}}\right)^{(2r'-r)}\\
&\leqslant L_1^2\cdot L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(2\left(\frac{m}{N-m}\right)^{k'}\right)^{(2r'-r)}\\
&= L_1^2\cdot L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(\frac{m}{N-m}\right)^{k'(2r'-r)}.
\end{aligned}
$$

□

By the setting of parameters, $L_1 = n^{(1+\tau)r'}\cdot 2^{-O(r')}$. By the choice of parameters, we get that L_1 is greater than $(1+\Gamma)^{2k'(2r'-r)}$. With a probability of at least 0.9, we get that

$$T_3(P|_V) \leqslant 20\cdot\mathbb{E}_{V'\leftarrow D}\left[T_3(P|_{V'})\right] \leqslant L_1^2\cdot L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(\frac{m}{N-m}\right)^{k'(2r'-r)}$$

and

$$\sum_{\substack{\alpha,\alpha'\in\mathscr{M}'\\ \beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha,\gamma)}} \left|A_m(\alpha,\beta)\cap A_m(\alpha',\gamma)\right| \leqslant \sum_{\substack{\alpha,\alpha'\in\mathscr{M}\\ \beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha,\gamma)}} \left|A_m(\alpha,\beta)\cap A_m(\alpha',\gamma)\right| = T_3(P|_V)$$

From the afore mentioned discussion and by using the Strong Inclusion Exclusion lemma (Lemma 4), we get the following.

$$\mathrm{PSSPD}_{r',m}(P|_V) \geqslant \left|\bigcup_{\substack{\alpha\in\mathscr{M}'\\ \beta\in M(\alpha)}} A_m(\alpha,\beta)\right| \geqslant \frac{L_2\cdot\binom{N-k}{m}}{2^{O(r')}\cdot\left(\frac{m}{N-m}\right)^{k'(2r'-r)}}.$$

Theorem 17. *Let n, d and δ be large enough integers such that $d \leqslant n^{0.01}$ and $\delta \leqslant n^{0.016}$. The polynomial $P|_V$ of degree $d = (2k'+3)\cdot r'$, over $n^{O(1)}$ variables $Y \sqcup Z$ such that any syntactically multi-δ-ic depth four circuit with bottom Z-support of at most $t = 2\times 10^{-5}k'$, computing it must have size $n^{\Omega(r')}$.*

Proof. Let $\Phi|_V$ be the multi-δ-ic depth four circuit of bottom Z support of at most t, that computes the polynomial $P|_V$. Thus,

$$\mathrm{PSSPD}_{r',m}(\Phi|_V) = \mathrm{PSSPD}_{r',m}(P|_V).$$

Combining the above discussion with Lemma 9, we get that

$$s \cdot \binom{\frac{2N\delta}{t}}{r'} \cdot \binom{N}{m + r't} \cdot (m + r't) \geqslant \frac{L_2 \cdot \binom{N-k}{m}}{2^{O(r')} \cdot \left(\frac{m}{N-m}\right)^{k'(2r'-r)}}.$$

Thus,

$$\begin{aligned}
s &\geqslant \frac{L_2 \cdot \binom{N-k}{m}}{2^{O(r')} \cdot \left(\frac{m}{N-m}\right)^{k'(2r'-r)} \cdot \binom{\frac{2N\delta}{t}}{r'} \cdot \binom{N}{m+r't} \cdot (m + r't)} \\
&= \frac{L_2}{2^{O(r')} \cdot \left(\frac{m}{N-m}\right)^{k'(2r'-r)} \cdot \binom{\frac{2N\delta}{t}}{r'}} \cdot \frac{\binom{N-k}{m}}{\binom{N}{m+r't}} \\
&\approx \frac{L_2}{2^{O(r')} \cdot \left(\frac{m}{N-m}\right)^{k'(2r'-r)} \cdot \binom{\frac{2N\delta}{t}}{r'}} \cdot \left(\frac{N-m}{N}\right)^{k} \cdot \left(\frac{m}{N-m}\right)^{r't} \\
&= \frac{1}{2^{O(r')} \cdot \binom{\frac{2N\delta}{t}}{r'}} \cdot \left(\frac{N-m}{m}\right)^{k'(2r'-r)-r't} && \text{Since } L_2 = \left(\frac{N}{N-m}\right)^{k} \\
&\geqslant \frac{1}{2^{O(r')}} \cdot \left(\frac{N-m}{m}\right)^{k'(2r'-r)-r't} \cdot \left(\frac{tr'}{2eN\delta}\right)^{r'} && \text{Since } \binom{n}{k} \leqslant \left(\frac{en}{k}\right)^{k} \\
&= \frac{1}{2^{O(r')}} \cdot \left(\frac{N-m}{m}\right)^{k'(2r'-r)-r't} \cdot \left(\frac{t}{n^{(1+\eta)}\delta}\right)^{r'} && \text{Since } N = O(n^{(1+\eta)} \cdot r') \\
&= \frac{1}{2^{O(r')}} \cdot \left(\frac{1+\Gamma}{1-\Gamma}\right)^{k'(2r'-r)-r't} \cdot \left(\frac{tr'}{n^{(1+\eta)}\delta}\right)^{r'} \\
&\geqslant \frac{1}{2^{O(r')}} \cdot (1+\Gamma)^{2(k'(2r'-r)-r't)} \cdot \left(\frac{t}{n^{(1+\eta)}\delta}\right)^{r'} && \text{Since } \frac{1}{1-\Gamma} \geqslant (1+\Gamma) \\
&\geqslant \frac{1}{2^{O(r')}} \cdot (n^{\varepsilon})^{4(1-\tau-10^{-5})r'} \cdot \left(\frac{t}{n^{(1+\eta)}\delta}\right)^{r'} \\
&= \frac{1}{2^{O(r')}} \cdot n^{(4\varepsilon(1-\tau-10^{-5})-(1+\eta))r'} \cdot \left(\frac{t}{\delta}\right)^{r'} \\
&\geqslant \frac{1}{2^{O(r')}} \cdot n^{0.017r'} \cdot \left(\frac{t}{\delta}\right)^{r'}.
\end{aligned}$$

Thus, for all $\delta \leqslant n^{0.016}$, we get that $s \geqslant n^{\Omega(r')}$. □

Putting it all together

Let Φ be a multi-δ-ic depth four circuit of size at most $n^{\frac{t}{2}}$ computing the $\mathsf{IMM}_{n,d}$ polynomial. From Lemma 8, we get that with a probability of at least $(1 - o(1))$ over $V \leftarrow D$, $\Phi|_V$ is a multi-δ-ic depth four circuit of bottom support at most t. Note that $\Phi|_V$ is of size at most $n^{\frac{t}{2}}$. From Theorem 17, $\Phi|_V$ must have size at least $n^{\Omega(r')} = n^{c_0 r'}$. From our choice of parameters, $c_0 r'$ is at most $\frac{t}{2}$. Thus, any multi-δ-ic depth four circuit computing $\mathsf{IMM}_{n,d}$ must have size at least $n^{\Omega(\sqrt{d})}$. We can now formally state our main theorem as follows.

Theorem 18. *Let n, d and δ be integers such that $d \leqslant n^{0.01}$ and $\delta \leqslant n^{0.016}$. Then any syntactically multi-δ-ic depth four circuit computing $\mathsf{IMM}_{n,d}$ must be of size $n^{\Omega(\sqrt{d})}$.*

5 Proof of Lemma 13

Proof of Item 1 of Lemma 13: Recall that $T_1(P|_V, \alpha) = \sum_{\beta \in M(\alpha)} |S_m(\alpha, \beta)|$. $S_m(\alpha, \beta)$ corresponds to the set of multilinear monomials γ of degree m whose support has zero intersection with the support of the monomial β. Thus, $|S_m(\alpha, \beta)| = \binom{N-k}{m}$. Thus,

$$T_1(P|_V, \alpha) = \sum_{\beta \in M(\alpha)} \binom{N-k}{m} = L_2 \cdot \binom{N-k}{m}.$$

Proof of Item 2 of Lemma 13: Recall that $T_2(P|_V, \alpha)$ corresponds to the sum of cardinalities of sets of monomials $S_m(\alpha, \beta) \cap S_m(\alpha, \gamma)$, for all $\beta, \gamma \in M(\alpha)$. Further for all $\beta, \gamma \in M(\alpha)$, $S_m(\alpha, \beta) \cap S_m(\alpha, \gamma)$ corresponds to set of multilinear monomials of degree m which are disjoint from both β and γ.

$$\begin{aligned}
T_2(P|_V, \alpha) &= \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} |S_m(\alpha, \beta) \cap S_m(\alpha, \gamma)| \\
&= \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \binom{N-k-\Delta(\beta,\gamma)}{m} \\
&= \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \binom{N-k-\Delta(\beta,\gamma)}{m} \frac{\binom{N-k}{m}}{\binom{N-k}{m}} \\
&= \binom{N-k}{m} \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \frac{\binom{N-k-\Delta(\beta,\gamma)}{m}}{\binom{N-k}{m}} \\
&\approx \binom{N-k}{m} \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}.
\end{aligned}$$

The quantity $\Delta(\beta, \gamma)$ could be as small as 1 and thus the quantity $\left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}$ could be as large as $\frac{N-m}{N}$. We now need the following proposition from [KS17b] to show that on average $\sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}$ cannot be too large. Even though the proof of Proposition 19 follows from Proposition 9.1 in [KS17b], we present it again in Section 6 for the sake of completeness and show that it is resilient under the change of parameters.

Proposition 19 (Proposition 9.1, [KS17b])**.** *For any monomial β in $M(\alpha)$ (where $M(\alpha)$ is the set of monomials in the support of $\sigma_Y(\partial_\alpha(P|_V))$),*

$$\mathbb{E}_{V \leftarrow D}\left[\sum_{\substack{\gamma \in M(\alpha) \\ \beta \neq \gamma}} \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}\right] \leqslant 2^{O(r')}.$$

Using Proposition 19, we get that

$$\begin{aligned}
\mathbb{E}_{V\leftarrow D}\left[T_2(P|_V,\alpha)\right] &\approx \mathbb{E}_{V\leftarrow D}\left[\binom{N-k}{m}\sum_{\substack{\beta,\gamma\in M(\alpha)\\ \beta\neq\gamma}}\left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}\right]\\
&= \binom{N-k}{m}\mathbb{E}_{V\leftarrow D}\left[\sum_{\beta\in M(\alpha)}\sum_{\substack{\gamma\in M(\alpha)\\ \beta\neq\gamma}}\left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}\right]\\
&= \binom{N-k}{m}\sum_{\beta\in M(\alpha)}\mathbb{E}_{V\leftarrow D}\left[\sum_{\substack{\gamma\in M(\alpha)\\ \beta\neq\gamma}}\left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}\right]\\
&\leqslant \binom{N-k}{m}\sum_{\beta\in M(\alpha)}\left(2^{O(r')}\right)\\
&\leqslant \binom{N-k}{m}\cdot L_2\cdot 2^{O(r')}.
\end{aligned}$$

Proof of Item 3 of Lemma 13: Recall that the term $T_3(P|_V,\alpha,\alpha')$ corresponds to the sum of cardinalities of sets of monomials $A_m(\alpha,\beta)\cap A_m(\alpha',\gamma)$, for all $\beta\in M(\alpha)$ and $\gamma\in M(\alpha')$. Note that $|A_m(\alpha,\beta)\cap A_m(\alpha',\gamma)|$ is upper bounded by the number of multilinear monomials of degree $m+k$ which are divisible by both β and γ.

$$\begin{aligned}
T_3(P|_V,\alpha,\alpha') &= \sum_{\substack{\beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha',\gamma)}}\left|A_m(\alpha,\beta)\cap A_m(\alpha',\gamma)\right|\\
&\leqslant \sum_{\substack{\beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha',\gamma)}}\binom{N-k-\Delta(\beta,\gamma)}{m-\Delta(\beta,\gamma)}\\
&= \sum_{\substack{\beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha',\gamma)}}\binom{N-k-\Delta(\beta,\gamma)}{m-\Delta(\beta,\gamma)}\frac{\binom{N-k}{m}}{\binom{N-k}{m}}\\
&= \binom{N-k}{m}\sum_{\substack{\beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha',\gamma)}}\frac{\binom{N-k-\Delta(\beta,\gamma)}{m-\Delta(\beta,\gamma)}}{\binom{N-k}{m}}\\
&= \binom{N-k}{m}\sum_{\substack{\beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha',\gamma)}}\frac{(N-k-\Delta(\beta,\gamma))!\cdot m!}{(m-\Delta(\beta,\gamma))!\cdot(N-k)!}\\
&\approx \binom{N-k}{m}\sum_{\substack{\beta\in M(\alpha)\\ \gamma\in M(\alpha')\\ (\alpha,\beta)\neq(\alpha',\gamma)}}\frac{N^k\cdot m^{\Delta(\beta,\gamma)}}{N^{k+\Delta(\beta,\gamma)}}
\end{aligned}$$

$$= \binom{N-k}{m} \sum_{\substack{\beta \in M(\alpha) \\ \gamma \in M(\alpha') \\ (\alpha,\beta) \neq (\alpha',\gamma)}} \left(\frac{m}{N}\right)^{\Delta(\beta,\gamma)}.$$

This above sum can be further refined into two cases as follows.

$$T_3^{=}(P|_V, \alpha) = \binom{N-k}{m} \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \left(\frac{m}{N}\right)^{\Delta(\beta,\gamma)}$$

$$\text{and } \forall\ \alpha \neq \alpha',\ T_3^{\neq}(P|_V, \alpha, \alpha') = \binom{N-k}{m} \sum_{\substack{\beta \in M(\alpha) \\ \gamma \in M(\alpha') \\ (\alpha,\beta) \neq (\alpha',\gamma)}} \left(\frac{m}{N}\right)^{\Delta(\beta,\gamma)}.$$

As in the case of $T_2(P|_V, \alpha)$, we can compute the expected value of $T_3^{=}(P|_V, \alpha)$ using Proposition 19 as follows.

$$\begin{aligned}
\mathbb{E}_{V \leftarrow D}\left[T_3(P|_V, \alpha)\right] &\approx \mathbb{E}_{V \leftarrow D}\left[\binom{N-k}{m} \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \left(\frac{m}{N}\right)^{\Delta(\beta,\gamma)}\right] \\
&\leqslant \mathbb{E}_{V \leftarrow D}\left[\binom{N-k}{m} \sum_{\substack{\beta,\gamma \in M(\alpha) \\ \beta \neq \gamma}} \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}\right] \qquad \left(\text{Since } m < \frac{N}{2}\right) \\
&\leqslant \binom{N-k}{m} \sum_{\beta \in M(\alpha)} \left(2^{O(r')}\right) \\
&\leqslant \binom{N-k}{m} \cdot L_2 \cdot 2^{O(r')}.
\end{aligned}$$

For any $\beta \in M(\alpha)$, let $T_3^{\neq}(P|_V, \alpha, \alpha', \beta)$ be defined as follows.

$$T_3^{\neq}(P|_V, \alpha, \alpha', \beta) = \binom{N-k}{m} \sum_{\substack{\gamma \in M(\alpha') \\ \beta \neq \gamma}} \left(\frac{m}{N}\right)^{\Delta(\beta,\gamma)}$$

and thus

$$T_3^{\neq}(P|_V, \alpha, \alpha') = \sum_{\beta \in M(\alpha)} T_3^{\neq}(P|_V, \alpha, \alpha', \beta).$$

For any $\beta \in M(\alpha), \gamma \in M(\alpha')$, let $\mathrm{BlockDiff}(\beta, \gamma)$ be defined as follows.

$$\mathrm{BlockDiff}(\beta, \gamma) = \left\{ i \mid \Delta(\beta^{(i)}, \gamma^{(i)}) = k' \right\}.$$

For any $\beta \in M(\alpha)$, let $C_V^t(\beta)$ be the set of monomials $\gamma \in M(\alpha')$ such that $(\alpha, \beta) \neq (\alpha', \gamma)$ and $|\mathrm{BlockDiff}(\beta, \gamma)| = t$. Note that for all $\gamma \in C_V^t(\beta)$, $\Delta(\beta, \gamma) \geqslant t \cdot k'$.

$$T_3^{\neq}(P|_V, \alpha, \alpha', \beta) = \binom{N-k}{m} \sum_{\substack{\gamma \in M(\alpha') \\ (\alpha,\beta) \neq (\alpha',\gamma)}} \left(\frac{m}{N}\right)^{\Delta(\beta,\gamma)}$$

$$
\begin{aligned}
&= \binom{N-k}{m} \sum_{\gamma \in M(\alpha')} \left(\frac{m}{N-m}\right)^{\Delta(\beta,\gamma)} \cdot \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)} \\
&\leqslant \binom{N-k}{m} \sum_{t=0}^{2r'-r} \sum_{\gamma \in C_V^t(\beta)} \left(\frac{m}{N-m}\right)^{t \cdot k'} \cdot \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)} \\
&\leqslant \binom{N-k}{m} \sum_{t=0}^{2r'-r} \left(\frac{m}{N-m}\right)^{t \cdot k'} \cdot \left(\sum_{\gamma \in C_V^t(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)} \right)
\end{aligned}
$$

Recall that $\Delta(\beta,\gamma) = \prod_{i \in [2r']} \Delta(\beta^{(i)}, \gamma^{(i)})$. Let $C_V^t(\beta)$ be expressed as a sum of all $C_V^{t,S}(\beta)$ for all sets S of size $2r' - r - t$ that indexes blocks i such that $\alpha_i \neq \alpha'_i$ and $\mathrm{BlockDiff}(\beta^{(i)}, \gamma^{(i)}) < k'$.

$$
\begin{aligned}
&\sum_{\gamma \in C_V^t(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)} \\
&= \sum_{\gamma \in C_V^t(\beta)} \prod_{i \in [r']} \left(\frac{N-m}{N}\right)^{\Delta(\beta^{(i)},\gamma^{(i)})} \\
&= \sum_{\substack{S \subseteq [2r'-r] \\ |S| = 2r'-r-t}} \sum_{\gamma \in C_V^{t,S}(\beta)} \prod_{i \in [r']} \left(\frac{N-m}{N}\right)^{\Delta(\beta^{(i)},\gamma^{(i)})} \\
&\leqslant \binom{2r'-r}{t} \cdot \left(\prod_{i \in S} \sum_{\gamma \in C_V^{t,S}(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta^{(i)},\gamma^{(i)})} \right) \cdot \left(\prod_{i \notin S} \sum_{\gamma \in C_V^{t,S}(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta^{(i)},\gamma^{(i)})} \right).
\end{aligned}
$$

From the definition, it follows that for all those $\gamma \in C_V^t(\beta)$ and $i \in S$, $\gamma^{(i)} \in A_V^{(i)}(\beta)$.

$$
\begin{aligned}
\mathbb{E}_{V \leftarrow D} \left[\prod_{i \in S} \sum_{\gamma \in C_V^{t,S}(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta^{(i)},\gamma^{(i)})} \right] &= \prod_{i \in S} \mathbb{E}_{V \leftarrow D} \left[\sum_{\gamma \in C_V^{t,S}(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta^{(i)},\gamma^{(i)})} \right] \\
&\leqslant \left(\frac{1}{n^\nu}\right)^{2r'-r-t}.
\end{aligned}
$$

For the blocks that are not indexed by S, we get the following.

$$
\mathbb{E}_{V \leftarrow D} \left[\prod_{i \notin S} \sum_{\gamma \in C_V^t(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta^{(i)},\gamma^{(i)})} \right] \leqslant 2^{O(r')}.
$$

Putting these together we get that

$$
\mathbb{E}_{V \leftarrow D} \left[\sum_{\gamma \in C_V^t(\beta)} \left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)} \right] \leqslant \binom{2r'-r}{t} \cdot \frac{2^{O(r')}}{n^{(2r'-r-t)\nu}}
$$

and

$$\begin{aligned}
&\mathbb{E}_{V\leftarrow D}\left[T_3^{\neq}(P|_V,\alpha,\alpha',\beta)\right]\\
&=\mathbb{E}_{V\leftarrow D}\left[\binom{N-k}{m}\sum_{t=0}^{2r'-r}\left(\frac{m}{N-m}\right)^{t\cdot k'}\cdot\left(\sum_{\gamma\in C_V^t(\beta)}\left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}\right)\right]\\
&=\binom{N-k}{m}\cdot\sum_{t=0}^{2r'-r}\left(\frac{m}{N-m}\right)^{t\cdot k'}\cdot\mathbb{E}_{V\leftarrow D}\left[\sum_{\gamma\in C_V^t(\beta)}\left(\frac{N-m}{N}\right)^{\Delta(\beta,\gamma)}\right]\\
&\leqslant\binom{N-k}{m}\cdot\sum_{t=0}^{2r'-r}\left(\frac{m}{N-m}\right)^{t\cdot k'}\cdot\binom{2r'-r}{t}\cdot\frac{2^{O(r')}}{n^{(2r'-r-t)\nu}}\\
&=\binom{N-k}{m}\cdot 2^{O(r')}\cdot\sum_{t=0}^{2r'-r}\binom{2r'-r}{t}\cdot\left(\frac{m}{N-m}\right)^{t\cdot k'}\cdot\frac{1}{n^{(2r'-r-t)\nu}}\\
&=\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(\left(\frac{m}{N-m}\right)^{k'}+\frac{1}{n^{\nu}}\right)^{2r'-r}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}_{V\leftarrow D}\left[T_3^{\neq}(P|_V,\alpha,\alpha')\right]&=\sum_{\beta\in M(\alpha)}\mathbb{E}_{V\leftarrow D}\left[T_3^{\neq}(P|_V,\alpha,\alpha',\beta)\right]\\
&\leqslant L_2\cdot\binom{N-k}{m}\cdot 2^{O(r')}\cdot\left(\left(\frac{m}{N-m}\right)^{k'}+\frac{1}{n^{\nu}}\right)^{2r'-r}.
\end{aligned}$$

6 Proof of Proposition 19

Recall that each $\beta\in M(\alpha)$ can be expressed as a product of $\beta^{(i)}$'s for all $i\in[2r']$. That is, for all $i\in[2r']$, $\beta^{(i)}$ corresponds to a *monomial* from the ith block. It is easy to see that $\Delta(\beta,\gamma)$ is equal to $\sum_{i\in[2r']}\Delta(\beta^{(i)},\gamma^{(i)})$ as the blocks are defined over disjoint sets of variables. Thus,

$$\mathbb{E}_{V\leftarrow D}\left[\sum_{\substack{\gamma\in M(\alpha)\\ \beta\neq\gamma}}D^{-\Delta(\beta,\gamma)}\right]\leqslant\prod_{i\in[r']}\mathbb{E}_{V\leftarrow D}\left[\sum_{\substack{\gamma\in M(\alpha)\\ \beta\neq\gamma}}D^{-\Delta(\beta^{(i)},\gamma^{(i)})}\right].$$

To prove Proposition 19, it is sufficient to show that each term of the right side of the above equation is at most 5.

Lemma 20. *For all $\beta\in M(\alpha)$ and all $i\in[2r']$,*

$$\mathbb{E}_{V\leftarrow D}\left[\sum_{\substack{\gamma\in M(\alpha)\\ \beta\neq\gamma}}D^{-\Delta(\beta^{(i)},\gamma^{(i)})}\right]\leqslant O(1).$$

Proof. Let $M^{(i)}(\alpha)$ be the set of monomials corresponding to the polynomial h_i after deriving $P|_V$ with $\alpha\in\mathscr{M}$. It is now clear that $M(\alpha)\subseteq\prod_{i=1}^{2r'}M^{(i)}(\alpha)$. For all $\beta^{(i)}\in M^{(i)}(\alpha)$, let

- $A_V^{(i)}(\beta^{(i)})$ be the set of $\gamma^{(i)}$ such that there is some $j \in [k'-1]$ such that $\beta^{(i,j)} = \gamma^{(i,j)}$ and $\beta^{(i,j+1)} = \gamma^{(i,j+1)}$.
- $B_V^{(i)}(\beta^{(i)})$ be the set of $\gamma^{(i)}$ such that there is some $j, j' \in [k']$ such that if $\beta^{(i,j)} = \gamma^{(i,j)}$ and $\beta^{(i,j')} \neq \gamma^{(i,j')}$ then $j < j'$.

Claim 21. *For all* $\beta^{(i)} \in M^{(i)}(\alpha)$*,*

$$\mathbb{E}_{V \leftarrow D}\left[\sum_{\gamma^{(i)} \in B_V^{(i)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})}\right] \leqslant 4.$$

Proof. Let $B_V^{(i)}(\beta^{(i)})$ be further partitioned into $k'+1$ sets $B_V^{(i,j)}(\beta^{(i)})$ for all $t \in [0, k']$ such that $B_V^{(i,t)}(\beta^{(i)})$ consists of all those $\gamma^{(i)}$'s such that $\beta^{(i,j)} = \gamma^{(i,j)}$ for all $j \leqslant t$ and $\beta^{(i,j)} \neq \gamma^{(i,j)}$ for all $j > t$. This means that every $\gamma^{(i)}$ in $B_V^{(i,t)}(\beta^{(i)})$ matches with $\beta^{(i)}$ at the first t positions and differ at every position after that. Thus, it is easy to see that there are at most $\prod_{j>t} \deg_V(X^{(i,j)})$. This is due to the fact that there are $\deg_V(X^{(i,j)})$ choices at every position $j > t$. Recall that $\prod_{j\in[k']} \deg_V(X^{(i,j)}) = D^{k'}$ and for all $\gamma^{(i)} \in B_V^{(i,t)}(\beta^{(i)})$, $\Delta(\beta^{(i)}, \gamma^{(i)}) = k' - t$. Thus,

$$\sum_{\gamma^{(i)} \in B_V^{(i,t)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} \leqslant \prod_{j>t} \deg_V(X^{(i,j)}) \cdot D^{-(k'-t)}$$

For $t = 0$, the above expression equals 1. For $t = 1$, the above expression is at most $\frac{D}{n^\eta}$. For $t \in [2, k' - \varepsilon' \log n]$, the above expression evaluates to

$$\begin{aligned}
\sum_{\gamma^{(i)} \in B_V^{(i,t)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} &\leqslant D^t \prod_{j>t} (\deg_V(X^{(i,j)}))^{-1} \\
&= \frac{D}{n^\eta} \prod_{j=2}^{t} \frac{D}{2} \\
&\leqslant \frac{D}{n^\eta} && \text{Since } D < 2.
\end{aligned}$$

For $t > k' - 2\log n$, we get that

$$\sum_{\gamma^{(i)} \in B_V^{(i,t)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} \leqslant \prod_{j>t} \deg_V(X^{(i,j)}) \cdot D^{-(k'-t)} = D^{-(k'-t)}.$$

Thus, we get that

$$\begin{aligned}
\sum_{\gamma^{(i)} \in B_V^{(i)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} &= \sum_{t=0}^{k'} \sum_{\gamma^{(i)} \in B_V^{(i,t)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} \\
&\leqslant 1 + (k' - \varepsilon' \log n) \cdot \frac{D}{n^\eta} + \sum_{t=k'-\varepsilon' \log n+1}^{k'} D^{-(k'-t)} \\
&\leqslant 2 + \sum_{t=0}^{\varepsilon' \log n} D^{-t}
\end{aligned}$$

$$\leqslant 2 + \frac{D}{D-1} \leqslant 4.$$

The second inequality in the above mathblock uses the fact that $2(k' - \varepsilon' \log n) < n^{\eta}$. This completes the proof of Claim 21. □

Claim 22. *For all* $\beta^{(i)} \in M^{(i)}(\alpha)$,

$$\mathbb{E}_{V \leftarrow D}\left[\sum_{\gamma^{(i)} \in A_V^{(i)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)}, \gamma^{(i)})}\right] \leqslant \frac{1}{n^{\nu}}.$$

Proof. For any $\beta^{(i)} \in M^{(i)}(\alpha)$ and $\gamma^{(i)}$, and for some $j \in [k']$ we call j an agreement switch if $\beta^{(i,j-1)} \neq \gamma^{(i,j-1)}$ and $\beta^{(i,j)} = \gamma^{(i,j)}$, and a disagreement switch if $\beta^{(i,j-1)} = \gamma^{(i,j-1)}$ and $\beta^{(i,j)} \neq \gamma^{(i,j)}$.

Let $A_V^{(i)}(\beta^{(i)})$ be further partitioned into k' sets $A_V^{(i,j)}(\beta^{(i)})$ for all $t \in [0, k']$ based on the number of switches such that $A_V^{(i,t)}(\beta^{(i)})$ contains all those $\gamma^{(i)}$ that contain t many disagreements with respect to $\beta^{(i)}$.

Let S_t be a t sized subset of $[k']$ and $b \in \{0,1\}$ be a bit. By specifying the first switch b and the switch locations S_t, we can precisely characterize any element in $A_V^{(i,t)}(\beta^{(i)})$. Let $T = \{d_1, d_2, \cdots, d_s\}$ be the set of s coordinates where $\gamma^{(i)} \in A_V^{(i,t)}(\beta^{(i)})$ disagrees with $\beta^{(i)}$ where $\gamma^{(i)}$ is characterized by (b, S_t). Note that there are at most $\deg_V(X^{(i,d_j)})$ many options for picking an edge to the next layer, that is, at most $\deg_V(X^{(i,d_j)})$ many possible labels for position $j \in [s]$.

Further, if there is a agreement switch, then the last disagreeing edge label must have the same end point as that of $\beta^{(i)}$. Since the edge end points are picked uniformly at random in D and $\beta^{(i)}$ being fixed, the probability that the correct endpoint is chosen is $\frac{1}{n}$.

Thus given a disagreement pattern T and the fact that there are q agreement switches, the number of $\gamma^{(i)}$'s that match the disagreement pattern and q agreement switches is at most $\prod_{j \in T} \deg_V(X^{(i,j)}) \cdot n^{-q}$. If there are t switches in a $\gamma^{(i)}$ the number of agreement switches can at most be $(t+1)/2$ and at least be $\max\{1, (t-1)/2\}$. This is due to the fact that the agreement and disagreement switches alternate. Thus the number of elements in $A_V^{(i,t)}(\beta^{(i)})$ with disagreement pattern T (denoted by $A_V^{(i,t,T)}(\beta^{(i)})$) is at most $\prod_{j \in T} \deg_V(X^{(i,j)}) \cdot n^{\max\{1,(t-1)/2\}}$. Recall that fixing (b, S_t) fixes an element of $A_V^{(i,t)}(\beta^{(i)})$ and thus it fixes a disagreement pattern. There are at most $2 \cdot \binom{k'}{t}$ many sets (b, S_t) and thus there are at most $2 \cdot \binom{k'}{t}$ disagreement patterns. Putting this all together we get that

$$\begin{aligned}
&\mathbb{E}_{V \leftarrow D}\left[\sum_{\gamma^{(i)} \in A_V^{(i)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)}, \gamma^{(i)})}\right] \\
&\leqslant \sum_{\substack{t \in [k'] \\ T \subseteq [k'] \\ \gamma^{(i)} \in A_V^{(i,t,T)}(\beta^{(i)})}} \left(\prod_{j \in T} \deg_V(X^{(i,j)})\right) \cdot n^{-\max\{1,(t-1)/2\}} \cdot D^{-|T|}.
\end{aligned}$$

Claim 23. *For all* $i \in [r']$ *and all* $T \subseteq [k']$, $\left(\prod_{j \in T} \deg_V(X^{(i,j)})\right) \cdot D^{-|T|} \leqslant n^{\varepsilon'}$.

Proof. It is easy to see that the product $\left(\prod_{j\in T} \deg_V(X^{(i,j)})\right)$ is maximized for $T = [k' - \varepsilon' \log n]$.

$$\begin{aligned}\max_T \left\{ \left(\prod_{j\in T} \deg_V(X^{(i,j)}) \right) \cdot D^{-|T|} \right\} &\leqslant \left(\prod_{j=1}^{k'-\varepsilon'\log n} \deg_V(X^{(i,j)}) \right) \cdot D^{\varepsilon'\log n - k'} \\ &= D^{k'} \cdot D^{\varepsilon'\log n - k'} \\ &= D^{\varepsilon'\log n} \leqslant n^{\varepsilon'}.\end{aligned}$$

□

Using this fact, we get that

$$\mathbb{E}_{V\leftarrow D}\left[\sum_{\gamma^{(i)}\in A_V^{(i)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} \right] \leqslant \sum_{t\in[k']} 2 \cdot \binom{k'}{t} \cdot \frac{n^{\varepsilon'}}{n^{-\max\{1,(t-1)/2\}}} \leqslant \frac{1}{n^{\nu}}.$$

The last inequality follows from our choice of k', ε', ν and n. □

By putting Claim 21 and Claim 22, we get the needed result.

$$\begin{aligned}&\mathbb{E}_{V\leftarrow D}\left[\sum_{\substack{\gamma\in\mathcal{M}(\alpha)\\ \beta\neq\gamma}} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} \right] \\ &= \mathbb{E}_{V\leftarrow D}\left[\sum_{\gamma\in A_V^{(i)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} \right] + \mathbb{E}_{V\leftarrow D}\left[\sum_{\gamma\in B_V^{(i)}(\beta^{(i)})} D^{-\Delta(\beta^{(i)},\gamma^{(i)})} \right] \\ &\leqslant \frac{1}{n^{\nu}} + 4 \leqslant 5.\end{aligned}$$

□

Acknowledgement

Research supported by CHE-PBC Post Doctoral Fellowship, and Binational Science Foundation Grant of mentor Noga Ron-Zewi. We would like to thank Nutan Limaye and Srikanth Srinivasan for patiently listening to a preliminary presentation of [Chi20] (upon which this work is based), Srikanth Srinivasan for his suggestions that eventually led to the current version of this work, Mrinal Kumar for helping us understand their work [KS17b] better, and Ramprasad Saptharishi for lucidly presenting the crux of the arguments in [KS17b] in his survey [Sap19].

References

[AKV18] Noga Alon, Mrinal Kumar, and Ben Lee Volk. Unbalancing sets and an almost quadratic lower bound for syntactically multilinear arithmetic circuits. In *CCC*, volume 102 of *LIPIcs*, pages 11:1–11:16. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.

[CELS18] Suryajith Chillara, Christian Engels, Nutan Limaye, and Srikanth Srinivasan. A near-optimal depth-hierarchy theorem for small-depth multilinear circuits. In *FOCS*, pages 934–945. IEEE Computer Society, 2018.

[Chi20] Suryajith Chillara. On Computing Multilinear Polynomials Using Multi-r-ic Depth Four Circuits. In Christophe Paul and Markus Bläser, editors, *37th International Symposium on Theoretical Aspects of Computer Science (STACS 2020)*, volume 154 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 47:1–47:16, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. URL: https://drops.dagstuhl.de/opus/volltexte/2020/11908, doi:10.4230/LIPIcs.STACS.2020.47.

[CLS18] Suryajith Chillara, Nutan Limaye, and Srikanth Srinivasan. A quadratic size-hierarchy theorem for small-depth multilinear formulas. In *ICALP*, volume 107 of *LIPIcs*, pages 36:1–36:13. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.

[CLS19] Suryajith Chillara, Nutan Limaye, and Srikanth Srinivasan. Small-depth multilinear formula lower bounds for iterated matrix multiplication with applications. *SIAM J. Comput.*, 48(1):70–92, 2019.

[CM19] Suryajith Chillara and Partha Mukhopadhyay. Depth-4 lower bounds, determinantal complexity: A unified approach. *computational complexity*, May 2019. URL: https://doi.org/10.1007/s00037-019-00185-4, doi:10.1007/s00037-019-00185-4.

[FLMS15] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth-4 formulas computing iterated matrix multiplication. *SIAM J. Comput.*, 44(5):1173–1201, 2015.

[GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *Journal of the ACM (JACM)*, 61(6):33, 2014.

[GRS19] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. *Essential coding theory*. 2019. URL: https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/.

[HS17] Sumant Hegde and Chandan Saha. Improved lower bound for multi-r-ic depth four circuits as a function of the number of input variables. *Proceedings of the Indian National Science Academy*, 83(4):907–922, 2017.

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

[KdOS19] Mrinal Kumar, Rafael Mendes de Oliveira, and Ramprasad Saptharishi. Towards optimal depth reductions for syntactically multilinear circuits. In *ICALP*, volume 132 of *LIPIcs*, pages 78:1–78:15. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2019.

[KLSS17] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An exponential lower bound for homogeneous depth four arithmetic formulas. *SIAM J. Comput.*, 46(1):307–335, 2017.

[KS17a] Neeraj Kayal and Chandan Saha. Multi-k-ic depth three circuit lower bound. *Theory Comput. Syst.*, 61(4):1237–1251, 2017.

[KS17b] Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. *SIAM J. Comput.*, 46(1):336–387, 2017.

[KST18] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. On the size of homogeneous and of depth-four formulas with low individual degree. *Theory of Computing*, 14(16):1–46, 2018. URL: http://www.theoryofcomputing.org/articles/v014a016, doi:10.4086/toc.2018.v014a016.

[NW97] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6(3):217–234, 1997. doi:10.1007/BF01294256.

[Raz04] Ran Raz. Multilinear-$\mathsf{NC}^2 \neq$ multilinear-NC^1. In *proceedings of Foundations of Computer Science (FOCS)*, pages 344–351, 2004. URL: https://doi.org/10.1109/FOCS.2004.42, doi:10.1109/FOCS.2004.42.

[Raz06] Ran Raz. Separation of multilinear circuit and formula size. *Theory of Computing*, 2(1):121–135, 2006. doi:10.4086/toc.2006.v002a006.

[RSY08] Ran Raz, Amir Shpilka, and Amir Yehudayoff. A lower bound for the size of syntactically multilinear arithmetic circuits. *SIAM Journal of Computing*, 38(4):1624–1647, 2008. doi:10.1137/070707932.

[RY08] Ran Raz and Amir Yehudayoff. Balancing syntactically multilinear arithmetic circuits. *Computational Complexity*, 17(4):515–535, 2008. doi:10.1007/s00037-008-0254-0.

[RY09] Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Computational Complexity*, 18(2):171–207, 2009. doi:10.1007/s00037-009-0270-8.

[Sap19] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity version 8.0.4. Github survey, 2019. URL: https://github.com/dasarpmar/lowerbounds-survey/releases/.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5:207–388, March 2010. URL: http://dx.doi.org/10.1561/0400000039.

A Upper Bound on $\mathrm{PSSPD}_{r',m}(\Psi)$

Recall that Ψ is a sum of at most s many products of polynomials $T^{(1)} + \cdots + T^{(s)}$ where each T_i is a syntactically multi-δ-ic product of polynomials of low monomial support.

We shall first prove a bound on $\mathrm{PSSPD}_{r',m}(T_i)$ for an arbitrary T_i and derive a bound on $\mathrm{PSSPD}_{r',m}(C)$ by using sub-additivity of the measure (cf. Proposition 1).

Let T be a syntactic multi-δ-ic product of polynomials $\tilde{Q}_1(Y,Z) \cdot \tilde{Q}_2(Y,Z) \cdot \ldots \cdot \tilde{Q}_D(Y,Z) \cdot R(Y)$ such that $\left|\mathrm{MonSupp}_Z\left(\tilde{Q}_i\right)\right| \leqslant t$. We will first argue that D is not too large since T is a syntactically multi-δ-ic product. We shall first pre-process the product T by doing the following procedure.

Repeat this process until all but at most one of the factors in T (except R) have a Z-support of at least $\frac{t}{2}$.

1. Pick two factors $\tilde{Q}_{i_1}$ and $\tilde{Q}_{i_2}$ such that they have the smallest Z-support amongst $\tilde{Q}_1, \cdots, \tilde{Q}_D$.
2. If both of them have support strictly less than $\frac{t}{2}$, merge these factors to obtain a new factor. Else, stop.

In the afore mentioned procedure, it is important to note that the monomial support in Z variables post merging will still be at most t since the factors being merged are of support strictly less than $\frac{t}{2}$. Henceforth, W.L.O.G we shall consider that every product gate at the top, in any multi-δ-ic depth four circuit to be in a processed form.

Let $T = Q_1(Y,Z) \cdot Q_2(Y,Z) \cdot \ldots \cdot Q_D(Y,Z) \cdot R(Y)$ be the product obtained after the preprocessing. Each of the Q_i has a Z-support of at least $\frac{t}{2}$. The total Z-support is at most $|Z|\,\delta = N\delta$ since T is a syntactically multi-δ-ic product. Thus t could at most be $\frac{2N\delta}{t}$.

Lemma 24. *Let N, r', m and t be positive integers such that $m + r't < \frac{N}{2}$. Let T be a processed syntactic multi-δ-ic product of polynomials $Q_1(Y,Z) \cdot Q_2(Y,Z) \cdot \ldots \cdot Q_D(Y,Z) \cdot R(Y)$ such that $|\mathrm{MonSupp}_Z(Q_i)| \leqslant t$. Then, $\mathrm{PSSPD}_{r',m}(T)$ is at most $\binom{D}{r'} \cdot \binom{N}{m+r't} \cdot (m + r't)$.*

Before proving Lemma 24, we shall first use it to show an upper bound on the dimension of the space of Projected Shifted Skew Partial derivatives of a depth four multi-δ-ic circuit of low bottom support.

Proof of Lemma 9. W.L.O.G we can assume that C be expressed as $\sum_{i=1}^{s} T^{(i)}$ such that $T^{(i)}$ is a processed syntactically multi-δ-ic product of polynomials with bottom support at most t with respect to Z variables. From Proposition 1, we get that $\mathrm{PSSPD}_{r',m}(C) \leqslant \sum_{i=1}^{s} \mathrm{PSSPD}_{r',m}(T^{(i)})$. From the afore mentioned discussion we know that the number of factors in $T^{(i)}$ with a non-zero Z-support is at most $\frac{2N\delta}{t}$. From Lemma 24, we get that $\mathrm{PSSPD}_{k,\ell}(T^{(i)})$ is at most $\binom{\frac{2N\delta}{t}}{r'} \cdot \binom{N}{m+r't} \cdot (m + r't)$. By putting all of this together, we get that

$$\mathrm{PSSPD}_{r',m}(C) \leqslant s \cdot \binom{\frac{2Nr}{t}}{r'} \cdot \binom{N}{m + r't} \cdot (m + r't)\,.$$

□

Proof of Lemma 24. We will first show by induction on r', the following.

$$\partial_Y^{=r'} T \subseteq \mathbb{F}\text{-span}\left\{ \bigcup_{S \in \binom{[D]}{D-r'}} \left\{ \prod_{i \in S} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{ML}}^{\leqslant r't} \cdot \mathbb{F}[Y] \right\} \right\}$$
$$\bigcup \mathbb{F}\text{-span}\left\{ \bigcup_{S \in \binom{[D]}{D-r'}} \left\{ \prod_{i \in S} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{NonML}}^{\leqslant r'\delta t} \cdot \mathbb{F}[Y] \right\} \right\}$$

The base case of induction for $r' = 0$ is trivial as T is already in the required form. Let us assume the induction hypothesis for all derivatives of order $< r'$. That is, $\partial_Y^{=r'-1} T$ can be expressed as a linear combination of terms of the form

$$h(Y,Z) = \prod_{i \in S} Q_i(Y,Z) \cdot h_1(Z) \cdot h_2(Y)$$

where S is a set of size $D - (r'-1)$, $h_1(Z)$ is a polynomial in Z variables of degree at most $(r'-1)\delta t$, and $h_2(Y)$ is some polynomial in Y variables. In fact, $h_1(Z)$ can be expressed as a linear combination of multilinear monomials of degree at most $(r'-1)t$, and non-multilinear monomials of degree at most $(r'-1)\delta t$.

For some $u \in [|Y|]$ and some fixed i_0 in S,

$$\begin{aligned}
\frac{\partial h(Y,Z)}{\partial y_u} &= \left(\sum_{j \in S} \prod_{\substack{i \in S \\ i \neq j}} Q_i(Y,Z) \cdot \frac{\partial Q_j(Y,Z)}{\partial y_u} \cdot h_1(Z) \cdot h_2(Y) \right) \\
&\qquad + \frac{\prod_{i \in S} Q_i}{Q_{i_0}} \cdot Q_{i_0}(Y,Z) \cdot h_1(Z) \cdot \frac{\partial h_2(Y)}{\partial y_k} \\
&\in \mathbb{F}\text{-span} \left\{ \prod_{\substack{i \in S \\ i \neq j}} Q_i(Y,Z) \cdot \frac{\partial Q_j(Y,Z)}{\partial y_u} \cdot h_1(Z) \cdot \mathbb{F}[Y] \mid j \in [S] \right\} \\
&\qquad \bigcup \mathbb{F}\text{-span} \left\{ \frac{\prod_{i \in S} Q_i}{Q_{i_0}} \cdot Q_{i_0}(Y,Z) \cdot h_1(Z) \cdot \mathbb{F}[Y] \right\} \\
&\subseteq \mathbb{F}\text{-span} \left\{ \bigcup_{T \in \binom{S}{|S|-1}} \left\{ \prod_{i \in T} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{ML}}^{t} \cdot h_1(Z) \cdot \mathbb{F}[Y] \right\} \right\} \\
&\qquad \bigcup \left\{ \bigcup_{T \in \binom{S}{|S|-1}} \left\{ \prod_{i \in T} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{NonML}}^{\leqslant t\delta} \cdot h_1(Z) \cdot \mathbb{F}[Y] \right\} \right\} \\
&\subseteq \mathbb{F}\text{-span} \left\{ \bigcup_{T \in \binom{[D]}{D-r'}} \left\{ \prod_{i \in T} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{ML}}^{\leqslant r't} \cdot \mathbb{F}[Y] \right\} \right\} \\
&\qquad \bigcup \mathbb{F}\text{-span} \left\{ \bigcup_{T \in \binom{[D]}{D-r'}} \left\{ \prod_{i \in T} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{NonML}}^{\leqslant r'\delta t} \cdot \mathbb{F}[Y] \right\} \right\}
\end{aligned}$$

The last inclusion follows from the fact that $h_1(Z)$ is a linear combination of multilinear monomials of degree at most $(r'-1)t$, and non-multilinear monomials of degree at most $(r'-1)\delta t$. From the discussion above we know that any polynomial in $\partial_Y^{=r'}(T)$ can be expressed as a linear combination of polynomials of the form $\frac{\partial h}{\partial y_u}$. Further every polynomial of the form $\frac{\partial h}{\partial y_u}$ belongs to the set

$$\begin{aligned}
W = \mathbb{F}\text{-span} &\left\{ \bigcup_{T \in \binom{[D]}{D-r'}} \left\{ \prod_{i \in T} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{ML}}^{\leqslant r't} \cdot \mathbb{F}[Y] \right\} \right\} \\
&\bigcup \mathbb{F}\text{-span} \left\{ \bigcup_{T \in \binom{[D]}{D-r'}} \left\{ \prod_{i \in T} Q_i(Y,Z) \cdot \mathbf{z}_{\mathrm{NonML}}^{\leqslant r'\delta t} \cdot \mathbb{F}[Y] \right\} \right\}.
\end{aligned}$$

Thus, we get that $\partial^{=r'}T$ is a subset of W. This completes the proof by induction.

From the afore mentioned discussion, we can now derive the following expressions.

$$\sigma_Y\left(\partial_Y^{=r'}T\right) \subseteq \mathbb{F}\text{-span} \left\{ \bigcup_{S \in \binom{[D]}{D-r'}} \left\{ \prod_{i \in S} \sigma_Y(Q_i) \cdot \mathbf{z}_{\mathrm{ML}}^{\leqslant r't} \right\} \right\}$$

$$\bigcup \mathbb{F}\text{-span}\left\{\bigcup_{S\in\binom{[D]}{D-r'}}\left\{\prod_{i\in S}\sigma_Y(Q_i)\cdot \mathbf{z}_{\text{NonML}}^{\leqslant r'\delta t}\right\}\right\}$$

$$\mathbf{z}^{=m}\cdot\sigma_Y\left(\partial_Y^{=r'}T\right)\subseteq \mathbb{F}\text{-span}\left\{\bigcup_{S\in\binom{[D]}{D-r'}}\left\{\prod_{i\in S}\sigma_Y(Q_i)\cdot \mathbf{z}_{\text{ML}}^{\leqslant m+r't}\right\}\right\}$$

$$\bigcup \mathbb{F}\text{-span}\left\{\bigcup_{S\in\binom{[D]}{D-r'}}\left\{\prod_{i\in S}\sigma_Y(Q_i)\cdot \mathbf{z}_{\text{NonML}}^{\leqslant m+r'\delta t}\right\}\right\}$$

$$\implies \mathbb{F}\text{-span}\left\{\text{mult}\left(\mathbf{z}^{=m}\cdot\sigma_Y\left(\partial_Y^{=r'}T\right)\right)\right\}\subseteq \mathbb{F}\text{-span}\left\{\bigcup_{S\in\binom{[D]}{D-r'}}\left\{\prod_{i\in S}\text{mult}(\sigma_Y(Q_i))\cdot \mathbf{z}_{\text{ML}}^{\leqslant m+r't}\right\}\right\}.$$

Thus we get that $\dim\left(\mathbb{F}\text{-span}\left\{\text{mult}\left(\mathbf{z}^{=m}\cdot\sigma_Y(\partial_Y^{=r'}T)\right)\right\}\right)$ is at most

$$\begin{aligned}
&\dim\left(\mathbb{F}\text{-span}\left\{\bigcup_{S\in\binom{[D]}{D-r'}}\left\{\prod_{i\in S}\text{mult}(\sigma_Y(Q_i))\cdot \mathbf{z}_{\text{ML}}^{\leqslant m+r't}\right\}\right\}\right)\\
\leqslant&\dim\left(\mathbb{F}\text{-span}\left\{\bigcup_{S\in\binom{[D]}{D-r'}}\left\{\prod_{i\in S}\text{mult}(\sigma_Y(Q_i))\right\}\right\}\right)\cdot\dim\left(\mathbb{F}\text{-span}\left\{\mathbf{z}_{\text{ML}}^{\leqslant m+r't}\right\}\right)\\
\leqslant&\binom{D}{D-r'}\cdot\sum_{i=0}^{m+r't}\binom{N}{i}\\
\leqslant&\binom{D}{r'}\cdot\binom{N}{m+r't}\cdot(m+r't)\qquad\text{(Since } m+r't<N/2\text{).}
\end{aligned}$$

□

B Missing Proofs

Proof of Lemma 5. There are $n^{2r'}$ elements in $\mathscr{P}$. It is easy to see that the volume of a Hamming Ball of radius $\Delta_0/2$ for vectors of length $2r'$ is at most $\sum_{i=0}^{\Delta_0}\binom{2r'}{i}\cdot n^i\leqslant\frac{\Delta_0}{2}\binom{2r'}{\Delta_0/2}n^{\Delta_0/2}$ and thus there are at most $\frac{\Delta_0}{2}\binom{2r'}{0.5\Delta_0}n^{0.5\Delta_0}$ many vectors $\mathbf{a}$ in that Hamming ball. Thus, there exists a packing of these Hamming balls in $\mathscr{P}$ with at least $\frac{2n^{2r'-0.5\Delta_0}}{\Delta_0\binom{2r'}{0.5\Delta_0}}$ many balls. □

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il