Lower Bounds for Noncommutative Circuits with Low Syntactic Degree

Pratik Shastri

The Institute of Mathematical Sciences (a CI of Homi Bhabha National Institute), Chennai, India

Abstract

Proving lower bounds on the size of noncommutative arithmetic circuits is an important problem in arithmetic circuit complexity. For explicit n variate polynomials of degree $\Theta(n)$, the best general bound has been stuck at $\Omega(n \log n)$ for over 40 years [Str73; BS83]. Recent work of Chatterjee and Hrubeš [CH23] has provided stronger $(\Omega(n^2))$ bounds for the restricted class of *homogeneous* circuits.

This paper extends these results to a broader class of circuits by using syntactic degree as a complexity measure. The syntactic degree of a circuit measures the extent to which high degree computation is used in the circuit. A homogeneous circuit computing a degree d polynomial can be assumed, without loss of generality, to have syntactic degree exactly equal to d [Fou+24]. We generalize this by considering circuits that are not necessarily homogeneous but have low syntactic degree. Specifically, for an explicit n variate, degree n polynomial f we show that any circuit with syntactic degree $O(n)$ computing f must have size $\Omega(n^{1+c})$ for some constant $c > 0$. We also show that any circuit with syntactic degree $o(n \log n / \log \log n)$ computing the same f must have size $n(\log n)^{\omega(1)}$. Finally, we analyze the circuit size required to compute f based on the *number* of types of gates that appear in the circuit, when gates are typed based on their syntactic degrees. We show that for a circuit if this number is $o(n \log n)$, then it must have size $\omega(n \log n)$.

1 Introduction

Let $\mathbb{F}$ be a field and $X = \{x_1, \ldots, x_n\}$ be a set of variables. A noncommutative polynomial $f(X)$ over $\mathbb{F}$ in the variables X is an $\mathbb{F}$-linear combination of noncommutative monomials in X (equivalently, one may think of these as *words* in $\{x_1, \ldots, x_n\}^*$). The set of all such polynomials forms a ring under addition and noncommutative multiplication, this is the noncommutative polynomial ring denoted by $\mathbb{F}\langle X \rangle$. In noncommutative arithmetic circuit complexity, the central object of study is the noncommutative arithmetic circuit. A noncommutative arithmetic circuit (we will often simply say circuit for short), is a directed acyclic graph whose leaves are labeled by variables (say from X) and constants from $\mathbb{F}$, and whose internal nodes, called gates, are labeled by either $+$ or $\times$. A circuit is said to have fan-in 2 if each gate has in-degree 2. Each gate in the circuit computes a polynomial in the natural way: a $+$ gate computes the sum of the polynomials computed at the children and a $\times$ gate computes the product. Importantly, since the product is noncommutative, each product gate comes with an ordering on the children. Since we will deal with fan-in 2 circuits, each such gate has a designated left child and a designated right child. We will say that a circuit has size s if the underlying graph has s vertices. We will say that a circuit has depth Δ if the longest leaf to root path in the circuit has length Δ. We designate a particular gate of the circuit to be the output gate, and so the polynomial computed by the circuit is defined to be the polynomial computed at this gate.

A central goal in this area is to prove lower bounds on the size of circuits computing explicit[1] polynomials. The most influential work in this area is that of Nisan [Nis91]. He showed exponential lower bounds on the size of *formulas* computing explicit polynomials. Formulas are circuits whose underlying graph is a tree. In order to show this, he introduced a complexity measure, now called Nisan Rank. Unfortunately, his methods do not provide such strong bounds for circuits. Indeed, the best lower bounds on the circuit size of an explicit, n variate degree $\Theta(n)$ polynomial are of the form $\Omega(n \log n)$ [Str73; BS83] and improving upon this has been an outstanding open problem for more than 40 years.

In a recent work [CH23], Chatterjee and Hrubeš improve upon this state of affairs assuming that the circuit is homogeneous. A circuit is said to be homogeneous if every gate in it computes a homogeneous polynomial. In their main result ([CH23], Theorem 14), they construct an n variate, degree $\Theta(n)$ polynomial f such that any *homogeneous* circuit computing f requires size $\Omega(n^2)$.

Note that noncommutative circuits can be *homogenized*, but this procedure incurs a multiplicative blowup of d^2 (where d is the degree of the polynomial) in size and therefore the results in the paper discussed above fall short of giving lower bounds for general (not necessarily homogeneous) circuits.

2 Our Results

In this paper, we remove the homogeneity restriction and replace it with a bound on the *syntactic degree*. Let Ψ be a circuit. For each gate g in Ψ, we define the syntactic degree $d(g)$ inductively as follows: for a leaf g labeled by a field constant, we define $d(g) = 0$. For a leaf labeled by a variable, we define $d(g) = 1$. For a sum gate $g = g_1 + g_2$ we define

[1] We will say that a polynomial $f \in \mathbb{F}\langle X \rangle$ is *explicit* if given a monomial $m \in X^*$, the coefficient of m in f can be computed in time polynomial in the degree of m, assuming field arithmetic over $\mathbb{F}$ is free of cost.

$d(g) = \max\{d(g_1), d(g_2)\}$. If $g = g_1 \times g_2$ is a product gate, we define $d(g) = d(g_1)+d(g_2)$. Note that for each gate g, the degree of the polynomial computed at g is at most $d(g)$. The syntactic degree of Ψ is then defined as the maximum over all gates g of Ψ of $d(g)$.

Notice that if we have a homogeneous circuit Ψ for a degree d polynomial, we can assume *without loss of generality* that the syntactic degree of Ψ is exactly d. This is because in such a circuit, each gate g either computes a polynomial of degree equal to $d(g)$, or it computes the 0 polynomial. This is well known [Fou+24] and easily proved, for example by induction. Therefore, homogeneity *forbids* (useful) computation of polynomials of degree larger than d. One may relax homogeneity by allowing the circuit to have syntactic degree slightly larger than the degree of the output. Such relaxations of homogeneity have been examined in the literature. For instance, in a recent result, the authors of [Fou+24] provided a *quasi*-homogenization result for commutative formulas. A formula is said to be quasi-homogeneous if the syntactic degree of the formula is a polynomially bounded (from above) by the degree of the output.

As mentioned before, Chatterjee and Hrubeš proved an $\Omega(n^2)$ lower bound on the size of any homogeneous noncommutative circuit computing some explicit polynomial on n variables of degree $\Theta(n)$. On the other hand, for general circuits, the best size lower bound for an n variate polynomial of degree $\Theta(n)$ is $\Omega(n \log n)$.

In the absence of stronger lower bounds for general circuits, one can ask the following question: can we prove $\omega(n \log n)$ lower bounds on the size of circuits computing an explicit n variate, degree $\Theta(n)$ polynomial where the circuit is not necessarily homogeneous, but has relatively small syntactic degree? Such circuits may compute intermediate inhomogeneous polynomials of degree higher than the degree of the output, but not much higher. We answer this question in the affirmative.

Remark 2.1. *Circuits with bounded syntactic degree are provably more powerful than their homogeneous counterparts: there exists an n variate, degree $\Theta(n)$ polynomial which can be computed by a circuit of size $n \cdot \mathsf{polylog}(n)$ and syntactic degree $O(n)$ but* any homogeneous circuit *computing it must have size $\Omega(n^2)$. This follows from the work of Chatterjee and Hrubeš (See [CH23], Theorem 14 and the construction in Section 5).*

Let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$. Our results apply to the palindrome polynomial $\mathrm{Pal}_{n,n}(X, Y) \in \mathbb{F}\langle X, Y\rangle$ where $\mathrm{Pal}_{n,d}(X, Y) \in \mathbb{F}\langle X, Y\rangle$ is defined as follows:

$$\mathrm{Pal}_{n,d}(X, Y) = \sum_{(i_1 \ldots, i_d) \in [n]^d} \left(\prod_{j=1}^{d} x_{i_j} \prod_{j=1}^{d} y_{i_{d+1-j}} \right)$$

A version of this polynomial was already introduced by Nisan in [Nis91] to separate circuits from formulas. Indeed, there exists a homogeneous circuit of size $O(n^2)$ that computes $\mathrm{Pal}_{n,n}(X, Y)$. We obtain this from the recursive expression $\mathrm{Pal}_{n,d} = \sum_{i=1}^{n} x_i \mathrm{Pal}_{n,d-1} y_i$. On the other hand, it follows from Nisan's work [Nis91] that any formula computing $\mathrm{Pal}_{n,n}(X, Y)$ requires size $n^{\Omega(n)}$. Since a fan-in two circuit with depth Δ can be converted into a formula of size $\leq 2^{\Delta}$, it follows that any circuit computing $\mathrm{Pal}_{n,n}(X,Y)$ must have depth, and therefore size, $\Omega(n \log n)$. In this paper we improve upon this lower bound for circuits with low syntactic degree. As alluded to before, our main results are the following.

Theorem 2.2. *(Stated below as Corollary 4.3) Let Ψ be a circuit computing $\mathrm{Pal}_{n,n}(X,Y)$ with syntactic degree $O(n)$. Then there exists a constant $c > 0$ such that Ψ has size $\Omega(n^{1+c})$.*

Theorem 2.2 offers a qualitative strengthening of the main theorem in the work of Chatterjee and Hrubeš ([CH23], Theorem 14), since their result applies to homogeneous circuits while ours applies to a provably stronger class of circuits (Remark 2.1).

We can also handle slightly larger syntactic degree, at the expense of obtaining a weaker lower bound (but still better than $\Omega(n \log n)$), as shown in the following theorem.

Theorem 2.3. *(Stated below as Corollary 4.4) Let Ψ be a circuit computing $\mathrm{Pal}_{n,n}(X,Y)$ with syntactic degree $o(n \log n / \log \log n)$. Then Ψ has size $n(\log n)^{\omega(1)}$.*

Finally, we analyze the circuit size required to compute $\mathrm{Pal}_{n,n}$ based on the *number* of different syntactic degrees that appear in the circuit. We show that for a circuit if this number is $o(n \log n)$, then it must have size $\omega(n \log n)$.

Theorem 2.4. *(Stated below as Corollary 4.5) Let Ψ be a circuit computing $\mathrm{Pal}_{n,n}(X,Y)$. Let $d' = |\{d(g) \mid g \text{ is a gate in } \Psi\}|$. If $d' = o(n \log n)$ Ψ has size $\omega(n \log n)$.*

Note that Theorem 2.4 applies to circuits with arbitrarily high syntactic degree. We only require that the *number* of distinct $d(g)$'s appearing in Ψ satisfies the constraints mentioned above. Also, note that that if $d' = \omega(n \log n)$, we trivially get an $\omega(n \log n)$ bound: if a circuit has $\omega(n \log n)$ types of gates then it must have $\omega(n \log n)$ gates. So, the only case where we do not obtain a lower bound asymptotically better than $\Omega(n \log n)$ is when d' and the size of Ψ are *both* $\Theta(n \log n)$. We do not know if such circuits exist for $\mathrm{Pal}_{n,n}$.

3 Proof Idea and Notations

We use the methods of Nisan [Nis91] to obtain our lower bounds. Specifically, we bound the dimension of the coefficient space of an X, Y separated polynomial computed by a circuit of low syntactic degree: Let $f \in \mathbb{F}\langle X, Y\rangle$ be a polynomial. We say f is X, Y separated if in each nonzero monomial of f, every X variable appears before every Y variable. Note that the $\mathrm{Pal}_{n,d}(X,Y)$ polynomials are X, Y separated. Now suppose f is X, Y separated. Thinking of f as a polynomial in the Y variables with coefficients that are elements of $\mathbb{F}\langle X\rangle$, we write

$$f = \sum_{m \in Y^*} c_m m$$

where $c_m \in \mathbb{F}\langle X\rangle$. Define $\mathsf{coeff_m}(f) \triangleq c_m$, and $\mathsf{Coeff}_X(f) \triangleq \{\mathsf{coeff_m}(f) \mid m \in Y^*\}$. We consider the dimension of the span of this set. For $f = \mathrm{Pal}_{n,n}(X,Y)$, this quantity is n^n. On the other hand, Nisan, in [Nis91], showed that if f is computed by a formula of size s, then this dimension is bounded above by $s^{O(1)}$. Since a circuit of size s can be converted into a formula of size $\leq 2^s$, his method applied directly gives a lower bound of $\Omega(n \log n)$ on the size of any circuit computing $\mathrm{Pal}_{n,n}(X,Y)$. We refine this analysis and show that if an X, Y separated polynomial is computed by a circuit with restrictions on the syntactic degrees appearing in it, we may obtain a better *upper* bound on the dimension of the span of the coefficients, by explicitly constructing a spanning set of polynomials. This gives the lower bounds. For future convenience, we set up some more notation. Let Ψ be a noncommutative circuit. Recall that for a gate g in Ψ, $d(g)$ denotes its syntactic degree. Define $D_\Psi \triangleq \{d(g) \mid g \text{ is a gate in } \Psi\}$. We say that a circuit Ψ is in *normal form* if no product gate has a child of syntactic degree 0. Since gates with syntactic degree 0 can only compute constants, we may assume without loss of generality, by pushing constants all the way down to the leaves, that a given circuit is in normal form. We allow leaves labeled by αx where $\alpha \in \mathbb{F}$ and x is a variable.

4 Lower Bound

In this section, we prove our main result. Theorems 2.2 and 2.3 follow as corollaries of Theorem 4.1, which we prove below.

Theorem 4.1. *Let Ψ be a size s, fan-in 2 circuit computing $\mathrm{Pal}_{n,d}(X,Y)$. Let $|D_\Psi| = d'$. Then, $s \geq (d'-1)n^{(d-2)/(d'-1)} - 2d' + 2$.*

Proof. For simplicity, we assume without loss of generality that Ψ is in normal form. For each gate $g \in \Psi$, we let $\hat{g}$ denote the polynomial computed at g in Ψ. For each gate g, we write

$$\hat{g} = \hat{g}_1 + \hat{g}_X + \hat{g}_Y + \hat{g}_{XY} + \hat{g}_{\text{other}}$$

where $\hat{g}_1$ denotes the constant term of $\hat{g}$, $\hat{g}_X$ denotes the sum (with coefficient) of non-constant monomials of $\hat{g}$ which contain only X variables, $\hat{g}_Y$ denotes the sum (with coefficient) of non-constant monomials of $\hat{g}$ which contain only Y variables, $\hat{g}_{XY}$ denotes the sum (with coefficient) of non-constant monomials of $\hat{g}$ in which all X variables appear *before* all the Y variables, and $\hat{g}_{\text{other}}$ denotes the sum (with coefficient) of all the other monomials.

Let $D_\Psi = \{l_1, \ldots, l_{d'}\}$ with $1 = l_1 < \ldots < l_{d'}$. For each $k \in [d']$, define $\mathcal{G}_k \triangleq \{g \mid g \text{ is a gate in } \Psi \text{ with } d(g) = l_k\}$. For each $k \in [d']$, we will build a (reasonably small) set $\mathcal{B}_k \subseteq \mathbb{F}\langle X\rangle$ of spanning polynomials such that for each $g \in \mathcal{G}_k$, $\mathsf{Coeff}_X(\hat{g}_1)$, $\mathsf{Coeff}_X(\hat{g}_X)$, $\mathsf{Coeff}_X(\hat{g}_Y)$, $\mathsf{Coeff}_X(\hat{g}_{XY}) \subseteq \mathsf{span}\{\mathcal{B}_k\}$ and for each $t < k, \mathcal{B}_t \subseteq \mathcal{B}_k$. We build these sets iteratively. Our base cases will be $k = 0$ with $\mathcal{B}_0 = \{1\}$ and $k = 1$ with $\mathcal{B}_1 = \{1, x_1, \ldots, x_n\}$ and $|\mathcal{B}_1| = n + 1$. Now suppose we have already constructed $\mathcal{B}_l$ for each $l < k$. First, set $\mathcal{B}_k \leftarrow \mathcal{B}_{k-1}$. Let $\mathcal{P}_k \subseteq \mathcal{G}_k$ be the set of product gates g in Ψ

with $d(g) = l_k$. For each such gate $g = g' \times g''$ with $d(g') = l_t, d(g'') = l_m$, we note that $t, m < k$ since Ψ is normal. Observe the following:

$$
\begin{aligned}
\hat{g}_1 &= \hat{g'}_1 \times \hat{g''}_1 \\
\hat{g}_X &= \hat{g'}_X \times \hat{g''}_X + \hat{g'}_X \times \hat{g''}_1 + \hat{g'}_1 \times \hat{g''}_X \\
\hat{g}_Y &= \hat{g'}_Y \times \hat{g''}_Y + \hat{g'}_Y \times \hat{g''}_1 + \hat{g'}_1 \times \hat{g''}_Y \\
\hat{g}_{XY} &= \hat{g'}_X \times \hat{g''}_Y + \hat{g'}_{XY} \times \hat{g''}_Y + \hat{g'}_X \times \hat{g''}_{XY} + \hat{g'}_{XY} \times \hat{g''}_1 + \hat{g'}_1 \times \hat{g''}_{XY}
\end{aligned}
$$

For each such $g \in \mathcal{P}_k$ (with $g = g' \times g''$), we update $\mathcal{B}_k$ as $\mathcal{B}_k \leftarrow \mathcal{B}_k \cup \{\hat{g}_X\} \cup \{\hat{g'}_X \times h \mid h \in \mathcal{B}_{k-1}\}$.

Claim 4.2. *For $k \in [d']$ and each gate $g \in \mathcal{G}_k$, we have that* $\mathsf{Coeff}_X(\hat{g}_1)$, $\mathsf{Coeff}_X(\hat{g}_X)$, $\mathsf{Coeff}_X(\hat{g}_Y)$, $\mathsf{Coeff}_X(\hat{g}_{XY}) \subseteq \mathsf{span}\{\mathcal{B}_k\}$

Proof. We show this by inducting on the depth of the gate g.

- If g is a leaf, it satisfies the claim by construction of $\mathcal{B}_0$ and $\mathcal{B}_1$.
- *g is a product gate:* Let $g = g' \times g''$ with $d(g) = l_k$. As mentioned earlier, if $d(g') = l_t, d(g'') = l_m$ then $t, m < k$ since Ψ is normal.
  1. $\mathsf{Coeff}_X(\hat{g}_1) \subseteq \mathsf{span}\{\mathcal{B}_k\}$: This is true since $\mathsf{Coeff}_X(\hat{g}_1)$ just contains a constant, $\mathcal{B}_1 \subseteq \mathcal{B}_k$ and $\mathcal{B}_1$ contains the constant 1.
  2. $\mathsf{Coeff}_X(\hat{g}_X) \subseteq \mathsf{span}\{\mathcal{B}_k\}$: This is true since $\mathsf{Coeff}_X(\hat{g}_X) = \{\hat{g}_X\}$ and $\hat{g}_X$ is in $\mathcal{B}_k$
  3. $\mathsf{Coeff}_X(\hat{g}_Y) \subseteq \mathsf{span}\{\mathcal{B}_k\}$: Again, this is true since $\mathsf{Coeff}_X(\hat{g}_Y)$ just contains a constant, $\mathcal{B}_1 \subseteq \mathcal{B}_k$ and $\mathcal{B}_1$ contains the constant 1.
  4. $\mathsf{Coeff}_X(\hat{g}_{XY}) \subseteq \mathsf{span}\{\mathcal{B}_k\}$: Consider a monomial $m \in \mathbb{F}\langle Y\rangle$ with coefficient $\mathsf{coeff}_m(\hat{g}_{XY}) \in \mathbb{F}\langle X\rangle$. We write $\mathsf{coeff}_m(\hat{g}_{XY}) = \mathsf{coeff}_m(\hat{g'}_X \times \hat{g''}_Y) + \mathsf{coeff}_m(\hat{g'}_{XY} \times \hat{g''}_Y) + \mathsf{coeff}_m(\hat{g'}_X \times \hat{g''}_{XY}) + \mathsf{coeff}_m(\hat{g'}_{XY} \times \hat{g''}_1) + \mathsf{coeff}_m(\hat{g'}_1 \times \hat{g''}_{XY})$. It suffices to show that each term in this expansion is individually in $\mathsf{span}\{\mathcal{B}_k\}$. Observe that $\mathsf{coeff}_m(\hat{g'}_X \times \hat{g''}_Y)$ is a scalar multiple of $\hat{g'}_X$ which is in $\mathcal{B}_k$ ($1 \in \mathcal{B}_{k-1}$ and so $\hat{g'}_X \times 1 \in \mathcal{B}_k$). Next, $\mathsf{coeff}_m(\hat{g'}_{XY} \times \hat{g''}_Y)$ is a linear combination of coefficients of the form $\mathsf{coeff}_{m'}(\hat{g'}_{XY})$ where m' is a prefix of m. By induction, for each such m' we have $\mathsf{coeff}_{m'}(\hat{g'}_{XY}) \in \mathsf{span}\{\mathcal{B}_t\} \subseteq \mathsf{span}\{\mathcal{B}_k\}$ and so $\mathsf{coeff}_m(\hat{g'}_{XY} \times \hat{g''}_Y) \in \mathsf{span}\{\mathcal{B}_k\}$. Further, $\mathsf{coeff}_m(\hat{g'}_X \times \hat{g''}_{XY})$ is just $\hat{g'}_X \times \mathsf{coeff}_m(\hat{g''}_{XY})$. Since $\mathsf{coeff}_m(\hat{g''}_{XY}) \in \mathsf{span}\{\mathcal{B}_m\} \subseteq \mathsf{span}\{\mathcal{B}_{k-1}\}$ by induction, $\mathsf{coeff}_m(\hat{g'}_X \times \hat{g''}_{XY}) \in \mathsf{span}\{\mathcal{B}_k\}$ by construction. Finally, by induction, $\mathsf{coeff}_m(\hat{g'}_{XY})$ and $\mathsf{coeff}_m(\hat{g''}_{XY})$ are in $\mathrm{span}\{\mathcal{B}_t\} \subseteq \mathrm{span}\{\mathcal{B}_k\}$ and $\mathrm{span}\{\mathcal{B}_m\} \subseteq \mathrm{span}\{\mathcal{B}_k\}$ respectively, and therefore so are $\mathsf{coeff}_m(\hat{g'}_{XY} \times \hat{g''}_1)$ and $\mathsf{coeff}_m(\hat{g'}_1 \times \hat{g''}_{XY})$.

- *g is a sum gate:* Let $g = g' + g''$ with $d(g) = l_k$, $d(g') = l_k$ and $d(g'') = l_t$ with $t \leq k$. We assume the claim for g' and g'', by induction on depth. That is, we assume that $\mathsf{Coeff}_X(\hat{g'}_1), \mathsf{Coeff}_X(\hat{g'}_X), \mathsf{Coeff}_X(\hat{g'}_Y), \mathsf{Coeff}_X(\hat{g'}_{XY}) \subseteq \mathsf{span}\{\mathcal{B}_k\}$ and $\mathsf{Coeff}_X(\hat{g''}_1)$, $\mathsf{Coeff}_X(\hat{g''}_X)$, $\mathsf{Coeff}_X(\hat{g''}_Y)$, $\mathsf{Coeff}_X(\hat{g''}_{XY}) \subseteq \mathsf{span}\{\mathcal{B}_t\} \subseteq \mathsf{span}\{\mathcal{B}_k\}$. The claim then follows from the fact that $\hat{g}_1 = \hat{g'}_1 + \hat{g''}_1$, $\hat{g}_X = \hat{g'}_X + \hat{g''}_X$, $\hat{g}_Y = \hat{g'}_Y + \hat{g''}_Y$ and $\hat{g}_{XY} = \hat{g'}_{XY} + \hat{g''}_{XY}$.

This finishes the proof Claim 4.2 $\square$

Let us now bound $|\mathcal{B}_{d'}|$. For each $k \in [d']$, define $s_k \triangleq |\mathcal{P}_k|$. Note that $\sum_{k=0}^{d'} s_k$ is precisely s, the size of Ψ. Observe that for each $k \geq 2$, we have by construction that $|\mathcal{B}_k| \leq s_k(|\mathcal{B}_{k-1}| + 1) + |\mathcal{B}_{k-1}| \leq (s_k + 1)(|\mathcal{B}_{k-1}| + 1)$. Define, for each $1 \leq k \leq d'$, $g_k \triangleq |\mathcal{B}_k| + 1$. This fixes $g_1 = n + 2$. Applying the inequality above, we get that for each $2 \leq k \leq d'$, $g_k \leq (s_k + 1)g_{k-1} + 1 \leq (s_k + 2)g_{k-1}$. Therefore, we see that $g_{d'} = |\mathcal{B}_{d'}| + 1 \leq (n+2)\left(\prod_{k=2}^{d'}(s_k + 2)\right)$.

Applying the AM-GM inequality, $|\mathcal{B}_{d'}| \leq (n+2)\left(\dfrac{s + 2d' - 2}{d' - 1}\right)^{d'-1}$. Since Ψ computes $\mathrm{Pal}_{n,d}$, we must also have that $|\mathcal{B}_{d'}| \geq n^d$. Therefore,

$$
\begin{aligned}
n^{d/(d'-1)} &\leq (n+2)^{1/(d'-1)}\left(\frac{s + 2d' - 2}{d' - 1}\right) \implies \\
s &\geq (d'-1)n^{(d/(d'-1)) - (\log(n+2)/((d'-1)\log n))} - 2d' + 2 \\
&\geq (d'-1)n^{(d-2)/(d'-1)} - 2d' + 2
\end{aligned}
$$

This finishes the proof of Theorem 4.1. $\square$

Corollary 4.3 (Restatement of Theorem 2.2)**.** *Let Ψ be a circuit computing the polynomial $\mathrm{Pal}_{n,n}(X,Y)$ with syntactic degree $O(n)$. Then there exists a constant $c > 0$ such that Ψ has size $\Omega(n^{1+c})$.*

Proof. Let $d' = |D_\Psi|$. Suppose the syntactic degree of Ψ is $\leq \alpha n$ for some constant α. Note that $\alpha \geq 2$ since the degree of $\mathrm{Pal}_{n,n}$ is $2n$. Also, note that by definition, $d' \leq \alpha n$ because d' is the number of distinct syntactic degrees appearing in Ψ and also that $d' \geq \log(2n) > 1$, because the degree of $\mathrm{Pal}_{n,n}(X,Y)$ is $2n$ and one may find a root to leaf path in the circuit on which the syntactic degree drops by a factor of at most half at each step. We split into two cases.

- *Case 1). $d' \leq n/2$:* In this case, we apply Theorem 4.1 and find that $s \geq (d'-1)n^2 - 2d' + 2 = \Omega(n^2 \log n)$.

- *Case 2). $n/2 < d' \leq \alpha n$:* In this case, we apply Theorem 4.1 and find that $s \geq (n/2)n^{(n-2)/(\alpha n - 1)} - 2d' + 2 = \Omega(n^{1+1/\alpha})$.

Since $\alpha \geq 2$, we have $s = \Omega(n^{1+1/\alpha})$ in both cases. $\square$

When the syntactic degree is slightly higher, we have the following corollary.

Corollary 4.4 (Restatement of Theorem 2.3)**.** *Let Ψ be a circuit computing the polynomial $\mathrm{Pal}_{n,n}(X,Y)$ with syntactic degree $o(n \log n / \log\log n)$. Then Ψ has size $\omega(n \log n)$.*

Proof. Again, we split into two cases.

- *Case 1). $d' \leq n/2$:* In this case, as before, we get $s = \Omega(n^2 \log n)$.
- *Case 2). $n/2 < d'$:* In this case, from Theorem 4.1 we get $s \geq (n/2)n^{\omega(\log\log n/\log n)}$ $\implies s \geq n(\log n)^{\omega(1)}$.

This finishes the proof. $\square$

Let us now give a more precise analysis based on the the number of distinct syntactic degrees appearing in the circuit, i.e., on the size of $|D_\Psi|$.

Corollary 4.5. *[Restatement of Theorem 2.4] Let Ψ be a circuit computing the polynomial $\mathrm{Pal}_{n,n}(X,Y)$. If $|D_\Psi| = o(n \log n)$, then Ψ has size $\omega(n \log n)$.*

Proof. Let the size of Ψ be s. Suppose $d' = d'(n) = |D_\Psi| = o(n \log n)$. Define $g(n) \triangleq n \log n / d'(n)$. Note that $g(n) = \omega(1)$. Applying Theorem 4.1, we find that

$$\begin{aligned} s &\geq (d'/2) \times n^{n/2d'} - O(d') \\ &\geq (d'/2) \times n^{n/2d'} - O(d') \\ &\geq \frac{n \log n}{2g(n)} \times 2^{g(n)/2} - o(n \log n) \\ &\geq \omega(n \log n) \end{aligned}$$

This finishes the proof.

$\square$

5 Discussion

In this paper, we proved $\omega(n \log n)$ lower bounds for noncommutative circuits of low syntactic degree. The question of proving such lower bounds for general circuits remains open. One possible avenue of attack for this problem is inventing a size preserving procedure for reducing a circuit's syntactic degree. As a concrete question, suppose we are given a circuit of size s and syntactic degree $d = \omega(n)$ that computes $\mathrm{Pal}_{n,n}$. Can we construct another circuit computing $\mathrm{Pal}_{n,n}$ with syntactic degree $O(n)$ and size $s \times n^{o(1)}$? An affirmative answer to this question, combined with our work, would immediately yield strong lower bounds for *general* circuits. Reducing syntactic degree is no harder (and perhaps easier) than homogenization, so one may expect a more efficient way to do it than simple homogenization. Another question that stems from this work is whether we can handle separately the case when the number of distinct syntactic degrees appearing in a circuit computing $\mathrm{Pal}_{n,n}$ is $\Theta(n \log n)$.

References

[BS83] Walter Baur and Volker Strassen. "The complexity of partial derivatives". In: *Theoretical Computer Science* 22.3 (1983), pp. 317–330. ISSN: 0304-3975. DOI: https://doi.org/10.1016/0304-3975(83)90110-X. URL: https://www.sciencedirect.com/science/article/pii/030439758390110X.

[CH23] Prerona Chatterjee and Pavel Hrubeš. "New Lower Bounds Against Homogeneous Non-Commutative Circuits". In: *38th Computational Complexity Conference (CCC 2023)*. Ed. by Amnon Ta-Shma. Vol. 264. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 13:1–13:10. ISBN: 978-3-95977-282-2. DOI: 10.4230/LIPIcs.CCC.2023.13. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CCC.2023.13.

[Fou+24] Hervé Fournier et al. "On the Power of Homogeneous Algebraic Formulas". In: *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. STOC 2024. Vancouver, BC, Canada: Association for Computing Machinery, 2024, pp. 141–151. ISBN: 9798400703836. DOI: 10.1145/3618260.3649760. URL: https://doi.org/10.1145/3618260.3649760.

[Nis91] Noam Nisan. "Lower Bounds for Non-Commutative Computation". In: *Proceedings of the Twenty-Third Annual ACM Symposium on Theory of Computing, May 6-8, 1991, New Orleans, Louisiana, USA*. Ed. by Shafi Goldwasser. ACM, 1991, pp. 410–418. ISBN: 0-89791-397-3. DOI: 10.1145/103418.103463. URL: https://doi.org/10.1145/103418.103463.

[Str73] Volker Strassen. "Die Berechnungskomplexität von elementarsymmetrischen Funktionen und von Interpolationskoeffizienten". In: *Numerische Mathematik* 20.3 (June 1973), pp. 238–251. ISSN: 0945-3245. DOI: 10.1007/BF01436566. URL: https://doi.org/10.1007/BF01436566.