On Lower Bounds for Multiplicative Circuits and Linear Circuits in Noncommutative Domains

V. Arvind[1], S. Raja[1], and A.V. Sreejith[2]

[1] The Institute of Mathematical Sciences (IMSc), Chennai, India
{arvind,rajas}@imsc.res.in
[2] Tata Institute of Fundamental Research (TIFR), Mumbai, India
sreejith@imsc.res.in

Abstract. In this paper we show some lower bounds for the size of multiplicative circuits computing multi-output functions in some *noncommutative* domains such as monoids and finite groups. We also introduce and study a generalization of linear circuits in which the goal is to compute MY where Y is a vector of indeterminates and M is a matrix whose entries come from *noncommutative* rings. We show some lower bounds in this setting as well.

1 Introduction

Let $(S, \circ)$ be a semigroup, i.e., S is a set closed under the binary operation $\circ$ which is associative. A natural multi-output computational model is a circuit over $(S, \circ)$. The circuit is given by a directed acyclic graph with input nodes labeled $x_1, ..., x_n$ of indegree 0 and output nodes $y_1, ..., y_m$ of outdegree 0.

The gates of the circuit all compute the monoid product. We assume that all gates have fanin 2. The size of the circuit is the number of nodes in it and it computes a function $f : S^n \to S^m$.

This provides a general setting to some well studied problems in circuit complexity. For example:

(1) If $S = \mathbb{F}_2$ and $\circ$ is addition in $\mathbb{F}_2$, the problem is one of computing $A\mathbf{x}$ for an $m \times n$ matrix over $\mathbb{F}_2$. The problem of giving an explicit A such that the size of any circuit for it is superlinear is a longstanding open problem. By means of counting arguments, we know that there exist such matrices A [11].

This problem has a rich literature with many interesting developments. Morgenstern [7] showed an $\Omega(n \log n)$ lower bound for the Hadamard matrix in the bounded coefficient model when $\mathbb{F} = \mathbb{C}$. Valiant [11] developed matrix rigidity as a means to attack the problem in the case of logarithmic depth circuits. In spite of many interesting results and developments, superlinear size lower bounds remain elusive over any field $\mathbb{F}$ even for the special case of log-depth circuits (Lokam's monograph [6] contains most of the recent results).

(2) When $S = \{0, 1\}$ and $\circ$ is the boolean OR, this problem is also well studied and due to its monotone nature it has explicit lower bounds of circuit size $n^{2-o(1)}$ (e.g. see section 3.4 in [3]).

E.A. Hirsch et al. (Eds.): CSR 2014, LNCS 8476, pp. 65–76, 2014.

A more restricted form is $S = (\mathbb{N}, +)$ called SUM circuits also well studied e.g. [3]. While for monotone settings (OR,SUM circuits) there are nontrivial lower bounds, in the commutative case for S we do not have strong lower bounds results. In this paper, we explore the case when $(S, \circ)$ is noncommutative and manage to prove strong lower bounds in some cases.

An interesting aspect is that the number of inputs can be restricted to just two: x_0, x_1. The explicit functions y_i, $1 \leq i \leq m$ are defined as words $y_i = y_{i1}y_{i2}...y_{in}$ where $y_{ij} \in \{x_0, x_1\}$ and $\{y_1, y_2, ..., y_m\}$ are explicitly defined. We show that any circuit $C : \{x_0, x_1\} \rightarrow \{y_1, y_2, ..., y_m\}$ is of size $\Omega(\frac{mn}{\log^2 n})$ in the following four settings:

1. When $(S, \circ)$ is the free monoid X^* for X such that $|X| \geq 2$.
2. When $(S, \circ)$ is the finite matrix semigroup over the boolean ring and matrices are of dimension $n^c \times n^c$ for some constant $c > 0$.
3. When $(S, \circ)$ is the free group G_X generated by $X = \{x_1, x_2, x_1^{-1}, x_2^{-1}\}$.
4. When $(S, \circ)$ is the permutation group where $S = S_N$ for $N = n^d$ for some constant $d > 0$.

In Section 6, we show lower bounds for a generalization of linear circuits model. In this model we allow coefficients to come from *noncommutative rings.*

2 Circuits over Free Monoids

We consider the free monoid X^* where X is a finite alphabet and the monoid operation is concatenation with the empty string ϵ as identity. The notion of a multiplicative circuits over a free monoid is also known in th area of data compression as a straight line program [5].

Notice that when X is a singleton set $X = \{1\}$ then $(1^*, \circ)$ is essentially the semigroup $(\mathbb{N}, +)$. We consider the simplest noncommutative setting with $X = \{0, 1\}$. In the problem, we consider circuits that take the "generating set" X as input and the m outputs $y_1, y_2, ..., y_m \in X^n$ (where n is the "input" parameter).

Since each y_i is of length n, clearly n gates are sufficient to compute each y_i and hence $O(mn)$ is an obvious upper bound for the circuit size. We will give an explicit set $y_1, y_2, ..., y_m \in \{0, 1\}^n$ so that $\Omega(\frac{mn}{\log^2 n})$ is the circuit size lower bound. We will let $m = n$ in the construction and it can be suitably generalized to larger values of m. We now explain the construction of the set $S = \{y_1, y_2, ..., y_m\} \subseteq \{0, 1\}^n$.

Construction of S

Consider the set $[n^2]$ of the first n^2 natural numbers. Each $i \in [n^2]$ requires $2\log_2 n$ bits to represent in binary. Initially let $D = [n^2]$.

```
for $i = 1, ..., n$ do
```

```
    pick the first $\lceil \frac{n}{2\log n} \rceil$ numbers from the current $D$, concatenate their binary
      representations to obtain $y_i$ and remove these numbers from $D$.
end for
```

This defines the set $S = \{y_1, y_2, ..., y_n\}$. Each y_i constructed has the property that y_i has $\geq \frac{n}{2\log n}$ distinct substrings of length $2\log n$. We show the following two result about these strings:

- For each $y_i \in S$ any concatenation circuit that generates y_i from input $X = \{0, 1\}$ requires size $\Omega(\frac{n}{\log^2 n})$.
- Any concatenation circuit that takes $X = \{0, 1\}$ as input and outputs $S = \{y_1, y_2, ..., y_n\}$ at n output gates requires size $\Omega(\frac{n^2}{\log^2 n})$.

Lemma 1. *Let $s \in X^n$ be any string where $|X| \geq 2$, such that the number of distinct substrings of s of length l is N. Then any concatenation circuit for s will require $\Omega(\frac{N}{l})$ gates.*

Proof. Let C be any circuit that computes the string s. Now each gate g of C computes some string s_g. Suppose $g = g_1 \circ g_2$ is a gate whose inputs are gates g_1, g_2.

Suppose s_{g_1} has k_1 distinct substrings of length l and s_{g_2} has k_2 distinct substrings of length l. Now, in s_g notice that the *new* substrings of length l (not occurring in s_{g_1} or s_{g_2}) could only arise as a concatenation of some suffix of s_{g_1} and prefix of s_{g_2} such that neither of them is the empty string. The number of such substrings is at most $l - 1$.

Hence, s_g can have at most $k_1 + k_2 + l - 1$ distinct substrings of length l. Thus, each new gate of C can generate at most $l - 1$ new substrings of length l. Since the output string s has N distinct length l substrings, it follows that number of gates in C is $\Omega(\frac{N}{l})$. □

Note the case not covered by the lemma: $|X| = 1$. In that case we know that every string of length n (for every n) has a concatenation circuit of size $\leq 2\log_2 n$ and the circuit exploits the fact that for each length l there is a unique string. Similar to Lemma 1 is known earlier (e.g. see Lemma 3 in [2]).

Theorem 1. *Let $S \subseteq \{0, 1\}^n$ be the explicit set of n strings defined above. Any concatenation circuit that takes $X = \{0, 1\}$ as input and outputs S at its n output gates will require size $\Omega(\frac{n^2}{\log^2 n})$.*

Proof. Let $S = \{y_1, y_2, ..., y_n\}$ as defined above. Notice that, each y_i can be generated by size n circuit. Let C be any concatenation circuit that takes $X = \{0, 1\}$ as inputs and at its n output gates generates $y_1, y_2, ..., y_n$ respectively. Let $C^{'}$ be a concatenation circuit obtained from C by adding $n - 1$ new gates such that $C^{'}$ outputs the concatenation $y = y_1 y_2 ... y_n$. By construction $size(C^{'}) = size(C) + n - 1$. The number of distinct length $2\log n$ strings in the string y is, by construction, $\geq \frac{n^2}{2\log n}$. This is because each y_i has $\geq \frac{n}{2\log n}$ distinct substrings and these are disjoint for different y_i. Hence by Lemma 1, $size(C^{'}) = \Omega(\frac{n^2}{\log^2 n})$ which implies $size(C) = \Omega(\frac{n^2}{\log^2 n})$. □

3 Circuits over Matrix Semigroups

The setting now is that of a finite monoid $(M, \circ)$ where M consisting of $p(n) \times p(n)$ matrices whose entries come from the Boolean semiring $\{0, 1, \vee, \wedge\}$. We will modify the lower bound of the previous section to make it work over $(M, \circ)$ which is a finite monoid.

Recall we constructed $S = \{y_1, y_2, ..., y_n\} \subseteq \{0, 1\}^n$. Let D_l be the set of all length l substrings of each $y_i \in S$. Let $D = \bigcup_{l=0}^{n} D_l$. Clearly $|D| = \sum_{l=0}^{n} |D_l| \leq n^3$. The matrices in M are $|D| \times |D|$. We now define two functions $f_0, f_1 : D \to D$ corresponding to the generating set $X = \{0, 1\}$ of the free monoid. For $b \in \{0, 1\}$, define

$$f_b(s) = \begin{cases} s \circ b & s \circ b \in D \\ \epsilon & \text{otherwise} \end{cases}$$

where $s \in \{0, 1\}^*$. These give rise to two matrices M_b, $b \in \{0, 1\}$. The rows and columns of M_b are indexed by elements of D and $M_b(s, s \circ b) = 1$ if $s \circ b \in D$ and $M_b(s, s^{'}) = 0$ if $s \circ b \neq s^{'}$. If $s \circ b \notin D$ then $M_b(s, \epsilon) = 1$ and $\forall s^{'} \neq \epsilon$, $M_b(s, s^{'}) = 0$.

Thus, we have defined a morphism, $\Phi : (X^*, \circ) \to (M, \circ)$ which maps $b \to M_b$, $b \in \{0, 1\}$ and by natural extension maps a string $s \in X^*$ to M_s. In particular, the set $S = \{y_1, y_2, ..., y_n\}$ defined in section 2 is mapped to $\hat{S} = \{M_{y_1}, M_{y_2}, ..., M_{y_n}\}$.

Theorem 2. *Any circuit over $(M, \circ)$ that takes M_0, M_1 as input and computes $\{M_{y_i} | y_i \in S\}$ at its n output gates is of size $\Omega(\frac{n^2}{\log^2 n})$.*

Proof. Let C be a circuit over $(M, \circ)$ computing M_{y_i}, $1 \leq i \leq n$ at the n output gates and input M_0, M_1. Consider the corresponding circuit $C^{'}$ over the free monoid X^* with input $X = \{0, 1\}$. Let g_i be the output gate of C computing M_{y_i}, $1 \leq i \leq n$. In $C^{'}$ let $w_i \in X^*$ be the word computed at g_i. We know that $M_{w_i} = M_{y_i}$ for $1 \leq i \leq n$. That means $M_{w_i}(\epsilon, y_i) = 1$. By definition of the matrices M_b, the only way this can happen is when $w_i = y_i \circ z_i$ for some $z_i \in X^*$ for each i. Now, let $C^{''}$ be a new circuit obtained from $C^{'}$ that outputs the concatenation of $w_1, w_2, ..., w_n$ in that order. Then $size(C^{''}) \leq size(C^{'}) + n - 1$. The output string by $C^{''}$ is of the form $y_1 \circ z_1 \circ y_2 \circ z_2 \circ ... \circ y_n \circ z_n$. Since the number of distinct substrings of length $2 \log n$ in $\{y_1, y_2, ..., y_n\}$ we know is $\geq \frac{n^2}{\log^2 n}$, it follows by Lemma 1 that $size(C^{''}) = \Omega(\frac{n^2}{\log^2 n})$. Consequently, $size(C) = size(C^{'}) = \Omega(\frac{n^2}{\log^2 n})$. This completes the proof. □

4 Circuits over Free Groups

We consider the free group G_X generated by the set $X = \{x_1, x_2, x_1^{-1}, x_2^{-1}\}$ consisting of x_1, x_2 and their inverses. The group operation is concatenation with the empty string ϵ as identity and the only cancellation rules we can repeatedly use are $x_i x_i^{-1} = x_i^{-1} x_i = \epsilon$ for $i \in \{1, 2\}$. Given a word $w \in X^*$ we can repeatedly apply these rules and obtain a *normal form* $w' \in G_X$ from it which cannot be simplified further. This normal form, by Church-Rosser property, is unique and independent of how we apply the rules.

Recall the set of binary strings we constructed in Section 2. Replacing 0 by x_1 and 1 by x_2 we obtain $S = \{y_1, y_2, ..., y_n\} \subseteq \{x_1, x_2\}^n \subseteq G_X$. Each word y_i constructed has the property that y_i has $\geq \frac{n}{2\log n}$ distinct subwords of length $2\log n$. These words are already in their normal forms.

Lemma 2. *Let $w \in G_X$ be any word where $X = \{x_1, x_2, x_1^{-1}, x_2^{-1}\}$, such that the number of distinct subwords of length l in its normal form w' is N. Then any concatenation circuit for w will require size $\Omega(\frac{N}{l})$ gates.*

Proof. Let C be any circuit that computes the word w. Now each gate g of C computes some word w_g and, as above, $w_g^{'}$ denotes its normal form.

Suppose $g = g_1 \circ g_2$ is a gate whose inputs are gates g_1, g_2. Then, by the Church-Rosser property of cancellations, the normal form for w_g satisfies

$$w_g' = (w_{g_1}' \circ w_{g_2}')'.$$

Suppose $w_{g_1}^{'}$ has k_1 distinct subwords of length l and $w_{g_2}^{'}$ has k_2 distinct subwords of length l. Now, in $w_g^{'}$ notice that the *new* subwords of length l (not occurring in $w_{g_1}^{'}$ or $w_{g_2}^{'}$) could only arise as a concatenation of some suffix of word $w_{g_1}^{'}$ and prefix of word $w_{g_2}^{'}$ such that neither of them is the empty string. The number of such new subwords is at most l. Hence, w_g' can have at most $k_1 + k_2 + l$ distinct subwords of length l.

Now, since the normal form w' for the output word w has N distinct length l subwords, it follows that number of gates in C is $\Omega(\frac{N}{l})$. □

Theorem 3. *Let $S \subseteq \{x_1, x_2\}^n \subseteq G_X$ be the explicit set of n words defined above. Any concatenation circuit that takes $X = \{x_1, x_2, x_1^{-1}, x_2^{-1}\}$ as input and outputs S at its n output gates will require size $\Omega(\frac{n^2}{\log^2 n})$.*

Proof. Let $S = \{y_1, y_2, ..., y_n\}$ as defined above and let C be any concatenation circuit that takes $X = \{x_1, x_2, x_1^{-1}, x_2^{-1}\}$ as inputs and at its n output gates generates $y_1, y_2, ..., y_n$ respectively. Let $C^{'}$ be a concatenation circuit obtained from C by adding $n-1$ new gates such that $C^{'}$ outputs the concatenation $y = y_1 y_2 ... y_n$. By construction $size(C^{'}) = size(C) + n - 1$. The number of distinct length $2\log n$ words in the words y is, by construction, $\geq \frac{n^2}{2\log n}$. This is because each y_i has $\geq \frac{n}{2\log n}$ distinct subwords and these are disjoint for different y_i. Hence by Lemma 2, $size(C^{'}) = \Omega(\frac{n^2}{\log^2 n})$ which implies $size(C) = \Omega(\frac{n^2}{\log^2 n})$. □

Remark 1. Let $\mathbf{M_0} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $\mathbf{M_1} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ be 2×2 matrices. Consider the infinite group G generated by these elements and their inverses over the field of rationals $\mathbb{Q}$. It is well known (e.g. see [4] for a nice complexity theoretic application) that the group G is isomorphic to the free group G_X, where the isomorphism is defined by $x_1 \rightarrow M_0$ and $x_2 \rightarrow M_1$. It follows that Theorem 3 also applies to the group G by setting $x_1 = M_0$ and $x_2 = M_1$.

5 Circuits over Permutation Groups

We now present a lower bound in the setting of finite groups. We will transform our free monoid construction to this setting. Recall the set of binary strings S we constructed in Section 2. To this end, we will define two permutations $\pi_0, \pi_1 \in S_N$ (where $N = poly(n)$ will be defined later). These permutations correspond to $X = \{0, 1\}$ and by multiplication the target output permutations are defined:
$G_S = \{\pi_{y_i} = \Pi_{j=1}^{n} \pi_{y_i[j]} | y_i \in S\}$, where $y_i[j]$ is the j-th bit of string y_i.

Definition of $\boldsymbol{\pi_0, \pi_1}$:

We pick r primes $p_1, p_2, ..., p_r$ where $r = n^2$ such that $n < p_1 < p_2 < ... < p_r < n^4$. The permutation π_0 is defined as the product of $r+1$ disjoint cycles, $\pi_0 = C_0.C_1...C_r$ where C_0, C_1 are of length p_1 and for $i \geq 2$, C_i is of length p_i. Similarly, $\pi_1 = C_0^{'}.C_1^{'}...C_r^{'}$ is a product of $r+1$ disjoint cycles with $C_0^{'}$ and $C_1^{'}$ of length p_1 and for $i \geq 2$, $C_i^{'}$ is of length p_i. Let supp(C) denote the set of points moved by C for a cycle C (i.e., if we write $C = (i_1 i_2 ... i_p)$ it means C maps i_1 to i_2 and so on i_p to i_1 and moves no other element of the domain. Hence, supp$(C) = \{i_1, i_2, ..., i_p\}$). In the construction above we pick the cycles C_i and $C_i^{'}$, $0 \leq i \leq r$ such that supp$(C_0) \cap$ supp$(C_0^{'}) = \{1\}$ and $\forall (i, j) \neq (0, 0)$ supp$(C_i) \cap$ supp$(C_j^{'}) = \phi$. The domain $[N]$ on which these permutations are defined is $\bigcup_{i=0}^{r}(\text{supp}(C_i) \cup \text{supp}(C_i^{'}))$. Note that $N \leq 4p_1 + 2\sum_{i=2}^{r} p_i = O(n^6)$. Thus, the problem we consider is that of designing a circuit over S_N that takes as input x_0, x_1 and outputs at the n output gates $\pi_{y_i} = \Pi_{j=1}^{n} x_{y_i[j]}$ where $y_i[j]$ is the j-th bit of string y_i for each $y_i \in S$.

Theorem 4. *Any circuit over the group $(S_N, \circ)$ that takes as input π_0, π_1 and computes $G_S = \{\pi_{y_i} | y_i \in S\}$ as output is of size $\Omega(\frac{n^2}{\log^2 n})$.*

Proof. Let C be the circuit that solves this problem of computing G_S from x_0, x_1. We fix the input as $x_0 = \pi_0$ and $x_1 = \pi_1$. Now, consider the corresponding concatenation circuit $C^{'}$ with input $x_0, x_1 \in X$. At each output gate g_i, $1 \leq i \leq m$, circuit $C^{'}$ computes some word $w_i \in X^*$ such that $\forall i$, $\pi_{w_i} = \pi_{y_i}$ where π_{w_i} is the permutation in S_N obtained by putting $x_0 = \pi_0$ and $x_1 = \pi_1$ in w_i. If $w_i = y_i$ for all i, then in fact $C^{'}$ as a concatenation circuit computes the set S at its output gates. This implies by Theorem 1 that $size(C^{'}) = \Omega(\frac{n^2}{\log^2 n})$ and $size(C) = \Omega(\frac{n^2}{\log^2 n})$.

Suppose $w_i \neq y_i$ at some output gate g_i. We can write $w_i = u \circ b_2 \circ s$ and $y_i = v \circ b_1 \circ s$ where $b_1 \neq b_2$. Assume, without loss of generality, that $b_1 = 0$ and $b_2 = 1$. Since $\pi_{w_i} = \pi_{y_i}$, we know $\pi_u \pi_{b_2} \pi_s = \pi_v \pi_{b_1} \pi_s$ (i.e., $\pi_u \pi_1 \pi_s = \pi_v \pi_0 \pi_s$). Let $\alpha \in [N]$ such that $\pi_s(\alpha) = 1$. In $\pi_{y_i} = \pi_v \pi_0 \pi_s$, the permutation π_0 will map 1 to $\beta \in C_0\backslash\{1\}$, whereas in $\pi_{w_i} = \pi_u\pi_1\pi_s$ the permutation π_1 maps 1 to $\gamma \in C_0^{'}\backslash\{1\}$. Since $|v| < n$ the point β cannot be moved back to 1 and subsequently to $C_0^{'}\backslash\{1\}$. This is because $p_1 > n$ and the length of cycle $C_0^{'}$ is p_1. Therefore by π_{y_i} the point α is mapped to some point in $C_0\backslash\{1\}$. Since π_{w_i} must map α to the same point and $\pi_1\pi_s$ has mapped α to a point in $\gamma \in C_0^{'}\backslash\{1\}$, π_u must have at least $p_1 > n$ occurrences of π_1 in it to move γ to 1 and subsequently to the final point in $C_0\backslash\{1\}$ (using some π_0 applications). We will now argue that this forces w_i to be a long string.

Pick any tuple of points $(\alpha_1, \alpha_2, ..., \alpha_r)$ where $\alpha_i \in C_i^{'}$, $1 \leq i \leq r$. Notice that only π_1 moves this tuple because α_i, $1 \leq i \leq r$ do not belong to supp(π_0). Since $p_1, ..., p_r$ are distinct primes, the permutation π_1 maps $(\alpha_1, \alpha_2, ..., \alpha_r)$ to a set of $\Pi_{i=1}^r p_i - 1$ distinct r-tuples before returning to $(\alpha_1, \alpha_2, ..., \alpha_r)$. Suppose there are l occurrences of π_1 in π_{y_i}, $l < n$. Thus, if $\pi_{y_i}(\alpha_1, \alpha_2, ..., \alpha_r) = (\beta_1, \beta_2, ..., \beta_r)$ then $\pi_1^l(\alpha_1, \alpha_2, ..., \alpha_r) = (\beta_1, \beta_2, ..., \beta_r)$. Then $\pi_{w_i}(\alpha_1, \alpha_2, ..., \alpha_r) = (\beta_1, \beta_2, ..., \beta_r)$. However we know number of occurrences of π_1 in π_{w_i} is some $k \geq n$ which means $\pi_{w_i}(\alpha_1, \alpha_2, ..., \alpha_r) = (\beta_1, \beta_2, ..., \beta_r) = \pi_1^k(\alpha_1, \alpha_2, ..., \alpha_r)$.

It follows that $\pi_1^{k-l}(\alpha_1, \alpha_2, ..., \alpha_r) = (\alpha_1, \alpha_2, ..., \alpha_r)$ which implies $k - l$ is a multiple of $\Pi_{i=1}^r p_i$. Hence $|w_i| \geq \Pi_{i=1}^r p_i$. This implies that the circuit needs at least $\log \Pi_{i=1}^r p_i$ multiplication gates to compute w_i. This gives, $size(C) \geq \log \Pi_{i=1}^r p_i \geq \log 2^{n^2} = n^2$.

Putting it together $size(C) = \Omega(\frac{n^2}{\log^2 n})$ in any case. This completes the proof. □

6 Linear Circuits over Rings

In this section we consider a generalization of the linear circuits model. In this generalization we allow the coefficients come from *noncommutative rings*. In principle, we can expect lower bounds could be easier to prove in this model. The circuits are more constrained when coefficients come from a noncommutative ring as fewer cancellations can take place. This is in the same spirit as Nisan's [8] work on lower bounds for noncommutative algebraic branching programs. However, in this paper we succeed in showing only some limited lower bounds. We leave open problems that might be more accessible than the notorious problems for linear circuits over fields.

Let $(R, +, \cdot)$ be an arbitrary ring (possibly noncommutative). A *linear circuit* over R takes n inputs $y_1, y_2, \ldots, y_n$ labeling the indegree 0 nodes of a directed acyclic graph. The circuit has m output nodes. Each edge of the graph is labeled by some element of the ring R. The indegree of each non-input node is two. Each node of the circuit computes a linear form $\sum_{i=1}^n \alpha_i y_i$ for $\alpha_i \in R$ as follows: the input node labeled y_i computes y_i. Suppose g is a node with incoming edges from nodes g_1 and g_2, and the edges (g_1, g) and (g_2, g) are labeled by α and β respectively. If g_1 and g_2 computes the linear forms ℓ_1 and ℓ_2 respectively, then g computes $\alpha\ell_1 + \beta\ell_2$. Thus, for an $m \times n$ matrix A over the ring R, the circuit computes $A\mathbf{y}$ at the m output gates.

When R is a field we get the well-studied linear circuits model [7,11,6]. However, no explicit superlinear size lower bounds are known for this model over fields (except for some special cases like the bounded coefficient model [7] or in the cancellation free case [1]).

When the coefficients to come from a *noncommutative* ring R, we prove lower bounds for certain restricted linear circuits. Suppose the coefficient ring is $R = \mathbb{F}\langle x_0, x_1\rangle$ consisting of polynomials over the field $\mathbb{F}$ in noncommuting variables x_0 and x_1.

Let $M \in \mathbb{F}^{n\times n}\langle x_0, x_1\rangle$ where x_0, x_1 are noncommuting variables and $Y = (y_1, y_2, \ldots, y_n)^T$ is a column vector of input variables. The first restriction we consider are *homogeneous* linear circuits over the ring $\mathbb{F}\langle x_0, x_1\rangle$ for computing MY. The restriction is that for every gate g in the circuit, if g has its two incoming edges from nodes g_1 and g_2, then the edges (g_1, g) and (g_2, g) are labeled by α and β respectively, where $\alpha, \beta \in \mathbb{F}\langle x_0, x_1\rangle$ are restricted to be *homogeneous polynomials* of same degree in the variables x_0 and x_1. It follows, as a consequence of this restriction, that each gate g of the circuit computes a linear form $\sum_{i=1}^{n} \alpha_i y_i$, where the $\alpha_i \in \mathbb{F}\langle x_0, x_1\rangle$ are all homogeneous polynomials of the same degree. Our goal is to construct an explicit matrix $M \in \mathbb{F}^{n\times n}\langle x_0, x_1\rangle$ such that MY can not be computed by any circuit C with size $O(n)$ and depth $O(\log n)$. We prove this by suitably generalizing Valiant's matrix rigidity method [11] as explained below.

Consider $n \times n$ matrices $\mathbb{F}^{n\times n}$ over field $\mathbb{F}$. The *support* of a matrix $A \in \mathbb{F}^{n\times n}$ is the set of locations $\mathrm{supp}(A) = \{(i, j) \mid A_{ij} \neq 0\}$.

Definition 1. *Let $\mathbb{F}$ be any field. The rigidity $\rho_r(\mathcal{A})$ of a deck of matrices $\mathcal{A} = \{A_1, A_2, \ldots, A_N\} \subseteq \mathbb{F}^{n\times n}$ is the smallest number t for which there are a set of t positions $S \subseteq [n] \times [n]$ and a deck of matrices $\mathcal{B} = \{B_1, B_2, \ldots, B_N\}$ such that for all i: $\mathrm{supp}(B_i) \subseteq S$ and the rank of $A_i + B_i$ is bounded by r. A collection $\mathcal{A} = \{A_1, A_2, \ldots, A_N\} \subseteq \mathbb{F}^{n\times n}$ is a* rigid deck *if $\rho_{\epsilon\cdot n}(\mathcal{A}) = \Omega(n^{2-o(1)})$, where $\epsilon > 0$ is a constant.*

Notice that for $N = 1$ this is precisely the notion of rigid matrices. We are interested in constructing *explicit* rigid decks: I.e. a deck $\mathcal{A}$ such that for each $k \in [N]$ and each $1 \leq i, j \leq n$ there is a polynomial (in n) time algorithm that outputs the $(i, j)^{th}$ entry of A_k. We describe an explicit deck of size $N = 2^{n^2}$ over any field $\mathbb{F}$ and use it to prove our first lower bound result. It is convenient to write the deck as $\mathcal{A} = \{A_m \mid m \in \{x_0, x_1\}^{n^2}\}$ with matrices A_m indexed by monomials m of degree n^2 in the noncommuting variables x_0 and x_1. The matrix A_m is defined as follows:

$$A_m[i, j] = \begin{cases} 1 \text{ if } m_{ij} = x_1 \\ 0 \text{ if } m_{ij} = x_0 \end{cases}$$

Note that all the matrices A_m in the deck $\mathcal{A}$ are in $\mathbb{F}^{n\times n}$. Clearly, $\mathcal{A}$ is an explicit deck. We prove that it is a rigid deck.

Lemma 3. *The deck $\mathcal{A} = \{A_m \mid m \in \{x_0, x_1\}^{n^2}\}$ is an explicit rigid deck for any field $\mathbb{F}$.*

Proof. Valiant [11] showed that almost all $n \times n$ 0-1 matrices over any field $\mathbb{F}$ have rigidity $\Omega(\frac{(n-r)^2}{\log n})$ for target rank r. In particular, for $r = \epsilon \cdot n$, over any field $\mathbb{F}$, there is a 0-1 matrix R for which we have $\rho_r(R) \geq \frac{\delta \cdot n^2}{\log n}$ for some constant $\delta > 0$ depending on ϵ.

We claim that for the deck $\mathcal{A}$ we have $\rho_{\epsilon n}(\mathcal{A}) \geq \frac{\delta \cdot n^2}{\log n}$. To see this, let $E = \{E_m \in \mathbb{F}^{n \times n} | m \in \{x_0, x_1\}^{n^2}\}$ be any collection of matrices such that $|\text{supp}(E_m)| \leq \frac{\delta n^2}{\log n}$ for each m. Since the deck $\mathcal{A}$ contains all 0-1 matrices, in particular $R \in \mathcal{A}$ and $R = A_m$ for some monomial m. From the rigidity of R we know that the rank of $R + E_m$ is at least ϵn. This proves the claim and the lemma follows. □

We now turn to the lower bound result for homogeneous linear circuits where the coefficient ring is $\mathbb{F}\langle x_0, x_1 \rangle$. We define an explicit $n \times n$ matrix M as

$$M_{ij} = (x_0 + x_1)^{(i-1)n+j-1} \cdot x_1 \cdot (x_0 + x_1)^{n^2 - ((i-1)n+j)}. \tag{1}$$

It is easy to see that we can express the matrix M as $M = \sum_{m \in \{x_0, x_1\}^{n^2}} A_m m$, where $\mathcal{A} = \{A_m \mid m \in \{x_0, x_1\}^{n^2}\}$ is the deck defined above.

Theorem 5. *Any homogeneous linear circuit C over the coefficient ring $\mathbb{F}\langle x_0, x_1 \rangle$ computing MY, for M defined above, requires either size $\omega(n)$ or depth $\omega(\log n)$.*

Proof. Assume to the contrary that C is a homogeneous linear circuit of size $O(n)$ and depth $O(\log n)$ computing MY. We know that by Valiant's graph-theoretic argument (see e.g. [6]) that in the circuit C there is a set of gates V of cardinality $s = \frac{c_1 n}{\log \log n} = o(n)$ such that at least $n^2 - n^{1+\delta}$, for $\delta < 1$, input-output pairs have all their paths going through V. Thus, we can write $M = B_1 B_2 + E$ where $B_1 \in \mathbb{F}^{n \times s}\langle x_0, x_1 \rangle$ and $B_2 \in \mathbb{F}^{s \times n}\langle x_0, x_1 \rangle$ and $E \in \mathbb{F}^{n \times n}\langle x_0, x_1 \rangle$, and $|\text{supp}(E)| \leq n^{1+\delta}$. By collecting the matrix coefficient of each monomial we can express M and E as

$$M = \sum_{m \in \{x_0, x_1\}^{n^2}} A_m m, \text{ and } E = \sum_{m \in \{x_0, x_1\}^{n^2}} E_m m,$$

where A_m are already defined and $| \cup_{m \in \{x_0, x_1\}^{n^2}} \text{supp}(E_m)| \leq n^{1+\delta}$. Now consider the matrix $B_1 B_2$. By collecting matrix coefficients of monomials we can write $B_1 B_2 = \sum_{m \in \{x_0, x_1\}^{n^2}} B_m m$.

We now analyze the matrices B_m. Crucially, by the homogeneity condition on the circuit C, we can partition $V = V_1 \cup V_2 \cup \ldots V_\ell$, where each gate g in V_i computes a linear form $\sum_{j=1}^n \gamma_j y_j$ and $\gamma_j \in \mathbb{F}\langle x_0, x_1 \rangle$ is a homogeneous degree d_i polynomial. Let $s_i = |V_i|, 1 \leq i \leq \ell$. Then we have $s = s_1 + s_2 + \ldots s_\ell$. Every monomial m has a unique prefix of length d_i for each degree d_i associated with the gates in V. Thus, we can write $B_m = \sum_{j=1}^{\ell} B_{m,j,1} B_{m,j,2}$, where $B_{m,j,1}$ is the $n \times s_j$ matrix corresponding to the d_j-prefix of m and $B_{m,j,2}$ is the $s_j \times n$ matrix corresponding to the $n^2 - d_j$-suffix of m. It follows that for each monomial m

the rank of B_m is bounded by s. Putting it together, for each monomial m we have $A_m = B_m + E_m$, where B_m is rank s and $|\cup_{m \in \{x_0,x_1\}^{n^2}} \text{supp}(E_m)| \leq n^{1+\delta}$. This contradicts the fact that $\mathcal{A}$ is a rigid deck. □

Remark 2. For the matrix $M = (m_{ij})$, as defined above, it does not seem that Shoup-Smolensky dimension method [10] can be used to prove a similar lower bound. To see this, suppose $\Gamma_M(n)$ is the set of all monomials of degree n in $\{m_{ij}\}$ and let $D_M(n)$ be the dimension of the vector space over $\mathbb{F}$ spanned by the set $\Gamma_M(n)$. The upper bound for $D_M(n)$ that we can show for a depth d and size $O(n)$ linear circuit over the ring $\mathbb{F}\langle x_0, x_1\rangle$ is as large as $(\frac{O(n)}{d})^{dn}$. This bound, unfortunately, is much larger than the bounds obtainable for the commutative case [10]. On the other hand, the lower bound for $D_M(n)$ is only $n^{\Theta(n)}$. Thus, we do not get a superlinear size lower bound for the size using Shoup-Smolensky dimensions when the coefficient ring is $\mathbb{F}\langle x_0, x_1\rangle$.

We next consider homogeneous depth 2 linear circuits. These are linear circuits of depth 2, where each addition gate can have unbounded fanin. More precisely, if g is an addition gate with inputs from $g_1, g_2, \ldots, g_t$ then the gate g computes $\sum_{i=1}^{t} \alpha_i g_i$, where each edge (g_i, g) is labeled by $\alpha_i \in \mathbb{F}\langle x_0, x_1\rangle$ such that $\alpha_i, 1 \leq i \leq t$ are all homogeneous polynomials of the same degree. We again consider the problem of computing MY for $M \in \mathbb{F}^{n\times n}\langle x_0, x_1\rangle$. The goal is to lower bound the number of wires in the linear circuit. This problem is also well studied for linear circuits over fields and only an explicit $\Omega(n \log^2 n / \log\log n)$ lower bound is known for it [6,9], although for random matrices the lower bound is $\Omega(n^2/\log n)$.

We show that for the explicit matrix M as defined above, computing MY by a depth 2 homogeneous linear circuit (with unbounded fanin) requires $\Omega(\frac{n^2}{\log n})$ wires.

Theorem 6. *Let $M \in \mathbb{F}_2^{n\times n}\langle x_0, x_1\rangle$ as defined in Equation 1. Any homogeneous linear circuit C of depth 2 computing MY requires $\Omega(\frac{n^2}{\log n})$ wires.*

Proof. Let C be a homogeneous linear circuit of depth 2 computing MY. Let $w(C)$ denote the number of wires in C. Let s be the number of gates in the middle layer of C. We can assume without loss of generality that, all input to output paths in C are of length 2 and hence pass through the middle layer. A *level 1* edge connects an input gate to a middle-layer gate and a *level 2* edge is from middle layer to output. Thus, we can factorize $M = M^{'} * M^{''}$ where the matrix $M^{'}$ is in $\mathbb{F}^{n\times s}\langle x_0, x_1\rangle$ and $M^{''}$ is in $\mathbb{F}^{s\times n}\langle x_0, x_1\rangle$, and the complexity of C is equivalent to total number of nonzero entries in $M^{'}$ and $M^{''}$. As before, write $M = \sum_{m\in\{x_0,x_1\}^{n^2}} A_m m$.

Given A_m for $m \in \{x_0, x_1\}^{n^2}$, we show how to extract from C a depth-2 linear circuit over the *field* $\mathbb{F}$, call it $C^{(m)}$, that computes A_m such that the number of wires in $C^{(m)}$ is at most the number of wires in C. Indeed, we do not add any new gate or wires in obtaining $C^{(m)}$ from C.

For each gate g in the middle layer, there are at most n incoming edges and n outgoing edges. As C is a homogeneous linear circuit we can associate a degree d_g

to gate g. Each edge (i,g) to g is labeled by a homogeneous degree-d_g polynomial $\alpha_{i,g}$ in $\mathbb{F}\langle x_0,x_1\rangle$. Likewise, each edge (g,j) from g to the output layer is labeled by a degree (n^2-d_g) homogeneous polynomial $\beta_{g,j}$. Let $m=m_1m_2$, where m_1 is of degree d_g and m_2 of degree n^2-d_g. For each incoming edge (i,g) to g we keep as label the coefficient of the monomial m_1 in $\alpha_{i,g}$ and for outgoing edge (g,j) from g we keep as label the coefficient of the monomial m_2 in $\beta_{g,j}$. We do this transformation for each gate g in the middle layer of C. This completes the description of the depth-2 circuit $C^{(m)}$. By construction it is clear that $C^{(m)}$ computes A_m and the number of wires $w(C^{(m)})$ in $C^{(m)}$ is bounded by $w(C)$ for each monomial $m\in\{x_0,x_1\}^{n^2}$. However, $\{A_m \mid m\in\{x_0,x_1\}^{n^2}\}$ is the set of all 0-1 matrices over $\mathbb{F}$ and it is known that there are $n\times n$ 0-1 matrices A_m such that any depth-2 linear circuit for it requires $\Omega(\frac{n^2}{\log n})$ wires (e.g. see [6]). Hence, the number of wires in C is $\Omega(\frac{n^2}{\log n})$. □

If we restrict the edge labels in the linear circuit computing MY to only *constant-degree polynomials*, then we can obtain much stronger lower bounds using Nisan's lower bound technique for noncommutative algebraic branching programs. We can define the matrix M as follows. Let $M_{ij}=w_{ij}w_{ij}^R$, where $w_{ij}\in\{x_0,x_1\}^{2\log n}$ and $1\le i,j\le n$ are all distinct monomials of degree $2\log n$. We refer to M as a palindrome matrix. All entries of M are distinct and note that each entry of MY can be computed using $O(n\log n)$ gates.

Theorem 7. *Any linear circuit over $\mathbb{F}\langle x_0,x_1\rangle$ computing MY, where edge labels are restricted to be constant-degree polynomials, requires size $\Omega(\frac{n^2}{\log n})$.*

Proof. Let C be such a linear circuit computing MY. Since edges can be labeled by constant-degree polynomials, we can first obtain a linear circuit C' computing MY such that each edge is labeled by a *homogeneous linear form.* The size $\text{size}(C')=O(\text{size}(C)\log n)$. From C', we can obtain a noncommutative algebraic branching program $\hat{C}$ that computes the palindrome polynomial $\sum_{w\in\{x_0,x_1\}^{2\log n}} ww^R$ such that $\text{size}(\hat{C})=O(\text{size}(C'))$. By Nisan's lower bound [8] $\text{size}(\hat{C})=\Omega(n^2)$, which implies $\text{size}(C)=\Omega(\frac{n^2}{\log n})$. □

Theorem 8. *Any linear circuit, whose edge labels are restricted to be either a homogeneous degree $4\log n$ polynomial <u>or</u> a scalar, computing MY requires $\Omega(n^2)$ size, where M is the palindrome matrix. Moreover, there is a matching upper bound.*

Proof. Let C be any linear circuit computing MY. Each entry m_{ij} of the matrix M can be written as sum of products of polynomials $m_{ij}=\sum_{\rho_{ij}}\prod_{e\in\rho_{ij}} l(e)$ where ρ_{ij} is a path from input y_j to output gate i in C and $l(e)$ is the label of edge e in C. Let S be set of all edge labels in C with degree $4\log n$ polynomial. Thus, each m_{ij} is a linear combinations of elements in the set S over $\mathbb{F}$. This implies that $m_{ij}\in Span(S)$ where $i\le i,j\le n$. Since all m_{ij} are distinct, $|S|\ge n^2$. Since fan in is 2, $size(C)\ge\frac{n^2}{2}=\Omega(n^2)$.

For upper bound, we use n^2 edges (n edges starting from each input y_i) each labeled by a corresponding monomial in M (of degree $4\log n$) and then we add relevant edges to get the output gates. Thus, upper bound is $O(n^2)$ for computing MY. □

Note that, since we have not used noncommutativity in the proof, Theorem 8 also holds in the commutative settings (we require $\Omega(n^2)$ entries of M to be distinct).

7 Concluding Remarks

For multiplicative circuits we could prove lower bounds only for large monoids and large groups. The main question here is whether we can prove lower bounds for an explicit function $f : S^n \to S^m$, for some constant size nonabelian group or monoid S.

We introduced the notion of rigidity for decks of matrices, but the only explicit example we gave was the trivial one with a deck of size 2^{n^2}. A natural question is to give explicit constructions for smaller rigid decks of $n \times n$ matrices, say of size $n!$ or less. Or is the construction of rigid decks of smaller size equivalent to the original matrix rigidity problem?

Acknowledgments. We are grateful to the referees for their detailed comments and useful suggestions.

References

1. Boyar, J., Find, M.G.: Cancellation-free circuits in unbounded and bounded depth. In: Gąsieniec, L., Wolter, F. (eds.) FCT 2013. LNCS, vol. 8070, pp. 159–170. Springer, Heidelberg (2013)
2. Charikar, M., Lehman, E., Liu, D., Panigrahy, R., Prabhakaran, M., Sahai, A., Shelat, A.: The smallest grammar problem. IEEE Transactions on Information Theory 51(7), 2554–2576 (2005)
3. Jukna, S., Sergeev, I.: Complexity of linear boolean operators. Foundations and Trends in Theoretical Computer Science 9(1), 1–123 (2013)
4. Lipton, R.J., Zalcstein, Y.: Word problems solvable in logspace. Journal of the ACM (JACM) 24(3), 522–526 (1977)
5. Lohrey, M.: Algorithmics on slp-compressed strings: A survey. Groups Complexity Cryptology 4(2), 241–299 (2012)
6. Lokam, S.V.: Complexity lower bounds using linear algebra. Foundations and Trends in Theoretical Computer Science 4(1-2), 1–155 (2009)
7. Morgenstern, J.: Note on a lower bound on the linear complexity of the fast fourier transform. Journal of the ACM (JACM) 20(2), 305–306 (1973)
8. Nisan, N.: Lower bounds for non-commutative computation (extended abstract). In: STOC, pp. 410–418 (1991)
9. Pudlak, P.: Large communication in constant depth circuits. Combinatorica 14(2), 203–216 (1994)
10. Shoup, V., Smolensky, R.: Lower bounds for polynomial evaluation and interpolation problems. Computational Complexity 6(4), 301–311 (1996)
11. Valiant, L.G.: Graph-theoretic arguments in low-level complexity. In: Gruska, J. (ed.) MFCS 1977. LNCS, vol. 53, pp. 162–176. Springer, Heidelberg (1977)