Lower bounds for circuits with MOD_m gates

Arkadev Chattopadhyay *
McGill University, Montreal
achatt3@cs.mcgill.ca

Navin Goyal *
McGill University, Montreal
navin@cs.mcgill.ca

Pavel Pudlák †
Czech Academy of Sciences, Prague
pudlak@math.cas.cz

Denis Thérien *
McGill University, Montreal
denis@cs.mcgill.ca

Abstract

Let $\mathrm{CC}_{o(n)}[m]$ *be the class of circuits that have size* $o(n)$ *and in which all gates are* MOD_m *gates.*

- *We show that* $\mathrm{CC}[m]$ *circuits cannot compute* MOD_q *in sub-linear size when* $m, q > 1$ *are co-prime integers. No non-trivial lower bounds were known before on the size of* $\mathrm{CC}[m]$ *circuits of constant depth for computing* MOD_q. *On the other hand, our results show circuits of type* $\mathrm{MAJ} \circ \mathrm{CC}_{o(n)}[m]$ *need exponential size to compute* MOD_q. *Using Bourgain's recent breakthrough result on estimates of exponential sums, we extend our bound to the case where small fan-in AND gates are allowed at the bottom of such circuits i.e. circuits of type* $\mathrm{MAJ} \circ \mathrm{CC}[m] \circ \mathrm{AND}_{\epsilon \log n}$*, where* $\epsilon > 0$ *is a sufficiently small constant.*
- $\mathrm{CC}[m]$ *circuits of constant depth need superlinear number of wires to compute both the* AND *and* MOD_q *functions. To prove this, we show that any circuit computing such functions has a certain connectivity property that is similar to that of superconcentration. We show a superlinear lower bound on the number of edges of such graphs extending results on superconcentrators.*

1 Introduction

Proving lower bounds on the size of boolean circuits needed to compute explicit functions is of fundamental importance in theoretical computer science. Since the problem has proved to be very hard in general, various restricted models of circuits have been considered. One of the most fruitful directions has been the study of small depth circuits. The result (the most general version of which appears in [14]) that circuits constructed using unrestricted fan-in OR, AND and NOT gates with constant depth (such circuits when restricted to polynomial size define the class AC^0) need exponential size to compute the PARITY function, remains a jewel of the area.

Smolensky [25], extending the work of Razborov [24], showed that sub-exponential size AC^0 circuits augmented with MOD_m gates (such circuits when restricted to polynomial size define the class $\mathrm{ACC}^0[m]$) cannot compute MOD_q if $(m, q) = 1$ and m is a prime power. However, the seemingly innocuous problem of extending these lower bounds to $\mathrm{ACC}^0[m]$ circuits for general m has remained open despite extensive efforts. Some of the difficulties that one faces in attempting to extend the Razborov-Smolensky method to general m are discussed in [2].

One of the main impediments in obtaining bounds for ACC^0, is understanding the power of circuits of constant depth having *only* MOD_m gates. The class of such circuits is denoted by $\mathrm{CC}^0[m]$ when the circuits are further restricted to have polynomial size. Since it is difficult to compute the MOD_m function using AND and OR gates, it is an interesting question to determine if small size $\mathrm{CC}^0[m]$ circuits can compute AND and OR. It is known that both AND and MOD_q functions are impossible to compute by constant depth circuits composed entirely of MOD_m gates when m is a prime power. In contrast, it is also known that depth two MOD_6 circuits can compute every boolean function in exponential size [4]. [18] makes the tempting conjecture that AND needs ex-

*Supported by NSERC and FQRNT.
†Supported by grant No. A1019401 of the Academy of Sciences, Czech Republic.

ponential size $\mathrm{CC}^0[m]$ circuits. A special case of a conjecture of Smolensky implies exponential lower bounds on size of such circuits computing MOD_q, whenever $(m, q) = 1$.

Most known lower bounds, e.g., [4, 17, 11, 10] work only for special classes of $\mathrm{CC}^0[m]$ circuits. We do not even know if the satisfiability problem (SAT) can be solved by depth-2 linear size $\mathrm{CC}[6]$ circuits, when the gates used are *generalized* MOD_6 gates [7] (see Section 2 for the definition of generalized MOD gate).

The currently best known lower bound on the size of $\mathrm{CC}^0[m]$ circuits computing AND is linear in the number of variables [28]; [26] proved a linear lower bound for a more complicated function. However, the methods of [28] and [26] do not seem to yield a linear lower bound for MOD_q. In fact, previous to this work, to the best of our knowledge, no linear lower bounds were known for MOD_q. The difficulty in proving such lower bounds may partly be explained by the fact mentioned above that depth two $\mathrm{CC}[m]$ circuits can compute all boolean functions if m has at least two different prime factors, but not if m is a prime power. The advantage of composites over prime powers in computing the AND and MOD_q functions is also witnessed in the closely related setting of polynomials over $\mathbb{Z}_m$ (see [3, 12]).

As a special case of CC^0, [4] considered $\mathrm{MOD}_p \circ \mathrm{MOD}_m$ circuits (those having depth two with a MOD_p gate at the output and a single layer of MOD_m gates at the input). A number of papers [4, 11, 27] showed exponential lower bounds for such circuits computing AND and MOD_r, where $(r, p) = (r, m) = 1$. [4] formulated the Constant Degree Hypothesis (CDH) whose special case asserts that circuits of the type $\mathrm{MOD}_p \circ \mathrm{MOD}_m \circ \mathrm{AND}_{O(1)}$ (layered depth-3 circuits with AND gates of constant fan-in in the input layer, MOD_m gates in the middle layer, and a MOD_p gate at the output) require exponentially many MOD_m gates to compute AND. Some progress towards proving CDH is made by [29, 11, 10]. While obtaining the general CDH remains wide open, previous to our work not even linear lower bounds on the number of MOD_m gates were known, without restricting the type of subcircuits rooted at each MOD_m gate.

It is known (see [5]) that ACC^0 can be simulated by depth three circuits using MAJORITY gates alone in quasi-polynomial size. [1] shows that quasi-polynomial size circuits of type $\mathrm{MAJ} \circ \mathrm{MOD}_m \circ \mathrm{AND}_{\mathrm{polylog}(n)}$ can simulate AC^0, for *every* $m > 1$. An interesting open question is if these circuits are powerful enough to simulate ACC^0. Recently, Bourgain [6] made a breakthrough by showing that such circuits required exponential size to compute MOD_q if the fan-in of the AND gates at the bottom were restricted to a constant, $(m, q) = 1$ and m is odd. [9] showed that there is a simple way to extend Bourgain's argument to all m.

While the number of gates has been the more popular measure of circuit size, number of wires has also been studied fairly extensively, e.g., [8, 22, 23, 16]. [8] show that, somewhat surpisingly, AC^0 circuits can compute $\mathrm{THRESHOLD}_k$ for any constant k using linear number of wires. Superlinear lower bounds on the number of wires for constant depth arithmetic circuits are proved in [23]. They also show similar lower bounds for boolean circuits computing n boolean functions with n inputs. [16] is able to give a superlinear bound on the number of wires in ACC^0 circuits computing a single boolean function. However, their method applies only to those functions that have high communication complexity. Consequently, their method fails to give bounds on simple functions like AND and MOD_q.

Our results. We need some definitions. $\mathrm{CC}[m]$ denotes the class of circuits consisting of MOD_m gates *without* any depth restriction; as mentioned before, $\mathrm{CC}^0[m]$ denotes the same class but now the circuits have constant depth; $\mathrm{CC}_{o(n)}[m]$ denotes the class $\mathrm{CC}[m]$ when the circuits are restricted to have size $o(n)$. Unless otherwise specified we always consider generalized MOD_m gates which we now define.

For every positive integer m, we define the boolean function $\mathrm{MOD}_m : \{0,1\}^n \to \{0,1\}$ in the following way: $\mathrm{MOD}_m(x) = 1$ iff $\sum_{i=1}^n x_i \neq 0 \pmod m$. For each $A \subseteq Z_m$, the *generalized* MOD_m^A boolean gate computes the following function : $\mathrm{MOD}_m^A(x) = 1$ iff $\sum_{i=1}^n x_i \in A$. The set A is called the accepting set of the MOD gate. We remark that the standard gate used in the literature is the one that has the accepting set $\{1, \ldots, m-1\}$. However, in circuits that we consider each gate would have its own accepting set that may or may not be the same as those of other gates.

Let q be a positive integer and $b \in \{0, \ldots, q-1\}$. Define the bth MOD_q-residue class of $\{0,1\}^n$ by

$$M_{n,q}(b) =$$

$$\{x = (x_1, \ldots, x_n) \in \{0,1\}^n \mid \sum_{i=1}^n x_i = b \mod q\}$$

Our lower bounds on the number of gates in $\mathrm{CC}[m]$ circuits follow from the following two results about boolean solutions to systems of linear equations over $\mathbb{Z}_m$. These results may be of independet interest.

Let $\mathcal{L} = \{\theta_1, \ldots, \theta_s\}$ be a set of s linear forms over $\mathbb{Z}_m$. For $v \in \mathbb{Z}_p^s$, let $K^{\mathcal{L}}(v)$ represent the set of points in $\{0,1\}^n$, that satisfy $\theta_i = v_i$ for all $1 \le i \le s$. Using simple estimates on exponential sums, we show the following :

Lemma 1 (Linear Uniformity Lemma) *Using the notation above, for all positive integers m, q with $(m,q) = 1$, there exists a positive constant $\gamma = \gamma(m,q) < 1$ such that for all n and linear mappings $\mathcal{L} : \mathbb{Z}_m^n \to \mathbb{Z}_m^s$,*

$$\big||K^{\mathcal{L}}(v) \cap M_{n,q}(b)| - |K^{\mathcal{L}}|/q\big| \le (2\gamma)^n. \quad (1)$$

for each $b \in \{0, \ldots, q-1\}$ and $v \in \mathbb{Z}_m^s$.

The above lemma shows that if $|K^{\mathcal{L}}(v)|$ is large compared to the RHS of (1), then every MOD_q residue class occurs with roughly the same frequency in $K^{\mathcal{L}}(v)$. In that case, $K^{\mathcal{L}}(v)$ looks random to a MOD_q counter. By combining two known results from additive number theory and Fourier analysis, we show that the set $K^{\mathcal{L}}(v)$ is indeed large, whenever it is non-empty.

Theorem 2 *For all positive integers m there exists a positive constant c such that the following holds. Let $\mathcal{L} : \mathbb{Z}_m^n \to \mathbb{Z}_m^s$ be a linear map. For any $v \in \mathbb{Z}_m^s$, if $K^{\mathcal{L}}(v)$ is non-empty, then*

$$|K^{\mathcal{L}}(v)| \ge \frac{2^n}{c^s}. \quad (2)$$

Lower bounds on the number of gates. In Section 3, we show that Lemma 1 implies

Lemma 3 *Consider any positive integers q, m that are co-prime to each other and numbers $a, b \in \{0, \ldots, q-1\}$. Then, for every $\mathrm{CC}[m]$ circuit C of size $o(n)$, we have*

$$\begin{aligned} &|\Pr_x[C(x) = 1 | x \in M_{n,q}(a)] \\ &- \Pr_x[C(x) = 1 | x \in M_{n,q}(b)]| \le 2^{-\Omega(n)}. \quad (3) \end{aligned}$$

Consider any boolean function f and two disjoint sets A, B, where $A \subseteq f^{-1}(1)$ and $B \subseteq f^{-1}(0)$. We say that a circuit C is an ϵ-discriminator for f with respect to A and B if

$$|\Pr_x[C(x) = 1 | x \in A] - \Pr_x[C(x) = 1 | x \in B]| \ge \epsilon.$$

The ϵ-discriminator lemma of Hajnal et al. [15] states that if a circuit with a MAJORITY gate at the output computes a function f and the fan-in of the output MAJORITY gate is s, then for every $A \subseteq f^{-1}(1)$ and $B \subseteq f^{-1}(0)$ at least one of the sub-circuits feeding into the output gate $1/s$-discriminates f. Hence, choosing $A = M_{n,q}(1)$ and $B = M_{n,q}(0)$ and using Lemma 3, we get the following where by a circuit of type $\mathrm{MAJ} \circ \mathrm{CC}_{o(n)}[m]$ we mean a circuit with a MAJ gate at the ouput with $\mathrm{CC}[m]$ circuits of sublinear size feeding into it.

Theorem 4 *Any circuit of type $\mathrm{MAJ} \circ \mathrm{CC}_{o(n)}[m]$ computing MOD_q requires the output gate to have fan-in $2^{\Omega(n)}$ if $(m,q) = 1$.*

In fact, our methods show the above theorem for the more general class of circuits of type $\mathrm{MAJ} \circ \mathrm{ANY} \circ \mathrm{CC}[m]$, where the sub-circuit rooted at each ANY gate is sub-linear in size and every input to an ANY gate is the output of a MOD_m gate.

We also prove a different kind of lower bound for $\mathrm{CC}[m]$ circuits. An intuition about $\mathrm{CC}[m]$ circuits of small size is that since the set of inputs that a single MOD_m gate accepts is large and cannot be concentrated in a small portion of the boolean hypercube, hence one would expect a similar situation to occur for a small $\mathrm{CC}[m]$ circuit. The Uniformity Lemma does not imply any lower bounds on the size of the support set of $\mathrm{CC}[m]$ circuits. Theorem 2 below gives a bound of this type

Theorem 5 *For every positive integer m there exists a positive constant c such that any boolean function with support size less than $2^n/c^s$ requires $\mathrm{CC}[m]$ circuits of size at least s.*

Thérien [28] implies a similar but weaker result that functions with support set of size less than $(\frac{\alpha(m)}{\alpha(m)-1})^n \frac{1}{\alpha(m)^s}$ require $\mathrm{CC}[m]$ circuits of size s.

In particular, such results imply that AND cannot be computed by sublinear size $\mathrm{CC}[m]$ circuits.

Lower bound for number of wires. We give super-linear lower bounds on the number of wires in $\mathrm{CC}^0[m]$ circuits computing AND and MOD_q. To state our result more precisely, define for $d = 1, 2, \ldots,$

$$\lambda_1(n) = \lceil \log_2 n \rceil,$$
$$\lambda_{d+1}(n) = \min\{i \in \mathbb{N};\ \lambda_d^{(i)}(n) \le 1\},$$

where the superscript i denotes the i-times iterated function.

Theorem 6 *For every q and d there exist $\delta > 0$ such that every circuit computing AND or MOD_q functions that has depth $d+1$ and uses only MOD_m gates, has at least $\delta n \lambda_d(n)$ wires.*

We consider the bounded depth directed graph of a boolean circuit. The proof of the above theorem involves first showing that such graphs must satisfy a certain connectivity property similar to that of superconcentrators. We next prove a superlinear lower bound on the number of edges in such graphs. This theorem is stronger than the lower bounds proved on bounded depth superconcentrators (when the depth of superconcentrator is even) and enables us to prove lower bounds on $\mathrm{CC}^0[m]$ circuits for which we cannot use superconcentrators.

Uniformity Lemma for equations of small degree. Finally in Section 5, we show an extension of the Linear Uniformity Lemma using recent breakthrough result on estimates of exponential sums by Bourgain [6]. This extension can then be used to derive a correponding generalization of Lemma 3 to circuits of type $\mathrm{CC}_{o(n)}[m] \circ \mathrm{AND}_{\epsilon \log n}$. More precisely,

Theorem 7 *For each $\epsilon \in (0,1)$ there exists a $\delta > 0$ such that every circuit of type $\mathrm{MAJ} \circ \mathrm{CC}_{o(n)}[m] \circ \mathrm{AND}_{\delta \log n}$ computing MOD_q requires the output gate to have fan-in $2^{\Omega(n^\epsilon)}$ if $(m,q)=1$.*

A related result was first proved by Hansen [13] after the preliminary vesrsion of this paper was submitted. We later observed that Theorem 7 followed easily from our work.

2 Properties of boolean solutions of systems of equations

2.1 Proof of the Uniformity Lemma

The proof of the Uniformity Lemma uses an exponential sum argument. Exponential sums have been previously used in similar contexts [6, 9]. As is standard, we use the notation $e_m(x)$ to denote $e^{2\pi i x/m}$, where i is the complex square root of -1.

Proof: [of Lemma 1] Suppose $K^{\mathcal{L}}(v)$ is non-empty. Then, $\theta(a) = v$ for some boolean vector a. Substituting $x_i = x_i - a_i$ and $b = b - \sum_{i=1}^n a_i$, for $1 \le i \le n$, we reduce to the case of v being the all-zero vector. We abbreviate $K^{\mathcal{L}}(0^s)$ to $K^{\mathcal{L}}$. We first write $|K^{\mathcal{L}} \cap M_{n,q}(b)|$ as an exponential sum and then estimate this exponential sum by grouping the terms appropriately.

$$|K^{\mathcal{L}} \cap M_{n,q}(b)| = \sum_{x\in\{0,1\}^n} \Big[\prod_{i=1}^{s} \big(\frac{1}{m}\sum_{j=0}^{m-1} \mathrm{e}_m(j\theta_i(x))\big)\big(\frac{1}{q}\sum_{j=0}^{q-1}\mathrm{e}_q(j(\sum_{k=1}^n x_k - b))\big)\Big]. \tag{4}$$

The above identity is immediate from the well-known and simple fact that $\frac{1}{m}\sum_{j=0}^{m-1}\mathrm{e}_m(ja)$ is 1 if $a=0$ and is 0 otherwise, for every positive integer m. We now rewrite the right hand side (RHS) in (4) as

$$\begin{aligned}(4) = & \sum_{x\in\{0,1\}^n} \frac{1}{q}\prod_{i=1}^{s}\big(\frac{1}{m}\sum_{j=0}^{m-1}\mathrm{e}_m(j\theta_i(x))\big) + \\ & + \sum_{x\in\{0,1\}^n}\Big[\prod_{i=1}^{s}\big(\frac{1}{m}\sum_{j=0}^{m-1}\mathrm{e}_m(j\theta_i(x))\big) \\ & \times \big(\frac{1}{q}\sum_{j=1}^{q-1}\mathrm{e}_q(j(\sum_{k=1}^n x_k - b))\big)\Big]. \end{aligned} \tag{5}$$

The first term in the RHS is easily seen to be $|K^{\mathcal{L}}|/q$. Hence we get

$$\begin{aligned} &\big||K^{\mathcal{L}} \cap M_{n,q}(b)| - |K^{\mathcal{L}}|/q\big| = \\ &\Big|\sum_{x\in\{0,1\}^n}\Big[\prod_{i=1}^{s}\big(\frac{1}{m}\sum_{j=0}^{m-1}\mathrm{e}_m(j\theta_i(x))\big) \\ &\times \big(\frac{1}{q}\sum_{j=1}^{q-1}\mathrm{e}_q(j(\sum_{k=1}^n x_k - b))\big)\Big]\Big| \end{aligned} \tag{6}$$

We now estimate the RHS of (6). To do this, let us multiply out the terms in the summand inside the abolute value and then sum the resulting terms. We obtain $m^s(q-1)$ terms after multiplying out the terms in the summand, each of which gives rise to a sum of the form

$$\begin{aligned} &\frac{\mathrm{e}_q(-jb)}{m^s q}\sum_{x\in\{0,1\}^n}\big[\mathrm{e}_m(j_1\theta_1(x) + \ldots + j_s\theta_r(x)) \\ &\times \mathrm{e}_q(j\sum_{k=1}^n x_k)\big]. \end{aligned} \tag{7}$$

where $(j_1,\ldots,j_s) \in \{0,\ldots,m-1\}^s$ and $j \in \{1,\ldots,q-1\}$.

Bounding the absolute value of the expression in the previous equation is standard (see, e.g., Bourgain [6]); however, we include the proof here as it is simple and thus makes the proof self-contained. Writing $a_1x_1 + \ldots + a_nx_n := j_1\theta_1(x) + \ldots + j_s\theta_s(x)$, using the trigonometric identity $1 + e^{i2\rho} = 2e^{i\rho}\cos(\rho)$, and taking absolute values, we have

$$\begin{aligned} |(7)| &= \Big|\frac{1}{m^s q}\prod_{i=1}^{n}(1 + \mathrm{e}_m(a_i)\mathrm{e}_q(j))\Big| \\ &= \Big|\frac{2^n}{m^s q}\prod_{i=1}^{n}\cos\big(\pi(\frac{a_i}{m} + \frac{j}{q})\big)\Big|. \end{aligned} \tag{8}$$

Let $\gamma = \max_{a_i \in \mathbb{Z}_q;\, j \in \mathbb{Z}_m} |\cos\left(\pi(\frac{a_i}{m} + \frac{j}{q}\right)|$. Since, m and q are co-prime and $j \neq 0$, it can be verified that $\gamma < 1$. Hence

$$|(8)| \leq \frac{2^n \gamma^n}{m^s q}. \tag{9}$$

Using the triangle inequality in the RHS of (6) and plugging in the bound of (9), we get

$$\left||K^{\mathcal{L}} \cap M_{n,q}(b)| - |K^{\mathcal{L}}|/q\right| \leq m^s(q-1)\frac{(2\gamma)^n}{m^s q}. \tag{10}$$

■

As mentioned in the introduction, the linear uniformity lemma can be generalized to systems of equations if the degrees of the equations are small. A precise statement appears in Section 5.

2.2 Size of a solution set

A simple averaging argument shows that for every $\mathcal{L} : \mathbb{Z}_m^n \to \mathbb{Z}_m^s$, there exists a $v \in \mathbb{Z}_m^s$ such that $K^{\mathcal{L}}(v)$ has size at least $2^n/m^s$. An interesting question is if every v for which $K^{\mathcal{L}}(v)$ is non-empty is of size close to the average size? We note that the results in [28] based on methods introduced in [4], imply a lower bound of $(\frac{\alpha}{\alpha-1})^n \cdot \frac{1}{\alpha^s}$ for $|K^{\mathcal{L}}(v)|$ when it is non-zero. This is still exponentially smaller than the average size. Theorem 2 gives a lower bound for every non-empty $K^{\mathcal{L}}(v)$ that is indeed close to the average size.

Proof:[of Theorem 2] To prove Theorem 2 we combine two existing results from the literature. First, we need a notion from additive combinatorics: for any abelian group G, the *Davenport constant* of G (denoted by $s(G)$) is the smallest integer k such that every sequence of elements of G of length at least k, has a non-empty subsequence that sums to zero. Olson [20] showed that there exists a connection between $s(G)$ and the set of boolean solutions to the equation $g_1x_1 + \ldots + g_nx_n = 0$ (denoted by $K(G, n)$), where each $g_i \in G$.

Theorem 8 (Olson's Theorem) $|K(G,n)| \geq \max\{1, 2^{n+1-s(G)}\}$.

Note that the group we are interested in is $\mathbb{Z}_m^s$, i.e. an equation in n variables over Z_m^s is equivalent to s equations over $\mathbb{Z}_m$ in the same set of variables. The argment at the beginning of the proof of the Uniformity Lemma to show that it suffices to prove the Lemma for $K^{\mathcal{L}}(0^s)$ applies to the setting of the following corolloary as well, and so Olson's Theorem above gives

Corollary 9 *Let $\mathcal{L} : \mathbb{Z}_m^n \to \mathbb{Z}_m^s$ be a linear map. Then for all $v \in \mathbb{Z}_m^s$ such that $K^{\mathcal{L}}(v)$ is non-empty we have $|K^{\mathcal{L}}(v)| \geq 2^{n+1-s(\mathbb{Z}_m^s)}$.*

To the best of our knowledge, determining $s(\mathbb{Z}_m^s)$ for $s \geq 3$ and arbitrary m, is an open question. However, the independent works of [19, 28] based on Fourier analysis, imply the following upper bound:

Theorem 10 *$s(\mathbb{Z}_m^s) \leq (m \log m)s$.*

Theorem 2 follows by combining Corollary 9 and bound on $s(\mathbb{Z}_m^s)$ given by Theorem 10.

■

3 Lower bounds on number of gates

Consider a $\mathrm{CC}[m]$ circuit C having s MOD_m gates $g_1, \ldots, g_s$. For each gate g_i, we form the linear form $\theta_i = \sum_{j=1}^n c_{i,j}x_j$, where $c_{i,j}$ is the number (modulo m) of copies of input bit x_j feeding into g_i. We thus get at most s non-trivial linear forms that give rise to the linear map $\theta : \{0,1\}^n \to \mathbb{Z}_m^s$. One can easily verify that if $\theta(x) = \theta(y)$, then C outputs the same value on x and y. Let $V \subseteq \mathbb{Z}_m^s$ be the set of those vectors which correspond to the circuit outputting 1 i.e. for every y in V, $\theta(x) = y$ implies that $C(x) = 1$. The size of V is at most m^s. Thus, we obtain the following:

$$\begin{aligned}&\big|\Pr_x[C(x) = 1 \wedge x \in M_{n,q}(a)] - \\ &\Pr_x[C(x) = 1 \wedge x \in M_{n,q}(b)]\big| = \\ &\Bigg|\sum_{y \in V} \big[\Pr_x[\theta(x) = y \wedge x \in M_{n,q}(a)] - \\ &\Pr_x[\theta(x) = y \wedge x \in M_{n,q}(b)]\big]\Bigg|\end{aligned} \tag{11}$$

Using (1) from the Linear Uniformity Lemma and the triangle inequality, one can easily show that the summand in the RHS of (11), for every $y \in V$ is at most $2\gamma^n$, where the constant γ is defined in the Uniformity Lemma. Combining this with the fact that $s = o(n)$, we obtain

$$(11) \leq |V| \cdot 2\gamma^n \leq m^s \cdot 2\gamma^n = 2^{-\Omega(n)}. \tag{12}$$

Since MOD_q is an almost balanced function, i.e. $|\Pr_x[x \in M_{n,q}(a)] - \Pr_x[x \in M_{n,q}(b)]| \leq 2^{-\Omega(n)}$, (12) implies Lemma 3.

As described in the Introduction, a routine argument using the ϵ-discriminator lemma yields the exponential lower bound on the fan-in of the output gate in circuits of type $\mathrm{MAJ} \circ \mathrm{CC}_{o(n)}[m]$ computing MOD_q.

4 Lower bound on the number of wires

In this section we prove superlinear lower bound on the number of wires needed in a CC^0 circuit to compute AND and MOD functions, namely Theorem 6 .

This section is organized as follows. After setting up some notation we prove a superlinear lower bound on the number of edges in bounded depth graphs with a certain connectivity property. The proof is then completed by showing that the circuits in Theorem 6 satisfy this property and hence have superlinear number of edges.

Notation. Let G be a finite directed acyclic graph with a distinguished set of indegree zero vertices V_0, which will be called *input vertices*. Let X be a subset of input vertices. We shall say that a subset of vertices S *separates* X, if for every two different input vertices $x, y \in X$, every vertex v and every pair of directed paths p, q starting in x and y respectively and ending in v, at least one of the paths must contain a vertex from S. S may contain input vertices.

We shall say that X is *ε-separable*, if there exists an S such that S separates X and $|S| \leq \varepsilon|X|$.

We shall say that G is *ε-inseparable*, if for every subset of input vertices X, if $|X| \geq 2$, then X is not ε-separable. ($\varepsilon < 1$, as X separates itself.)

Define, for $d = 1, 2, \ldots$,

$$\lambda_1(n) = \lceil \log_2 n \rceil,$$

$$\lambda_{d+1}(n) = \min\{i \in \mathbb{N};\ \lambda_d^{(i)}(n) \leq 1\},$$

where the superscript i denotes the i-times iterated function.[1]

We can now state the theorem about graphs that we will use for our lower bound on the number of wires.

Theorem 11 *For every $\varepsilon > 0$ and every integer $d \geq 1$, there exists $\delta > 0$ such that for all n, if G has depth d, n inputs and it is ε-inseparable, then it has at least $\delta n \lambda_d(n)$ edges.*

We shall prove a stronger version of this theorem. For a set of inputs X of G, define

$$s(X) = \min\{|S|;\ S \text{ separates } X\}.$$

Let n be the number of input vertices, let $2 \leq t \leq n$, and $\varepsilon > 0$. We shall say that G is *weakly t, ε-inseparable*, if for all k, $t \leq k \leq n$,

$$\mathop{\mathbf{E}}_{|X|=k}(s(X)) > \varepsilon k.$$

[1]Note that the functions λ_i defined in [23] are different.

The greater generality (in particular, the bound on the expectation, instead of an absolute bound) is needed for the proof.

Theorem 12 *For every $\varepsilon > 0$ and every integer $d \geq 1$, there exists $\delta > 0$ such that for every $2 \leq t \leq n$, every weakly t, ε-inseparable G of depth d with n input vertices has at least $\delta n \lambda_d(\frac{n}{t})$ edges.*

This theorem is proved by induction on the depth d. We shall assume w.l.o.g. that G is stratified into levels $V_0, V_1, \cdots, V_d$ and edges are only between consecutive levels. The following two lemmas formalize the induction base and the induction step.

Lemma 13 *For every $\varepsilon > 0$, there exists $\delta > 0$ such that if G has depth 1, has n input vertices and it is weakly t, ε-inseparable, where $2 \leq t \leq n$, then it has more than $\delta n \log \frac{n}{t}$ edges.*

Proof: Suppose G is weakly t, ε-inseparable. Let $v_1, v_2, \ldots$ be all vertices on the level 1 (the level 0 being the input vertices) ordered by the decreasing indegrees $d_1 \geq d_2 \geq \ldots$. For $t \leq q \leq \frac{\varepsilon n}{2}$ consider the undirected graph H_q with the set of vertices being the input vertices of G and edges (x, y) such that $x \to v_i, y \to v_i$ in G for some $i > q$. Thus H_q has $m \leq \sum_{i>q} \binom{d_i}{2}$ edges. Let X be a random subset of inputs of cardinality $k = \lceil \frac{2q}{\varepsilon} \rceil$ (thus $t \leq k \leq n$). The expected number of edges on X is $\frac{m}{\binom{n}{2}} \binom{k}{2}$.

Observe that if there are ℓ edges of H_q on X, then $s(X) \leq \ell + q$ (take the vertices $v_1, \cdots, v_q$ and one vertex from each edge). Thus we have

$$\frac{m}{\binom{n}{2}} \binom{k}{2} + q \geq \mathbf{E}(s(x)) > \varepsilon k.$$

Since $q \leq \varepsilon k/2$, we have

$$\frac{m}{\binom{n}{2}} \binom{k}{2} > \frac{\varepsilon k}{2}.$$

Substituting for m and simplifying we get

$$\sum_{i>q} \frac{\binom{d_i}{2}}{\binom{n}{2}} > \frac{\varepsilon}{k-1}.$$

Since $d_i \leq n$, we can estimate $\frac{\binom{d_i}{2}}{\binom{n}{2}} \leq \frac{d_i^2}{n^2}$. Thus we get

$$\sum_{i>q} \frac{d_i^2}{n^2} > \frac{\varepsilon}{k-1} = \frac{\varepsilon}{\lceil \frac{2q}{\varepsilon} \rceil - 1} \geq \frac{\varepsilon^2}{2q}.$$

By Lemma 4 of [21], this implies

$$\sum_i \frac{d_i}{n} \geq \delta_1 \log \frac{\lfloor \frac{\varepsilon n}{2} \rfloor}{t},$$

for some $\delta_1 > 0$ depending only on ε. Hence if $t = o(n)$, we get

$$\sum_i d_i \geq \delta n \log \frac{n}{t}.$$

Otherwise use the trivial lower bound εt on the number of edges. ∎

Lemma 14 *For every integer $d \geq 1$, reals $\varepsilon > 0$, and $\gamma > 0$, there exists $\delta > 0$ such that for every n, if*

> *(i) for every $2 \leq t \leq n$, every weakly $t, \frac{\varepsilon}{2}$-inseparable G of depth d with n input vertices has at least $\gamma n \lambda_d(\frac{n}{t})$ edges,*

then

> *(ii) for every $2 \leq t \leq n$, every weakly t, ε-inseparable G of depth $d+1$ with n input vertices has at least $\delta n \lambda_{d+1}(\frac{n}{t})$ edges.*

Proof: Suppose (i) holds true. Let G be weakly t, ε-inseparable directed graph with depth $d+1$ and n input vertices.

Let us briefly sketch the idea of the proof before doing detailed computations. We would like to distinguish two cases: either there are a lot of vertices of high degree on the first level, or not. In the first case there are, clearly, many edges. In the second case we can delete the vertices on the first level that have large degrees, connect inputs directly to the second level and then we can apply (i) to the resulting depth d graph. However, this does not quite work, as after deleting the vertices with high degree, the degrees of the remaining vertices on level 1 are still too large. Therefore we have to consider also vertices with intermediate degrees. If the number of those vertices would be small, then a random set of inputs would meet only a few edges connected to them.

Let $deg(v)$ denote the indegree of a vertex v. Let t be given, $2 \leq t \leq n$. Put $r = \frac{n}{t}$,

$$A_0 = \{v \in V_1;\ \deg(v) > \lambda_d(r)\},$$
$$A_i = \{v \in V_1;\ \lambda_d^{(i+1)}(r) < \deg(v) \leq \lambda_d^{(i)}(r)\},$$
for $i \geq 1$.

Let E denote the set of edges of G.

Claim. For every i, $1 \leq i \leq \lambda_{d+1}(r)/2 - 3$, at least one of the following three inequalities is satisfied:

1. $|A_0 \cup \cdots \cup A_{i-1}| \geq \frac{\varepsilon}{4} \frac{n}{\lambda_d^{(i+1)}(r)}$;
2. $|\{(u,v) \in E;\ u \in V_0, v \in A_i \cup A_{i+1} \cup A_{i+2}\}| \geq \frac{\varepsilon}{4} n$;
3. $|\{(u,v) \in E;\ u,v \notin A_0 \cup \cdots \cup A_{i+2}\}| \geq \gamma n \frac{\lambda_d^{(i+2)}(r)}{\lambda_d^{(i+3)}(r)}$.

Proof of Claim. Let i be given and suppose that conditions (1) and (2) are false. Let $n/\lambda_d^{(i+1)}(r) \leq k \leq n$. Observe that $n/\lambda_d^{(i+1)}(r) = n/\lambda_d^{(i+1)}(n/t) \geq t$, since $\lambda_d(x) \leq x$ for all x. Let $X \subseteq V_1$ be a random subset of size k. We shall show that if we remove from G all edges incident with $A_0 \cup \cdots \cup A_{i+2}$, then

$$\mathbf{E}(s'(X)) > \frac{\varepsilon}{2} k,$$

where $s'(X)$ denotes $s(X)$ in the modified graph, which we shall denote by G'.

Indeed, let $a = |A_0 \cup \cdots \cup A_{i-1}|$, $b(X) = |\{(u,v) \in E;\ u \in X, v \in A_i \cup A_{i+1} \cup A_{i+2}\}|$. Then

$$s(X) \leq a + b(X) + s'(X).$$

Hence

$$\begin{aligned}\mathbf{E}(s'(X)) &\geq \mathbf{E}(s(X) - b(X) - a)\\ &= \mathbf{E}(s(X)) - \mathbf{E}(b(X)) - a.\end{aligned}$$

By non-1, $a < \frac{\varepsilon}{4} \frac{n}{\lambda_d^{(i+1)}(r)} \leq \frac{\varepsilon}{4} k$. By non-2, we have $\mathbf{E}(b(X)) < \frac{\varepsilon}{4} k$, (each edge from $\{(u,v) \in E;\ u \in V_0, v \in A_i \cup A_{i+1} \cup A_{i+2}\}$ is chosen with probability k/n; use the linearity of expectation).

Thus G' is weakly $n/\lambda_d^{(i+1)}(r), \frac{\varepsilon}{2}$-inseparable.

We shall further modify G' by removing all edges between V_1 and V_2 and adding, for every path (u,v,w) in G' with $u \in V_0, v \in V_1, w \in V_2$, the edge (u,w). The resulting graph will be denoted by G''. It has depth d (the first level being $V_1 \cup V_2$, the second level being V_3 etc.) and at most $\lambda_d^{(i+3)}(r)$-times more edges.

Furthermore, G'' is also weakly $n/\lambda_d^{(i+1)}(r), \frac{\varepsilon}{2}$-inseparable. To see that, observe that if X is a set of inputs (in G' and G'') and S is a separating set for X in G'', then S is a separating set for X also in G'. Indeed, let S be a separating set for X in G'' and let $(v_0, \cdots, v_j)$ and $(u_0, \cdots, u_j)$ be two paths in G', $v_0, u_0 \in X$, $v_0 \neq u_0$ and $v_j = u_j$. Then if $j = 1$, these paths are also paths in G'', and if $j > 1$, $(v_0, v_2, \cdots, v_j)$ and $(u_0, u_2, \cdots, u_j)$ are paths in G''. In both cases they contain an element from S, whence the original pair of paths also contains an element from S. Thus separating sets are at least as large in G'' as in G'.

By the assumption (i), G'' must have at least $\gamma n \lambda_d(\lambda_d^{(i+1)}(r)) = \gamma n \lambda_d^{(i+2)}(r)$ edges. Hence G' has at least $\gamma n \lambda_d^{(i+2)}(r)/\lambda_d^{(i+3)}(r)$ edges, which proves 3. This finishes the proof of the Claim.

To finish the proof of Lemma 14, we shall use the inequality

$$\frac{\lambda_d^{(i)}(r)}{\lambda_d^{(i+1)}(r)} \geq \tfrac{1}{2}\lambda_{d+1}(r),$$

for every $i \leq \lambda_{d+1}(r)/2 - 1$, which was proved in [21] as Lemma 5. By the Claim it suffices to consider the following three cases.

1. Suppose for some $i \leq \lambda_{d+1}(r)/2 - 3$ the condition (i) of Claim is satisfied. Then, since every $v \in A_0 \cup \cdots \cup A_{i-1}$ has degree $> \lambda_d^{(i)}(r)$, the number of edges in G is at least

$$\frac{\varepsilon}{4}\frac{n}{\lambda_d^{(i+1)}(r)}\lambda_d^{(i)}(r) \geq \frac{\varepsilon}{8}n\lambda_{d+1}(r).$$

2. Suppose for *all* $i \leq \lambda_{d+1}(r)/2 - 3$ the condition (ii) of Claim is satisfied. Then the number of edges of G is at least

$$\tfrac{1}{3}(\lambda_{d+1}(r)/2 - 3)\frac{\varepsilon}{4}n = \Omega(n\lambda_{d+1}(r)).$$

3. Suppose for some $i \leq \lambda_{d+1}(r)/2 - 3$ the condition (iii) of Claim is satisfied. Then the number of edges of G is at least

$$\gamma n \frac{\lambda_d^{(i+2)}(r)}{\lambda_d^{(i+3)}(r)} \geq \tfrac{1}{2}\gamma n \lambda_{d+1}(r).$$

■

Proof:[Proof of Theorem 6] Without loss of generality, it suffices to show lower bound for the function $\overline{x}_1 \wedge \ldots \overline{x}_n$ instead of the AND function. Thus assume that the circuits C computes function $F(x_1, \ldots, x_n)$ which is either the function $\overline{x}_1 \wedge \ldots \overline{x}_n$ or the function MOD_q where $(m, q) = 1$.

Let $0 < \varepsilon < \gamma$, let $\delta > 0$ be given by Theorem 11 for these ε and d. Suppose that the circuit has $< \delta n \lambda_d(n)$ edges. Then, by Theorem 11, there exists a set of inputs X which is ε-separated in the depth d graph obtained by removing the output gate from the circuit. Let S be the separating set augmented with the output gate. Then S is a separating set in the whole circuit and $|S| \leq \varepsilon|X| + 1$. We may moreover require that $|X| \geq \log n$, thus if n is sufficiently large, $|S| \leq \gamma|X|$.

Furthermore, for every $v \in S$, disconnect v from its inputs and set it to be the constant equal to the boolean value computed at v when all inputs are 0. Let C' be the resulting circuit. Let $v \in S$ and let w be an input gate of v in C. Then in C', the gate w only depends on at most one input from X, because S is a separating set. Thus if we put back the original MOD_m gate on v, the boolean function computed at v will be some MOD_m function E_v: that is, there is a linear form $\theta(x_1, \ldots, x_n)$ over $\mathbb{Z}_m$, and $A \subseteq \{0, \ldots, m-1\}$ such that $E_v(x) = 1$ iff $\theta(x) \in A$.

Thus in order to get a contradiction with the assumption that C computes $F(x_1, \ldots, x_n)$ we need only find a boolean assignment $a \neq 0^n$ of $x_1, \ldots, x_n$ such that the variables outside X are set to 0 and the following holds: For every $v \in S$

$$E_v(a) = E_v(0^n), \tag{13}$$

but $F(a) \neq F(0^n)$.

On the left hand side of (13) we replace each boolean function $E_v(\cdot)$ by its underlying linear form that takes values in $\mathbb{Z}_m$.

Then if the resulting linear system over $\mathbb{Z}_m$ is satisfied then so is (13). The assumption that F is either $\overline{x}_1 \wedge \ldots \overline{x}_n$ or MOD_q with $(q, m) = 1$ guarantees the existence of a boolean solution $a \neq 0^n$ to this system such that $F(a) \neq F(0^n)$ by the Linear Uniformity Lemma. Thus C cannot compute AND or MOD_q. ■

5 Generalization to low degree polynomials

In this section we generalize Theorem 4 to Theorem 7 by extending the Linear Uniformity Lemma to systems of equations involoving low degree polynomials. We shall need the following estimate from [9]:

Fact 15 *Let q, m be any relatively prime numbers. Further, let $\phi(x) = \phi(x_1, \ldots, x_n)$ be any polynomial of degree d with coefficients in $\mathbb{Z}_m$. Then there exists $0 < \alpha(m, q) < 1$, such that*

$$\sum_{x \in \{0,1\}^n} e_m(\phi(x)) e_q(a \sum_{i=1}^{n} x_i) \leq 2^n e^{-\frac{\alpha n}{(m 2^m)^d}}, \tag{14}$$

whenever $a \neq 0 \mod q$.

We now show that one can easily use (14) to extend our Linear Uniformity Lemma to systems of polynomial equations of small degree.

Let $\mathcal{S} = \{\phi_1, \ldots, \phi_r\}$ be a set of s polynomials over $\mathbb{Z}_m$, where ϕ_i has degree d_i. Let $\Delta = \Delta(\mathcal{S}) = \max_{1 \le i \le s} d_i$ be the maximum degree among all polynomials in $\mathcal{S}$. For $v \in \mathbb{Z}_m^s$, let $K^{\mathcal{S}}(v)$ represent the set of points in $\{0,1\}^n$, that satisfy $\phi_i = v_i$ for all $1 \le i \le s$. We show the following :

Lemma 16 (General Uniformity Lemma) *Using the notation above, for all positive integers m, q, with $(m, q) = 1$, there exist constants $\alpha, \beta > 0$ such that for all polynomial mapping $\mathcal{S} : \mathbb{Z}_m^n \to \mathbb{Z}_m^s$, we have*

$$\big||K^{\mathcal{S}}(v) \cap M_{n,q}(b)| - |K^{\mathcal{S}}|/q\big| \le \left(\frac{2}{e^{\alpha/\beta^{\Delta}}}\right)^n. \tag{15}$$

for each $b \in \{0, \ldots, q-1\}$ and vector $v \in \mathbb{Z}_m^s$.

Proof: One can easily mimick the first few steps of the proof of the Uniformity Lemma from Section 2 to obtain the following :

$$\begin{aligned} &\big||K^{\mathcal{S}} \cap M_{n,q}(b)| - |K^{\mathcal{S}}|/q\big| = \\ &\Bigg| \sum_{x \in \{0,1\}^n} \Bigg[\prod_{i=1}^{s} \Big(\frac{1}{m} \sum_{j=0}^{m-1} \mathrm{e}_m(j\phi_i(x))\Big) \\ &\times \Big(\frac{1}{q} \sum_{j=1}^{q-1} \mathrm{e}_q\big(j(\sum_{k=1}^{n} x_k - b)\big)\Big) \Bigg] \Bigg|. \end{aligned} \tag{16}$$

As in Section 2, we multiply out the terms in the summand above. This gives us $m^s(q-1)$ terms, each of which is of the form below:

$$\begin{aligned} &\frac{\mathrm{e}_q(-jb)}{m^s q} \sum_{x \in \{0,1\}^n} \big[\mathrm{e}_m(j_1\phi_1(x) + \ldots + j_s\phi_s(x)) \\ &\times \mathrm{e}_q(j \sum_{k=1}^{n} x_k)\big], \end{aligned} \tag{17}$$

where $j \in [q-1]$ and $(j_1, \ldots, j_s) \in \{0, 1, \ldots, m-1\}^s$. The degree of the form $j_1\phi_1(x) + \ldots + j_s\phi_s(x)$ is at most $\Delta(\mathcal{S})$ for every $(j_1, \ldots, j_s)$. Using the bound on (14) in Fact 15, one can write

$$(17) \le \frac{q-1}{q}\Big(\frac{2}{e^{\alpha/\beta^{\Delta}}}\Big)^n \tag{18}$$

It can be easily verified that the bound of (18) easily yields (15) in the General Uniformity Lemma. ■

We apply an argument similar to that applied to prove Lemma 3 from the Linear Uniformity Lemma (see Section 3), to inequlaity (15). This shows that for every $0 < \epsilon < 1$, there exists a constant $\delta = \delta(m, q, \epsilon) > 0$ such that

$$\begin{aligned} &|\Pr_x[C(x) = 1 | x \in M_{n,q}(a)] \\ &- \Pr_x[C(x) = 1 | x \in M_{n,q}(b)]| \le 2^{-\Omega(n)}, \end{aligned} \tag{19}$$

where C is circuit of type $\mathrm{CC}_{o(n)}[m] \circ \mathrm{AND}_{\delta \log n}$. Using the ϵ-discriminator lemma we immediately get Theorem 7.

Note that our method again yields a bound for a more general class of circuits i.e. of type $\mathrm{MAJ} \circ \mathrm{ANY} \circ \mathrm{CC}_{o(n)}[m] \circ \mathrm{AND}_{\delta \log n}$ where the sub-circuit rooted at each ANY gate uses sub-linear number of MOD_m gates and each input of an ANY gate is the output of a MOD_m gate. Using Hastad's switch Lemma and our results above, one can show that super-polynomial size is needed when the bottom AND gates are replaced by poly-sized AC^0 circuits.

We believe that circuits of type $\mathrm{MOD}_m \circ \mathrm{MOD}_m \circ \mathrm{AND}_{O(1)}$ need exponential number of MOD_m gates to compute the AND function. We cannot even show a linear lower bound on the number of MOD_m gates at the moment. To show that, it would be sufficient to show that $K^{\mathcal{S}}$ is large if $\mathcal{S}$ has sub-linear number of equations. This looks like a problem of independent interest.

Acknowledgements

We thank Kristoffer Arnsfelt Hansen for pointing out some inconsistencies in the preliminary version of this paper that led to a better presentation of our results. We also thank him for drawing our attention to his paper [13].

References

[1] E. Allender A note on the power of threshold circuits *FOCS* (1989) 580–584

[2] D. Barrington. Some problems involving Razborov-Smolensky polynomials. *Boolean function complexity* (Durham, 1990) London Math. Soc. Lecture Note Ser., 169 (1992) 109–128. Cambridge Univ. Press, Cambridge.

[3] D. A. M. Barrington, R. Beigel, S. Rudich. Representing Boolean Functions as Polynomials Modulo Composite Numbers. *Computational Complexity* 4 (1994) 367–382.

[4] D. Barrington, H. Straubing and D. Thérien. Non-uniform automata over groups. *Information and Computation* 89(2) (1990) 109–132.

[5] R. Beigel and J. Tarui On ACC *Computational Complexity* 4 (1994) 350–366

[6] J. Bourgain. Estimation of certain exponential sums arising in complexity theory. *C. R. Acad. Sci. Paris*, Ser I 340 (2005), no. 9, 627–631.

[7] H. Caussinus. A Note on a Theorem of Barrington, Straubing and Thérien. *Inf. Process. Lett.* 58(1) (1996) 31–33.

[8] A. K. Chandra, S. Fortune, R. J. Lipton. Lower Bounds for Constant Depth Circuits for Prefix Problems. *ICALP* 1983 109–117.

[9] F. Green, A. Roy and H. Straubing. Bounds on an exponential sum arising in boolean circuit complexity. *C. R. Acad. Sci. Paris*, Ser I 341 (2005), 279–282.

[10] V. Grolmusz. A Degree-Decreasing Lemma for (MOD-q - MOD-p) Circuits. *Discrete Mathematics and Theoretical Computer Science*, 4(2) (2001) 247–254.

[11] V. Grolmusz, G. Tardos. Lower Bounds for (MOD-p - MOD-m) Circuits. *SIAM J. Comput.* 29(4) (2000) 1209–1222.

[12] K. A. Hansen. On Modular Counting with Polynomials. *Computational Complexity Conference 2006* 202–209.

[13] K. A. Hansen. Lower Bounds for Circuits with Few Modular Gates using Exponential Sums. *Electronic Colloquium on Computational Complexity*, TR06-079.

[14] J. Håstad. Computational limitations on small depth circuits. Ph.D Thesis, MIT, (1986).

[15] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy and G.Turán. Threshold circuits of bounded depth. *J. Comput. Syst. Sci* 46(2) (1993) 129–154.

[16] M. Koucký, P. Pudlák and D. Thérien. Boundeddepth circuits: Separating wires from gates. *37th ACM STOC,* 2005, pp.257-265.

[17] M. Krause, S. Waack. Variation Ranks of Communication Matrices and Lower Bounds for Depth Two Circuits Having Symmetric Gates with Unbounded Fan-In. *FOCS* 777–782 (1991).

[18] P. McKenzie, P. Péladeau, D. Thérien. NC^1: The Automata-Theoretic Viewpoint. *Computational Complexity* 1: 330-359 (1991).

[19] R. Meshulam. An uncertainty inequality and zero subsums. *Discrete Mathematics* (84) 1990, 197–200.

[20] J.E. Olson. A combinatorial problem on finite abelian groups, II. *J. Number Theory* 1 (1969), 195–199.

[21] P. Pudlák. Communication in bounded depth circuits. *Combinatorica* 14(2), pp.203–216, 1994.

[22] P. Ragde and A. Wigderson. Linear-size constant-depth polylog-threshold circuits. *Information Processing Letters*, 39 (1991), 143–146.

[23] R. Raz and A. Shpilka. Lower bounds for matrix product in bounded depth circuits with arbitrary gates. *SIAM J. on Computing* 32(2), 2003, pp.488-513.

[24] A. Razborov. Lower bounds on the size of bounded-depth networks over a complete basis with logical addition. *Mathematical Notes of the Academy of Sci. of the USSR*, 41(4): 333–338, 1987.

[25] R. Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. *19th STOC* (1987), 77–82.

[26] R. Smolensky. On interpolation by analytic functions with special properties and some weak lower bounds on the size of circuits with symmetric gates. *31st FOCS* (1990), 628–631.

[27] H. Straubing, D. Thérien. A Note on MODp-MODm Circuits. accepted in Theory of Computing Systems.

[28] D. Thérien. Circuits constructed with MOD_q gates cannot compute AND in sublinear size. Comput. Complexity 4 (1994), no. 4, 383–388.

[29] P. Y. Yan, I. Parberry. Exponential Size Lower Bounds for Some Depth Three Circuits Inf. Comput. 112(1): 117–130 (1994)