Linear Systems Over Finite Abelian Groups

Arkadev Chattopadhyay
Department of Computer Science
University of Toronto
Toronto, Canada
arkadev@cs.toronto.edu

Shachar Lovett
School of Mathematics
Institute for Advanced Study
Princeton, USA
slovett@math.ias.edu

***Abstract*—We consider a system of linear constraints over any finite Abelian group G of the following form: $\ell_i(x_1, \ldots, x_n) \equiv \ell_{i,1}x_1 + \cdots + \ell_{i,n}x_n \in A_i$ for $i = 1, \ldots, N$ and each $A_i \subset G$, $\ell_{i,j}$ is an element of G and x_i's are Boolean variables. Our main result shows that the subset of the Boolean cube that satisfies these constraints has exponentially small correlation with the $\mathbf{MOD}_q$ boolean function, when the order of G and q are co-prime numbers.**

Our work extends the recent result of Chattopadhyay and Wigderson (FOCS'09) who obtain such a correlation bound for linear systems over cyclic groups whose order is a product of two distinct primes or has at most one prime factor. Our result also immediately yields the first exponential bounds on the size of boolean depth-four circuits of the form $\mathbf{MAJ} \circ \mathbf{AND} \circ \mathbf{ANY}_{O(1)} \circ \mathbf{MOD}_m$ for computing the $\mathbf{MOD}_q$ function, when m, q are co-prime. No superpolynomial lower bounds were known for such circuits for computing any explicit function.

This completely solves an open problem posed by Beigel and Maciel (Complexity'97).

***Keywords*-lower bounds; boolean circuit complexity; modular gates; composite moduli; exponential sums;**

I. Introduction

A fundamental open problem in theoretical computer science is to understand the computational power of counting modulo composite numbers. For example, we do not know if hard problems like SATISFIABILITY have efficient depth-three circuits comprising only MOD_6 gates. This is in contrast to the classical theorem of Razborov [1] and Smolensky [2] that says constant-depth circuits having only AND, OR and MOD_m gates cannot compute even the MOD_q function in sub-exponential size, i.e. in size $2^{n^{o(1)}}$, when m, q are co-prime and when m has only one prime factor. Smolensky [2] conjectured that this theorem extends to all m, but despite a series of attempts over two decades, this conjecture remains wide open. While Smolensky's conjecture easily implies that not all functions computable in deterministic linear time have efficient ACC^0 circuits, the best one can prove is the recent breakthrough result of Williams [3] who showed that non-deterministic exponential time does not have efficient ACC^0 circuits.

A recent result of Hansen and Koucky [4] shows that every function in $\mathrm{ACC}^0[m]$ is in fact computed by a quasipolynomial size circuit of the canonical form $\mathrm{OR} \circ \mathrm{AND} \circ \mathrm{CC}^0[m]$, where $\mathrm{CC}^0[m]$ refers to the class of constant-depth circuits just comprising MOD_m gates. Given this characterization, a natural first step towards proving Smolensky's conjecture is to verify it for depth-three circuits $\mathrm{OR} \circ \mathrm{AND} \circ \mathrm{MOD}_m$ with just one layer of MOD_m gates at the base. However, this step was long identified by Beigel and Maciel [5] as a barrier. They observed that there are no known techniques to prove strong lower bounds on the size of such depth-three circuits when the MOD_m gates at the base are generalized in the following sense: each such gate has an associated *accepting set* A and the gates output 1 iff the sum of the input bits evaluates to an element in A modulo m. Interestingly, if each gate at the base had a singleton accepting set then they could prove very strong lower bounds, but their methods failed for general accepting sets. The problem of handling such general accepting sets is not specific to their work but is well known to researchers. For example, it is not hard to show that depth-two circuits having MOD_m gates cannot compute even the AND function in sub-exponential size when the output gate has a singleton accepting set, while it is consistent with our current knowledge that such circuits in linear size compute SATISFIABILITY for an appropriate choice of accepting set for the output gate (see Caussinus [6]).

It is known that the choice of accepting sets makes a non-trivial difference in the closely related world of polynomial representation of boolean functions. For example, polynomials over the ring $\mathbb{Z}_m$, need degree $\Omega(n)$ to compute simple functions like AND, OR and MOD_q when the accepting set is a singleton. However, Barrington, Beigel and Rudich [7] gave an elegant and surprising construction for computing AND and OR with polynomials of degree $O(n^{1/t})$ having proper

accepting sets, where m has t distinct prime factors. Hansen [8] showed that judicious choice of accepting sets affords similar advantage for computing MOD_q. No superlogarithmic lower bound on the degree of polynomials with general accepting sets is known for computing any function in NP. On the positive side, the construction of Barrington *et al* has led to other interesting constructions outside of circuit complexity. For example, all known bounds on explicit constructions of Ramsey graphs can be achieved using it [9]. Further, a series of recent breakthroughs in constructing more efficient locally decodable codes [10], [11] have crucially relied on the construction of Barrington *et al* .

Recently, Chattopadhyay and Wigderson [12] attacked this depth-three question by naturally considering a system of linear constraints of the form $\ell_i \in A_i$ for $i = 1, \ldots, N$, where each $A_i \subset \mathbb{Z}_m$ and ℓ_i's are linear forms. Their main result gives an exponentially small upper bound on the correlation of the Boolean solution set of any such system with the MOD_q function, when m is either a prime power or is a product of two distinct primes like 6. This implied the first exponential lower bounds on the size of depth-three circuits of the form $\mathrm{MAJ}\circ\mathrm{AND}\circ\mathrm{MOD}_m^A$ for computing MOD_q, for such m. No superpolynomial lower bounds were known before on the size of such circuits.

Our Work: We extend the result of Chattopadhyay and Wigderson to arbitrary, fixed m. More generally, our main result is the following: for any system of linear constraints $\mathcal{L}$, let $B_{\mathcal{L}}$ be the set of points in the Boolean cube that satisfies $\mathcal{L}$. The *correlation* of a set S of boolean points with the MOD_q function, denoted by $\mathrm{Corr}\big(S, \mathrm{MOD}_q\big)$, is defined as $\max_{a,b}\big|\Pr_x[x \in S \wedge \mathrm{MOD}_q^{\{a\}}(x) = 1] - \Pr_x[x \in S \wedge \mathrm{MOD}_q^{\{b\}}(x) = 1]\big|$.

Theorem 1 (Main Theorem)**.** *Let m be an arbitrary fixed positive integer and $\mathcal{L}$ be a system of linear constraints over n variables of the following form: $\ell_i\big(x_1, \ldots, x_n\big) \in A_i$ for $i = 1, \ldots, t$, where each $A_i \subset \mathbb{Z}_m$ and ℓ_i is a linear form over $\mathbb{Z}_m$. Then, $Corr\big(B_{\mathcal{L}}, \mathrm{MOD}_q\big) \leq 2^{-\Omega(n)}$ when m and q are co-prime.*

Furthermore, it generalizes to finite Abelian groups:

Theorem 2. *Let G be any finite and fixed Abelian group and $\mathcal{L}$ be a system of linear constraints with n boolean variables where the coefficients of each constraint are elements of G. Then, $Corr\big(B_{\mathcal{L}}, \mathrm{MOD}_q\big) \leq 2^{-\Omega(n)}$ when the order of G and q are co-prime.*

A direct consequence of our Theorem 2, obtained by an easy application of the so call ϵ-discriminator Lemma of Hajnal et.al. [13] (restated equivalently in Section II of this article), is the following exponential lower bound on the size of boolean circuits:

Corollary 3. *Let GMOD_m denote mod-m gates with general accepting sets. Then, depth-four circuits of the form MAJ $\circ$ AND $\circ$ $\mathrm{ANY}_{O(1)} \circ \mathrm{GMOD}_m$ require exponential fan-in at the output Majority gate to compute the MOD_q function, if m, q are co-prime.*

Beigel and Maciel [5] identified the problem of proving lower bounds for depth-three circuits of the form $\mathrm{MAJ}\circ\mathrm{AND}\circ\mathrm{GMOD}_m$ as an important next step towards understanding circuits having modular gates. Corollary 3 completely solves this problem by obtaining the first strong lower bounds for such circuits.

In the language of Barrington and Thérien [14], the result of Hansen and Koucky implies that one way of proving Smolensky's conjecture is to show that functions computed by systems of programs over finite solvable groups do not correlate well with the MOD_q function if q is co-prime with the order of the group. Our result takes the first step in this direction by verifying this for Abelian groups.

Proof: Note that for any $s = O(1)$, $\mathrm{ANY}_s \circ \mathrm{GMOD}_m \subset \mathrm{GMOD}_{(\mathbb{Z}_m)^s}$. Let $G = (\mathbb{Z}_m)^s$. Theorem 2 gives that the MOD_q has exponentially small correlation with any function in $\mathrm{AND}\circ\mathrm{GMOD}_G$. Hence $\mathrm{MAJ}\circ\mathrm{AND}\circ\mathrm{GMOD}_G$ circuits that compute the MOD_q function require exponential size. ■

No strong lower bounds were known for computing any explicit function by such circuits.

Our Technique: In the world of arithmetic circuits, Grigoriev and Razborov [15] introduced the ingenious notion of communication rank for linear systems over a finite fields. Chattopadhyay and Wigderson [12] generalized this notion to systems over $\mathbb{Z}_m$ for an arbitrary composite m. Using this notion, [12] showed that if a system has high rank then it is highly unsatisfiable over the boolean cube and if they have low rank, exploiting estimates of exponential sums by Bourgain, they showed that the correlation of the solution set to the MOD_q function is small. For technical reasons, their analysis of the low rank case only worked if m was a product of two distinct primes or had just one prime factor.

In this work, we realize that in order to work with arbitrary composites, it is convenient to consider more general systems of linear constraints. We consider constraints in which the accepting set is itself a function of a constant number of variables as opposed to being a fixed set as in the work of Chattopadhyay and Wigderson. This leads us to further generalize the notion of communication rank to facilitate analysis of such

linear systems. In particular, we consider an iterative simplification process of linear systems where this new notion of rank plays a crucial role. This simplification, driven by our Lemma 8 in Section III, is the key new ingredient of our work that allows us to work with arbitrary modulus m. A rough description of the main idea is as follows: either our system has large communication rank in which case it is highly unsatisfiable or it has low rank in which case we simplify it in the following sense. Each constraint in the simplified system has either a singleton accepting set or the number of variables on which the accepting set depends is one less than before or the system is over a modulus m' that is less than m. A repeated application of this procedure yields a nice structural result: every generalized linear system L over $\mathbb{Z}_m$ over n variables can be decomposed into at most $t = 2^{\epsilon n}$ linear systems $L_1, \ldots, L_t$ where each L_i is either satisfied by an exponentially small fraction of the points in the boolean cube or L_i is the intersection of L_i^0 and L_i^1 where every constraint in L_i^0 has a singleton accepting set and each constraint of L_i^1 corresponds to a $k = k(m)$-junta.

It is already known from the work of Chattopadhyay and Wigderson, restated in Lemma 7 of this work, that subsytems that are intersections of singleton systems over $\mathbb{Z}_m$ and junta systems have exponentially small correlation with MOD_q. The subsystems of our decomposition that have poor satisfiability cannot, by definition, correlate with a much more balanced function like MOD_q. Since there are only few subsytems in the decomposition, an easy application of the union bound finishes the argument.

Paper organization: We give basic definitions and recall the necessary background in Section II. We prove our result for arbitrary cyclic groups $\mathbb{Z}_m$ in Section III. For lack of space, we defer the proof for general Abelian groups to the full version of this paper.

II. Preliminaries

Let $\mathbb{Z}_q := \{0, \ldots, q-1\}$ and $\mathbb{N} = \{0, 1, 2, \ldots\}$ denote the natural numbers. We study the correlation of subsets $S \subset \{0,1\}^N$ with the sum modulo q. It is natural and convenient to estimate this quantity using the q-th roots of unity. Let $\mathrm{e}_q(y) := \exp\left(2\pi i y/q\right)$, where i denotes the complex square-root of unity. We will use in the paper the following definition for correlation:

$$\mathrm{Corr}(S, \mathrm{MOD}_q) := \max_{b \in \mathbb{Z}_q \setminus \{0\}} \left| \mathbb{E}_{x \in \{0,1\}^n} [1_S(x) \cdot e_q(b(x_1 + \ldots + x_n))] \right|,$$

where 1_S is the indicator function of S. It is straightforward to verify that

$$\max_{a,b \in \mathbb{Z}_q} \big| \Pr_x[x \in S \wedge \mathrm{MOD}_q^{\{a\}}(x) = 1] - \Pr_x[x \in S \wedge \mathrm{MOD}_q^{\{b\}}(x) = 1] \big| \leq 2 \cdot \mathrm{Corr}(S, \mathrm{MOD}_q),$$

so our definition indeed captures the more intuitive definition of having elements of S being approximately equidistributed modulo q. For a family of subsets $\mathcal{S}_N = \{S \subset \{0,1\}^N\}$ we define their correlation with sums modulo q as the maximal correlation for $S \in \mathcal{S}_N$.

The simple tool that we use for lower bounding the size of our circuits for computing MOD_q is the so-called ϵ-Discriminator Lemma, introduced by Hajnal et.al.[13]. We state here a specialized version of it that is particularly convenient for our work, and has been also used in earlier works (see for example [16], [17]).

Lemma 4 (Discriminator Lemma)**.** *Let C be a circuit that has a MAJORITY gate at its output that is being fed by t subcircuits $C_1, \ldots, C_t$. If C computes MOD_q, then there exists a subcircuit C_i, such that Corr$\big(C_i, \mathrm{MOD}_q\big) = \Omega(1/t)$.*

In order to estimate the correlation of solution sets of linear systems with MOD_q function, we will need estimates of exponential sums that were first obtained in the work of Bourgain [16] and refined progressively in further works [18], [19], [20]. We state the most refined estimate below:

Theorem 5 ([20])**.** *Let m, q be two fixed positive coprime integers and let P be any n-variate multilinear polynomial of degree d with coefficients in $\mathbb{Z}_m$ and b be any number non-zero modulo q. Then, there exists a constant $\beta = \beta(m, q)$ such that the following holds:*

$$\left| \mathbb{E}_{x \in \{0,1\}^n} \left[\mathrm{e}_m\big(P(x)\big) \mathrm{e}_q\left(b \sum_i x_i \right) \right] \right| \leq \exp(-\beta^d n). \tag{1}$$

We point out that the above estimate fails to give anything non-trivial when the degree d of the polynomial P is more than $\log n$. Finding exponentially small upper bounds for the exponential sum in (1) for $d > \log n$, even when m is prime, remains a very interesting open problem.

III. Linear systems over cyclic groups

We study in this section systems of linear equations with arbitrary accepting sets over arbitrary (constant) moduli.

Definition 1 (Linear system with accepting sets)**.** *A linear equation in n binary variables over $\mathbb{Z}_m$ with an accepting set is the set of solutions (over $\{0,1\}^n$) to an equation of the form*

$$E = \{x \in \{0,1\}^n : \sum a_i x_i \text{ (mod } m) \in A\},$$

where $a_1, \ldots, a_n \in \mathbb{Z}_m$ and $A \subsetneq \mathbb{Z}_m$.

A linear system in n binary variables over $\mathbb{Z}_m$ with accepting sets is the set of common solutions to several such equations, i.e. it is a subset of $\{0,1\}^n$ of the form

$$L = E_1 \cap \ldots \cap E_N =$$
$$\{x \in \{0,1\}^n : \sum a_{i,j} x_j \text{ (mod } m) \in A_i \ \ \forall 1 \leq i \leq N\}$$

where $a_{i,j} \in \mathbb{Z}_m$ and $A_i \subsetneq \mathbb{Z}_m$. We denote by $\mathcal{L}_n(m)$ the family of all such linear systems (where we do not distinguish the number of equations).

We aim to bound the correlation of the solution set of linear systems with the MOD_q function. Our main approach is to iteratively simplify the system. In order for us to define these simplifications, we need some further definitions of more general systems of linear forms. We first define the special case of a linear system all of whose accepting sets are singletons, i.e. consist of a single value.

Definition 2 (Singleton linear systems)**.** *A linear equation in n binary variables over $\mathbb{Z}_m$ with a single accepting value is the set of solutions (over $\{0,1\}^n$) to an equation of the form*

$$E = \{x \in \{0,1\}^n : \sum a_i x_i \equiv b \text{ (mod } m)\},$$

where $a_1, \ldots, a_n, b \in \mathbb{Z}_m$.

A singleton linear system in n binary variables over $\mathbb{Z}_m$ is the set of common solutions to several such equations, i.e. it is a subset of $\{0,1\}^n$ of the form

$$L = E_1 \cap \ldots \cap E_N =$$
$$\{x \in \{0,1\}^n : \sum a_{i,j} x_j \equiv b_i \text{ (mod } m) \ \ \forall 1 \leq i \leq N\},$$

where $a_{i,j}, b_j \in \mathbb{Z}_m$. We denote by $\mathcal{L}_n^{Sing}(m)$ the family of all such linear systems (where we do not distinguish the number of equations).

We will also need the following generalized notions of linear systems with accepting sets which depend on a few variables in an arbitrary manner.

Definition 3 (Linear systems with accepting sets of sparsity k)**.** *A linear equation in n binary variables over $\mathbb{Z}_m$ with an accepting set of sparsity k is the set of solutions (over $\{0,1\}^n$) to a linear equation with an accepting set which depends on k of the variables, i.e. to an equation of the form*

$$E = \{x \in \{0,1\}^n : \sum a_i x_i \text{ (mod } m) \in A(x_{i_1}, \ldots, x_{i_k})\},$$

where $a_1, \ldots, a_n \in \mathbb{Z}_m$, $i_1, \ldots, i_k \in [n]$ and each set $A(x_{i_1}, \ldots, x_{i_k})$ is a subset of $\mathbb{Z}_m$ for every setting of $x_{i_1}, \ldots, x_{i_k} \in \{0,1\}^k$. Moreover, we require that the accepting-set function A is not trivial, i.e. $A(x_{i_1}, \ldots, x_{i_k}) \subsetneq \mathbb{Z}_m$ for at least one setting of $x_{i_1}, \ldots, x_{i_k}$.

A linear system in n binary variables over $\mathbb{Z}_m$ with accepting sets of sparsity k is the set of common solutions to several such equations, i.e. it is a subset of $\{0,1\}^n$ of the form

$$L = E_1 \cap \ldots \cap E_N =$$
$$\{x \in \{0,1\}^n : \sum a_{i,j} x_j \text{ (mod } m) \in A_i(x_{i_1}, \ldots, x_{i_k}) \ \ \forall i = 1, \ldots, N\}$$

where $a_{i,j}, b_j \in \mathbb{Z}_m$, and each accepting set-function A_i is not trivial. We denote by $\mathcal{L}_n(m,k)$ the family of all such linear systems (where we do not distinguish the number of equations in the system).

Note that $\mathcal{L}_n(m,0) = \mathcal{L}_n(m)$. For $k > 1$, we also allow modulus $m = 1$, in which case we interpret each equation in $\mathcal{L}_n(1,k)$ as $E = \{x : 0 \in A(x_{i_1}, \ldots, x_{i_k})\}$, where set function A is not trivial. The linear system is the set of common solutions to several such equations.

We now define the most general linear system, which will be simplified iteratively in the proof of our main theorem. It will be the intersection of systems over several moduli ℓ which divide m. Let $\ell \div m$ denote "ℓ divides m". We define linear systems which are intersections of linear systems in $\mathcal{L}_n^{Sing}(m)$ and $\mathcal{L}_n(\ell,k)$ for several $\ell \div m$. For a modulus m define $\mathrm{div}(m) = \{1 \leq \ell \leq m : \ell \div m\}$ to be the set of (not necessarily prime) factors of m. We will maintain a sparsity function $\kappa : \mathrm{div}(m) \to \mathbb{N} \cup \{-\infty\}$ which will specify the allowed sparsity for each $\ell \div m$. That is, we will have $\mathcal{L}_n(\ell, \kappa(\ell))$ systems for all $\ell \in \mathrm{div}(m)$ such that $\kappa(\ell) \geq 0$, where $\kappa(\ell) = -\infty$ means we have no $\mathcal{L}_n(\ell, \cdot)$ system. Note that if $\ell \div m$ and $\kappa : \mathrm{div}(\ell) \to \mathbb{N} \cup \{-\infty\}$ then $\mathcal{L}_n(\ell, \kappa(\cdot)) \subset \mathcal{L}_n(m, \kappa(\cdot))$ since equations modulo ℓ can always be lifted to equations modulo m by multiplying them by m/ℓ.

Definition 4 (Linear systems with general accepting sets over several moduli)**.** *Let m be a modulus and let $\kappa : \mathrm{div}(m) \to \mathbb{N} \cup \{-\infty\}$. We define $\mathcal{L}_n(m, \kappa(\cdot))$ as follows: $L \in \mathcal{L}_n(m, \kappa(\cdot))$ if there exists $L^{Sing} \in$*

$\mathcal{L}_n^{Sing}(m)$ *and* $L^\ell \in \mathcal{L}_n(\ell, \kappa(\ell))$ *for all* $\ell \in \mathrm{div}(m)$ *such that* $\kappa(\ell) \geq 0$ *and*

$$L = L^{Sing} \cap \bigcap_{\ell \in \mathrm{div}(m):\kappa(\ell)\geq 0} L^\ell.$$

Theorem 1 follows from the following theorem for κ defined as $\kappa(m) = 0$ and $\kappa(\ell) = -\infty$ for all $\ell \in \mathrm{div}(m) \setminus \{m\}$.

Theorem 6 (Correlation bound for $\mathcal{L}_n(m, \kappa(\cdot))$ and MOD_q)**.** *Let* m, q *be co-prime and* $\kappa : \mathrm{div}(m) \to \mathbb{N} \cup \{-\infty\}$*. Let* $L \in \mathcal{L}_n(m, \kappa(\cdot))$ *be any linear system. Then*

$$\mathrm{Corr}(L, \mathrm{MOD}_q) \leq \exp(-n/c),$$

where $c = c_6(m, q, \kappa(\cdot))$*. Crucially,* c *does not depend on* n*.*

The proof of Theorem 6 follows from induction over $\kappa(\cdot)$. The following two Lemmas specify the base case and the inductive step.

Lemma 7 (Base case)**.** *Let* m, q *be co-prime and let* $k \geq 0$ *be a sparsity. Let* $L = L' \cap L''$ *where* $L' \in \mathcal{L}_n^{Sing}(m)$ *and* $L'' \in \mathcal{L}_n(1, k)$*. Then*

$$\mathrm{Corr}(L, \mathrm{MOD}_q) \leq \exp(-n/\beta^k),$$

where $\beta = \beta(m, k)$ *is as given in Theorem 5.*

The Lemma above is implicit in the work of Chattopadhyay and Wigderson and points out why we call such linear systems simple. It is obvious that if we could decompose a given linear system into unions of a few such simple systems, we would obtain our desired correlation bounds by the union bound. The next lemma, the main inductive step, roughly shows that the only obstacle from having such a nice decomposition is the existence of subsystems that are satisfied by an exponentially small fraction of the points of the cube.

Lemma 8 (Simplification process for $\mathcal{L}_n(m, k)$)**.** *For any* $m, k \geq 0$*, there exists* $c = c_8(m, k)$ *such that for any* $L \in \mathcal{L}_n(m, k)$ *and any* $1 \leq r \leq n$*, one of the following must hold:*

1) $\mathrm{Pr}_{x\in\{0,1\}^n}[x \in L] \leq \exp(-r/c)$.
2) *There exist* $L_1, \ldots, L_R \in \mathcal{L}_n(m, \kappa(\cdot))$ *such that* $L = L_1 \cup \ldots \cup L_R$ *for* $R \leq \exp(cr)$*, and* $\kappa : \mathrm{div}(m) \to \mathbb{N} \cup \{-\infty\}$ *is given as*
   a) *If* $k > 0$ *then* $\kappa(m) = k - 1$ *and* $\kappa(\ell) = k + m \log m$ *for all* $\ell \in \mathrm{div}(m) \setminus \{m\}$.
   b) *If* $k = 0$ *then* $\kappa(m) = -\infty$ *and* $\kappa(\ell) = k + m \log m$ *for all* $\ell \in \mathrm{div}(m) \setminus \{m\}$.

We first prove Theorem 6 given Lemmas 7 and 8. We then proceed to prove Lemmas 7 and 8.

Proof of Theorem 6 given Lemmas 7 and 8: Define a lexicographic order on $\kappa : \mathrm{div}(m) \to \mathbb{N} \cup \{-\infty\}$: $\kappa > \kappa'$ if there exists $\ell_0 \in \mathrm{div}(m)$ such that $\kappa(\ell) = \kappa'(\ell)$ for all $\ell > \ell_0$ and $\kappa(\ell_0) > \kappa'(\ell_0)$ [1].

The base case of $\kappa(\ell) = -\infty$ for all $\ell > 1$ is given by Lemma 7. For the inductive step, let $\ell_{\max} > 1$ be maximal such that $\kappa(\ell_{\max}) \geq 0$. Let $L \in \mathcal{L}_n(m, \kappa(\cdot))$. Then

$$L = L^{Sing} \cap \bigcap_{\ell \in \mathrm{div}(m):\kappa(\ell)\geq 0} L^\ell,$$

where $L^{Sing} \in \mathcal{L}_n^{Sing}(m)$ and $L^\ell \in \mathcal{L}_n(\ell, \kappa(\ell))$. Apply Lemma 8 for $L^{\ell_{\max}}$. Let $c = c_8(\ell_{\max}, \kappa(\ell_{\max}))$ and let $r = n/c^*$ be a parameter to be determined later. One of the following must hold:

1) $\mathrm{Pr}_{x\in\{0,1\}^n}[x \in L^{\ell_{\max}}] \leq \exp(-r/c) = \exp(-n/(cc^*))$. Hence, $\mathrm{Pr}_{x\in\{0,1\}^n}[x \in L] \leq \exp(-n/(cc^*))$ and $\mathrm{Corr}(L, \mathrm{MOD}_q) \leq \exp(-n/(cc^*))$.
2) There exists $L_1^{\ell_{\max}}, \ldots, L_R^{\ell_{\max}} \in \mathcal{L}_n(\ell_{\max}, \kappa_1(\cdot)) \subset \mathcal{L}_n(m, \kappa_1(\cdot))$ such that $L^{\ell_{\max}} = L_1^{\ell_{\max}} \cup \ldots \cup L_R^{\ell_{\max}}$ for $R \leq \exp(cr) = \exp(n(c/c^*))$, where $\kappa_1 < \kappa$ is given by $\kappa_1(\ell_{\max}) < \kappa(\ell_{\max})$ and $\kappa_1(\ell) = \kappa(\ell_{\max}) + m^2$ for $\ell < \ell_{\max}$. Define

$$L_i := L^{Sing} \cap \bigcap_{\ell \in \mathrm{div}(m):\ell<\ell_{\max},\kappa(\ell)\geq 0} L^\ell \cap L_i^{\ell_{\max}}.$$

We have $L = L_1 \cup \ldots \cup L_R$ where $L_i \in \mathcal{L}_n(m, \kappa'(\cdot))$, with $\kappa'(\cdot)$ defined as $\kappa'(\ell_{\max}) = \kappa_1(\ell_{\max})$ and $\kappa'(\ell) = \max(\kappa(\ell), \kappa_1(\ell))$ for $\ell < \ell_{\max}$. Note that $\kappa' < \kappa$, so we can apply the induction hypothesis for $L_1, \ldots, L_R$:

$$\begin{aligned}
&\mathrm{Corr}(L, \mathrm{MOD}_q) = \\
&\max_{b\in\mathbb{Z}_q\setminus\{0\}} \left|\mathbb{E}_{x\in\{0,1\}^n}[1_L(x) \cdot \omega_q(b(x_1 + \ldots + x_n))]\right| \\
&= \max_{b\in\mathbb{Z}_q\setminus\{0\}} \left|\sum_{i=1}^{R} \mathbb{E}_{x\in\{0,1\}^n}[1_{L_i}(x) \cdot \omega_q(b(x_1 + \ldots + x_n))]\right| \\
&\leq \sum_{i=1}^{R} \max_{b\in\mathbb{Z}_q\setminus\{0\}} \left|\mathbb{E}_{x\in\{0,1\}^n}[1_{L_i}(x) \cdot \omega_q(b(x_1 + \ldots + x_n))]\right| \\
&\leq |R| \cdot \mathrm{Corr}(\mathcal{L}_n(m, \kappa'(\cdot)), \mathrm{MOD}_q) \\
&\leq \exp((c/c^* - 1/c_6(m, \kappa'(\cdot)))n),
\end{aligned}$$

where crucially we used the fact that $L_1, \ldots, L_R$ are disjoint.

Setting c^* to be a large enough constant (say $c^* = 2c \cdot c_6(m, \kappa'(\cdot))$) concludes the proof. ∎

[1] As $\mathbb{N}^d$ is a well founded set for all $d \geq 1$ this defines a proper Noetherian induction. An explicit bound can be derived using the explicit bounds on the growth of $\kappa(\cdot)$ given by Lemma 8.

A. Proof of base case: Lemma 7

Our argument here essentially is taken from [12]. Let $L = L' \cap L''$ with $L' \in \mathcal{L}_n^{Sing}(m)$ and $L'' \in \mathcal{L}_n(1,k)$. That is,

$$L' = \{x \in \{0,1\}^n : \sum a_{i,j}x_j \equiv b_i \pmod m \quad \forall i = 1,\ldots,N'\}$$

$$L'' = \{x \in \{0,1\}^n : 0 \in A_i(x_{v(i,1)},\ldots,x_{v(i,k)}) \quad \forall i = 1,\ldots,N''\},$$

where $a_{i,j}, b_i \in \mathbb{Z}_m$, $v(i,j) \in [n]$ and $A_i(z_1,\ldots,z_k) \subset \mathbb{Z}_m$.

Define $P_i(x) := \sum a_{i,j}x_j - b_i$ to be linear functions over $\mathbb{Z}_m$ for $i \in [N']$ so that

$$1_{L'}(x) = \prod_{i=1}^{N'} 1_{P_i(x)=0}.$$

Define $Q_i(x)$ to be polynomials over $\mathbb{Z}_m$ of degree at most k such that $Q_i(x) = 0$ iff $0 \in A_i(x_{v(i,1)},\ldots,x_{v(i,k)})$, so that

$$1_{L''}(x) = \prod_{i=1}^{N''} 1_{Q_i(x)=0}.$$

Using the fact that for $z \in \mathbb{Z}_m$ we have $1_{z=0} = \frac{1}{m}\sum_{a=0}^{m-1} \mathrm{e}_m(a \cdot z)$ we get

$$1_L(x) = \left(\prod_{i=1}^{N'} \frac{1}{m}\sum_{a=0}^{m-1} \mathrm{e}_m\big(a \cdot P_i(x)\big)\right) \times \left(\prod_{j=1}^{N''} \frac{1}{m}\sum_{b=0}^{m-1} \mathrm{e}_m\big(b \cdot Q_j(x)\big)\right)$$
$$= \frac{1}{m^{N'+N''}} \sum \mathrm{e}_m\left(\sum_{i=1}^{N'} a_i \cdot P_i(x) + \sum_{j=1}^{N''} b_j \cdot Q_j(x)\right).$$

where the last summation is over $a_1,\ldots,a_{N'},b_1,\ldots,b_{N''} \in \mathbb{Z}_m$. The bound for the correlation between L and MOD_q now follows from Theorem 5 since all $P_i(x), Q_j(x)$ are polynomials of degree at most k, and so are all linear combinations of them.

B. Proof of inductive step: Lemma 8

We define a notion of rank of $L \in \mathcal{L}_n(m,k)$ which is appropriate for our purposes and whose origins lie in the elegant work of Grigoriev and Razborov [15]. Chattopadhyay and Wigderson [12] generalized the Grigoriev-Razborov notion to deal with linear systems of type $\mathcal{L}_n(m)$. We further generalize it to deal with systems in $\mathcal{L}_n(m,k)$, where k is a constant non-negative integer. For the sake of consistency with earlier work, we call this notion the *communication rank* of the linear system. Fix some equations $E_1,\ldots,E_N$ of sparsity k such that $L = E_1 \cap \ldots \cap E_N$, where each E_i is given by

$$E_i = \{x \in \{0,1\}^n : \sum a_{i,j}x_j \pmod m \in A_i(x_{v(i,1)},\ldots,x_{v(i,k)})\},$$

The definition of communication rank will in fact depend on the specific $E_1,\ldots,E_N$ chosen.

Definition 5 (Communication Rank)**.** *Let $L \in \mathcal{L}_n(m,k)$ given by $L = E_1 \cap \ldots \cap E_N$. We say that a subset of equations $I = \{i_1,\ldots,i_r\} \subset [N]$ is s-*wise independent *if the following conditions hold. Let $V_i = \{v(i,1),\ldots,v(i,k)\}$ for $i \in I$ be the set of variables on which A_i depends. We first require that all sets $V_{i_1},\ldots,V_{i_r}$ be disjoint. We also require that there exist subsets of variables $J_{p,t} \subset [n]$ of size $|J_{p,t}| = |I| = r$, where p ranges over the distinct prime factors of m and $t = 1,\ldots,s$, such that:*

1) *All sets $J_{p,t}$ and $V_{i_1},\ldots,V_{i_r}$ are pair-wise disjoint.*
2) *Let $M_{p,t}$ be the following $r \times r$ matrix over $\mathbb{F}_p$: if $J_{p,t} = \{j_1,\ldots,j_r\}$ then the (x,y)-entry of $M_{p,t}$ is given by a_{i_x,j_y} (modulo p), i.e. $M_{p,t}$ is the $r \times r$ minor given by the rows of I and the columns of $J_{p,t}$. We require that for any prime factor p of m, and any $t = 1,\ldots,s$, the matrix $M_{p,t}$ has full rank modulo p.*

The s-wise communication rank of L modulo m, denoted by $\mathrm{ccrank}_s^m(L)$, is the maximal r for which this holds for some $I \subset [N]$ of size $|I| = r$.

If $k = 0$, then the above definition exactly corresponds to the notion of communication rank used by Chattopadhyay and Wigderson. We use the following result that follows from their work:

Lemma 9 (Implicit in Chattopadhyay-Wigderson [12])**.** *Let $L \in \mathcal{L}_n(m)$ have ccrank_m^m communication rank at least r. Then*

$$\Pr_{x \in \{0,1\}^n}\left[\bigwedge_{i=1}^{N} \ell_i(x) \in A_i\right] \leq \exp\big(-r/c_9(m)\big),$$

where each $A_i \subsetneq \mathbb{Z}_m$ is an arbitrary set.

Remark 1. *This lemma appears in [12] with the restriction that m has no repeated prime factors. We show in the appendix that this restriction can be lifted by a slight modification of the argument in [12]. In fact, we generalize it to all Abelian groups.*

We next show an easy corollary of the above lemma for systems in $\mathcal{L}_n(m,k)$, when $k > 0$.

Lemma 10. *Let $L \in \mathcal{L}_n(m,k)$ have m-wise communication rank at least r. Then*

$$\Pr_{x\in\{0,1\}^n}[x \in L] \leq \exp(-r/c_{10}(m,k)).$$

Proof: Let $I = \{i_1,\ldots,i_r\}$ be a set of indices corresponding to independent equations. We focus entirely on the sub-system indexed by this set. Since sets $V_{i_1},\ldots,V_{i_r}$ are disjoint, we can sample all $x_j \in \cup_{i\in I}V_i$ and guarantee, by the Chernoff bound, that with probability at least $1-exp(-r/2^k)$ we will get $\Omega(r/2^k)$ non-trivial accepting sets left. Thus, after sampling, we are left with an ordinary sub-system in $\mathcal{L}_{n'}(m)$ whose rank is $\Omega(r/2^k)$, where $n' \geq n - rk$. Applying Lemma 9 to this sub-system, the argument follows by setting $c_9(m,k) = \theta(2^k c_8(m))$. ∎

Lemma 10 shows that if the linear system has high communication rank, then its correlation with MOD_q is small as the size of the solution set is a very small fraction of the boolean cube. We next deal with the complementary case, where the communication rank is small.

We start off by a convenient structural result about such systems, stated in Chattopadhyay and Wigderson that generalizes a lemma of Grigoriev and Razborov.

Lemma 11 (Restatement of Lemma 13 of [12])**.** *Let m have w distinct prime factors $p_1,\ldots,p_w$. Consider an ordinary linear system $L \in \mathcal{L}_n(m)$ such that $ccrank^m_s(L) = r$. Then, there exists a set I of at r linear forms satisfying the following condition: for every linear form ℓ in L, there exists a prime p_j such that $\ell \equiv \ell_I + \ell_0$ (mod p_j), where ℓ_I is in the $\mathbb{Z}_{p_j}$-linear span of I and ℓ_0 is ws-sparse.*

We will also need the following simple claim.

Claim 12. *Let $\ell(x_1,\ldots,x_n) \in A$ (mod m) be any linear constraint. Let p be a prime factor of m, and assume that $\ell \equiv \ell_I + \ell_0$ (mod p) where ℓ_0 is supported on variables $x_{i_1},\ldots,x_{i_k}$. For $a \in \mathbb{Z}_m$ let $B_a = \{x \in \{0,1\}^n : \ell_I(x) \equiv a$ (mod m)$\}$. Then for every $a \in \mathbb{Z}_m$ and values for $x_{i_1},\ldots,x_{i_k} \in \{0,1\}$, there exists a linear form $\ell'(x)$ and a set $A_a(x_{i_1},\ldots,x_{i_k}) \subset \mathbb{Z}_{m/p}$ such that*

$$B_a \cap \{x \in \{0,1\}^n : \ell(x) \text{ (mod } m) \in A\} =$$

$$B_a \cap \{x \in \{0,1\}^n : \ell'(x) \text{ (mod } m/p) \in A_a(x_{i_1},\ldots,x_{i_k})\}.$$

Proof: Let $S = \{x_1,\ldots,x_k\}$ be the variables in the support of ℓ_0, and $T = [n] \setminus S$ be the remaining variables. For $x \in \{0,1\}^n$ let $x^S \in \{0,1\}^S$ and $x^T \in \{0,1\}^T$ denote its restriction to the corresponding variables sets. Note that $\ell_0(x) = \ell_0(x^S)$. Partition $\ell(x) = \ell^S(x^S) + \ell^T(x^T)$ to a linear form over x^S and a linear form over x^T, and similarly $\ell_I(x) = \ell^S_I(x^S) + \ell^T_I(x^T)$. Note that by assumption $\ell^T \equiv \ell^T_I \pmod p$, hence all the coefficients of $\ell^T - \ell^T_I$ are divisible by p. We now define $\ell'(x) = \ell'(x^T) = (\ell^T(x^T) - \ell^T_I(x^T))/p \pmod{m/p}$. Note that $\ell(x) \equiv p\ell'(x^T) + \ell_I(x) + \ell^S(x^S) - \ell^S_I(x^S) \pmod m$. Consider any assignment for $x_{i_1},\ldots,x_{i_k}$, and let $b := a + \ell^S(x^S) - \ell^S_I(x^S) \pmod m$. We set $A_a(x_{i_1},\ldots,x_{i_k}) = \{(z-b)/p \pmod{m/p} : z \in A,\ z \equiv b \pmod p\}$. ∎

We now state the simplification lemma for $\mathcal{L}_n(m,k)$ systems of low communication rank.

Lemma 13. *Let $L \in \mathcal{L}_n(m,k)$ have m-wise communication rank at most r. Then there exist $L_1,\ldots,L_R \in L_n(m,\kappa(\cdot))$ such that $L = L_1 \cup \ldots \cup L_R$ where $R \leq \exp((k + \log m + m\log m)r)$ and $\kappa : \mathrm{div}(m) \to \mathbb{N} \cup \{-\infty\}$ is defined as*

1) *If $k > 0$ then $\kappa(m) = k-1$ and $\kappa(\ell) = k + m\log m$ for all $\ell \in \mathrm{div}(m) \setminus \{m\}$.*
2) *If $k = 0$ then $\kappa(m) = -\infty$ and $\kappa(\ell) = k + m\log m$ for all $\ell \in \mathrm{div}(m) \setminus \{m\}$.*

Proof: Let $L = \bigwedge_{i=1}^N \{x \in \{0,1\}^n : \ell_i(x) \pmod m \in A_i(x_{v(i,1),\ldots,x_{v(i,k)}})\}$. Denote the number of distinct prime factors of m by $w \leq \log m$. Assume that L has m-wise communication rank of at most r. Let I be the set of r equations given by Lemma 11. Consider the set of variables $W \subset [n]$ given by

$$W = \bigcup_{i\in I} V_i \cup \bigcup_{p,t} J_{p,t}.$$

We have $|W| \leq r(k + mw)$. We will consider all possible assignments to variables in W and all possible values for equations in I. For $\alpha \in \{0,1\}^W$ and $\beta \in \mathbb{Z}_m^I$ define

$$B_{\alpha,\beta} := \{x \in \{0,1\}^n : \forall w \in W, x_w = \alpha_w \text{ and } \forall i \in I, \ell_i(x) \equiv \beta_i \pmod m\}.$$

Note that $\{0,1\}^n = \cup_{\alpha\in\{0,1\}^W,\beta\in\mathbb{Z}_m^I} B_{\alpha,\beta}$. We will show that for any setting of α,β we have

$$B_{\alpha,\beta} \cap L \in \mathcal{L}_n(m,\kappa(\cdot)).$$

Hence, $L = L_1 \cup \ldots \cup L_R$ for $R = 2^{|W|}m^{|I|} \leq \exp((k + \log mm\log m)r)$.

We partition the set of rows outside of I into two parts. I' is the set of all rows i' such that $V_{i'}$ intersects

W. Let I'' be the set of all other rows not in I'. Note that if $k = 0$ then I' is empty.

Consider first rows $i' \in I'$. Note that after fixing values for elements in W the sparsity of $A_{i'}$ reduces by at least one, hence they are equivalent to linear forms in $\mathcal{L}_n(m, k-1)$. Consider next a row $i'' \in I''$. By Lemma 11 there exists a prime factor p of m with $\ell_{i''} \equiv \ell_I + \ell_0 \pmod{p}$, where ℓ_I is in the $\mathbb{Z}_p$-span of rows in I and ℓ_0 is wm-sparse and supported on variables in $x_{u(i'',1)}, \ldots, x_{u(i'',wm)}$. We will add these wm variables to the accepting set of row i''. Thus, under every fixing of linear forms in I over $\mathbb{Z}_m$, by Claim 12 every linear constraint in I'' simplifies to one over a modulus m/p for some factor p of m with an accepting set of sparsity at most $k + mw$. Combining these arguments, for any row $i \in [N]$ and any assignment $\alpha \in \{0,1\}^W, \beta \in \mathbb{Z}_m^I$ we have

$$L \cap B_{\alpha,\beta} = L_{\text{Sing}}^{(\alpha,\beta)} \cap L_m^{(\alpha,\beta)} \cap \bigcap_{p \text{ prime factor of } m} L_{m/p}^{(\alpha,\beta)},$$

where $L_{\text{Sing}}^{(\alpha,\beta)} \in \mathcal{L}_m^{Sing}$ is the constraints $x_w = \alpha_w$ for $w \in W$ (which can equivalently by stated modulo m as $x_w \in \{0,1\}$) and $\ell_i(x) \equiv \beta_i \pmod{m}$ for $i \in I$; $L_m^{(\alpha,\beta)} \in \mathcal{L}_n(m, k-1)$ is given by simplified equations for $i' \in I'$; and $L_{m/p}(\alpha,\beta) \in \mathcal{L}_n(m/p, k+mw)$ is given by equations for $i'' \in I''$. Note that if $k = 0$ then L_m does not appear. Thus $L \cap B_{\alpha,\beta} \in \mathcal{L}_n(m, \kappa(\cdot))$ and the lemma follows.

■

Lemma 8 follows immediately by combining Lemma 10 and Lemma 13, concluding this section.

REFERENCES

[1] A. A. Razborov, "Lower bounds on the size of bounded depth circuits over a complete basis with logical addition," *Mathematical Notes*, vol. 41, pp. 333–338, 1987, 10.1007/BF01137685. [Online]. Available: http://dx.doi.org/10.1007/BF01137685

[2] R. Smolensky, "Algebraic methods in the theory of lower bounds for boolean circuit complexity," in *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, ser. STOC '87. New York, NY, USA: ACM, 1987, pp. 77–82. [Online]. Available: http://doi.acm.org/10.1145/28395.28404

[3] R. Williams, "Non-uniform ACC circuit lower bounds," 2010, preprint.

[4] K. A. Hansen and M. Koucky, "A new characterization of ACC^0 and probabilistic CC^0," *Computational Complexity, Annual IEEE Conference on*, vol. 0, pp. 27–34, 2009.

[5] R. Beigel and A. Maciel, "Upper and lower bounds for some depth-3 circuit classes," in *Proceedings of the 12th Annual IEEE Conference on Computational Complexity*. Washington, DC, USA: IEEE Computer Society, 1997, pp. 149–. [Online]. Available: http://portal.acm.org/citation.cfm?id=791230.792290

[6] H. Caussinus, "A note on a theorem of barrington, straubing and thrien," *Information Processing Letters*, vol. 58, no. 1, pp. 31 – 33, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/B6V0F-3VS5HKD-H/2/8c89e4ab169f92ff81918fda0f6816be

[7] D. Barrington, R. Beigel, and S. Rudich, "Representing boolean functions as polynomials modulo composite numbers," *Computational Complexity*, vol. 4, pp. 367–382, 1994, 10.1007/BF01263424. [Online]. Available: http://dx.doi.org/10.1007/BF01263424

[8] K. Hansen, "On modular counting with polynomials," in *Computational Complexity, 2006. CCC 2006. Twenty-First Annual IEEE Conference on*, 0 2006.

[9] P. Gopalan, "Constructing ramsey graphs from boolean function representations," in *Computational Complexity, 2006. CCC 2006. Twenty-First Annual IEEE Conference on*, 0 2006.

[10] K. Efremenko, "3-query locally decodable codes of subexponential length," in *Proceedings of the 41st annual ACM symposium on Theory of computing*, ser. STOC '09. New York, NY, USA: ACM, 2009, pp. 39–44. [Online]. Available: http://doi.acm.org/10.1145/1536414.1536422

[11] Z. Dvir, P. Gopalan, and S. Yekhanin, "Matching vector codes," 2010, manuscript. [Online]. Available: ./DvirGopalanYekhanin10.pdf

[12] A. Chattopadhyay and A. Wigderson, "Linear systems over composite moduli," in *Foundations of Computer Science, 2009. FOCS '09. 50th Annual IEEE Symposium on*, 2009, pp. 43 –52.

[13] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy, and G. Turán, "Threshold circuits of bounded depth," *J.Computer.System.Sciences*, vol. 46, no. 2, pp. 129–154, 1993.

[14] D. Barrington and D. Therien, "Finite monoids and the fine structure of nc1," in *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, ser. STOC '87. New York, NY, USA: ACM, 1987, pp. 101–109. [Online]. Available: http://doi.acm.org/10.1145/28395.28407

[15] D. Grigoriev and A. Razborov, "Exponential complexity lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields," in *Foundations of Computer Science, 1998. Proceedings.39th Annual Symposium on*, Nov. 1998, pp. 269 –278.

[16] J. Bourgain, "Estimation of certain exponential sums arising in complexity theory," *Comptes Rendus Mathematique*, vol. 340, no. 9, pp. 627 – 631, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/B6X1B-4G1WY89-1/2/5b1282e1368e39a773d356ede8e7cf55

[17] A. Chattopadhyay, N. Goyal, P. Pudlák, and D. Thérien, "Lower bounds for circuits with MOD_m gates," in *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2006, pp. 709–718.

[18] F. Green, A. Roy, and H. Straubing, "Bounds on an exponential sum arising in boolean circuit complexity," *Comptes Rendus Mathematique*, vol. 341, no. 5, pp. 279 – 282, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/B6X1B-4GX64X8-2/2/13530bedc0482c4c7d6834c8bf7627f1

[19] E. Viola and A. Wigderson, "Norms, XOR lemmas, and lower bounds for $GF(2)$ polynomials and multiparty protocols," *Theory of Computing*, pp. 137–168, 2008.

[20] A. Chattopadhyay, "Discrepancy and the power of bottom fan-in in depth-three circuits," in *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 449–458. [Online]. Available: http://portal.acm.org/citation.cfm?id=1333875.1334217