On the Correlation Between Parity and Modular Polynomials

Anna Gál[*] and Vladimir Trifonov

Dept. of Computer Science, University of Texas at Austin,
Austin, TX 78712-1188, USA
{panni, vladot}@cs.utexas.edu

Abstract. We consider the problem of bounding the correlation between parity and modular polynomials over $\mathbb{Z}_q$, for arbitrary odd integer $q \geq 3$. We prove exponentially small upper bounds for classes of polynomials with certain linear algebraic properties. As a corollary, we obtain exponential lower bounds on the size necessary to compute parity by depth-3 circuits of certain form. Our technique is based on a new representation of the correlation using exponential sums.

Our results include Goldmann's result [Go] on the correlation between parity and degree one polynomials as a special case. Our general expression for representing correlation can be used to derive the bounds of Cai, Green, and Thierauf [CGT] for symmetric polynomials, using ideas of the [CGT] proof. The classes of polynomials for which we obtain exponentially small upper bounds include polynomials of large degree and with a large number of terms, that previous techniques did not apply to.

1 Introduction

In this paper, we study the correlation between the MOD_2 function and Boolean functions computed by depth-2 circuits with a MOD_q gate at the top (for odd q), and AND gates at the input level (called $MOD_q \circ AND$ circuits). The Boolean function $MOD_m : \{0,1\}^n \to \{0,1\}$ is defined to be 0 when the sum of the input bits is divisible by m, and 1 otherwise. For every $MOD_m \circ AND$ circuit there is a multilinear polynomial P over $\mathbb{Z}_m$ such that on inputs $\mathbf{x} \in \{0,1\}^n$, the output of the circuit is 0 if and only if $P(\mathbf{x})$ is 0 modulo m. There is a straightforward way to associate such a polynomial with each circuit, using the inputs associated with the AND gates to form the monomials. This polynomial is called the *defining polynomial* of the circuit, and its degree is the largest fan-in of the AND gates in the circuit. Thus, depth-2 circuits of the above form correspond to polynomials over the ring $\mathbb{Z}_m$.

The *correlation* $C(f_1, f_2)$ between two Boolean functions $f_1, f_2 : \{0,1\}^n \to \{0,1\}$ is defined as $C(f_1, f_2) = \frac{1}{2^n} \sum_{\mathbf{x} \in \{0,1\}^n} (-1)^{f_1(\mathbf{x})} (-1)^{f_2(\mathbf{x})}$. Our interest in this question is motivated by its relevance to circuit complexity lower bounds. In addition, we believe that the question is also interesting on its own right.

[*] Supported in part by NSF Grant CCF-0430695 and an Alfred P. Sloan Research Fellowship.

R. Královic and P. Urzyczyn (Eds.): MFCS 2006, LNCS 4162, pp. 387–398, 2006.

1.1 Bounded Depth Circuits

Proving lower bounds on the size of Boolean circuits for specific functions is one of the central problems in complexity theory. It is also considered to be notoriously difficult, since for example, superpolynomial lower bounds on the size of Boolean circuits computing a function from the complexity class NP would imply that P $\neq$ NP. However, even much weaker (e.g. superlinear) lower bounds seem to remain out of reach of the current techniques. Imposing various restrictions on the circuits and developing lower bound methods for restricted circuit models has received a lot of attention in the last few decades. The hope is to extend such techniques, and develop new methods that are applicable towards stronger and stronger models.

One of the circuit models that has been extensively studied is bounded depth circuits. The results of [Aj, FSS, Ha, Yao85] show that the parity function cannot be computed by AC^0 circuits (constant depth polynomial size circuits with AND, OR, NOT gates). Barrington [Ba] defined the class $\mathrm{ACC}^0 = \cup_q \mathrm{ACC}^0(q)$, where $\mathrm{ACC}^0(q)$ denotes the class of constant depth, polynomial size circuits with AND,OR, NOT and MOD_q gates. Smolensky [Sm] proved that $MOD_r \notin \mathrm{ACC}^0(p^k)$ when p and r are distinct primes. The power of $\mathrm{ACC}^0(q)$ circuits when q is not a prime power is much less understood. For example, it is not known if all of NP can be computed by depth-3 $\mathrm{ACC}^0(6)$ circuits.

Depth-3 circuits can be surprisingly powerful. Allender [Al] proved that AC^0 is contained in the class of depth-3 circuits of quasipolynomial ($2^{(\log n)^{O(1)}}$) size with a $MAJORITY$ gate at the top, MOD_2 gates in the middle, and AND gates of $(\log n)^{O(1)}$ fan-in at the input level. (Such circuits are referred to as $MAJ \circ MOD_2 \circ AND_{(\log n)^{O(1)}}$ circuits.) Yao [Yao90] proved that ACC^0 is contained in the class of depth-3 threshold circuits of quasipolynomial ($2^{(\log n)^{O(1)}}$) size with AND gates of $(\log n)^{O(1)}$ fan-in at the input level. But it remains open if ACC^0 is contained in the class of quasipolynomial ($2^{(\log n)^{O(1)}}$) size $MAJ \circ MOD_q \circ AND_{(\log n)^{O(1)}}$ circuits, for some fixed q. In other words, it is not known whether Allender's result [Al] can be extended to ACC^0. (Essentially this question was asked by Green in [Gr02].) A recent result of Bourgain [Bo05] implies that parity requires exponential size $MAJ \circ MOD_q \circ AND_{\epsilon \log n}$ circuits, for any odd q and ϵ depending on q. It is not clear how to extend Bourgain's result to larger fan-in AND gates. The results of Håstad and Goldmann [HG] imply that a function in ACC^0 (the generalized inner product function) requires exponential size $MAJ \circ MOD_2 \circ AND_{O(\log n)}$ circuits. The results of Razborov and Wigderson [RW] imply that some function in ACC^0 requires $n^{\Omega(\log n)}$ size $MAJ \circ MOD_2 \circ AND$ circuits. This result was recently extended to circuits with arbitrary AC^0 circuits in place of the AND gates by Hansen and Miltersen [HM]. [RW] and [HM] build on the results of [HG]. However, the method in [HG] applies for arbitrary symmetric gates in the middle layer. Thus, in view of Yao's result [Yao90], these results cannot be directly extended to obtain exponential lower bounds for computing an ACC^0-function by $MAJ \circ MOD_q \circ AND_{(\log n)^{O(1)}}$ circuits.

Other combinations of threshold, MOD and AND gates in depth-3 circuits and other definitions of MOD gates have been also considered, and in some of these models exponential lower bounds have been proved for functions in ACC^0 (see e.g. [BM, Gro, GT, KP]). The power of $MAJ \circ MOD \circ AND_{(\log n)^{O(1)}}$ circuits remains less understood.

1.2 Correlation and Circuit Lower Bounds

Obtaining exponential lower bounds for $MAJ \circ MOD_q \circ AND$ circuits under various restrictions has received considerable attention in the last few years (e.g. [AB, CGT, Gr99, Gr02, Go]. The starting point of all these papers, including [HG] which considers the more general $MAJ \circ SYM \circ AND$ circuits, is the following special case of a lemma of [HMPST].

Lemma 1. *(Lemma 3.3, [HMPST]) Let $f : \{0,1\}^n \rightarrow \{0,1\}$ be a Boolean function computed by a Boolean circuit with a fan-in m (unweighted) threshold gate on top, taking the results of the subcircuits $C_1, \ldots, C_m$ as inputs to the threshold gate. Let $g_i : \{0,1\}^n \rightarrow \{0,1\}$ be the Boolean function computed by the subcircuit C_i ($i = 1, \ldots, m$). Let f be a balanced function, i.e. $|f^{-1}(0)| = |f^{-1}(1)|$. Then for at least one of the subcircuits (for some $1 \leq i \leq m$), the absolute value of the correlation $|C(f, g_i)|$ is at least $1/m$.*

Thus, upper bounds on the absolute value of the correlation of a balanced function f with arbitrary functions that can be computed by circuits of a given class $\mathcal{C}$, imply lower bounds on the fan-in of the $MAJORITY$ gate in $MAJ \circ \mathcal{C}$ type circuits for computing f.

In particular, proving that the absolute value of the correlation of parity with modular polynomials over $\mathbb{Z}_q$ of certain type is exponentially small, implies exponential lower bounds on the size of the corresponding $MAJ \circ MOD_q \circ AND$ circuits. Smolensky's results [Sm] imply that for p and r distinct primes, the absolute value of the correlation of the MOD_r function and low degree polynomials over $\mathbb{Z}_{p^k}$ is at most $\frac{1}{n^{1/2-o(1)}}$. Note that the technique of [Sm] does not yield smaller bounds on the absolute value of the correlation even for degree 2 and very sparse polynomials, and it cannot be applied over $\mathbb{Z}_q$, if q is not a prime power. It is also curious to note that on the other hand, by Ajtai's [Aj] result we know that the absolute value of the correlation of parity with functions in AC^0 is exponentially small, and it remains exponentially small even allowing superpolynomial number of gates [Ha]. Cai, Green and Thierauf [CGT] proved that the absolute value of the correlation of parity with symmetric polynomials of degree $(\log n)^{O(1)}$ over $\mathbb{Z}_q$ for q odd, is exponentially small (at most $2^{-n^{\Omega(1)}}$). This was generalized by Green [Gr99] to proving similar exponentially small upper bounds on the absolute value of the correlation of the MOD_p function with symmetric polynomials of degree $(\log n)^{O(1)}$ over $\mathbb{Z}_q$ when p is a prime that does not divide q.

Extending these bounds to allowing non-symmetric polynomials posed a significant challenge. The degree 1 case was solved by Goldmann [Go], who proved that the absolute value of the correlation of MOD_p and MOD_q when p has a prime factor that does not divide q, is at most $2^{-\Omega(n)}$. Alon and Beigel [AB] showed that the absolute value of the correlation of parity with degree 2 polynomials over $\mathbb{Z}_q$ for odd q, is at most $2^{-(\log n)^{\epsilon}}$ for some constant $\epsilon < 1$, and for degree $O(1)$ polynomials the absolute value of the correlation is $o(1)$. Note that the bounds of [AB] are weaker than the $\frac{1}{n^{1/2-o(1)}}$ upper bounds implied by Smolensky's results [Sm], but [Sm] is applicable only when q is a prime power. The first improvement over the bounds of [Sm] and [AB] for non-symmetric polynomials of degree greater than 1 was achieved by Green [Gr02]. Green [Gr02] proved that the absolute value of the correlation of parity with degree 2 polynomials over $\mathbb{Z}_3$ is at most $2^{-\Omega(n)}$. The method used in [Gr02] very specifically relies on the degree being at most 2 and $q = 3$, and appears to be not applicable to other degrees or other values of q. A breakthrough was achieved by Bourgain [Bo05], proving that for q odd, and p, q relatively prime, the absolute value of the correlation between MOD_p and degree d polynomials over $\mathbb{Z}_q$ is exponentially small for $d < \epsilon \log n$, where ϵ depends on p and q. Bourgain's result was generalized by Green, Roy and Straubing [GRS] to arbitrary (not necessarily odd) q and p, q relatively prime.

While Bourgain's result resolves the question about the correlation between parity and modular polynomials of degree up to $\epsilon \log n$, it leaves open the question described in the previous section about whether Allender's result [Al] can be extended to ACC^0. To obtain sufficiently strong lower bounds for depth 3 circuits of the desired type by bounding correlation, we would need to be able to provide estimates on the correlation for up to polylogarithmic degree polynomials.

1.3 Our Approach

We suggest a new approach to estimate the correlation of parity with modular polynomials over $\mathbb{Z}_q$ that is applicable to arbitrary odd q, and provides improvements over the previous bounds for several classes of polynomials.

The starting point of our approach is a representation of the correlation using exponential sums. Exponential sums have been used to estimate correlation in several previous papers starting with the results of Cai, Green and Thierauf [CGT] for symmetric polynomials and also in [Gr99, Gr02, Bo05, GRS]. We give a representation of the correlation of parity with polynomials over $\mathbb{Z}_q$ using exponential sums in a very general setting. The novelty of our representation is that it allows to use certain linear algebraic properties of the terms of the corresponding polynomials. We also present a general expression for representing correlation, that can be used to yield our results as well as to derive the bounds of Cai, Green, and Thierauf [CGT] for symmetric polynomials, using ideas of the [CGT] proof. The two approaches can be viewed in a unifying framework as working with different components of our expression.

We are able to evaluate the exponential sums involved in this representation under various conditions, and we obtain exponentially small upper bounds on the absolute value of the correlation between parity and modular polynomials of certain type. Interestingly, the classes of polynomials for which we prove exponentially small bounds include polynomials of very large degree and polynomials with very large number of terms as well (as long as they satisfy some other, linear algebraic conditions). All previous methods assumed small degree to obtain exponentially small upper bounds on correlation with parity, thus could not be used to obtain our results. Moreover, some of our results yield exponentially small upper bounds on the absolute value of the correlation with parity over every nonempty subset of the variables.

Due to space limitations, all proofs are omitted from this extended abstract.

2 Exact Representations of the Correlation

2.1 Notation

For $r \in \mathbb{Z}^+$ and $z \in \mathbb{Z}$ define $\delta_r(z)$ to be 1, if $r|z$, and -1, otherwise. We will use $x \equiv_r y$ to denote $r|(x-y)$. $\equiv_r$ is extended to vectors by applying the congruence on every coordinate. That is, we use the notation $\mathbf{x} \equiv_r \mathbf{y}$ to indicate that $x_i \equiv_r y_i$ for every coordinate. The exponent function is extended to vectors in a component-wise manner, that is, $c^{\mathbf{x}}$ denotes $(c^{x_1}, \ldots, c^{x_n})$.

We will denote with $\mathbf{0}$ the all 0's vector, where the dimension of the vector will be understood by the context. Similarly $\mathbf{1}$ is the all 1's vector. We use $\mathbf{1}_n$ to denote the all 1's vector of length n, we omit indicating the length when it is clear from the context. Vectors will be assumed to be in a column form and $\mathbf{x}^T$ is the row vector corresponding to a column vector $\mathbf{x}$. Similarly M^T is the transpose of a matrix M. For two vectors $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}^T\mathbf{y}$ is the usual inner product of the two vectors, that is, $\mathbf{x}^T\mathbf{y} = \sum_i x_i y_i$. For a matrix M and a vector $\mathbf{x}$, $M\mathbf{x}$ is the product of M and $\mathbf{x}$. Unless indicated otherwise, all sums and products are over the integers $\mathbb{Z}$.

The following notation for sets will be used: $[r] = \{1, \ldots, r\}$, $[0, r] = \{0, \ldots, r\}$, and $[0, 1) = \{a \in \mathbb{R} : 0 \leq a < 1\}$.

Let $h : \mathbb{Z}^n \to \mathbb{Z}$ be an arbitrary integer valued function and let $\mathbf{g} \in \{0,1\}^n$. We use the following notation.

$$C(\mathbf{g}, h) := 2^{-n} \sum_{\mathbf{x} \in \{0,1\}^n} \delta_2(\mathbf{g}^T\mathbf{x})\delta_q(h(\mathbf{x}))$$

We wish to estimate how well $MOD_q \circ AND$ circuits approximate parity. Let $f : \{0,1\}^n \to \{0,1\}$ be the function computed by a $MOD_q \circ AND$ circuit, and let P_f be the defining polynomial of the circuit. Then $(-1)^{f(\mathbf{x})} = \delta_q(P_f(\mathbf{x}))$ for $\mathbf{x} \in \{0,1\}^n$, and with our notation, the correlation between parity and f is equal to $C(\mathbf{1}, P_f)$. In general, our methods apply to estimating the correlation for the parity over arbitrary subsets of the input bits. Thus, we are interested in estimating $C(\mathbf{g}, P)$ for a multilinear polynomial $P(x_1, \ldots, x_n)$ with integer coefficients and a vector $\mathbf{g} \in \{0,1\}^n$.

Notice that we do not identify $\{0,1\}$ with $\mathbb{Z}_2$, i.e. arithmetic with numbers from $\{0,1\}$ is done in $\mathbb{Z}$, unless indicated otherwise. For $M \in \{0,1\}^{m \times n}$, $\mathrm{rk}_2(M)$ denotes the rank of M over $\mathbb{Z}_2$.

2.2 Exponential Sums

Following [CGT, Gr99, Gr02], we use exponential sums to represent the correlation. We give a representation of the correlation of parity with modular polynomials by exponential sums for arbitrary degree and arbitrary odd $q \geq 3$. Moreover, our representation applies to parity taken over arbitrary subsets of the input variables.

Let $\omega_q = e^{2\pi\mathrm{i}/q} = \cos 2\pi/q + \mathrm{i}\sin 2\pi/q$, the principal q-th root of unity, and $\bar{\omega} = \omega^{-1}$, the complex conjugate of ω. We omit q from the subscript of ω when it is clear from the context. We have the following lemma, which gives an alternative expression for the correlation $C(\mathbf{g}, h)$ between integer valued functions and parity.

Lemma 2. *Let $h : \mathbb{Z}^n \to \mathbb{Z}$, $\mathbf{g} \in \{0,1\}^n$, and let $q \geq 3$ be an odd integer. Then*

$$C(\mathbf{g}, h) = -\nu + \frac{2^{-(n-1)}}{q} \sum_{t=0}^{q-1} \sum_{\mathbf{x} \in \{0,1\}^n} (-1)^{\mathbf{g}^T\mathbf{x}} \omega^{th(\mathbf{x})},$$

where ν is 0 if $\mathbf{g} \neq \mathbf{0}$, and 1 otherwise.

Definition 1. *For $t \in [0, q-1]$, $h : \mathbb{Z}^n \to \mathbb{Z}$, and $\mathbf{g} \in \{0,1\}^n$ define*

$$C_t(\mathbf{g}, h) = 2^{-n} \sum_{\mathbf{x} \in \{0,1\}^n} (-1)^{\mathbf{g}^T\mathbf{x}} \omega^{th(\mathbf{x})}.$$

Notice that, if $\mathbf{g} \neq \mathbf{0}$, by the triangle inequality applied to the expression in Lemma 2, there exists $t \in [0, q-1]$ such that $|C(\mathbf{g}, h)| \leq 2|C_t(\mathbf{g}, h)|$. Hence, if $\mathbf{g} \neq \mathbf{0}$ and we can obtain an exponentially small bound on $|C_t(\mathbf{g}, h)|$ for every $t \in [0, q-1]$, then we will have an exponentially small bound on $|C(\mathbf{g}, h)|$. We can show that the converse is also true in some sense: if $|C(\mathbf{g}, h+c)|$ is exponentially small for every $c \in [0, q-1]$, then $|C_t(\mathbf{g}, h)|$ is exponentially small as well for every $t \in [0, q-1]$.

2.3 Matrix Notation

Let $P(\mathbf{x})$ be a multilinear polynomial with integer coefficients. First we will construct a multilinear polynomial $Q(\mathbf{y})$ with integer coefficients and with the same degree as $P(\mathbf{x})$ such that, for $\mathbf{x} \in \{0,1\}^n$, $P(\mathbf{x}) \equiv_q Q((-1)^{\mathbf{x}})$. Recall that $(-1)^{\mathbf{x}}$ denotes $((-1)^{x_1}, \ldots, (-1)^{x_n})$.

For $q \geq 3$ odd, there exists a unique integer $\rho \in [q-1]$ such that $2\rho \equiv_q 1$. For $z \in \mathbb{Z}$, let $l(z) = \rho(1-z)$ and extend l to vectors in a component-wise manner, that is, $l(\mathbf{y}) = (\rho(1-y_1), \ldots, \rho(1-y_n))$. Notice that for $\mathbf{x} \in \{0,1\}^n$

$$\mathbf{x} \equiv_q l((-1)^{\mathbf{x}}). \tag{1}$$

Define $Q(\mathbf{y}) = P(l(\mathbf{y}))$. Since l (considered as a univariate polynomial) is linear with integer coefficients, Q is a multilinear polynomial with integer coefficients of the same degree as P. Also, by (1), for every $\mathbf{x} \in \{0,1\}^n$ we have $P(\mathbf{x}) \equiv_q Q((-1)^{\mathbf{x}})$. Thus

$$C(\mathbf{g}, P) = 2^{-n} \sum_{\mathbf{x} \in \{0,1\}^n} \delta_2(\mathbf{g}^T\mathbf{x})\delta_q(Q((-1)^{\mathbf{x}})). \tag{2}$$

Our next goal will be to express $Q((-1)^{\mathbf{x}})$ using a linear transformation. Since Q is multilinear with integer coefficients, we can write it as $Q(\mathbf{y}) = \sum_{I \subseteq [n]} c_I y_I$, where $c_I \in \mathbb{Z}$ and $y_I = \prod_{i \in I} y_i$. Let $M \in \{0,1\}^{m \times n}$ be the matrix whose rows are the incidence vectors of the subsets $I \subseteq [n]$, each repeated ($c_I \bmod q$) times. (The incidence vector of the empty set is the all zero row, and $y_\emptyset = 1$ for any $\mathbf{y}$.) Notice that the degree of Q (and therefore the degree of P) is at most d if and only if M has at most d 1's per row. For $\mathbf{x} \in \{0,1\}^n$ we have

$$Q((-1)^{\mathbf{x}}) \equiv_q \mathbf{1}^T (-1)^{M\mathbf{x}}. \tag{3}$$

We use the following notation:

$$C(\mathbf{g}, M) := 2^{-n} \sum_{\mathbf{x} \in \{0,1\}^n} \delta_2(\mathbf{g}^T \mathbf{x}) \delta_q(\mathbf{1}^T (-1)^{M\mathbf{x}}).$$

Then, using (2) and (3), we obtain the following.

Lemma 3. *Let P and Q be multilinear polynomials with integer coefficients such that $P(\mathbf{x}) \equiv_q Q((-1)^{\mathbf{x}})$ for $\mathbf{x} \in \{0,1\}^n$. Let M correspond to Q according to the above mapping, and let $\mathbf{g} \in \{0,1\}^n$. Then $C(\mathbf{g}, P) = C(\mathbf{g}, M)$*

Before we proceed, consider the following example. Let $Q(\mathbf{y}) = \prod_{i=1}^n y_i - 1$. Then M consists of a single row of all 1's, and $q-1$ copies of the all zero row. The corresponding correlation is $C(\mathbf{1}, M) = 1$, since $\delta_q(\mathbf{1}^T(-1)^{M\mathbf{x}}) = \delta_q(\prod_{i=1}^n (-1)^{x_i} + q - 1) = \delta_2(\mathbf{1}^T \mathbf{x})$ for every $\mathbf{x} \in \{0,1\}^n$ (and every $q \geq 3$).

Definition 2. *For $t \in [0, q-1]$, $M \in \{0,1\}^{m \times n}$, and $\mathbf{g} \in \{0,1\}^n$ define*

$$C_t(\mathbf{g}, M) = 2^{-n} \sum_{\mathbf{x} \in \{0,1\}^n} (-1)^{\mathbf{g}^T \mathbf{x}} \omega^{t \mathbf{1}^T (-1)^{M\mathbf{x}}}.$$

With this notation, using Lemma 2

$$C(\mathbf{g}, M) = -\nu + \frac{2}{q} \sum_{t=0}^{q-1} C_t(\mathbf{g}, M), \tag{4}$$

where ν is 0 if $\mathbf{g} \neq \mathbf{0}$, and 1 otherwise.

Remark 1. Our methods directly apply to estimating the correlation between parity and $MOD_q \circ MOD_2$ circuits. In this case, given a $MOD_q \circ MOD_2$ circuit, there is a multilinear polynomial Q with integer coefficients such that on inputs $\mathbf{x} \in \{0,1\}^n$ the output of the circuit is 0 if and only if $Q((-1)^{\mathbf{x}})$ is 0 modulo q. Let M be the matrix corresponding to the polynomial Q as above. Then the correlation between the output of the circuit and the parity of the subset of variables corresponding to the vector $\mathbf{g}$ is equal to $C(\mathbf{g}, M)$.

2.4 Main Lemma

It is immediate from (4) by the triangle inequality that for $\mathbf{g} \neq \mathbf{0}$, $|C(\mathbf{g}, M)|$ is exponentially small if $|C_t(\mathbf{g}, M)|$ is exponentially small for every $t \in [0, q-1]$.

Thus, we will be concerned with giving bounds on $|C_t(\mathbf{g},M)|$, and use them to bound $|C(\mathbf{g},M)|$ using (4).

Definition 3. *For $M \in \{0,1\}^{m\times n}$ and $\mathbf{g} \in \{0,1\}^n$, define*

$$I(M) = \left\{\mathbf{z} \in \{0,1\}^n : \exists\, \mathbf{y} \in \{0,1\}^m \text{ s.t. } M^T\mathbf{y} \equiv_2 \mathbf{z}\right\},$$
$$K(M,\mathbf{g}) = \left\{\mathbf{y} \in \{0,1\}^m : M^T\mathbf{y} \equiv_2 \mathbf{g}\right\}.$$

The following lemma is our main technical tool for obtaining bounds on the correlation based on linear algebraic properties of the polynomials.

Lemma 4. *Let $t \in [0,q-1]$, $M \in \{0,1\}^{m\times n}$, and $\mathbf{g} \in \{0,1\}^n$. Then*

$$C_t(\mathbf{g},M) = 2^{-m} \sum_{\mathbf{y}\in K(M,\mathbf{g})} (\omega^t - \bar{\omega}^t)^{|\mathbf{y}|}(\omega^t + \bar{\omega}^t)^{m-|\mathbf{y}|},$$

where $|\mathbf{y}|$ is the number of 1's in $\mathbf{y}$.

2.5 A More General Framework

Definition 4. *For $\mathbf{g} \in \{0,1\}^n$, $A \in \mathbb{N}^{m\times n}$, and $\mathbf{b} \in \mathbb{N}^m$, define*

$$\kappa(\mathbf{g},A,\mathbf{b}) = \sum_{\mathbf{x}\in\{0,1\}^n : A\mathbf{x}=\mathbf{b}} (-1)^{\mathbf{g}^T\mathbf{x}}.$$

For $\mathbf{z} = (z_1,\ldots,z_m)$ define the following polynomial over $z_1,\ldots,z_m$.

$$T(\mathbf{g},A,\mathbf{z}) = \sum_{\mathbf{b}\in I_A} \kappa(\mathbf{g},A,\mathbf{b}) z_1^{b_1}\cdot\ldots\cdot z_m^{b_m},$$

where $I_A = \{\mathbf{b} \in \mathbb{N}^m : 0 \le b_i \le \sum_{j=1}^n a_{ij},\ \text{for}\ i \in [m]\}$.

First note that this definition includes as a special case the definition of Krawtchouk polynomials [Sz]. To see this take $m = 1$ and A to be an all 1's row of length n. Then $\kappa(\mathbf{g}, \mathbf{1}_n^T, k) = K_k^{(n)}(|\mathbf{g}|)$, where $K_k^{(n)}(l)$ is the k-th Krawtchouk polynomial, i.e. $K_k^{(n)}(l) = \sum_{i=0}^k (-1)^i \binom{l}{i}\binom{n-l}{k-i}$ which is the coefficient of y^k of the polynomial $(1-y)^l(1+y)^{n-l}$.

In our definition $\kappa(\mathbf{g},A,\mathbf{b})$ is not necessarily a polynomial except in special cases, but it gives the coefficient of the monomial $z_1^{b_1}\cdot\ldots\cdot z_m^{b_m}$ of the polynomial $T(\mathbf{g},A,\mathbf{z})$, which can be written in the following form.

$$T(\mathbf{g},A,\mathbf{z}) = \prod_{j\in[n]:g_j=1}\left(1 - \prod_{i\in[m]} z_i^{a_{ij}}\right)\prod_{j\in[n]:g_j=0}\left(1 + \prod_{i\in[m]} z_i^{a_{ij}}\right). \qquad (5)$$

(This expression can be verified by expanding the right side of (5) and then grouping the terms in $z_1,\ldots,z_m$ of the same form together.) Thus, in some sense the functions $\kappa(\mathbf{g},A,\mathbf{b})$ are analogues of the Krawtchouk polynomials in a more general setting.

As we have seen before, for $\mathbf{g} \neq \mathbf{0}$, and arbitrary $h : \mathbb{Z}^n \to \mathbb{Z}$, the correlation $|C(\mathbf{g},h)|$ is exponentially small if $|C_t(\mathbf{g},h)|$ is exponentially small for every $t \in [0,q-1]$. We give the following general expression for $C_t(\mathbf{g},h)$.

Lemma 5. *Let $q, r \in \mathbb{N}^+$ and $t \in [0, q-1]$. Let $h : \mathbb{Z}^n \to \mathbb{Z}$ be such that there exists $A \in \mathbb{N}^{m\times n}$ and $G : \mathbb{N}^m \to \mathbb{Z}$ such that for every $\mathbf{x} \in \{0,1\}^n$ and $\mathbf{z} \in \mathbb{N}^m$, if $\mathbf{z} \equiv_r A\mathbf{x}$, then $h(\mathbf{x}) \equiv_q G(\mathbf{z})$. Then*

$$C_t(\mathbf{g}, h) = 2^{-n} \sum_{\mathbf{y}\in[0,r-1]^m} T(\mathbf{g}, A, \omega_r^{\mathbf{y}})\phi_{r,q,t}(\mathbf{y}, G)\ , \tag{6}$$

where $\phi_{r,q,t}(\mathbf{y}, G) = r^{-m} \sum_{\mathbf{z}\in[0,r-1]^m} \bar{\omega}_r^{\mathbf{y}^T\mathbf{z}} \omega_q^{tG(\mathbf{z})}$.

This lemma can be used to derive our main lemma (Lemma 4) that we use to exploit the linear algebraic properties of the polynomials when estimating correlation, as well as the bounds of Cai, Green and Thierauf [CGT] for symmetric polynomials. Interestingly, the statement yields these two arguments by working with different parts of the expression. To obtain our results in this paper we carefully estimate $\phi_{r,q,t}$, but we set things up so that for T we only have one possible nonzero value, and we just have to argue about when is T nonzero. To obtain the bounds of [CGT], we carefully estimate T, and use only a trivial bound on $\phi_{r,q,t}$, namely that $|\phi_{r,q,t}| \leq 1$. The key to derive the bounds of [CGT] from Lemma 5 is to show that for symmetric polynomials the matrix $A = \mathbf{1}_n^T$ with only one row and certain small odd r have the desired properties.

Note that all our expressions so far have been precise and we obtained exact representations of the correlation between parity and modular polynomials. Next we consider cases where we can obtain exponentially small upper bounds on the absolute value of our expressions.

3 Bounds Based on the Linear Algebraic Structure of the Polynomials

Lemma 4 allows us to obtain estimates on the correlation of the polynomial $P(\mathbf{x})$ with parity, based on the linear algebraic properties of the matrix $M \in \{0,1\}^{m\times n}$ considered as a matrix over $\mathbb{Z}_2$. Recall that to obtain M, first P is transformed to another polynomial Q, and the rows of M are defined by the terms of Q as described in Section 2.3. Also note that our methods can be used to estimate the correlation of modular polynomials and parity over arbitrary subsets of the variables. Parity is taken over the coordinates that are 1 in the vector $\mathbf{g}$, taking parity of all the variables corresponds to using $\mathbf{g} = \mathbf{1}$.

An immediate consequence of Lemma 4 is that $C_t(\mathbf{g}, M) = 0$, if $K(M, \mathbf{g}) = \emptyset$. Hence if $\mathbf{0} \neq \mathbf{g} \notin I(M)$, then $C(\mathbf{g}, M) = 0$. Thus we get the following interesting statement.

Theorem 1. *Let P and Q be multilinear polynomials with integer coefficients such that $P(\mathbf{x}) \equiv_q Q((-1)^{\mathbf{x}})$ for $\mathbf{x} \in \{0,1\}^n$. Let M be the matrix corresponding to Q, and let $\mathbf{g} \in \{0,1\}^n$. If the rows of the matrix M do not span the vector $\mathbf{g}$ over $\mathbb{Z}_2$, that is when $\mathbf{g} \notin I(M)$, then the correlation $C(\mathbf{g}, P) = 0$.*

This theorem extends the well known fact that if a polynomial P does not depend on all the variables over which we take parity, then the correlation between parity and P is zero.

Estimating $C_t(\mathbf{g}, M)$ when $\mathbf{g} \in I(M)$ is a challenging task in general. We prove that $|C_t(\mathbf{g}, M)|$ is exponentially small for certain classes of matrices M. Then, (4) can be used to obtain upper bounds on $|C(\mathbf{g}, M)|$. Note that while our estimates of $|C_t(\mathbf{g}, M)|$ apply to arbitrary $\mathbf{g} \in \{0,1\}^n$, and give good bounds even for $\mathbf{g} = \mathbf{0}$, the bounds on $|C(\mathbf{g}, M)|$ are interesting only for $\mathbf{g} \neq \mathbf{0}$, since in (4) $\nu = 1$ for $\mathbf{g} = \mathbf{0}$. Notice that if $\mathbf{g} \neq \mathbf{0}$, then $C_0(\mathbf{g}, M) = 0$, so it is enough to estimate $|C_t(\mathbf{g}, M)|$ for $t \in [q-1]$.

First we consider the class of non-singular matrices over $\mathbb{Z}_2$.

Theorem 2. *Let $M \in \{0,1\}^{n \times n}$ be a non-singular matrix over $\mathbb{Z}_2$, and $\mathbf{g} \in \{0,1\}^n$. Let $q \geq 3$ be an odd integer, and let $t \in [q-1]$. There exists $\gamma = \gamma(q) \in [0,1)$ (depending only on q) such that $|C_t(\mathbf{g}, M)| \leq \gamma^n$.*

It is interesting to note that Theorem 2 gives exponentially small upper bounds on the absolute value of the correlation with parity for polynomials possibly with arbitrarily large degree that previous techniques did not apply to. It is also interesting that we get exponentially small correlation with respect to parity over arbitrary nonempty subsets of the variables. On the other hand, Theorem 2 does not apply for example to all degree one polynomials, since repeating rows (according to the coefficients of Q) means that the matrix is singular. We are able to extend our results to a much larger class of matrices, that also includes all degree one polynomials. First we consider an extension of the non-singularity condition, next we state our results with respect to arbitrary matrices that have a partition into submatrices with not too much overlap between the subspaces spanned by their rows.

Definition 5. *A matrix $M \in \{0,1\}^{m \times n}$ is* block non-singular *over $\mathbb{Z}_2$ if M can be partitioned into submatrices $M_1, \ldots, M_k$ with $M_i \in \{0,1\}^{m_i \times n}$ for $i \in [k]$, such that $\sum_i^k rk_2(M_i) = rk_2(M) = n$.*

Note that the above definition implies that the linear subspaces over $\mathbb{Z}_2$ spanned by the rows of the different blocks are disjoint, except containing the $\mathbf{0}$ vector. In other words, the row-space of the matrix M is the direct sum of the row-spaces of the submatrices in the partition.

Theorem 3. *Let $M \in \{0,1\}^{m \times n}$ be a block non-singular matrix over $\mathbb{Z}_2$. Let $q \geq 3$ be an odd integer, and let $t \in [q-1]$. Given $\mathbf{g} \in \{0,1\}^n$, let $\ell(\mathbf{g})$ be the smallest number of blocks in the partition that contribute a nonzero vector to obtaining $\mathbf{g}$ as a linear combination over $\mathbb{Z}_2$ of the rows of M. There exists $\gamma = \gamma(q) \in [0,1)$ (depending only on q) such that $|C_t(\mathbf{g}, M)| \leq \gamma^{\ell(\mathbf{g})}$.*

Note that if M corresponds to an arbitrary degree one polynomial, then M is block non-singular, and $\ell(\mathbf{g}) = |\mathbf{g}|$ for any $\mathbf{g} \in \{0,1\}^n$. Thus, the above theorem contains Goldmann's result [Go] on the correlation between parity and degree one polynomials as a special case.

Given a matrix M, the bounds on the correlation obtained by Theorem 3 depend via $\ell(\mathbf{g})$ on over which subset of variables the parity is taken. On the other hand, Theorem 3 can be extended to yield exponentially small bounds on the correlation as long as the subspaces over $\mathbb{Z}_2$ spanned by the blocks do not overlap too much and many blocks are needed to span $\mathbf{g}$, that is, when $\sum_{i=1}^{k} rk_2(M_i) - rk_2(M)$ is relatively small and $\ell(\mathbf{g})$ is relatively large.

If we further restrict the class of polynomials and allow only coefficients relatively prime to q or 0 modulo q, we obtain a statement that gives the same (potentially exponentially small) upper bound on the absolute value of the correlation with the parity of *every* nonempty subset of the variables.

We say that a matrix $M \in \{0,1\}^{m\times n}$ is *nontrivial* if the polynomial Q it represents is not identically 0 modulo q over $\{-1,1\}^n$. Typically we are only interested in estimating $C_t(\mathbf{g}, M)$ for nontrivial M. Moreover, we can assume without loss of generality that every submatrix formed by a subset of the rows of M is nontrivial: deleting the rows of a trivial submatrix cannot change the value of $C_t(\mathbf{g}, M)$.

Theorem 4. *Let $M \in \{0,1\}^{m\times n}$, and $\mathbf{g} \in \{0,1\}^n$. Assume that every submatrix formed by a subset of the rows of M is nontrivial. Let $M_1, \dots, M_k$ be an arbitrary partition of the nonzero rows of M into blocks, and let $r = max_{i\in[k]} rk_2(M_i)$. Let $q \geq 3$ be an odd integer, and let $t \in [q-1]$. Assume that M corresponds to a polynomial such that all coefficients are either relatively prime to q or 0 modulo q. Then there exists $\gamma = \gamma(q,r) \in [1/2, 1)$ (depending only on q and r) such that $|C_t(\mathbf{g}, M)| \leq 2^{\sum_{i=1}^{k} rk_2(M_i) - rk_2(M)} \gamma^k$.*

If q is prime, the extra condition we consider does not impose any restrictions, but for composite q it is essential: we have an example of a (mod 15) polynomial that otherwise satisfies the conditions of the theorem, but has coefficients not relatively prime to 15 and has constant correlation with parity.

Having $rk_2(M_i) \leq r$ for each block is not essential, it just makes the theorem simpler to state. It is enough for getting exponentially small upper bounds that a large number of blocks has small rank. Note that this holds for example if M is block non-singular with sufficiently many blocks, and in this case the correlation with parity over *every* nonempty subset of variables is exponentially small.

References

[Aj] M. Ajtai. Σ_1^1-formulae on finite structures. *Ann. Pure Appl. Logic* 24, 1983, 1-48.

[Al] E. Allender. A note on the power of threshold circuits. *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, 1989, 580–584.

[AB] N. Alon, R. Beigel. Lower bounds for approximations by low degree polynomials over $\mathbb{Z}_m$, Proceedings of the 16th Annual IEEE Conference on Computational Complexity, 2001, 184–187

[Ba] D. Barrington. Bounded-width polynomial size branching programs recognize exactly those languages in NC_1. *Journal of Computer and System Sciences* 38 (1989), 150-164.

[BM] R. Beigel, A. Maciel. Upper and lower bounds for some depth-3 circuit classes, in *Proceedings of the 30th Symposium on Foundations of Computer Science*, 1989, 580–584.

[Bo05] J. Bourgain. Estimation of certain exponential sums arising in complexity theory, *C.R. Acad. Sci. Paris* Ser. I 340 (2005) pp. 627-631.

[CGT] J.-Y. Cai, F. Green, and T. Thierauf. On the correlation of symmetric functions. *Mathematical Systems Theory* 29 (1996), 245–258.

[FSS] M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial hierarchy. *Mathematical Systems Theory* 17, 1984, 13–27.

[Gr99] F. Green. Exponential sums and circuits with single threshold gate and mod-gates , *Theory Computing Systems* 32 (1999), 453-466.

[Gr02] F. Green. The Correlation between parity and quadratic polynomials mod 3, *Proceedings of the 17th Annual IEEE Conference on Computational Complexity*, 2002, 65–72.

[GRS] F. Green, A. Roy, H. Straubing. Bounds on an exponential sum arising in boolean circuit complexity, *C.R. Acad. Sci. Paris* Ser. I 340 (2005).

[Go] M. Goldmann. A note on the power of majority gates and modular gates, *Information Processing Letters* **53** (1995), 321–327

[Gro] V. Grolmusz. A weight-size tradeoff for circuits with mod m gates, in *Proceedings of the 26th ACM Symposium on the Theory of Computing*, 1994, 68–74.

[GT] V. Grolmusz, G. Tardos. Lower bounds for $MOD_p - MOD_m$ circuits, *SIAM J. Comput.*, 29, No. 4, 2000, 1209–1222.

[HMPST] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy, and G. Turán. Threshold Circuits of bounded depth, *Proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science*, 1987, 99-110

[HM] K. Hansen and P. Miltersen. Some meet-in-the-middle circuit lower bounds. in Proceedings of MFCS, 2004, 334–345.

[Ha] J. Håstad. *Computational Limitations of Small-Depth Circuits.* MIT Press, 1986.

[HG] J. Håstad and M. Goldmann. On the power of small depth threshold circuits. *Computational Complexity* 1(2), 1991, 113–129.

[KP] M. Krause, P. Pudlák. On the computational power of depth 2 circuits with threshold and modulo gates, in *Proceedings of the 26th ACM Symposium on the Theory of Computing*, 1994, 48–57.

[LN] R. Lidl and H. Niederreiter. Introduction to Finite Fields and Their Applications, Cambridge University Press.

[RW] A. Razborov and A. Wigderson. $n^{\Omega(\log n)}$ Lower bounds on the size of depth-3 threshold circuits with AND gates at the bottom. *Information Processing Letters* 45, 1993, 303–307.

[Sm] R. Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity, *Proceedings of the 19th Annual ACM Symposium on Theory of Computi ng*, 1987, 77–82

[Sz] G. Szegö. Orthogonal Polynomials. American Mathematical Society, 1939.

[Yao85] A. Yao. Separating the polynomial hierarchy by oracles. *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science*, 1985, 1–10.

[Yao90] A. Yao. On *ACC* and threshold circuits. *Proceedings of the 31th Annual IEEE Symposium on Foundations of Computer Science*, 1990, 619–627.

[Zb] S. Zabek. *Sur la périodicité modulo m des suites de nombres* $\binom{n}{k}$. Ann. Univ. Mariae Curie Sklodowska, A10 (1956), pp. 37–47.