Electronic Colloquium on Computational Complexity, Report No. 107 (2006)

An improved bound on correlation between polynomials over Z_m and MOD_q

Arkadev Chattopadhyay *
McGill University, Montreal, Canada
achatt3cs.mcgill.ca

Abstract

Let $m, q > 1$ be two integers that are co-prime and A be any subset of Z_m. Let P be any multi-variate polynomial of degree d in n variables over Z_m. We show that the MOD_q boolean function on n variables has correlation at most $\exp(-\Omega(n/(m2^{m-1})^d))$ with the boolean function f defined by $f(x) = 1$ iff $P(x) \in A$ for all $x \in \{0,1\}^n$. This improves on the bound of $\exp(-\Omega(n/(m2^m)^d))$ obtained in the breakthrough work of Bourgain [3] and Green et al. [9]. Our calculation is also slightly shorter than theirs.

Our result immediately implies the bound of $\exp(-\Omega(n/4^d))$ for the special case of $m = 2$. This bound was first reported in the recent work of Viola [11]. [11] states that it is not clear how to extend their method to general m.

1 Introduction

Understanding the computational power of constant depth circuits made of MAJORITY and MOD counting gates remains a very important and challenging open problem. Such circuits of even depth three have surprising power. Allender [1] shows that all functions in AC^0 (circuits using AND and OR gates of constant depth and polynomial size) can be computed by quasi-polynomial sized circuits of type $\mathrm{MAJ} \circ \mathrm{MOD}_m \circ \mathrm{AND}_{\mathrm{polylog}(n)}$ i.e. circuits with a MAJORITY gate at the output, MOD_m gates at the middle layer and AND gates of poly-log fan-in at the input layer, where $m > 1$ is any integer. It is of considerable interest to determine if such circuits are powerful enough to simulate the class ACC^0 i.e. circuits of constant depth and poly-size that use MOD_q gates in addition to AND and OR gates, for any fixed $q > 1$.

The study of upper bounds on correlation of boolean functions computed by polynomials of degree d over Z_m with a function f is motivated by the fact that such bounds yield a lower bound on the size of circuits of type $\mathrm{MAJ} \circ \mathrm{MOD}_m \circ \mathrm{AND}_d$ computing f. This is done by using the so called ϵ-discriminator lemma of Hajnal et al. [10]. A long line of

*supported by a NSERC scholarship and research grants of Prof. D. Thérien

ISSN 1433-8092

research (see [2, 4, 7, 8, 6]) sought to establish that polynomials of any constant degree d over Z_m have small correlation with MOD_q if m and q are co-prime. In a breakthrough work, Bourgain [3] showed an upper bound of $\exp(-\Omega(n/(m2^m)^d))$ on the correlation, for odd m. Green et al. [9] later modified Bourgain's proof to show that the bound holds for general m.

Our results Let $m > 1$ be any integer, and let P be any multi-linear polynomial of degree d over Z_m in n variables. For any $a \in Z_m$, let $K_n^P(a) = \{x \in \{0,1\}^n | P(x) = a\}$. Further, for any integers $q > 0$ and $0 \le b < q$, let

$$M_{n,q}(b) = \{x \in \{0,1\}^n | \sum_{i=1}^{n} x_i \equiv b \,(\mathrm{mod}\ q)\}$$

Our main technical lemma is the following :

Lemma 1 *Let $m, q > 1$ be integers that are co-prime. Then, there exists a constant $\beta = \beta(m,q)$, such that for every polynomial P of degree d over Z_m and for each $a \in Z_m$ and $0 \le b < q$, the following holds:*

$$\big||K_n^P(a) \cap M_{n,q}(b)| - \frac{1}{q}|K_n^P(a)|\big| \le exp(-\frac{\beta n}{(m2^{m-1})^d}) \tag{1}$$

The above lemma can be used to derive upper bounds on correlation between functions computed by polynomials over Z_m and MOD_q. A function f is computed by a polynomial P over Z_m if there exists an *accepting* set $A \subseteq Z_m$, such that for all $x \in \{0,1\}^n$, $f(x) = 1$ iff $P(x) \in A$. The MOD_q boolean function is defined in the following simple way : $\mathrm{MOD}_q(x) = 0$ iff $x \in M_{n,q}(0)$. We define the correlation between boolean functions f and g as the quantity below :

$$Corr(f,g) = \big|\Pr_x[f(x) = 1|g(x) = 1] - \Pr_x[f(x) = 1|g(x) = 0]\big| \tag{2}$$

where we are considering the uniform distribution over $\{0,1\}^n$.

Theorem 2 *For every pair of co-prime positive integers $m, q > 1$ and function f computed by a polynomial of degree d over Z_m, $Corr(f, MOD_q)$ is at most $exp(-\Omega(\frac{n}{(m2^{m-1})^d}))$.*

2 Proof of bound

Following [5], we will write $|K_n^P(a) \cap M_{n,q}(b)|$ as an exponential sum. Let $e_m(y)$ denote $\exp(\frac{2\pi j y}{m})$, where j is the complex square root of -1. Recall the following elementary fact : $\frac{1}{m}\sum_{a=0}^{m-1} e_m(ay)$ is 1 if $y = 0$ and is 0 otherwise. Then, it can be easily verified that

$$|K_n^P(a)\cap M_{n,q}(b)| = \sum_{x\in\{0,1\}^n}\Big(\frac{1}{m}\sum_{\alpha=0}^{m-1} e_m\big(\alpha(P(x)-a)\big)\Big)\Big(\frac{1}{q}\sum_{\beta=0}^{q-1} e_q\big(\beta(x_1+\cdots+x_n-b)\big)\Big) \tag{3}$$

Expanding the sum inside the second multiplicand and treating the case of $\beta = 0$ separately, one gets

$$(3) = \frac{1}{q}\sum_{x\in\{0,1\}^n}\Big(\frac{1}{m}\sum_{\alpha=0}^{m-1} e_m\big(\alpha(P(x)-a)\big)\Big) + \frac{1}{mq}\sum_{\alpha\in[m],\beta\in[q]-\{0\}} S_n^{m,q}(\alpha,\beta,P)e_m(-a\alpha)e_q(-b\beta) \tag{4}$$

where,

$$S_n^{m,q}(\alpha,\beta,P) = \sum_{x\in\{0,1\}^n} e_m(\alpha P(x))\cdot e_q\big(\beta(x_1+\cdots+x_n)\big) \tag{5}$$

Observing that the first sum in (4) is simply $\frac{1}{q}|K_n^P(a)|$ and $|e_m(-a\alpha)| = |e_q(-b\beta)| = 1$, we get :

$$\big||K_n^P(a)\cap M_{n,q}(b)| - \frac{1}{q}|K_n^P(a)|\big| \le \frac{1}{mq}\sum_{\alpha\in[m],\beta\in[q]-\{0\}} |S_n^{m,q}(\alpha,\beta,P)| \tag{6}$$

Lemma 1 gets proved by the bound on $|S_n^{m,q}(\alpha,\beta,P)|$ provided below.

Lemma 3 *For each pair of co-prime integers $m, q > 1$ there exists a constant $\beta = \beta(q)$ such that for every polynomial P of degree $d > 0$ in Z_m and numbers $\alpha \in [m]$, $\beta \in [q]-\{0\}$, the following holds :*

$$|S_n^{m,q}(\alpha,\beta,P)| \le exp\Big(-\frac{\beta n}{(m2^{m-1})^d}\Big) \tag{7}$$

Before we begin our formal calculations, we remind the reader that a slightly weaker estimate of $|S_n^{m,q}(\alpha,\beta,P)|$ was first obtained by Bourgain [3] and later generalized by Green et al [9]. The case when P is a linear polynomial was essentially dealt with in [4] and forms our base case just as in [3, 9].

In order to explain the intuition behind our proof of Lemma 3, we develop some definitions and notations. Let $f : \{0,1\}^n \to Z_m$ be any function. Consider any set $I \subseteq [n]$.

Note that each binary vector v of length $|I|$ can be thought of as a partial assignment to the input variables of f by assigning v to the variables in I in a natural way. Let $f^{I(v)}$ be the subfunction of f on variables not indexed in I induced by the partial assignment v to variables indexed in I. For any sequence $Y = \{y_1, \ldots, y_t\}$ having t boolean vectors from $\{0,1\}^n$, let f_Y be the function defined by $f_Y(x) = f(x) + \sum_{i=1}^{t} f(x \oplus y_i)$, where the sum is taken in Z_m. Let $I[Y] \subseteq [n]$ be the set of those indices on which every vector in Y is zero and $J[Y]$ be just the complement of $I[Y]$. Then, the following observation will be very useful in our calculation :

Observation 4 *Let P be a polynomial of degree d in n variables over Z_m for any $m > 1$. Then, for each sequence Y of $m-1$ boolean vectors in $\{0,1\}^n$, the polynomial $P_Y^{J[Y](v)}$ is a polynomial of degree $d-1$ in variables from $I[Y]$ for each vector $v \in \{0,1\}^{|J[Y]|}$.*

Proof:[of Lemma 3] We drop the superscript from $S_n^{m,q}$ to avoid clutter in the following discussion.We shall induce on the degree d of the polynomial. Our IH is that there exists a positive real constant $\mu_{d-1} < 1$ such that for all polynomials R of degree at most $d-1$ and for all $n \geq 0$ we have $|S_n(\alpha, \beta, R)| \leq 2^n \mu_{d-1}^n$. The base case of $d = 0$ is easily verified and is dealt with in earlier works on correlation. Note that μ_0 depends only on q. Our inductive step will yield a relationship between μ_{d-1} and μ_d that will also give us our desired explicit bound of (7).

As in [3, 9], we raise S_n to its mth power. Our point of departure from the earlier techniques, is to write $(S_n)^m$ in a different way.

$$(S_n)^m =$$

$$\sum_{y^1,\ldots,y^{m-1} \in \{0,1\}^n} \sum_{x \in \{0,1\}^n} e_m\Big(P(x) + \sum_{j=1}^{m-1} P(x \oplus y^j)\Big) e_q\Big(\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} (x_i \oplus y_i^1) + \cdots + \sum_{i=1}^{n} (x_i \oplus y_i^{m-1})\Big) \tag{8}$$

Let Y be the sequence of length $m-1$ formed by a given set of vectors $y^1, \ldots, y^{m-1}$. We denote by u and v respectively the projection of x to $I[Y]$ and $J[Y]$. Let n_I and n_J be the cardinality of $I[Y]$ and $J[Y]$ (note that $n_I + n_J = n$) . Then, one can verify

$$(8) =$$

$$\sum_{y^1,\ldots,y^{m-1} \in \{0,1\}^n} \sum_{v \in \{0,1\}^{n_J}} e_m\big(Q^{y^1,\ldots,y^{m-1}}(v)\big) e_q(n_J) \sum_{u \in \{0,1\}^{n_I}} e_m\big(P_Y^{I[Y](v)}(u)\big) e_q\big(m \sum_{i=1}^{n_I} u_i\big) \tag{9}$$

where $Q^{y^1,\ldots,y^{m-1}}$ is some polynomial that is determined by $y^1, \ldots, y^{m-1}$ and polynomial P.

The key thing to note is that Observation 4 implies $P_Y^{I[Y](v)}$ to be a polynomial of degree at most $d-1$ over u for every sequence $Y = y^1, \ldots, y^{m-1}$ and every vector v. Thus, the inside sum of (9) over the variable u can be estimated using our inductive hypothesis. Noting that the number of sequences Y for which $|I_Y| = k$ is exactly $\binom{n}{k}(2^{m-1}-1)^{n-k}$ and using the triangle inequality with the binomial theorem, we get.

$$|S_n|^m \leq \sum_{k=0}^{n} \binom{n}{k} (2^{m-1}-1)^{n-k} 2^{n-k} 2^k \mu_{d-1}^k \; = \; 2^{nm} \left(1 - \frac{1-\mu_{d-1}}{2^{m-1}}\right)^n \tag{10}$$

The rest of the calulation proceeds exactly as in Green et. al. [9]. We repeat it here for the sake of self-containment. Taking the mth root of both sides of (10), using the inequality $(1-x)^{1/m} \leq 1 - x/m$ if $0 \leq x < 1$ amd $m > 1$ after rearranging, we obtain

$$1 - \mu_d \geq \frac{1-\mu_{d-1}}{m2^{m-1}} \geq \frac{1-\mu_0}{\left(m2^{m-1}\right)^d} \tag{11}$$

Substituting $\beta = 1 - \mu_0$, one gets $\mu_d \leq \exp\left(-\frac{\beta}{(m2^{m-1})^d}\right)$. This immediately yields (7) in Lemma 3. ∎

We now show that Theorem 2 follows from Lemma 1 easily. Note that the argument we present has been used in a slightly more general setting in [5] (see proof of their Lemma 3).

Proof: [of Theorem 2] Let P be a polynomial of degree d computing f over Z_m with an accepting set A. Then, using the definition of correlation as given in (2), we can write

$$Corr(f, \mathrm{MOD}_q) \leq \sum_{a \in A} \left| \Pr_x[P(x) = a | x \notin M_{n,q}(0)] - \Pr_x[P(x) = a | x \in M_{n,q}(0)] \right| \tag{12}$$

Since MOD_q is an almost balanced function i.e. for any b $|\Pr_x[x \in M_{n,q}(b)] - \frac{1}{q}| = 2^{-\Omega(n)}$, we can rewrite (12) as

$$\begin{aligned} &\text{RHS of (12)} \leq \\ &2^{-\Omega(n)} + \left(\frac{q}{q-1}\right) \sum_{a \in A} \left| \Pr_x[P(x) = a \wedge x \notin M_{n,q}(0)] - (q-1) \Pr_x[P(x) = a \wedge x \in M_{n,q}(0)] \right| \end{aligned} \tag{13}$$

which implies the following :

RHS of (13) $\leq$

$$2^{-\Omega(n)} + (\frac{q}{q-1}) \sum_{a \in A} \sum_{b \in [q]-\{0\}} | \Pr_x[P(x) = a \wedge x \in M_{n,q}(b)] - \Pr_x[P(x) = a \wedge x \in M_{n,q}(0)]| \tag{14}$$

Using the bound of (1) and the triangle inequality, we get

$$\text{RHS of (14)} \leq (2q^2 m) \cdot \exp(-\frac{\beta n}{(m 2^{m-1})^d}) + 2^{-\Omega(n)} \tag{15}$$

which gives us our bound. ∎

Acknowledgements

The author thanks Fred Green for several discussions on this problem and for reading the first draft of the proof of Lemma 1. We also thank Navin Goyal for pointing out the recent report of Viola [11].

References

[1] E. Allender. A note on the power of Threshold circuits. In *FOCS*, pages 580–584, 1989.

[2] N. Alon and R. Beigel. Lower bounds for approximations by low degree polynomials over z_m. In *IEEE Sixteenth Annual Conference on Computational Complexity*, pages 184–187, 2001.

[3] J. Bourgain. Estimation of certain exponential sums arising in complexity theory. *C.R. Acad. Sci. Paris, Ser.1*, 340:627–631, 2005.

[4] J.-Y. Cai, F. Green, and T. Thierauf. On the correlation of symmetric functions. *Math. Systems Theory*, 29:245–258, 1996.

[5] A. Chattopadhyay, N. Goyal, P. Pudlák, and D. Thérien. Lower bounds for circuits with MOD_m gates. In *FOCS*, 2006.

[6] E. Duenez, S. Miller, A. Roy, and H. Straubing. Incomplete quadratic exponential sums in several variables. *J. Number Theory*, 116(1):168–199, 2006.

[7] F. Green. Exponential sums and circuits with a single threshold gate and mod-gates. *Theory Comput. Systems*, 32:453–466, 1999.

[8] F. Green. The correlation between parity and quadratic polynomials mod 3. *J. Comput. System. Sci.*, 69(1):28–44, 2004.

[9] F. Green, A. Roy, and H. Straubing. Bounds on an exponential sum arising in Boolean circuit complexity. *C.R. Acad. Sci. Paris, Ser.1*, 341:279–282, 2005.

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

[11] E. Viola. New correlation bounds for GF(2) polynomials using Gowers uniformity. In *ECCC*, number 97, 2006.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/