Polynomial approximations over $\mathbb{Z}/p^k \mathbb{Z}$

Abhishek Bhrushundi∗ Prahladh Harsha† Srikanth Srinivasan‡
August 23, 2015

Abstract

We study approximation of Boolean functions by low-degree polynomials over the ring $\mathbb{Z}/p^k \mathbb{Z}$. More precisely, given a Boolean function $F : \{0, 1\}^n \rightarrow \{0, 1\}$, define its $k$-lift to be $F_k : \{0, 1\}^n \rightarrow \{0, p^k - 1\}$ by $F_k(x) = p^k - F(x) \mod p^k$. We consider the fractional agreement of $F_k$ with degree $d$ polynomials from $\mathbb{Z}/p^k \mathbb{Z}[x_1, \ldots, x_n]$, which we call $\gamma_{d,k}(F)$.

Our results are the following:

• Power of this model: We observe that as $k$ increases, $\gamma_{d,k}(F)$ cannot decrease. We give two kinds of examples where $\gamma_{d,k}(F)$ actually increases. The first is an infinite family of functions $F$ such that $\gamma_{2,2}(F) - \gamma_{2,1}(F) \geq \Omega(1)$. The second is an infinite family of functions $F$ such that $\gamma_{d,1}(F) \leq \frac{1}{2} + o(1)$ — as small as possible — but $\gamma_{d,3}(F) \geq \frac{1}{2} + \Omega(1)$.

• Limitations of this model: We show that the Majority function $\text{Maj}_n$ satisfies

$$\gamma_{d,k}(\text{Maj}_n) \leq \frac{1}{2} + \frac{O(d)}{\sqrt{n}}$$

irrespective of the value of $k$, strengthening classical results of Szegedy and Smolensky. Previous results only yielded $\gamma_{d,k}(\text{Maj}_n) \leq \frac{1}{2} + \frac{O(pd^k)}{\sqrt{n}}$ for $k > 1$.

We observe that the model we study subsumes the model of non-classical polynomials and in particular, proving limitations on our model also proves upper bounds on the agreement of non-classical polynomials with Boolean functions. In particular, we can use our second result to show that no non-classical polynomial of degree $d$ can agree with $\text{Maj}_n$ on more than a $\frac{1}{2} + \frac{O(d)}{\sqrt{n}}$ fraction of its inputs, confirming a conjecture of Bhowmick and Lovett [In Proc. 30th Computational Complexity Conf., pages 72-87, 2015].

Our result also strengthens results about weak representations of the Majority function modulo large prime powers.

∗Department of Computer Science, Rutgers University, USA. abhishek.bhr@cs.rutgers.edu. Work done while the author was visiting the Tata Institute of Fundamental Research. Research supported in part by UGC-ISF grant 1399/4.

†Tata Institute of Fundamental Research, India. prahladh@tifr.res.in. Research supported in part by UGC-ISF grant 1399/4.

‡Department of Mathematics, Indian Institute of Technology, Bombay, India. srikanth@math.iitb.ac.in.
1 Introduction

Many lower bound results in circuit complexity are proved by showing that any small sized circuit in a given circuit class can be approximated by a function from a simple computational model (e.g., small depth circuits by low-degree polynomials) and subsequently showing that this is not possible for some suitable “hard function”.

A classic case in point is the work of Razborov [16] which shows lower bounds for $\text{AC}^0[\oplus]$, the class of constant depth circuits made up of AND, OR and $\oplus$ gates. Razborov shows that any small $\text{AC}^0[\oplus]$ circuit $C$ can be well approximated by a low-degree multivariate polynomial $Q(x_1,\ldots,x_n) \in \mathbb{F}_2[x_1,\ldots,x_n]$ in the sense that

$$\Pr_{x \sim \{0,1\}^n} [Q(x) \neq C(x)] = o(1).$$

The next step in the proof is to show that the hard function, on the other hand, does not have any such approximation. Razborov does this for a suitable symmetric function, Smolensky [18] for the MOD$_q$ function (for constant odd $q$), and Szegedy [20] and Smolensky [19] for the Majority function $\text{Maj}_n$ on $n$ bits.

Given the importance of the above lower bound, polynomial approximations in other domains and metrics have been intensely investigated and have resulted in interesting combinatorial constructions and error-correcting codes [10, 7], learning algorithms [14, 11] and more recently in the design of algorithms for combinatorial problems [24, 1] as well.

To describe the model of polynomial approximation considered in this paper, we first recall the Razborov [16] model of polynomial approximation. Given a Boolean function $F : \{0,1\}^n \to \{0,1\}$ and degree $d \leq n$, Razborov considers the largest $\gamma$ such that there is a polynomial $Q \in \mathbb{F}_2[x_1,\ldots,x_n]$ that has agreement at least $\gamma$ with $F$ (i.e., $\Pr_{x} [Q(x) = F(x)] \geq \gamma$). Call this $\gamma_d(F)$. In this notation, Szegedy [20] and Smolensky’s [19] results for the Majority function can be succinctly stated as

$$\gamma_d(\text{Maj}_n) \leq \frac{1}{2} + \frac{O(d)}{\sqrt{n}}.$$

We consider a generalization of the above model to rings $\mathbb{Z}/2^k\mathbb{Z}$ in the following simple manner. To begin with, we consider the ring $\mathbb{Z}/4\mathbb{Z}$. Given a Boolean function $F$, let $F_2 : \{0,1\}^n \to \{0,2\} \subseteq \mathbb{Z}/4\mathbb{Z}$ be the 2-lift of $F$ defined as $F_2(x) : = 2^{2^F(x)}$ (i.e., $F_2(x) : = 0$ if $F(x) = 0$ and $F_2(x) : = 2$ otherwise). Once again, we can define $\gamma_{d,2}(F)$ to be the largest $\gamma$ such that there exists a multivariate polynomial $Q_2 \in \mathbb{Z}/4\mathbb{Z}[x_1,\ldots,x_n]$ that has agreement $\gamma$ with $F_2$. Note that $\gamma_{d,2}(F) \geq \gamma_d(F)$ since if, for instance, $Q(x) = x_1x_2 + x_3 \in \mathbb{F}_2[x_1,\ldots,x_n]$ has agreement $\gamma$ with $F$, then $Q_2 := 2(x_1x_2 + x_3) \in \mathbb{Z}/4\mathbb{Z}[x_1,\ldots,x_n]$ also has the same agreement $\gamma$ with $F_2$. Hence, proving upper bounds for $\gamma_{d,2}(F)$ is at least as hard as proving upper bounds for $\gamma_d(F)$.

More generally, we can extend these definitions to $\gamma_{d,k}(F)$, the agreement of $F_k$, the $k$-lift of $F$, defined as $F_k(x) : = 2^{k^F(x)} \mod 2^k$, with polynomials from $\mathbb{Z}/2^k\mathbb{Z}[x_1,\ldots,x_n]$. It is not hard to show that $\gamma_{d,k+1}(F) \geq \gamma_{d,k}(F)$ and hence as $k$ increases, the problem of proving upper bounds on $\gamma_{d,k}(F)$ can only get harder.

Our motivation for this model comes from a recent work of Bhowmick and Lovett [6], who study the maximum agreement between non-classical polynomials of degree $d$ and a Boolean function $F$, which is similar to $\gamma_{d,d}(F)$ (see Section 5 for an exact translation between the above model
and non-classical polynomials). In particular, non-classical polynomials of degree \(d\) can be considered a subset of the degree \(d\) polynomials in \(\mathbb{Z}/2^d\mathbb{Z}[x_1, \ldots, x_n]\). With respect to correlation\(^1\), Bhowmick and Lovett showed that there exist non-classical polynomials (and hence polynomials in \(\mathbb{Z}/2^d\mathbb{Z}[x_1, \ldots, x_n]\)) of logarithmic degree that have very good correlation with the \(\text{Maj}_n\) function. With respect to agreement, they show that low-degree non-classical polynomials can only have small agreement with the Majority function. Their results stated in our language, imply that

\[
\gamma_{d,k}(\text{Maj}_n) \leq \frac{1}{2} + \frac{O(d \cdot 2^d)}{\sqrt{n}}. \tag{1.1}
\]

In particular, if \(d > \log n\), this result unfortunately does not give any non-trivial bound on the maximum agreement between non-classical polynomials of degree \(d\) and the \(\text{Maj}_n\) function. Bhowmick and Lovett, however, conjectured that this result could be improved and left open the question of whether non-classical polynomials of degree \(d\) can do any better than classical polynomials of degree \(d\) (c.f., [20]) in approximating the Majority function. More generally, they informally conjectured that although non-classical polynomials achieve better correlation with Boolean functions than their classical counterparts, they possibly do not approximate Boolean functions any better than classical polynomials. Our work stems from trying to answer these questions.

### 1.1 Our results

We prove the following results about agreement of Boolean functions with polynomials over the ring \(\mathbb{Z}/2^k\mathbb{Z}^\perp\):

1. We show that for \(\text{Maj}_n\), the majority function on \(n\) bits, and any \(d, k \in \mathbb{Z}^+\),

\[
\gamma_{d,k}(\text{Maj}_n) \leq \frac{1}{2} + \frac{O(d)}{\sqrt{n}},
\]

\(^2\) confirming the Bhowmick-Lovett conjecture. To the best of our knowledge, this is the first improvement on (1.1) in this setting and in particular, the first such result where there is no dependence on \(k\) on the right hand side.

This result easily extends to the setting of polynomial approximation by polynomials \(Q \in \mathbb{Z}/p^k\mathbb{Z}[x_1, \ldots, x_n]\) (where \(F_k\) now maps \(\{0, 1\}^n\) to \(\{0, p^k-1\}\)) and more generally to \(Q \in \mathbb{Z}/m\mathbb{Z}[x_1, \ldots, x_n]\) for any \(m \geq 2\) (where Boolean functions now map \(\{0, 1\}^n\) to \(\{0, a\}\) for some non-zero \(a \in \mathbb{Z}/m\mathbb{Z}\)).

Our result also implies bounds for weak representations [3] of the majority function. In particular, we show that any polynomial over \(\mathbb{Z}/p^k\mathbb{Z}\) of degree \(d\) can weakly represent majority on at most a \(\frac{1}{2} + \frac{O(d)}{\sqrt{n}}\) fraction of points. This strengthens classical results of Szegedy [20] and Smolensky [19] who gave bounds of the form \(\frac{1}{2} + \frac{O(dp^d)}{\sqrt{n}}\) for \(k > 1\).

\(^1\)The correlation between \(F, G : \{0, 1\}^n \rightarrow \mathbb{Z}/2^k\mathbb{Z}\) is defined to be \(\mathbb{E}_x[\omega^{F(x) - G(x)}]\) where \(\omega\) is the primitive \(2^k\)th root of unity in \(\mathbb{C}\). If \(F, G\) are \(\{0, 2^{k-1}\}\)-valued, then this quantity is exactly \(2\gamma - 1\) where \(\gamma\) is the agreement between \(F\) and \(G\). Otherwise, however, it does not measure agreement.

\(^2\)the constant in the \(O(\cdot)\) is an absolute constant.
2. The above result demonstrates that polynomials over the ring $\mathbb{Z}/2^k\mathbb{Z}$ cannot approximate the Majority function any better than polynomials over $\mathbb{Z}/2\mathbb{Z}$. But do there exist Boolean functions for which agreement can increase by increasing $k$? That is, do there exist Boolean $F$ such that $\gamma_{d,k}(F) > \gamma_{d,1}(F)$?

It is not hard to show that this is impossible for $d = 1$. Further, it can be shown that if $\gamma_{d,k}(F) > 1 - \frac{1}{2^k}$, then $\gamma_{d,k}(F) = \gamma_{d,1}(F)$. Keeping this in mind, the first place where we can expect larger $k$ to show better agreement is $\gamma_{2,2}$ vs. $\gamma_{2,1}$. Our next result shows that there are indeed separating examples in this regime.

(a) There exist infinitely many $n$ and $F : \{0,1\}^n \to \{0,1\}$ such that $\gamma_{2,1}(F) \leq 5/8 + o(1)$ but $\gamma_{2,2}(F) \geq 3/4$.

Note that since $F$ is Boolean, $\gamma_{d,k}(F) \geq 1/2$ for any $d,k$. We then ask if there exist Boolean functions $F$ such that $\gamma_{d,1}(F)$ is more or less the trivial bound of $1/2$, while $\gamma_{d,k}(F)$ is significantly larger even for $d' < d$ and $k > 1$. In this context, we show the following result.

(b) Fix any $\ell \geq 2$. For large enough $n$, there is a Boolean function $F : \{0,1\}^n \to \{0,1\}$ such that $\gamma_{2^{\ell-1},1}(F) \leq 1/2 + o(1)$ but $\gamma_{2^{\ell-1}+2^{\ell-2},3}(F) \geq 9/16 - o(1)$.

This answers an informal question of Bhowmick and Lovett who ask if non-classical polynomials of degree $d$ can approximate Boolean functions better than classical polynomials of the same degree.

1.2 Related work

As we mentioned earlier, polynomial approximations of Boolean functions have been intensely investigated. We refer the reader to the surveys of Beigel [4] and Viola [23] for a more in-depth account of these results.

We concentrate on a generalization of our model that already exists in the literature [3]. Given any $m \geq 2$, we say that a polynomial $P \in \mathbb{Z}/m\mathbb{Z}[x_1,\ldots,x_m]$ weakly represents a Boolean function $F : \{0,1\}^n \to \{0,1\}$ at $x \in \{0,1\}^n$ if $P(x) = 0 \iff F(x) = 0$. We say $P$ strongly represents $F$ at $x$ if $P(x) = F(x)$. Note that if $\gamma_{d,k}(F) \geq \gamma$, then in particular, there exists a polynomial $P \in \mathbb{Z}/2^k\mathbb{Z}[x_1,\ldots,x_n]$ of degree $d$ that weakly represents $F$ at a $\gamma$ fraction of the inputs from $\{0,1\}^n$. Thus, this is a stronger model than ours.

Razborov [16] and Smolensky [18] both proved lower bounds for strongly representing a function $F$ at many inputs over a field. Their results also imply a lower bound for weakly representing $F$ modulo an integer $m \geq 2$ when $m$ is either a prime $p$ or of the form $p^k$. This involves converting a polynomial $P \in \mathbb{Z}/m\mathbb{Z}[x_1,\ldots,x_n]$ that weakly represents $F$ at $S \subseteq \{0,1\}^n$ to a polynomial $\tilde{P} \in \mathbb{Z}/p\mathbb{Z}[x_1,\ldots,x_n]$ that strongly represents $F$ at $S$ (see, e.g.,[4, Theorem 18]. Bhowmick and Lovett [6] also use similar ideas). However, $\deg(\tilde{P}) = \deg(P) \cdot (m-1)$ in general, which makes this method unsuitable for large $m$.

To the best of our knowledge, Szegedy [20] was the first to prove a lower bound for weak representations that also works for polynomials modulo large $m$. Specifically, when $m$ is a prime, Szegedy shows that degree $d$ polynomials in $\mathbb{Z}/m\mathbb{Z}[x_1,\ldots,x_n]$ cannot weakly represent $\text{Maj}_n$, the
majority function at more than a $\frac{1}{2} + O(d) \sqrt{n}$ fraction of points\(^3\) (also proved later by Smolensky \cite{19}). Szegedy also proves similar results for composite \(m\) but comparable results are obtained only when \(m = O(1)\) and a prime power. For instance, when \(m\) is a product of 2 distinct primes, then the upper bound on the agreement with with degree \(d\) polynomials (in the sense of weak representation of majority) degrades to $\frac{3}{4} + o(1)$. (See also the discussion after Theorem 25 in Beigel’s survey \cite{4}.)

Tsai \cite{22} shows lower bounds of the degree required to weakly represent the majority function (and also other functions) at all inputs in \(\{0,1\}^n\). He shows that for any modulus \(m \geq 2\), any polynomial that weakly represents majority must have degree at least \(n/2\).

Our results are incomparable to those of Szegedy \cite{20} and Tsai \cite{22} since our bounds are for a somewhat weaker model. However, we note that for the case when \(m\) is a prime power, we can actually show that our bounds also hold for weakly representing the Majority function (see Remark 3.5). We leave the problem of generalizing this to all composite moduli as an open question.

### 1.3 Proof Ideas

#### 1.3.1 Agreement of low-degree polynomials with \(\text{Maj}_n\)

We start by recalling a variant\(^4\) of the arguments of Szegedy \cite{20} and Smolensky \cite{19} for upper bounding \(\gamma_{d,1}(\text{Maj}_n)\).

We will use crucially the notion of an interpolating set for polynomials of degree \(D\) in \(\mathbb{F}_2[x_1, \ldots, x_n]\), which is a subset \(I \subseteq \{0,1\}^n\) such that the only polynomial of degree \(D\) that vanishes at all points in \(I\) is the zero polynomial. A simple counting argument tells us that any interpolating set for degree \(D\) must have size at least \(\sum_{i=0}^{D} \binom{n}{i}\), which is the size of any Hamming ball of radius \(D\). It is also known that this bound is tight, since any Hamming ball of radius \(D\) in \(\{0,1\}^n\) is an interpolating set for degree \(D\). Further, the proofs of these facts extend straightforwardly to polynomials from \(\mathbb{Z}/2^k\mathbb{Z}\) as well.

Now we show how to upper bound \(\gamma_{d,1}(\text{Maj}_n)\). Say that a polynomial \(P \in \mathbb{F}_2[x_1, \ldots, x_n]\) of degree \(d\) computes \(\text{Maj}_n\) correctly on the inputs in \(S_P \subseteq \{0,1\}^n\) where \(|S_P|/2^n \geq \frac{1}{2} + \varepsilon\). We first find a non-zero degree \(D\) (\(D\) as small as possible) polynomial \(Q \in \mathbb{F}_2[x_1, \ldots, x_n]\) such that \(Q\) is zero at all points of \(S_P\). To be able to do this, we need that \(S_P\) is not an interpolating set for degree \(D\), which will be the case if \(D\) is chosen so that the Hamming ball of radius \(D\) is large enough: in particular, \(D = \frac{2}{\varepsilon} - O(\varepsilon \sqrt{n})\) will do.

Consider the polynomial \(R = Q \cdot P\). On any input \(x \in \text{Maj}_n^{-1}(0)\), \(R(x) = 0\) since either \(x \not\in S_P\), and hence \(Q(x) = 0\), or \(x \in S_P\), which implies that \(P(x) = 0\). Secondly, since Hamming balls are interpolating sets, we can also say that there is a point \(x_0\) of Hamming weight \(> n/2\) (i.e. a point in the Hamming ball of radius \(n/2 - 1\) around the all 1s vector) such that \(Q(x_0) \neq 0\). Hence we have

\[
R(x_0) = Q(x_0)P(x_0) \neq 0.
\]

Therefore, \(R\) is a non-zero polynomial of degree at most \(\text{deg}(Q) + \text{deg}(P)\) that vanishes at all inputs of Hamming weight \(< n/2\). Again appealing to the fact that Hamming balls are interpolating sets, we can conclude that there is a point \(x_0\) that vanishes at all \(\{0,1\}^n\) points except \(S_P\), which implies that \(\gamma_{d,1}(\text{Maj}_n)\) must be at least \(\frac{1}{2} + \varepsilon\).

\(^3\)The constant in the \(O(\cdot)\) is independent of \(m\).

\(^4\)This is essentially a “dual” view of their argument.
lating sets, this implies that $\deg(Q) + \deg(P) \geq n/2$ and hence that $\varepsilon \leq O(d/\sqrt{n})$, finishing the proof of the theorem.

The problem in extending the above to the setting of polynomials in $\mathbb{Z}/2^k\mathbb{Z}$ is that (1.2) does not work any more, since the product of two non-zero elements in $\mathbb{Z}/2^k\mathbb{Z}$ can be zero. In particular, it could be the case that $Q(x_0)$ is even and $P(x_0) = \text{Maj}_{n,k}(x_0) = 2^{k-1}$, in which case their product is 0.

To overcome this, we instead try to find a $Q$ that vanishes on $S_F$ and moreover $Q(x_0)$ is odd for some $x_0 \in S_F$. We say that $S_F$ is forcing for degree $D$ if this cannot be done. Note that this notion is different from the notion of interpolating sets: every interpolating set is of course forcing, but the converse is not true (see Example 3.3). Our main technical lemma (Lemma 3.4) is a tight lower bound on the size of forcing sets for degree $D$. We show that any such set must have size at least $\sum_{i=0}^{D} \binom{n}{i}$, which is tight even for interpolating sets as mentioned above. The proof of this lemma is via a shifting argument.

### 1.3.2 Separation results

We begin by outlining the ideas used for proving the existence of Boolean functions $F$ for which $\gamma_{2,2}(F) > \gamma_{2,1}(F)$. We consider Boolean functions in $2n$ variables $x_1, \ldots, x_n, y_1, \ldots, y_n$. It can be shown that for any Boolean function $F$, $\gamma_{2,2}(F) = \gamma_{2,1}(F)$ unless $\gamma_{2,2}(F) \leq 3/4$. If we restrict our attention to Boolean functions $F$ for which this bound is tight i.e. $\gamma_{2,2}(F) \geq 3/4$, the Schwartz-Zippel lemma dictates that the quadratic polynomials over $\mathbb{Z}/4\mathbb{Z}$ that have maximum possible agreement with $F_2$ (the 2-lift of $F$) are polynomials of the form $L_1(x,y)L_2(x,y) + 2Q$ where $Q$ is some quadratic and $L_1,L_2$ are linear functions.

Motivated by the above, we try to work our way backwards: we begin with a quadratic polynomial $P$ over $\mathbb{Z}/4\mathbb{Z}$ of the above form and construct a Boolean function $F$ whose 2-lift has agreement 3/4 with $P$, hoping that $F$ has $< 3/4$ agreement with quadratic polynomials over $\mathbb{F}_2$. In particular, we choose $P(x,y) = (\sum_{1 \leq i \leq n} x_i) (\sum_{1 \leq i \leq n} y_i)$, and define $F$ as

$$F(x,y) = \begin{cases} 0 & \text{if } P(x,y) = 0 \\ 1 & \text{if } P(x,y) = 2 \\ H(x,y) & \text{otherwise,} \end{cases}$$

where $H(x,y)$ is a Boolean function.

By construction, $\gamma_{2,2}(F) \geq 3/4$. Suppose $S = \{(x,y) \mid P(x,y) \in \{0,2\}\}$. We first ensure that $F$ has $\leq \frac{1}{2} + o(1)$ agreement with quadratic polynomials over $\mathbb{F}_2$ when restricted to $S$, by carefully choosing $H(x,y)$ to be a function that’s “hard” for these polynomials (e.g., a random function). Using the fact that $F$ restricted to $S$ is a high rank quadratic polynomial, we then show that, over the set $S$, $F$ has “sufficiently” low agreement with quadratic polynomials, to prove that $\gamma_{2,1}(F) \leq 5/8 + o(1)$.

We now sketch the ideas behind the second example: i.e. that there are Boolean functions $F$ such that $\gamma_{d,1}(F) \leq \frac{1}{2} + o(1)$ but $\gamma_{d',k} \geq \frac{1}{2} + \Omega(1)$ for some $d' < d$. We consider $F = S_{2'}(x)$, the elementary symmetric polynomial of degree $2'$ from $\mathbb{F}_2[x_1, \ldots, x_n]$, which is known to satisfy $\gamma_{2'-1,1}(F) \leq \frac{1}{2} + o(1)$ for any constant $\ell$ [9, 2].

\footnote{Though this choice of $F$ may seem somewhat arbitrary, it is not: it can be shown that for $d',k$ small enough in}
It is a well-known fact that for \( x \in \{0, 1\}^n \), \( S_2(x) \) is the \((\ell + 1)\)th least significant bit of \(|x|\) (the Hamming weight of \( x \)) written in base 2. Thus, to come up with \( P \in \mathbb{Z}/2^k\mathbb{Z}[x_1, \ldots, x_n] \) that has good agreement with \( S_{2'\ell,k} \), we need to choose \( P \) such that the value of \( P(x) \) has some information about the \((\ell + 1)\)th least significant bit of \(|x|\). A natural choice for \( P \) is the elementary symmetric polynomial of degree \( d' \) (for some \( d' < d \)). Note that \( P(x) = \sum_{|T| = d'} \prod_{i \in T} x_i = \left( \frac{|x|}{d'} \right) \).

It can be shown that the values of \( \left( \frac{|x|}{d'} \right) \) modulo \( 2^k \) depend on the values of \(|x|\) modulo \( 2^{\lfloor \lg d' \rfloor} \cdot 2^k \). Hence, if \( 2^{\lfloor \lg d' \rfloor} \cdot 2^k \geq 2^{\ell+1} \), then there is a possibility that \( P(x) \) might have non-trivial agreement with \( S_{2'}(x) \). Unfortunately, not every choice of \( d', k \) satisfying the above constraint works, but we show that for \( d' = 2^{\ell-1} + 2^{\ell-2} \) and \( k = 3 \), the polynomial \( P(x) \) does indeed have agreement \( \frac{1}{2} + \Omega(1) \) with \( S_{2',k}(x) \). In order to analyze the value of \( P(x) \), we use a classic theorem of Kummer that characterizes the largest power of a prime dividing a given binomial coefficient.

### 1.4 Organisation

We start with some preliminaries in Section 2. In Section 3, we prove upper bounds for \( \gamma_{d,k}^{(p)}(\text{Maj}_n) \). Next, in Section 4, we show some separation results. In Section 5, we discuss how our model relates to non-classical polynomials.

### 2 Preliminaries

We will always think of \( n \in \mathbb{N} \) as a growing parameter. \( \mathbb{Z}/2\mathbb{Z} \) and \( \mathbb{F}_2 \) are used interchangeably, and are identified with \( \{0, 1\} \) in the natural sense. \( \oplus \) denotes addition modulo 2.

For \( x \in \{0, 1\}^n \), \(|x|\) denotes the Hamming weight of \( x \), and for \( i \geq 0 \), \(|x|_i\) is the \((i + 1)\)th least significant bit of \(|x|\) in base 2. For \( d \in \mathbb{N} \), we use \( \{0, 1\}^n_{\leq d} \) (resp. \( \{0, 1\}^n_{= d} \)) to denote the set of elements in \( \{0, 1\}^n \) of Hamming weight at most \( d \) (resp. exactly \( d \)).

We use \( \mathcal{F}_n \) to denote the collection of all Boolean functions defined on \( \{0, 1\}^n \).

#### 2.1 Elementary symmetric polynomials

Recall that for \( t \geq 1 \), the elementary symmetric polynomial of degree \( t \), \( S_t(x_1, \ldots, x_n) \), is defined as

\[
S_t(x_1, \ldots, x_n) = \oplus_{1 \leq a_1 < \ldots < a_t \leq n} x_{a_1} \ldots x_{a_t}.
\]

This may be interpreted as

\[
S_t(x_1, \ldots, x_n) = \left( \frac{|x|}{t} \right) \mod 2. \tag{2.1}
\]

A direct consequence of Lucas’s theorem (see, e.g., [12, Section 1.2.6, Ex. 10]) and (2.1) is the following:

**Fact 2.1.** For every \( \ell \geq 0 \), \( S_{2\ell}(x) = |x|_\ell \). More generally, \( S_t(x) = \prod_i |x|_i \) where the product runs over all \( i \geq 0 \) such that the \((i + 1)\)th least significant bit of the binary expansion of \( t \) is 1.

comparison to \( d \), any \( F \) choose must be a counterexample to the Inverse conjecture for Gowers Norms over finite fields. It was established in [15, 9] that for constant \( \ell \), \( S_{2\ell} \) is indeed such a polynomial and this is why we also consider \( S_{2\ell} \).
The following result follows from the work of Green and Tao [9, Theorem 11.3], who build upon the ideas of Alon and Beigel [2].

**Theorem 2.2 (Alon-Beigel [2])**. Fix \( \ell \geq 0 \). Then, for every multilinear polynomial \( P \in \mathbb{F}_2[x_1, \ldots, x_n] \) of degree at most \( 2^\ell - 1 \), we have
\[
\Pr_x[S_2^\ell(x) = P(x)] \leq \frac{1}{2} + o(1)
\]

**Theorem 2.2** has a nice corollary:

**Corollary 2.3**. For every fixed \( \ell \geq 0 \), the functions \( \{S_2^i(x)\}_{0 \leq i \leq \ell} \) are almost balanced and almost uncorrelated, i.e.
\begin{itemize}
  \item \( \forall 0 \leq i \leq \ell, |\Pr[S_2^i(x) = 0] - \Pr[S_2^i(x) = 1]| = o(1) \)
  \item \( \forall a_0, \ldots, a_\ell \in \{0, 1\}, |\Pr[\bigwedge_{0 \leq i \leq \ell} S_2^i(x) = a_i] - \frac{1}{2^\ell} | = o(1) \).
\end{itemize}

Combining **Corollary 2.3** with **Fact 2.1**, we get another useful fact:

**Fact 2.4**. Let \( x \) be uniformly distributed over \( \{0, 1\}^n \). Then, for every fixed \( r \geq 1 \), the random variables \( \{|x|_i\}_{0 \leq i \leq r-1} \) are almost uniform and almost \( r \)-wise independent i.e.
\begin{itemize}
  \item \( \forall 0 \leq i \leq r-1, |\Pr[|x|_i = 0] - \Pr[|x|_i = 1]| = o(1). \)
  \item \( \forall (a_0, \ldots, a_{r-1}) \in \{0, 1\}^r, |\Pr[(|x|_0, \ldots, |x|_{r-1}) = (a_0, \ldots, a_{r-1})] - \frac{1}{2^r} | = o(1). \)
\end{itemize}

### 2.2 Boolean functions and polynomials over \( \mathbb{Z}/p^k\mathbb{Z} \)

Let \( p \) be a prime and \( k \geq 1 \). Given an \( F \in \mathcal{F}_n \), we define the \((p^k, p^{k-1})\)-lift of \( F \) to be the function \( F^{(p)}_k : \{0, 1\}^n \rightarrow \mathbb{Z}/p^k\mathbb{Z} \) defined as follows. For any \( x \in \{0, 1\}^n \),
\[
F^{(p)}_k(x) = \begin{cases} 
0 & \text{if } F(x) = 0, \\
p^{k-1} & \text{otherwise.}
\end{cases}
\]

We call \( p^{k-1} \) the accepting element of \( F^{(p)}_k \).

We use the notation \( \mathcal{P}^{(p)}_d \) to denote the set of all multilinear polynomials \( Q \in \mathbb{Z}/p^k\mathbb{Z}[x_1, \ldots, x_n] \). Further, for \( d \in \mathbb{N} \), \( \mathcal{P}^{(p)}_{d,k} \) will denote the set of multilinear polynomials of degree at most \( d \).

For functions \( F, G : D \rightarrow R \) for some domain \( D \) and range \( R \), define the **agreement of \( F \) and \( G \)** — denoted \( \text{agr}(F, G) \) — to be the fraction of inputs where they agree: i.e.,
\[
\text{agr}(F, G) = \Pr_{x \sim D} [F(x) = G(x)].
\]

Let \( \Delta(F, G) = 1 - \text{agr}(F, G) \).

We will consider how well multilinear polynomials of a certain degree can approximate Boolean functions in the above sense. More precisely, for any Boolean function \( F \in \mathcal{F}_n \), we define
\[
\gamma^{(p)}_{d,k}(F) = \max_{Q \in \mathcal{P}^{(p)}_{d,k}} \text{agr}(F^{(p)}_k, Q).
\]
**Remark 2.5.** Note that it also makes sense to consider the \((p^k,a)\)-lift of a function, where the accepting element is a non-zero \(a \in \mathbb{Z}/p^k\mathbb{Z}\). However, if we are interested in maximizing agreement with degree-\(d\) polynomials, then it is best to consider \(a = p^{k-1}\). The reason for this is that every non-zero \(a \in \mathbb{Z}/p^k\mathbb{Z}\) divides \(p^{k-1}\): \(ab = p^{k-1}\) for some \(b \in \mathbb{Z}/p^k\mathbb{Z}\). Hence, if a polynomial \(P\) has agreement \(\gamma\) with the \((p^k,a)\)-lift \(F'\) of \(F\), then the polynomial \(b \cdot F\) also has agreement at least \(\gamma\) with \(b \cdot F' = F_k^{(p)}\).

Usually the prime \(p\) will be clear from context, in which case we omit \(p\) and write \(F_k\) for \(F_k^{(p)}\), \(P_{d,k}\) for \(P_{d,k}^{(p)}\) and \(\gamma_{d,k}(F)\) for \(\gamma_{d,k}^{(p)}(F)\).

Following [8], we call a set \(I \subseteq \{0,1\}^n\) an interpolating set for \(P_{d,k}^{(p)}\) if the only polynomial \(P \in P_{d,k}^{(p)}\) that vanishes at all points in \(I\) is zero everywhere. Formally, for any \(P \in P_{d,k}^{(p)}\),

\[
(\forall x \in I \ P(x) = 0) \Rightarrow (\forall y \in \{0,1\}^n \ P(y) = 0)
\]

A set \(U \subseteq \{0,1\}^n\) is said to be an unconstrained set for \(P_{d,k}^{(p)}\) if given any function \(f : U \to \mathbb{Z}/p^k\mathbb{Z}\), there is a polynomial \(P \in P_{d,k}^{(p)}\) such that \(P|_U = f\). (See also [19, Section 2.4] for a similar definition.)

We state now a number of standard facts regarding Boolean functions and polynomials over \(\mathbb{Z}/p^k\mathbb{Z}\). The omitted proofs are either easy or well-known.

Unless mentioned otherwise, let \(n,m,d,k\) be any integers satisfying \(n \geq 1, d \geq 0, k \geq 1\) and \(p\) a prime.

**Fact 2.6.** Any polynomial \(Q \in P_{d,k}\) satisfies the following:

1. (Schwartz-Zippel) If \(Q\) is non-zero, then \(\Pr_{x \in \{0,1\}^n}[Q(x) \neq 0] \geq \frac{1}{2^d}\).
2. \(Q\) is the zero polynomial iff \(Q(x) = 0\) for all \(x \in \{0,1\}^n\).
3. (Möbius Inversion) Say \(Q(x) = \sum_{|S| \leq d} c_S x_S\), where \(c_S \in \mathbb{Z}/p^k\mathbb{Z}\) and \(x_S\) denotes \(\prod_{i \in S} x_i\). Then, \(c_S = \sum_{T \subseteq S} (-1)^{|S| - |T|} Q(1_T)\) where \(1_T \in \{0,1\}^n\) is the characteristic vector of \(T\).
4. (\(\{0,1\}^n_{\leq d}\) is an interpolating set) \(Q\) vanishes at all points in \(\{0,1\}^n\) iff \(Q\) vanishes at all points of \(\{0,1\}^n_{\leq d}\). By shifting the origin to any point of \(\{0,1\}^n\), the same is true of any Hamming ball of radius \(D\) in \(\{0,1\}^n\).
5. (\(\{0,1\}^n_{\leq d}\) is an unconstrained set) Given any \(f : \{0,1\}^n_{\leq d} \to \mathbb{Z}/p^k\mathbb{Z}\), there is a \(Q \in P_{d,k}\) such that \(Q|_{\{0,1\}^n_{\leq d}} = f\). By shifting the origin to any point of \(\{0,1\}^n\), the same is true of any Hamming ball of radius \(D\) in \(\{0,1\}^n\).

Point 2 follows from point 1, and points 4 and 5 from point 3.

**Fact 2.7.** Fix any \(F \in F_n\).

1. \(\gamma_{d,k}(F) \geq \frac{1}{2}\).
2. \(\gamma_{d,k+1}(F) \geq \gamma_{d,k}(F)\).
3. \(\gamma_{d,k}(F) > 1 - \frac{1}{2^d} \Rightarrow \gamma_{d,k}(F) = \gamma_{d,1}(F)\).
4. $\gamma_{1,k}(F) = \gamma_{1,1}(F)$.

Proof: 1 is trivial since there is a constant polynomial that has agreement at least $\frac{1}{2}$ with $F_k$.

Point 2: Say $P \in \mathcal{P}_{d,k}$ has agreement $\alpha$ with $F_k$. Then, $p \cdot P$ (interpreted naturally as a polynomial in $\mathcal{P}_{d,k+1}$) has agreement $\alpha$ with $F_{k+1}$.

For point 3, consider a polynomial $Q \in \mathcal{P}_{d,k}$ that achieves the maximum agreement $\alpha > 1 - \frac{1}{2^k}$ with $F_k$. Let $Q' \in \mathcal{P}_{d,1}$ be the polynomial obtained from $Q$ by dropping all its co-efficients modulo $p^{k-1}$. Note that for any $x$, $Q(x) \in \{0, p^{k-1}\}$ implies that $Q'(x) = 0$ (in the ring $\mathbb{Z}/p^{k-1}\mathbb{Z}$). Hence, the probability that $Q'$ is zero is at least $\alpha > 1 - \frac{1}{2^k}$. Fact 2.6 point 1 implies that $Q'$ must be the zero polynomial. Equivalently, all of the coefficients of $Q$ are divisible by $p^{k-1}$ and hence $Q$ can be naturally identified with $p^{k-1} \cdot Q''$ for some $Q'' \in \mathbb{Z}/p\mathbb{Z}[x_1, \ldots, x_n]$. It is easy to check that $\text{agr}(Q'', F_1) = \alpha$ and hence we have $\gamma_{d,1}(F) \geq \gamma_{d,k}(F)$. On the other hand, from point 2, we already know that $\gamma_{d,1}(F) \leq \gamma_{d,1}(F)$. Hence we are done.

Point 4 follows from points 1 and 3. \hfill \square

3 Upper bounds for $\gamma_{d,k}(\text{Maj}_n)$

In this section, we show a strong upper bound on $\gamma_{d,k}(\text{Maj}_n)$ where $\text{Maj}_n$ denotes the Majority function on $n$ bits.

Theorem 3.1. For any $k \geq 1$, $d \in \mathbb{Z}^+$ and prime $p$, $\gamma_{d,k}^{(p)}(\text{Maj}_n) \leq \frac{1}{2} + \frac{10d}{\sqrt{n}}$.

Remark 3.2. As has been noted in the literature (see, e.g., [5]), the above theorem is close to tight for $k = 1$ and $p = 2$, since there are polynomials of degree $d$ from $\mathbb{F}_2[x_1, \ldots, x_n]$ that compute $\text{Maj}_n$ correctly on $\frac{1}{2} + \Omega(\frac{d}{\sqrt{n}})$ fraction of inputs. Thus, we see that going modulo composites does not offer us any significant advantage in approximating the Majority function when $p = 2$. For larger $p$, this theorem is tight in the sense of weak representations (see Remark 3.5).

For the rest of this section, we fix the prime $p$. We will need some definitions and facts about $\mathcal{P}_{D,k}$.

We use $\pi$ to denote the unique ring homomorphism from $\mathbb{Z}/p^k\mathbb{Z}$ to $\mathbb{Z}/p\mathbb{Z}$. Its kernel $\pi^{-1}(0) = \{a \in \mathbb{Z}/p^k\mathbb{Z} \mid p^k a = 0\}$ is the set of non-invertible elements in $\mathbb{Z}/p^k\mathbb{Z}$.

We call a set $F \subseteq \{0,1\}^n$ forcing for $\mathcal{P}_{D,k}$ if any polynomial $P \in \mathcal{P}_{D,k}$ that vanishes over $F$ is forced to take a value in $\pi^{-1}(0)$ at all points $x \in \{0,1\}^n$. Formally,

$$ (\forall x \in F \ P(x) = 0) \Rightarrow (\forall y \in \{0,1\}^n \ \pi(P(y)) = 0). $$

Define the polynomial $\pi(P) \in \mathbb{Z}/p\mathbb{Z}[x_1, \ldots, x_n]$ to be the polynomial obtained by applying the map $\pi$ to each of the coefficients of $P$. Since a multilinear polynomial in $\mathbb{Z}/p^k\mathbb{Z}[x_1, \ldots, x_n]$ is the zero polynomial iff it vanishes at all points of $\{0, 1\}^n$ (by Fact 2.6), we see that $F$ is forcing iff

$$ (\forall x \in F \ P(x) = 0) \Rightarrow (\pi(P) = 0). $$

Note that any interpolating set for $\mathcal{P}_{D,k}$ (see Section 2 for the definition) is forcing for $\mathcal{P}_{D,k}$, but the converse need not be true, as the following example shows.

---

6We define the majority function as $\text{Maj}_n(x) = 1$ iff $|x| > n/2$. 

9
Example 3.3. Let \( p = 2 \) in this example. Take \( n = 3, D = 1, \) and \( k = 2. \) Consider the set \( F \) consisting of all inputs of Hamming weight exactly 2 and all 0s input. We first argue that \( F \) is not an interpolating set for \( \mathcal{P}_{1,2}. \) Consider the polynomial \( P(x) = 2(x_1 + x_2 + x_3) \in \mathbb{Z}/4\mathbb{Z}[x_1, x_2, x_3]: \) it is easy to see that \( P \) vanishes at all points in \( F \) but \( P(1, 1, 1) = 2 \) and hence \( F \) is not interpolating. However, we can also argue that \( F \) is forcing for \( \mathcal{P}_{1,2}. \) Note that for \( a \in \mathbb{Z}/4\mathbb{Z}, \) \( \pi(a) = 0 \) iff \( 2a = 0. \) Consider an arbitrary \( P \in \mathcal{P}_{1,3}. \) We can write \( P = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3. \) Since \( P \) vanishes over \( F, \) we have the following:

\[
\begin{align*}
a_0 &= 0 \\
a_0 + a_1 + a_2 &= 0 \\
a_0 + a_1 + a_3 &= 0 \\
a_0 + a_2 + a_3 &= 0.
\end{align*}
\]

Adding 3 times the first equation with the remaining tells us that \( 2(a_0 + a_1 + a_2 + a_3) = 2P(1, 1, 1) = 0. \) This implies that \( \pi_2(P(1, 1, 1)) = 0. \) From here, we can easily derive that each of \( 2P(0, 0, 1), 2P(0, 1, 0), \) and \( 2P(1, 0, 0) \) are 0 as well. By way of example, we show the case of \( P(0, 0, 1). \)

\[
2P(0, 0, 1) = 2(a_0 + a_3) \\
= 2(a_0 + a_1 + a_2 + a_3) + 2(a_0 + a_1 + a_2) + 2a_0 \\
= 2P(1, 1, 1) + 2P(1, 1, 0) + 2P(0, 0, 0) = 0.
\]

Hence, \( F \) is a forcing set.

The following lemma is the main technical step in the proof of Theorem 3.1.

Lemma 3.4. If \( F \) is forcing for \( \mathcal{P}_{D,k}, \) then \( |F| \geq |\{0, 1\}_D^n| = \binom{n}{\leq D}. \)

Note that the lower bound obtained in Lemma 3.4 for the size of forcing sets is tight, since by Fact 2.6, \( \{0, 1\}_D^n \) is forcing (and in fact interpolating) for \( \mathcal{P}_{D,k}. \)

We first use Lemma 3.4 to prove Theorem 3.1.

Proof of Theorem 3.1. We assume throughout that \( 1 \leq d \leq \frac{\sqrt{n}}{10}, \) otherwise, there is nothing to prove. Let \( \text{Maj}_{n,k} : \{0, 1\}^n \rightarrow \mathbb{Z}/p^k\mathbb{Z} \) be the \((p^k, p^{k-1})\)-lift of the \( \text{Maj}_n \) function. Let \( P \in \mathcal{P}_{d,k} \) be arbitrary and let \( S_P = \{ x \in \{0, 1\}^n \mid P(x) = \text{Maj}_{n,k}(x) \}. \) We want to show that \( |S_P| \leq 2^n \cdot \left( \frac{1}{2} + \frac{10d}{\sqrt{n}} \right). \) We will argue by contradiction. So assume that \( |S_P| > 2^n \cdot \left( \frac{1}{2} + \frac{10d}{\sqrt{n}} \right). \)

Let \( E_P \) be the complement of \( S_P, \) i.e. the set of points where \( P \) makes an error in computing \( \text{Maj}_{n,k}. \) We have \( |E_P| < 2^n \left( \frac{1}{2} - \frac{10d}{\sqrt{n}} \right). \) We will try to find a degree \( D \) (for suitable \( D \leq \lfloor n/2 \rfloor \)) polynomial \( Q \) such that \( Q \) vanishes at all points in \( E_P \) but has the property that \( Q(x) \) is a unit (i.e. \( \pi(Q(x)) \neq 0 \)) for some \( x \in \{0, 1\}^n. \) To be able to do this, we need the fact that \( E_P \) is not forcing for \( \mathcal{P}_{D,k}. \) By Lemma 3.4, if \( E_P \) is indeed forcing for \( \mathcal{P}_{D,k}, \) then
Remark 3.5. In the proof of the above theorem, it is not hard to see that the function $|E_p| \geq \sum_{i=0}^{D} \binom{n}{i} = \left( \sum_{i=0}^{[n/2]} \binom{n}{i} \right) - \sum_{i=D+1}^{[n/2]} \binom{n}{i}$
\[ \geq 2^{n-1} - \left( \frac{n}{2} - D \right) \cdot \binom{n}{[n/2]} \]
\[ \geq 2^n \left( \frac{1}{2} - \frac{2([n/2] - D)}{\sqrt{n}} \right) = 2^n \left( \frac{1}{2} - \frac{4d}{\sqrt{n}} \right) \]

where the last equality follows if we choose $D = \lfloor n/2 \rfloor - 2d$. This contradicts our upper bound on the size of $|E_p|$. Hence, $E_p$ cannot be forcing for $\mathcal{P}_{d,k}$. In particular, we can find $Q$ that vanishes on $E_p$ and furthermore, $\pi(Q(x)) \neq 0$ for some $x \in \{0,1\}^n$.

We now claim that $\pi(Q(x_0)) \neq 0$ for some $x_0$ of Hamming weight $> n/2$. To see this, consider the polynomial $Q_1 = \pi(Q)$. By construction of $Q$, we know that $Q_1$ is a non-zero polynomial of degree $D$. Hence, by Fact 2.6, $Q_1$ is non-zero when restricted to the Hamming ball of radius $D < n/2$ around the all 1s vector. In particular, this implies that there is an input $x_0$ of Hamming weight $> n/2$ where $Q_1(x_0)$ is non-zero and hence $\pi(Q(x_0)) \neq 0$, or equivalently $p^{k-1}Q(x_0) \neq 0$. Fix this $x_0$ for the remainder of the proof. Note that $x_0 \notin E_p$ since $Q$ vanishes on $E_p$.

Now, consider the polynomial $R(x) = Q(x) \cdot P(x)$. We first show that $R(x) = 0$ for all $x$ of Hamming weight $\leq n/2$. Consider any $x$ of Hamming weight $\leq n/2$. If $x \in E_p$, then $R(x) = 0$ since $Q(x) = 0$. On the other hand, if $x \notin E_p$, then $P(x) = \text{Maj}_{n,k}(x) = 0$ since $x$ has Hamming weight $\leq n/2$. Thus, $R$ vanishes at all inputs of Hamming weight $\leq n/2$.

Since the degree of $R$ is at most $\deg(Q) + \deg(P) = D + d = ([n/2] - 2d) + d \leq [n/2] - d$ and $R$ vanishes at all inputs of $\{0,1\}^{n/2}$, this implies (by Fact 2.6) that $R$ must be 0 everywhere. However, at $x_0$, $R(x_0) = Q(x_0)P(x_0) = Q(x_0)\text{Maj}_{n,k}(x_0) = p^{k-1}Q(x_0) \neq 0$. This yields the desired contradiction. \hfill \qed

Remark 3.5. In the proof of the above theorem, it is not hard to see that the function $\text{Maj}_{n,k}(x)$ can be replaced by any function $F : \{0,1\}^n \rightarrow \mathbb{Z}/p^k\mathbb{Z}$ that is 0 at all points of Hamming weight $\leq n/2$ and non-zero at all points of Hamming weight $> n/2$. Thus, the above theorem actually shows something stronger: recall [3] that a polynomial $P$ weakly represents a Boolean function $F$ if $P(x) = 0 \iff F(x) = 0$. Then any polynomial $P \in \mathcal{P}_{d,k}$ can weakly represent the Majority function on at most a $\frac{1}{2} + \frac{10d}{\sqrt{n}}$ fraction of all inputs. Currently, we are unable to prove this stronger theorem for all composites (not just prime powers), since the proof of Lemma 3.4 crucially uses the fact that we are working modulo a prime power. We leave the problem of extending this result to general composites as an open question.

Theorem 3.1 also easily implies an analog for arbitrary composites. We refer the reader to Appendix B in the appendix for details.

We now prove Lemma 3.4.

3.1 Proof of Lemma 3.4

We need the following fact about rings of the form $\mathbb{Z}/m\mathbb{Z}$. Since we are unable to find a reference that gives us this statement exactly, we sketch how to derive it from standard results in Appendix A.
Fact 3.6. Say \( R = \mathbb{Z}/m\mathbb{Z} \). Let \( A z = b \) be a system of inhomogeneous linear equations over \( R \) with \( A \in R^{M \times N} \) and \( b \in R^M \). Then \( A z = b \) does not have a solution over \( R^N \) if and only if there is a \( y \in R^M \) such that \( y^T A = 0 \) and \( y^T b \neq 0 \).

This leads to the an interesting statement for polynomials. Sat that \( F \subseteq \{0, 1\}^n \) is forcing at \( x \in \{0, 1\}^n \) if for any multilinear \( Q \in \mathcal{P}_{D,K} \),

\[
(\forall u \in F \ (Q(u) = 0) \Rightarrow \pi(Q(x)) = 0
\]

Corollary 3.7. Fix any \( F \subseteq \{0, 1\}^n, x \in \{0, 1\}^n \) and \( D \in \mathbb{N} \). Then, \( F \) is forcing at \( x \) for \( \mathcal{P}_{D,K} \) if and only if there are \( a^n_u \in \mathbb{Z}/p^k\mathbb{Z} \) (\( u \in F \)) such that for any \( Q \in \mathcal{P}_{D,K} \),

\[
p^{k-1}Q(x) = \sum_{u \in F} a^n_u Q(u).
\]

Proof. Consider the problem of finding a \( Q \in \mathcal{P}_{D,K} \) such that \( Q(x) = 1 \) but \( Q(u) = 0 \) for all \( u \in F \). Let \( R := \mathbb{Z}/p^k\mathbb{Z} \). Each of the constraints \( Q(x) = 1 \) and \( Q(u) = 0 \) (\( u \in F \)) is an \( R \)-linear constraint over variables that correspond to co-efficients of \( Q \). More precisely, the existence of such a \( Q \) is equivalent to the existence of a solution to the system of equations \( A z = b \) where \( A \) and \( b \) are defined as follows.

The matrix \( A \) lies in \( R^{M \times N} \) where \( M = |F| + 1 \) and \( N = \binom{n}{N} \). The first row of \( A \) is labelled by \( x \) and the remaining rows by \( u \in F \). The columns of \( A \) are labelled by \( S \subseteq [n] \) of size at most \( D \). The entry corresponding to \( (v, S) \) is \( \prod_{i \in S} v_i \).

The vector \( b \in R^n \) is just the vector that has a 1 in the first entry and is 0 elsewhere.

Since \( F \) is forcing at \( x \) for \( \mathcal{P}_{D,K} \), the system of equations \( A z = b \) has no solution. By Fact 3.6, we then have a \( y \) such that \( y^T A = 0 \) but \( y^T b \neq 0 \). Note that the latter condition is just the condition that \( y_1 \), the first entry of \( y \), is non-zero. Since every non-zero element in \( R \) divides \( p^{k-1} \), we can assume that \( y_1 = p^{k-1} \) by multiplying it with a scalar \( \alpha \in R \) if necessary.

Now, \( y^T A = 0 \) can be rephrased as

\[
y_1 \prod_{i \in S} x_i = - \sum_{u \in F} y_u \prod_{i \in S} u_i
\]

for each \( S \subseteq [n] \) of size at most \( D \). Let \( a^n_u = -y_u \) for each \( u \in F \). Taking suitable linear combinations of the above equations, we get the claimed constraint for each \( Q \in \mathcal{P}_{D,K} \).

We also need the following easy consequence of Fact 2.6.

Lemma 3.8. Fix any \( D, k \geq 1 \). A set \( F \subseteq \{0, 1\}^n \) is forcing for \( \mathcal{P}_{D,K} \) if and only if \( F \) is forcing at each \( x \in \{0, 1\}^n \).

Proof. One of the directions of the lemma is trivial: if \( F \) is forcing for \( \mathcal{P}_{D,K} \), then \( F \) is forcing at every \( x \in \{0, 1\}^n \) and hence in particular at \( x \in \{0, 1\}^n_{\leq D} \).

For the converse, consider an \( F \) that is forcing at each \( x \in \{0, 1\}^n \). Let \( Q \in \mathcal{P}_{D,K} \) be any polynomial that vanishes over \( F \). Then, by definition \( \pi(Q(x)) = 0 \) for each \( x \in \{0, 1\}^n_{\leq D} \). This implies that \( Q_1 = \pi(Q) \) is a degree \( D \) polynomial in \( \mathcal{P}_{D,1} \) that vanishes at all \( x \in \{0, 1\}^n_{\leq D} \). By Fact 2.6, this implies that \( Q_1 \) vanishes at all \( x \in \{0, 1\}^n \). Thus, \( \pi(Q(x)) = 0 \) for all \( x \in \{0, 1\}^n \). Since \( Q \in \mathcal{P}_{D,K} \) was arbitrary, we have shown that \( F \) is forcing for \( \mathcal{P}_{D,K} \).
Proof of Lemma 3.4. We want to show that any $F \subseteq \{0,1\}^n$ such that $|F| < (\leq_D^n)$ cannot be forcing for $P_{D,k}$. For the sake of brevity, we will simply use “forcing” to denote “forcing for $P_{D,k}$”.

Let us start with the very special case when $F \subseteq \{0,1\}^n \leq_D$. Since $|F| < (\leq_D^n)$, $F$ must be a strict subset of $\{0,1\}^n \leq_D$. Fix any $x_0 \in \{0,1\}^n \leq_D \setminus F$ and consider the function $f : F \cup \{x_0\} \to \mathbb{Z}/p^k\mathbb{Z}$ defined by:

$$f(x) = \begin{cases} 0 & \text{if } x \in F, \\ 1 & \text{if } x = x_0. \end{cases}$$

By Fact 2.6, we know that there is a polynomial $Q$ of degree at most $D$ such that $Q|_{F \cup \{x_0\}} = f$. Hence, $F$ cannot be forcing.

Now, given any $F \subseteq \{0,1\}^n$ of size less than $(\leq_D^n)$ that is forcing, we show that there is another set $F' \subseteq \{0,1\}^n \leq_D$ of size at most the size of $F$ that is also forcing. This will contradict our previous argument that showed that no such $F'$ exists, and hence the proof will be complete.

We transform $F$ to $F'$ by a simple shifting argument. If $F \subseteq \{0,1\}^n \leq_D$ already, then there is nothing to do. Otherwise, the set $F_{>D} := F \setminus \{0,1\}^n \leq_D$ is non-empty. We show how to produce an $F''$ of size at most the size of $F$ such that $F''$ is also forcing and further, the set $F''_{>D}$ is a strict subset of $F_{>D}$. By iterating this argument, we will be able to produce the desired set $F'$.

By Corollary 3.7, we know that for each $x \in \{0,1\}^n$, we have $\alpha^x_u \in \mathbb{Z}/p^k\mathbb{Z}$ ($u \in F$) such that for any $Q \in P_{D,k}$

$$p^{k-1}Q(x) = \sum_{u \in F} \alpha^x_u Q(u) = \sum_{u \in F_{\leq_D}} \alpha^x_u Q(u) + \sum_{u \in F_{>D}} \alpha^x_u Q(u)$$

where $F_{\leq_D} := F \setminus F_{>D}$. Let $A = \{\alpha^x_u | x \in \{0,1\}^n_{\leq_D}, u \in F_{>D}\}$.

Any $\alpha \in \mathbb{Z}/p^k\mathbb{Z}$ can be written as $\alpha = p^\ell \cdot r$ for some $\ell \leq k$ and $r \in \mathbb{Z}/p^k\mathbb{Z}$ is invertible; here $\ell$ is unique and we call it the level of $\alpha$, denoting it by $\ell(\alpha)$. Note that for $\alpha, \beta \in \mathbb{Z}/p^k\mathbb{Z}$, $\alpha \mid \beta$ iff $\ell(\alpha) \leq \ell(\beta)$.

Among all the elements of $A$, we choose the element of the lowest level. Say that this element is $\alpha = \alpha^y_v \in \{0,1\}^n_{\leq_D}, v \in F_{>D}$ and let $\ell = \ell(\alpha)$. We take $F'' = F \setminus \{v\}$ and $F''_{>D} \subsetneq F_{>D}$. It is clear that $F''$ has size at most the size of $F$ and that $F''_{>D} \subsetneq F_{>D}$. We only need to show that $F''$ continues to be a forcing set.

By Corollary 3.7 and Lemma 3.8, it suffices to show that, for each $y \in \{0,1\}^n_{\leq_D}$, we have $\beta^y_u (u \in F'')$ such that for any $Q \in P_{D,k}$

$$p^{k-1}Q(y) = \sum_{u \in F} \beta^y_u Q(u). \quad (3.1)$$

Note that we already have

$$p^{k-1}Q(y) = \sum_{u \in F} \alpha^y_u Q(u) = \sum_{u \in F \setminus \{v\}} \alpha^y_u Q(u) + \alpha^y_v Q(v) \quad (3.2)$$

for any $Q \in P_{D,k}$. Since $\alpha^y_v \in A$, we know that $\ell(\alpha^y_v) \geq \ell$. Thus, $\alpha^y_v = a_0 \alpha^x_v$ for some $a_0 \in \mathbb{Z}/p^k\mathbb{Z}$.
On the other hand, we also have
\[ p^{k-1}Q(x) = \sum_{u \in F \setminus \{v\}} \alpha_u^x Q(u) + \alpha_v^x Q(v) \]
\[ \Rightarrow \alpha_v^x Q(v) = p^{k-1}Q(x) - \sum_{u \in F \setminus \{v\}} \alpha_u^x Q(u) \]
\[ \Rightarrow \alpha_v^x Q(v) = \alpha_0 p^{k-1}Q(x) - \sum_{u \in F \setminus \{v\}} \alpha_0 \alpha_u^x Q(u) \]

Substituting in (3.2), we obtain
\[ p^{k-1}Q(y) = \sum_{u \in F \setminus \{v\}} \alpha_u^y Q(u) + \alpha_0 p^{k-1}Q(x) - \sum_{u \in F \setminus \{v\}} \alpha_0 \alpha_u^x Q(u) \]
\[ = \left( \sum_{u \in F \setminus \{v\}} (\alpha_u^y - \alpha_0 \alpha_u^x) Q(u) \right) + \alpha_0 p^{k-1}Q(x) \]

Thus, defining \( \beta_u^y = \alpha_u^y - \alpha_0 \alpha_u^x \) for \( u \in F \setminus \{x\} \) and \( \beta_x^y = \alpha_0 p^{k-1} \), we have (3.1). (This assumes that \( x \notin F \). If \( x \in F \), then we should actually set \( \beta_x^y = \alpha_0 p^{k-1} + \alpha_x^y - \alpha_0 \alpha_x^x \).) \( \square \)

4 Some separation results

In this section, we fix the prime \( p = 2 \). Recall that we use \( F_k \) to denote \( F_k(2) \), \( \mathcal{P}_{d,k} \) for \( \mathcal{P}_{d,k}^{(2)} \), and \( \gamma_{d,k}(F) \) for \( \gamma_{d,k}^{(2)}(F) \).

4.1 Symmetric functions as separating examples

We know from Theorem 2.2 that, for every fixed \( \ell \geq 2 \), \( \gamma_{2^{\ell-1},1}(S_{2^{\ell}}) \leq \frac{1}{2} + o(1) \). In contrast, the main result of this section shows that

**Theorem 4.1.** For every fixed \( \ell \geq 2 \), \( \gamma_{d,3}(S_{2^{\ell}}) \geq \frac{9}{16} - o(1) \), where \( d = 2^{\ell-1} + 2^{\ell-2} \).

Notice that \( 2^{\ell-1} + 2^{\ell-2} \leq 2^\ell - 1 \) for \( \ell \geq 2 \). This implies that, for \( \ell \geq 2 \), \( S_{2^{\ell}}(x) \) is an example of a function \( F \) for which there exist \( k, d \in \mathbb{N} \) such that \( \gamma_{d,1}(F) \leq \frac{1}{2} + o(1) \) but \( \gamma_{d,k}(F) \geq \frac{1}{2} + \Omega(1) \) for some \( d' \leq d \).

**Proof of Theorem 4.1.** Fact 2.1 from Section 2 tells us that
\[ S_{2^{\ell}}(x) = |x|_\ell. \]

Thus, \( S_{2^{\ell},3}(x) \), the \((8,4)\)-lift of \( S_{2^{\ell}}(x) \), is given by
\[ S_{2^{\ell},3}(x) = \begin{cases} 4 & \text{if } |x|_\ell = 1 \\ 0 & \text{otherwise} \end{cases} \] (4.1)
Fix \( d = 2^{\ell-1} + 2^{\ell-2} \) and consider the following polynomial in \( \mathbb{Z}/8\mathbb{Z}[x_1, \ldots, x_n] \)

\[
P(x) = \sum_{T \subseteq [d]} \prod_{i \in T} x_i.
\]

To prove the theorem, it suffices to show that

\[
\Pr[P(x) = S_{2^{\ell},3}(x)] \geq \frac{1}{2} + \frac{1}{16} - o(1)
\]

Clearly, \( P(x) = \left(\frac{|x|}{d}\right) \mod 8 \), and

\[
P(x) = \begin{cases} 
0 & \text{if } 8 \mid \left(\frac{|x|}{d}\right) \\
4 & \text{if } 4 \mid \left(\frac{|x|}{d}\right) \text{ but } 8 \nmid \left(\frac{|x|}{d}\right)
\end{cases}
\] (4.2)

The following theorem due to Kummer (see, e.g., [12, Section 1.2.6, Ex. 11]) determines the largest power of a prime that divides a binomial coefficient.

**Theorem 4.2 (Kummer).** Let \( p \) be a prime and \( N, M \in \mathbb{N} \) such that \( N \geq M \). Suppose \( r \) is the largest integer such that \( p^r \mid \binom{N}{M} \). Then \( r \) is equal to the number of borrows required when subtracting \( M \) from \( N \) in base \( p \).

Let \( B(x) \) be the number of borrows required when subtracting \( d \) from \( |x| \). Rewriting (4.2) in terms of \( B(x) \) using Kummer’s theorem, we get

\[
P(x) = \begin{cases} 
4 & \text{if } B(x) = 2 \\
0 & \text{if } B(x) \geq 3
\end{cases}
\] (4.3)

We will need the following lemma.

**Lemma 4.3.** \( P(x) = S_{2^{\ell},3}(x) \) if

1. \( |x|_{\ell-2} = 0 \), or
2. \( (|x|_{\ell-2}, |x|_{\ell-1}, |x|_{\ell}, |x|_{\ell+1}) = (1,0,0,0) \).

**Proof.** Since \( d = 2^{\ell-1} + 2^{\ell-2} \), all the bits of \( d \) except \( d_{\ell-1} \) and \( d_{\ell-2} \) are zero. An important observation is that, when subtracting \( d \) from \( |x| \), no borrows are required by the bits \( |x|_i \), \( 0 \leq i \leq \ell - 3 \).

Using the above observation, the reader can verify that when \( (|x|_{\ell-2}, |x|_{\ell-1}, |x|_{\ell}, |x|_{\ell+1}) = (1,0,0,0) \) the number of borrows required is at least 3 i.e. \( B(x) \geq 3 \), which in turn implies that \( P(x) = 0 \). Since \( |x|_\ell = 0 \), \( S_{2^{\ell},3}(x) = 0 \). This proves the second part of the lemma.

To prove the first part, suppose \( |x|_{\ell-2} = 0 \). Since \( d_{\ell-1} = d_{\ell-2} = 1 \), it follows that both \( |x|_{\ell-2} \) and \( |x|_{\ell-1} \) will need to borrow when subtracting \( d \) from \( |x| \). As argued before, no borrows are required by the bits before (i.e. less significant than) \( |x|_{\ell-2} \), and thus the total number of borrows required by the bits \( |x|_i \), \( 0 \leq i \leq \ell - 1 \), is two.

Note that the bit \( |x|_{\ell-1} \) borrows from \( |x|_\ell \). Consider the following case analysis:
• Case $|x|_ℓ = 1$: $|x|_ℓ$ will not need to borrow since $d_ℓ = 0$. In fact, none of the bits after (i.e. more significant than) $|x|_ℓ$ will need to borrow, and thus $B(x) = 2$. This implies that $P(x) = 4$. We also have $S_{2',3}(x) = 4$ and hence $P(x) = S_{2',3}(x)$.

• Case $|x|_ℓ = 0$: $|x|_ℓ$ will require a borrow and this means $B(x) \geq 3$. This would imply $P(x) = 0$. Since $|x|_ℓ = 0$, it follows that $P(x) = S_{2',3}(x)$.

This completes the proof. □

By Lemma 4.3, we have

$$\Pr[P(x) = S_{2',3}(x)] \geq \Pr[|x|_{ℓ−2} = 0] + \Pr[|x|_{ℓ−2}, |x|_{ℓ−1}, |x|_ℓ, |x|_{ℓ+1} = (1,0,0,0)]$$  (4.4)

Using Fact 2.4 from Section 2, we have

$$\Pr[|x|_{ℓ−2} = 0] \geq \frac{1}{2} - o(1)$$
$$\Pr[|x|_{ℓ−2}, |x|_{ℓ−1}, |x|_ℓ, |x|_{ℓ+1} = (1,0,0,0)] \geq \frac{1}{16} - o(1)$$

which, together with (4.4), implies

$$\Pr[P(x) = S_{2',3}(x)] \geq \frac{1}{2} + \frac{1}{16} - o(1).$$
□

4.2 A separation at $k = 2$

In this section, we show that there are functions $F$ for which $γ_{2,2}(F) > γ_{2,1}(F)$.

**Theorem 4.4.** For any odd $n$, there exists a function $F \in F_{2n}$ such that $γ_{2,2}(F) \geq \frac{3}{4}$ but $γ_{2,1}(F) \leq \frac{5}{8} + o(1)$.

This result is interesting because it shows that there is a separation at the first place where it is possible to have one (Recall that $γ_{1,k}(F) = γ_{1,1}(F)$ for any $F \in F_n$ from Fact 2.7).

We start with some basic preliminaries about $F_2$-quadratics, i.e. polynomials from $P_{2,1}$. Given $Q(x_1, \ldots, x_n) \in P_{2,1}$, we define the rank of $Q$ to be the least $r$ such that we can write $Q$ in the form

$$Q(x) = \bigoplus_{i=1}^r ℓ_iℓ'_i \oplus L$$  (4.5)

where $ℓ_i, ℓ'_i (i \in [r])$ and $L$ are linear functions in $x_1, \ldots, x_n$. We use $\text{rk}(Q)$ to denote the rank of $Q$.

The following are standard facts about ranks of quadratic polynomials (see, e.g., [13, Chapter 6]).

**Fact 4.5.** 1. The rank is sub-additive: $\text{rk}(Q_1 \oplus Q_2) \leq \text{rk}(Q_1) + \text{rk}(Q_2)$.

2. If the $ℓ_i, ℓ'_i (i \in [r])$ in (4.5) form a set of $2r$ linearly independent polynomials, then $\text{rk}(Q) = r$. 

16
3. If $Q$ has rank $k$, then

$$\Pr_x[Q(x) \neq 0] \geq \frac{1}{2} - \frac{1}{2^{k+1}}$$

Let us begin the proof of Theorem 4.4. We first define a family of Boolean functions on $\{0, 1\}^{2n}$, where $n \equiv 1 \pmod{2}$. We denote the $2n$ variables by $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. We use $|x|$ to denote the Hamming weight of $x$.

Let $S = \{(x, y) \mid |x|, |y| \equiv 1 \pmod{2}\}$. Given any function $H : S \to \{0, 1\}$, we define Boolean function $F_H(x_1, \ldots, x_n, y_1, \ldots, y_n)$ as follows:

$$F_H(x, y) = \begin{cases} 0 & \text{if } |x| \cdot |y| \equiv 0 \pmod{4}, \\ 1 & \text{if } |x| \cdot |y| \equiv 2 \pmod{4}, \\ H(x, y) & \text{otherwise.} \end{cases}$$

First of all, let us note that for any choice of $H$, we have:

**Lemma 4.6.** $\gamma_{2,2}(F_H) \geq \frac{3}{4}$.

**Proof.** Consider the polynomial $P(x, y) = (\sum_i x_i) \cdot (\sum_j y_j) \in \mathcal{P}_{2,2}$. It can be checked that if $|x| \cdot |y| \equiv 0 \pmod{2}$, then $P(x, y) = F_{H,2}(x, y)$ (recall that $F_{H,2}$ is the $(4, 2)$-lift of $F_H$) and hence, the probability that $P(x, y) \neq F_{H,2}(x, y)$ is the probability that $|x|$ and $|y|$ are both odd, which is $1/4$. This gives the claim. \qed

We will now argue that $\gamma_{2,1}(F_H) \leq \frac{5}{8} + o(1)$. It will be helpful to describe an alternate basis for the space of $(\mathbb{F}_2)$-linear functions over $\{0, 1\}^n$. Consider the linear functions

$$u_1(x) = \bigoplus_i x_i$$

$$u_2(x) = x_1 \oplus \bigoplus_{i \geq 3} x_i$$

$$u_3(x) = x_2 \oplus \bigoplus_{i \geq 3} x_i$$

$$u_4(x) = x_3 \oplus \bigoplus_{i \geq 5} x_i$$

$$u_5(x) = x_4 \oplus \bigoplus_{i \geq 5} x_i$$

$$\vdots$$

$$u_{n-1}(x) = x_{n-2} \oplus x_n$$

$$u_n(x) = x_{n-1} \oplus x_n$$

It is easy to check that these linear functions are linearly independent and hence form a basis for all linear functions on $\{0, 1\}^n$. In particular, any polynomial $P(x_1, \ldots, x_n)$ of degree $d$ can be alternately written as a degree $d$ polynomial $P'(u_1, \ldots, u_n)$. We similarly define linear functions $v_1, \ldots, v_n$ of the $y$ variables.

To complete the description of the function $F_H$, we show how to choose $H$.

**Lemma 4.7.** There is an $H : S \to \{0, 1\}$ such that for all restrictions $q|_S$ of $q \in \mathcal{P}_{2,1}$, we have $\Delta(H, q|_S) \geq \frac{1}{2} - o(1)$. 

17
Proof. The proof is a trivial union bound. The number of quadratic polynomials $q$ is at most $2^{O(n^2)}$. For each such $q$, the expected number of locations $x \in S$ where $H(x) \neq q(x)$ is $|S|/2$. By a Chernoff bound, the probability that this number is at most $|S|/2 - |S|^{2/3}$ is $\exp(-\Omega(|S|^{1/3})) = \exp(-2^{\Omega(n)})$. A union bound over all the possible $q$ tells us that with probability $1 - o(1)$ over the choice of $H$, every $q|_S$ differs from $H$ in at least $|S|/2 - |S|^{2/3} = |S|\left(\frac{1}{2} - o(1)\right)$ locations. This completes the proof.  

We fix any $H$ so that the above holds. This completes the description of $F_H$. We are now ready to show the main result, which implies Theorem 4.4.

**Lemma 4.8.** For any $q \in \mathcal{P}_{2,1}$, we have $\Delta(F_H, q) \geq 3/8 - o(1)$. Hence, $\gamma_{2,1}(F_H) \leq 5/8 + o(1)$.

**Proof.** We write $q$ as a quadratic polynomial with variables $u_1, \ldots, u_n, v_1, \ldots, v_n$. Hence, we have

$$q = c \oplus L(u, v) \oplus B(u, v) \oplus Q(u) \oplus R(v)$$

where $c$ is the constant term, $L(u, v)$ is the homogeneous degree 1 part, $B(u, v)$ is the bilinear\(^7\) part, and $Q(u)$ and $R(v)$ are the degree 2 polynomials in $u$ and $v$ respectively.

We can further expand $Q(u)$ as $u_1 \ell(u) + Q'(u)$ where $Q'(u)$ is the sum of all terms that don’t involve $u_1$ and $\ell(u_2, \ldots, u_n)$ is a linear function. Similarly $R(v) = v_1 \lambda(v) + R'(v)$.

For the sake of simplicity, we assume that $c = 0$ and $L(u, v) = 0$, though our proof works even when this is not the case.

For any $b_1, b_2 \in \{0, 1\}$, define $G_{b_1, b_2}(u_2, \ldots, u_n, v_2, \ldots, v_n)$ to be the polynomial that represents $F_H$ on the inputs where $u_1 = b_1$ and $u_2 = b_2$. Also, let

$$\Delta_{b_1, b_2} = \Pr_{u, v}[G_{b_1, b_2}(u, v) \oplus q(u, v) \neq 0 \mid u_1 = b_1, v_1 = b_2]$$

**Observation 4.9.** $G_{0, 0} = 0$, $G_{0, 1}$ is a rank-$\frac{n-1}{2}$ quadratic in $u_2, \ldots, u_n$, and $G_{1, 0}$ is a rank-$\frac{n-1}{2}$ quadratic in $v_2, \ldots, v_n$.

**Proof of Observation 4.9.** The claim about $G_{0, 0}$ is trivial. To prove the claim about $G_{0, 1}$ (the claim about $G_{1, 0}$ is similar), note that restricted to $u_1 = 0$ and $v_1 = 1$, we know that $|x|$ is even and $|y|$ is odd. Hence, $F_H(x, y)$ evaluates to 1 if $|x| \equiv 2 \mod 4$ and 0 if $|x| \equiv 0 \mod 4$ (note that the setting of $y$ is irrelevant). Restricted to even weight inputs $x \in \{0, 1\}^n$, we can see that $F_H(x, y) = S_2(x)$, the degree-2 elementary symmetric polynomial in the $x$ variables (see Section 2).

It is not hard to check that $S_2(x) = u_2u_3 \oplus u_4u_5 \oplus \cdots \oplus u_{n-1}u_n \oplus L'(u)$, where $L'$ is a linear function. This is a rank-$\frac{n-1}{2}$ quadratic in $u_2, \ldots, u_n$.  

Clearly, we have

$$\Delta(F_H, q) = \frac{1}{4} (\Delta_{0, 0} + \Delta_{0, 1} + \Delta_{1, 0} + \Delta_{1, 1})$$

$$= \frac{1}{4} (\Delta_{0, 0} + \Delta_{0, 1} + \Delta_{1, 0} + \Delta(H, q|_{u_1=v_1=1}))$$

$$\geq \frac{1}{4} (\Delta_{0, 0} + \Delta_{0, 1} + \Delta_{1, 0} + \frac{1}{2} - o(1))$$

(4.6)
where the inequality is a consequence of Lemma 4.7. So it suffices to prove that $\Delta_{0,0} + \Delta_{0,1} + \Delta_{1,0} \geq 1 - o(1)$. We do this now.

**Claim 4.10.** $\Delta_{0,0} + \Delta_{0,1} + \Delta_{1,0} \geq 1 - o(1)$.

**Proof of Claim 4.10.** By substituting the values of $u_1$ and $v_1$ in the expression for $q$, we obtain

- $P_{0,0} := G_{0,0} \oplus q_{|u_1=v_1=0} = 0 \oplus Q'(u) \oplus R'(v) \oplus B_{0,0}(u, v) = Q'(u) \oplus R'(v) \oplus B_{0,0}(u, v)$,
- $P_{0,1} := G_{0,1}(u) \oplus q_{|u_1=0,v_1=1} = G_{0,1}(u) \oplus Q'(u) \oplus R'(v) \oplus B_{0,1}(u, v)$,
- $P_{1,0} := G_{1,0}(v) \oplus q_{|u_1=1,v_1=0} = Q'(u) \oplus G_{1,0}(v) \oplus R'(v) \oplus B_{1,0}(u, v)$,

where $B_{0,0}, B_{0,1}$, and $B_{1,0}$ are polynomials of individual degree at most 1 in $u$ and $v$. Note that $\Delta_{b_1,b_2} = \Pr_{u,v}[P_{b_1,b_2}(u, v) \neq 0]$.

To argue about these probabilities, we use point 3 of Fact 4.5. The proof is a case analysis on the tuple $(\rk(Q'), \rk(R'))$. W.l.o.g., we assume that $\rk(Q') \leq \rk(R')$. Let $k = (n - 1)/2$.

- **Case 1 ($\rk(Q') \geq k/2$):** In particular, this implies that $\rk(R') \geq \rk(Q') \geq k/2$ and hence both $Q'$ and $R'$ are high rank quadratics. For each setting of $u$ (say) in $P_{0,0}$, we get a quadratic in $v$ of rank at least $k/2$ and hence by Fact 4.5, $\Delta_{0,0} \geq \frac{1}{2} - o(1)$. In a similar way, it is possible to show that each of $\Delta_{1,0}$ and $\Delta_{0,1}$ is also at least $\frac{1}{2} - o(1)$. Thus, we have the claim in this case.

- **Case 2 ($\rk(Q') \leq k/2$ and $\rk(R') \geq k/2$):** In this case, for each setting of $u$, the polynomials $P_{0,0}$ and $P_{0,1}$ are rank-$k/2$ quadratics in $v$ and hence, by Fact 4.5, we have $\Delta_{0,0}, \Delta_{0,1} \geq \frac{1}{2} - o(1)$. Hence, the claim follows in this case.

- **Case 3 ($\rk(R') \leq k/2$):** We also have $\rk(Q') \leq k/2$. In this case, by the sub-additivity of rank (see Fact 4.5) and Observation 4.9, the quadratic polynomials $G_{0,1} \oplus Q'$ and $G_{1,0} \oplus R'$ have rank at least $k/2$ each. Thus, the polynomials $P_{0,1}$ and $P_{1,0}$ are rank-$k/2$ quadratics for each setting of $v$ and $u$ respectively. Thus, each of $\Delta_{0,1}$ and $\Delta_{1,0}$ is at least $\frac{1}{2} - o(1)$. So we are done.

5 **Connection to non-classical polynomials**

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denote the one dimensional torus, and let $|\cdot|$ denote the natural map from $\mathbb{F}_2$ to $\{0, 1\} \subset \mathbb{Z}$.

Depending on the context, we will think of a Boolean function $F$ as $F : \mathbb{F}_2^n \to \mathbb{F}_2$ or $F : \mathbb{F}_2^n \to \{0, 1/2\} \subset \mathbb{T}$, by identifying $\mathbb{F}_2$ with $\{0, 1/2\}$.

Tao and Ziegler [21] define non-classical polynomials as follows:
Definition 5.1 ([21]). A function $F : \mathbb{F}_2^n \rightarrow \mathbb{T}$ is said to be a non-classical polynomial of degree $\leq d$ if it has the following form:

$$F(x_1, \ldots, x_n) = \alpha + \sum_{0 \leq e_1, \ldots, e_n \leq 1, \sum e_i + (k-1) \leq d} \frac{c_{e_1, \ldots, e_n, k} \prod |x_i|^{e_i} \mod 1}{2^k}$$

Here $\alpha \in \mathbb{T}$, and $c_{e_1, \ldots, e_n, k} \in \{0, 1\}$. $\alpha$ is called the shift of $F$, and the largest $k$ such that $c_{e_1, \ldots, e_n, k} \neq 0$ is called the depth of $F$.

Since we are interested in the agreement of a non-classical polynomial with Boolean (i.e. $\{0, 1/2\}$-valued) functions, we will only consider polynomials with shift $\alpha = \frac{A}{2^k}$, where $k$ is the depth of the polynomial and $A \in \{0, \ldots , 2^k - 1\}$.

Remark 5.2. Classical polynomials are non-classical polynomials with $\alpha \in \{0, 1/2\}$ and depth $= 1$.

The following lemma relates our model to non-classical polynomials:

Lemma 5.3. Let $F$ be a Boolean function, and $d, k \in \mathbb{Z}^+, d \geq k$.

1. If there is a non-classical polynomial $P$ of degree $d$ and depth $k$ satisfying $	ext{agr}(F, P) = \gamma$, then there is a $P' \in \mathcal{P}(2, d)_{d,k}$ satisfying $	ext{agr}(F_k, P') = \gamma$, where $F_k$ is the $(2^k, 2^k - 1)$-lift of $F$.

2. If there is a $P \in \mathcal{P}(2, d)_{d,k}$ satisfying $	ext{agr}(F_k, P) = \gamma$, then there is a non-classical polynomial $P'$ of degree $\leq d + k - 1$ satisfying $	ext{agr}(F, P') = \gamma$.

The first part of Lemma 5.3 implies the following corollary of Theorem 3.1:

Corollary 5.4. Let $F : \mathbb{F}_2^n \rightarrow \mathbb{T}$ be a non-classical polynomial of degree $d$. Then,

$$\Pr_{x \in \mathbb{F}_2^n}[\text{Maj}_n(x) = F(x)] \leq \frac{1}{2} + O\left(\frac{d}{\sqrt{n}}\right).$$

This proves a conjecture of Bhowmick and Lovett [6].

The following is a consequence of Theorem 2.2 and the first part of Lemma 5.3:

Corollary 5.5. Let $\ell \geq 2$. Then, for every classical polynomial $P : \mathbb{F}_2^n \rightarrow \mathbb{T}$ of degree $\leq 2\ell - 1$,

$$\Pr_{x \in \mathbb{F}_2^n}[P(x) = S_{2\ell}(x)] \leq \frac{1}{2} + o(1).$$

On the other hand, the second part of Lemma 5.3 and Theorem 4.1 imply

Corollary 5.6. For every $\ell \geq 2$, there is a non-classical polynomial $F : \mathbb{F}_2^n \rightarrow \mathbb{T}$ of degree $\leq 2\ell - 1 + 2\ell - 2 + 2$ and depth $3$ such that

$$\Pr_{x \in \mathbb{F}_2^n}[F(x) = S_{2\ell}(x)] \geq \frac{9}{16} - o(1).$$

Noting that $2\ell - 1 + 2\ell - 2 + 2 < 2^\ell$ for $\ell \geq 4$, Corollary 5.5 and Corollary 5.6 imply the following:
Theorem 5.7. There is a Boolean function $F : \mathbb{F}_2^n \to \{0, 1/2\}$ and $d \geq 1$, such that for every classical polynomial $P$ of degree at most $d$, we have

$$\Pr_{x \in \mathbb{F}_2^n} [F(x) = P(x)] \leq \frac{1}{2} + o(1),$$

but there is a non-classical polynomial $P'$ of degree $d' \leq d$ satisfying

$$\Pr_{x \in \mathbb{F}_2^n} [F(x) = P'(x)] \geq \frac{1}{2} + \Omega(1).$$

This answers an informal question of Bhowmick and Lovett [6].

We now prove Lemma 5.3.

Proof of Lemma 5.3. Fix $F$, $d$, and $k$ for the rest of the proof.

Proof of 1: Let $P$ be a non-classical polynomial of degree $d$ and depth $k$ with $\text{agr}(F, P) = \gamma$. It is not hard to verify that $P$ can be written in the following form (See, e.g., proof of Lemma 2.2 in [6]):

$$P(x) = \frac{P''(x)}{2^k} \pmod{1}$$

where $P''(x) \in \mathbb{Z}[x_1, \ldots, x_n]$ is of degree $d$.

Suppose $P''(x) = \sum_{S \subseteq [n]} c_S \prod_{i \in S} x_i$. Choose $P' \in \mathcal{P}_{d,k}^{(2)}$, $P'(x) = \sum_{S \subseteq [n]} c'_S \prod_{i \in S} x_i$, satisfying

$$\forall S \subseteq [n], c'_S \equiv c_S \pmod{2^k}.$$ 

By our choice of $P'$, we have that, for every $x \in \{0, 1\}^n$ and $a \in \{0, \ldots, 2^k - 1\}$,

$$P(x) = \frac{a}{2^k} \iff P'(x) = a.$$ 

It follows that $\text{agr}(F_k, P') = \gamma$.

Proof of 2: Let $P \in \mathcal{P}_{d,k}^{(2)}$ such that $\text{agr}(F_k, P) = \gamma$. Using arguments similar to above, we can find a $P'' \in \mathbb{Z}[x_1, \ldots, x_n]$ of degree $d$ such that $P''(x) \equiv P(x) \pmod{2^k}$, for all $x \in \{0, 1\}^n$.

Define $P'(x)$ as

$$P'(x) = \frac{P''(x)}{2^k} \pmod{1}.$$ 

By comparing to the form in Definition 5.1, it is easy to see that $P'(x)$ is a non-classical polynomial of degree at most $d + k - 1$. Furthermore, we have that, for all $x \in \{0, 1\}^n$ and $a \in \{0, \ldots, 2^k - 1\}$,

$$P(x) = a \iff P'(x) = \frac{a}{2^k}.$$ 

The completes the proof. 

6 Acknowledgements

We would like to thank David Barrington for taking the time to explain Szegedy’s [20] result to us. We would also like to thank Swagato Sanyal for helpful discussions.
References


A Proof Sketch of Fact 3.6

The following is a well-known statement about solving linear Diophantine equations (e.g. see [17, Corollary 4.1c]).

**Theorem A.1.** Fix any $\tilde{A} \in \mathbb{Z}^{M \times N}$ that is full row-rank and $\tilde{b} \in \mathbb{Z}^M$. Then, the system $\tilde{A}\tilde{z} = \tilde{b}$ has no solution in $\mathbb{Z}^N$ if and only if there is a $\tilde{y} \in \mathbb{Q}^M$ such that $\tilde{y}^T \tilde{A}$ is integral but $\tilde{y}^T \tilde{b}$ is not integral.

We now prove Fact 3.6. One of the directions is trivial. For the other direction, we proceed as follows. Assume that $Az = b$ has no solution over $\mathbb{Z}/m\mathbb{Z}$.

Identify $\mathbb{Z}/m\mathbb{Z}$ with $\{0, 1, \ldots, m - 1\}$ in the natural way and using this identification, obtain $\tilde{A} \in \mathbb{Z}^{M \times N}$ from $A$ and $\tilde{b} \in \mathbb{Z}^M$ from $b$. Note that $Az = b$ has a solution over $\mathbb{Z}/m\mathbb{Z}$ iff $\tilde{A}\tilde{z} + mIz' = \tilde{b}$ has a solution over $\mathbb{Z}$, where $z'$ is a new set of $N$ variables and $I$ is the $N \times N$ identity matrix. Defining $\tilde{B} = [\tilde{A}mI]$, we see that $Az = b$ has a solution over $\mathbb{Z}/m\mathbb{Z}$ iff $\tilde{B}\tilde{w} = \tilde{b}$ has a solution over $\mathbb{Z}$. Since we assumed that $Az = b$ has no solution over $\mathbb{Z}/m\mathbb{Z}$, Theorem A.1 implies that there is a $\tilde{y} \in \mathbb{Q}^M$ such that $\tilde{y}^T \tilde{B}$ is integral but $\tilde{y}^T \tilde{b}$ is not integral.

Since $\tilde{y}^T \tilde{B}$ is integral, in particular $m\tilde{y}^T \tilde{I} = m\tilde{y}^T$ is integral. Let $y \in \mathbb{Z}/m\mathbb{Z}^M$ be obtained by dropping the entries of $m\tilde{y}$ modulo $m$. It is easy to check that $y$ satisfies the required properties.
B Theorem 3.1 for arbitrary composites

Let $m \in \mathbb{N}$ such that $m > 1$. Given an $F \in \mathcal{F}_n$ and any non-zero $a \in \mathbb{Z}/m\mathbb{Z}$, we define the $(m, a)$-lift of $F$ to be the function $F^{(m,a)} : \{0,1\}^n \to \mathbb{Z}/m\mathbb{Z}$ defined as follows. For any $x \in \{0,1\}^n$,

$$F^{(m,a)}(x) = \begin{cases} 0 & \text{if } F(x) = 0, \\ a & \text{otherwise}. \end{cases}$$

We use the notation $\mathcal{P}^{(m)}$ to denote the set of all multilinear polynomials $Q \in \mathbb{Z}/m\mathbb{Z}[x_1, \ldots, x_n]$. Further, for $d \in \mathbb{N}$, $\mathcal{P}_d^{(m)}$ will denote the set of multilinear polynomials of degree at most $d$. For any Boolean function $F \in \mathcal{F}_n$ and any non-zero $a \in \mathbb{Z}/m\mathbb{Z}$, we define

$$\gamma_d^{(m,a)}(F) = \max_{Q \in \mathcal{P}_d^{(m)}} \text{agr}(F^{(m,a)}, Q).$$

The following lemma allows us to bound $\gamma_d^{(m,a)}$ in terms of $\gamma_d^{(p)}$ for various primes $p$.

**Lemma B.1.** Say $m = \prod_{i \in [t]} p_i^{k_i}$ is factorization of $m$ into distinct primes. Then, $\gamma_d^{(m,a)}(F) \leq \max_{i \in [t]} \gamma_d^{(p_i)}(F)$.

**Proof.** We fix any prime $p_i$ such that $p_i^{k_i} \nmid a$ (such an $i$ must exist as $a$ is non-zero modulo $m$). Let $P \in \mathcal{P}_d^{(m)}$ be the polynomial that achieves maximum agreement $\alpha$ with $F^{(m,a)}$. Dropping the coefficients of $P$ modulo $m_i := p_i^{k_i}$ gives a polynomial $P' \in \mathcal{P}_d^{(m_i)}$ such that $\text{agr}(P', F^{(m_i,b)}) \geq \alpha$ where $b \equiv a \pmod{m_i}$ is a non-zero element of $\mathbb{Z}/p_i^{k_i}\mathbb{Z}$. Hence, $\gamma_d^{(m_i,b)} \geq \alpha$. As mentioned in Remark 2.5, we have $\gamma_d^{(p_i)} \geq \gamma_d^{(m_i,b)}$. This proves the lemma.

Lemma B.1 and Theorem 3.1 together imply the following.

**Corollary B.2.** For any $m \geq 2$ and $a \in \mathbb{Z}/m\mathbb{Z}$ non-zero, $\gamma_d^{(m,a)}(\text{Maj}_n) \leq \frac{1}{2} + \frac{10d}{\sqrt{n}}$. 

24