

A Dense Model Theorem for the Boolean Slice

Gil Kalai *

Noam Lifshitz †

Dor Minzer ‡

Tamar Ziegler §

Abstract

The (low soundness) linearity testing problem for the middle slice of the Boolean cube is as follows. Let $\varepsilon > 0$ and f be a function on the middle slice on the Boolean cube, such that when choosing a uniformly random quadruple $(x, y, z, x \oplus y \oplus z)$ of vectors of $2n$ bits with exactly n ones, the probability that $f(x \oplus y \oplus z) = f(x) \oplus f(y) \oplus f(z)$ is at least $1/2 + \varepsilon$. The linearity testing problem, posed by [5], asks whether there must be an actual linear function that agrees with f on $1/2 + \varepsilon'$ fraction of the inputs, where $\varepsilon' = \varepsilon'(\varepsilon) > 0$.

We solve this problem, showing that f must indeed be correlated with a linear function. To do so, we prove a dense model theorem for the middle slice of the Boolean hypercube for Gowers uniformity norms. Specifically, we show that for every $k \in \mathbb{N}$, the normalized indicator function of the middle slice of the Boolean hypercube $\{0, 1\}^{2n}$ is close in Gowers norm to the normalized indicator function of the union of all slices with weight $t = n \pmod{2^{k-1}}$. Using our techniques we also give a more general ‘low degree test’ and a biased rank theorem for the slice.

1 Introduction

The middle slice of the Boolean hypercube $\{0, 1\}^{2n}$ is given by $\mathcal{U}_{2n} = \{x \in \{0, 1\}^{2n} \mid |x| = n\}$. A fundamental problem in a subarea of theoretical computer science known as property testing concerns determining whether properties of functions can be determined efficiently by examining only few of their inputs. One striking classical result of this kind is the BLR theorem [2] which shows that the question whether a function $f: \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is linear or far from any linear function can be determined by evaluation f on only three random inputs of the form $(x, y, x \oplus y)$. David, Dinur, Goldenberg, Kindler, and Shinkar [5] solved the corresponding problem for the slice by showing that if $(x, y, x \oplus y)$ are a uniformly random triple among such triples with $x, y, x \oplus y \in \mathcal{U}_{2n}$ and $f: \mathcal{U}_{2n} \rightarrow \mathbb{F}_2$ satisfies $\Pr[f(x \oplus y) = f(x) \oplus f(y)] \geq 1 - \varepsilon$, then f agrees with a linear function on $1 - O(\varepsilon)$ fraction of the inputs.

They then posed the corresponding problem in the so called ‘low soundness’ regime, also known as the 1% regime, where one wishes to show that if a function passes the test with probability significantly larger than $1/2$, then f is correlated with a linear function. The main result of this paper is a result along these lines of a closely related test.

Theorem 1.1. *For all $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon) \in \mathbb{N}$ such that the following holds for $n \geq n_0$. If a function $f: \mathcal{U}_{2n} \rightarrow \{0, 1\}$ satisfies that*

$$\Pr_{x,y,z} [f(x) \oplus f(y) \oplus f(z) = f(x \oplus y \oplus z) \mid x, y, z, x \oplus y \oplus z \in \mathcal{U}_{2n}] \geq \frac{1}{2} + \varepsilon,$$

*Einstein Institute of Mathematics, Hebrew University of Jerusalem and Efi Arazi School of Computer Science, Reichman University. Supported by ERC grant 834735 and by an ISF grant 2669/21.

†Einstein Institute of Mathematics, Hebrew University of Jerusalem. Supported by ISF grant 1980/22.

‡Department of Mathematics, Massachusetts Institute of Technology. Supported by NSF CCF award 2227876 and NSF CAREER award 2239160.

§Einstein Institute of Mathematics, Hebrew University of Jerusalem. Supported by ISF grant 2112/20.

then there exists $S \subseteq [2n]$ and $b \in \{0, 1\}$ such that

$$\Pr_x \left[f(x) = b + \bigoplus_{i \in S} x_i \mid x \in \mathcal{U}_{2n} \right] \geq \frac{1}{2} + \frac{\sqrt{\varepsilon}}{400}.$$

1.1 Method

Intuitively, one may think of the Boolean slice as very similar to the Boolean hypercube, as one only imposes a constraint on the Hamming weight of the vectors. Indeed, there is some truth to it; as far as low-degree functions are concerned (more specifically, degree $o(\sqrt{n})$ functions), the two domains are almost interchangeable; this is formalized in [10, 9] as an invariance principle (see also [3] for a simplification and extensions). In general however, proving results for the slice is more challenging, even when these results are concerned with low-degree functions; see for example [8, 7]. For results regarding high degree functions, the situation is even worse as there are high-degree functions that certainly distinguish between the two domains, and therefore the argument in the slice has to be significantly different from the argument on the hypercube. The problem of linearity testing over the slice is an example problem which concerns high degree functions (in the sense of Fourier analytic degree), and as mentioned prior to this paper only partial results in the 99% regime were known [5].

Our main tool is a variation of the dense model theorem, which was famously used by Green and Tao [11] to show the primes contain an arbitrary long arithmetic progression. Roughly speaking, the dense model theorem allows transferring arithmetic properties of dense subsets of an Abelian group G to analogue properties of dense subsets of arbitrary sets $S \subseteq G$ with the property that $\frac{|G|1_S}{|S|}$ has a small U_k -norm.

The dense model provides us with another avenue of comparing the Boolean slice \mathcal{U}_{2n} with the Boolean hypercube and deducing results about the former from the latter. Here and throughout, for an integer $s \geq 1$, the Gowers uniformity norm U_s of a function $f: \{0, 1\}^n \rightarrow \mathbb{C}$ is defined as

$$\|f\|_{U_s} = \left(\mathbb{E}_{x, h_1, \dots, h_s \in \{0, 1\}^n} \left[\prod_{T \subseteq [s]} C^{|T|} f \left(x + \sum_{i \in T} h_i \right) \right] \right)^{\frac{1}{2^s}},$$

where C denotes complex conjugation. For $k \in \mathbb{N}$ thought of as a constant, suppose that $n = a \pmod{2^{k-1}}$ and define

$$\mathcal{D}_{2n,k} = \left\{ x \in \{0, 1\}^{2n} \mid |x| = a \pmod{2^{k-1}} \right\}.$$

Note that $\mathcal{D}_{2n,k} \supseteq \mathcal{U}_{2n}$, and while the measure of \mathcal{U}_{2n} is vanishing with n (more specifically, it is $\Theta(1/\sqrt{n})$), the measure of $\mathcal{D}_{2n,k}$ is constant (roughly standing at $\frac{1}{2^{k-1}} + o(1)$). The main tool of our paper establishes that $\mathcal{D}_{2n,k}$ is a dense model for \mathcal{U}_{2n} for Gowers' uniformity norms, and more precisely:

Theorem 1.2. *For all $k \in \mathbb{N}, \varepsilon > 0$ there exists $n_0 > 0$, such that if $n > n_0$ we have*

$$\left\| \frac{1_{\mathcal{U}_{2n}}}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} - \frac{1_{\mathcal{D}_{2n,k}}}{\mathbb{E}[1_{\mathcal{D}_{2n,k}}]} \right\|_{U_k} \leq \varepsilon.$$

A variation of the Green Tao dense model theorem allows us to deduce Theorem 1.1 from Theorem 1.2. Indeed, we use the dense model theorem of [6].¹

¹We remark that for $k = 2$ one could use the result of Conlon, Fox and Zhao [4], but the linear forms condition becomes difficult to check for $k \geq 3$.

1.2 Other applications

1.2.1 Low Degree Testing

A second application of our method concerns the problem of testing polynomial of higher-degree, also in the low-soundness regime. For an integer $d \in \mathbb{N}$, consider the d -Gowers' test over the slice:

1. Sample $x, h_1, \dots, h_d \in \{0, 1\}^{2n}$ conditioned on $x \oplus \bigoplus_{i \in T} h_i \in \mathcal{U}_{2n}$ for all $T \subseteq [d]$.
2. Check that $\sum_{T \subseteq [d]} f(x \oplus \bigoplus_{i \in T} h_i) = 0$.

We prove the following result, asserting that if a Boolean function f passes the d -Gowers' test with probability $1/2 + \varepsilon$, then it must be correlated with a non-classical polynomial of degree at most $d - 1$.²

Theorem 1.3. *For all $\varepsilon > 0$, $d \in \mathbb{N}$ there are $n_0 \in \mathbb{N}$ and $\delta > 0$ such that the following holds for $n \geq n_0$. Suppose that $f: \mathcal{U}_{2n} \rightarrow \{0, 1\}$ passes the d -Gowers' test with probability at least $\frac{1}{2} + \varepsilon$. Then there exists a non-classical polynomial $p: \{0, 1\}^{2n} \rightarrow [0, 1]$ of degree at most $d - 1$ such that*

$$\left| \mathbb{E}_{x \in \mathcal{U}_{2n}} \left[(-1)^{f(x)} e^{2\pi i p(x)} \right] \right| \geq \delta.$$

Theorem 1.3 gives an answer to a question of [5] regarding degree d testing over the slice.

1.2.2 Biased Rank

A third application of our method is concerned with the biased rank problem. In this scenario, we show that a low-degree polynomial P on the slice which is biased must be of small rank. More precisely,

Theorem 1.4. *Let $P: \{0, 1\}^{2n} \rightarrow [0, 1]$ be a non-classical polynomial of degree d such that $|\mathbb{E}_{x \in \mathcal{U}_{2n}} [(-1)^{P(x)}]| \geq \delta$, and suppose that $n = a \pmod{2^d}$. Then there is $L = L(\delta, d)$ such that for some j and some $\Gamma: [0, 1]^L \rightarrow [0, 1]$*

$$P(x) - \frac{j(|x| - a)}{2^d} = \Gamma(Q_1(x), \dots, Q_L(x)) \pmod{1}$$

for all $x \in \{0, 1\}^{2n}$, where Q_i are non-classical polynomials of degree strictly smaller than d . In particular, for $x \in \mathcal{U}_{2n}$ we have that

$$P(x) = \Gamma(Q_1(x), \dots, Q_L(x)) \pmod{1}.$$

1.3 Other related works

We remark that the Gowers' uniformity norms have been used in theoretical computer science in numerous other contexts, such as PCP [17], communication complexity [20], property testing [1] and more; we refer the reader to [12] for a survey. There are also a few connections between dense model theorems and theoretical computer science [16, 14, 19].

²We remark that a random function passes the test with probability $1/2$, and therefore the natural question is what can be said about a function that passes the test with probability noticeably bigger than $1/2$.

2 Preliminaries

In this section we present a few necessary facts from analysis of Boolean functions. We refer the reader to [15] for a more systematic presentation. We consider the Boolean hypercube $\{0, 1\}^n$ equipped with the uniform measure, and define an inner product for functions over $\{0, 1\}^n$ as

$$\langle f, g \rangle = \mathbb{E}_x \left[f(x) \overline{g(x)} \right]$$

for all $f, g: \{0, 1\}^n \rightarrow \mathbb{C}$. For each $S \subseteq [n]$ we may define the Fourier character $\chi_S: \{0, 1\}^n \rightarrow \{-1, 1\}$ by $\chi_S(x) = \prod_{i \in S} (-1)^{x_i}$. It is a standard fact that $\{\chi_S\}_{S \subseteq [n]}$ is an orthonormal basis for $L_2(\{0, 1\}^n)$, and thus any $f: \{0, 1\}^n \rightarrow \mathbb{R}$ can be written as $f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \chi_S(x)$ where $\widehat{f}(S) = \langle f, \chi_S \rangle$.

Definition 2.1. For an integer $1 \leq d \leq n$ and a function $f: \{0, 1\}^n \rightarrow \mathbb{R}$, we define the level d weight of f to be

$$W_{\leq d}[f] = \sum_{|S| \leq d} \left| \widehat{f}(S) \right|^2.$$

Our proof uses the level d -inequality, which asserts a function $f: \{0, 1\}^n \rightarrow \{-1, 0, 1\}$ that has small ℓ_1 -norm must have small level d weight.

Lemma 2.2. Suppose that $f: \{0, 1\}^n \rightarrow \{-1, 0, 1\}$ has $\alpha = \mathbb{E}[|f|]$. Then for all $d \leq n$,

$$W_{\leq d}[f] \leq \alpha^2 \log^{O(d)}(1/\alpha).$$

Lastly, we will use the notion of discrete derivatives defined as follows:

Definition 2.3. For a function $f: \{0, 1\}^n \rightarrow \mathbb{C}$ and a direction $h \in \{0, 1\}^n$, the discrete derivative of f in direction h is the function $\partial_h f: \{0, 1\}^n \rightarrow \mathbb{C}$ defined by

$$\partial_h f(x) = \overline{f(x \oplus h)} f(x).$$

For directions $h_1, \dots, h_d \in \{0, 1\}^n$, we define

$$\partial_{h_1, \dots, h_d} f(x) = (\partial_{h_1} \partial_{h_2} \cdots \partial_{h_d}) f(x).$$

Non-classical polynomials: Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the torus. A function $p: \{0, 1\}^n \rightarrow \mathbb{T}$ is called a degree d non-classical polynomial if $\partial_{h_1, \dots, h_{d+1}} p = 0$ for all directions $h_1, \dots, h_{d+1} \in \{0, 1\}^n$; here $\partial_h p(x) = p(x + h) - p(x)$ is the standard notion of discrete derivative. We will not use this notion of derivative and hence there will be no confusion regarding which notion of derivative is used.

We note that if p is a degree d non-classical polynomial, then $\partial_{h_1, \dots, h_{d+1}} e^{2\pi i p} \equiv 1$ for all all directions $h_1, \dots, h_{d+1} \in \{0, 1\}^n$; the derivative now is as in Definition 2.3.

3 Proof of Theorem 1.2

3.1 Auxiliary Tools

We begin by presenting a few tools that we need in the proof of our main result. We begin with the following fact, asserting that if we know the Hamming weight of x, z and $x \oplus z$ modulo 2^j , then we know the Hamming weight of $x \wedge z$ modulo 2^{j-1} .

Fact 3.1. Suppose that $|x \oplus z| = b \pmod{2^j}$, $|x| = c \pmod{2^j}$ and $|z| = d \pmod{2^j}$. Then $c + d - b$ is divisible by 2 and $|x \wedge z| = \frac{c+d-b}{2} \pmod{2^{j-1}}$.

Proof. Note that $|x \oplus z| = |x \vee z| - |x \wedge z| = |x| + |z| - 2|x \wedge z|$, and the result follows from re-arranging and dividing by 2. \square

It will be convenient for us to think of a vector $x \in \{0, 1\}^{2n}$ also as a subset of $[2n]$, namely as $\text{supp}(x)$. For vectors $x_1, \dots, x_t \in \{0, 1\}^{2n}$, the algebra generated by them $\mathcal{B} = \mathcal{B}[\text{supp}(x_1), \dots, \text{supp}(x_t)]$ consists of vectors that correspond to sets that can be formed by $\text{supp}(x_1), \dots, \text{supp}(x_t)$ by taking unions, intersections and complements (in other words, it is the σ -algebra generated by these sets). Suppose that $x_1, \dots, x_t \in \{0, 1\}^{2n}$ are vectors so that in $\mathcal{B} = \mathcal{B}[\text{supp}(x_1), \dots, \text{supp}(x_t)]$, each atom has at least $n/2^{t+1}$ elements. It is a standard computation that sampling $x \in \{0, 1\}^{2n}$ uniformly we have that $x \in \mathcal{U}_{2n}$ with probability $\Theta(1/\sqrt{n})$. We would like to say that the events that $x \oplus z \in \mathcal{U}_{2n}$ for all $z \in \text{Span}(x_1, \dots, x_t)$ are almost independent, and in particular that

$$\Pr_x [x \oplus z \in \mathcal{U}_{2n} \text{ for all } z \in \text{Span}(x_1, \dots, x_t)] \leq O_t(\sqrt{n}^{-2^t}).$$

The following lemma shows that this is indeed the case.

Lemma 3.2. In the setting above, for all $n_z \in \mathbb{N}$

$$\Pr_x [|x \oplus z| = n_z \text{ for all } z \in \text{Span}(x_1, \dots, x_t)] \leq O_t(n^{-2^{t-1}}).$$

Proof. Let $B \in \mathcal{B}$ be an atom, and without loss of generality $B = \text{supp}(x_1) \cap \dots \cap \text{supp}(x_t)$. We show that if this event holds, then $|x \wedge x_1 \wedge \dots \wedge x_t|$ has to be equal to some specific value. Indeed, note that as in Fact 3.1 for all $z \in \text{Span}(x_1, \dots, x_t)$

$$|x \wedge z| = \frac{1}{2} (|x| + |z| - |x \oplus z|),$$

hence the Hamming weight of $x \wedge z$ has to be a specific value if the event in question holds. Also, we know that for $z \in \text{Span}(x_2, \dots, x_t)$ we have

$$|x \wedge (x_1 \oplus z)| = |x \wedge (x_1 \cup z)| - |x \wedge x_1 \wedge z| = |x \wedge x_1| + |x \wedge z| - 2|x \wedge x_1 \wedge z|,$$

and hence the Hamming weight of $x \wedge x_1 \wedge z$ is determined. Noting that

$$|(x \wedge x_1) \oplus z| = |x \wedge x_1| + |z| - 2|x \wedge x_1 \wedge z|$$

we get that the Hamming weight of $(x \wedge x_1) \oplus z$ is determined for all $z \in \text{Span}(x_2, \dots, x_t)$. Iterating the argument with $x' = x \wedge x_1$ gives that the Hamming weight of $(x' \wedge x_2) \oplus z$ is determined for all $z \in \text{Span}(x_3, \dots, x_t)$, and repeating this argument shows that the Hamming weight of $|x \wedge x_1 \wedge \dots \wedge x_t|$ is determined.

For each $b \in \{-1, 1\}^t$, let $x_b \in \{0, 1\}^{2n}$ be the vector $\bigwedge_{i=1}^t v_i$ where $v_i = x_i$ if $b_i = 1$ and else $v_i = 1 - x_i$. From the above argument, it follows that for each b there is a number n_b such that if the event in question holds, then $|x \wedge x_b| = n_b$. Thus, the probability in question is at most

$$\Pr_x [|x \wedge x_b| = n_b \ \forall b \in \{-1, 1\}^t] = \prod_{b \in \{-1, 1\}^t} \Pr_x [|x \wedge x_b| = n_b],$$

where the last transition holds because the supports of the vectors x_b are disjoint and so the random variables $x \wedge x_b$ are independent. As the support of x_b is in the algebra generated by x_1, \dots, x_t , we have that $|x_b| \geq n/2^{t+1}$ and hence

$$\Pr_x [|x \wedge x_b| = n_b] \leq O\left(\sqrt{\frac{2^t}{n}}\right),$$

and plugging this above finishes the proof. \square

We will also need to consider sub-events of the event in Lemma 3.2 and argue that they are also almost independent. More precisely:

Lemma 3.3. *Let $Z \subseteq \text{Span}(x_1, \dots, x_t)$. Then*

$$\Pr_x [x \oplus z \in \mathcal{U}_{2n} \ \forall z \in Z] \leq n^{-|Z|/2} (\log n)^{O_t(1)}.$$

Proof. Consider $z' \in \overline{Z} := \text{Span}(x_1, \dots, x_t) \setminus Z$, and note that by Chernoff's bound the probability that $|x \oplus z'|$ is outside the range between $n - \sqrt{n} \log n$ and $n + \sqrt{n} \log n$ is at most $2^{-\Omega(\log^2 n)}$. Thus, the probability in question is at most

$$\Pr_x [x \oplus z \in \mathcal{U}_{2n} \ \forall z \in Z, |x \oplus z'| \in [n - \sqrt{n} \log n, n + \sqrt{n} \log n] \ \forall z' \in \overline{Z}] + 2^{t-\Omega(\log^2 n)}. \quad (1)$$

Denote $q = \Pr_x [x \oplus z \in \mathcal{U}_{2n} \ \forall z \in Z, |x \oplus z'| \in [n - \sqrt{n} \log n, n + \sqrt{n} \log n] \ \forall z' \in \overline{Z}]$. Trivially,

$$q = \sum_{(n_{z'})_{z' \in \overline{Z}} \in [n - \sqrt{n} \log n, n + \sqrt{n} \log n]^{\overline{Z}}} \Pr_x [x \oplus z \in \mathcal{U}_{2n} \ \forall z \in Z, |x \oplus z'| = n_{z'} \ \forall z' \in \overline{Z}],$$

and as by Lemma 3.2 each summand is at most $O_t(n^{-2^{t-1}})$ we get that

$$q \leq (2\sqrt{n} \log n + 1)^{|\overline{Z}|} O_t(n^{-2^{t-1}}) \leq n^{-|Z|/2} (\log n)^{O_t(1)}.$$

Plugging this into (1) finishes the proof. \square

3.2 The Main Argument

The proof of Theorem 1.2 is by induction. To be more precise, fix n and k as in the theorem. We prove by induction on k' the following result:

Theorem 3.4. *For all $1 \leq k' \leq k$ we have that*

$$\left\| \frac{1_{\mathcal{U}_{2n}}}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} - \frac{1_{\mathcal{D}_{2n,k}}}{\mathbb{E}[1_{\mathcal{D}_{2n,k}}]} \right\|_{U_{k'}}^{2^{k'}} \leq \frac{(\log n)^{C_{k',k}}}{n},$$

where $C_{k',k}$ is a constant depending only on k', k .

The base case of Theorem 3.4, namely the case that $k' = 1$, is clear as the U_1 norm is the absolute value of the expected value of the function, which in our case is 0. We now move on to the inductive hypothesis and fix $k' \geq 2$ and assume the statement holds for $k' - 1$. We show that the statement holds for k' with $C_{k',k} \leq C_{k'-1,k} + O_k(1)$. Throughout this section, we will denote $\mathcal{D}_{2n} = \mathcal{D}_{2n,k}$ for ease of notations.

For simplicity of presentation, denote $f = \frac{1_{\mathcal{U}_{2n}}}{\mathbb{E}[1_{\mathcal{U}_{2n}}]}$ and $g = \frac{1_{\mathcal{D}_{2n}}}{\mathbb{E}[1_{\mathcal{D}_{2n}}]}$. By definition of Gowers' uniformity norms, we have that

$$\|f - g\|_{U_{k'}}^{2^{k'}} = \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[\|\partial_{x_1, \dots, x_{k'-2}}(f - g)\|_{U_2}^4 \right].$$

Consider the algebra \mathcal{B} generated by $\text{supp}(x_i)$ for $i = 1, \dots, k' - 2$, and let E be the event that each atom there has at least $n/2^k$ elements. By Chernoff's bound we have that $\Pr[\bar{E}] \leq 2^{-\Omega(n)}$, and as $\|f\|_\infty, \|g\|_\infty \leq n$ we get that

$$\begin{aligned} \|f - g\|_{U_{k'}}^{2^{k'}} &\leq 2^{-\Omega(n)} n^{\Theta(2^k)} + \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[\|\partial_{x_1, \dots, x_{k'-2}}(f - g)\|_{U_2}^4 1_E \right] \\ &\leq O_k \left(\frac{1}{n} \right) + \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_S \left| \widehat{\partial_{x_1, \dots, x_{k'-2}}(f - g)}(S) \right|^4 \right]. \end{aligned} \quad (2)$$

We focus on the second sum now. Consider the function $h(x) = \partial_{x_1, \dots, x_{k'-2}}(f - g)(x)$, and define the subgroup $G \subseteq S_{2n}$ as

$$G = \{ \pi \in S_{2n} \mid \pi(B) = B \ \forall B \in \mathcal{B} \},$$

and note that h is symmetric under G . For a subset $I \subseteq [2n]$, define

$$\text{orb}_I = \{ \pi(I) \mid \pi \in G \},$$

so that we get that $\widehat{h}(I) = \widehat{h}(I')$ if $I' \in \text{orb}_I$. In our upper bound of (2) we partition the characters S into 3 types, and upper bound the contribution of each one of them separately. This partition will depend on the size of the orbit of the character S under G , and towards this end we have the following claim.

Claim 3.5. *Let $S \subseteq [2n]$, and let $d = \min_{B \in \mathcal{B}} |S \Delta B|$.*

1. *If $d = 0$, then $|\text{orb}_S| = 1$.*
2. *If $1 \leq d \leq 2^{100 \cdot k}$, then $|\text{orb}_S| \geq \frac{n}{2^{101k}}$.*
3. *If $d > 2^{100 \cdot k}$, then $|\text{orb}_S| \geq \Omega_k(n^{100 \cdot 2^k})$.*

Proof. For the first item, if $d = 0$ then $S \in \mathcal{B}$, and it is clear that $\text{orb}_S = \{S\}$.

For the second item, write $S = B \Delta I$ where $B \in \mathcal{B}$ and $|I| = d$. Thus, $\pi(S) = B \Delta \pi(I)$ for $\pi \in G$. Take some $i \in I$ and consider the atom of \mathcal{B} in which i lies, say B' . For any $j \in B'$ we may find $\pi_{i,j} \in G$ such that $\pi_{i,j}(i) = j$. As $|B'| \geq n/2^k$, we conclude that there are at least $\frac{n}{d2^k}$ distinct sets among $B \Delta \pi_{i,j}(I)$.

For the third item, write $S = B \Delta I$ again for $B \in \mathcal{B}$ and I of size d . Write $I = \{i_1, \dots, i_d\}$, and let B_1, \dots, B_d be the atoms of \mathcal{B} that i_1, \dots, i_d lie in, respectively. For all distinct $j_1 \in B_1, \dots, j_d \in B_d$ we may find $\pi_{i_1, \dots, i_d, j_1, \dots, j_d} \in G$ such that $\pi_{i_1, \dots, i_d, j_1, \dots, j_d}(i_\ell) = j_\ell$ for all $\ell = 1, \dots, d$. It follows that the orbit of S has size at least $\frac{n}{2^k} \cdot \left(\frac{n}{2^k} - 1 \right) \cdots \left(\frac{n}{2^k} - 100 \cdot 2^k \right) \geq \Omega_k(n^{100 \cdot 2^k})$. \square

Define

$$T_1 = \mathcal{B}, \quad T_2 = \left\{ S \subseteq [2n] \mid 1 \leq \min_{B \in \mathcal{B}} |S \Delta B| \leq 2^{100 \cdot k} \right\}, \quad T_3 = \left\{ S \subseteq [2n] \mid \min_{B \in \mathcal{B}} |S \Delta B| > 2^{100 \cdot k} \right\},$$

so that by (2) we have that

$$\begin{aligned} \|f - g\|_{U_{k'}}^{2^{k'}} &\leq \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_1} |\widehat{h}(S)|^4 \right] + \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_2} |\widehat{h}(S)|^4 \right] \\ &\quad + \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_3} |\widehat{h}(S)|^4 \right] + O_k \left(\frac{1}{n} \right). \end{aligned} \quad (3)$$

We now upper bound each term on the right hand side of (3). We start with the contribution from T_2 .

Claim 3.6. $\mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_2} |\widehat{h}(S)|^4 \right] \leq \frac{(\log n)^{O_k(1)}}{n}.$

Proof. Clearly we have that

$$\mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_2} |\widehat{h}(S)|^4 \right] = \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{B \in \mathcal{B}} \sum_{1 \leq |S \Delta B| \leq 2^{100 \cdot k}} |\widehat{h}(S)|^4 \right].$$

Denote $W_B = \sum_{1 \leq |S \Delta B| \leq 2^{100 \cdot k}} |\widehat{h}(S)|^2$. As by Claim 3.5 the orbit of each S has size at least $n/2^{101k}$, we get that

$$\mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_2} |\widehat{h}(S)|^4 \right] \leq \frac{2^{101k}}{n} \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{B \in \mathcal{B}} W_B^2 \right]. \quad (4)$$

Fix $x_1, \dots, x_{k'-2}$ satisfying E and $B \in \mathcal{B}$, and inspect W_B . Opening up the definition of h , by the triangle inequality $|\widehat{h}(S)|$ is upper bounded by a sum of $2^{2^{k'-2}}$ terms of the form $|\mathbb{E}_x [F_1(x) \cdots F_{2^{k'-2}}(x) \chi_S(x)]|$ where writing $\text{Span}(x_1, \dots, x_{k'-2}) = \{z_1, \dots, z_{2^{k'-2}}\}$, each F_i is either $f(x + z_i)$ or $g(x + z_i)$. Thus, by Cauchy-Schwarz

$$W_B \leq 2^{2^{k'-2}} \sum_{F_1, \dots, F_{2^{k'-2}}} W_{\leq 2^{100 \cdot k}} [F_1 \cdots F_{2^{k'-2}} \chi_B].$$

Fix a choice of $F_1, \dots, F_{2^{k'-2}}$. Namely, fix a subset $Z \subseteq \text{Span}(x_1, \dots, x_{k'-2})$, and say that

$$F_1 \cdots F_{2^{k'-2}}(x) = \prod_{z \in Z} f(x \oplus z) \prod_{z' \in \overline{Z}} g(x \oplus z).$$

Write $P(x) = \prod_{z \in Z} 1_{x \oplus z \in \mathcal{U}_{2n}} \prod_{z \in \overline{Z}} 1_{x \oplus z \in \mathcal{D}_{2n}}$. Then

$$\begin{aligned} W_{\leq 2^{100 \cdot k}} [F_1 \cdots F_{2^{k'-2}} \chi_B] &= \frac{1}{\mu(\mathcal{U}_{2n})^{2|Z|} \mu(\mathcal{D}_{2n})^{2|\overline{Z}|}} W_{\leq 2^{100 \cdot k}} [P \chi_B] \\ &\leq \frac{1}{\mu(\mathcal{U}_{2n})^{2|Z|} \mu(\mathcal{D}_{2n})^{2|\overline{Z}|}} \mathbb{E}[P]^2 \log^{O_k(1)}(\mathbb{E}[P]). \end{aligned}$$

where the last transition is by Lemma 2.2. By Lemma 3.3 we have $\mathbb{E}[P] \leq n^{-|Z|/2}(\log n)^{O_k(1)}$, and by inspection $\mu(\mathcal{D}_{2n}) \geq \Omega_k(1)$, $\mu(\mathcal{U}_{2n}) \geq \Omega(n^{-1/2})$. Thus, using the monotonicity of the function $z \rightarrow z^2 \log^{O_k(1)}(1/z)$ in the interval $[0, c]$ where $c = c(k) > 0$ is an absolute constant, we get that

$$W_{\leq 2^{100 \cdot k}}[F_1 \cdots F_{2^{k-2}} \chi_B] \leq (\log n)^{O_k(1)}.$$

It follows that

$$W_B \leq (\log n)^{O_k(1)}, \quad (5)$$

and plugging (5) into (4) gives that

$$\mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_2} \left| \hat{h}(S) \right|^4 \right] \leq \frac{(\log n)^{O_k(1)}}{n}.$$

as required. \square

Next, we bound the contribution from T_3 .

$$\text{Claim 3.7. } \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_3} \left| \hat{h}(S) \right|^4 \right] \leq O_k \left(\frac{1}{n} \right).$$

Proof. Fix $x_1, \dots, x_{k'-2}$ satisfying E and fix $S' \in T_3$. Note that

$$|\text{orb}_{S'}| \left| \hat{h}(S') \right|^2 = \sum_{S \in \text{orb}_{S'}} \left| \hat{h}(S) \right|^2 \leq \|h\|_2^2,$$

where we used Parseval's equality. Note that $\|f - g\|_\infty \leq n$, and so $\|h\|_2^2 \leq \|f - g\|_\infty^{2^{k-1}} \leq n^{2^{k-1}}$. Concluding, we get via Claim 3.5 that

$$\left| \hat{h}(S') \right|^2 \leq \frac{n^{2^{k-1}}}{|\text{orb}_{S'}|} \leq O_k(n^{-99 \cdot 2^k}).$$

It follows that

$$\mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_3} \left| \hat{h}(S) \right|^4 \right] \leq O_k(n^{-99 \cdot 2^k}) \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[\sum_{S \in T_3} \left| \hat{h}(S) \right|^2 \right] \leq O_k(n^{-99 \cdot 2^k}) \mathbb{E}_{x_1, \dots, x_{k'-2}} [\|h\|_2^2],$$

which is at most $O_k(n^{-98 \cdot 2^k}) \leq O_k \left(\frac{1}{n} \right)$ using the upper bound $\|h\|_2^2 \leq n^{2^{k-1}}$. \square

We end by upper bounding the contribution from T_1 :

$$\text{Claim 3.8. } \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_1} \left| \hat{h}(S) \right|^4 \right] \leq \frac{(\log n)^{C_{k'-1,k} + O_k(1)}}{n}.$$

Proof. Fix $x_1, \dots, x_{k'-2}$ satisfying E . We observe that for all $S \in T_1$ it holds that $\left| \hat{h}(S) \right| = \left| \hat{h}(\emptyset) \right|$. To see that, it suffices to show that $\chi_S(x)$ is constant on all x such that $h(x) \neq 0$. It suffices to prove this assertion on any atom S of \mathcal{B} , and for the sake of simplicity of notation we take $S = \text{supp}(x_1) \cap \dots \cap \text{supp}(x_{k'-2})$.

By definition of h , if $h(x) \neq 0$, then we must have that $x \oplus z \in \mathcal{D}_{2n}$ for all $z \in \text{Span}(x_1, \dots, x_{k'-2})$. Thus, we get that for any such x it holds that $|x \oplus z| = a \pmod{2^k}$, i.e. the Hamming weight of $x \oplus z$ is constant modulo 2^k . Applying Fact 3.1 on $x \oplus z$, x and z , we conclude that the Hamming weight of $x \wedge z$ modulo 2^{k-1} depends only on $|z|$. In particular, for $z \in \text{Span}(x_2, \dots, x_{k'-2})$ the Hamming weight of $x \wedge x_1$ and $x \wedge (x_1 \oplus z)$ is constant modulo 2^{k-1} , and as

$$|x \wedge (x_1 \oplus z)| = |x \wedge (x_1 \cup z)| - |x \wedge x_1 \wedge z| = |x \wedge z| + |x \wedge z| - 2|x \wedge x_1 \wedge z|$$

we get that the Hamming weight of $|x \wedge x_1 \wedge z|$ modulo 2^{k-2} depends only on $|z|$. Continuing in this fashion, we conclude that the Hamming weight of $|x \wedge x_1 \wedge \dots \wedge x_{k-2}|$ modulo 2 is constant, and hence $\chi_S(x)$ is constant.

Thus, we get that

$$\mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_1} \left| \hat{h}(S) \right|^4 \right] \leq 2^{2^{k-2}} \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[\left| \hat{h}(\emptyset) \right|^4 \right].$$

Observe now that using the notations of Claim 3.6 we have that

$$\left| \hat{h}(\emptyset) \right|^2 \leq W_\emptyset \leq (\log n)^{O_k(1)}$$

where we used (5). Thus,

$$\mathbb{E}_{x_1, \dots, x_{k'-2}} \left[1_E \cdot \sum_{S \in T_1} \left| \hat{h}(S) \right|^4 \right] \leq (\log n)^{O_k(1)} \mathbb{E}_{x_1, \dots, x_{k'-2}} \left[\left| \hat{h}(\emptyset) \right|^2 \right] = (\log n)^{O_k(1)} \|f - g\|_{U_{k'-1}}^{2^{k'-1}}.$$

Upper bounding $\|f - g\|_{U_{k'-1}}^{2^{k'-1}}$ by $\frac{(\log n)^{C_{k'-1,k}}}{n}$ via the induction hypothesis finishes the proof. \square

Plugging Claims 3.8, 3.6, 3.7 into (3) finishes the proof of the inductive step, and thereby the proof of Theorem 3.4.

4 Applications

In this section we use Theorem 1.2 to prove a few applications in property testing. Towards this end, we use [6, Corollary 4.4], which we specialize below for the group $(\{0, 1\}^n, \oplus)$.

Theorem 4.1. *For all $k \in \mathbb{N}$ and $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds. Suppose that $\nu: \{0, 1\}^n \rightarrow [0, \infty)$ is a function such that $\|\nu - 1\|_{U_{2k}} \leq \delta$. Then for all $f: \{0, 1\}^n \rightarrow \mathbb{R}$ such that $|f(x)| \leq \nu(x)$ pointwise, there exists $\tilde{f}: \{0, 1\}^n \rightarrow [-1, 1]$ such that*

$$\|f - \tilde{f}\|_{U_k} \leq \varepsilon.$$

4.1 Proof of Theorem 1.1

In this section we prove Theorem 1.1, and for that we first need the following claim which is a version of that theorem for bounded functions on $\{0, 1\}^n$.

Claim 4.2. Suppose that a function $f: \{0, 1\}^n \rightarrow [-1, 1]$ satisfies that

$$\mathbb{E}_{x,y,z \in \{0,1\}^n} [f(x)f(y)f(z)f(x \oplus y \oplus z)] \geq \varepsilon.$$

Then there exists $S \subseteq [n]$ such that $|\widehat{f}(S)| \geq \sqrt{\varepsilon}$.

Proof. Plugging in the Fourier expansion of f , we get that

$$\begin{aligned} \varepsilon &\leq \mathbb{E}_{x,y,z \in \{0,1\}^n} [f(x)f(y)f(z)f(x \oplus y \oplus z)] = \sum_{S \subseteq [n]} \widehat{f}(S)^4 \leq \max_S \widehat{f}(S)^2 \sum_{S \subseteq [n]} \widehat{f}(S)^2 \\ &= \max_S \widehat{f}(S)^2 \|f\|_2^2, \end{aligned}$$

which is at most $\max_S \widehat{f}(S)^2$ as $\|f\|_2 \leq 1$. \square

We also need the Cauchy-Schwarz-Gowers inequality (see for example [13, Lemma 4.2])

Lemma 4.3. Suppose that $\{f_S\}_{S \subseteq [r]}$ are complex valued functions. Then

$$\left| \mathbb{E}_{x,y_1,\dots,y_r} \left[\prod_{S \subseteq [r]} C^{k-|S|} f_S \left(x + \sum_{i \in S} y_i \right) \right] \right| \leq \prod_{S \subseteq [r]} \|f_S\|_{U^r}.$$

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Fix $k = 2$, $\varepsilon > 0$ and take δ from Theorem 4.1 for $\varepsilon/10^6$. Take a function $f: \mathcal{U}_{2n} \rightarrow \{0, 1\}$ and define $f': \{0, 1\}^{2n} \rightarrow \mathbb{R}$ by $f'(x) = (-1)^{f(x)} 1_{x \in \mathcal{U}_{2n}} \frac{\mathbb{E}[1_{\mathcal{D}_{2n,2k}}]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]}$. Note that

$$\begin{aligned} \mathbb{E}_{x,y,z} [f'(x)f'(y)f'(z)f'(x \oplus y \oplus z)] &= \mathbb{E}[1_{\mathcal{D}_{2n,2k}}]^4 \frac{\Pr_{x,y,z \in \{0,1\}^{2n}} [x, y, z, x \oplus y \oplus z \in \mathcal{U}_{2n}]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]^4} \\ &\cdot \left(2 \Pr_{x,y,z} [f(x) + f(y) + f(z) = f(x \oplus y \oplus z) \mid x, y, z, x \oplus y \oplus z \in \mathcal{U}_{2n}] - 1 \right). \end{aligned}$$

By the premise the second term is at least 2ε , and a direct calculation shows that

$$\Pr_{x,y \in \{0,1\}^{2n}} [x, y, z, x \oplus y \oplus z \in \mathcal{U}_{2n}] \geq \mathbb{E}[1_{\mathcal{U}_{2n}}]^4,$$

and hence $\mathbb{E}_{x,y,z} [f'(x)f'(y)f'(z)f'(x \oplus y \oplus z)] \geq 2\mathbb{E}[1_{\mathcal{D}_{2n,2k}}]\varepsilon$. As $\mathbb{E}[1_{\mathcal{D}_{2n,2k}}] \geq \frac{1}{10}$ we get that

$$\mathbb{E}_{x,y,z} [f'(x)f'(y)f'(z)f'(x \oplus y \oplus z)] \geq \frac{2\varepsilon}{10^4}. \tag{6}$$

Take

$$\nu = \frac{\mathbb{E}[1_{\mathcal{D}_{2n,2k}}]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} 1_{\mathcal{U}_{2n}} + 1 - 1_{\mathcal{D}_{2n,2k}},$$

and note that ν is non-negative and that $|f'(x)| \leq \nu(x)$ pointwise. By Theorem 1.2 we have

$$\|\nu - 1\|_{U_{2k}} = \left\| \frac{\mathbb{E}[1_{\mathcal{D}_{2n,2k}}]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} 1_{\mathcal{U}_{2n}} - 1_{\mathcal{D}_{2n,2k}} \right\|_{U_{2k}} \leq \left\| \frac{1}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} 1_{\mathcal{U}_{2n}} - \frac{1_{\mathcal{D}_{2n,2k}}}{\mathbb{E}[1_{\mathcal{D}_{2n,2k}}]} \right\|_{U_{2k}} = o(1),$$

and hence for sufficiently large n_0 , if $n \geq n_0$ we have that $\|\nu - 1\|_{U_{2k}} \leq \delta$. Applying Theorem 4.1 we get that there is a function $\tilde{f}: \{0, 1\}^n \rightarrow [-1, 1]$ with $\|f' - \tilde{f}\|_{U_k} \leq \frac{\varepsilon}{10^6}$. Define the function $\Delta(x) = f'(x) - \tilde{f}(x)$; by the triangle inequality we may bound

$$\begin{aligned}
& \left| \mathbb{E}_{x,y,z} [f'(x)f'(y)f'(z)f'(x \oplus y \oplus z)] - \mathbb{E}_{x,y,z} [\tilde{f}(x)\tilde{f}(y)\tilde{f}(z)\tilde{f}(x \oplus y \oplus z)] \right| \\
&= \left| \mathbb{E}_{x,y,z} [(\tilde{f} + \Delta)(x)(\tilde{f} + \Delta)(y)(\tilde{f} + \Delta)(z)(\tilde{f} + \Delta)(x \oplus y \oplus z)] - \mathbb{E}_{x,y,z} [\tilde{f}(x)\tilde{f}(y)\tilde{f}(z)\tilde{f}(x \oplus y \oplus z)] \right| \\
&\leq \mathbb{E}_{x,y,z} [4\Delta(x)\tilde{f}(y)\tilde{f}(z)\tilde{f}(x \oplus y \oplus z) + 6\Delta(x)\Delta(y)\tilde{f}(z)\tilde{f}(x \oplus y \oplus z) + 4\Delta(x)\Delta(y)\Delta(z)\tilde{f}(x \oplus y \oplus z)] \\
&+ \mathbb{E}_{x,y,z} [\Delta(x)\Delta(y)\Delta(z)\Delta(x \oplus y \oplus z)]. \tag{7}
\end{aligned}$$

We use Lemma 4.3 to bound each expectation by $\|\Delta\|_{U_2}$, and we take the first expression for example. Note that the distributions of $(x, y, z, x \oplus y \oplus z)$ is the same as of $(x, x \oplus y, x \oplus z, x \oplus y \oplus z)$ and that Δ, \tilde{f} are real valued, so by Lemma 4.3 the expectation of the first term is at most

$$\|\Delta\|_{U_2} \|\tilde{f}\|_{U_2}^3 \leq \|\Delta\|_{U_2},$$

where we used the fact that \tilde{f} is 1-bounded. Therefore, combining (6) (7) we get that

$$\mathbb{E}_{x,y,z} [\tilde{f}(x)\tilde{f}(y)\tilde{f}(z)\tilde{f}(x \oplus y \oplus z)] \geq \frac{2\varepsilon}{10^4} - 15\|\Delta\|_{U_2} \geq \frac{\varepsilon}{10^4}.$$

By Claim 4.2 we conclude that there exists $S \subseteq [2n]$ such that $|\hat{\tilde{f}}(S)| \geq \frac{\sqrt{\varepsilon}}{100}$, and so

$$|\hat{f}'(S)| \geq |\hat{\tilde{f}}(S)| - \|f' - \tilde{f}\|_1 \geq |\hat{\tilde{f}}(S)| - \|\Delta\|_{U_2} \geq \frac{\sqrt{\varepsilon}}{200}.$$

Define $L_S = \bigoplus_{i \in S} x_i$, and suppose without loss of generality that $\hat{f}'(S)$ is positive (otherwise we multiply \tilde{f} by -1). Note that

$$\frac{\sqrt{\varepsilon}}{200} \leq \hat{f}'(S) = \mathbb{E}[1_{\mathcal{D}_{2n,2k}}] \mathbb{E}_{x \in \mathcal{U}_{2n}} [(-1)^{L_S(x) + f(x)}] = \mathbb{E}[1_{\mathcal{D}_{2n,2k}}] \left(2 \Pr_{x \in \mathcal{U}_{2n}} [f(x) = L_S(x)] - 1 \right).$$

Rearranging and using $\mathbb{E}[1_{\mathcal{D}_{2n,2k}}] \leq 1$ finishes the proof. \square

4.2 Proof of Theorem 1.3

We first need the following elementary conclusion of Theorem 1.2.

Fact 4.4.

$$\Pr_{x, h_1, \dots, h_d \in \{0, 1\}^{2n}} \left[x \oplus \bigoplus_{i \in T} h_i \in \mathcal{U}_{2n} \ \forall T \subseteq [d] \right] \geq \mathbb{E}[1_{\mathcal{U}_{2n}}]^{2^d}.$$

Proof. Define $f(x) = \frac{1_{x \in \mathcal{U}_{2n}}}{\mathbb{E}[1_{x \in \mathcal{U}_{2n}}]}$, and note that the left hand side is equal to

$$\mathbb{E}[1_{\mathcal{U}_{2n}}]^{2^d} \mathbb{E}_{x, h_1, \dots, h_d} \left[\prod_{T \subseteq [d]} f \left(x \oplus \bigoplus_{i \in T} h_i \right) \right] = \mathbb{E}[1_{\mathcal{U}_{2n}}]^{2^d} \|f\|_{U_d}^{2^d} \geq \mathbb{E}[1_{\mathcal{U}_{2n}}]^{2^d} \|f\|_{U_1}^{2^d} = \mathbb{E}[1_{\mathcal{U}_{2n}}]^{2^d},$$

where we used the fact that $\|f\|_{U_d} \geq \|f\|_{U_1}$ and that $\|f\|_{U_1} = 1$. \square

Towards the proof of Theorem 1.3 we shall need the inverse Gowers theorem over fields with small characteristics due to [18].

Theorem 4.5. *For all $\varepsilon > 0$ there is $\delta > 0$ such that if $f: \{0, 1\}^n \rightarrow [-1, 1]$ satisfies that $\|f\|_{U_d} \geq \varepsilon$. Then there exists a non classical polynomial $p: \{0, 1\}^n \rightarrow \mathbb{T}$ of degree at most $d - 1$ such that*

$$|\langle f, e^{2\pi i p} \rangle| \geq \delta.$$

We are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. We take parameters

$$0 < \xi \ll \eta \ll \delta \ll \varepsilon \ll d^{-1} = k^{-1} \leq 1.$$

Take a function $f: \mathcal{U}_{2n} \rightarrow \{0, 1\}$ and define $f': \{0, 1\}^{2n} \rightarrow \mathbb{R}$ by

$$f'(x) = (-1)^{f(x)} 1_{x \in \mathcal{U}_{2n}} \frac{\mathbb{E}[1_{\mathcal{D}_{2n,2d}}]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]}.$$

Note that

$$\begin{aligned} \|f'\|_{U_d}^{2^d} &= \mathbb{E}[1_{\mathcal{D}_{2n,2d}}]^{2^d} \frac{\Pr_{x, h_1, \dots, h_d \in \{0,1\}^{2n}} \left[x \oplus \bigoplus_{i \in T} h_i \in \mathcal{U}_{2n} \ \forall T \subseteq [d] \right]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]^{2^d}} \\ &\cdot \left(2 \Pr_{x, h_1, \dots, h_d \in \{0,1\}^{2n}} \left[\sum_T f(x \oplus \bigoplus_{i \in T} h_i) = 0 \mid x \oplus \bigoplus_{i \in T} h_i \in \mathcal{U}_{2n} \ \forall T \subseteq [d] \right] - 1 \right). \end{aligned}$$

The first term is at least $\Omega_d(1)$, the second term is at least 1 by Fact 4.4 and the third term is at least 2ε by the premise. Thus, $\|f'\|_{U_d}^{2^d} \geq \Omega_d(\varepsilon)$. Take

$$\nu = \frac{\mathbb{E}[1_{\mathcal{D}_{2n,2d}}]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} 1_{\mathcal{U}_{2n}} + 1 - 1_{\mathcal{D}_{2n,2d}},$$

and note that ν is non-negative and that $|f'(x)| \leq \nu(x)$ pointwise. By Theorem 1.2 we have

$$\|\nu - 1\|_{U_{2d}} = \left\| \frac{\mathbb{E}[1_{\mathcal{D}_{2n,2d}}]}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} 1_{\mathcal{U}_{2n}} - 1_{\mathcal{D}_{2n,2d}} \right\|_{U_{2d}} \leq \left\| \frac{1}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} 1_{\mathcal{U}_{2n}} - \frac{1_{\mathcal{D}_{2n,2d}}}{\mathbb{E}[1_{\mathcal{D}_{2n,2d}}]} \right\|_{U_{2d}} = o(1),$$

and hence for sufficiently large n_0 , if $n \geq n_0$ we have that $\|\nu - 1\|_{U_{2d}} \leq \xi$. Applying Theorem 4.1 we get that there is a function $\tilde{f}: \{0, 1\}^n \rightarrow [-1, 1]$ with $\|f' - \tilde{f}\|_{U_d} \leq \eta$. It follows from the triangle inequality that

$$\|\tilde{f}\|_{U_d} \geq \|f'\|_{U_d} - \|f' - \tilde{f}\|_{U_d} \geq \Omega_d(\varepsilon^{2^{-d}}) - \eta \geq \Omega_d(\varepsilon^{2^{-d}}).$$

By Theorem 4.5 we conclude that there exists a non-classical polynomial $p: \{0, 1\}^n \rightarrow \mathbb{T}$ of degree at most $d - 1$ such that

$$|\langle \tilde{f}, e^{2\pi i p} \rangle| \geq \delta.$$

Thus, by the triangle inequality

$$|\langle f', e^{2\pi i p} \rangle| \geq |\langle \tilde{f}, e^{2\pi i p} \rangle| - |\langle f' - \tilde{f}, e^{2\pi i p} \rangle| \geq \delta - |\langle f' - \tilde{f}, e^{2\pi i p} \rangle|.$$

Note that

$$\left| \left\langle f' - \tilde{f}, e^{2\pi i p} \right\rangle \right| = \|(f' - \tilde{f})e^{-2\pi i p}\|_{U_1} \leq \|(f' - \tilde{f})e^{-2\pi i p}\|_{U_d} = \|f' - \tilde{f}\|_{U_d},$$

where in the last transition we used the fact that p is a non-classical polynomial of degree at most $d - 1$ and hence its order d derivatives vanish. As $\|f' - \tilde{f}\|_{U_d} \leq \eta$ we conclude that $|\langle f', e^{2\pi i p} \rangle| \geq \frac{\delta}{2}$, and so

$$\left| \mathbb{E}_{x \in \mathcal{U}_{2n}} \left[(-1)^{f(x)} e^{-2\pi i p(x)} \right] \right| \geq \frac{\delta}{2}. \quad \square$$

4.3 Proof of Theorem 1.4

The proof requires [18, Theorem 1.20], stated below for convenience.

Theorem 4.6. *Let $\delta > 0$, and suppose that $p: \{0, 1\}^{2n} \rightarrow \mathbb{T}$ is a non-classical polynomial of degree d such that $\|e^{2\pi i p}\|_{U_d} \geq \delta$. Then there exists $L = L(\delta, d)$, non-classical polynomials Q_1, \dots, Q_L of degree at most $d - 1$ and $F: [0, 1]^L \rightarrow \mathbb{T}$ such that for all $x \in \{0, 1\}^{2n}$ we have $P(x) = F(Q_1(x), \dots, Q_L(x))$.*

We are now ready to prove Theorem 1.4, and we assume for convenience of notations that $n \equiv 0 \pmod{2^d}$. Note that

$$\begin{aligned} \left| \mathbb{E}_x \left[e^{2\pi i P(x)} \left(\frac{1_{\mathcal{U}_{2n}}(x)}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} - \frac{1_{\mathcal{D}_{2n,d+1}}(x)}{\mathbb{E}[1_{\mathcal{D}_{2n,d+1}}]} \right) \right] \right| &= \left\| e^{2\pi i P} \left(\frac{1_{\mathcal{U}_{2n}}}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} - \frac{1_{\mathcal{D}_{2n,d+1}}}{\mathbb{E}[1_{\mathcal{D}_{2n,d+1}}]} \right) \right\|_{U_1} \\ &\leq \left\| e^{2\pi i P} \left(\frac{1_{\mathcal{U}_{2n}}}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} - \frac{1_{\mathcal{D}_{2n,d+1}}}{\mathbb{E}[1_{\mathcal{D}_{2n,d+1}}]} \right) \right\|_{U_{d+1}} \\ &= \left\| \frac{1_{\mathcal{U}_{2n}}}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} - \frac{1_{\mathcal{D}_{2n,d+1}}}{\mathbb{E}[1_{\mathcal{D}_{2n,d+1}}]} \right\|_{U_{d+1}} \\ &= o(1), \end{aligned}$$

where we used Theorem 1.2. Thus, as the assumption of the theorem implies that $\left| \mathbb{E}_x \left[\frac{e^{2\pi i P(x)}}{\mathbb{E}[1_{\mathcal{U}_{2n}}]} 1_{\mathcal{U}_{2n}}(x) \right] \right| \geq \delta$, we conclude that

$$\left| \mathbb{E}_x \left[\frac{e^{2\pi i P(x)}}{\mathbb{E}[1_{\mathcal{D}_{2n,d+1}}]} 1_{\mathcal{D}_{2n,d+1}}(x) \right] \right| \geq \delta - o(1) \geq \frac{\delta}{2}.$$

By definition,

$$1_{\mathcal{D}_{2n,d+1}}(x) = 1_{|x|=0 \pmod{2^d}} = \frac{1}{2^d} \sum_{j=0}^{2^d-1} e^{2\pi i \frac{j|x|}{2^d}},$$

and plugging this above implies that there exists j such that $\left| \mathbb{E}_x \left[e^{2\pi i (P(x) + j|x|/2^d)} \right] \right| \geq \frac{\delta}{2}$. Note that the polynomial $P(x) + j|x|/2^d$ is a non-classical polynomial of degree d . The conclusion is now immediate from Theorem 4.6.

5 Acknowledgements

We thank Zach Hunter for carefully reading the paper, giving us useful feedback and pointing a simplification in the proof of Fact 4.4. We thank an anonymous referee for helpful comments.

References

- [1] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing low-degree polynomials over $gf(2)$. In *International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 188–199. Springer, 2003.
- [2] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/ correcting with applications to numerical problems. *J. Comput. Syst. Sci.*, 47(3):549–595, 1993.
- [3] Mark Braverman, Subhash Khot, Noam Lifshitz, and Dor Minzer. An invariance principle for the multi-slice, with applications. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 228–236. IEEE, 2021.
- [4] David Conlon, Jacob Fox, and Yufei Zhao. The green-tao theorem: an exposition. *EMS Surveys in Mathematical Sciences*, 1(2):249–282, 2014.
- [5] Roe David, Irit Dinur, Elazar Goldenberg, Guy Kindler, and Igor Shinkar. Direct sum testing. *SIAM J. Comput.*, 46(4):1336–1369, 2017.
- [6] Pandelis Dodos and Vassilis Kanellopoulos. Uniformity norms, their weaker versions, and applications. *Acta Arithmetica*, 203(3):251–270, 2022.
- [7] Yuval Filmus. Friedgut–Kalai–Naor theorem for slices of the boolean cube. *Chic. J. Theor. Comput. Sci.*, 2016, 2016.
- [8] Yuval Filmus. Junta threshold for low degree boolean functions on the slice. *Electron. J. Comb.*, 30(1), 2023.
- [9] Yuval Filmus, Guy Kindler, Elchanan Mossel, and Karl Wimmer. Invariance principle on the slice. *ACM Trans. Comput. Theory*, 10(3):11:1–11:37, 2018.
- [10] Yuval Filmus and Elchanan Mossel. Harmonicity and invariance on slices of the boolean cube. In *CCC 2016*, volume 50 of *LIPICS*, pages 16:1–16:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.
- [11] Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Annals of mathematics*, pages 481–547, 2008.
- [12] Hamed Hatami, Pooya Hatami, and Shachar Lovett. Higher-order fourier analysis and applications. *Foundations and Trends® in Theoretical Computer Science*, 13(4):247–448, 2019.
- [13] Hamed Hatami and Shachar Lovett. Higher-order fourier analysis of and the complexity of systems of linear forms. *Geometric And Functional Analysis*, 21(6):1331–1357, 2011.
- [14] Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil Vadhan. Computational differential privacy. In *Annual International Cryptology Conference*, pages 126–142. Springer, 2009.
- [15] Ryan O’Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.
- [16] Omer Reingold, Luca Trevisan, Madhur Tulsiani, and Salil Vadhan. Dense subsets of pseudorandom sets. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 76–85. IEEE, 2008.

- [17] Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and pcps. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of Computing*, pages 11–20, 2006.
- [18] Terence Tao and Tamar Ziegler. The inverse conjecture for the gowers norm over finite fields in low characteristic. *Annals of Combinatorics*, 16(1):121–188, 2012.
- [19] Luca Trevisan. Dense model theorems and their applications. In *Theory of Cryptography Conference*, pages 55–57. Springer, 2011.
- [20] Emanuele Viola and Avi Wigderson. Norms, xor lemmas, and lower bounds for $gf(2)$ polynomials and multiparty protocols. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pages 141–154. IEEE, 2007.