Electronic Colloquium on Computational Complexity, Report No. 6 (2025)

Biased Linearity Testing in the 1% Regime

Subhash Khot* Kunal Mittal†

February 3, 2025

Abstract

We study linearity testing over the p-biased hypercube $(\{0,1\}^n, \mu_p^{\otimes n})$ in the 1% regime. For a distribution ν supported over $\{x \in \{0,1\}^k : \sum_{i=1}^k x_i = 0 \pmod 2\}$, with marginal distribution μ_p in each coordinate, the corresponding k-query linearity test $\mathrm{Lin}(\nu)$ proceeds as follows: Given query access to a function $f : \{0,1\}^n \to \{-1,1\}$, sample $(x_1, \ldots, x_k) \sim \nu^{\otimes n}$, query f on $x_1, \ldots, x_k$, and accept if and only if $\prod_{i\in[k]} f(x_i) = 1$.

Building on the work of Bhangale, Khot, and Minzer (STOC '23), we show, for $0 < p \leqslant \frac{1}{2}$, that if $k \geqslant 1 + \frac{1}{p}$, then there exists a distribution ν such that the test $\mathrm{Lin}(\nu)$ works in the 1% regime; that is, any function $f : \{0,1\}^n \to \{-1,1\}$ passing the test $\mathrm{Lin}(\nu)$ with probability $\geqslant \frac{1}{2} + \epsilon$, for some constant $\epsilon > 0$, satisfies $\Pr_{x\sim\mu_p^{\otimes n}}[f(x) = g(x)] \geqslant \frac{1}{2} + \delta$, for some linear function g, and a constant $\delta = \delta(\epsilon) > 0$.

Conversely, we show that if $k < 1 + \frac{1}{p}$, then no such test $\mathrm{Lin}(\nu)$ works in the 1% regime. Our key observation is that the linearity test $\mathrm{Lin}(\nu)$ works if and only if the distribution ν satisfies a certain pairwise independence property.

1 Introduction

A function $f : \{0,1\}^n \to \{-1,1\}$ is said to be linear over $\mathbb{F}_2$[1], if there exists a set $S \subseteq [n]$, such that $f(x) = \prod_{i\in S} (-1)^{x_i}$; this function is denoted by χ_S. The classical linearity testing problem, asks, given query access[2] to a function $f : \{0,1\}^n \to \{-1,1\}$, to distinguish between the following two cases[3]:

1. f is a linear function.
2. f is far from being linear; that is, for every linear function χ_S, the functions f and χ_S disagree on many points.

*Department of Computer Science, Courant Institute of Mathematical Sciences, New York University. E-mail: khot@cs.nyu.edu. Research supported by NSF Award CCF-1422159, NSF Award CCF-2130816, and the Simons Investigator Award.

†Department of Computer Science, Princeton University. E-mail: kmittal@cs.princeton.edu. Research supported by NSF Award CCF-2007462, and the Simons Investigator Award.

[1] by identifying the range $\mathbb{F}_2$ with $\{-1,1\}$, under the map $b \to (-1)^b$

[2] the algorithm is allowed to ask/query the value of $f(x)$ at any $x \in \{0,1\}^n$

[3] the algorithm is allowed to answer arbitrarily for functions f which violate both the conditions

ISSN 1433-8092

Linearity testing was first studied by Blum, Luby, and Rubinfeld, who gave a very simple 3-query test for this problem [BLR93]. This test, known as the BLR test, proceeds in the following manner: Sample $x, y \sim \{0,1\}^n$ uniformly and independently; query f at x, y, and $x \oplus y$, and accept if and only if $f(x \oplus y) = f(x) \cdot f(y)$. Observe that this test accepts all linear functions with probability 1. Blum, Luby and Rubinfeld proved that any function f passing this test with high probability ($1 - \delta$, for some small $\delta > 0$), must agree with some linear function χ_S on most (at least $1 - O(\delta)$ fraction of) points in $\{0,1\}^n$. This result, with the acceptance/agreement probability close to 1, is known as the 99%-regime of the test.

It was shown later [BCH+96, KLX10] that the above result extends to the 1% regime as well; more precisely, for every $\delta \in [0,1]$, and $f : \{0,1\}^n \to \{-1,1\}$ such that

$$\mathop{\mathbb{E}}_{x,y \sim \{0,1\}^n} [f(x) \cdot f(y) \cdot f(x \oplus y)] \geqslant \delta,$$

there exists $S \subseteq [n]$ such that $\mathbb{E}_{x \sim \{0,1\}^n} [f(x) \cdot \chi_S(x)] \geqslant \delta$.

The above test is of fundamental importance in theoretical computer science, and has several applications; for example, it is one of the ingredients in the proof of the celebrated PCP theorem [FGL+96, AS98, ALM+98]. Furthermore, the analysis of the BLR test by Bellare et al. [BCH+96] is one of the early uses of Fourier analysis over the boolean hypercube, an area which now plays a crucial role in many diverse subfields of mathematics and computer science, like complexity theory, harness of approximation, learning theory, coding theory, social choice theory, etc. [O'D14].

In this work, we are interested in the problem of linearity testing over the p-biased hypercube. For $p \in (0,1)$, we denote by μ_p the p-biased distribution on $\{0,1\}$, which assigns probability p to 1, and $1-p$ to 0. The p-biased hypercube refers the set $\{0,1\}^n$, with the n-fold product measure $\mu_p^{\otimes n}$. Linearity testing, in this p-biased setting, asks to distinguish between linear functions, and functions which are far (with respect to the p-biased measure) from being linear.

The 99% regime of this problem is well-understood [KS09, DFH19], and a simple 4-query test works in this case (see Example 4 below). The question for the 1% regime turns out to be significantly more challenging for any $p \neq 1/2$, and was wide open until a recent work of Bhangale, Khot and Minzer [BKM23b] made significant progress. In particular, for every $p \in \left(\frac{1}{3}, \frac{2}{3}\right)$, they give a 4-query test that works in the 1% regime.

Building upon the work of Bhangale, Khot and Minzer, we consider a very general class of tests, where, very roughly, some k queries $x_1, \ldots, x_k \in \{0,1\}^n$, satisfying $\sum_{i \in [k]} x_i = 0 \pmod 2$ are chosen, and the test accepts $f : \{0,1\}^n \to \{-1,1\}$ if $\prod_{i \in [k]} f(x_i) = 1$. We shall require the following definitions:

Definition 1. *(Class of Distributions) For $k \in \mathbb{N}$, $p \in (0,1)$, we define $\mathcal{D}(p,k)$ to be the class of all distributions ν on $\{0,1\}^k$ having μ_p as the marginal distribution on each coordinate $i \in [k]$, and such that $\mathrm{supp}(\nu) \subseteq \left\{x \in \{0,1\}^k : \sum_{i=1}^k x_i = 0 \pmod 2\right\}$. We say that such a distribution ν has* full even-weight support*, if the above inclusion is an equality.*

For a distribution $\nu \in \mathcal{D}(p,k)$, we say that $i \in [k]$ is a pairwise independent coordinate*, if for each $j \in [k], j \neq i$, it holds that $\mathbb{E}_{X \sim \nu} [X_i \cdot X_j] = p^2$. We say that ν is* pairwise independent*, if all its coordinates are pairwise independent.*

Definition 2. *(Class of Linearity Tests) For a distribution $\nu \in \mathcal{D}(p,k)$, we define a corresponding linearity test, denoted by* $\mathrm{Lin}(\nu)$*, as follows. Given query access to a function $f : \{0,1\}^n \to \{-1,1\}$: Sample*[4] *$x = (x_1, \ldots, x_k) \sim \nu^{\otimes n}$, and accept if and only if $f(x_1) \cdot f(x_2) \cdots f(x_k) = 1$.*

Note that every linear function passes such a test with probability 1[5]. More strongly, each query in $\nu^{\otimes n}$ having marginal distribution $\mu_p^{\otimes n}$ ensures that functions that are close to linear (with respect to p-biased measure) are also accepted with high probability; in the property testing literature, such tests are called tolerant. Furthermore, this is a very general class of linearity tests, containing many of the mentioned previously tests, as demonstrated by the following examples:

Example 3. *The BLR test uses ν to be uniform over $\left\{x \in \{0,1\}^3 : x_1 + x_2 + x_3 = 0 \pmod 2\right\}$.*

Example 4. *The 4-query p-biased test of [DFH19] (for the 99% regime) uses a distribution ν over $\{0,1\}^4$ of the following form: With probability p_0, set all coordinates to 0; with probability p_1, set all coordinates to 1; and with probability $1 - p_0 - p_1$, sample uniformly from the set $\left\{x \in \{0,1\}^4 : x_1 + x_2 + x_3 + x_4 = 0 \pmod 2\right\}$. Note that each coordinate has bias $p_1 + \frac{1}{2}(1 - p_0 - p_1)$, and p_0, p_1 are chosen so that this equals p.*

In this work, we analyze the precise conditions under which tests in Defintion 2 work for linearity testing, in the 1% regime. Our main result (proven in Section 6) is the following:

Theorem 5. *Let $p \in (0,1)$.*

1. *For every integer $k > 1 + \frac{1}{\min\{p,1-p\}}$, there exists a distribution $\nu \in \mathcal{D}(p,k)$, such that the test* $\mathrm{Lin}(\nu)$ *is a k-query linearity test over the p-biased hypercube, for the 1% regime.*

   That is, for every $\epsilon > 0$, there exists a $\delta > 0$, such that for every large $n \in \mathbb{N}$, and every function $f : \{0,1\}^n \to [-1,1]$ satisfying

$$\left| \mathop{\mathbb{E}}_{(X_1,\ldots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i\in[k]} f(X_i) \right] \right| \geqslant \epsilon,$$

   there exists a set $S \subseteq [n]$, such that $\left|\mathbb{E}_{X\sim\mu_p^{\otimes n}}\left[f(X) \cdot \chi_S(X)\right]\right| \geqslant \delta$.

2. *The above point also holds for all integers $k \geqslant 3$ with $p = \frac{1}{k-1}$, and for all even integers $k \geqslant 4$ with $p = 1 - \frac{1}{k-1}$.*

3. *Conversely, for every positive integer $k < 1 + \frac{1}{\min\{p,1-p\}}$, and every distribution $\nu \in \mathcal{D}(p,k)$, the test* $\mathrm{Lin}(\nu)$ *fails in the 1% regime.*

[4] here, by $x = (x_1, \ldots, x_k) \sim \nu^{\otimes n}$, we mean that for each $j \in [n]$, sample $(x_1^{(j)}, \ldots, x_k^{(j)}) \sim \nu$ independently (also see Section 2 for notation).

[5] When k is even, affine functions of the form $\pm\chi_S$ also pass the test with probability 1. In this work, we shall ignore the distinction between these functions and linear functions.

That is, there exists a constant $\alpha > 0$, such that for every large $n \in \mathbb{N}$, there exists a function $f : \{0,1\}^n \to \{-1,1\}$ satisfying

$$\left| \mathop{\mathbb{E}}_{(X_1,\ldots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i\in[k]} f(X_i) \right] \right| \geqslant \alpha,$$

and such that for every $S \subseteq [n]$, it holds that $\left|\mathbb{E}_{X\sim\mu_p^{\otimes n}} [f(X) \cdot \chi_S(X)]\right| \leqslant o_n(1)$.

Remark 6. *Note that the above theorem does not discuss the case when $k \geqslant 5$ is an odd integer, and $p = 1 - \frac{1}{k-1}$. This case is very interesting and is discussed in more detail in Section 6.1. Informally speaking, the test corresponding to the "natural" distribution $\nu \in \mathcal{D}(p,k)$, in this case, ensures correlation with a character of $\mathbb{Z}/(k-1)\mathbb{Z}$, and not a linear function χ_S (that is, a character of $\mathbb{Z}/2\mathbb{Z}$). In Section 6.1, we also present an alternative test to get around this.*

Next, we shall describe the main technical results we prove along the way to prove Theorem 5. We start by stating (a generalized version of) the main linearity testing result of Bhangale, Khot and Minzer [BKM23b]:

Theorem 7. *(General version proved later as Theorem 33) Let $k \geqslant 3$ be a positive integer, and let $p \in (0,1)$, $\epsilon \in (0,1]$ be constants, and let $\nu \in \mathcal{D}(p,k)$ be a distribution with full even-weight support (see Definition 1). Then, there exists constants $\delta > 0$, $d \in \mathbb{N}$ (possibly depending on k, p, ϵ, ν), such that for every large enough $n \in \mathbb{N}$, the following is true:*

Let $f : \{0,1\}^n \to [-1,1]$ be a function such that

$$\left| \mathop{\mathbb{E}}_{(X_1,\ldots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i=1}^{k} f(X_i) \right] \right| \geqslant \epsilon\,.$$

Then, there exists a set $S \subseteq [n]$, and a polynomial $g : \{0,1\}^n \to \mathbb{R}$ of degree at most d and with 2-norm $\mathbb{E}_{X\sim\mu_p^{\otimes n}} [g(X)^2] \leqslant 1$, such that

$$\left| \mathop{\mathbb{E}}_{X\sim\mu_p^{\otimes n}} [f(X) \cdot \chi_S(X) \cdot g(X)] \right| \geqslant \delta.$$

Moreover, if the distribution ν has some pairwise independent coordinate, then we may assume $g \equiv 1$; that is, f correlates with a linear function χ_S.

We remark that Bhangale, Khot and Minzer only consider the case $k = 4$, and only show $g \equiv 1$ in the case that all coordinates of ν are pairwise independent. However, their proofs extend to the more general setting of Theorem 7; we give an outline of this proof in Section 7. Furthermore, we note that we are able to analyze the linearity test for a class of distributions which is much larger than the class of full even-weight support distributions; these distributions, in some sense, *contain the BLR test*, and are formally defined in Section 7.

In the above work, the authors ask whether the conclusion $g \equiv 1$ can be obtained without assumption that ν has a pairwise independent coordinate. We show this is not possible, and in fact the assumption that ν has a pairwise independent coordinate is necessary.

Theorem 8. *(Restated and proved later as Theorem 24) Let $k \in \mathbb{N}$, $p \in (0,1)$, and let $\nu \in \mathcal{D}(p,k)$ be a distribution having no pairwise independent coordinate (see Definition 1).*

Then, there exists a constant $\alpha > 0$, such that for every large enough $n \in \mathbb{N}$, there exists a function $f : \{0,1\}^n \to [-1,1]$ such that

1. $\left|\mathbb{E}_{X\sim\nu^{\otimes n}}\left[\prod_{i=1}^{k} f(X_i)\right]\right| \geqslant \alpha$.
2. *For every $S \subseteq [n]$, it holds that* $\left|\mathbb{E}_{X\sim\mu_p^{\otimes n}}\left[f(X)\cdot\chi_S(X)\right]\right| \leqslant o_n(1)$.

Moreover, if the distribution ν is such that $\eta := \max_{i,j\in[k],i\neq j} \Pr_{X\sim\nu}[X_i = X_j] < 1$ (that is, no two coordinates are almost surely equal), the above holds for a function f with range $\{-1,1\}$.

Remark 9.

1. *The assumption $\eta < 1$ in the second part of the Theorem 8 is necessary. For example, if the i^{th} and j^{th} coordinates of ν are equal, then, for functions f with range $\{-1,1\}$, the terms $f(X_i)$ and $f(X_j)$ cancel out in the product $\mathbb{E}_{X\sim\nu^{\otimes n}}\left[\prod_{i=1}^{k} f(X_i)\right]$. In particular, the test is equivalent to the $(k-2)$-query test with coordinates i,j removed from ν, and this new distribution may possibly satisfy the conditions of Theorem 7.*
2. *The function f we construct in Theorem 8 does not correlate well with any linear function, although, as possibly required by Theorem 7, it does correlate well with some constant degree function.*
3. *The above theorem, answers in the* negative *a question of [BKM23b], who ask if*

$$\left|\mathop{\mathbb{E}}_{(X,Y,Z,W)\sim\nu^{\otimes n}}\left[g_1(X)\cdot g_2(Y)\cdot g_3(Z)\cdot g_4(W)\right]\right| = o_n(1)$$

   for distributions $\nu \in \mathcal{D}(p,4)$ with full even-weight support, and $g_1,\ldots,g_4 : \{0,1\}^n \to \mathbb{R}$ bounded, noise stable, and resilient functions.
4. *It is an easy check that the distribution ν from Example 4 cannot have a pairwise independent coordinate, unless $p = 1/2$. This shows that for $p \neq \frac{1}{2}$, simple tests that work in the 99% regime fail to work in the 1% regime.*
5. *Recall that every $\nu \in \mathcal{D}(p,k)$ satisfies $\sum_i X_i = 0 \pmod 2$ almost surely, for $X \sim \nu$. We never use this in the proof of the above theorem, and the conclusion holds without it.*

Very roughly speaking, in the proof of the above theorem we first construct a counter-example function in Gaussian space which "passes the test" with decent probability, while having zero expectation; this function is then converted to a boolean function using the Central Limit Theorem and a rounding procedure. Along the way, we prove a simple characterization for a random vector to have an independent coordinate, which we believe to be of independent interest, and is stated as follows:

Proposition 10. *(Restated formally and proved later as Proposition 19) Let $X = (X_1, \ldots, X_k)$ be a k-dimensional multivariate Gaussian random vector, such that for each $i \in [k]$, the marginal is $X_i \sim \mathcal{N}(0,1)$. Then, the following are equivalent:*

1. *For every "nice" function $f : \mathbb{R} \to \mathbb{R}$ satisfying $\mathbb{E}_{Z\sim\mathcal{N}(0,1)}[f(Z)] = 0$, it holds that $\mathbb{E}[f(X_1) \cdot f(X_2) \cdots f(X_k)] = 0$.*
2. *There exists $i \in [k]$ such that X_i is independent of $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)$.*

Finally, to use the above theorems (Theorem 7 and Theorem 8), we analyze the tradeoff between the number of queries k and the bias p, such that a distribution $\nu \in \mathcal{D}(p,k)$ with some pairwise independent coordinate exists. In particular, we prove the following (restated and proved later as Proposition 25 and Proposition 27):

Proposition 11. *Let $k \in \mathbb{N}$, $p \in (0,1)$. Then, there exists a distribution $\nu \in \mathcal{D}(p,k)$ with some pairwise independent coordinate if and only if $k \geqslant 1 + \frac{1}{\min\{p,1-p\}}$.*

We note that the above generalizes the parameter setting for both the BLR test, corresponding to $p = \frac{1}{2}$, $k = 3$, and the case of $p \in \left(\frac{1}{3}, \frac{2}{3}\right)$, $k = 4$ considered in [BKM23b].

1.1 Related work

The problem of linearity testing has been extensively studied, starting with the work of Blum, Luby and Rubinfeld [BLR93], who gave a test for the uniform distribution, in the 99% regime. The analysis of their test was later extended to the 1% regime [BCH+96, KLX10]. Tests for linearity have been also been studied in the low-randomness regime, and in the setting of non-abelian groups [BSSVW03, BoCLR08, SW06].

For the p-biased case, in the 99% regime, Halevy and Kushilevitz [HK07] gave a 3-query linearity test, that only uses random samples from the p-biased distribution! However, the test is not tolerant, makes queries that are not distributed according to $\mu_p^{\otimes n}$, and hence may reject functions that are very close to linear (with respect to the p-biased measure). Tolerant testers were analyzed later [KS09, DFH19]. More strongly, the work of Dinur, Filmus and Harsha [DFH19] gives 2^d-query tolerant tester for p-biased testing of degree d functions over $\mathbb{F}_2$, a problem which has been well studied over the uniform distribution [AKK+05, BKS+10].

As a part of their work on approximability of satisfiable constraint satisfaction problems [BKM22, BKM23a, BKM23b, BKM24a, BKM24b], Bhangale, Khot and Minzer study the p-biased version of linearity testing, in the 1% regime. As mentioned before, they give a 4-query test for $p \in \left(\frac{1}{3}, \frac{2}{3}\right)$.

David, Dinur, Goldberg, Kindler and Shinkar [DDG+17] study linearity testing on the k-slice (vectors of hamming-weight k), denoted by $L_{k,n}$, of the n-dimensional boolean hypercube, for even integers k. They show that if $f : \{0,1\}^n \to \{-1,1\}$ is such that $f(x \oplus y) = f(x)f(y)$ with probability $1 - \epsilon$ over $x, y, x \oplus y$ (conditioned on all lying in $L_{k,n}$), then f agrees with a linear function on $1 - \delta$ fraction of $L_{k,n}$, where $\delta = \delta(\epsilon) \to 0$ as $\epsilon \to 0$. In a recent work, Kalai, Lifshitz, Minzer and Ziegler [KLZM24] prove a similar result for the $n/2$-slice, in the 1% regime.

1.2 Organization of the paper

We start by presenting some preliminaries in Section 2. In Section 3, we prove a variant of Theorem 8 over the Gaussian distribution, which then is used in Section 4 to prove Theorem 8. In Section 5, we analyze the tradeoff between the bias p and the number of queries k, for the existence of a valid linearity test. Combining all results, we prove Theorem 5 in Section 6. In Section 7, we outline of the proof of Theorem 7.

2 Preliminaries

We use exp to denote the exponential function, given by $\exp(x) = e^x$ for $x \in \mathbb{R}$.

Let $\mathbb{N} = \{1, 2, \ldots\}$ be the set of natural numbers. For each $n \in \mathbb{N}$, we use $[n]$ to denote the set $\{1, 2, \ldots, n\}$. For non-negative functions $f, g : \mathbb{N} \to \mathbb{R}$, we say that $f(n) = o_n(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

For a probability distribution ν on $\mathcal{X}$, we use $\mathrm{supp}(\nu)$ to denote its support. For $n \in \mathbb{N}$, we use $\nu^{\otimes n}$ to denote the n-fold product distribution on $\mathcal{X}^n$. In particular, we shall be interested in the case when $\mathcal{X} \subseteq \mathbb{R}^k$ for some $k \in \mathbb{N}$. In this case, for vectors $x \in \mathbb{R}^{kn}$, we shall use subscripts for indices in $[k]$ and superscripts for indices in $[n]$; that is, for each $i \in [k], j \in [n]$, we use $x_i^{(j)}$ to denote the $(i,j)^{\text{th}}$ coordinate of x. Further, for each $i \in [k]$, we use x_i to denote the vector $\left(x_i^{(1)}, \ldots, x_i^{(n)}\right) \in \mathbb{R}^n$, and similarly for each $j \in [n]$, we use $x^{(j)}$ to denote the vector $\left(x_1^{(j)}, \ldots, x_k^{(j)}\right) \in \mathbb{R}^k$.

For $k \in \mathbb{N}$, let S_k denote the group of all permutations on $[k]$. For each $\pi \in S_k$, $x \in \mathbb{R}^k$, we use x_π to denote $\left(x_{\pi(1)}, \ldots, x_{\pi(k)}\right) \in \mathbb{R}^k$. With this notation, we define the symmetrization of functions over $\mathbb{R}^k$:

Definition 12. *For any function $f : \mathbb{R}^k \to \mathbb{R}$, we define its symmetrization as the function $\mathrm{Sym}(f) : \mathbb{R}^k \to \mathbb{R}$, given by $\mathrm{Sym}(f)(x) = \sum_{\pi\in S_k} f(x_\pi)$.*

We shall use the following facts from probability theory:

Fact 13. *(Chebyshev's Inequality; see [Dur19] for reference) Let X be a random variable such that $\mathbb{E}[X^2] < \infty$. Then, for any any $a > 0$,*

$$\Pr\left[|X - \mathbb{E}[X]| \geqslant a\right] \leqslant \frac{\mathrm{Var}[X]}{a^2}.$$

Fact 14. *(Hoeffding's Inequality [Hoe63]) Let $X_1, \ldots, X_n$ be independent random variables such that $a_i \leqslant X_i \leqslant b_i$ almost surely, and let $S = \sum_{i=1}^n X_i$. Then, for all $t > 0$,*

$$\Pr\left[|S - \mathbb{E}[S]| \geqslant t\right] \leqslant 2 \cdot \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

Theorem 15. *(Multivariate Central Limit Theorem; see [Dur19] for reference) Let $X^{(1)}, X^{(2)}, \ldots$ be $\mathbb{R}^k$-valued i.i.d. random vectors, with mean zero, and finite a covariance matrix $\Sigma \in \mathbb{R}^{k\times k}$*

given by $\Sigma_{i,j} = \mathbb{E}\left[X_i^{(1)} \cdot X_j^{(1)}\right]$. *If* $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X^{(i)}$, *then,* $S_n \xrightarrow{\mathcal{D}} Z$ *as* $n \to \infty$, *for* $Z \sim \mathcal{N}(0, \Sigma)$. *That is, for every bounded continuous function* $H : \mathbb{R}^k \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}\left[H(S_n)\right] = \mathop{\mathbb{E}}_{Z\sim\mathcal{N}(0,\Sigma)}\left[H(Z)\right].$$

We shall also use the following fact about zeros of polynomials:

Lemma 16. *Let* $p_1, \ldots, p_r : \mathbb{R}^k \to \mathbb{R}$ *be non-zero polynomials. Then, there exists* $y \in \mathbb{R}^k$ *such that for each permutation* $\pi \in S_k$, *and each* $i \in [r]$, *it holds that* $p_i(y_\pi) \neq 0$.

Proof Sketch. The zero-set of any non-zero polynomial has measure zero, with respect to the Lebesgue measure on $\mathbb{R}^k$. Hence, by sub-additivity, the set of points $y \in \mathbb{R}^k$ violating the statement of the lemma is of measure zero as well. □

Next, we give some basic results about the probabilist's Hermite polynomials. The reader is referred to Chapter 11 in [O'D14] for more details.

Definition 17. *The Hermite polynomials* $(H_j)_{j\in\mathbb{Z}_{\geqslant 0}}$ *are univariate polynomials, with* H_j *a monic polynomial of degree* j, *satisfying the power series expression*

$$\exp\left(tx - \frac{t^2}{2}\right) = \sum_{j=0}^{\infty} \frac{1}{j!} \cdot H_j(x) \cdot t^j, \quad \text{for } t, x \in \mathbb{R}.$$

Note that the series above is absolutely convergent, with $\sum_{j=0}^{\infty} \frac{1}{j!} \cdot |H_j(x)| \cdot |t|^j \leqslant \exp\left(|t| \cdot |x| + \frac{t^2}{2}\right)$ *for each* $t, x \in \mathbb{R}$.

Lemma 18. *Let* $k \in \mathbb{N}$ *and* $s_1, s_2, \ldots, s_k \in \mathbb{Z}_{\geqslant 0}$, *and let* $\Sigma \in \mathbb{R}^{k\times k}$ *be a positive semi-definite matrix such that* $\Sigma_{i,i} = 1$ *for each* i. *For* $V = \Sigma - I$, *it holds that*

$$\mathop{\mathbb{E}}_{X\sim\mathcal{N}(0,\Sigma)}\left[H_{s_1}(X_1) \cdots H_{s_k}(X_k)\right] = \frac{s_1! \cdots s_k!}{d! \cdot 2^d} \cdot \left[\left(t^\top V\ t\right)^d : t_1^{s_1} \cdots t_k^{s_k}\right]$$

where $s_1 + \cdots + s_k = 2d$, *and* $\left[\left(t^\top V\ t\right)^d : t_1^{s_1} \cdots t_k^{s_k}\right]$ *denotes the coefficient of* $t_1^{s_1} \cdots t_k^{s_k}$ *in the polynomial* $\left(t^\top V\ t\right)^d$. *Also, the above expectation is zero when* $s_1 + \cdots + s_k$ *is odd.*

Proof. Recall that the moment generating function of a multivariate Gaussian distribution is given by

$$\mathop{\mathbb{E}}_{X\sim\mathcal{N}(0,\Sigma)}\left[\exp\left(t_1X_1 + \ldots t_kX_k\right)\right] = \exp\left(\frac{1}{2} \cdot t^\top \Sigma\ t\right),$$

for each $t \in \mathbb{R}^k$. Multiplying the above by $\exp(-\frac{1}{2} \cdot t^\top t)$, and plugging in the power series in

Definition 17, we get for each $t \in \mathbb{R}^k$ that

$$\begin{aligned}
\sum_{d=0}^{\infty} \frac{1}{d! \cdot 2^d} \cdot \left(t^\top V\ t\right)^d &= \exp\left(\frac{1}{2} \cdot t^\top V\ t\right) \\
&= \mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[\exp\left(\left(t_1 X_1 - \frac{t_1^2}{2}\right) + \cdots + \left(t_k X_k - \frac{t_k^2}{2}\right)\right)\right] \\
&= \mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[\left(\sum_{s_1=0}^{\infty} \frac{1}{s_1!} \cdot H_{s_1}(X_1) \cdot t_1^{s_1}\right) \cdots \left(\sum_{s_k=0}^{\infty} \frac{1}{s_k!} \cdot H_{s_k}(X_k) \cdot t_k^{s_k}\right)\right] \\
&= \sum_{s_1,\ldots,s_k \geqslant 0} \frac{t_1^{s_1} \cdots t_k^{s_k}}{s_1! \cdots s_k!} \cdot \mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[H_{s_1}(X_1) \cdots H_{s_k}(X_k)\right].
\end{aligned}$$

Note that since the power series in Definition 17 is absolutely convergent, all steps above of interchanging limits and expectations are valid by the dominated convergence theorem. Finally, comparing coefficients, we have the desired result. □

3 A Gaussian Variant

The first step towards proving Theorem 8 is to prove a Gaussian variant, stated below:

Proposition 19. *Let $k \in \mathbb{N}$, and let $\Sigma \in \mathbb{R}^{k \times k}$ be a symmetric positive semi-definite matrix such that:*

1. *For each $i \in [k]$, it holds that $\Sigma_{i,i} = 1$.*
2. *The matrix $V = \Sigma - I$ has no row/column as all zeros.*

Then, there exists a Lipschitz continuous function $f : \mathbb{R} \to [-1, 1]$ such that:

$$\mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,1)} [f(X)] = 0, \quad \text{and} \quad \left| \mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[\prod_{i \in [k]} f(X_i)\right] \right| > 0.$$

3.1 Symmetric Powers of Polynomials

Before we prove the above proposition, we first prove a lemma about (symmetrization of) powers of multivariate polynomials. We show that if a polynomial $q(x_1, \ldots, x_k)$ depends on all the variables $x_1, \ldots, x_k$, then some power $\mathrm{Sym}(q^d)$ (see Definition 12) contains a monomial divisible by $x_1 x_2 \cdots x_k$.

Lemma 20. *Let $k \in \mathbb{N}$, and let $q : \mathbb{R}^k \to \mathbb{R}$ be a polynomial such that for each $i \in [k]$, the polynomial $\ell_i = \partial_i q$ is not identically zero. Then, there exists some $d \in \mathbb{N}$, and positive integers $s_1, \ldots, s_k \in \mathbb{N}$ such that the coefficient of $x_1^{s_1} \cdot x_2^{s_2} \cdots x_k^{s_k}$ in the polynomial $\mathrm{Sym}(q^d)$ is non-zero.*

We start by proving the following lemma about derivates of powers of q.

Lemma 21. *Let $k \in \mathbb{N}$, and let $q : \mathbb{R}^k \to \mathbb{R}$ be a polynomial. For each $i \in [k]$, let $\ell_i = \partial_i q$.*

Then, for every $s = (s_1, \ldots, s_k) \in \mathbb{Z}^k_{\geqslant 0}$ with $|s| = \sum_{i\in[k]} s_i$, there exist polynomials $p_0, \ldots, p_{|s|}$, with $p_{|s|} = \prod_{i\in[k]} \ell_i^{s_i}$, such that for each $d \geqslant |s|$, it holds that

$$\partial_1^{s_1} \cdot \partial_2^{s_2} \cdots \partial_k^{s_k} \left(q^d\right) = q^{d-|s|} \cdot \left(\sum_{i=0}^{|s|} d^i \cdot p_i \right).$$

Proof. The proof is by induction on $|s|$. For the base case, if $|s| = 0$, we have $s = (0, 0, \ldots, 0)$, and $p_0 = 1$ satisfies the statement of the lemma.

For the inductive step, consider any $s = (s_1, \ldots, s_k) \in \mathbb{Z}^k_{\geqslant 0}$ with $|s| = \sum_{i\in[k]} s_i > 0$. Without loss of generality, by symmetry, we can assume that $s_1 > 0$. By the inductive hypothesis applied to $(s_1 - 1, s_2 \ldots, s_k)$, we have the existence of polynomials $p_0, \ldots, p_{|s|-1}$, with $p_{|s|-1} = \ell_1^{s_1-1} \cdot \prod_{i=2}^k \ell_i^{s_i}$, and such that for each $d \geqslant |s| - 1$, we have

$$\partial_1^{s_1-1} \cdot \partial_2^{s_2} \cdots \partial_k^{s_k} \left(q^d\right) = q^{d-|s|+1} \cdot \left(\sum_{i=0}^{|s|-1} d^i \cdot p_i \right).$$

Now, if $d \geqslant |s|$, differentiating the above with respect to x_1, we get

$$\begin{aligned}
\partial_1^{s_1} \cdot \partial_2^{s_2} \cdots \partial_k^{s_k} \left(q^d\right) &= q^{d-|s|} \cdot \left((d - |s| + 1) \cdot \ell_1 \cdot \sum_{i=0}^{|s|-1} d^i \cdot p_i \right) + q^{d-|s|+1} \cdot \left(\sum_{i=0}^{|s|-1} d^i \cdot \partial_1(p_i) \right) \\
&= q^{d-|s|} \cdot \left(\sum_{i=1}^{|s|} d^i \cdot \ell_1 \cdot p_{i-1} + \sum_{i=0}^{|s|-1} d^i \cdot ((-|s| + 1) \cdot \ell_1 \cdot p_i + q \cdot \partial_1 p_i) \right) \\
&= q^{d-|s|} \cdot \left(\sum_{i=0}^{|s|} d^i \cdot \tilde{p}_i \right),
\end{aligned}$$

where the polynomials $\tilde{p}_1, \ldots, \tilde{p}_{|s|}$ do not depend on d, and are such that $\tilde{p}_{|s|} = p_{|s|-1} \cdot \ell_1 = \prod_{i\in[k]} \ell_i^{s_i}$, as desired. □

With the above lemma in hand, next we shall consider the symmetrization operation applied to derivatives of powers of q.

Lemma 22. *Let $k \in \mathbb{N}$, and let $q : \mathbb{R}^k \to \mathbb{R}$ be a polynomial such that for each $i \in [k]$, the polynomial $\ell_i = \partial_i q$ is not identically zero.*

Then, for each large enough even integer $d \in \mathbb{N}$, the polynomial $\mathrm{Sym}\left(\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(q^d\right)\right)$ *is not identically zero.*

Proof. By applying Lemma 21 on $s = (2, 2, \ldots, 2)$, we have the existence of polynomials $p_0, \ldots, p_{2k}$, with $p_{2k} = \prod_{i\in[k]} \ell_i^2$, such that for each $d \geqslant 2k$, it holds that $\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(q^d\right) = q^{d-2k} \cdot \left(\sum_{i=0}^{2k} d^i \cdot p_i\right)$.

By Lemma 16, let $y \in \mathbb{R}^k$ be such that y (and its permutations) don't lie in the zero set of any of the polynomials $\ell_1, \ldots, \ell_k, q$. We define

$$A = \min_{\pi \in S_k} \left[\prod_{i \in [k]} \ell_i \left(y_\pi\right)^2 \right] > 0, \quad B = \max_{0 \leqslant i \leqslant 2k-1,\ \pi \in S_k} |p_i(y_\pi)| \geqslant 0.$$

Then, for any even integer $d \geqslant \max\left\{2k, \frac{4kB}{A}\right\}$, it holds that

$$\begin{aligned}
\operatorname{Sym}\left(\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(q^d\right)\right)(y) &= \sum_{\pi \in S_k} q\left(y_\pi\right)^{d-2k} \cdot \left(d^{2k} \cdot \prod_{i \in [k]} \ell_i \left(y_\pi\right)^2 + \sum_{i=0}^{2k-1} d^i \cdot p_i\left(y_\pi\right) \right) \\
&\geqslant \sum_{\pi \in S_k} q\left(y_\pi\right)^{d-2k} \cdot \left(d^{2k} \cdot A - \sum_{i=0}^{2k-1} d^i \cdot B \right) \\
&\geqslant \left(\sum_{\pi \in S_k} q\left(y_\pi\right)^{d-2k} \right) \cdot \left(d^{2k} \cdot A - 2k \cdot d^{2k-1} \cdot B \right) \\
&\geqslant \left(\sum_{\pi \in S_k} q\left(y_\pi\right)^{d-2k} \right) \cdot \frac{d^{2k} A}{2} > 0.
\end{aligned}$$

Hence, for even integers $d \geqslant \max\left\{2k, \frac{4kB}{A}\right\}$, the polynomial $\operatorname{Sym}\left(\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(q^d\right)\right)$ is not identically zero. □

Finally, we prove the main lemma of this section.

Proof of Lemma 20. Let $k \in \mathbb{N}$, and let $q : \mathbb{R}^k \to \mathbb{R}$ be a polynomial such that for each $i \in [k]$, the polynomial $\ell_i = \partial_i q$ is not identically zero. It suffices to prove that for some $d \in \mathbb{N}$, the polynomial $\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(\operatorname{Sym}(q^d)\right)$ is not identically zero, since then the coefficient of some monomial divisible by $x_1^2 \cdot x_2^2 \cdots x_k^2$ is non-zero.

For each polynomial $p : \mathbb{R}^k \to \mathbb{R}$, and each $\pi \in S_k$, we shall use p_π to denote the polynomial given by $p_\pi(x) = p(x_\pi)$. Then, for all $s_1, \ldots, s_k \in \mathbb{Z}_{\geqslant 0}$, we have that $\partial_1^{s_1} \cdot \partial_2^{s_2} \cdots \partial_k^{s_k}\left(p_\pi\right) = \left(\partial_{\pi^{-1}(1)}^{s_1} \cdot \partial_{\pi^{-1}(2)}^{s_2} \cdots \partial_{\pi^{-1}(k)}^{s_k}\left(p\right)\right)_\pi$.

By the above, we have that for each $d \in \mathbb{N}$,

$$\begin{aligned}
\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(\operatorname{Sym}(q^d)\right) &= \partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(\sum_{\pi \in S_k} q_\pi^d \right) \\
&= \sum_{\pi \in S_k} \partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(q_\pi^d\right) \\
&= \sum_{\pi \in S_k} \left(\partial_{\pi^{-1}(1)}^2 \cdot \partial_{\pi^{-1}(2)}^2 \cdots \partial_{\pi^{-1}(k)}^2 \left(q^d\right)\right)_\pi \\
&= \sum_{\pi \in S_k} \left(\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(q^d\right)\right)_\pi \\
&= \operatorname{Sym}\left(\partial_1^2 \cdot \partial_2^2 \cdots \partial_k^2 \left(q^d\right)\right).
\end{aligned}$$

Now, the result follows from Lemma 22. □

3.2 Proving the Gaussian Variant

We start by proving a slight variant of Proposition 19, where we allow f to be an arbitrary (possibly unbounded) polynomial.

Lemma 23. *Let $k \in \mathbb{N}$, and let $\Sigma \in \mathbb{R}^{k \times k}$ be a symmetric positive semi-definite matrix such that:*

1. *For each $i \in [k]$, it holds that $\Sigma_{i,i} = 1$.*
2. *The matrix $V = \Sigma - I$ has no row/column as all zeros.*

Then, there exists a polynomial $f : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}_{X \sim \mathcal{N}(0,1)}[f(X)] = 0$, and

$$\left| \mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[\prod_{i \in [k]} f(X_i) \right] \right| > 0.$$

Proof. For $s = (s_1, \ldots, s_k) \in \mathbb{N}^k$ and $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k$, let $f_{s,\alpha} : \mathbb{R} \to \mathbb{R}$ be the polynomial defined by $f_{s,\alpha}(x) = \alpha_1 H_{s_1}(x) + \cdots + \alpha_k H_{s_k}(x)$, where the polynomials H_{s_i} are Hermite polynomials (see Definition 17). Observe that since $s_1, \ldots, s_k \geqslant 1$, this polynomial satisfies $\mathbb{E}_{X \sim \mathcal{N}(0,1)}[f(X)] = 0$.

Suppose, for the sake of contradiction, that for every $s \in \mathbb{N}^k$, $\alpha \in \mathbb{R}^k$, it holds that

$$\mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[\prod_{i \in [k]} f_{s,\alpha}(X_i) \right] = \mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[\prod_{i \in [k]} \sum_{j \in [k]} \alpha_j H_{s_j}(X_i) \right] = 0.$$

Observe that for every $s \in \mathbb{N}^k$, the above expression can be written as a multivariate polynomial in $\alpha_1, \ldots, \alpha_k$. If the polynomial vanishes for all $\alpha \in \mathbb{R}^k$, the coefficient of $\alpha_1 \cdot \alpha_2 \cdots \alpha_k$ must be zero; that is,

$$\sum_{\pi \in S_k} \mathop{\mathbb{E}}_{X \sim \mathcal{N}(0,\Sigma)} \left[\prod_{i \in [k]} H_{s_{\pi(i)}}(X_i) \right] = 0.$$

Now, applying Lemma 18, we get that for each $d \in \mathbb{N}$, and each $s_1, \ldots, s_k \geqslant 1$ with $s_1 + \cdots + s_k = 2d$,

$$\sum_{\pi \in S_k} \left[\left(t^\top V\ t\right)^d : t_1^{s_{\pi(1)}} \cdots t_k^{s_{\pi(k)}} \right] = \sum_{\pi \in S_k} \left[\left(t_\pi^\top V\ t_\pi\right)^d : t_1^{s_k} \cdots t_k^{s_k} \right] = \left[\mathrm{Sym}\left(\left(t^\top V\ t\right)^d \right) : t_1^{s_k} \cdots t_k^{s_k} \right] = 0.$$

Note that the assumption that V has no zero row/column implies that for every $i \in [k]$, the polynomial $\partial_i \left(t^\top V\ t\right)$ is not identically zero. By Lemma 20, this is a contradiction. □

With the above, we now prove Proposition 19 via a standard truncation argument.

Proof of Proposition 19. By Lemma 23, we know that there exists a polynomial $f : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}_{X \sim \mathcal{N}(0,1)}[f(X)] = 0$, and $\left| \mathbb{E}_{X \sim \mathcal{N}(0,\Sigma)} \left[\prod_{i \in [k]} f(X_i) \right] \right| > 0$.

For each integer $M \in \mathbb{N}$, we define the truncated function $f_M : \mathbb{R} \to [-M, M]$ by

$$f_M(x) = f(x) \cdot \mathbb{1}_{|f(x)| \leqslant M} + M \cdot \mathbb{1}_{f(x) > M} - M \cdot \mathbb{1}_{f(x) < -M}.$$

Also, let $g_M : \mathbb{R} \to [-2M, 2M]$, be given by $g_M(x) = f_M(x) - \mathbb{E}_{X \sim \mathcal{N}(0,1)}[f_M(X)]$. Observe that

1. For every M, it holds that $\mathbb{E}_{X \sim \mathcal{N}(0,1)}[g_M(X)] = 0$.
2. For every M, the function g_M is bounded and Lipschitz continuous.
3. For every $x \in \mathbb{R}$, $f_M(x) \to f(x)$ as $M \to \infty$. Further, since $|f_M(x)| \leqslant |f(x)|$ for each $x \in \mathbb{R}, M \in \mathbb{N}$, by the dominated convergence theorem, we have $\mathbb{E}_{X \sim \mathcal{N}(0,1)}[f_M(x)] \to \mathbb{E}_{X \sim \mathcal{N}(0,1)}[f(x)] = 0$ as $M \to \infty$. This implies that for each $x \in \mathbb{R}$, $g_M(x) \to f(x)$ as $M \to \infty$.

   Also, for each $x \in \mathbb{R}, M \in \mathbb{N}$, we have $|g_M(x)| \leqslant |f(x)| + \mathbb{E}_{X \sim \mathcal{N}(0,1)}[|f(X)|]$. Hence, by the dominated convergence theorem, we have that $\mathbb{E}_{X \sim \mathcal{N}(0,\Sigma)}\left[\prod_{i \in [k]} g_M(X_i)\right] \to \mathbb{E}_{X \sim \mathcal{N}(0,\Sigma)}\left[\prod_{i \in [k]} f(X_i)\right] \neq 0$ as $M \to \infty$.

By the above, for some large enough M, the function $\frac{1}{2M} \cdot g_M : \mathbb{R} \to [-1, 1]$ satisfies the desired properties. □

4 Linearity Testing Requires Pairwise Independence

In this section, we prove Theorem 8, which is restated below.

Theorem 24. *Let $k \in \mathbb{N}$, $p \in (0, 1)$, and let $\nu \in \mathcal{D}(p, k)$ be a distribution having no pairwise independent coordinate (see Definition 1). Then, there exists a constant $\alpha > 0$, such that for every large enough $n \in \mathbb{N}$, there exists a function $f : \{0, 1\}^n \to [-1, 1]$ such that*

1. $\left|\mathbb{E}_{X \sim \nu^{\otimes n}}\left[\prod_{i=1}^{k} f(X_i)\right]\right| \geqslant \alpha$.
2. *For every $S \subseteq [n]$, it holds that $\left|\mathbb{E}_{X \sim \mu_p^{\otimes n}}[f(X)\chi_S(X)]\right| \leqslant o_n(1)$.*

Moreover, if the distribution ν is such that $\eta := \max_{i,j \in [k], i \neq j} \Pr_{X \sim \nu}[X_i = X_j] < 1$ (that is, no two coordinates are almost surely equal), the above holds for a function f with range $\{-1, 1\}$.

The remainder of this section is devoted to the proof of Theorem 24. In Section 4.1, we prove the first part of the theorem, dealing with functions with range $[-1, 1]$. Then, in Section 4.2, we show how to round to functions with range $\{-1, 1\}$.

4.1 Function with Range $[-1, 1]$

Let $k \in \mathbb{N}$, $p \in (0,1)$, and let $\nu \in \mathcal{D}(p,k)$ be a distribution having no pairwise independent coordinate. Let $\Sigma \in \mathbb{R}^{k \times k}$ be the (normalized) covariance matrix corresponding to the distribution ν, given by, $\Sigma_{i,j} = \mathbb{E}_{X \sim \nu}\left[\frac{(X_i - p)\cdot(X_j - p)}{p - p^2}\right]$. Observe that the matrix Σ satisfies the conditions of Proposition 19, and hence there exists a function $h : \mathbb{R} \to [-1,1]$ such that

1. $\mathbb{E}_{Z \sim \mathcal{N}(0,1)}[h(Z)] = 0$
2. The function $H : \mathbb{R}^k \to [-1,1]$ given by $H(x) = \prod_{i \in [k]} h(x_i)$ is such that
$$\alpha := \frac{1}{2} \cdot \left| \mathop{\mathbb{E}}_{Z \sim \mathcal{N}(0,\Sigma)} [H(Z)] \right| > 0.$$
3. The function h is K-Lipschitz for some $K > 0$; in particular, both h and H are bounded continuous functions.

Consider any large $n \in \mathbb{N}$. We define $f : \{0,1\}^n \to [-1,1]$ by
$$f(x) = h\left(\frac{1}{\sqrt{n}} \cdot \sum_{j=1}^{n} \frac{x^{(j)} - p}{\sqrt{p - p^2}}\right),$$

The function f satisfies the two properties in the theorem statement, as follows:

- Let $X \sim \nu^{\otimes n}$, and let $Y = (Y_1, \ldots, Y_k)$ be a $\{0,1\}^k$-valued random vector, defined as $Y_i = \frac{1}{\sqrt{n}} \cdot \sum_{j=1}^{n} \frac{X_i^{(j)} - p}{\sqrt{p - p^2}}$.

  Let $F : \{0,1\}^{kn} \to [-1,1]$ be given by $F(x) = \prod_{i \in [k]} f(x_i)$. Since H is continuous and bounded, we have by the Multivariate CLT (Theorem 15) that
$$\left| \mathbb{E}[F(X)] - \mathop{\mathbb{E}}_{Z \sim \mathcal{N}(0,\Sigma)} [H(Z)] \right| = \left| \mathbb{E}[H(Y)] - \mathop{\mathbb{E}}_{Z \sim \mathcal{N}(0,\Sigma)} [H(Z)] \right| \leqslant o_n(1).$$
  Hence, for large n, we get $\left|\mathbb{E}_{X \sim \nu^{\otimes n}}\left[\prod_{i=1}^{k} f(X_i)\right]\right| \geqslant 2\alpha - o_n(1) \geqslant \alpha$, as desired.

- Consider any subset $S \subseteq [n]$, and let $T \subseteq S$ be any subset of size $|T| = \min\left\{\lfloor n^{1/4} \rfloor, |S|\right\}$. Let $\tilde{f} : \{0,1\}^n \to [-1,1]$ be defined by $\tilde{f}(X) = h\left(\frac{1}{\sqrt{n - |T|}} \cdot \sum_{j \in [n] \setminus T} \frac{x^{(j)} - p}{\sqrt{p - p^2}}\right)$; note that this function only depends on the coordinates of x outside the set T. Further, for each

$x \in \{0,1\}^n$, by the Lipschitz bound on h, we get

$$\begin{aligned}\left|f(x) - \tilde{f}(x)\right| &\leqslant K \cdot \left| \frac{1}{\sqrt{n}} \cdot \sum_{j=1}^{n} \frac{x^{(j)} - p}{\sqrt{p-p^2}} - \frac{1}{\sqrt{n-|T|}} \cdot \sum_{j \in [n] \setminus T} \frac{x^{(j)} - p}{\sqrt{p-p^2}} \right| \\ &\leqslant \frac{K}{\sqrt{p-p^2}} \cdot \left(\frac{|T|}{\sqrt{n}} + (n-|T|) \cdot \left| \frac{1}{\sqrt{n-|T|}} - \frac{1}{\sqrt{n}} \right| \right) \\ &\leqslant \frac{K}{\sqrt{p-p^2}} \cdot \left(\frac{|T|}{\sqrt{n}} + \frac{n-|T|}{\sqrt{n}} \cdot \frac{|T|}{n} \right) \\ &\leqslant \frac{K}{\sqrt{p-p^2}} \cdot \frac{2\,|T|}{\sqrt{n}} = o_n(1),\end{aligned}$$

where we used that $(1-t)^{-1/2} \leqslant 1 + t$ for each $t \in [0, 1/2]$.

Now, for $X \sim \mu_p^{\otimes n}$, we have

$$\begin{aligned}\left| \mathop{\mathbb{E}}_{X} [f(X) \cdot \chi_S(X)] \right| &\leqslant \left| \mathop{\mathbb{E}}_{X} \left[\tilde{f}(X) \cdot \chi_S(X) \right] \right| + o_n(1) \\ &= \left| \mathop{\mathbb{E}}_{X} \left[\tilde{f}(X) \cdot \chi_{S \setminus T}(X) \right] \cdot \mathop{\mathbb{E}}_{X} [\chi_T(X)] \right| + o_n(1) \\ &= \left| \mathop{\mathbb{E}}_{X} \left[\tilde{f}(X) \cdot \chi_{S \setminus T}(X) \right] \right| \cdot |1-2p|^{|T|} + o_n(1).\end{aligned}$$

If $|S| \geqslant \lfloor n^{1/4} \rfloor$, then $|1-2p|^{|T|} = o_n(1)$. Otherwise, we have that $S = T$, and by the Central Limit Theorem (see Theorem 15) , the first term in the above product equals

$$\left| \mathop{\mathbb{E}}_{X} \left[\tilde{f}(X) \right] \right| = \left| \mathop{\mathbb{E}}_{X} \left[\tilde{f}(X) \right] - \mathop{\mathbb{E}}_{Z \sim \mathcal{N}(0,1)} [h(Z)] \right| = o_n(1). \qquad \square$$

4.2 Rounding to a Function with Range $\{-1, 1\}$

Now, we shall prove the second part of Theorem 24.

Let $k \in \mathbb{N}$, $p \in (0,1)$, and let $\nu \in \mathcal{D}(p,k)$ be a distribution having no pairwise independent coordinate. Further suppose that the distribution ν is such that

$$\eta := \max_{i,j \in [k], i \neq j} \Pr_{X \sim \nu} [X_i = X_j] < 1.$$

Let $\alpha > 0$ be as obtained in Section 4.1. Consider any large $n \in \mathbb{N}$, and let $f : \{0,1\}^n \to [-1,1]$ be the function obtained in Section 4.1.

Let $g : \{0,1\}^n \to \{-1,1\}$ be a random function, defined as $g(x) = \begin{cases} 1, & w.p.\ \frac{1+f(x)}{2} \\ -1, & w.p.\ \frac{1-f(x)}{2} \end{cases}$, independently for each $x \in \{0,1\}^n$. Observe that this satisfies $\mathbb{E}_g [g(x)] = f(x)$ for each $x \in \{0,1\}^n$. We will show that the function g satisfies the two desired properties with probability $1 - o_n(1)$, and hence by the probabilistic method, this guarantees the existence of a non-random g as desired. This is done as follows:

1. Let $F, G : \{0,1\}^{kn} \to [-1,1]$ be defined as $F(x) = \prod_{i\in[k]} f(x_i)$ and $G(x) = \prod_{i\in[k]} g(x_i)$. Let $X, Y \sim \nu^{\otimes n}$ be independent (of each other and of g) and let E be the event that $X_1, \ldots, X_k, Y_1, \ldots, Y_k$ are all distinct. Then, by a union bound, we have that $\Pr\left[\bar{E}\right] \leqslant 2 \cdot \binom{k}{2} \cdot \eta^n + k^2 \cdot \left(p^2 + (1-p)^2\right)^n = o_n(1)$, and hence

$$\begin{aligned}\left| \mathop{\mathbb{E}}_{g} \mathop{\mathbb{E}}_{X\sim\nu^{\otimes n}} [G(X)] - \mathop{\mathbb{E}}_{X\sim\nu^{\otimes n}} [F(X)] \right| &\leqslant \Pr\left[\bar{E}\right] + \left| \mathop{\mathbb{E}}_{g} \mathop{\mathbb{E}}_{X,Y} [G(X)\cdot \mathbb{1}_E] - \mathop{\mathbb{E}}_{X} [F(X)] \right| \\ &\leqslant \Pr\left[\bar{E}\right] + \left| \mathop{\mathbb{E}}_{X,Y} [F(X)\cdot \mathbb{1}_E] - \mathop{\mathbb{E}}_{X} [F(X)] \right| \\ &\leqslant 2\Pr\left[\bar{E}\right] = o_n(1).\end{aligned}$$

   Similarly, we have

$$\begin{aligned}\left| \mathop{\mathbb{E}}_{g} \left[\mathop{\mathbb{E}}_{X} [G(X)]\right]^2 - \left[\mathop{\mathbb{E}}_{X} [F(X)]\right]^2 \right| &= \left| \mathop{\mathbb{E}}_{g} \mathop{\mathbb{E}}_{X,Y} [G(X)\cdot G(Y)] - \mathop{\mathbb{E}}_{X,Y} [F(X)\cdot F(Y)] \right| \\ &\leqslant 2\Pr\left[\bar{E}\right] = o_n(1).\end{aligned}$$

   Letting $\beta = \left|\mathbb{E}_X [F(X)]\right| \geqslant \alpha$, we get $\mathrm{Var}_g\left[\mathbb{E}_X [G(X)]\right] \leqslant \beta^2 + o_n(1) - (\beta - o_n(1))^2 = o_n(1)$. Hence, by Chebyshev's inequality (Fact 13), we have $\left|\mathbb{E}_X [G(X)]\right| \geqslant \frac{\alpha}{2}$ with probability $1 - o_n(1)$.

2. Fix $S \subseteq [n]$. Let $X \sim \mu_p^{\otimes n}$, and let $W = \mathbb{E}_X [\chi_S(X)\cdot g(X)] = \sum_{x\in\{0,1\}^n} \Pr[X = x] \cdot \chi_S(x)\cdot g(x)$. Observe that W is a sum of 2^n independent and bounded random variables, and such that $\mathbb{E}_g[W] = \mathbb{E}_X [\chi_S(X)\cdot f(X)]$. For $q = \max\{p, 1-p\} < 1$, it holds that $\sum_x (2\Pr[X = x])^2 \leqslant 4q^n \cdot \sum_x \Pr[X = x] = 4q^n$, and by Hoeffding's inequality (Fact 14), we have for each $t > 0$ that

$$\Pr\left[|W - \mathbb{E}[W]| \geqslant t\right] \leqslant 2 \cdot \exp\left(-\frac{2t^2}{4q^n}\right).$$

   Let $t = q^{n/4}$. Then, with probability at least $1 - o_n(2^{-n})$, it holds that $|W| = \left|\mathbb{E}_X [\chi_S(X)\cdot g(X)]\right| \leqslant \left|\mathbb{E}_X [\chi_S(X)\cdot f(X)]\right| + q^{n/4} = o_n(1)$.

   Now, a union bound over $S \subseteq [n]$ shows that with probability $1 - o_n(1)$, the above holds for every $S \subseteq [n]$. □

5 Queries vs. Bias Tradeoff

In this section, we analyze the relation between p (the bias) and k (the number of queries) for the existence of a distribution $\nu \in \mathcal{D}(p,k)$ with some pairwise independent coordinate, and with full even-weight support (see Definition 1).

5.1 Query Lower Bound

We prove a lower bound on k in terms of the p, as follows:

Proposition 25. *Let $k \in \mathbb{N}$, $p \in (0,1)$, and let $\nu \in \mathcal{D}(p,k)$ be a distribution that has some pairwise independent coordinate. Then, it holds that $k \geqslant 3$ and $\frac{1}{k-1} \leqslant p \leqslant 1 - \frac{1}{k-1}$.*

Proof. Let $X \sim \nu$, and let $i \in [k]$ be a pairwise independent coordinate under ν.

For $Z = \sum_{j \neq i} X_j$, we have by linearity of expectation, that $\mathbb{E}[X_i \cdot Z] = (k-1)p^2$. On the other hand, observe that if $X_i = 1$, then $Z = 1 \pmod 2$ and so $Z \geqslant 1$. Hence,

$$p = \mathbb{E}[X_i \cdot 1] \leqslant \mathbb{E}[X_i \cdot Z] = (k-1)p^2,$$

and we have $(k-1)p \geqslant 1$; in particular, this shows $k \geqslant 3$.

For the upper bound on p, we consider the following cases:

- k is odd: In this case, if $X_i = 1$, then $Z = 1 \pmod 2$ and so $Z \leqslant k-2$. Hence,

  $$(k-1)p^2 = \mathbb{E}[X_i \cdot Z] \leqslant \mathbb{E}[X_i \cdot (k-2)] = p(k-2),$$

  and we have $(k-1)p \leqslant (k-2)$, as desired.

- k is even: In this case, observe that the distribution of the random variable $(1 - X_1, \ldots, 1 - X_k)$ also satisfies the hypothesis of the proposition, with p replaced by $1-p$. Hence, the above proof gives us $(k-1) \cdot (1-p) \geqslant 1$, as desired. □

Remark 26. *The proof of Proposition 25 also shows that for $k > 3$ and $p \in \left\{\frac{1}{k-1}, 1 - \frac{1}{k-1}\right\}$, any distribution satisfying the assumptions of Proposition 25 cannot have full even-weight support. This is because if $p \in \left\{\frac{1}{k-1}, 1 - \frac{1}{k-1}\right\}$, in all cases in the above proof, the random variable Z must be constant under some value of X_i (either $X_i = 0$ or $X_i = 1$); this cannot be the case for a distribution with full even-weight support when $k > 3$.*

5.2 Query Upper Bound

In this subsection, we shall prove the following proposition.

Proposition 27. *Let $k \geqslant 3$ be a positive integer, and let $p \in \left[\frac{1}{k-1}, 1 - \frac{1}{k-1}\right]$ (note that this interval is non-empty for $k \geqslant 3$).*

Then, there exists a permutation-invariant[6] and pairwise independent distribution $\nu(k,p) \in \mathcal{D}(p,k)$ (see Definition 1). Furthermore, if $k = 3$ or if $p \notin \left\{\frac{1}{k-1}, 1 - \frac{1}{k-1}\right\}$, then there exists such a distribution with full even-weight support.

The proof involves various cases, considered below in Lemma 28 and Lemma 29.

Lemma 28. *Let $k \geqslant 4$ be a positive integer, and let $p \in \left[\frac{1}{k-1}, \frac{2}{k-1}\right) \cup \left(1 - \frac{2}{k-1}, 1 - \frac{1}{k-1}\right]$ (note that this interval is contained in $\left[\frac{1}{k-1}, 1 - \frac{1}{k-1}\right]$ for $k \geqslant 4$). Then, there exists a pairwise independent distribution $\nu(k,p) \in \mathcal{D}(p,k)$.*

Moreover, if $p \notin \left\{\frac{1}{k-1}, 1 - \frac{1}{k-1}\right\}$, then there exists such a distribution with full even-weight support.

[6] we say that a distribution ν over $\{0,1\}^k$ is *permutation-invariant*, if for $X = (X_1, \ldots, X_k) \sim \nu$, and any permutation $\pi : [k] \to [k]$, the distribution of $(X_{\pi(1)}, \ldots, X_{\pi(k)})$ is the same as ν.

Proof. Let $k \geqslant 4$ be a positive integer, and let $p \in \left[\frac{1}{k-1}, \frac{2}{k-1}\right) \cup \left(1-\frac{2}{k-1}, 1-\frac{1}{k-1}\right]$. Let $s = \lfloor \frac{k}{2} \rfloor$; we shall exhibit a vector $q = (q_0, q_1, \ldots, q_s) \in [0,1]^{s+1}$ satisfying:

$$\sum_{i=0}^{s} \binom{k}{2i} \cdot q_i = 1, \quad \sum_{i=1}^{s} \binom{k-1}{2i-1} \cdot q_i = p, \quad \sum_{i=1}^{s} \binom{k-2}{2i-2} \cdot q_i = p^2.$$

The distribution $\nu(p,k)$ is then defined by assigning probability $\begin{cases} q_{|x|/2}, & |x| = 0 \pmod 2 \\ 0, & |x| = 1 \pmod 2 \end{cases}$ to the point $x \in \{0,1\}^k$, where $|x| = \sum_{i=1}^{k} x_i$. Note that the above properties correspond to $\nu(k,p)$ being a valid probability distribution supported on even-hamming-weight vectors, having marginals μ_p, and pairwise independent coordinates. Furthermore, the distribution $\nu(p,k)$ has full even-weight support if and only if each $q_i \in (0,1)$.

The vector q is defined as follows in different cases (for brevity, we omit the verification of the above properties):

1. $k \geqslant 5$ is odd, $p \in \left[\frac{1}{k-1}, \frac{2}{k-1}\right)$: Let $q_0 = 1 + \frac{kp^2}{2} - \frac{k^2 p}{2(k-1)}$, $q_1 = \frac{(k-2)p-(k-1)p^2}{(k-1)(k-3)}$, $q_{(k-1)/2} = \frac{(k-1)p^2-p}{(k-1)(k-3)}$, and zero otherwise.

2. $k \geqslant 5$ is odd, $1-p \in \left[\frac{1}{k-1}, \frac{2}{k-1}\right)$: Let $q_0 = 1 + \frac{kp^2}{k-3} - \frac{k(2k-5)p}{(k-1)(k-3)}$, $q_{(k-3)/2} = \frac{3(k-2)p-3(k-1)p^2}{(k-1)(k-2)(k-3)}$, $q_{(k-1)/2} = \frac{(k-1)p^2-(k-4)p}{2(k-1)}$, and zero otherwise.

3. $k \geqslant 4$ is even, $p \in \left[\frac{1}{k-1}, \frac{2}{k-1}\right)$: Let $q_0 = \frac{(k-1)p^2-(k+1)p+2}{2}$, $q_1 = \frac{p-p^2}{k-2}$, $q_{k/2} = \frac{(k-1)p^2-p}{k-2}$, and zero otherwise.

4. $k \geqslant 4$ is even, $1-p \in \left[\frac{1}{k-1}, \frac{2}{k-1}\right)$: In this case, we define $\nu(k,p)$ to be the distribution obtained by flipping each coordinate of $\nu(k, 1-p)$.

Next, we show that if $p \notin \left\{\frac{1}{k-1}, 1-\frac{1}{k-1}\right\}$, then such a distribution $\nu(p,k)$ with full even-weight support exists. We only need to do this for the first three cases, as the procedure described in the fourth case preserves the property of full even-weight support.

The same argument applies in all cases, and we present it for the first case: that is when $k \geqslant 5$ is odd, and $p \in \left(\frac{1}{k-1}, \frac{2}{k-1}\right)$. We observe if $p \neq \frac{1}{k-1}$, each of the probabilities $q_0, q_1, q_{(k-1)/2}$ above lie in the interval $(0,1)$. Now, consider the equations

$$\sum_{i=0}^{s} \binom{k}{2i} \cdot \tilde{q}_i = 0, \quad \sum_{i=1}^{s} \binom{k-1}{2i-1} \cdot \tilde{q}_i = 0, \quad \sum_{i=1}^{s} \binom{k-2}{2i-2} \cdot \tilde{q}_i = 0.$$

In these equations, the variables $\tilde{q}_0, \tilde{q}_1, \tilde{q}_{(k-1)/2}$ are linearly independent, and hence, there exists a vector $\tilde{q} \in \mathbb{R}^{s+1}$ satisfying these equations, which has all coordinates equal to 1, other than possibly $\tilde{q}_0, \tilde{q}_1, \tilde{q}_{(k-1)/2}$. Then, for some small $\delta > 0$, the vector $q + \delta \cdot \tilde{q}$ has all coordinates in $(0,1)$, and satisfies the required properties. □

Lemma 29. *Let $k \geqslant 6$ be a positive integer, and let $p \in \left[\frac{2}{k-1}, 1-\frac{2}{k-1}\right] \setminus \left\{\frac{1}{2}\right\}$ (note that this interval is non-empty for $k \geqslant 6$). There, there exists a pairwise independent distribution $\nu(k,p) \in \mathcal{D}(p,k)$ with full even-weight support.*

Proof. Let $k \geqslant 6$ be a positive integer, and let $p \in \left[\frac{2}{k-1}, 1-\frac{2}{k-1}\right]$, $p \neq \frac{1}{2}$. That is, for $q = \min\{p, 1-p\} < \frac{1}{2}$, we have $k \geqslant 1 + \frac{2}{q}$. Let ℓ be the smallest odd integer satisfying $\ell > 1 + \frac{1}{q} > 3$. Note that this satisfies $4 \leqslant \ell \leqslant 3 + \frac{1}{q} < 1 + \frac{2}{q} \leqslant k$, and we have $q \in \left(\frac{1}{\ell-1}, \frac{2}{\ell-1}\right)$.

By Lemma 28, there exist pairwise independent distributions $\nu(\ell, p)$ and $\nu(\ell, 1-p)$, with full even-weight support. Let $\tilde{\nu}_0 = \nu(\ell, p)$, and let $\tilde{\nu}_1$ be the distribution obtained by flipping each coordinate of $\nu(\ell, 1-p)$. Since ℓ is odd, for each $b \in \{0,1\}$, it holds that $\tilde{\nu}_b$ has pairwise independent coordinates, each with marginal μ_p, and such that $\mathrm{supp}(\tilde{\nu}_b) = \left\{x \in \{0,1\}^k : \sum_{i=1}^k x_i = b \pmod 2\right\}$. Finally, we define $X \sim \nu(k,p)$ via the following random process: Let $(X_{\ell+1}, \ldots, X_k) \sim \mu_p^{\otimes(k-\ell)}$, and with $Z = \sum_{i=\ell+1}^k X_i \pmod 2$, we let $(X_1, \ldots, X_\ell) \sim \tilde{\nu}_Z$. It is an easy check that this distribution satisfies the required properties. □

Finally, we prove Proposition 27.

Proof of Proposition 27. Note that it suffices to find such a distribution that is not necessarily permutation invariant, since averaging the distribution over all permutations preserves pairwise independence and full even-weight support.

If $p = 1/2$, for any $k \geqslant 3$, we let $\nu(k,p)$ be the uniform distribution on the set $\left\{x \in \{0,1\}^k : \sum_{i=1}^k x_i = 0 \pmod 2\right\}$.

Now, for $k = 3$, it must hold that $p = 1/2$, in which case $\nu(k,p)$ is as above. For $k = 4$ or $k = 5$, and $p \neq 1/2$, it must hold that $p \in \left[\frac{1}{k-1}, \frac{2}{k-1}\right) \cup \left(1-\frac{2}{k-1}, 1-\frac{1}{k-1}\right]$, and the result follows from Lemma 28. For $k \geqslant 6$ and $p \neq \frac{1}{2}$, the result follows from Lemma 28 and Lemma 29. □

6 Putting Everything Together

We are now ready to prove our main result.

Proof of Theorem 5. Let $p \in (0,1)$.

1. Consider any positive integer $k > 1 + \frac{1}{\min\{p,1-p\}} \geqslant 3$ (or $k = 3$ with $p = \frac{1}{2}$). By Proposition 27, there exists a pairwise independent distribution $\nu \in \mathcal{D}(p,k)$ with full even-weight support. The result now follows by Theorem 33.

2. Suppose that $k \geqslant 3$ with $p = \frac{1}{k-1}$, or $k \geqslant 4$ is even with $p = 1 - \frac{1}{k-1}$. In these cases, we observe that the distribution $\nu \in \mathcal{D}(p,k)$ constructed in Lemma 28 is pairwise independent, and contains BLR (see Definition 32):

   (a) If $k \geqslant 3, p = \frac{1}{k-1}$, the distribution ν contains all vectors in $\{0,1\}^k$ of hamming-weights 0 and 2 in its support. In this case, Definition 32 is satisfied with $\tilde{b} = 0$ and $\tilde{z}$ as the all-zeros vector.

   (b) If $k \geqslant 4$ is even, and $p = 1 - \frac{1}{k-1}$, the distribution ν contains all vectors in $\{0,1\}^k$ of hamming-weights $k-2$ and k in its support. In this case, Definition 32 is satisfied with $\tilde{b} = 1$ and $\tilde{z}$ as the all-ones vector.

The result now follows by Theorem 33.

3. Suppose that $k < 1+\frac{1}{\min\{p,1-p\}}$ is a positive integer, and let $\nu \in \mathcal{D}(p,k)$. We perform the following operation on the distribution ν: if $i,j \in [k],\ i \neq j$ are such that $\Pr_{X\sim\nu}[X_i = X_j] = 1$, we remove coordinates i,j from ν, and repeat until no such pairs remain.

   Finally, we are left with a distribution $\tilde{\nu}$ on $\tilde{k} \leqslant k$ coordinates. We consider the following two cases:

   (a) Suppose that $\tilde{k} = 0$. In this case, for every $n \in \mathbb{N}$, and every $f : \{0,1\}^n \to \{-1,1\}$, it holds that $\mathbb{E}_{X\sim\nu^{\otimes n}}\left[\prod_{i=1}^{k} f(X_i)\right] = 1$, since the k terms in the product cancel out in pairs. Hence, it suffices to show the existence of a function $f : \{0,1\}^n \to \{-1,1\}$ satisfying $\left|\mathbb{E}_{X\sim\mu_p^{\otimes n}}\left[f(X)\cdot\chi_S(X)\right]\right| \leqslant o_n(1)$ for every $S \subseteq [n]$. Note that a (uniformly) random function $f : \{0,1\}^n \to \{-1,1\}$ satisfies this with high probability, by an argument similar to the one at the end of Section 4.2 (a random function can be thought of as rounding the constant zero function as in Section 4.2).

   (b) Now, suppose that $\tilde{k} \neq 0$. Then, it holds that $\tilde{\nu} \in \mathcal{D}(p,\tilde{k})$, and by Proposition 25, we have that $\tilde{\nu}$ has no pairwise independent coordinate. Now, by Theorem 24 there exists a constant $\alpha > 0$, such that for every large $n \in \mathbb{N}$, there exists a function $f : \{0,1\}^n \to \{-1,1\}$ such that

$$\left|\mathop{\mathbb{E}}_{(X_1,\ldots,X_k)\sim\nu^{\otimes n}}\left[\prod_{i\in[k]} f(X_i)\right]\right| = \left|\mathop{\mathbb{E}}_{(X_1,\ldots,X_{\tilde{k}})\sim\tilde{\nu}^{\otimes n}}\left[\prod_{i\in[\tilde{k}]} f(X_i)\right]\right| \geqslant \alpha,$$

   and such that $\left|\mathbb{E}_{X\sim\mu_p^{\otimes n}}\left[f(X)\cdot\chi_S(X)\right]\right| \leqslant o_n(1)$ for every $S \subseteq [n]$. □

6.1 A Corner Case

In the above proof, we leave the case of odd $k \geqslant 5$ and $p = 1 - \frac{1}{k-1}$. This turns out to be very interesting, and we discuss it next. For the remainder of this section, we fix such a k and p.

In this case, the pairwise independent distribution $\nu \in \mathcal{D}(p,k)$ constructed in Lemma 28, is supported on vectors of hamming weights 0 and $k-1$ (and does not contain BLR as in Definition 32). In particular, for every $x \in \mathrm{supp}(\nu)$, it holds that $\sum_{i=1}^{k} x_i = 0 \pmod{k-1}$. For this reason, as we show next, the best we can expect from the test $\mathrm{Lin}(\nu)$, is to guarantee correlation with a character over $\mathbb{Z}/(k-1)\mathbb{Z}$, and this is indeed true.

Definition 30. *(Characters over $\mathbb{Z}/(k-1)\mathbb{Z}$) Let ω be a primitive $(k-1)^{th}$ root of unity. For every $0 \leqslant r \leqslant k-2$, we define the function $\phi_r : \{0,1\} \to \mathbb{C}$ as $\phi_r(x) = \omega^{rx}$.*

For every $n \in \mathbb{N}$, and every integers $0 \leqslant r^{(1)},\ldots,r^{(n)} \leqslant k-2$, we define the product character $\phi_{r^{(1)},\ldots,r^{(n)}} : \{0,1\}^n \to \mathbb{C}$ by $\phi_{r^{(1)},\ldots,r^{(n)}}(x) = \prod_{j=1}^{n} \phi_{r^{(j)}}(x^{(j)}) = \omega^{\sum_{j=1}^{n} r^{(j)}x^{(j)}}$.

Now, consider the test $\mathrm{Lin}(\nu)$. Observe that any character $f = \phi_{r^{(1)},\ldots,r^{(n)}}$ passes this test with probability 1:

$$\mathop{\mathbb{E}}_{X\sim\nu^{\otimes n}}\left[\prod_{i\in[k]} f(X_i)\right] = \prod_{j=1}^{n} \mathop{\mathbb{E}}_{Y\sim\nu}\left[\prod_{i\in[k]} \phi_{r^{(j)}}(Y_i)\right] = \prod_{j=1}^{n} \mathop{\mathbb{E}}_{Y\sim\nu}\left[\omega^{r^{(j)}\cdot\left(\sum_{i\in[k]} Y_i\right)}\right] = 1.$$

Next, we claim that characters explain the success of $\mathrm{Lin}(\nu)$ for any function f:

Theorem 31. *For every constant $\epsilon > 0$, there exists a constant $\delta > 0$ such that for every large enough $n \in \mathbb{N}$, the following is true:*

Let $f : \{0,1\}^n \to [-1,1]$ be a function such that $\left|\mathbb{E}_{X\sim\nu^{\otimes n}}\left[\prod_{i=1}^{k} f(X_i)\right]\right| \geqslant \epsilon$. Then, there exist integers $0 \leqslant r^{(1)},\ldots,r^{(n)} \leqslant k-2$, such that

$$\left|\mathop{\mathbb{E}}_{X\sim\mu_p^{\otimes n}}\left[f(X)\cdot\phi_{r^{(1)},\ldots,r^{(n)}}(X)\right]\right| \geqslant \delta.$$

Proof. The result follows from the work of Bhangale, Khot, Liu and Minzer [BKLM24a, BKLM24b], and we omit the details. Very roughly speaking, the proof follows a similar strategy as in Section 7: first show that f has good correlation with a character under random restrictions; then, use this to show that f has good correlation with character times a low-degree function; finally, use that ν is pairwise independent to get rid of the low-degree function. □

Finally, we present an alternative solution to deal with this corner case of odd $k \geqslant 5$ and $p = 1 - \frac{1}{k-1}$. Instead of the test $\mathrm{Lin}(\nu)$, we can perform the following test:

Let $f : \{0,1\}^n \to [-1,1]$, and let $\nu' \in \mathcal{D}(1-p,k) = \mathcal{D}(\frac{1}{k-1},k)$ be the pairwise independent distribution from Lemma 28.

1. Sample $X = (X_1,\ldots,X_k) \sim \nu'^{\otimes n}$.
2. Let X' be the vector obtained by negating each of the kn coordinates of X.
3. Query f on $X'_1,\ldots,X'_k$ and accept if and only if $\prod_{i\in[k]} f(X'_i) = 1$.

Each query X'_i of the above test is distributed according to $\mu_p^{\otimes n}$, and the analysis of the test simply follows from the analysis for $\mathrm{Lin}(\nu')$ in Theorem 5. The drawback here, though, is that the test does not accept all linear functions with probability 1, but only functions of the form $(-1)^{|S|}\cdot\chi_S$, for $S \subseteq [n]$.

7 Analysis of the Linearity Test

In this section, we shall state and prove a generalized version of Theorem 7. The proof follows the work of Bhangale, Khot and Minzer [BKM23b], and hence we only give a rough outline (skipping many of the technical points), pointing out the places where the proof differs from the above work. We start with the following definition:

Definition 32. *Let $k \geqslant 3, p \in (0,1)$, and let $\nu \in \mathcal{D}(p,k)$ be a distribution. We say that ν* contains BLR*, if there exists some $\tilde{b} \in \{0,1\}$, $\tilde{z} \in \{0,1\}^{k-3}$, such that*

$$\left\{(x_1,\ x_2,\ x_1 \oplus x_2 \oplus \tilde{b},\ \tilde{z}) : x_1, x_2 \in \{0,1\}\right\} \subseteq \mathrm{supp}(\nu) \subseteq \{0,1\}^k .$$

Furthermore, for technical reasons, we shall also require that

$$\mathrm{span}_{\mathbb{F}_2}(\mathrm{supp}(\nu)) = \left\{ x \in \{0,1\}^k : \sum_{i=1}^{k} x_i = 0 \pmod 2 \right\}.$$

Observe that any ν with full even-weight support contains BLR (with $\tilde{b} = 0$, and $\tilde{z}$ the all-zeros vector). With this, we state the following generalization of Theorem 7:

Theorem 33. *Let $k \geqslant 3$ be a positive integer, and let $p \in (0,1)$, $\epsilon \in (0,1]$ be constants, and let $\nu \in \mathcal{D}(p,k)$ be a distribution containing BLR (see Definition 32). Then, there exists constants $\delta > 0$, $d \in \mathbb{N}$ (possibly depending on k, p, ϵ, ν), such that for every large enough $n \in \mathbb{N}$, the following is true:*

Let $f : \{0,1\}^n \to [-1,1]$ be a function such that

$$\left| \mathop{\mathbb{E}}_{(X_1,\ldots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i=1}^{k} f(X_i) \right] \right| \geqslant \epsilon .$$

Then, there exists a set $S \subseteq [n]$, and a polynomial $g : \{0,1\}^n \to \mathbb{R}$ of degree at most d and with 2-norm $\mathbb{E}_{X\sim\mu_p^{\otimes n}}[g(X)^2] \leqslant 1$, such that

$$\left| \mathop{\mathbb{E}}_{X\sim\mu_p^{\otimes n}} [f(X) \cdot \chi_S(X) \cdot g(X)] \right| \geqslant \delta.$$

Moreover, if the distribution ν has some pairwise independent coordinate, then we may assume $g \equiv 1$; that is, f correlates with a linear function χ_S.

The remainder of this section is devoted to the proof of the above theorem. Let $k \geqslant 3$ be an integer, and let $p \in (0,1)$, $\epsilon \in (0,1]$ be constants, and let $\nu \in \mathcal{D}(p,k)$ be a distribution containing BLR (see Definition 32). Also, let $f : \{0,1\}^n \to [-1,1]$ be a function such that

$$\left| \mathop{\mathbb{E}}_{X=(X_1,\ldots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i=1}^{k} f(X_i) \right] \right| \geqslant \epsilon . \tag{1}$$

Step 1: Large Fourier Coefficient under Random Restriction.

We note that the proof of this step is where we differ from [BKM23b].

Since the distribution $\nu \in \mathcal{D}(p,k)$ contains BLR, we can write $\nu = (1-\beta)\cdot\nu' + \beta\cdot\mu$, for some small constant $0 < \beta < \frac{1}{2}\min\{p, 1-p\}$, some distribution ν' over $\{0,1\}^k$, and with μ the uniform distribution over $\left\{(x_1, x_2, x_1 \oplus x_2 \oplus \tilde{b}, \tilde{z}) : x_1, x_2 \in \{0,1\}\right\}$, where $\tilde{b}, \tilde{z}$ are as in Definition 32. Using this, we can describe choosing $X \sim \nu^{\otimes n}$ as the following two step

process. First choose a set $I \subseteq [n]$, denoted $I \sim_{1-\beta} [n]$, by choosing $i \in I$ with probability $1-\beta$, independently for each $i \in [n]$. Then, choose $Z \sim \nu'^{\otimes I}$ and $Y \sim \mu^{\bar{I}}$, and set $X = (Y, Z)$.

With the above, we can prove that the function f satisfies the property of having a large fourier coefficient under random restrictions; the reader is referred to [O'D14] for an introduction to Fourier analysis over the hypercube.

Lemma 34. *With $\delta = \epsilon/2$, it holds that*

$$\Pr_{I\sim_{1-\beta}[n],\ Z\sim\nu'^{\otimes I}}\left[\exists S \subseteq [n]\setminus I :\ \left|\widehat{f_{I\to Z_1}}(S)\right| \geqslant \delta\ \right] \geqslant \delta.$$

Here, $f_{I\to Z_1}$ refers to the restriction of the function f, with the variables in I set to Z_1.

Proof. By Equation 1, we have

$$\begin{aligned}
\epsilon &\leqslant \left| \mathop{\mathbb{E}}_{X=(X_1,\ldots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i=1}^{k} f(X_i)\right] \right| \\
&= \left| \mathop{\mathbb{E}}_{I\sim_{1-\beta}[n],\ Z\sim\nu'^{\otimes I}} \mathop{\mathbb{E}}_{Y\sim\mu^{\otimes\bar{I}}} \left[\prod_{i=1}^{k} f_{I\to Z_i}(Y_i)\right] \right| \\
&\leqslant \mathop{\mathbb{E}}_{I\sim_{1-\beta}[n],\ Z\sim\nu'^{\otimes I}} \left| \mathop{\mathbb{E}}_{Y\sim\mu^{\otimes\bar{I}}} \left[\prod_{i=1}^{k} f_{I\to Z_i}(Y_i)\right] \right|
\end{aligned}$$

Observe that in the above expression, the random variables $Y_4, \ldots, Y_k$ are constants (determined by $\tilde{z}$). Now, using a (classical) Fourier analytic argument to analyze the BLR linearity test over the uniform distribution (see Chapter 1 of [O'D14]), we get

$$\begin{aligned}
\epsilon &\leqslant \mathop{\mathbb{E}}_{I\sim_{1-\beta}[n],\ Z\sim\nu'^{\otimes I}} \left| \mathop{\mathbb{E}}_{Y\sim\mu^{\otimes\bar{I}}} \left[\prod_{i=1}^{3} f_{I\to Z_i}(Y_i)\right] \right| \\
&= \mathop{\mathbb{E}}_{I\sim_{1-\beta}[n],\ Z\sim\nu'^{\otimes I}} \left| \sum_{S\subseteq\bar{I}} \widehat{f_{I\to Z_1}}(S)\cdot\widehat{f_{I\to Z_2}}(S)\cdot\widehat{f_{I\to Z_3}}(S)\cdot(-1)^{\tilde{b}\cdot|S|} \right| \\
&\leqslant \mathop{\mathbb{E}}_{I\sim_{1-\beta}[n],\ Z\sim\nu'^{\otimes I}} \left[\max_{S\subseteq\bar{I}} \left|\widehat{f_{I\to Z_1}}(S)\right|\right] \\
&\leqslant \Pr_{I\sim_{1-\beta}[n],\ Z\sim\nu'^{\otimes I}} \left[\exists S \subseteq \bar{I} :\ \left|\widehat{f_{I\to Z_1}}(S)\right| \geqslant \epsilon/2\right] + \epsilon/2.
\end{aligned}$$

□

Step 2: Direct Product Test

Using Theorem 1.1 in [BKM23b], by Lemma 34 we get the existence of constants $d \in \mathbb{N}, \delta' > 0$, a set $S \subseteq [n]$, and a polynomial $g : \{0,1\}^n \to \mathbb{R}$ of degree at most d, and with 2-norm $\mathbb{E}_{X\sim\mu_p^{\otimes n}}[g(X)^2] \leqslant 1$, such that

$$\left| \mathop{\mathbb{E}}_{X\sim\mu_p^{\otimes n}} [f(X)\cdot\chi_S(X)\cdot g(X)] \right| \geqslant \delta'.$$

This proves the first part of Theorem 7. It remains to show that if ν has some pairwise independent coordinate, it is possible to remove the function g in the above expression.

Step 3: List Decoding.

This step follows Section 4.2 and Section 4.3 in [BKM23b].

Using an iterative list-decoding process, we can find a constant $r \in \mathbb{N}$, and functions $\chi_{S_1}, \dots, \chi_{S_r}$, and constant degree polynomials $g_1, \dots, g_r$, such that it is possible to "replace" f by $\sum_{i \in [r]} \chi_{S_i} \cdot g_i$ in Equation 1 (and lose at most some constant factor in ϵ). Now, this implies that for some constant $\epsilon' > 0$, and some indices $j_1, \dots, j_k \in [r]$, we have

$$\left| \mathop{\mathbb{E}}_{(X_1,\dots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i=1}^{k} \chi_{S_{j_i}}(X_i) g_{j_i}(X_i) \right] \right| \geqslant \epsilon'. \tag{2}$$

We remark that for the next step, some extra structure on S_{j_i}'s is needed, and ensuring that it holds requires the condition on $\mathrm{span}_{\mathbb{F}_2}(\mathrm{supp}(\nu))$ in Definition 32.

Step 4: Invariance Principle Argument.

This step follows Section 4.4, Section 4.5, and Section 4.6 in [BKM23b].

Assume, for the sake of contradiction, that f is not correlated well with any χ_S; that is, $\mathbb{E}_{X\sim\mu_p^{\otimes n}}[f(X)\cdot\chi_S(X)] \leqslant o_n(1)$ for each $S \subseteq [n]$. Using this, it can be shown, roughly, that for each $i \in [k]$, the expectation $\mathbb{E}_{X\sim\mu_p^{\otimes n}}\left[\chi_{S_{j_i}}(X) g_{j_i}(X)\right] \leqslant o_n(1)$; note that for this conclusion to hold, we might have to modify S_{j_i}'s and g_{j_i}'s, however it is possible to do so while maintaining Equation 2.

Now, by an invariance principle argument [MOO10, Mos10, Mos20], very roughly, it is possible to replace the expectation in Equation 2 over $(X_1, \dots, X_k) \sim \nu^{\otimes n}$, by an expectation over $(Z_1, \dots, Z_k) \sim \mathcal{N}(0, \Sigma)^{\otimes n}$, where $\Sigma \in \mathbb{R}^{k\times k}$ is the (normalized) covariance matrix of ν. Finally, we use that some coordinate X_{i^*} is pairwise independent of each X_i, for $i \neq i^*$. Since the Gaussian distribution is determined by its covariance matrix, this implies that Z_{i^*} is mutually independent of $(Z_i)_{i\neq i^*}$. We have

$$\begin{aligned}
\epsilon' &\leqslant \left| \mathop{\mathbb{E}}_{X=(X_1,\dots,X_k)\sim\nu^{\otimes n}} \left[\prod_{i=1}^{k} \chi_{S_{j_i}}(X_i) g_{j_i}(X_i) \right] \right| \\
&\approx \left| \mathop{\mathbb{E}}_{Z=(Z_1,\dots,Z_k)\sim\mathcal{N}(0,\Sigma)^{\otimes n}} \left[\prod_{i=1}^{k} \chi_{S_{j_i}}(Z_i) g_{j_i}(Z_i) \right] \right| \\
&\approx \left| \mathop{\mathbb{E}}_{Z_{i^*}\sim\mathcal{N}(0,1)^{\otimes n}} \left[\chi_{S_{j_{i^*}}}(Z_{i^*}) g_{j_{i^*}}(Z_{i^*}) \right] \right| \cdot \left| \mathop{\mathbb{E}}_{Z} \left[\prod_{i\in[k], i\neq i^*} \chi_{S_{j_i}}(Z_i) g_{j_i}(Z_i) \right] \right| \\
&\approx \left| \mathop{\mathbb{E}}_{X_{i^*}\sim\mu_p^{\otimes n}} \left[\chi_{S_{j_{i^*}}}(X_{i^*}) g_{j_{i^*}}(X_{i^*}) \right] \right| \cdot \left| \mathop{\mathbb{E}}_{Z} \left[\prod_{i\in[k], i\neq i^*} \chi_{S_{j_i}}(Z_i) g_{j_i}(Z_i) \right] \right| \\
&\leqslant o_n(1),
\end{aligned}$$

which is a contradiction. □

Acknowledgements

We thank Amey Bhangale, Yang P. Liu, and Dor Minzer for discussions that helped this project. Amey and Dor politely declined to be co-authors.

References

[AKK+05] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing Reed-Muller codes. *IEEE Trans. Inform. Theory*, 51(11):4032–4039, 2005. 6

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. 2

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: a new characterization of NP. *J. ACM*, 45(1):70–122, 1998. 2

[BCH+96] Mihir Bellare, Don Coppersmith, Johan Håstad, Marcos Kiwi, and Madhu Sudan. Linearity testing in characteristic two. *IEEE Trans. Inform. Theory*, 42(6, part 1):1781–1795, 1996. (also in SFCS 1995). 2, 6

[BKLM24a] Amey Bhangale, Subhash Khot, Yang P. Liu, and Dor Minzer. On approximability of satisfiable k-CSPs: VI. 2024. Available at https://arxiv.org/pdf/2411.15133. 21

[BKLM24b] Amey Bhangale, Subhash Khot, Yang P. Liu, and Dor Minzer. On approximability of satisfiable k-CSPs: VII. 2024. Available at https://arxiv.org/pdf/2411.15136. 21

[BKM22] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-CSPs: I. In *STOC*, pages 976–988, 2022. 6

[BKM23a] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-CSPs: II. In *STOC*, pages 632–642, 2023. 6

[BKM23b] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-CSPs: III. In *STOC*, pages 643–655, 2023. 2, 4, 5, 6, 21, 22, 23, 24

[BKM24a] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-CSPs: IV. In *STOC*, pages 1423–1434, 2024. 6

[BKM24b] Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-CSPs: V. *Electron. Colloquium Comput. Complex.*, TR24-129, 2024. 6

[BKS+10] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. Optimal testing of Reed-Muller codes. In *FOCS*, pages 488–497, 2010. 6

[BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *J. Comput. System Sci.*, 47(3):549–595, 1993. (also in STOC 1990). 2, 6

[BoCLR08] Michael Ben-or, Don Coppersmith, Mike Luby, and Ronitt Rubinfeld. Non-abelian homomorphism testing, and distributions close to their self-convolutions. *Random Structures Algorithms*, 32(1):49–70, 2008. (also in APPROX-RANDOM 2004). 6

[BSSVW03] Eli Ben-Sasson, Madhu Sudan, Salil Vadhan, and Avi Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *STOC*, pages 612–621, 2003. 6

[DDG+17] Roee David, Irit Dinur, Elazar Goldenberg, Guy Kindler, and Igor Shinkar. Direct sum testing. *SIAM J. Comput.*, 46(4):1336–1369, 2017. (also in ITCS 2015). 6

[DFH19] Irit Dinur, Yuval Filmus, and Prahladh Harsha. Analyzing Boolean functions on the biased hypercube via higher-dimensional agreement tests. In *SODA*, pages 2124–2133, 2019. 2, 3, 6

[Dur19] Rick Durrett. *Probability—theory and examples*. Cambridge University Press, Cambridge, 2019. Fifth edition. 7

[FGL+96] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996. 2

[HK07] Shirley Halevy and Eyal Kushilevitz. Distribution-free property-testing. *SIAM J. Comput.*, 37(4):1107–1138, 2007. (also in APPROX-RANDOM 2003, 2005). 6

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963. 7

[KLX10] Tali Kaufman, Simon Litsyn, and Ning Xie. Breaking the ϵ-soundness bound of the linearity test over GF(2). *SIAM J. Comput.*, 39(5):1988–2003, 2010. (also in APPROX-RANDOM 2008). 2, 6

[KLZM24] Gil Kalai, Noam Lifshitz, Tamar Ziegler, and Dor Minzer. A dense model theorem for the boolean slice. In *FOCS*, 2024. (to appear). 6

[KS09] Swastik Kopparty and Shubhangi Saraf. Tolerant linearity testing and locally testable codes. In *APPROX-RANDOM*, pages 601–614. 2009. 2, 6

[MOO10] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Ann. of Math. (2)*, 171(1):295–341, 2010. 24

[Mos10] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geom. Funct. Anal.*, 19(6):1713–1756, 2010. 24

[Mos20] Elchanan Mossel. Gaussian bounds for noise correlation of resilient functions. *Israel J. Math.*, 235(1):111–137, 2020. 24

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014. 2, 8, 23

[SW06] Amir Shpilka and Avi Wigderson. Derandomizing homomorphism testing in general groups. *SIAM J. Comput.*, 36(4):1215–1230, 2006. (also in STOC 2004). 6

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il