Optimal Testing of Reed-Muller Codes

Arnab Bhattacharyya*, Swastik Kopparty†, Grant Schoenebeck‡, Madhu Sudan§, David Zuckerman¶
*Computer Science and Artificial Intelligence Laboratory
MIT, Cambridge, MA
Email: abhatt@mit.edu
†Computer Science and Artificial Intelligence Laboratory
MIT, Cambridge, MA
Email: swastik@mit.edu
‡Department of Computer Science
University of California-Berkeley
Email: grant@cs.berkeley.edu
§Microsoft Research
Cambridge MA
Email: madhu@mit.edu
¶Computer Science Department
University of Texas at Austin
Austin TX
Email: diz@cs.utexas.edu

***Abstract*—We consider the problem of testing if a given function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is close to any degree d polynomial in n variables, also known as the Reed-Muller testing problem. Alon et al. [1] proposed and analyzed a natural 2^{d+1}-query test for this problem. This test turned out to be intimately related to the Gowers norm. Alon et. al. showed that this test accepts every degree d polynomial with probability 1, while it rejects functions that are $\Omega(1)$-far with probability $\Omega(1/(d2^d))$. We give an asymptotically optimal analysis of this test, and show that it rejects functions that are (even only) $\Omega(2^{-d})$-far with $\Omega(1)$-probability (so the rejection probability is a universal constant independent of d and n). This implies a tight relationship between the $(d+1)^{\text{st}}$-Gowers norm of a function and its maximal correlation with degree d polynomials, when the correlation is close to 1.**

Our proof works by induction on n and yields a new analysis of even the classical Blum-Luby-Rubinfeld [2] linearity test, for the setting of functions mapping $\mathbb{F}_2^n$ to $\mathbb{F}_2$. The optimality follows from a tighter analysis of counterexamples to the "inverse conjecture for the Gowers norm" constructed by [3], [4].

Our result has several implications. First, it shows that the Gowers norm test is tolerant, in that it also accepts close codewords. Second, it improves the parameters of an XOR lemma for polynomials given by Viola and Wigderson [5]. Third, it implies a "query hierarchy" result for property testing of affine-invariant properties. That is, for every function $q(n)$, it gives an affine-invariant property that is testable with $O(q(n))$-queries, but not with $o(q(n))$-queries, complementing an analogous result of [6] for graph properties.

***Keywords*-Sublinear-time algorithms, Property testing, Reed-Muller codes, Gowers norm, Low-degree tests**

AB: Work was partially supported by a DOE Computational Science Graduate Fellowship and NSF Awards 0514771, 0728645, and 0732334.

SK: Work was partially done while author was a summer intern at Microsoft Research New England and partially supported by NSF Grant CCF-0829672.

GS: Work was partially done while author was a summer intern at Microsoft Research New England and partially supported by a National Science Foundation Graduate Fellowship.

DZ: Work was partially done while the author consulted at Microsoft Research New England, and partially supported by NSF Grants CCF-0634811 and CCF-0916160.

I. Introduction

Can the proximity of a function to a low-degree polynomial be estimated by sampling the function in few places? Variants of this question have been studied in two different communities for different purposes.

A. Gowers norm

In the additive combinatorics community, this issue arose in Gowers' notable improvement of Szemerédi's theorem, that any subset of the integers with positive density has infinitely long arithmetic progressions. To make his advance, Gowers introduced his uniformity norms, now usually called Gowers norms. The motivation for these norms is that if a function f has degree d, then its derivative in direction a, $f(x+a) - f(x)$, has degree at most $d-1$. Hence the $(d+1)$-fold derivative is 0. Thus, a natural test to decide if a function f has degree d is to set $k = d+1$, evaluate the k-fold derivative of f in k random directions, and accept only if the derivative

evaluates to 0. This is what we call the k^{th} Gowers norm test, $T_{\text{GN}(k)}$, for $k = d + 1$.

Our paper focuses on the field $\mathbb{F}_2$ of two elements, and we now restrict to this case. The k^{th} Gowers norm of $f : \mathbb{F}_2 \to \mathbb{F}_2$, denoted $\|f\|_{U^k}$, is given by the expression

$$\|f\|_{U^k} \stackrel{\text{def}}{=} (\Pr[T_{\text{GN}(k)} \text{ accepts}] - \Pr[T_{\text{GN}(k)} \text{ rejects}])^{\frac{1}{2^k}}.$$

Gowers [7] (see also [8]) showed that the correlation of f to the closest degree d polynomial is at most $\|f\|_{U^{d+1}}$. The well-known Inverse Conjecture for the Gowers Norm states that some sort of converse holds: if $\|f\|_{U^{d+1}} = \Omega(1)$, then the correlation of f to some degree d polynomial is $\Omega(1)$. Lovett et al. [4] and Green and Tao [3] disproved this conjecture as stated, but a modification of the conjecture remains open, and was recently proven in high characteristic [9], [10]. These conjectures and the Gowers norms have been extremely influential. For example, Green and Tao [11] used the Gowers norms over the integers to prove that the primes contain arbitrarily long arithmetic progressions.

Study of the Gowers norms over $\mathbb{F}_2$ has led to impressive results in theoretical computer science. Samorodnitsky and Trevisan [12] used Gowers norms to obtain very strong PCPs for Unique-Games-hard languages. This implied that Maximum Independent Set in graphs of maximum degree Δ could not be approximated within $\Delta/\text{polylog}(\Delta)$ under the Unique Games Conjecture. Using Gowers norms, Bogdanov and Viola [13] gave a pseudorandom generator fooling low-degree polynomials over $\mathbb{F}_2$. They could only prove their result under the inverse conjecture for the Gowers norm, but later Lovett [14] and Viola [15] used related ideas to prove an unconditional result. Finally, Viola and Wigderson [5] used Gowers norms to prove "XOR" lemmas for correlation to low-degree polynomials and to low communication protocols.

B. Local testing of Reed-Muller codes

Traditionally the Gowers norm is used in what Green and Tao call the 1% setting, where the correlation of a function to its closest low-degree polynomial is non-negligible but small. The 99% setting, where the correlation is close to 1, was addressed by Alon, Kaufman, Krivelevich, Litsyn, and Ron [1], and is the focus of our work. More precisely, Alon et al. considered the question of testing if a Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, given by an oracle, is close to a degree d multivariate polynomial. They proposed a test that turned out to be a variation of the Gowers norm test[1], where all the derivative directions are linearly independent, and showed that this test suffices for that setting. Thus, their analysis gave the only known relationship between the Gowers norm and the proximity to low-degree polynomials in the 99% setting.

However, their analysis was not optimal. In this work, we give an improved, asymptotically optimal, analysis of the Gowers norm test. This gives a tight connection with the Gowers norm in the 99% setting. Before we elaborate, let us introduce our framework.

Our question is also called testing of Reed-Muller codes, which are codes based on low-degree polynomials. The Reed-Muller codes are parameterized by two parameters: n, the number of variables, and d, the degree parameter. The Reed-Muller codes consist of all functions from $\mathbb{F}_2^n \to \mathbb{F}_2$ that are evaluations of polynomials of degree at most d. We use $\text{RM}(d, n)$ to denote this class, i.e., $\text{RM}(d, n) = \{f : \mathbb{F}_2^n \to \mathbb{F}_2 | \deg(f) \leq d\}$.

The proximity of functions is measured by the (fractional Hamming) distance. Specifically, for functions $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$, we let the *distance* between them, denoted by $\delta(f, g)$, be the quantity $\Pr_{x \leftarrow_U \mathbb{F}_2^n}[f(x) \neq g(x)]$. For a family of functions $\mathcal{F} \subseteq \{g : \mathbb{F}_2^n \to \mathbb{F}_2\}$ let $\delta(f, \mathcal{F}) = \min\{\delta(f, g) | g \in \mathcal{F}\}$. We say f is δ-close to $\mathcal{F}$ if $\delta(f, \mathcal{F}) \leq \delta$ and δ-far otherwise.

Let $\delta_d(f) = \delta(f, \text{RM}(d, n))$ denote the distance of f to the class of degree d polynomials. The goal of Reed-Muller testing is to "test", with "few queries" of f, whether $f \in \text{RM}(d, n)$ or f is far from $\text{RM}(d, n)$. Specifically, for a function $q : \mathbb{Z}^+ \times \mathbb{Z}^+ \times (0, 1] \to \mathbb{Z}^+$, a *$q$-query tester* for the class $\text{RM}(d, n)$ is a randomized oracle algorithm T that, given oracle access to some function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and a proximity parameter $\delta \in (0, 1]$, queries at most $q = q(d, n, \delta)$ values of f and accepts $f \in \text{RM}(d, n)$ with probability 1, while if $\delta(f, \text{RM}(d, n)) \geq \delta$ it rejects with probability at least, say, 2/3. The function q is the *query complexity* of the test and the main goal here is to minimize q, as a function possibly of d, n and δ. We denote the test T run using oracle access to the function f by T^f.

As mentioned earlier, Alon et al. [1] gave a tester with query complexity $O(\frac{d}{\delta} \cdot 4^d)$. Their tester consists of repetitions of a basic test, which we denote T_{GN}. T_{GN} is a modification of the Gowers norm test $T_{\text{GN}(d+1)}$ so that the $(d + 1)$-fold derivatives are evaluated in $d + 1$ random *linearly independent* directions. This modified tester, whose rejection probability differs from that of the original Gowers norm tester by only a constant factor, can be described alternatively as follows. Given oracle access to f, T_{GN} selects a random $(d + 1)$-dimensional affine subspace A, and accepts if f restricted to A is a degree d polynomial. This requires 2^{d+1} queries of f (since that is the number of points contained in A). Alon et al. show that if $\delta(f) \geq \delta$ then T_{GN} rejects f with probability $\Omega(\delta/(d \cdot 2^d))$. Their final tester then

[1]Gowers originally defined the Gowers norm as a norm on functions from $\mathbb{Z}_n \to \mathbb{C}$. Green and Tao later extended the definition of this norm to functions from an arbitrary abelian group G to $\mathbb{C}$. In the case of $G = \mathbb{F}_2^n$, this norm is closely related to the success probability of the Gowers norm test T_{GN}.

simply repeated T_{GN} $O(\frac{d}{\delta} \cdot 2^d)$ times and accepted if all invocations of T_{GN} accepted. The important feature of this result is that the number of queries is independent of n, the dimension of the ambient space. Alon et al. also show that any tester for $\mathrm{RM}(d, n)$ must make at least $\Omega(2^d + 1/\delta)$ queries. Thus their result was tight to within almost quadratic factors, but left a gap open. We close this gap in this work.

C. Main Result

Our main result is an optimal analysis of the Gowers norm test, up to constants. We show that if $\delta_d(f) \geq 0.1$, in fact even if it's at least $0.1 \cdot 2^{-d}$, then in fact the Gowers norm test rejects with probability lower bounded by some *absolute constant*. We now formally state our main theorem.

Theorem 1: There exists a constant $\epsilon_1 > 0$ such that for all d, n, and for all functions $f : \mathbb{F}_2^n \to \mathbb{F}_2$, we have[2]

$$\Pr[T_{\mathrm{GN}}^f \text{ rejects}] \geq \min\{2^d \cdot \delta_d(f), \epsilon_1\}.$$

Therefore, to reject functions δ-far from $\mathrm{RM}(d, n)$ with constant probability, a tester can repeat the test T_{GN} at most $O(1/\min\{2^d\delta_d(f), \epsilon_1\}) = O(1 + \frac{1}{2^d\delta})$ times, making the total query complexity $O(2^d + 1/\delta)$. This query complexity is asymptotically tight in view of the earlier mentioned lower bound in [1].

Our error analysis is also asymptotically tight. Note that our theorem effectively states that functions that are accepted by T_{GN} with constant probability (close to 1) are (very highly) correlated with degree d polynomials. To get a qualitative improvement, one could hope that every function that is accepted by T_{GN} with probability strictly greater than half is somewhat correlated with a degree d polynomial. Such stronger statements, however, are effectively ruled out by the counterexamples to the "inverse conjecture for the Gowers norm" given by [4], [3]. Since the analysis given in these works does not match our parameters asymptotically, we show (see the full version of this paper [16]) how an early analysis due to the authors of [4] can be used to show the asymptotic tightness of the parameters of Theorem 1.

Our analysis of the Gowers norm test implies a tight relationship between the Gowers norm and distance to degree d in the 99% setting. In particular, we show the following theorem.

Theorem 2: There exists $\epsilon > 0$ such that if $\|f\|_{U^{d+1}} \geq 1 - \epsilon/2^d$, then $\delta_d(f) = \Theta(1 - \|f\|_{U^{d+1}})$.

For comparison, the best previous lower bound comes from the Alon et al. work, whose result can be interpreted as showing that there exists $\epsilon > 0$ such that if $\|f\|_{U^{d+1}} \geq 1 - \epsilon/4^d$, then $\delta_d(f) = O(4^d(1 - \|f\|_{U^{d+1}}))$.

[2]For a tester T and a function f, the notation T^f indicates the execution of T with oracle access to f.

Before explaining our technique, we describe some applications of our result.

D. Tolerant testing of RM codes

Parnas, Ron, and Rubinfeld [17] introduced the notion of tolerant testing, and Guruswami and Rudra [18] studied this in the coding theoretic setting. Standard testers are required to reject strings that are far from codewords, but are not required to accept strings that are close to codewords. A tolerant tester is required to accept close codewords. In particular, for a code with minimum (relative) distance $\delta_{\min}$, there exists constants c_1 and c_2 such that the test must accept strings within distance $c_1\delta_{\min}$ with probability at least 2/3 (called the acceptance condition), and reject strings that are at least $(c_2\delta_{\min})$-far with probability at least 2/3 (called the rejection condition).

Any tester which satisfies the rejection condition must make at least $\Omega(1/\delta_{\min})$ queries. We observe that a tester that satisfies the rejection condition and makes $C/\delta_{\min}$ queries for a constant C is also tolerant. This follows because a string with distance $\delta_{\min}/(3C)$ will be rejected with probability at most 1/3. It even suffices to have the rejection condition with a constant probability (instead of 2/3), because the test can be repeated a constant number of times to boost the probability to 2/3.

In particular, for Reed-Muller codes, $\delta_{\min} = 2^{-d}$, and so, the Gowers norm test is also tolerant. No tolerant tester for binary Reed-Muller codes appears to have been known.

Theorem 3: T_{GN} is a tolerant tester for $\mathrm{RM}(d, n)$.

E. XOR lemma for low-degree polynomials

As mentioned earlier, Viola and Wigderson [5] used the Gowers norm and the Alon et al. analysis to give an elegant "hardness amplification" result for low-degree polynomials. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be such that $\delta_d(f)$ is noticeably large, say ≥ 0.1. Viola and Wigderson showed how to use this f to construct a $g : \mathbb{F}_2^m \to \mathbb{F}_2$ such that $\delta_d(g)$ is significantly larger, around $\frac{1}{2} - 2^{-\Omega(m)}$. In their construction, $g = f^{\oplus t}$, the t-wise XOR of f, where $f^{\oplus t} : (\mathbb{F}_2^n)^t \to \mathbb{F}_2$ is given by:

$$f^{\oplus t}(x_1, \ldots, x_t) = \sum_{i=1}^{t} f(x_i).$$

In particular, they showed that if $\delta_d(f) \geq 0.1$, then $\delta_d(f^{\oplus t}) \geq 1/2 - 2^{-\Omega(t/4^d)}$. Their proof proceeded by studying the rejection probabilities of T_{GN} on the functions f and $f^{\oplus t}$. The analysis of the rejection probability of T_{GN} given by [1] was a central ingredient in their proof. By using our improved analysis of the rejection probability of T_{GN} from Theorem 1 instead, we get the following improvement.

Theorem 4: Let ϵ_1 be as in Theorem 1. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. Then

$$\delta_d(f^{\oplus t}) \geq \frac{1-(1-2\min\{\epsilon_1/4, 2^{d-2}\cdot\delta_d(f)\})^{t/2^d}}{2}.$$

In particular, if $\delta_d(f) \geq 0.1$, then $\delta_d(f^{\oplus t}) \geq 1/2 - 2^{-\Omega(t/2^d)}$.

F. Query hierarchy for affine-invariant properties

Our result falls naturally in the general framework of property testing [2], [19], [20]. Goldreich et al. [6] asked an interesting question in this broad framework: Given an ensemble of properties $\mathcal{F} = \{\mathcal{F}_N\}_N$ where $\mathcal{F}_N$ is a property of functions on domains of size N, which functions correspond to the query complexity of some property? That is, for a given complexity function $q(N)$, is there a corresponding property $\mathcal{F}$ such that $\Theta(q(N))$-queries are necessary and sufficient for testing membership in $\mathcal{F}_N$? This question is interesting even when we restrict the class of properties being considered.

For completely general properties, this question is easy to solve. For graph properties, Goldreich et al. [6] show that for every efficiently computable function $q(N) = O(N)$, there is a graph property for which $\Theta(q(N))$ queries are necessary and sufficient (on graphs on $\Omega(\sqrt{N})$ vertices). Thus this gives a "hierarchy theorem" for query complexity.

Our main theorem settles the analogous question in the setting of "affine-invariant" properties. Given a field $\mathbb{F}$, a property $\mathcal{F} \subseteq \{\mathbb{F}^n \to \mathbb{F}\}$ is said to be affine-invariant if for every $f \in \mathcal{F}$ and affine map $A : \mathbb{F}^n \to \mathbb{F}^n$, the composition of f with A, i.e, the function $f \circ A(x) = f(A(x))$, is also in $\mathcal{F}$. Affine-invariant properties seem to be the algebraic analog of graph-theoretic properties and generalize most natural algebraic properties (see Kaufman and Sudan [21]).

Since the Reed-Muller codes form an affine-invariant family, and since we have a tight analysis for their query complexity, we can get the affine-invariant version of the result of [6]. Specifically, given any (reasonable) query complexity function $q(N)$, consider N that is a power of two and consider the class of functions on $n = \log_2 N$ variables of degree at most $d = \lceil \log_2 q(N) \rceil$. We have that membership in this family requires $\Omega(2^d) = \Omega(q(N))$-queries, and on the other hand $O(2^d) = O(q(N))$-queries also suffice, giving an ensemble of properties $\mathcal{P}_N$ (one for every $N = 2^n$) that is testable with $\Theta(q(N))$-queries.

Theorem 5: For every $q : \mathbb{N} \to \mathbb{N}$ that is at most linear, there is an affine-invariant property that is testable with $O(q(n))$ queries (with one-sided error) but is not testable in $o(q(n))$ queries (even with two-sided error). Namely, this property is membership in $\mathrm{RM}(\lceil \log_2 q(n) \rceil, n)$.

G. Technique

Our main theorem (Theorem 1) is obtained by a novel proof that gives a (yet another!) new analysis even of the classical linearity test of Blum, Luby, Rubinfeld [2]. The heart of our proof is an inductive argument on n, the dimension of the ambient space. While proofs that use induction on n have been used before in the literature on low-degree testing (see, for instance, [22], [23], [24]), they tend to have a performance guarantee that degrades significantly with n. Indeed no inductive proof was known even for the case of testing linearity of functions from $\mathbb{F}_2^n \to \mathbb{F}_2$ that showed that functions at $\Omega(1)$ distance from linear functions are rejected with $\Omega(1)$ probability. (We note that the original analysis of [2] as well as the later analysis of [25] do give such bounds, but they do not use induction on n.) In the process of giving a tight analysis of the [1] test for Reed-Muller codes, we thus end up giving a new (even if weaker) analysis of the linearity test over $\mathbb{F}_2^n$. Below we give the main idea behind our proof.

Consider a function f that is δ-far from every degree d polynomial. For a "hyperplane", i.e., an $(n-1)$-dimensional affine subspace A of $\mathbb{F}_2^n$, let $f|_A$ denote the restriction of f to A. We first note that the test can be interpreted as first picking a random hyperplane A in $\mathbb{F}_2^n$ and then picking a random $(d+1)$-dimensional affine subspace A' within A and testing if $f|_{A'}$ is a degree d polynomial. Now, if on every hyperplane A, $f|_A$ is still δ-far from degree d polynomials then we would be done by the inductive hypothesis. In fact our hypothesis gets weaker as $n \to \infty$, so that we can even afford a few hyperplanes where $f|_A$ is not δ-far. The crux of our analysis is when for several (but just $O(2^d)$) hyperplanes $f|_A$ is close to some degree d polynomial P_A. In this case we manage to "sew" the different polynomials P_A (each defined on some $(n-1)$-dimensional subspace within $\mathbb{F}_2^n$) into a degree d polynomial P that agrees with *all* the P_A's. We then show that this polynomial is close to f, completing our argument.

To stress the novelty of our proof, note that this is not a "self-correction" argument as in [1], where one defines a natural function that is close to P, and then works hard to prove it is a polynomial of appropriate degree. In contrast, our function is a polynomial by construction and the harder part (if any) is to show that the polynomial is close to f. Moreover, unlike other inductive proofs, our main gain is in the fact that the new polynomial P has degree no greater than that of the polynomials given by the induction.

Organization of this paper:: We prove our main theorem, Theorem 1, in Section II assuming three lemmas, two of which study the rejection probability of the k-dimensional affine subspace test, and another that

relates the rejection probability of the basic $(d+1)$-dimensional affine subspace test to that of the k-dimensional affine subspace test. These three lemmas are proved in the following section, Section III.

We give the relationship to the Gowers norm in Section V, and we prove our improved hardness amplification theorem, Theorem 4, in Section VI. Some proofs that are abbreviated in this version may be found in more detail in the full version of this paper [16].

II. Proof of Main Theorem

In this section, we prove Theorem 1. We start with an overview of our proof. Recall that a k-flat is an affine subspace of dimension k, and a hyperplane is an $(n-1)$-flat.

The proof of the main theorem proceeds as follows. We begin by studying a variant of the basic tester T_{GN}, which we call $T_{d,k}$ or the *k-flat test*. For an integer $k \geq d+1$, $T^f_{d,k}$ picks a uniformly random k-flat in $\mathbb{F}_2^n$, and accepts if and only if the restriction of f to that flat has degree at most d. In this language, the tester T_{GN} of interest to us is $T_{d,d+1}$. To prove Theorem 1, we first show that for $k \approx d+10$, the tester $T^f_{d,k}$ rejects with constant probability if $\delta_d(f)$ is $\Omega(2^{-d})$ (see Lemma 9). We then relate the rejection probabilities of $T^f_{d,k}$ and T^f_{GN} (see Lemma 10).

The central ingredient in our analysis is thus Lemma 9 which is proved by induction on n, the dimension of the ambient space. Recall that we want to show that the two quantities (1) $\delta_d(f)$ and (2) $\Pr[T^f_{d,k}$ rejects], are closely related. We consider what happens to f when restricted to some hyperplane A. Denote such a restriction by $f|_A$. For a hyperplane A we consider the corresponding two quantities (1) $\delta_d(f|_A)$ and (2) $\Pr[T^{f|_A}_{d,k}$ rejects]. The inductive hypothesis tells us that these two quantities are closely related for each A. Because of the local nature of tester $T_{d,k}$, it follows easily that $\Pr[T^f_{d,k}$ rejects] is the average of $\Pr[T^{f|_A}_{d,k}$ rejects] over all hyperplanes A. The main technical content of Lemma 9 is that there is a similar tight relationship between $\delta_d(f)$ and the numbers $\delta_d(f|_A)$ as A varies over all hyperplanes A. This relationship suffices to complete the proof. The heart of our analysis focuses on the case where for many hyperplanes (about 2^k of them, independent of n), the quantity $\delta_d(f|_A)$ is very small (namely, for many A, there is a polynomial P_A of degree d that is very close to $f|_A$). In this case, we show how to "sew" together the polynomials P_A to get a polynomial P on $\mathbb{F}_2^n$ that is also very close to f. In contrast to prior approaches which yield a polynomial P with larger degree than that of the P_A's, our analysis crucially preserves this degree, leading to the eventual tightness of our analysis.

We now turn to the formal proof.

A. Preliminaries

We begin by formally introducing the k-flat test and some related notation.

Definition 6 (k-flat test $T_{d,k}$): The test $T^f_{d,k}$ picks a random k-flat $A \subseteq \mathbb{F}_2^n$ and accepts if and only if $f|_A$ (f restricted to A) is a polynomial of degree at most d.

The rejection probability of $T^f_{d,k}$ is denoted $\mathsf{Rej}_{d,k}(f)$. In words, this is the probability that $f|_A$ is not a degree d polynomial when A is chosen uniformly at random among all k-flats of F_2^n.

Although we don't need it for our argument, we note that $T_{\mathrm{GN}} = T_{d,d+1}$ accepts if and only if the 2^{d+1} evaluations $f|_A$ sum to 0. The following folklore proposition shows that for $k \geq d+1$, $T_{d,k}$ has perfect completeness.

Proposition 7: For every $k \geq d+1$, $\delta_d(f) = 0$ if and only if $\mathsf{Rej}_{d,k}(f) = 0$.

B. Key Lemmas

We now state our three key lemmas, and then use them to finish the proof of Theorem 1. The first is a simple lemma that says if the function is sufficiently close to a degree d polynomial, then the rejection probability is linear in its distance from degree d polynomials.

Lemma 8: For every k, ℓ, d such that $k \geq \ell \geq d+1$, if $\delta(f) = \delta$ then $\mathsf{Rej}_{d,k}(f) \geq 2^\ell \cdot \delta \cdot (1 - (2^\ell - 1)\delta)$. In particular, if $\delta \leq 2^{-(d+2)}$ then $\mathsf{Rej}_{d,k}(f) \geq \min\{\frac{1}{8}, 2^{k-1} \cdot \delta\}$.

The proof uses pairwise-independence in a straightforward way to argue that, with good probability, the randomly chosen flat will contain exactly one point where f and the closest degree d polynomial differ, in which case the test will reject. Details are omitted from this version of the paper.

The next lemma is at the heart of our analysis and allows us to lower bound the rejection probability when the function is bounded away from degree d polynomials.

Lemma 9: There exist positive constants $\beta < 1/4, \epsilon_0, \gamma$ and c such that the following holds for every d, k, n, such that $n \geq k \geq d+c$. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be such that $\delta(f) \geq \beta \cdot 2^{-d}$. Then $\mathsf{Rej}_{d,k}(f) \geq \epsilon_0 + \gamma \cdot 2^d/2^n$.

The final lemma relates the rejection probabilities of different dimensional tests.

Lemma 10: For every n, d and $k \geq k' \geq d+1$, and every $f : \mathbb{F}_2^n \to \mathbb{F}_2$, we have

$$\mathsf{Rej}_{d,k'}(f) \geq \mathsf{Rej}_{d,k}(f) \cdot 2^{-(k-k')}.$$

Given the three lemmas above, Theorem 1 follows easily as shown below.

Proof of Theorem 1: Let ϵ_0 and c be as in Lemma 9. We prove the theorem for $\epsilon_1 = \epsilon_0 \cdot 2^{-(c-1)}$. First note that if $\delta(f) \leq 2^{-(d+2)}$, then we are done by Lemma 8. So assume $\delta(f) \geq 2^{-(d+2)} \geq \beta \cdot 2^{-d}$, where β is the constant from Lemma 9. By Lemma 9, we know

that $\mathsf{Rej}_{d,d+c}(f) \geq \epsilon_0$. Lemma 10 now implies that $\mathsf{Rej}_{d,d+1}(f) \geq \epsilon_0 \cdot 2^{-(c-1)}$, as desired. ∎

III. Analysis of the k-flat test

Throughout this section we fix d, so we suppress it in the subscripts and simply use $\delta(f) = \delta_d(f)$ and $\mathsf{Rej}_k(f) = \mathsf{Rej}_{d,k}(f)$.

A. Lemma 9: When f is bounded away from $\mathrm{RM}(d,n)$

The main idea of the proof of Lemma 9 is to consider the restrictions of f on randomly chosen "hyperplanes", i.e., $(n-1)$-flats. If on an overwhelmingly large fraction (which will be quantified in the proof) of hyperplanes, our function is far from degree d polynomials, then the inductive hypothesis suffices to show that f will be rejected with high probability (by the k-flat test). The interesting case is when the restrictions of f to several hyperplanes are close to degree d polynomials. In Lemma 12 we use the close polynomials on such hyperplanes to construct a polynomial that has significant agreement with f on the union of the hyperplanes.

We start by first fixing some terminology. We say A and B are *complementary* hyperplanes if $A \cup B = \mathbb{F}_2^n$. Recalling that a hyperplane is the set of points $\{x \in \mathbb{F}_2^n | L(x) = b\}$ where $L : \mathbb{F}_2^n \to \mathbb{F}_2$ is a nonzero linear function and $b \in \mathbb{F}_2$, we refer to L as the linear part of the hyperplane. We say that hyperplanes $A_1, \ldots, A_\ell$ are linearly independent if the corresponding linear parts are independent. The following proposition lists some basic facts about hyperplanes that we use. The proof is omitted.

Proposition 11 (Properties of hyperplanes): We have the following:

1) There are exactly $2^{n+1} - 2$ distinct hyperplanes in $\mathbb{F}_2^n$.
2) Among any $2^\ell - 1$ distinct hyperplanes, there are at least ℓ independent hyperplanes.
3) There is an affine invertible transform that maps independent hyperplanes $A_1, \ldots, A_\ell$ to the hyperplanes $x_1 = 0, x_2 = 0, \ldots, x_\ell = 0$.

We are now ready to prove Lemma 9. We first recall the statement.

Lemma 9 (recalled): There exist positive constants $\beta < 1/4, \epsilon_0, \gamma$ and c such that the following holds for every d, k, n, such that $n \geq k \geq d + c$. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be such that $\delta(f) \geq \beta \cdot 2^{-d}$. Then $\mathsf{Rej}_{d,k}(f) \geq \epsilon_0 + \gamma \cdot 2^d/2^n$.

Proof of Lemma 9: We prove the lemma for every $\beta < 1/24$, $\epsilon_0 < 1/8$, $\gamma \geq 72$, and c such that $2^c \geq \max\{4\gamma/(1-8\epsilon_0), \gamma/(1-\epsilon_0), 2/\beta\}$. (In particular, the choices $\beta = 1/25$, $\epsilon_0 = 1/16$, $\gamma = 72$ and $c = 10$ work.)

The proof uses induction on $n - k$. When $n = k$ we have $\mathsf{Rej}_k(f) = 1 \geq \epsilon_0 + \gamma \cdot 2^{d-k}$ as required, because $2^c \geq \frac{\gamma}{1-\epsilon_0}$. So we move to the inductive step.

Let $\mathcal{H}$ denote the set of hyperplanes in $\mathbb{F}_2^n$. Let $N = 2(2^n - 1)$ be the cardinality of $\mathcal{H}$. Let $\mathcal{H}^*$ be the set of all the hyperplanes $A \in \mathcal{H}$ such that $\delta(f|_A, \mathrm{RM}(d, n-1)) < \beta \cdot 2^{-d}$. Let $K = |\mathcal{H}^*|$.

Now because a random k-flat of a random hyperplane *is* a random k-flat, we have $\mathsf{Rej}_k(f) = \mathbb{E}_{A \in \mathcal{H}}[\mathsf{Rej}_k(f|_A)]$. By the induction hypothesis, for any $A \in \mathcal{H} \setminus \mathcal{H}^*$, we have $\mathsf{Rej}_k(f|_A) \geq \epsilon_0 + \gamma \cdot \frac{2^d}{2^{n-1}}$. Thus, $\mathsf{Rej}_k(f) \geq \epsilon_0 + \gamma \cdot \frac{2^d}{2^{n-1}} - K/N$. We now take cases on whether K is large or small:

1) **Case 1:** $K \leq \gamma \cdot 2^d$.
   In this case, $\mathsf{Rej}_k(f) \geq \epsilon_0 + \gamma \cdot 2^d/2^{n-1} - K/N \geq \epsilon_0 + \gamma \cdot 2^d/2^n$ as desired.
2) **Case 2:** $K > \gamma \cdot 2^d$.
   Lemma 12 (below) shows that in this case, $\delta(f) \leq \frac{3}{2}\beta \cdot 2^{-d} + 9/(\gamma 2^d) \stackrel{\text{def}}{=} \delta_0$, provided $\beta \cdot 2^{-d} < 2^{-(d+2)}$ (which holds since $\beta < 1/24 < 1/4$). Note that since $\beta < 1/24$ and $9/\gamma < 1/8$, we get $\delta_0 < 2^{-(d+2)}$ and so Lemma 8 implies that $\mathsf{Rej}_k(f) \geq \min\{2^{k-1} \cdot \delta(f), \frac{1}{8}\} \geq \min\{2^{k-1} \cdot \beta \cdot 2^{-d}, \frac{1}{8}\}$. It is easy to check that both these quantities above are at least $\epsilon_0 + \gamma/2^{(c+1)} \geq \epsilon_0 + \gamma 2^d/2^n$. This completes the proof of the lemma.

∎

Lemma 12: For $f : \mathbb{F}_2^n \to \mathbb{F}_2$, let $A_1, \ldots, A_K$ be hyperplanes such that $f|_{A_i}$ is α-close to some degree d polynomial on A_i. If $K > 2^{d+1}$ and $\alpha < 2^{-(d+2)}$, then $\delta(f) \leq \frac{3}{2}\alpha + 9/K$.

Proof: Let P_i be the degree d polynomial such that $f|_{A_i}$ is α-close to P_i.

Claim 13: If $4\alpha < 2^{-d}$ then for every pair of hyperplanes A_i and A_j, we have $P_i|_{A_i \cap A_j} = P_j|_{A_j \cap A_i}$.

Proof: If A_i and A_j are complementary then this is vacuously true. Otherwise, $|A_i \cap A_j| = |A_i|/2 = |A_j|/2$. So $\delta(f|_{A_i \cap A_j}, P_i|_{A_i \cap A_j}) \leq 2\delta(f|_{A_i}, P_i) \leq 2\alpha$ and similarly $\delta(f|_{A_i \cap A_j}, P_j|_{A_i \cap A_j}) \leq 2\alpha$. So $\delta(P_i|_{A_i \cap A_j}, P_j|_{A_i \cap A_j}) \leq 4\alpha < 2^{-d}$. But these are both degree d polynomials and so if their proximity is less than 2^{-d} then they must be identical. ∎

Let $\ell = \lfloor \log_2(K+1) \rfloor$. Thus $\ell > d$. By Proposition 11 there are at least ℓ linearly independent hyperplanes among $A_1, \ldots, A_K$. Without loss of generality let these be $A_1, \ldots, A_\ell$. Furthermore, by an affine transformation of coordinates, for $i \in [\ell]$ let A_i be the hyperplane $\{x \in \mathbb{F}_2^n \mid x_i = 0\}$. For $i \in [\ell]$ extend P_i to a function on all of $\mathbb{F}_2^n$ by making P_i independent of x_i. We will sew together $P_1, \ldots, P_\ell$ to get a polynomial close to f.

Let us write all functions from $\mathbb{F}_2^n \to \mathbb{F}_2$ as polynomials in n variables $x_1, \ldots, x_\ell$ and $\mathbf{y}$ where $\mathbf{y}$ denotes the last $n - \ell$ variables. For $i \in [\ell]$ and $S \subseteq [\ell]$, let $P_{i,S}(\mathbf{y})$ be the monomials of P_i which contain x_i for $i \in S$, and no x_j for $j \notin [\ell]$. That is, $P_{i,S}(\mathbf{y})$ are polynomials such that $P_i(x_1, \ldots, x_\ell, \mathbf{y}) = \sum_{S \subseteq [\ell]} P_{i,S}(\mathbf{y}) \prod_{j \in S} x_j$. Note that

the degree of $P_{i,S}$ is at most $d - |S|$. (In particular, if $|S| > d$, then $P_{i,S} = 0$.) Note further that since P_i is independent of x_i, we have that $P_{i,S} = 0$ if $i \in S$.

Claim 14: For every $S \subseteq [\ell]$ and every $i, j \in [\ell] - S$, $P_{i,S}(\mathbf{y}) = P_{j,S}(\mathbf{y})$.

Proof: Note that $P_i|_{A_i \cap A_j}(\mathbf{x}, \mathbf{y}) = \sum_{S \subseteq [\ell]-\{i,j\}} P_{i,S}(\mathbf{y}) \prod_{m \in S} x_m$. Similarly $P_j|_{A_i \cap A_j}(\mathbf{x}, \mathbf{y}) = \sum_{S \subseteq [\ell]-\{i,j\}} P_{j,S}(\mathbf{y}) \prod_{m \in S} x_m$. Since the two functions are equal (by Claim 13), we have that every pair of coefficients of $\prod_{m \in S} x_m$ must be the same. We conclude that $P_{i,S} = P_{j,S}$. ∎

Claim 14 above now allows us to define, for every $S \subsetneq [\ell]$, the polynomial $P_S(\mathbf{y})$ as the unique polynomial $P_{i,S}$ where $i \notin S$. We define

$$P(x_1, \ldots, x_\ell, \mathbf{y}) = \sum_{S \subsetneq [\ell]} P_S(\mathbf{y}) \prod_{j \in S} x_j.$$

By construction, the degree of P is at most d. This is the polynomial that we will eventually show is close to f.

Claim 15: For every $i \in [K]$, $P|_{A_i} = P_i|_{A_i}$.

Proof: First note that for each $i \in [\ell]$, $P|_{A_i} = P_i|_{A_i}$. This is because the coefficients of the two polynomials become identical after substituting $x_i = 0$ (recall that A_i is the hyperplane $\{x \in \mathbb{F}_2^n \mid x_i = 0\}$).

Now consider general $i \in [K]$. For any point $x \in A_i \cap (\bigcup_{j=1}^{\ell} A_j)$, letting $j^* \in [\ell]$ be such that $x \in A_{j^*}$, we have $P_i(x) = P_{j*}(x)$ (by Claim 13) and $P_{j^*}(x) = P(x)$ (by what we just showed, since $j^* \in [\ell]$). Thus P and P_i agree on all points in $A_i \cap (\bigcup_{j=1}^{\ell} A_j)$. Now since $\ell > d$, we have that $|A_i \cap (\bigcup_{j=1}^{\ell} A_j)|/|A_i| \geq 1 - 2^{-\ell} > 1 - 2^{-d}$, and since $P|_{A_i}$ and $P_i|_{A_i}$ are both degree d polynomials, we conclude that $P|_{A_i}$ and $P_i|_{A_i}$ are identical. Thus for all $i \in [K]$, $P|_{A_i} = P_i|_{A_i}$. ∎

We will show below that P is close to f, by considering all the hyperplanes $A_1, \ldots, A_K$. If these hyperplanes uniformly covered F_2^n, then we could conclude $\delta(f, P) \leq \alpha$, as f is α-close to P on each hyperplane. Since the A_i don't uniformly cover $\mathbb{F}_2^n$, we'll argue that almost all points are covered approximately the right number of times, which will be good enough. To this end, let

$$\textsc{Bad} = \{z \in \mathbb{F}_2^n | z \text{ is contained in less than } K/3 \text{ of the hyperplanes } A_1, \ldots, A_K\}.$$

Let $\tau = |\textsc{Bad}|/2^n$.

Claim 16: $\delta(f, P) \leq 3/2 \cdot \alpha + \tau$.

Proof: Consider the following experiment: Pick $z \in \mathbb{F}_2^n$ and $i \in [K]$ uniformly and independently at random and consider the probability that "$z \in A_i$ and $f(z) \neq P_i(z)$". A simple analysis (whose details we omit in this version of the paper), using the fact that $P|_{A_i} = P_i$, shows that

$$\frac{1}{3} \cdot (\delta(f, P) - \tau) \leq \Pr_{z,i}[z \in A_i \ \&\ f(z) \neq P_i(z)] \leq \frac{1}{2} \cdot \alpha.$$

∎

Claim 17: $\tau \leq 9/K$.

The proof is a straightforward "pairwise independence" argument, which we omit in this version of the paper.

The lemma follows from the last two claims above. ∎

IV. Lemma 10: Relating different dimensional tests

Lemma 18: Let $k \geq d + 1$ and let $f : \mathbb{F}_2^{k+1} \to \mathbb{F}_2$ have degree greater than d. Then $\mathsf{Rej}_{d,k}(f) \geq 1/2$.

Proof: Assume for contradiction that there is a strict majority of hyperplanes A on which $f|_A$ has degree d. Then there exists two complementary hyperplanes A and $\bar{A}$ such that $f|_A$ and $f|_{\bar{A}}$ both have degree d. We can interpolate a polynomial P of degree at most $d+1$ that now equals f everywhere. If P is of degree d, we are done, so assume P has degree exactly $d+1$ and let P_h be the homogeneous degree $d+1$ part of P (i.e,, $P = P_h + Q$ where $\deg(Q) \leq d$ and P_h is homogeneous). Now consider all hyperplanes A such that $f|_A = P|_A$ has degree at most d. Since these form a strict majority, there are at least $\frac{1}{2}(2^{k+2} - 2) + 1 > 2^{k+1} - 1$ such hyperplanes. It follows that there are at least $k+1 \geq d+2$ linearly independent hyperplanes such that this condition holds. By an affine transformation we can assume these hyperplanes are of the form $x_1 = 0, \ldots, x_{d+2} = 0$. But then $\prod_{i=1}^{d+2} x_i$ divides P_h which contradicts the fact that the degree of P_h is at most $d + 1$. ∎

Lemma 19: Let $n \geq k \geq d + 1$ and let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ have degree greater than d. Then $\mathsf{Rej}_{d,k}(f) \geq 2^{k-n}$.

Proof: The proof is a simple induction on n. The base case of $n = k$ is trivial. Now assume for $n-1$. Pick a random hyperplane A. With probability at least $1/2$ (by the previous lemma), $f|_A$ is not a degree d polynomial. By the inductive hypothesis, a random k-flat of A will now detect that $f|_A$ is not of degree d with probability 2^{k-n+1}. We conclude that a random k-flat of $\mathbb{F}_2^n$ yields a function of degree greater than d with probability at least 2^{k-n}. ∎

We now have all the pieces needed to prove Lemma 10.

Lemma 10 (recalled): For every n, d and $k \geq k' \geq d + 1$, and every $f : \mathbb{F}_2^n \to \mathbb{F}_2$, we have

$$\mathsf{Rej}_{d,k'}(f) \geq \mathsf{Rej}_{d,k}(f) \cdot 2^{-(k-k')}.$$

Proof of Lemma 10: We view the k'-flat test as the following process: first pick a random k-flat A_1 of $\mathbb{F}_2^n$, then pick a random k'-flat A of A_1, and accept iff $f|_A$ is a degree d polynomial. Note that this is completely equivalent to the k'-flat test.

To analyze our test, we first consider the event that $f|_{A_1}$ is not a degree d polynomial. The probability that this happens is $\mathsf{Rej}_{d,k}(f)$. Now conditioned on the event that $f|_{A_1}$ is not a degree d polynomial, we can now

use Lemma 19 to conclude that the probability that $(f|_{A_1})|_A$ is not a degree d polynomial is at least $2^{-(k-k')}$. We conclude that $\mathsf{Rej}_{d,k'}(f) \geq \mathsf{Rej}_{d,k}(f) \cdot 2^{-(k-k')}$. The lemma follows. ∎

V. Gowers norms

Our main theorem can be interpreted as giving a tight relationship between the Gowers norm of a function f and its proximity to some low degree polynomial. In this section, we describe this relationship.

We start by recalling the definition of the test $T^f_{\mathrm{GN}(k)}$ and the Gowers norm $\|f\|_{U^k}$. On oracle access to function f, the test $T_{\mathrm{GN}(k)}$ picks x_0 and directions $a_1, \ldots, a_k$ uniformly and independently in $\mathbb{F}_2^n$ and accepts if and only if $f|_A$ is a degree $k-1$ polynomial, where $A = \{x_0 + \mathrm{span}(a_1, \ldots, a_k)\}$. The Gowers norm is given by the expression

$$\|f\|_{U^k} \stackrel{\mathrm{def}}{=} (\Pr[T^f_{\mathrm{GN}(k)} \text{ accepts}] - \Pr[T^f_{\mathrm{GN}(k)} \text{ rejects}])^{\frac{1}{2^k}}.$$

Our main quantity of interest is the correlation of f with degree d polynomials, i.e., the quantity $1 - 2\delta_d(f)$.

Our theorem relating the Gowers norm to the correlation is given below.

Theorem 20: There exists $\epsilon > 0$ such that if $\|f\|_{U^{d+1}} \geq 1 - \epsilon/2^d$, then $\delta_d(f) = \Theta(1 - \|f\|_{U^{d+1}})$.

To prove the theorem we first relate the rejection probability of the test $T_{\mathrm{GN}(d+1)}$ with that of the test T_{GN}.

Proposition 21: For every $n \geq d+1$ and for every f, $\Pr[T^f_{\mathrm{GN}(d+1)} \text{rejects}] \geq \frac{1}{4} \cdot \Pr[T^f_{\mathrm{GN}} \text{rejects}]$.

Proof: We show that with probability at least $1/4$, the a_i are linearly independent. Consider picking d independent vectors $a_1, \ldots, a_d$ in $\mathbb{F}_2^n$. For fixed $\beta_1, \ldots, \beta_d \in \mathbb{F}_2$ (not all zero), the probability that $\sum_i \beta_i a_i = 0$ is at most 2^{-n}. Taking the union bound over all sequences $\beta_1, \ldots, \beta_d$ we find that the probability that $a_1, \ldots, a_d$ have a linear dependency is at most $2^{d-n} \geq \frac{1}{2}$ if $n \geq d+1$. For any fixed $a_1, \ldots, a_d$, the probability that $a_{d+1} \in \mathrm{span}(a_1, \ldots, a_d)$ is also at most $\frac{1}{2}$. Thus we find with probability at least $1/4$, the vectors $a_1, \ldots, a_{d+1}$ are linearly independent provided $n \geq d+1$. The proposition follows since the rejection probability of $T^f_{\mathrm{GN}(d+1)}$ equals the rejection probability of T^f_{GN} times the probability that $a_1, \ldots, a_{d+1}$ are linearly independent. ∎

We are now ready to prove Theorem 20.

Proof of Theorem 20: The proof is straightforward given our main theorem and the work of Gowers et al. [7], [8]. As mentioned earlier, Gowers already showed that $1 - 2\delta_d(f) \leq \|f\|_{U^{d+1}}$ [7], [8], i.e., $\delta_d(f) \geq (1 - \|f\|_{U^{d+1}})/2$.

For the other direction, suppose $\|f\|_{U^{d+1}} = 1 - \gamma$, where $\gamma \leq \epsilon/2^d$ for small enough ϵ. Let ρ denote the rejection probability of $T^f_{\mathrm{GN}(d+1)}$. By Proposition 21 we have $\rho \geq \frac{1}{4} \cdot \mathsf{Rej}_{d,d+1}(f)$. By choosing ϵ small enough, we also have $1 - 2\rho = \|f\|^{2^{d+1}}_{U^{d+1}} > 1 - \epsilon_1/2$, i.e., $\rho < \epsilon_1/4$, so $\mathsf{Rej}_{d,d+1}(f) < \epsilon_1$. Thus, by Theorem 1,

$$\begin{aligned}
\delta_d(f) &\leq \frac{1}{2^d}\mathsf{Rej}_{d,d+1}(f)(f) \\
&\leq \frac{1}{2^{d-2}}\rho \\
&= \frac{1}{2^{d-1}}(1 - \|f\|^{2^{d+1}}_{U^{d+1}}) \\
&= \frac{1}{2^{d-1}}(1 - (1-\gamma)^{2^{d+1}}) \\
&\leq \frac{1}{2^{d-1}}(1 - (1 - O(2^{d+1}\gamma))) \\
&= O(\gamma),
\end{aligned}$$

as required. ∎

VI. XOR lemma for low-degree polynomials

In this section, we reproduce an argument of Viola and Wigderson [5] and show, taking our improved Reed-Muller test into account, an improved XOR lemma for low-degree polynomials.

A crucial feature of the test $T_{\mathrm{GN}(k)}$ (that is not a feature of the k-flat test for $k > d+1$) is that the rejection probability of $f^{\oplus t}$ can be exactly expressed as a rapidly growing (in t) function of the rejection probability of f. Let $\mathsf{Rej}^0_d(f)$ denote the rejection probability of $T^f_{\mathrm{GN}(d+1)}$. Then we have:

Proposition 22:

$$(1 - 2\mathsf{Rej}^0_d(f^{\oplus t})) = (1 - 2\mathsf{Rej}^0_d(f))^t.$$

Proof: We first note that the proposition is equivalent to showing that $\|f^{\oplus t}\|_{U^{d+1}} = (\|f\|_{U^{d+1}})^t$. It is a standard fact (e.g., Fact 2.6 in [5]) that for functions f, g on disjoint sets of inputs, $\|f(x) + g(y)\|_{U^{d+1}} = \|f(x)\|_{U^{d+1}} \cdot \|g(y)\|_{U^{d+1}}$. This immediately yields the proposition. ∎

We also use the following well-known relationship between the Gowers norm and the correlation of a function to the class of degree d polynomials. (We state it in terms of the rejection probability of the test $T_{\mathrm{GN}(d+1)}$.)

Lemma 23 ([7], [8]):

$$1 - 2\delta_d(g) \leq (1 - 2\mathsf{Rej}^0_d(g))^{\frac{1}{2^d}}.$$

We are now ready to prove Theorem 4 which we recall below.

Theorem 4 (recalled): Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. Then

$$\delta_d(f^{\oplus t}) \geq \frac{1 - (1 - 2\min\{\epsilon_1/4, 2^{d-2} \cdot \delta_d(f)\})^{t/2^d}}{2}.$$

In particular, if $\delta_d(f) \geq 0.1$, then $\delta_d(f^{\oplus t}) \geq \frac{1 - 2^{-\Omega(t/2^d)}}{2}$.

Proof of Theorem 4: By Theorem 1 and Proposition 21,

$$\mathsf{Rej}^0_d(f) \geq \min\{\epsilon_1/4, 2^{d-2} \cdot \delta_d(f)\}.$$

Thus by Proposition 22,

$$(1-2\mathsf{Rej}^0_d(f^{\oplus t}))^{\frac{1}{2^d}} = (1-2\mathsf{Rej}^0_d(f))^{\frac{t}{2^d}} \leq (1-2\min\{\epsilon_1/4, 2^{d-2}\cdot\delta_d(f)\})^{\frac{t}{2^d}}.$$

Finally, Lemma 23 shows that

$$\delta_d(f^{\oplus t}) \geq \frac{1-\left(1-2\min\{\epsilon_1/4, 2^{d-2}\cdot\delta_d(f)\}\right)^{\frac{t}{2^d}}}{2}.$$

■

Conclusions

We gave an optimal analysis of a natural test for low-degree polynomials over $\mathbb{F}_2$, and in the process determined the optimal query complexity for testing this class of properties.

Low-degree tests over general fields have also been studied extensively [19], [26], [27]. It would be interesting to determine the optimal query complexity for testing low-degree polynomials over other fields too.

Acknowledgments

Thanks to Alex Samorodnitsky and Shachar Lovett for sharing some of the unpublished parts of their work [4] and allowing us to present parts of their proof in the full version of this paper. Thanks to Alex also for numerous stimulating discussions from the early stages of this work, and to Jakob Nordström for bringing some of the authors together on this work.

References

[1] N. Alon, T. Kaufman, M. Krivelevich, S. Litsyn, and D. Ron, "Testing Reed-Muller codes," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 4032–4039, 2005.

[2] M. Blum, M. Luby, and R. Rubinfeld, "Self-testing/correcting with applications to numerical problems," *J. Comp. Sys. Sci.*, vol. 47, pp. 549–595, 1993, earlier version in STOC'90.

[3] B. Green and T. Tao, "The distribution of polynomials over finite fields, with applications to the Gowers norms," *Contributions to Discrete Mathematics*, vol. 4, no. 2, pp. 1–36, 2009.

[4] S. Lovett, R. Meshulam, and A. Samorodnitsky, "Inverse conjecture for the Gowers norm is false," in *Proc. 40th Annual ACM Symposium on the Theory of Computing*. New York, NY, USA: ACM, 2008, pp. 547–556.

[5] E. Viola and A. Wigderson, "Norms, XOR lemmas, and lower bounds for GF(2) polynomials and multiparty protocols," in *Proc. 22nd Annual IEEE Conference on Computational Complexity*, June 2007, pp. 141–154.

[6] O. Goldreich, M. Krivelevich, I. Newman, and E. Rozenberg, "Hierarchy theorems for property testing," in *APPROX-RANDOM*. Springer-Verlag, 2009, pp. 504–519.

[7] W. T. Gowers, "A new proof of Szeméredi's theorem," *Geom. funct. anal.*, vol. 11, no. 3, pp. 465–588, 2001.

[8] B. Green and T. Tao, "An inverse theorem for the Gowers U^3-norm," *Proc. Edin. Math. Soc.*, vol. 51, pp. 73–153, 2008.

[9] V. Bergelson, T. Tao, and T. Ziegler, "An inverse theorem for the uniformity seminorms associated with the action of $\mathbb{F}_p^\infty$," *Geom. funct. anal.*, vol. 19, no. 6, pp. 1539–1596, 2010.

[10] T. Tao and T. Ziegler, "The inverse conjecture for the Gowers norm over finite fields via the correspondence principle," *Analysis & PDE (to appear)*, preprint available at http://arxiv.org/abs/0810.5527.

[11] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions," *Ann. of Math. (2)*, vol. 167, no. 2, pp. 481–547, 2008.

[12] A. Samorodnitsky and L. Trevisan, "Gowers uniformity, influence of variables, and pcps," in *Proc. 38th Annual ACM Symposium on the Theory of Computing*. New York, NY, USA: ACM, 2006, pp. 11–20.

[13] A. Bogdanov and E. Viola, "Pseudorandom bits for polynomials," in *Proc. 48th Annual IEEE Symposium on Foundations of Computer Science*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 41–51.

[14] S. Lovett, "Unconditional pseudorandom generators for low degree polynomials," *Theory of Computing*, vol. 5, no. 1, pp. 69–82, 2009.

[15] E. Viola, "The sum of D small-bias generators fools polynomials of degree D," *Computational Complexity*, vol. 18, no. 2, pp. 209–217, 2009.

[16] A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman, "Optimal testing of Reed-Muller codes," *Electronic Colloquium in Computational Complexity*, vol. TR09-086, October 2009.

[17] M. Parnas, D. Ron, and R. Rubinfeld, "Tolerant property testing and distance approximation," *J. Comp. Sys. Sci.*, vol. 72, no. 6, pp. 1012–1042, 2006.

[18] V. Guruswami and A. Rudra, "Tolerant locally testable codes," in *APPROX-RANDOM*. Springer-Verlag, 2005, pp. 306–317.

[19] R. Rubinfeld and M. Sudan, "Robust characterizations of polynomials with applications to program testing," *SIAM J. on Comput.*, vol. 25, pp. 252–271, 1996.

[20] O. Goldreich, S. Goldwasser, and D. Ron, "Property testing and its connection to learning and approximation," *Journal of the ACM*, vol. 45, pp. 653–750, 1998.

[21] T. Kaufman and M. Sudan, "Algebraic property testing: the role of invariance," in *Proc. 40th Annual ACM Symposium on the Theory of Computing*. New York, NY, USA: ACM, 2008, pp. 403–412.

[22] L. Babai, L. Fortnow, and C. Lund, "Non-deterministic exponential time has two-prover interactive protocols," *Computational Complexity*, vol. 1, no. 1, pp. 3–40, 1991.

[23] L. Babai, L. Fortnow, L. A. Levin, and M. Szegedy, "Checking computations in polylogarithmic time," in *Proc. 23rd Annual ACM Symposium on the Theory of Computing*. New York: ACM Press, 1991, pp. 21–32.

[24] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy, "Interactive proofs and the hardness of approximating cliques," *Journal of the ACM*, vol. 43, no. 2, pp. 268–292, 1996.

[25] M. Bellare, D. Coppersmith, J. Håstad, M. Kiwi, and M. Sudan, "Linearity testing over characteristic two," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1781–1795, November 1996.

[26] T. Kaufman and D. Ron, "Testing polynomials over general fields," *SIAM J. on Comput.*, vol. 36, no. 3, pp. 779–802, 2006.

[27] C. S. Jutla, A. C. Patthak, A. Rudra, and D. Zuckerman, "Testing low-degree polynomials over prime fields," in *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004, pp. 423–432.