Lecture Notes on Low Degree Tests

Oded Goldreich*

April 7, 2016

Summary: Let $\mathcal{F}$ be a finite field of prime cardinality, and let $d < |\mathcal{F}|/2$ and $m \in \mathbb{N}$. These notes present a low-degree tester that, given oracle access to a function $f : \mathcal{F}^m \to \mathcal{F}$, queries the function at $d+2$ points and satisfies the following conditions:

1. If f is an m-variate polynomial of (total) degree d, then the tester accepts with probability 1.
2. If f is δ-far from the set of m-variate polynomials of (total) degree d, then the tester rejects with probability at least $\min(0.5\delta, \Omega(d^{-2}))$.

The sequence of queries is generated by selecting at random $\overline{x}$ and $\overline{h}$ uniformly in $\mathcal{F}^m$, and using $\overline{x} + i\overline{h}$ as the i^{th} query.

These notes are based on the work of Rubinfeld and Sudan [18]; specificlally, Section 4 is based on [18, Sec. 4], whereas Section 3 is based on [18, Apdx.].

1 A brief introduction

Polynomials of bounded individual degree and of bounded total degree are the most natural classes of functions over the vector space $\mathcal{F}^m$, where $\mathcal{F}$ is a finite field and $m \in \mathbb{N}$. Indeed, such polynomials are ubiquitous in this context, and linear functions over $\mathcal{F}$ are an important special case.

For a finite field $\mathcal{F}$ and any $m \in \mathbb{N}$, any function $f : \mathcal{F}^m \to \mathcal{F}$ can be written as a polynomial of individual degree $|\mathcal{F}| - 1$; that is, as a polynomial that has degree at most $|\mathcal{F}| - 1$ in each variable, and hence has total degree $m \cdot (|\mathcal{F}| - 1)$ (see Exercise 1). Hence, one may say that f is a low degree polynomial if it has degree that is specifically lower than that. Specifically, in this chapter, we call f a low degree polynomial if it has (total) degree at most d, where $d < |\mathcal{F}|/2$ is a parameter. Testing whether a function is a low degree is a natural computational problem, which has direct applications to several areas of the theory of computation, most notably to the design of PCPs and error correcting codes.

Notation: Fixing a finite field $\mathcal{F}$ and an integer m, we often distinguish m-dimentional vectors over $\mathcal{F}$ from elements of $\mathcal{F}$ by overlining the former. In particular, $e\overline{v}$ denotes the scalar multiplication of the vector $\overline{v} \in \mathcal{F}^m$ by the scalar $e \in \mathcal{F}$; that is, if $\overline{v} = (v_1, ..., v_m)$, then $e\overline{v} = (ev_1, ..., ev_m)$.

*Department of Computer Science, Weizmann Institute of Science, Rehovot, Israel.

2 A kind of intuition (which may be skipped)

In this section, we attempt to provide some intuition for the construction of low degree tests. We start with the univariate case, and then move to the multivariate case.

2.1 The univariate case

For $d \ll |\mathcal{F}|$, a natural way of testing whether $f : \mathcal{F} \to \mathcal{F}$ is a (univariate) polynomial of degree (at most) d is to check that the values of f at $d+2$ distinct random points match some degree d polynomial. Before analyzing this tester, note that it uses $d+2$ queries to the function f, whereas the size of the field $\mathcal{F}$ may be much larger. Note that this tester can be viewed as first finding (by extrapolation) the (unique) degree d polynomial that fits the values of f on the first $d+1$ points, and then checking that this polynomial agrees with f on the $d+2^{\text{nd}}$ point.

The analysis of this tester relies on the fact that the distance of f from the set of polynomials of degree d is upper-bounded by the distance of f to the (unique) degree d polynomial f' that fits the values of f on the first $d+1$ points. Now, since the $d+2^{\text{nd}}$ point is uniformly distributed among the other $|\mathcal{F}|-(d+1)$ points of $\mathcal{F}$, it follows that this point hits a point of disagreement (between f and f') with probability at least $\frac{\delta(f,f')\cdot|\mathcal{F}|}{|\mathcal{F}|-(d+1)} > \delta(f,f')$, which is at least the distance of f from being a polynomial of degree d. (Indeed, the foregoing analysis is oblivious of the distribution of the first $d+1$ points, which may even be fixed; it only requires that the $d+2^{\text{nd}}$ point is uniformly distributed (conditioned on the prior points).)

An alternative low degree test. Confining ourselves to the case of finite fields of prime cardinality (where the field $\mathcal{F}$ consists of the set $\mathbb{Z}_{|\mathcal{F}|} = \{0,1,...,|\mathcal{F}|-1\}$ with addition and multiplication modulo $|\mathcal{F}|$), we consider an alternative low degree test (for the univariate case), since this test will be implicitly used later. The tester selects uniformly $r, s \in \mathcal{F}$, and checks that the values of f at $r, r+s, ..., r+(d+1)\cdot s$ match some degree d polynomial. For starters, one can show that, for any $s \neq 0$, it holds that f is a degree d polynomial if and only if for every $r \in \mathcal{F}$ the values of f at $r, r+s, ..., r+(d+1)\cdot s$ match some degree d polynomial.[1] But *how does the rejection probability of this tester relate to the distance of f from the set of degree d polynomials of degree d?*

The answer to the foregoing question follows as a special case of the analysis of the tester outlined below for the class of low degree m-variate polynomials. Indeed, we would welcome a simpler analysis of the univariate case (or a simple reduction of the multivariate case to the univariate case). But, at this point, we wish to proceed with the intuition.

[1] Obviously, if f has degree d, then its values at any subset of $\mathcal{F}$ match a degree d polynomial. As usual throughout this lecture, the opposite direction is considerably less obvious, and its proof is outlined next. That is, we wish to show that, for any $s \in \mathcal{F} \setminus \{0\}$, if for every $r \in \mathcal{F}$ the values of f at $r, r+s, ..., r+(d+1)\cdot s$ match some degree d polynomial, then f is a polynomial of degree d. We start by letting f_r denote the (unique) degree d polynomial that agrees with f on the points $r, r+s, ..., r+d\cdot s$, and observe that (by the hypothesis) it holds that $f_r(r+(d+1)\cdot s) = f(r+(d+1)\cdot s)$. This implies that $f_r = f_{r+s}$, since $f_{r+s}(r+s+d\cdot s) = f(r+s+d\cdot s)$ (by the definition of f_{r+s}), whereas f_r and f_{r+s} are degree d polynomials (which were shown to agree on the $d+1$ points $r+s, ..., r+(d+1)\cdot s$). Using the fact that $(r-r')/s \in \mathcal{F}$ for every $r, r' \in \mathcal{F}$, we infer that all the f_r's are identical, and the claim follows since $f(r) = f_r(r)$ for each $r \in \mathcal{F}$ (by the definition of f_r). We mention that this local characterization of low degree polynomials (which refers to a fixed $s \in \mathcal{F} \setminus \{0\}$) does *not* yield a good tester: see Exercise 2.

2.2 The multivariate case

We now turn to the case of m-variate functions $f : \mathcal{F}^m \to \mathcal{F}$. The first observation here is that f is a degree d polynomial if and only if its values on each line in $\mathcal{F}^m$ can be described by a univariate polynomial of degree d, where a line in $\mathcal{F}^m$ is a ($|\mathcal{F}|$-long) sequence of the form $(\overline{x} + i\overline{h})_{i\in\mathcal{F}}$ such that $\overline{x}, \overline{h} \in \mathcal{F}^m$. One can readily verify that if $f : \mathcal{F}^m \to \mathcal{F}$ is a degree d polynomial, then its values on each line can be described by a univariate polynomial of degree d; that is, the function $f_{\overline{x},\overline{h}} : \mathcal{F} \to \mathcal{F}$ defined as $f_{\overline{x},\overline{h}}(z) = f(\overline{x} + z\overline{h})$ is a polynomial of degree d in z. The opposite direction is less obvious, but it is indeed true (see Theorem 1).

At this point, a natural suggestion is to test that $f : \mathcal{F}^m \to \mathcal{F}$ is of degree d by considering the values of f on a random line in $\mathcal{F}^m$. Recall that if f is not of degree d, then there exists a line such that the values of f on this line do not fit a degree d polynomial. But *if f is ϵ-far from being a degree d polynomial, then how far are its values on a random line from fitting a degree d univariate polynomial*?

The answer to the latter question is far from being obvious. Nevertheless, it is known that the expected distance (of these values from a univariate polynomial) is $\Omega(\epsilon)$, where the expectation is over all possible lines with uniform probability distribution (cf. [10, 4, 13]). In Section 4, we will show a lower bound of $\min(\Omega(\epsilon), \Omega(d^{-2}))$, but we have no real intuition to offer (beyond attempting to present the technical proof in words, an attempt we shall not venture). The actual analysis of the foregoing (low degree) tester mimics the analysis of the linearity tester, but is more complex (in some of its details). Specifically, we define a "self corrected" version of the tested function and show that if the test rejects with small probability, then this corrected version is a polynomial of degree d that is close to the tested function.[2]

As in the case of linearity testing, the only intuition we shall offer is an illustration as to why the "self-corrected" version of the function is a low degree polynomial that is relatively close to the function. The illustration will refer to a function that is obtained by slightly corrupting a low degree polynomial, and so it will only illustrate that the voting scheme emplyed when constructing the self-corrected version makes sense.

2.3 Linking the above intuition to the actual proof

The actual tester, presented in Section 4, tests that a function $f : \mathcal{F}^m \to \mathcal{F}$ is a polynomial of degree (at most) d by checking whether the values of the function restricted to a random line fit

[2] The following outline of the actual analysis is not supposed to be verifiable at this point. It is provided here mainly in order to evoke the analogy to the analysis of the linearity tester (which was presented in Chapter **??**).

Assuming that f is rejected with probability $\rho < 1/O(d^2)$, we shall show that f is 2ρ-close to a low degree polynomial, by taking the following steps (as in the analysis of the linearity tester):

- First we define a "self-corrected" version, denoted f', of the function f such that $f'(x)$ is the most frequent vote cast by the lines passing through x.
- Next, we show that f' is 2ρ-close to f, by using the fact that the vote of a specific line regarding $x \in \mathcal{F}^m$ was defined such that it equals $f(x)$ if and only if the test does not reject when examining (the values on) this line.
- Then, we show that there is a strong majority among the votes (for each point), by lower bounding the collision probability of the random variable that represents a vote of a random line.
- Finally, we show that f' is a low degree polynomial.

The last two steps are performed by showing that each of the corresponding claims can be written as the conjunction of relatively few events that are each related to the check performed by the tester, and using the hypothesis that the rejection probability of the tester (i.e., ρ) is sufficiently small.

a degree d univariate polynomial, where the latter check is performed by considering the values of this restriction on the first $d+2$ points. However, the fact that this line is random means that its starting point as well as the gap between its points are random. Specifically, considering the first $d+2$ values of f on the line $(\overline{x}+i\overline{h})_{i\in\mathcal{F}}$ is analogous to considering the values of f at the points $(r\cdot\overline{x}'+i\cdot s\cdot\overline{h}')_{i=0,...,d+1}$ such that $\overline{x}'=r^{-1}\cdot\overline{x}$ and $\overline{h}'=s^{-1}\cdot\overline{h}$; and, indeed, if $\overline{x},\overline{h}$ are uniformly distributed in $\mathcal{F}^m$, then so are $\overline{x}',\overline{h}'$. Hence, the tester for the m-variate case actually invokes the ("alternative") tester of the univariate case.[3] Furthermore, we use an explicit expression (i.e., Eq. (2)) that captures the decision of the latter tester; that is, we spell out the relation among the aforementioned $d+2$ values of a univariate function such that this relation holds if and only if the univariate function has degree (at most) d.

The analysis of the tester for the m-variate case combines *elements of a reduction* to the univariate case *with an analysis of a specific tester for the univariate case*. Moreover, we refer to the specific expression (i.e., Eq. (2)) used by the univariate tester in making its decision. We stress that this analysis does not present an explicit reduction of the m-variate case to the univariate case, although such reductions can be found elsewhere (see, e.g. [10, 4, 13]). These choices are made in order to make the analysis more concrete and hopefully more clear.

Note: For sake of simplicity, we focus on the case of finite fields of prime cardinality. In this case, the field $\mathcal{F}$ consists of the set $\mathbb{Z}_{|\mathcal{F}|}=\{0,1,...,|\mathcal{F}|-1\}$ with addition and multiplication modulo $|\mathcal{F}|$. In the general case (of arbitrary finite fields), the sequence $(\overline{x}+i\overline{h})_{i=0}^{d+1}$ is replaced by the sequence $(\overline{x}+e\overline{h},\overline{x}+e_1\overline{h},...,\overline{x}+e_{d+1}\overline{h})$, where e is uniformly distributed in $\mathcal{F}$, the e_i's are fixed (distinct) field elements, and the α_i's used in the extrapolation formula (i.e., Eq. (2)) are determined accordingly.

Teaching note: Section 3 provides proofs of two basic facts about polynomials (specifically, Theorems 1 and 2); it is highly technical and offers no intuition (for the reason that the author has none to offer). Unfortunately, these two facts (or rather their combination stated in Corollary 3) are necessary preliminaries for Section 4, which presents the analysis of the tester (which was outlined above). Fortunately, reading Section 4 only requires reading the statement of Corollary 3, and the reader may skip its proof, which is the bulk of Section 3.

3 Background

Throughout this lecture, $\mathcal{F}$ is a finite field of prime cardinality, and d,m are integers such that $d<|\mathcal{F}|/2$. We consider functions $f:\mathcal{F}^m\to\mathcal{F}$, and the class $\mathcal{P}_{m,d}$ of m-variate polynomials of total degree (at most) d. Such functions are called low degree polynomials, because their (total) degree is significantly smaller than $|\mathcal{F}|$.

As shown next, f is in $\mathcal{P}_{m,d}$ if and only if its restriction to each line in $\mathcal{F}^m$ can be represented as a univariate polynomial of degree d, where a line in $\mathcal{F}^m$ is a sequence of the form $L_{\overline{x},\overline{h}}\stackrel{\text{def}}{=}(\overline{x}+i\overline{h})_{i\in\mathcal{F}}$ for $\overline{x},\overline{h}\in\mathcal{F}^m$, and saying that the restriction of f to the line $L_{\overline{x},\overline{h}}$ is represented by the univariate polynomial p means that $p(i)=f(\overline{x}+i\overline{h})$ for every $i\in\mathcal{F}$. Hence, *the global condition of being a*

[3] The point made here is that although the test is described as inspecting the points $0,1,...,d+1$ on a random line, it is actually equivalent to a test that inspecting the points $r,r+s,...,r+(d+1)\cdot s$ on a random line, where s and r are uniformly and independently distributed in $\mathcal{F}$.

degree d polynomial is characterized as the conjunction of $|\mathcal{F}^m|^2$ local conditions, where each local condition refers to the value of the function on $|\mathcal{F}|$ points (on a line in $\mathcal{F}^m$).

Theorem 1 (local characterization of multivariate polynomials): *Let $|\mathcal{F}| > 2d$. The function $f : \mathcal{F}^m \to \mathcal{F}$ is in $\mathcal{P}_{m,d}$ if and only if for every $\overline{x}, \overline{h} \in \mathcal{F}^m$ there exists a degree-d univariate polynomial $p_{\overline{x},\overline{h}}$ such that $p_{\overline{x},\overline{h}}(i) = f(\overline{x} + i\overline{h})$ for every $i \in \mathcal{F}$.*

Proof: Clearly, the restriction of $f \in \mathcal{P}_{m,d}$ to any line in $\mathcal{F}^m$ can be represented as a univariate polynomial of degree d, since for every fixed $\overline{x} = (x_1, ..., x_m) \in \mathcal{F}^m$ and $\overline{h} = (h_1, ..., h_m) \in \mathcal{F}^m$ it holds that $f(\overline{x} + z\overline{h}) = f(x_1 + zh_1, ..., x_m + zh_m)$ is a univariate polynomial of degree d in z.

The opposite direction is not straightforward: it asserts that if the restriction of f to each line in $\mathcal{F}^m$ can be represented as a univariate polynomial of degree d, hereafter referred to as the lines-condition, then $f \in \mathcal{P}_{m,d}$. This claim is proved by induction on m, where the base case (of $m = 1$) is trivial. In the induction step (i.e., going from $m-1$ to m), given an m-variate polynomial $f : \mathcal{F}^m \to \mathcal{F}$ that satisfies the lines-condition, we need to show that $f \in \mathcal{P}_{m,d}$. Towards this end, for every fixed $e \in \mathcal{F}$, we consider the $(m-1)$-variate polynomial f_e defined by $f_e(x_1, ..., x_{m-1}) = f(x_1, ..., x_{m-1}, e)$. By the induction hypothesis, f_e is an $(m-1)$-variate polynomial of degree d (since the restriction of f_e to any line in $\mathcal{F}^{m-1}$ is a degree d univariate polynomial).[4] The following claim implies that f is a polynomial of total degree at most $2d$.

Claim 1.1 (the degree of f is at most $2d$): *For every $e \in \{0, 1, ..., d\}$, let δ_e be the unique degree d univariate polynomial that satisfies $\delta_e(e) = 1$ and $\delta_e(e') = 0$ for every $e' \in \{0, 1, ..., d\} \setminus \{e\}$. Then, $f(\overline{x}) = \sum_{e=0}^{d} \delta_e(x_m) f_e(x_1, ..., x_{m-1})$. Hence, f has degree at most d in x_m, whereas its total degree in $x_1, ..., x_{m-1}$ is at most d.*

Proof: Fixing any $e_1, ..., e_{m-1} \in \mathcal{F}$, we first observe that $g_{e_1,...,e_{m-1}}(x) = f(e_1, ..., e_{m-1}, x)$ is a degree d univariate polynomial in x, since $g_{e_1,...,e_{m-1}}$ describes the restriction of f to the line $L_{(e_1,...,e_{m-1},0),(0,...,0,1)}$ (and f satisfies the lines-condition). Next, we show that $f(e_1, ..., e_{m-1}, x) = \sum_{e=0}^{d} \delta_e(x) f_e(e_1, ..., e_{m-1})$. This holds since each side of the equation is a degree d univariate polynomial in x, whereas these two polynomials agree on $d+1$ points (specifically, for every $e' \in \{0, 1, ..., d\}$, it holds that $\sum_{e=0}^{d} \delta_e(e') f_e(e_1, ..., e_{m-1})$ equals $f_{e'}(e_1, ..., e_{m-1}) = f(e_1, ..., e_{m-1}, e'))$. ■

To show that f is actually of degree d, we consider for each $\overline{h} \in \mathcal{F}^m$ the univariate polynomial $g_{\overline{h}}(z) = f(z\overline{h})$. On the one hand, $\deg(g_{\overline{h}}) \leq d$ for every $\overline{h} \in \mathcal{F}^m$, since $g_{\overline{h}}$ describes the values of f on the line $L_{\overline{0},\overline{h}}$. On the other hand, we shall show next that $\deg(g_{\overline{h}}) = \deg(f)$ for some $\overline{h} \in \mathcal{F}^m$, and $\deg(f) \leq d$ will follow.

Claim 1.2 (the degree of some $g_{\overline{h}}$ equals the degree of f): *There exists $\overline{h} \in \mathcal{F}^m$ such that $\deg(g_{\overline{h}}) = \deg(f)$.*

Proof: We actually prove that, with probability at least $1 - \frac{\deg(f)}{|\mathcal{F}|} > 0$ over the choice of $\overline{h} \in \mathcal{F}^m$, it holds that $\deg(g_{\overline{h}}) = \deg(f)$, where the inequality uses $\deg(f) \leq 2d < |\mathcal{F}|$ (established by Claim 1.1). To prove this claim. consider the coefficient of $z^{\deg(f)}$ in $f(z\overline{h})$. This non-zero coefficient is a polynomial in $\overline{h}$ of total degree at most $\deg(f)$, whereas any non-zero polynomial of degree d' evaluates to zero on at most a $d'/|\mathcal{F}|$ fraction of the points (see Exercise 3). ■

[4] This is the case since the restriction of f_e to any line in $\mathcal{F}^{m-1}$ constitutes a restriction of f to a corresponding line in $\mathcal{F}^m$, whereas f satisfies the lines-condition. In other words, if f satisfies the lines-condition, then so does f_e.

Having proved Claim 1.2, the theorem follows (since for this $\overline{h}$ it holds that $\deg(f) = \deg(g_{\overline{h}}) \leq d$). ■

Notation. For $i = 0, 1, .., d+1$, let $\alpha_i = (-1)^{i+1} \cdot \binom{d+1}{i}$. The α_i's (or rather their values reduced modulo $|\mathcal{F}|$) are viewed as elements of $\mathcal{F}$.

Theorem 2 (local characterization of univariate polynomials): *A univariate polynomial $g : \mathcal{F} \to \mathcal{F}$ has degree $d < |\mathcal{F}|$ if and only if for every $e \in \mathcal{F}$ it holds that*

$$\sum_{i=0}^{d+1} \alpha_i \cdot g(e+i) \;=\; 0 \tag{1}$$

The essence of Theorem 2 is that g is of degree d if and only if its values at $d+1$ points determine its value at any other point. The specific formulation of Theorem 2 refers to $d+2$ specific points (and to a specific extrapolation formula determined by the α_i's). This specific formulation relies on the hypothesis that $0, 1, ..., d+1$ are distinct field elements, which holds since $|\mathcal{F}|$ is postulated to be a *prime* (and $|\mathcal{F}| \geq d+2$).[5] In the general case (i.e., for an arbitrary finite field $\mathcal{F}$ of size at least $d+2$), the sequence $(e+i)_{i=0}^{d+1}$ is replaced by the sequence $(e, e_1, ..., e_{d+1})$, where the e_i's are fixed (distinct) field elements, e varies, and the α_i's used in Eq. (1) are determined accordingly (depending on the e_i's and e).[6] In both cases, the global condition of being a degree d univariate polynomial is characterized as the conjunction of $|\mathcal{F}|$ local conditions, where each local condition refers to the value of the function on $d+2$ points (whereas d may be much smaller than $|\mathcal{F}|$).

Proof: We shall first show that g has degree exactly $d > 0$ if and only if $g'(x) \stackrel{\text{def}}{=} g(x+1) - g(x)$ has degree exactly $d-1$, and then use this fact in order to establish the main claim (i.e., the claim of the theorem) by induction on d.

Given $g : \mathcal{F} \to \mathcal{F}$, we consider the ("derivative") function $g' : \mathcal{F} \to \mathcal{F}$ defined by $g'(x) \stackrel{\text{def}}{=} g(x+1) - g(x)$. We first show that *$g$ has degree exactly $d > 0$ if and only if g' has degree exactly $d-1$*. Writing $g(x) = \sum_{j=0}^{d} c_j \cdot x^j$, where $c_d \neq 0$, we get

$$\begin{aligned} g'(x) &= \sum_{j=0}^{d} c_j \cdot (x+1)^j \;-\; \sum_{j=0}^{d} c_j \cdot x^j \\ &= \sum_{j=0}^{d} c_j \cdot ((x+1)^j - x^j) \\ &= \sum_{j=1}^{d} c_j \cdot \sum_{k=0}^{j-1} \binom{j}{k} \cdot x^k \end{aligned}$$

It follows that the degree of g' is at most $d-1$, whereas the coefficient of x^{d-1} equals $c_d \cdot \binom{d}{d-1} = c_d \cdot d \neq 0$, where the inequality uses $c_d \neq 0$ and $d \in \{1, ..., |\mathcal{F}| - 1\}$. The secondary claim follows.

[5] The case of $|\mathcal{F}| = d+1$ holds trivially, since every function over $\mathcal{F}$ is a polynomial of degree at most $|\mathcal{F}| - 1$.

[6] See Exercise 4. Indeed, in the case that $|\mathcal{F}|$ is a prime, we used $e_i = e + i$ for every $i = 1, ..., d+1$, and the α_i's were independent of e. Hence, in that case we used e_i's that vary with e rather than fixed e_i's. This difference mirrors the difference between the two different testers for the univariate case presented in Section 2.

Teaching note: Roei Tell suggests to make the rest of the proof more transparent by explicitly introducing iterative derivatives, proving that g is of degree d if and only if its $d+1^{\text{st}}$ derivative is identically zero, and showing that this derivative equals $\sum_{i=0}^{k}(-1)^{k-i}\cdot\binom{k}{i}\cdot g(x+i)$. This strategy is detailed in Exercise 5. The author prefers not to introduce an additional notion for the sake of a proof of the current nature, and notes that the actual arguments are analogous.

We now prove the main claim (i.e., the characterization of univariate polynomials via Eq. (1)) by induction on d. For the base case (i.e., $d=0$) we observe that g is a constant function if and only if $-g(e)+g(e+1)=0$ holds for every $e\in\mathcal{F}$. For the induction step (i.e., going from $d-1$ to d), we use the fact that g has degree $d>0$ if and only if g' has degree $d-1$. Using the induction hypothesis, the latter condition coincides with $\sum_{i=0}^{d}(-1)^{i+1}\cdot\binom{d}{i}\cdot g'(e+i) = 0$ for every $e\in\mathcal{F}$. Hence, g has degree d if and only if (for every $e\in\mathcal{F}$)

$$\sum_{i=0}^{d}(-1)^{i+1}\cdot\binom{d}{i}\cdot(g(e+i+1)-g(e+i)) \;=\; 0.$$

Finally, note that

$$\begin{aligned}
&\sum_{i=0}^{d}(-1)^{i+1}\cdot\binom{d}{i}\cdot(g(e+i+1)-g(e+i))\\
&= \sum_{i=0}^{d}(-1)^{i+1}\cdot\binom{d}{i}\cdot g(e+i+1) \;-\; \sum_{i=0}^{d}(-1)^{i+1}\cdot\binom{d}{i}\cdot g(e+i)\\
&= \sum_{j=1}^{d+1}(-1)^{j}\cdot\binom{d}{j-1}\cdot g(e+j) \;+\; \sum_{i=0}^{d}(-1)^{i}\cdot\binom{d}{i}\cdot g(e+i)\\
&= g(e)+(-1)^{d+1}\cdot g(e+d+1) \;+\; \sum_{i=1}^{d}(-1)^{i}\cdot\left(\binom{d}{i-1}+\binom{d}{i}\right)\cdot g(e+i)\\
&= -\sum_{i=0}^{d+1}(-1)^{i+1}\binom{d+1}{i}\cdot g(e+i)
\end{aligned}$$

and the inductive claim follows. ■

Combining Theorems 1 and 2, we get:

Corollary 3 *Let $|\mathcal{F}|>2d$ and $\alpha_i=(-1)^{i+1}\cdot\binom{d+1}{i}$. The function $f:\mathcal{F}^m\to\mathcal{F}$ is in $\mathcal{P}_{m,d}$ if and only if for every $\overline{x},\overline{h}\in\mathcal{F}^m$ it holds that*

$$\sum_{i=0}^{d+1}\alpha_i\cdot f(\overline{x}+i\overline{h}) \;=\; 0 \tag{2}$$

Proof: Clearly (by Theorem 2),[7] any $f\in\mathcal{P}_{m,d}$ satisfies Eq. (2), for every $\overline{x},\overline{h}\in\mathcal{F}^m$. When proving the opposite direction, for every line $L=L_{\overline{x},\overline{h}}$, we use Eq. (2) on the sequence $((\overline{x}+e\overline{h})+i\overline{h})_{i=0}^{d+1}$,

[7] Indeed, we also use the easy direction of Theorem 1.

for each $e \in \mathcal{F}$, and infer (by Theorem 2) that the restriction of f to L is a univariate polynomial of degree d. Specifically, for every line $L = L_{\overline{x},\overline{h}}$, we consider the function $g_L(z) = f(\overline{x} + z\overline{h})$ and infer $\sum_{i=0}^{d+1} \alpha_i g_L(e+i) = 0$ (for each $e \in F$) by using $\sum_{i=0}^{d+1} \alpha_i f((\overline{x} + e\overline{h}) + i\overline{h}) = 0$. (We complete the proof by using the non-obvious direction of Theorem 1.) ■

4 The tester

Recall that we consider functions $f : \mathcal{F}^m \to \mathcal{F}$, where $\mathcal{F}$ be a finite field of prime cardinality, and the class $\mathcal{P}_{m,d}$ of m-variate polynomials of total degree d, which is considered "low" since $d < |\mathcal{F}|/2$.

The characterization provided in Corollary 3 asserts that the global condition $f \in \mathcal{P}_{m,d}$ can be decomposed into $|\mathcal{F}^m|^2$ local conditions, where each local condition refers to the value of f at $d+2$ points in $\mathcal{F}^m$. Such a decomposition, yielding a characterization via a conjunction of many local conditions, is a highly non-obvious phenomenon. It is even more non-obvious that the corresponding characterization is robust in the sense that the fraction of unsatisfied local conditions is related to the distance of the object from the global condition.[8]

A parenthetical discussion. Note that while a characterization states a qualitative dichotomy (i.e., X holds if and only if Y holds), a robust characterization is a quantitative version that relates the "level of violation" of each of its "sides" (i.e., X is "δ-close to being satisfied" if and only if Y is "ρ-close to being satisfied"). The notion of closeness used here need not coincide with the notion of closeness used throughout this course. Still, in the specific case discussed here there is a correspondence: What we shall show is that f is δ-close to $\mathcal{P}_{m,d}$ if and only if a $1 - \Theta_d(\delta)$ fraction of the local conditions concerning f are satisfied, where the notation Θ_d hides factors that depend (polynomially) on d. Actually, we shall only show that if $f \in \mathcal{P}_{m,d}$ then all local conditions are satisfied, whereas if f is δ-far from $\mathcal{P}_{m,d}$ then at least a $\min(\Omega(\delta), \Omega_d(1))$ fraction of the local conditions are unsatisfied.[9]

The foregoing discussion leads to the following tester, which selects a local condition at random among the $|\mathcal{F}^m|^2$ conditions stated in Corollary 3 (or rather in the characterization provided by it).

Algorithm 4 (testing whether f is in $\mathcal{P}_{m,d}$): *Select uniformly, $\overline{x}, \overline{h} \in \mathcal{F}^m$, query f at the points $\overline{x}, \overline{x} + \overline{h}, ..., \overline{x} + (d+1)\overline{h}$ and accept if and only if these values satisfy* Eq. (2). *That is, the tester accepts if and only if*

$$\sum_{i=0}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}) \;=\; 0 \tag{3}$$

where $\alpha_i = (-1)^{i+1} \cdot \binom{d+1}{i}$.

The test checks whether the degree d univariate polynomial that interpolates the values of f on the first $d+1$ points on a random line agrees with the value assigned by f to the $d+2^{\text{nd}}$ point (on that

[8] Artificial examples where a local characterization is not robust are easy to generate; for example, we can augment any local characterization by many copies of the same local conditions (or insignificant variants of the same condition). Natural examples also exist: one such example is provided by Exercise 2.

[9] The reader can easily verify that if f is δ-close to $\mathcal{P}_{m,d}$, then at most a $O_d(\delta)$ fraction of the local conditions are unsatisfied. This follows from the fact that each of the $d+2$ queries made by the following tester is uniformly distributed in $\mathcal{F}^m$.

line). In other words, the test checks whether the value extrapolated for the $d+2^{\text{nd}}$ point based on the first $d+1$ points matches the actual value of that point (according to f itself). The fact that we use "evenly spaced" points as the $d+2$ points on the (random) line is inessential to the validity of this tester, but it allows to present an explicit extrapolation formula (in the case that $|\mathcal{F}|$ is prime).

4.1 Analysis of the tester

Recall that (by Corollary 3) $f \in \mathcal{P}_{m,d}$ if and only if Eq. (3) holds for every $\overline{x}, \overline{h} \in \mathcal{F}^m$. At times, it will be useful to write Eq. (3) as $f(\overline{x}) = \sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h})$, which asserts that the value of $f \in \mathcal{P}_{m,d}$ at $\overline{x}$ is determined (via extrapolation) by the value of f on $d+1$ points on the line $L_{\overline{x},\overline{h}}$.

Theorem 5 (analysis of Algorithm 4): *Let $\delta_0 = 1/(d+2)^2$. Then, Algorithm 4 is a* (one-sided error) *proximity oblivious tester with detection probability* $\min(\delta, \delta_0)/2$*, where δ denotes the distance of the given function from $\mathcal{P}_{m,d}$.*

> **Teaching note:** The following proof uses the strategy used in the ("full") analysis of the linearity tester of Blum, Luby, and Rubinfeld [8], as presented in the lecture on that topic. Indeed, the implementation of this strategy is more complex in the current setting (of low degree testing).

Proof: By (the easier direction of) Corollary 3, each $f \in \mathcal{P}_{m,d}$ is accepted by the tester with probability 1. Hence, the theorem follows by proving that if f is at distance δ from $\mathcal{P}_{m,d}$, then it is accepted by the tester with probability at most $1 - \min(\delta, \delta_0)/2$. Towards this goal, we denote by ρ the probability that f is rejected, and show that if $\rho < \delta_0/2$, then f is 2ρ-close to $\mathcal{P}_{m,d}$.[10] This is shown by presenting a function g and proving that g is 2ρ-close to f and that g is in $\mathcal{P}_{m,d}$.

The intuition underlying the proof is that the hypothesis regarding f (i.e., that it is rejected with probability $\rho < \delta_0/2$) implies that f can be modified (or "corrected") into a low degree polynomial by modifying f on relatively few values (i.e., on at most $2\rho \cdot |\mathcal{F}^m|$ values). Specifically, the hypothesis that $\mathbf{Pr}_{\overline{x},\overline{h} \in \mathcal{F}^m}[f(\overline{x}) \neq \sum_{i \in [d+1]} \alpha_i \cdot f(\overline{x} + i\overline{h})] = \rho < 1/2(d+2)^2$ suggests that a "corrected" version of f that is determined according to the most frequent value of $\sum_{i \in [d+1]} \alpha_i \cdot f(\overline{x} + i\overline{h})$, when considering all possible choices of $\overline{h} \in \mathcal{F}^m$, is a polynomial of degree d that is relatively close to f. Suppose, for illustration, that f is obtained by selecting an arbitrary degree d polynomial p and corrupting it on relatively few points (say on less than $|\mathcal{F}^m|/2(d+1)$ points). Then, the corrected version of f will equal p (since for a random $\overline{h} \in \mathcal{F}^m$, with probability at least $1 - (d+1) \cdot \rho > 1/2$ it holds that $\sum_{i \in [d+1]} \alpha_i \cdot f(\overline{x} + i\overline{h}) = \sum_{i \in [d+1]} \alpha_i \cdot p(\overline{x} + i\overline{h})$) and both claims hold (i.e., p is a polynomial of degree d that is relatively close to f). Needless to say, we cannot start with the foregoing assumption[11], but should rather start from an arbitrary f that satisfies

$$\mathbf{Pr}_{\overline{x},\overline{h} \in \mathcal{F}^m} \left[\sum_{i=0}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}) = 0 \right] = 1 - \rho, \tag{4}$$

We now turn to the actual proof.

[10] Hence, either $\rho \geq \delta_0/2$ or $\rho \geq \delta/2$, which implies $\rho \geq \min(\delta, \delta_0)/2$ as claimed.

[11] The gap between this illustration and the actual proof is reflected in the fact that the illustration refers to $\delta < 1/2(d+1)$, whereas the actual proof uses $\rho < 1/2(d+2)^2$.

Recall that assuming that $\rho < \delta_0/2$, we intend to present a function $g : \mathcal{F}^m \to \mathcal{F}$ and prove that g is 2ρ-close to f and that g is in $\mathcal{P}_{m,d}$. In accordance with the foregoing discussion, we define $g(\overline{x})$ as the most likely value of $\sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h})$, when $\overline{h}$ is uniformly distributed. In other words, letting $\mathtt{MAJ}_{e \in S}\{v_e\}$ denote the most frequently occurring value of v_e when $e \in S$ (with ties broken arbitrarily), we define

$$g(\overline{x}) \stackrel{\text{def}}{=} \mathtt{MAJ}_{\overline{h} \in \mathcal{F}^m} \left\{ \sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}) \right\} \tag{5}$$

Indeed, by Eq. (4), the function g is likely to agree with f on a random $\overline{x} \in \mathcal{F}^m$, and so g is likely to satisfy Eq. (3) on random $\overline{x}, \overline{h} \in \mathcal{F}^m$. However, we need much stronger assertions than the one just made, and stronger assertions will indeed be provided by the following claims.

Claim 5.1 (closeness): *The function g is 2ρ-close to f.*

Proof: This is merely an averaging argument, which counts as bad any point $\overline{x}$ such that Eq. (3) is satisfied by at most half of the possible $\overline{h}$'s, while noting that otherwise g agrees with f on $\overline{x}$. Details follow.

Let B denote the set of $\overline{x}$'s such that Eq. (3) is satisfied by at most half of the possible $\overline{h}$'s; that is, $\overline{x} \in B$ if and only if

$$\mathbf{Pr}_{\overline{h} \in \mathcal{F}^m} \left[\sum_{i=0}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}) = 0 \right] \leq 0.5.$$

By Eq. (4), $\mathbf{Pr}_{\overline{x} \in \mathcal{F}^m}[\overline{x} \in B] \leq 2\rho$, because otherwise $\mathbf{Pr}_{\overline{x},\overline{h} \in \mathcal{F}^m} \left[\sum_{i=0}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}) \neq 0 \right]$ is greater than $2\rho \cdot 0.5$. On the other hand, for every $\overline{x} \in \mathcal{F}^m \setminus B$, it holds that

$$\mathbf{Pr}_{\overline{h} \in \mathcal{F}^m} \left[f(\overline{x}) = \sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}) \right] > 0.5,$$

which implies that $f(\overline{x})$ is the majority value (obtained by the r.h.s of the foregoing random variable) and hence $f(\overline{x}) = g(\overline{x})$. ■

Recall that $g(\overline{x})$ was defined to equal the most frequent value of $\sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h})$, where frequencies were taken over all possible $\overline{h} \in \mathcal{F}^m$. Hence, $g(\overline{x})$ occurs with frequency at least $1/|\mathcal{F}|$ (yet, we saw, in the proof of Claim 5.1, that on at least $1 - 2\rho$ of the $\overline{x}$'s it holds that $g(\overline{x})$ is the majority value). We next show that $g(\overline{x})$ is much more frequent: it occurs in a strong majority (for every $\overline{x}$).

Claim 5.2 (strong majority): *For every $\overline{x} \in \mathcal{F}^m$, it holds that*

$$\mathbf{Pr}_{\overline{h} \in \mathcal{F}^m} \left[g(\overline{x}) = \sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}) \right] \geq 1 - 2(d+1)\rho.$$

Proof: For each $\overline{x} \in \mathcal{F}^m$, we consider the random variable $Z_{\overline{x}}(\overline{h})$ defined to equal $\sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h})$, where the probability space is uniform over the choice of $\overline{h} \in \mathcal{F}^m$. By Eq. (4), we have $\mathbf{Pr}_{\overline{x} \in \mathcal{F}^m}[f(\overline{x}) = Z_{\overline{x}}] = 1 - \rho$, which means that for typical $\overline{x}$ the value $Z_{\overline{x}}$ is almost always a

fixed value (i.e., $f(\overline{x})$), which implies that $Z_{\overline{x}} = g(\overline{x})$ with high probability. However, we want to establish such a statement for any $\overline{x}$, not only for typical ones.

Fixing any $\overline{x} \in \mathcal{F}^m$, the idea is to lower-bound the collision probability of $Z_{\overline{x}}$, which equals $\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}[Z_{\overline{x}}(\overline{h}_1) = Z_{\overline{x}}(\overline{h}_2)]$. (If this lower bound is greater than half, then the same lower bound would holds for $\mathbf{Pr}[Z_{\overline{x}} = g(\overline{x})]$.) Recalling that $Z_{\overline{x}}(\overline{h}) = \sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h})$, we consider

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}_1) = \sum_{i=1}^{d+1} \alpha_i \cdot f(\overline{x} + i\overline{h}_2)\right]. \tag{6}$$

The key observation is that each point on each of these two lines (i.e., $L_{\overline{x},\overline{h}_1}$ and $L_{\overline{x},\overline{h}_2}$)[12] is uniformly distributed in $\mathcal{F}^m$, and hence we can apply Eq. (4) to such a point using a random direction. Furthermore, we can use the direction $\overline{h}_2$ (resp., $\overline{h}_1$) for the points on $L_{\overline{x},\overline{h}_1}$ (resp., $L_{\overline{x},\overline{h}_2}$), which allows to express each of the two sums in Eq. (6) by the same double summation (see Figure 1, which illustrates that the j^{th} point on the line $L_{\overline{x}+i\overline{h}_1,\overline{h}_2}$ coincides with the i^{th} point on the line $L_{\overline{x}+j\overline{h}_2,\overline{h}_1}$).[13] As shown below, it follows that the collision probability of $Z_{\overline{x}}$ is lower bounded by $1-2(d+1)\cdot\rho$, and consequently the most frequent value of $Z_{\overline{x}}$, which is $g(\overline{x})$, occurs with probability at least $1 - 2(d+1)\rho$.

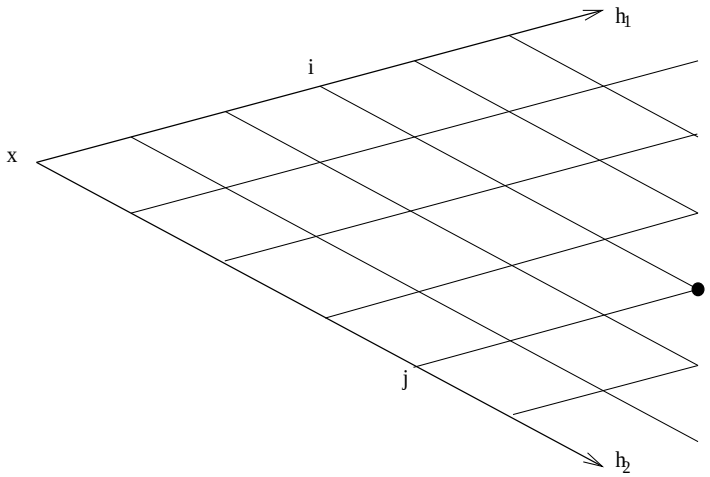

Figure 1: The lattice spanned by $\overline{h}_1$ and $\overline{h}_2$, and the point $\overline{x} + i\overline{h}_1 + j\overline{h}_2$.

We now turn to the actual proof, where an arbitrary $\overline{x} \in \mathcal{F}^m$ is fixed (for the entire proof). For every $i, j \in [d+1]$, if $\overline{h}_1$ and $\overline{h}_2$ are uniformly and independently distributed in $\mathcal{F}^m$, then so are $\overline{x} + i\overline{h}_1$ and $j\overline{h}_2$ (resp., $\overline{x} + j\overline{h}_2$ and $i\overline{h}_1$). By Eq. (4), for every $i \in [d+1]$, it follows that,

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[f(\overline{x} + i\overline{h}_1) = \sum_{j=1}^{d+1} \alpha_j \cdot f((\overline{x} + i\overline{h}_1) + j\overline{h}_2)\right] \;=\; 1 - \rho, \tag{7}$$

and likewise for every $j \in [d+1]$,

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[f(\overline{x} + j\overline{h}_2) = \sum_{i=1}^{d+1} \alpha_i \cdot f((\overline{x} + j\overline{h}_2) + i\overline{h}_1)\right] \;=\; 1 - \rho. \tag{8}$$

[12] Recall that $L_{\overline{x},\overline{h}} = (\overline{x} + i\overline{h})_{i\in\mathcal{F}}$.

[13] Indeed, this merely uses $(\overline{x} + i\overline{h}_1) + j\overline{h}_2 = (\overline{x} + j\overline{h}_2) + i\overline{h}_1$.

Hence, using a union bound (over $i \in [d+1]$ (resp., $j \in [d+1]$)), we have

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{i=1}^{d+1}\alpha_i f(\overline{x}+i\overline{h}_1)=\sum_{i=1}^{d+1}\sum_{j=1}^{d+1}\alpha_i\alpha_j\cdot f(\overline{x}+i\overline{h}_1+j\overline{h}_2)\right] \geq 1-(d+1)\cdot\rho \quad (9)$$

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{j=1}^{d+1}\alpha_j f(\overline{x}+j\overline{h}_2)=\sum_{j=1}^{d+1}\sum_{i=1}^{d+1}\alpha_i\alpha_j\cdot f(\overline{x}+i\overline{h}_1+j\overline{h}_2)\right] \geq 1-(d+1)\cdot\rho, \quad (10)$$

which implies (by a union bound on Eq. (9)&(10)) that

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{i=1}^{d+1}\alpha_i f(\overline{x}+i\overline{h}_1)=\sum_{j=1}^{d+1}\alpha_j f(\overline{x}+j\overline{h}_2)\right] \geq 1-2(d+1)\rho. \quad (11)$$

Note that the two summations in Eq. (11) represent two independent (and identically distributed) random variables, which are functions of $\overline{h}_1$ and $\overline{h}_2$ respectively. Furthermore, each of these summations is distributed identically to the random variable $Z = Z_{\overline{x}}(\overline{h}) \stackrel{\text{def}}{=} \sum_{i=1}^{d+1}\alpha_i f(\overline{x}+i\overline{h})$, which is a function of a uniformly distributed $\overline{h} \in \mathcal{F}^m$. This means that the collision probability of Z (which equals $\sum_u \mathbf{Pr}[Z=u]^2$) is at least $1-2(d+1)\rho$, which implies that the most frequent value occurs in Z with probability at least $1-2(d+1)\rho$ (since if v is the most frequent value assigned to Z then $\sum_u \mathbf{Pr}[Z=u]^2 \leq \sum_u \mathbf{Pr}[Z=v]\cdot\mathbf{Pr}[Z=u] = \mathbf{Pr}[Z=v]$). ■

Using Claim 5.2, we now show that $g \in \mathcal{P}_{m,d}$. This follows by combining Claim 5.3 with the characterization of $\mathcal{P}_{m,d}$.

Claim 5.3 ($g \in \mathcal{P}_{m,d}$): *For every $\overline{x},\overline{h} \in \mathcal{F}^m$, it holds that $\sum_{i=0}^{d+1}\alpha_i\cdot g(\overline{x}+i\overline{h})=0$.*

Proof: As in the proof of the analogous claim in the analysis of the linearity test, we prove the claim by considering a fictitious probabilistic expression regarding the event $\sum_{i=0}^{d+1}\alpha_i\cdot g(\overline{x}+i\overline{h})=0$, when $\overline{x}$ and $\overline{h}$ are fixed. That is, fixing any $\overline{x},\overline{h}\in\mathcal{F}^m$, we prove that $\sum_{i=0}^{d+1}\alpha_i\cdot g(\overline{x}+i\overline{h})=0$ by showing that $\mathbf{Pr}_{\overline{h}_1,\overline{h}_2}[\sum_{i=0}^{d+1}\alpha_i\cdot g(\overline{x}+i\overline{h})=0]>0$. (The random directions $\overline{h}_1$ and $\overline{h}_2$ will be used to set-up a lattice of random points and argue about them in a way that is similar to the proof of Claim 5.2, although the specific lattice and the arguments will be different.)[14]

Fixing any $\overline{x},\overline{h}\in\mathcal{F}^m$ and using Claim 5.2, we infer that, for each $i\in\{0,1,...,d+1\}$, it holds that

$$\mathbf{Pr}_{\overline{h}'\in\mathcal{F}^m}\left[g(\overline{x}+i\overline{h})=\sum_{j=1}^{d+1}\alpha_j\cdot f((\overline{x}+i\overline{h})+j\overline{h}')\right] \geq 1-2(d+1)\rho. \quad (12)$$

Rather than using the same direction $\overline{h}'$ for each i, we use pairwise independent directions such that the direction $\overline{h}_1+i\overline{h}_2$ is used for approximating $g(\overline{x}+i\overline{h})$, which means that we extrapolate (at the point $\overline{x}+i\overline{h}$) according to the line $L_i = L_{\overline{x}+i\overline{h},\overline{h}_1+i\overline{h}_2}$. Hence, the j^{th} point on the line L_i is $(\overline{x}+i\overline{h})+j\cdot(\overline{h}_1+i\overline{h}_2)$, which can be written as $(\overline{x}+j\overline{h}_1)+i\cdot(\overline{h}+j\overline{h}_2)$ (see Figure 2). Now, by

[14] In particular, in the proof of Claim 5.2 we used the lattice points $\overline{x}+i\overline{h}_1+j\overline{h}_2$ for $i,j\in[d+1]$, whereas here we shall use the lattice points $\overline{x}+i\overline{h}+j\overline{h}_1+ij\overline{h}_2$ for $(i,j)\in\{0,1,...,d+1\}\times[d+1]$.

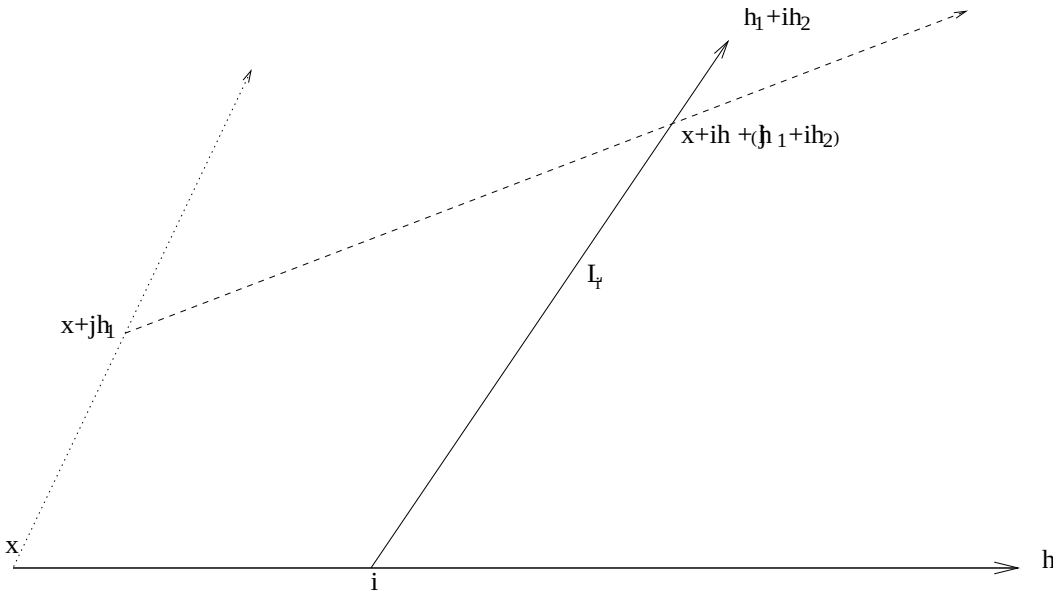

Figure 2: The j^{th} point on the (solid) line $L_i = L_{\overline{x}+i\overline{h},\overline{h}_1+i\overline{h}_2}$ is reached as the i^{th} point on the (dashed) line $L_{\overline{x}+j\overline{h}_1,\overline{h}+j\overline{h}_2}$.

the Eq. (4), for every $j \in [d+1]$ it holds that

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{i=0}^{d+1}\alpha_i\cdot f((\overline{x}+j\overline{h}_1)+i\cdot(\overline{h}+j\overline{h}_2))=0\right] \;=\; 1-\rho. \tag{13}$$

since $\overline{x}+j\overline{h}_1$ and $\overline{h}+j\overline{h}_2$ are uniformly and independently distributed in $\mathcal{F}^m$. (This fact as well as the rest of the argument will be farther detailed below.) Now, when all equalities captured in Eq. (12)&(13) hold, which happens with probability at least $1-(d+2)\cdot 2(d+1)\rho-(d+1)\cdot\rho$, we get

$$\begin{aligned}\sum_{i=0}^{d+1}\alpha_i\cdot g(\overline{x}+i\overline{h}) &= \sum_{i=0}^{d+1}\alpha_i\cdot\sum_{j=1}^{d+1}\alpha_j\cdot f((\overline{x}+i\overline{h})+j\cdot(\overline{h}_1+i\overline{h}_2))\\ &= \sum_{j=1}^{d+1}\alpha_j\cdot\sum_{i=0}^{d+1}\alpha_i\cdot f((\overline{x}+j\overline{h}_1)+i\cdot(\overline{h}+j\overline{h}_2))\\ &= \sum_{j=1}^{d+1}\alpha_j\cdot 0\end{aligned}$$

(where the first equality uses Eq. (12) with $\overline{h}' = \overline{h}_1 + i\overline{h}_2$, and the last one uses Eq. (13)). The claim follows by noting that the event in question (i.e., $\sum_{i=0}^{d+1}\alpha_i\cdot g(\overline{x}+i\overline{h})=0$) is fixed, and so if it occurs with positive probability (according to an analysis carried through in some auxiliary probability space), then it simply holds.

We now turn to the actual proof, which just repeats the foregoing argument while using more precise formulations. Fixing arbitrary $\overline{x},\overline{h}\in\mathcal{F}^m$, let $\overline{h}_1$ and $\overline{h}_2$ be uniformly and independently

distributed in $\mathcal{F}^m$. For every $i \in \{0,1,...,d+1\}$, using Claim 5.2, while noting that $\overline{h}_1 + i\overline{h}_2$ is uniformly distributed in $\mathcal{F}^m$, we get

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[g(\overline{x}+i\overline{h}) = \sum_{j=1}^{d+1}\alpha_j \cdot f((\overline{x}+i\overline{h}) + j(\overline{h}_1 + i\overline{h}_2))\right] \;\geq\; 1 - 2(d+1)\rho. \tag{14}$$

On the other hand, for every $j \in [d+1]$, noting that $\overline{x} + j\overline{h}_1$ and $\overline{h} + j\overline{h}_2$ are uniformly and independently distributed in $\mathcal{F}^m$, and using Eq. (4), we get

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{i=0}^{d+1}\alpha_i \cdot f((\overline{x}+j\overline{h}_1) + i(\overline{h}+j\overline{h}_2)) = 0\right] \;=\; 1-\rho. \tag{15}$$

Note that the argument to f (i.e., $(\overline{x}+j\overline{h}_1) + i(\overline{h}+j\overline{h}_2)$) can be written as $(\overline{x}+i\overline{h}) + j(\overline{h}_1+i\overline{h}_2)$. Hence, we get

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{j=1}^{d+1}\alpha_j \sum_{i=0}^{d+1}\alpha_i \cdot f((\overline{x}+i\overline{h}) + j(\overline{h}_1+i\overline{h}_2)) = 0\right] \;\geq\; 1-(d+1)\cdot\rho. \tag{16}$$

Combining Eq. (14)&(16), we get

$$\begin{aligned}&\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{i=0}^{d+1}\alpha_i g(\overline{x}+i\overline{h}) = \sum_{i=0}^{d+1}\alpha_i \sum_{j=1}^{d+1}\alpha_j \cdot f((\overline{x}+i\overline{h}) + j(\overline{h}_1+i\overline{h}_2)) = 0\right]\\ &\geq\quad 1-(d+2)\cdot 2(d+1)\rho - (d+1)\cdot\rho.\end{aligned}$$

Using $(2d+5)\cdot(d+1)\rho < 1$ (which follows from $\rho \leq 1/2(d+2)^2$), we get

$$\mathbf{Pr}_{\overline{h}_1,\overline{h}_2\in\mathcal{F}^m}\left[\sum_{i=0}^{d+1}\alpha_i g(\overline{x}+i\overline{h}) = 0\right] \;>\; 0 \tag{17}$$

and the claim follows (since $\sum_{i=0}^{d+1}\alpha_i g(\overline{x}+i\overline{h}) = 0$ is independent of the choice of $\overline{h}_1, \overline{h}_2 \in \mathcal{F}^m$).[15] ■

Combining Claims 5.1 and 5.3 with the characterization of $\mathcal{P}_{m,d}$ (i.e., Corollary 3),[16] it follows that f is 2ρ-close to $\mathcal{P}_{m,d}$. ■

4.2 Digest (or an abstraction)

We wish to spell out what is actually being used in the proof of Theorem 5. The proof refers to a test for functions of the form $f : D \to R$, where in our application $D = \mathcal{F}^m$ and $R = \mathcal{F}$ (and $t = d+1$), that checks a condition of the form $f(x) = F(f(y_1),...,f(y_t))$, where x is uniformly distributed in D and F is a fixed function. Indeed, at this point we assume nothing about the

[15] Recall that $\overline{x}, \overline{h} \in \mathcal{F}^m$ are fixed. Hence, the probability in Eq. (17) is either 0 and 1, whereas the lower bound rules out 0.

[16] Indeed, here we use the harder direction of Corollary 3.

distribution of $(y_1, ..., y_t)$ conditioned on x, hereafter denoted Y_x. First, a self-corrected version of f, denoted g, is defined by letting $g(x)$ be the most frequent value of $F(f(y_1), ..., f(y_t))$, when $(y_1, ..., y_t) \leftarrow Y_x$. Claim 5.1 holds in this generic setting; that is, if the test rejects with probability ρ, then g is 2ρ-close to f. In the proofs of Claims 5.2 and 5.3, we used additional features of Y_x, detailed next.

One such feature, which is used in both proofs, is that for every $x \in D$ and $i \in [t]$, the i^{th} element in Y_x is uniformly distributed in D. To state the other feature used in the proof of Claim 5.2, we let $Y_x(\omega)$ denote the value of Y_x when ω is a point in the probability space Ω that underlies Y_x (i.e., $Y_x : \Omega \to D^t$). The proof of Claim 5.2 boils down to lower-bounding the collision probability of $F(Y_x)$, for any x, and it uses the hypothesis (or fact) that for every $i, j \in [t]$ and $\omega_1, \omega_2 \in \Omega$ it holds that *the i^{th} element of $Y_v(\omega_1)$ equals the j^{th} element of $Y_u(\omega_2)$, where u is the i^{th} element of $Y_x(\omega_1)$ and v is the j^{th} element of $Y_x(\omega_2)$*. This feature holds when $D = \Omega$ is an additive group and the i^{th} element of $Y_x(\omega)$ equals $x + i\omega$, which is indeed the case in our application.[17]

In the proof of Claim 5.3 we use a more complex feature, which presumes that $D = \Omega$ and views it is an additive group. The actual feature is that for every $i, j \in [t]$ and $\omega, \omega_1, \omega_2 \in \Omega$ it holds that the j^{th} element of $Y_{x+i\omega}(\omega_1 + i\omega_2)$ equals the i^{th} element of $Y_{x+j\omega_1}(\omega + j\omega_2)$, which holds when the i^{th} element of $Y_x(\omega)$ equals $x + i\omega$ (since $(x + i\omega) + j(\omega_1 + i\omega_2)$ equals $(x + j\omega_1) + i(\omega + j\omega_2)$).

5 Chapter notes

We mention that low-degree tests play a key role in the construction of PCP systems, starting with the "first generation" of such constructions [5, 6, 9, 3, 2].

The analysis of Algorithm 4 provided in Theorem 5 is probably not tight. An improved analysis of a related low-degree tester appeared in [10]. This tester selects uniformly $\overline{x}, \overline{h} \in \mathcal{F}^m$ and $i \in \mathcal{F}$, queries f at $\overline{x}, \overline{x}+\overline{h}, ..., \overline{x}+d\overline{h}$ and $\overline{x}+i\overline{h}$, and accepts if and only if there exists a degree d univariate polynomial that agrees with these $d+2$ values (i.e., a polynomial p such that $p(j) = f(\overline{x} + j\overline{h})$ for every $j \in \{0, 1, ..., d, i\}$).[18] Friedl and Sudan [10] showed that the foregoing tester is a (one-sided error) proximity oblivious tester with detection probability $\min(0.124, \delta/2)$, where δ denotes the distance of the given function from $\mathcal{P}_{m,d}$ (and 0.124 can be replaced by any constant c_0 smaller than 1/8).[19]

The low error regime. Our presentation has focused on the "high error regime"; that is, we have only guaranteed small detection probability (e.g., in [10] the detection probability is smaller than 1/8). Equivalently, we asserted that if f is accepted with high probability (i.e., $\alpha = 1 - \rho > 7/8$), then it is close (i.e., 2ρ-close) to $\mathcal{P}_{m,d}$. Subsequent research regarding low degree testing refers to the "low error regime" where one asks what can be said about a function that is accepted with

[17] In that case, $v + i\omega_1 = x + j\omega_2 + i\omega_1 = u + j\omega_2$.

[18] **Advanced comment:** Alternatively, this tester may be viewed as checking whether the degree d univariate polynomial that fits the values of the first $d+1$ points on the (random) line agrees with the value assigned to a random point on this line. In the context of PCP, this tester is often described as having access to two oracles: the function $f : \mathcal{F}^m \to \mathcal{F}$, which is called a "point oracle", and a "line oracle" that assigns a degree d univariate polynomial to each line in $\mathcal{F}^m$ (i.e., the line-oracle is a function from $(\mathcal{F}^m)^2$ to $\mathcal{F}^{d+1}$). In such a case, it is called a line-vs-point tester. We mention that a plane-vs-point tester was also considered (cf. [17]): The plane-oracle assigns to each plane in $\mathcal{F}^m$ (which described by three points in $\mathcal{F}^m$) a degree d bivariate polynomial, which is supposed to describe the value of f when restricted to this plane.

[19] In addition, it is required that $|\mathcal{F}| > c \cdot d$ (rather than $|\mathcal{F}| > 2d$), where c is a constant that depends on c_0.

probability at least 0.01 (or so).[20] It turns out that in this case the function is 0.9934-close to $\mathcal{P}_{m,d}$; that is, if f is accepted with probability at least 0.01, then it agree with some degree d polynomial on at least 0.0066 fraction of the domain. In general, as shown in [4, 17] (using different tests of query complexity poly(d)), if f is accepted with probability at least α, then f is $(1-\Omega(\alpha))$-close to $\mathcal{P}_{m,d}$ (i.e., f agree with some degree d polynomial on at least $\Omega(\alpha)$ fraction of the domain).[21]

Small fields. So far, we have focused on the case of large fields; that is, we assumed that the field is larger than the degree bound (i.e., $|\mathcal{F}| > d$).[22] But, for multivariate polynomials, the case of small fields makes sense too. Alon *et al.* [1] studied the case of the two-element field, denoted GF(2), and presented a low-degree tester of query complexity that is exponential in the degree bound.[23] They also observed that exponential (in the degree bound) query complexity is required in this case. The case of fields of intermediate size (i.e., $|\mathcal{F}| \in (2, d]$) was studied by Kaufman and Ron [15] and Jutla *et al.* [14], who showed that the query complexity in this case is $|\mathcal{F}|^{\Theta(\ell)}$, where $\ell = \lceil (d+1)/(|\mathcal{F}|-1) \rceil$ if $|\mathcal{F}|$ is a prime (and $\ell = \lceil (d+1)/(p^s - p^{s-1}) \rceil$ if $|\mathcal{F}| = p^s$ for a prime p).[24]

Robust characterization. We have alluded to the notion of a robust characterization in some of our intuitive discussions (most conspicuously at the beginning of Section 4), but refrained from using it in the actual proofs. The notions of local charcaterization and its robustness were put forward by Rubinfeld and Sudan [18], and are briefly reviewed in the historical notes section of the first lecture. The interested reader is referred to these two texts.[25] We mention that some subsequent studies of low-degree tests are conducted in terms of the "robustness" of various local characterizations (see, e.g., [10, 4, 17]). For example, the robustness of the "line tester" was defined as the minimum, over all $f \notin \mathcal{P}_{m,d}$, of the ratio of the expected distance of the restriction of f to a random line from $\mathcal{P}_{1,d}$ (i.e., univariate degree d polynomials) versus the distance of f from $\mathcal{P}_{m,d}$.

Invariances. The class $\mathcal{P}_{m,d}$ is invariant under full rank affine transformation on the functions' domain. That is, for every $f : \mathcal{F}^m \to \mathcal{F}$ and any full rank affine transformation $T : \mathcal{F}^m \to \mathcal{F}^m$ it holds that $f \in \mathcal{P}_{m,d}$ if and only if $f \circ T \in \mathcal{P}_{m,d}$. A general study of the complexity of testing properties that are invariant under affine transformation was initiated by Kaufman and Sudan [16], and is surveyed in [19].[26]

[20] The terms "high" and "low" ("error regimes"), refer to the case that $f \notin \mathcal{P}_{m,d}$ and (rightfully) consider the acceptance probability in these cases as an error probability. Hence, accepting a function (not in $\mathcal{P}_{m,d}$) with probability 0.9 is considered to be in the high error regime, whereas accepting this function with probability 0.01 is in the low error regime.

[21] **Advanced comment:** These results assume that $|\mathcal{F}| \geq \text{poly}(d)$, whereas [10] only assumes $|\mathcal{F}| \geq \Theta(d)$. We mention that [13] only requires $|\mathcal{F}| \geq (1+\Omega(1)) \cdot d$.

[22] Actually, we focused on the case that $|\mathcal{F}| > 2d$, which does not cover the special case of $|\mathcal{F}| = 2$ and $d = 1$. We mention that this special case of $|\mathcal{F}| = 2$ and $d = 1$ can be viewed as a special case of group homomorphism testing, which is considered in the previous lecture (i.e., the homomorphism is from the group $\mathcal{F}^m$ to the group $\mathcal{F}$).

[23] **Advanced comment:** They actually presented a proximity-oblivious tester that, for a degree bound d, makes 2^{d+1} queries and has detection probability $\delta/2^q$, where δ denotes the distance of the tested function from being a degree d polynomial. It turns our that their tester has detection probability $\Omega(\delta)$; see [7] (as well as [11] which presents an analogous result for fields of intermediate size).

[24] The latter case is only analyzed in [15].

[25] **Advanced comment:** In the aforementioned historical notes the notion of locality was presented as referring to constant size neighborhoods, but the notion extends to neighborhoods of size poly(d). Actually, the notion extends to neighborhoods of any size that is significantly smaller than the tested object.

[26] Be warned that there have been many subsequent (to [19]) developments in this direction.

Exercises

The following exercises elaborate on comments made in the main text.

Exercise 1 (low degree extensions): *Show that for a finite field $\mathcal{F}$ and any $m \in \mathbb{N}$, any function $f : \mathcal{F}^m \to \mathcal{F}$ can be written as a polynomial of individual degree $|\mathcal{F}| - 1$. More generally, show that for any $H \subseteq \mathcal{F}$ and any function $f : H^m \to \mathcal{F}$ there exists a polynomial p of individual degree $|H| - 1$ such that $p(\overline{x}) = f(\overline{x})$ for every $\overline{x} \in H^m$.*

Guideline: For every $a \in H$, let $\delta_a : \mathcal{F} \to \mathcal{F}$ be such that $\delta_a(z) = \prod_{b \in H \setminus \{a\}} (x-b)/(a-b)$. Consider $p(x_1, ..., x_m) = \sum_{a_1,...,a_m \in H} f(a_1, ..., a_m) \cdot \prod_{i \in [m]} \delta_{a_i}(x_i)$.

Exercise 2 (a failed attempt regarding the univariate case): *Note that in the case of $m = 1$, Algorithm 4 amounts to selecting $r, s \in \mathcal{F}$ uniformly at random, and checking that the values of $f : \mathcal{F} \to \mathcal{F}$ at $r, r+s, ..., r+(d+1) \cdot s$ match some degree d polynomial. Consider the algorithm that selects r uniformly in $\mathcal{F}$, and checks that the values of f at $r, r+1, ..., r+d+1$ match some degree d polynomial. Show that this algorithm does not yield a good tester in the sense that, for $|\mathcal{F}| \gg d$, there exists a function $f : \mathcal{F} \to \mathcal{F}$ that is 0.499-far from being of degree d, whereas the algorithm rejects it with probability $O(d/|\mathcal{F}|)$.*

Guideline: Let $p_1, p_2 : \mathcal{F} \to \mathcal{F}$ be two distinct polynomials of degree d, and let $f(x) = p_1(x)$ if $x \in \{1, ..., \lfloor |\mathcal{F}|/2 \rfloor\}$ and $f(x) = p_2(x)$ otherwise. Then, f is $(0.5 - (d+1)/|\mathcal{F}|)$-far from being a polynomial of degree d, whereas the algorithm rejects f with probability at most $2(d+1)/|\mathcal{F}|$.[27]

Exercise 3 (The Schwartz–Zippel Lemma):[28] *Let $p : \mathcal{F}^m \to \mathcal{F}$ be a non-zero m-variate polynomial of total degree d over a finite field $\mathcal{F}$. Prove that $\mathbf{Pr}_{x \in \mathcal{F}^m}[p(x) = 0] \leq d/|\mathcal{F}|$.*

Guideline: Use induction on the number of variables, m. The base case of $m = 1$ follows by the fact that $p \not\equiv 0$ has at most d roots. In the induction step, assuming that p depends on its last variable, write $p(x) = \sum_{i=0}^{d} p_i(x_1, ..., x_{m-1}) \cdot x_m^i$, where p_i is an $(m-1)$-variate polynomial of degree at most $d - i$, and let i be the largest integer such that p_i is non-zero. Then, using $x' = (x_1, ..., x_{m-1})$, observe that

$$\mathbf{Pr}_{x \in \mathcal{F}^m}[p(x) = 0] \leq \mathbf{Pr}_{x' \in \mathcal{F}^{m-1}}[p_i(x') = 0] + \mathbf{Pr}_{x' \in \mathcal{F}^{m-1}}[p_i(x') \neq 0] \cdot \mathbf{Pr}_{x \in \mathcal{F}^m}[p(x) = 0 | p_i(x') \neq 0],$$

and that, for any fixed x' such that $p_i(x') \neq 0$, the value of $p(x)$ is a non-zero polynomial of degree i in x_m.

Exercise 4 (local characterization of low degree univariate polynomials in the case of general finite fields): *Let $\mathcal{F}$ be an arbitrary finite field and $d < |\mathcal{F}| - 1$. Suppose that $e_1, ..., e_{d+1}$ are distinct field elements. Prove that there exist a sequence of tuples $(\alpha_1^{(e)}, ..., \alpha_{d+1}^{(e)})_{e \in \mathcal{F}}$, where $\alpha_i^{(e)} \in \mathcal{F}$, such that $g : \mathcal{F} \to \mathcal{F}$ is a univariate polynomial of degree d if and only if for every $e \in \mathcal{F}$ it holds that*

$$g(e) \;=\; \sum_{i=1}^{d+1} \alpha_i^{(e)} \cdot g(e_i). \tag{18}$$

[27] The first claim holds because for every polynomial p of degree d there exists $i \in \{1, 2\}$ such that p agrees with p_i on at most d points, which implies that $\delta(p, f) \geq \delta(p, p_i) \geq (\lfloor |F|/2 \rfloor - d)/|\mathcal{F}|$. The second claim holds because the algorithm may reject only if $\{r, r+1, ..., r+d+1\}$ has a non-trivial; intersection with $\{1, ..., |\mathcal{F}|/2\}$.

[28] A more general version is presented in the lecture notes on testing Juntas.

Guideline: First, show that there exists a unique degree d polynomial p that agrees with g on $e_1, ..., e_{d+1}$, by writing $p(x) = \sum_{i=0}^{d} c_i x^i$ and observing that

$$\begin{pmatrix} g(e_1) \\ g(e_2) \\ \vdots \\ g(e_{d+1}) \end{pmatrix} = \begin{pmatrix} 1 & e_1 & \cdots & e_1^d \\ 1 & e_2 & \cdots & e_2^d \\ \vdots & \vdots & \cdots & \vdots \\ 1 & e_d & \cdots & e_{d+1}^d \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_d \end{pmatrix} \tag{19}$$

holds.[29] Furthermore, the c_i's can be expressed as a linear combination of the $g(e_i)$'s. Next, observe that g is a degree d polynomial if and only if $g(e) = \sum_{i=0}^{d} c_i \cdot e^i$ for every $e \in \mathcal{F}$. Finally, set the $\alpha_i^{(e)}$'s accordingly.

Exercise 5 (iterative derivatives and Theorem 2)[30]: *Recall that the proof of Theorem 2 referred to the derivatives of functions $g : \mathcal{F} \to \mathcal{F}$. Here we explicitly define* iterative derivatives*, denoted $\partial^{(i)}$, such that the value of $\partial^{(1)}g = \partial g$ at x equals $g(x+1) - g(x)$ and $\partial^{(i+1)}g = \partial\partial^{(i)}g$ (where $\partial^{(0)}g = g$). Recall that in the first part of the proof of Theorem 2 we showed that, for every $d > 0$, it holds that g has degree d if and only if ∂g has degree $d-1$. Prove the following two facts:*

1. *For every $d \geq 0$ and $g : \mathcal{F} \to \mathcal{F}$, it holds that g has degree d if and only if the function $\partial^{(d+1)}g$ is identically zero.*
2. *For every $k \geq 0$ and $g : \mathcal{F} \to \mathcal{F}$, it holds that the value of $\partial^{(k)}g$ at x equals*

$$\sum_{i=0}^{k} (-1)^{k-i} \cdot \binom{k}{i} \cdot g(x+i).$$

Observe that the combination of these facts establishes Theorem 2.

Guideline: Both facts can be proved by induction (on d and k, resp.).

References

[1] N. Alon, M. Krivelevich, T. Kaufman, S. Litsyn, and D. Ron. Testing low-degree polynomials over GF(2). In *Proceedings of the 7th RANDOM*, Springer LNCS, Vol. 2764, pages 188–199, 2003.

[2] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy. Proof Verification and Intractability of Approximation Problems. *Journal of the ACM*, Vol. 45, pages 501–555, 1998. Preliminary version in *33rd FOCS*, 1992.

[3] S. Arora and S. Safra. Probabilistic Checkable Proofs: A New Characterization of NP. *Journal of the ACM*, Vol. 45, pages 70–122, 1998. Preliminary version in *33rd FOCS*, 1992.

[29] Recall that the matrix in Eq. (19), which is the Vandermonde matrix, is full rank.

[30] The following alternative presentation of the second part of the proof of Theorem 2 was suggested to us by Roei Tell.

[4] S. Arora and M. Sudan. Improved Low-Degree Testing and its Applications. *Combinatorica*, Vol. 23(3), pages 365–426, 2003. Preliminary version in *29th STOC*, 1997.

[5] L. Babai, L. Fortnow, and C. Lund. Non-Deterministic Exponential Time has Two-Prover Interactive Protocols. *Computational Complexity*, Vol. 1, No. 1, pages 3–40, 1991. Preliminary version in *31st FOCS*, 1990.

[6] L. Babai, L. Fortnow, L. Levin, and M. Szegedy. Checking Computations in Polylogarithmic Time. In *23rd ACM Symposium on the Theory of Computing*, pages 21–31, 1991.

[7] A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman. Optimal Testing of Reed-Muller Codes. In *51st IEEE Symposium on Foundations of Computer Science*, pages 488–497, 2010.

[8] M. Blum, M. Luby and R. Rubinfeld. Self-Testing/Correcting with Applications to Numerical Problems. *Journal of Computer and System Science*, Vol. 47, No. 3, pages 549–595, 1993.

[9] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating Clique is almost NP-complete. *Journal of the ACM*, Vol. 43, pages 268–292, 1996. Preliminary version in *32nd FOCS*, 1991.

[10] K. Friedl and M. Sudan. Some Improvements to Total Degree Tests. In the proceedings of the *3rd Israel Symp. on Theory of Computing and Systems*, pages 190–198, 1995. Revision posted on *CoRR*, 1307.3975, 2013.

[11] E. Haramaty, A. Shpilka, and M. Sudan. Optimal Testing of Multivariate Polynomials over Small Prime Fields. *SIAM Journal on Computing*, Vol. 42 (2), pages 536–562, 2013. Preliminary version in *52nd FOCS*, 2011.

[12] O. Goldreich (ed.). *Property Testing: Current Research and Surveys*. Springer, LNCS, Vol. 6390, 2010.

[13] A. Guo, E. Haramaty, and M. Sudan. Robust testing of lifted codes with applications to low-degree testing. *ECCC*, TR15-043, 2015.

[14] C. S. Jutla, A. C. Patthak, A. Rudra, and D. Zuckerman. Testing low-degree polynomials over prime fields. *Random Structures and Algorithms*, Vol. 35 (2), pages 163–193, 2009. Preliminary version in *45th FOCS*, 2004.

[15] T. Kaufman and D. Ron. Testing polynomials over general fields. *SIAM Journal on Computing*, 35(3):779–802, 2006. Preliminary version in *45th FOCS*, 2004.

[16] T. Kaufman and M. Sudan. Algebraic Property Testing: The role of Invariance. In *40th ACM Symposium on the Theory of Computing*, pages 4-3–412, 2008.

[17] R. Raz and S. Safra. A Sub-Constant Error-Probability Low-Degree Test, and a Sub-Constant Error-Probability PCP Characterization of NP. In the proceedings of the *29th ACM Symposium on the Theory of Computing*, pages 475–484, 1997.

[18] R. Rubinfeld and M. Sudan. Robust characterization of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2), pages 252–271, 1996.

[19] M. Sudan. Invariances in Property Testing. In [12].