# Chapter 7

# LOCAL TESTING OF REED-MULLER CODES

From this chapter onwards, we will switch gears and talk about property testing of codes.

## *7.1 Introduction*

A *low degree tester* is a probabilistic algorithm which, given a degree parameter $t$ and oracle access to a function $f$ on $n$ arguments (which take values from some finite field $\mathbb{F}$), has the following behavior. If $f$ is the evaluation of a polynomial on $n$ variables with total degree at most $t$, then the low degree tester must accept with probability one. On the other hand, if $f$ is "far" from being the evaluation of some polynomial on $n$ variables with degree at most $t$, then the tester must reject with constant probability. The tester can query the function $f$ to obtain the evaluation of $f$ at any point. However, the tester must accomplish its task by using as few probes as possible.

Low degree testers play an important part in the construction of Probabilistically Checkable Proofs (or PCPs). In fact, different parameters of low degree testers (for example, the number of probes and the amount of randomness used) directly affect the parameters of the corresponding PCPs as well as various inapproximability results obtained from such PCPs ([36, 5]). Low degree testers also form the core of the proof of MIP = NEXPTIME in [9].

Blum, Luby, and Rubinfeld designed the first low degree tester, which handled the linear case, i.e., $t = 1$ ([21]), although with a different motivation. This was followed by a series of works that gave low degree testers that worked for larger values of the degree parameter ([93, 42, 7]). However, these subsequent results as well as others which use low degree testers ([9, 43]) only work when the degree is smaller than size of the field $\mathbb{F}$. Alon et al. proposed a low degree tester for any nontrivial degree parameter over the binary field $\mathbb{F}_2$ [1].

A natural open problem was to give a low degree tester for all degrees for finite fields of size between two and the degree parameter. In this chapter we (partially) solve this problem by presenting a low degree test for multivariate polynomials over any prime field $\mathbb{F}_p$.

### *7.1.1 Connection to Coding Theory*

The evaluations of polynomials in $n$ variables of degree at most $t$ are well known *Reed-Muller codes* (note that when $n = 1$, we have the Reed-Solomon codes, which we considered in Chapter 6). In particular, the evaluation of polynomials in $n$ variables of degree

at most $t$ over $\mathbb{F}_q$ is the Reed-Muller code or $\mathrm{RM}_q(t,n)$ with parameters $t$ and $n$. These codes have length $q^n$ and dimension $\binom{n+t}{n}$ (see [28, 29, 69] for more details). Therefore, a function has degree $t$ if and only if (the vector of evaluations of) the function is a valid codeword in $\mathrm{RM}_q(n,t)$. In other words, low degree testing is equivalent to locally testing Reed-Muller codes.

### *7.1.2 Overview of Our Results*

It is easier to define our tester over $\mathbb{F}_3$. To test if $f$ has degree at most $t$, set $k = \lceil \frac{t+1}{2} \rceil$, and let $i = (t+1) \pmod 2$. Pick $k$-vectors $y_1, \cdots, y_k$ and $b$ from $\mathbb{F}_3^n$, and test if

$$\sum_{c \in \mathbb{F}_3^k; c=(c_1,\cdots,c_k)} c_1^i f(b + \sum_{j=1}^{k} c_j y_j) = 0,$$

where for notational convenience we use $0^0 = 1$ (and we will stick to this convention throughout this chapter). We remark here that a polynomial of degree at most $t$ always passes the test, whereas a polynomial of degree greater than $t$ gets caught with non-negligible probability $\alpha$. To obtain a constant rejection probability we repeat the test $\Theta(1/\alpha)$ times.

The analysis of our test follows a similar general structure developed by Rubinfeld and Sudan in [93] and borrows techniques from [93, 1]. The presence of a doubly-transitive group suffices for the analysis given in [93]. Essentially we show that the presence of a doubly-transitive group acting on the coordinates of the dual code does indeed allow us to localize the test. However, this gives a weaker result. We use techniques developed in [1] for better results, although the adoption is not immediate. In particular the interplay between certain geometric objects described below and their polynomial representations plays a pivotal role in getting results that are only about a quadratic factor away from optimal query complexity.

In coding theory terminology, we show that Reed-Muller codes over prime fields are locally testable. We further consider a new basis of Reed-Muller code over prime fields that in general differs from the minimum weight basis. This allows us to present a novel exact characterization of the multivariate polynomials of degree $t$ in $n$ variables over prime fields. Our basis has a clean geometric structure in terms of *flats* [69], and unions of parallel flats but with different weights assigned to different parallel flats[1]. The equivalent polynomial and geometric representations allow us to provide an almost optimal test.

#### *Main Result*

Our results may be stated quantitatively as follows. For a given integer $t \geqslant (p-1)$ and a given real $\varepsilon > 0$, our testing algorithm queries $f$ at $O\left(\frac{1}{\varepsilon} + t \cdot p^{\frac{2t}{p-1}+1}\right)$ points to determine

[1]The natural basis given in [28, 29] assigns the same weight to each parallel flat.

whether $f$ can be described by a polynomial of degree at most $t$. If $f$ is indeed a polynomial of degree at most $t$, our algorithm always accepts, and if $f$ has a relative Hamming distance at least $\varepsilon$ from every degree $t$ polynomial, then our algorithm rejects $f$ with probability at least $\frac{1}{2}$. (In the case $t < (p-1)$, our tester still works but more efficient testers are known). Our result is almost optimal since any such testing algorithm must query $f$ in at least $\Omega(\frac{1}{\varepsilon} + p^{\frac{t+1}{p-1}})$ many points (see Corollary 7.5).

We extend our analysis also to obtain a *self-corrector* for $f$ (as defined in [21]), in case the function $f$ is reasonably close to a degree $t$ polynomial. Specifically, we show that the value of the function $f$ at any given point $x \in \mathbb{F}_p^n$ may be obtained with good probability by querying $f$ on $\Theta(p^{t/p})$ random points. Using pairwise independence we can achieve even higher probability by querying $f$ on $p^{O(t/p)}$ random points and using majority logic decoding.

### *7.1.3 Overview of the Analysis*

The design of our tester and its analysis follows the following general paradigm first formalized by Rubinfeld and Sudan [93]. The analysis also uses additional ideas used in [1]. In this section, we review the main steps involved.

The first step is coming up with an *exact characterization* for functions that have low degree. The characterization identifies a collection of subsets of points and a predicate such that an input function is of low degree if and only if for every subset in the collection, the predicate is satisfied by the evaluation of the function at the points in the subset. The second step entails showing that the characterization is a *robust characterization*, that is, the following natural tester is indeed a local tester (see section 2.3 for a formal definition): Pick one of the subsets in the collection uniformly at random and check if the predicate is satisfied by the evaluation of the function on the points in the chosen subset. Note that the number of queries made by the tester is bounded above by the size of the largest subset in the collection.

There is a natural characterization for polynomials of low degree using their alternative interpretation as a RM code. As RM code is a linear code, a function is of low degree if and only if it is orthogonal to every codeword in the dual of the corresponding RM code. The problem with the above characterization is that the resulting local tester will have to make as many queries as the maximum number of non-zero position in any dual codeword, which can be large. To get around this problem, instead of considering all codewords in the dual of the RM code, we consider a collection of dual codewords that have few non-zero positions. To obtain an exact characterization, note that this collection has to generate the dual code.

We use the well known fact that the dual of a RM code is a RM code (with different parameters). Thus, to obtain a collection of dual codewords with low weight that generate the dual of a RM code it is enough to find low weight codewords that generate every RM code. To this end we show that the characteristic vector of any affine subspace (also called a

*flat* in RM terminology [69]) generates certain RM codes. To complete the characterization, we show that any RM code can be generated by flats and certain weighted characteristic vectors of affine subspaces (which we call *pseudoflats*). To prove these we look at the affine subspaces as the intersection of (a fixed number of) hyperplanes and alternatively represent the characteristic vectors as polynomials.

To prove that the above exact characterization is robust we use the *self-correcting* approach ([21, 93]). Given an input $f$ we define a related function $g$ as follows. The value of $g(x)$ is defined to be the most frequently occurring value, or *plurality*, of $f$ at correlated random points. The major part of the analysis is to show that if $f$ disagrees from all low degree polynomials in a lot of places then the tester rejects with high probability.

The analysis proceeds by first showing that $f$ and $g$ agree on most points. Then we show that if the tester rejects with low enough probability then $g$ is a low degree polynomial. In other words, if $f$ is far enough from all low degree polynomials, then the tester rejects with high probability. To complete the proof, we take care of the case when $f$ is close to some low degree polynomial separately.

### *7.2 Preliminaries*

Throughout this chapter, we use $p$ to denote a prime and $q$ to denote a prime power ($p^s$ for some positive integer $s$) to be a prime power. In this chapter, we will mostly deal with prime fields. We therefore restrict most definitions to the prime field setting.

For any $t \in [n(q-1)]$, let $\mathcal{P}_t$ denote the family of all functions over $\mathbb{F}_q^n$ that are polynomials of total degree at most $t$ (and w.l.o.g. individual degree at most $q-1$) in $n$ variables. In particular $f \in \mathcal{P}_t$ if there exists coefficients $a_{(e_1,\cdots,e_n)} \in \mathbb{F}_q$, for every $i \in [n]$, $e_i \in \{0, \cdots, q-1\}$, $\sum_{i=1}^n e_i \leqslant t$, such that

$$f = \sum_{(e_1,\cdots,e_n)\in\{0,\cdots,q-1\}^n; 0\leqslant\sum_{i=1}^n e_i\leqslant t} a_{(e_1,\cdots,e_n)} \prod_{i=1}^n x_i^{e_i}. \tag{7.1}$$

The codeword corresponding to a function will be the evaluation vector of $f$. We recall the definition of the (Primitive) Reed-Muller code as described in [69, 29].

**Definition 7.1.** *Let $V = \mathbb{F}_q^n$ be the vector space of $n$-tuples, for $n \geqslant 1$, over the field $\mathbb{F}_q$. For any $k$ such that $0 \leqslant k \leqslant n(q-1)$, the $k^{th}$ order Reed-Muller code $\mathrm{RM}_q(k,n)$ is the subspace of $\mathbb{F}_q^{|V|}$ of all $n$-variable polynomial functions (reduced modulo $x_i^q - x_i$) of degree at most $k$.*

This implies that the code corresponding to the family of functions $\mathcal{P}_t$ is $\mathrm{RM}_q(t,n)$. Therefore, a characterization for one will simply translate into a characterization for the other.

We will be using terminology defined in Section 2.3. We now briefly review the definitions that are relevant to this chapter. For any two functions $f, g : \mathbb{F}_q^n \to \mathbb{F}_q$, the relative

distance $\delta(f,g) \in [0,1]$ between $f$ and $g$ is defined as $\delta(f,g) \stackrel{def}{=} \Pr_{x \in \mathbb{F}_q^n}[f(x) \neq g(x)]$. For a function $g$ and a family of functions $F$ (defined over the same domain and range), we say $g$ is $\varepsilon$- *close* to $F$, for some $0 < \varepsilon < 1$, if, there exists an $f \in F$, where $\delta(f,g) \leqslant \varepsilon$. Otherwise it is $\varepsilon$- *far* from $F$.

A one sided testing algorithm (*one-sided tester*) for $\mathcal{P}_t$ is a probabilistic algorithm that is given query access to a function $f$ and a distance parameter $\varepsilon$, $0 < \varepsilon < 1$. If $f \in \mathcal{P}_t$, then the tester should always accept $f$ (perfect completeness), and if $f$ is $\varepsilon$-*far* from $\mathcal{P}_t$, then with probability at least $\frac{1}{2}$ the tester should reject $f$.

For vectors $x, y \in \mathbb{F}_p^n$, the dot (scalar) product of $x$ and $y$, denoted $x \cdot y$, is defined to be $\sum_{i=1}^{n} x_i y_i$, where $w_i$ denotes the $i^{th}$ co-ordinate of $w$.

To motivate the next notation which we will use frequently, we give a definition.

**Definition 7.2.** *For any $k \geqslant 0$, a $k$-*flat *in $\mathbb{F}_p^n$ is a $k$-dimensional affine subspace. Let $y_1, \cdots, y_k \in \mathbb{F}_p^n$ be linearly independent vectors and $b \in \mathbb{F}_p^n$ be a point. Then the subset*

$$L = \{\sum_{i=1}^{k} c_i y_i + b | \forall i \in [k]\ c_i \in \mathbb{F}_p\}$$

*is a $k$-dimensional flat. We will say that $L$ is generated by $y_1, \cdots, y_k$ at $b$. The incidence vector of the points in a given $k$-flat will be referred to as the codeword corresponding to the given $k$-flat.*

Given a function $f : \mathbb{F}_p^n \to \mathbb{F}_p$, for $y_1, \cdots, y_l, b \in \mathbb{F}_p^n$ we define

$$T_f^0(y_1, \cdots, y_l, b) \stackrel{def}{=} \sum_{c=(c_1,\cdots,c_l) \in \mathbb{F}_p^l} f(b + \sum_{i \in [l]} c_i y_i), \tag{7.2}$$

which is the sum of the evaluations of function $f$ over an $l$-flat generated by $y_1, \cdots, y_l$, at $b$. Alternatively, as we will see later in Observation 7.4, this can also be interpreted as the dot product of the codeword corresponding to the $l$-flat generated by $y_1, \cdots, y_l$ at $b$ and that corresponding to the function $f$.

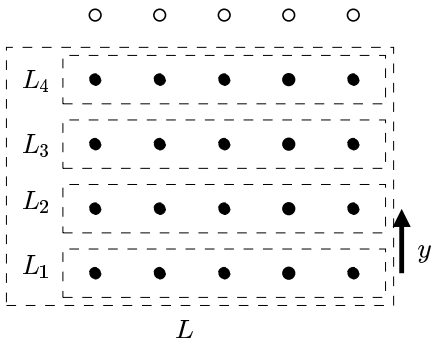
While $k$-flats are well-known, we define a new geometric object, called a pseudoflat. A $k$-pseudoflat is a union of $(p-1)$ parallel $(k-1)$-flats.

**Definition 7.3.** *Let $L_1, L_2, \cdots, L_{p-1}$ be parallel $(k-1)$-flats ($k \geqslant 1$), such that for some $y \in \mathbb{F}_p^n$ and all $t \in [p-2]$, $L_{t+1} = y + L_t$, where for any set $S \subseteq \mathbb{F}_p^n$ and $y \in \mathbb{F}_p^n$, $y + S \stackrel{def}{=} \{x + y | x \in S\}$. We define a $k$-*pseudoflat *to be the union of the set of points $L_1$ to $L_{p-1}$. Further, given an $r$ (where $1 \leqslant r \leqslant p-2$) and a $k$-pseudoflat, we define a $(k,r)$-*pseudoflat vector *as follows. Let $I_j$ be the incidence vector of $L_j$ for $j \in [p-1]$. Then the $(k,r)$-pseudoflat vector is defined to be $\sum_{j=1}^{p-1} j^r I_j$. We will also refer to the $(k,r)$-pseudoflat vector as a codeword.*

*Let $L$ be a $k$-pseudoflat. Also, for $j \in [p-1]$, let $L_j$ be the $(k-1)$-flat generated by $y_1, \cdots, y_{k-1}$ at $b + j \cdot y$, where $y_1, \cdots, y_{k-1}$ are linearly independent. Then we say that the*

*$(k, r)$-pseudoflat vector corresponding to $L$ as well as the pseudoflat $L$, are* generated by $y, y_1, \cdots, y_{k-1}$ *at* $b$ exponentiated along $y$.

See Figure 7.1 for an illustration of the Definition 7.3.




| | | | | |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 4 | 4 | 4 | 4 | 4 |
| 3 | 3 | 3 | 3 | 3 |
| 2 | 2 | 2 | 2 | 2 |
| 1 | 1 | 1 | 1 | 1 |

$(2, 1)$-pseudoflat vector corresponding to $L$

Figure 7.1: Illustration of a $k$-pseudoflat $L$ defined over $\mathbb{F}_p^n$ with $k = 2, p = 5$ and $n = 5$. Picture on the left shows the points in $L$ (recall that each of $L_1, \ldots, L_4$ are 1-flats or lines). Each $L_i$ (for $1 \leqslant i \leqslant 4$) has $p^{k-1} = 5$ points in it. The points in $L$ are shown by filled circles and the points in $\mathbb{F}_5^5 \setminus L$ are shown by unfilled circles. The picture on the right is the $(2, 1)$-pseudoflat corresponding to $L$.

Given a function $f : \mathbb{F}_p^n \to \mathbb{F}_p$, for $y_1, \cdots, y_l, b \in \mathbb{F}_p^n$, for all $i \in [p-2]$, we define

$$T_f^i(y_1, \cdots, y_l, b) \stackrel{def}{=} \sum_{c=(c_1,\cdots,c_l)\in\mathbb{F}_p^l} c_1^i \cdot f(b + \sum_{j\in[l]} c_j y_j). \tag{7.3}$$

As we will see later in Observation 7.5, the above can also be interpreted as the dot product of the codeword corresponding to the $(l, r)$-pseudoflat vector generated by $y_1, \cdots, y_l$ at $b$ exponentiated along $y_1$ and the codeword corresponding to the function $f$.

### *7.2.1 Facts from Finite Fields*

In this section we spell out some facts from finite fields which will be used later. We begin with a simple lemma.

**Lemma 7.1.** *For any $t \in [q-1]$, $\sum_{a\in\mathbb{F}_q} a^t \neq 0$ if and only if $t = q-1$.*

*Proof.* First note that $\sum_{a\in\mathbb{F}_q} a^t = \sum_{a\in\mathbb{F}_q^*} a^t$. Observing that for any $a \in \mathbb{F}_q^*$, $a^{q-1} = 1$, it follows that $\sum_{a\in\mathbb{F}_q^*} a^{q-1} = \sum_{a\in\mathbb{F}_q^*} 1 = -1 \neq 0$.

Next we show that for all $t \neq q-1$, $\sum_{a \in \mathbb{F}_q^*} a^t = 0$. Let $\alpha$ be a generator of $\mathbb{F}_q^*$. The sum can be re-written as $\sum_{i=0}^{q-2} \alpha^{it} = \frac{\alpha^{t(q-1)}-1}{\alpha^t-1}$. The denominator is non-zero for $t \neq q-1$ and thus, the fraction is well defined. The proof is complete by noting that $\alpha^{t(q-1)} = 1$. □

This immediately implies the following lemma.

**Lemma 7.2.** *Let $t_1, \cdots, t_l \in [q-1]$. Then*

$$\sum_{(c_1,\cdots,c_l)\in(\mathbb{F}_q)^l} c_1^{t_1} c_2^{t_2} \cdots c_l^{t_l} \neq 0 \text{ if and only if } t_1 = t_2 = \cdots = t_l = q-1. \tag{7.4}$$

*Proof.* Note that the left hand side can be rewritten as $\prod_{i\in[l]} \left(\sum_{c_i \in \mathbb{F}_q} c_i^{t_i}\right)$. □

We will need to transform products of variables to powers of linear functions in those variables. With this motivation, we present the following identity.

**Lemma 7.3.** *For each $k$, s.t. $0 < k \leqslant (p-1)$ there exists $c_k \in \mathbb{F}_p^*$ such that*

$$c_k \prod_{i=1}^{k} x_i = \sum_{i=1}^{k} (-1)^{k-i} S_i \quad \textit{where} \quad S_i = \sum_{\emptyset \neq I \subseteq [k]; |I|=i} \left(\sum_{j\in I} x_j\right)^k. \tag{7.5}$$

*Proof.* Consider the right hand side of the (7.5). Note that all the monomials are of degree exactly $k$. Also note that $\prod_{i=1}^{k} x_i$ appears only in the $S_k$ and nowhere else. Now consider any other monomial of degree $k$ that has a support of size $j$, where $0 < j < k$: w.l.o.g. assume that this monomial is $M = x_1^{i_1} x_2^{i_2} \cdots x_j^{i_j}$ such that $i_1 + \cdots + i_j = k$. Now note that for any $I \supseteq [j]$, $M$ appears with a coefficient of $\binom{k}{i_1, i_2, \ldots, i_j}$ in the expansion of $(\sum_{\ell \in I} x_\ell)^k$. Further for every $i \geqslant j$, the number of choices of $I \supseteq [j]$ with $|I| = i$ is exactly $\binom{k-j}{k-i}$. Therefore, summing up the coefficients of $M$ in the various summands $S_i$ (along with the $(-1)^{k-i}$ factor), we get that the coefficient of $M$ in the right hand side of (7.5) is

$$\begin{aligned}
\binom{k}{i_1, i_2, \ldots, i_j} \left(\sum_{i=j}^{k} (-1)^{k-i} \binom{k-j}{k-i}\right) &= \binom{k}{i_1, i_2, \ldots, i_j} \left(\sum_{\ell=0}^{k-j} (-1)^{k-j-\ell} \binom{k-j}{k-j-\ell}\right) \\
&= \binom{k}{i_1, i_2, \ldots, i_j} (1-1)^{k-j} \\
&= 0.
\end{aligned}$$

Moreover, it is clear that $c_k = \binom{k}{1,1,\ldots,1} = k! \pmod p$ and $c_k \neq 0$ for the choice of $k$. □

### 7.3 Characterization of Low Degree Polynomials over $\mathbb{F}_p$

In this section we present an exact characterization for the family $\mathcal{P}_t$ over prime fields. Specifically we prove the following:

**Theorem 7.1.** *Let $t = (p-1)\cdot k + r$. (Note $0 \leqslant r \leqslant p-2$.) Let $i = p-2-r$. Then a function $f$ belongs to $\mathcal{P}_t$, if and only if for every $y_1, \cdots, y_{k+1}, b \in \mathbb{F}_p^n$, we have*

$$T_f^i(y_1, \cdots, y_{k+1}, b) = 0 \tag{7.6}$$

As mentioned previously, a characterization for the family $\mathcal{P}_t$ is equivalent to a characterization for $\mathrm{RM}_p(t,n)$. It turns out that it is easier to characterize $\mathcal{P}_t$ when viewed as $\mathrm{RM}_p(t,n)$. Therefore our goal is to determine whether a given word belongs to the RM code. Since we deal with a linear code, a simple strategy will then be to check whether the given word is orthogonal to all the codewords in the dual code. Though this yields a characterization, this is computationally inefficient. Note however that the dot product is linear in its input. Therefore checking orthogonality with a basis of the dual code suffices. To make it computationally efficient, we look for a basis with small weights. The above theorem essentially is a clever restatement of this idea.

We recall the following useful lemma which can be found in [69].

**Lemma 7.4.** $\mathrm{RM}_q(k,n)$ *is a linear code with block length $q^n$ and minimum distance $(R+1)q^Q$ where $R$ is the remainder and $Q$ the quotient resulting from dividing $(q-1)\cdot n - k$ by $(q-1)$. Then* $\mathrm{RM}_q(k,n)^\perp = \mathrm{RM}_q((q-1)\cdot n - k - 1, n)$.

Since the dual of a RM code is again a RM code (of appropriate order), we therefore need the generators of RM code (of arbitrary order). We first establish that flats and pseudoflats (of suitable dimension and exponent) indeed generate the Reed-Muller code. We then end the section with a proof of Theorem 7.1 and a few remarks.

We begin with few simple observations about flats. Note that an $l$-flat $L$ is the intersection of $(n-l)$ hyperplanes in general position. Equivalently, it consists of all points $v$ that satisfy $(n-l)$ linear equations over $\mathbb{F}_p$ (i.e., one equation for each hyperplane): $\forall i \in [n-l] \quad \sum_{j=1}^n c_{ij}x_j = b_i$ where $c_{ij}, b_i$ defines the $i^{th}$ hyperplane (i.e., $v$ satisfies $\sum_{j=1}^n c_{ij}v_j = b_i$). General position means that the matrix $\{c_{ij}\}$ has rank $(n-l)$. Note that then the characteristic function (and by abuse of notation the incidence vector) of $L$ can be written as

$$\prod_{i=1}^{n-l}(1-(\sum_{j=1}^{n} c_{ij}x_j - b_i)^{p-1}) = \begin{cases} 1 & \text{if } (v_1, \cdots, v_l) \in L \\ 0 & \text{otherwise} \end{cases} \tag{7.7}$$

We now record a lemma here that will be used later in this section.

**Lemma 7.5.** *For $k \geqslant l$, the incidence vector of any $k$-flat is a linear sum of the incidence vectors of $l$-flats.*

*Proof.* Let $k = l+r$ and let $W$ be an $k$-flat. We want to show that it is generated by a linear combination of $l$ flats.

Let $W$ be generated by $y_1, \cdots, y_{l-1}, w_1, \cdots, w_{r+1}$ at $b$. For each non-zero vector $c_i = \langle c_{i1}, \ldots, c_{i(r+1)} \rangle$ in $\mathbb{F}_p^{r+1}$ define:

$$v_i = \sum_{j=1}^{r+1} c_{ij} w_j.$$

Clearly there are $(p^{r+1} - 1)$ such $v_i$. Now for each $i \in [p^{r+1} - 1]$, define an $l$-flat $L_i$ generated by $y_1, \cdots, y_{l-1}, v_i$ at $b$. Denote the incidence vector of a flat $V$ by $1_V$, then we claim that

$$1_W = (p-1) \sum_{i=1}^{p^{r+1}-1} 1_{L_i}. \tag{7.8}$$

Since the vectors $y_1, \ldots, y_{l-1}, w_1, \ldots, w_{r+1}$ are all linearly independent, we can divide the proof in three sub cases:

- $v \in W$ is of the form $b + \sum_{i=1}^{l-1} e_i y_i$, for some $e_1, \ldots, e_{l-1} \in \mathbb{F}_p$: Then each flat $L_i$ contributes 1 to the right hand side of (7.8), and therefore, the right hand side is $(p-1)(p^{r+1} - 1) = 1$ in $\mathbb{F}_p$.

- $v \in W$ is of the form $b + \sum_{i=1}^{r+1} d_i w_i$ for some $d_1, \ldots, d_{r+1} \in \mathbb{F}_p$: Then the flats $L_j$ that contribute have $V_j = a \cdot \sum_{i=1}^{r+1} d_i w_i$, for $a = 1, \ldots, p-1$. Therefore, the right hand side of (7.8) is $(p-1)^2 = 1$ in $\mathbb{F}_p$.

- $v \in W$ is of the form $b + \sum_{i=1}^{l-1} e_i y_i + \sum_{i=1}^{r+1} d_i w_i$: Then the flats $L_j$ that contribute have $V_j = a \cdot \sum_{i=1}^{r+1} d_i w_i$, for $a = 1, \ldots, p-1$. Therefore, the right hand side of (7.8) is $(p-1)^2 = 1$ in $\mathbb{F}_p$.

□

As mentioned previously, we give an explicit basis for $\mathrm{RM}_p(r, n)$. For the special case of $p = 3$, our basis coincides with the min-weight basis given in [29].[2] However, in general, our basis differs from the min-weight basis provided in [29].

The following Proposition shows that the incidence vectors of flats form a basis for the Reed-Muller code of orders that are multiples of $(p-1)$.

**Proposition 7.6.** $\mathrm{RM}_p((p-1)(n-l), n)$ *is generated by the incidence vectors of the* $l$*-flats.*

[2]The equations of the hyperplanes are slightly different in our case; nonetheless, both of them define the same basis generated by the min-weight codewords.

*Proof.* We first show that the incidence vectors of the $l$-flats are in $\mathrm{RM}_p((p-1)(n-l),n)$. Recall that $L$ is the intersection of $(n-l)$ independent hyperplanes. Therefore using (7.7), $L$ can be represented by a polynomial of degree at most $(n-l)(p-1)$ in $x_1,\cdots,x_n$. Therefore the incidence vectors of $l$-flats are in $\mathrm{RM}_p((p-1)(n-l),n)$.

We prove that $\mathrm{RM}_p((p-1)(n-l),n)$ is generated by $l$-flats by induction on $n-l$. When $n-l=0$, the code consists of constants, which is clearly generated by $n$-flats i.e., the whole space.

To prove for an arbitrary $(n-l)>0$, we show that any monomial of total degree $d \leqslant (p-1)(n-l)$ can be written as a linear sum of the incidence vectors of $l$-flats. Let the monomial be $x_1^{e_1}\cdots x_s^{e_s}$. Rewrite the monomials as $\underbrace{x_1\cdots x_1}_{e_1 \text{ times}}\cdots\underbrace{x_s\cdots x_t}_{e_s \text{ times}}$. Group into products of $(p-1)$ (not necessarily distinct) variables as much as possible. Rewrite each group using (7.5) with $k=(p-1)$. For any incomplete group of size $d'$, use the same equation by setting the last $(p-1-d')$ variables to the constant 1. After expansion, the monomial can be seen to be a sum of products of at most $(n-l)$ linear terms raised to the power of $p-1$. We can add to it a polynomial of degree at most $(p-1)(n-l-1)$ so as to represent the resulting polynomial as a sum of polynomials, each polynomial as in (7.7). Each such non-zero polynomial is generated by a $t$ flat, $t \geqslant l$. By induction, the polynomial we added is generated by $(l+1)$ flats. Thus, by Lemma 7.5 our given monomial is generated by $l$-flats. □

This leads to the following observation:

**Observation 7.4.** *Consider an $l$-flat generated by $y_1,\cdots,y_l$ at $b$. Denote the incidence vector of this flat by $I$. Then the right hand side of (7.2) may be identified as $I\cdot f$, where $I$ and $f$ denote the vector corresponding to respective codewords and $\cdot$ is the dot (scalar) product.*

To generate a Reed-Muller code of any arbitrary order, we need pseudoflats. Note that the points in a $k$-pseudoflat may alternatively be viewed as the space given by the union of intersections of $(n-k-1)$ hyperplanes, where the union is parameterized by another hyperplane that does not take one particular value. Concretely, it is the set of points $v$ which satisfy the following constraints over $\mathbb{F}_p$:

$$\forall i \in [n-k-1]\ \sum_{j=1}^{n} c_{ij}x_j = b_i;\ \text{and} \quad \sum_{j=1}^{n} c_{n-k,j}x_j \neq b_{n-k}.$$

Thus the values taken by the points of a $k$-pseudoflat in its corresponding $(k,r)$-pseudoflat vector is given by the polynomial

$$\prod_{i=1}^{n-k-1}\left(1-\left(\sum_{j=1}^{n} c_{ij}x_j - b_i\right)^{p-1}\right)\cdot\left(\sum_{j=1}^{n} c_{n-k,j}x_j - b_{n-k}\right)^{r} \tag{7.9}$$

**Remark 7.1.** *Note the difference between (7.9) and the basis polynomial in [29] that (along with the action of the affine general linear group) yields the min-weight codewords:*

$$h(x_1, \cdots, x_m) = \prod_{i=1}^{k-1}(1 - (x_i - w_i)^{p-1}) \prod_{j=1}^{r}(x_k - u_j),$$

*where* $w_1, \cdots, w_{k-1}, u_1, \cdots, u_r \in \mathbb{F}_p$.

The next lemma shows that the code generated by the incidence vectors of $l$-flats is a subcode of the code generated by the $(l, r)$-pseudoflats vectors.

**Claim 7.7.** *The* $(l, r)$*-pseudoflats vectors, where* $l \geqslant 1$ *and* $r \in [p-2]$*, generate a code containing the incidence vectors of* $l$*-flats.*

*Proof.* Let $W$ be the incidence vector of an $l$-flat generated by $y_1, \cdots, y_l$ at $b$. Since pseudoflat vector corresponding to an $l$-pseudoflat (as well as a flat) assigns the same value to all points in the same $(l-1)$-flat, we can describe $W$ (as well as any $(l, \cdot)$-pseudoflat vector) by giving its values on each of its $p$ $l-1$-flats. In particular, $W = \langle 1, \ldots, 1 \rangle$. Let $L_j$ be a pseudoflat generated by $y_1, \cdots, y_l$ exponentiated along $y_1$ at $b + j \cdot y_1$, for each $j \in \mathbb{F}_p$, and let $V_j$ be the corresponding $(l, r)$- pseudoflat vector. By Definition 7.3, $V_j$ assigns a value $i^r$ to the $(l-1)$-flat generated by $y_2, \cdots, y_l$ at $b + (j+i)y$. Rewriting them in terms of the values on its $l-1$-flats yields that $V_j = \langle (p-j)^r, (p-j+1)^r, \cdots, (p-j+i)^r, \cdots, (p-j-1)^r \rangle \in \mathbb{F}_p^p$. Let $\lambda_j$ denote $p$ variables for $j = 0, 1, \cdots, p-1$, each taking values in $\mathbb{F}_p$. Then a solution to the following system of equations

$$1 = \sum_{j \in \mathbb{F}_p} \lambda_j (i-j)^r \quad \text{for every } 0 \leqslant l \leqslant p-1$$

implies that $W = \sum_{j=0}^{p-1} \lambda_j V_j$, which suffices to establish the claim. Consider the identity

$$1 = (-1) \sum_{j \in \mathbb{F}_p} (j+i)^r j^{p-1-r}$$

which may be verified by expanding and applying Lemma 7.1. Setting $\lambda_j$ to $(-1)(-j)^{p-1-r}$ establishes the claim. □

The next Proposition complements Proposition 7.6. Together they say that by choosing pseudoflats appropriately, Reed-Muller codes of any given order can be generated. This gives an equivalent representation of Reed-Muller codes. An exact characterization then follows from this alternate representation.

**Proposition 7.8.** *For every* $r \in [p-2]$*, the linear code generated by* $(l, r)$*-pseudoflat vectors is equivalent to* $\mathrm{RM}_p((p-1)(n-l)+r, n)$.

*Proof.* For the forward direction, consider an $l$-pseudoflat $L$. Its corresponding $(l,r)$-pseudoflat vector is given by an equation similar to (7.9). Thus the codeword corresponding to the evaluation vector of this flat can be represented by a polynomial of degree at most $(p-1)(n-l)+r$. This completes the forward direction.

Since monomials of degree at most $(p-1)(n-l)$ are generated by the incidence vectors of $l$-flats, Claim 7.7 will establish the proposition for such monomials. Thus, to prove the other direction of the proposition, we restrict our attention to monomials of degree at least $(p-1)(n-l)+1$ and show that these monomials are generated by $(l,r)$-pseudoflats vectors. Now consider any such monomial. Let the degree of the monomial be $(p-1)(n-l)+r'$ ($1 \leqslant r' \leqslant r$). Rewrite it as in Proposition 7.6. Since the degree of the monomial is $(p-1)(n-l)+r'$, we will be left with an incomplete group of degree $r'$. We make any incomplete group complete by adding 1's (as necessary) to the product. We then use Lemma 7.3 to rewrite each (complete) group as a linear sum of $r^{th}$ powered terms. After expansion, the monomial can be seen to be a sum of product of at most $(n-l)$ degree $(p-1)^{th}$ powered linear terms and a $r^{th}$ powered linear terms. Each such polynomial is generated either by an $(l,r)$-pseudoflat vector or an $l$-flat. Claim 7.7 completes the proof. □

The following is analogous to Observation 7.4.

**Observation 7.5.** *Consider an $l$-pseudoflat, generated by $y_1, \cdots, y_l$ at $b$ exponentiated along $y_1$. Let $E$ be its corresponding $(l,r)$-pseudoflat vector. Then the right hand side of (7.3) may be interpreted as $E \cdot f$.*

Now we prove the exact characterization.

**Proof of Theorem 7.1**: The proof directly follows from Lemma 7.4, Proposition 7.6, Proposition 7.8 and Observation 7.4 and Observation 7.5. Indeed by Observation 7.4, Observation 7.5 and (7.6) are essentially tests to determine whether the dot product of the function with every vector in the dual space of $\mathrm{RM}_p(t,n)$ evaluates to zero. □

**Remark 7.2.** *One can obtain an alternate characterization from Remark 7.1 which we state here without proof.*

*Let $t=(p-1)\cdot k+R$ (note $0<R\leqslant(p-2)$). Let $r=p-R-2$. Let $W \subseteq \mathbb{F}_p$ with $|W|=r$. Define the polynomial $g(x) \stackrel{def}{=} \prod_{\alpha \in W}(x-\alpha)$ if $W$ is non-empty; and $g(x)=1$ otherwise. Then a function belong to $\mathcal{P}_t$ if and only if for every $y_1, \cdots, y_{k+1}, b \in \mathbb{F}_p^n$, we have*

$$\sum_{c_1 \in \mathbb{F}_p \setminus W} g(c_1) \sum_{(c_2,\cdots,c_{k+1}) \in \mathbb{F}_p^k} f(b+\sum_{i=1}^{k+1} c_i \cdot y_i) = 0.$$

*Moreover, this characterization can also be extended to certain degrees for more general fields, i.e., $\mathbb{F}_{p^s}$ (see the next remark).*

**Remark 7.3.** *The exact characterization of low degree polynomials as claimed in [42] may be proved using duality. Note that their proof works as long as the dual code has a min-weight basis (see [29]). Suppose that the polynomial has degree $d \leqslant q-q/p-1$, then*

*the dual of* $\mathrm{RM}_q(d,n)$ *is* $\mathrm{RM}_q((q-1)n-d-1,n)$ *and therefore has a min-weight basis. Note that then the dual code has min-weight* $(d+1)$*. Therefore, assuming the minimum weight codewords constitute a basis (that is, the span of all codewords with the minimum Hamming weight is the same as the code), any* $d+1$ *evaluations of the original polynomial on a line are dependent and vice-versa.*

## 7.4 A Tester for Low Degree Polynomials over $\mathbb{F}_p^n$

In this section we present and analyze a one-sided tester for $\mathcal{P}_t$. The analysis of the algorithm roughly follows the proof structure given in [93, 1]. We emphasize that the generalization from [1] to our case is not straightforward. As in [93, 1] we define a self-corrected version of the (possibly corrupted) function being tested. The straightforward adoption of the analysis given in [93] gives reasonable bounds. However, a better bound is achieved by following the techniques developed in [1]. In there, they show that the self-corrector function can be interpolated with overwhelming probability. However their approach appears to use special properties of $\mathbb{F}_2$ and it is not clear how to generalize their technique for arbitrary prime fields. We give a clean formulation which relies on the flats being represented through polynomials as described earlier. In particular, Claims 7.14, 7.15 and their generalizations appear to require our new polynomial based view.

### 7.4.1 Tester in $\mathbb{F}_p$

In this subsection we describe the algorithm when underlying field is $\mathbb{F}_p$.
**Algorithm Test-$\mathcal{P}_t$ in $\mathbb{F}_p$**

0. Let $t=(p-1)\cdot k+R$, $0\leqslant R<p-1$. Denote $r=p-2-R$.
1. Uniformly and independently at random select $y_1,\cdots,y_{k+1},b\in\mathbb{F}_p^n$.
2. If $T_f^r(y_1,\cdots,y_{k+1},b)\neq 0$, then *reject*, else *accept*.

**Theorem 7.2.** *The algorithm* **Test-$\mathcal{P}_t$ in $\mathbb{F}_p$** *is a one-sided tester for $\mathcal{P}_t$ with a success probability at least min$(c(p^{k+1}\varepsilon), \frac{1}{2(k+7)p^{k+2}})$ for some constant $c>0$.*

**Corollary 7.3.** *Repeating the algorithm* **Test-$\mathcal{P}_t$ in $\mathbb{F}_p$** *for $\Theta(\frac{1}{p^{k+1}\varepsilon}+kp^k)$ times, the probability of error can be reduced to less than $1/2$.*

We will provide a general proof framework. However, for the ease of exposition we prove the main technical lemmas for the case of $\mathbb{F}_3$. The proof idea in the general case is similar and the details are omitted. Therefore we will essentially prove the following.

**Theorem 7.4.** *The algorithm* **Test-$\mathcal{P}_t$ in $\mathbb{F}_3$** *is a one-sided tester for $\mathcal{P}_t$ with success probability at least min$(c(3^{k+1}\varepsilon), \frac{1}{2(t+7)3^{t/2+1}})$ for some constant $c>0$.*

### 7.4.2 *Analysis of Algorithm* Test-$\mathcal{P}_t$

In this subsection we analyze the algorithm described in Section 7.4.1. From Claim 7.1 it is clear that if $f \in \mathcal{P}_t$, then the tester accepts. Thus, the bulk of the proof is to show that if $f$ is $\varepsilon$-*far* from $\mathcal{P}_t$, then the tester rejects with significant probability. Our proof structure follows that of the analysis of the test in [1]. In particular, let $f$ be the function to be tested for membership in $\mathcal{P}_t$. Assume we perform Test $T_f^i$ for an appropriate $i$ as required by the algorithm described in Section 7.4.1. For such an $i$, we define $g_i : \mathbb{F}_p^n \to \mathbb{F}_p$ as follows: For $y \in F_p^n, \alpha \in \mathbb{F}_p$, denote $p_{y,\alpha} = \Pr_{y_1,\cdots,y_{k+1}}[f(y) - T_f^i(y - y_1, y_2, \cdots, y_{k+1}, y_1) = \alpha]$. Define $g_i(y) = \alpha$ such that $\forall \beta \neq \alpha \in \mathbb{F}_p, p_{y,\alpha} \geqslant p_{y,\beta}$ with ties broken arbitrarily. With this meaning of plurality, for all $i \in [p-2] \cup \{0\}$, $g_i$ can be written as:

$$g_i(y) \;=\; \text{plurality}_{y_1,\cdots,y_{k+1}} \left[ f(y) - T_f^i(y - y_1, y_2, \cdots, y_{k+1}, y_1) \right]. \tag{7.10}$$

Further we define

$$\eta_i \stackrel{def}{=} Pr_{y_1,\cdots,y_{k+1},b}[T_f^i(y_1, \cdots, y_{k+1}, b) \neq 0] \tag{7.11}$$

The next lemma follows from a Markov-type argument.

**Lemma 7.9.** *For a fixed $f : \mathbb{F}_p^n \to \mathbb{F}_p$, let $g_i, \eta_i$ be defined as above. Then, $\delta(f, g_i) \leqslant 2\eta_i$.*

*Proof.* If for some $y \in \mathbb{F}_p^n$, $\Pr_{y_1,\cdots,y_{k+1}}[f(y) = f(y) - T_f^i(y - y_1, y_2, \cdots, y_{k+1}, y_1)] > 1/2$, then $g(y) = f(y)$. Thus, we only need to worry about the set of elements $y$ such that $\Pr_{y_1,\cdots,y_{k+1}}[f(y) = f(y) - T_f^i(y - y_1, y_2, \cdots, y_{k+1}, y_1)] \leqslant 1/2$. If the fraction of such elements is more than $2\eta_i$ then that contradicts the condition that

$$\begin{aligned}
\eta_i &= \Pr_{y_1,\cdots,y_{k+1},b}[T_f^i(y_1, \cdots, y_{k+1}, b) \neq 0] \\
&= \Pr_{y_1,y_2,\cdots,y_{k+1},b}[T_f^i(y_1 - b, y_2, \cdots, y_{k+1}, b) \neq 0] \\
&= \Pr_{y,y_1,\cdots,y_{k+1}}[f(y) \neq f(y) - T_f^i(y - y_1, y_2, \cdots, y_{k+1}, y_1)].
\end{aligned}$$

Therefore, we obtain $\delta(f, g_i) \leqslant 2\eta_i$. □

Note that $\Pr_{y_1,\cdots,y_{k+1}}[g_i(y) = f(y) - T_f^i(y - y_1, y_2, \cdots, y_{k+1}, y_1)] \geqslant \frac{1}{p}$. We now show that this probability is actually much higher. The next lemma gives a weak bound in that direction following the analysis in [93]. For the sake of completeness, we present the proof.

**Lemma 7.10.** $\forall y \in \mathbb{F}_p^n$, $\Pr_{y_1,\cdots,y_{k+1} \in \mathbb{F}_p^n}[g_i(y) = f(y) - T_f^i(y - y_1, y_2, \cdots, y_{k+1}, y_1)] \geqslant 1 - 2p^{k+1}\eta_i$.

*Proof.* We will use $I, J, I', J'$ to denote $(k+1)$ dimensional vectors over $\mathbb{F}_p$. Now note that

$$\begin{aligned}
-g_i(y) &= \text{Plurality}_{y_1,\cdots,y_{k+1}\in\mathbb{F}_p^n}[\sum_{I\in\mathbb{F}_p^{k+1};I\neq\langle 1,0,\cdots,0\rangle} I_1^i f(I_1(y-y_1)+\sum_{t=2}^{k+1} I_t y_t + y_1)] \\
&= \text{Plurality}_{y-y_1,y_2,\cdots,y_{k+1}\in\mathbb{F}_p^n}[\sum_{I\in\mathbb{F}_p^{k+1};I\neq\langle 0,\cdots,0\rangle} (I_1+1)^i f(I_1(y-y_1) \\
&\quad + \sum_{t=2}^{k+1} I_t y_t + y)] \\
&= \text{Plurality}_{y_1,\cdots,y_{k+1}\in\mathbb{F}_p^n}[\sum_{I\in\mathbb{F}_p^{k+1};I\neq\langle 0,\cdots,0\rangle} (I_1+1)^i f(\sum_{t=1}^{k+1} I_t y_t + y)] \qquad (7.12)
\end{aligned}$$

Let $Y = \langle y_1, \cdots, y_{k+1}\rangle$ and $Y' = \langle y'_1, \cdots, y'_{k+1}\rangle$. Also we will denote $\langle 0, \cdots, 0\rangle$ by $\mathbf{0}$. Now note that

$$\begin{aligned}
1-\eta_i \leqslant\ & \Pr_{y_1,\cdots,y_{k+1},b}[T_f^i(y_1,\cdots,y_{k+1},b)=0] \\
=\ & \Pr_{y_1,\cdots,y_{k+1},b}[\sum_{I\in\mathbb{F}_p^{k+1}} I_1^i f(b+I\cdot Y)=0] \\
=\ & \Pr_{y_1,\cdots,y_{k+1},b}[f(b+y_1)+\sum_{I\in\mathbb{F}_p^{k+1};I\neq\langle 1,0,\cdots,0\rangle} I_1^i f(b+I\cdot Y)=0] \\
=\ & \Pr_{y_1,\cdots,y_{k+1},y}[f(y)+\sum_{I\in\mathbb{F}_p^{k+1};I\neq\langle 1,0,\cdots,0\rangle} I_1^i f(y-y_1+I\cdot Y)=0] \\
=\ & \Pr_{y_1,\cdots,y_{k+1},y}[f(y)+\sum_{I\in\mathbb{F}_p^{k+1};I\neq\langle 0,\cdots,0\rangle} (I_1+1)^i f(y+I\cdot Y)=0]
\end{aligned}$$

Therefore for any given $I\neq\mathbf{0}$ we have the following:

$$\Pr_{Y,Y'}[f(y+I\cdot Y)=\sum_{J\in\mathbb{F}_p^{k+1};J\neq\mathbf{0}} -(J_1+1)^i f(y+I\cdot Y+J\cdot Y')] \geqslant 1-\eta_i$$

and for any given $J\neq\mathbf{0}$,

$$\Pr_{Y,Y'}[f(y+J\cdot Y')=\sum_{I\in\mathbb{F}_p^{k+1};I\neq\mathbf{0}} -(I_1+1)^i f(y+I\cdot Y+J\cdot Y')] \geqslant 1-\eta_i.$$

Combining the above two and using the union bound we get,

$$\begin{aligned}
&\Pr_{Y,Y'}\left[\sum_{I\in\mathbb{F}_p^{k+1};I\neq\mathbf{0}}(I_1+1)^i f(y+I\cdot Y)\right]\\
&\quad=\sum_{I\in\mathbb{F}_p^{k+1};I\neq\mathbf{0}}\ \sum_{J\in\mathbb{F}_p^{k+1};J\neq\mathbf{0}} -(I_1+1)^i(J_1+1)^i f(y+I\cdot Y+J\cdot Y')\\
&\quad=\sum_{J\in\mathbb{F}_p^{k+1};J\neq\mathbf{0}}(J_1+1)^i f(y+J\cdot Y')]\\
&\quad\geqslant 1-2(p^{k+1}-1)\eta\geqslant 1-2p^{k+1}\eta_i
\end{aligned}\tag{7.13}$$

The lemma now follows from the observation that the probability that the same object is drawn from a set in two independent trials lower bounds the probability of drawing the most likely object in one trial: Suppose the objects are ordered so that $p_i$ is the probability of drawing object $i$, and $p_1 \geqslant p_2 \geqslant \cdots$. Then the probability of drawing the same object twice is $\sum_i p_i^2 \leqslant \sum_i p_1 p_i \leqslant p_1$. □

However, when the degree being tested is larger than the field size, we can improve the above lemma considerably. The following lemma strengthens Lemma 7.10 when $t \geqslant p-1$ or equivalently $k \geqslant 1$. We now focus on the $\mathbb{F}_3$ case. The proof appears in Section 7.4.3.

**Lemma 7.11.** $\forall y \in \mathbb{F}_3^n$, $\Pr_{y_1,\cdots,y_{k+1}\in\mathbb{F}_3^n}[g_i(y) = f(y) - T_f^i(y-y_1, y_2, \cdots, y_{k+1}, y_1)] \geqslant 1-(4k+14)\eta_i$.

Lemma 7.11 will be instrumental in proving the next lemma, which shows that sufficiently small $\eta_i$ implies $g_i$ is the self-corrected version of the function $f$ (the proof appears in Section 7.4.4).

**Lemma 7.12.** *Over $\mathbb{F}_3$, if $\eta_i < \frac{1}{(4k+14)3^{k+1}}$, then the function $g_i$ belongs to $\mathcal{P}_t$ (assuming $k \geqslant 1$).*

By combining Lemma 7.9 and Lemma 7.12 we obtain that if $f$ is $\Omega(1/(k3^k))$-far from $\mathcal{P}_t$ then $\eta_i$ is at least $\Omega(1/(k3^k))$. We next consider the case in which $\eta_i$ is small. By Lemma 7.9, in this case, the distance $\delta = \delta(f,g)$ is small. The next lemma shows that in this case the test rejects $f$ with probability that is close to $3^{k+1}\delta$. This follows from the fact that in this case, the probability over the selection of $y_1, \cdots, y_{k+1}, b$, that among the $3^{k+1}$ points $\sum_i c_i y_i + b$ (where $c_1, \ldots, c_{k+1} \in \mathbb{F}_3$), the functions $f$ and $g$ differ in precisely one point, is close to $3^{k+1} \cdot \delta$. Observe that if they do, then the test rejects.

**Lemma 7.13.** *Suppose $0 \leqslant \eta_i \leqslant \frac{1}{(4k+14)3^{k+1}}$. Let $\delta$ denote the relative distance between $f$ and $g$ and $\ell = 3^{k+1}$. Then, when $y_1, \cdots, y_{k+1}, b$ are chosen randomly, the probability that for exactly one point $v$ among the $\ell$ points $\sum_i c_i y_i + b$ (where $(c_1, \ldots, c_{k+1}) \in \mathbb{F}_3^{k+1}$), $f(v) \neq g(v)$ is at least $\left(\frac{1-\ell\delta}{1+\ell\delta}\right)\ell\delta$.*

Observe that $\eta_i$ is at least $\Omega(3^{k+1}\delta)$. The proof of Lemma 7.13 is deferred to Section 7.4.5.

**Proof of Theorem 7.4:** Clearly if $f$ belongs to $\mathcal{P}_t$, then by Claim 7.1 the tester accepts $f$ with probability 1.

Therefore let $\delta(f, \mathcal{P}_t) \geqslant \varepsilon$. Let $d = \delta(f, g_r)$, where $r$ is as in algorithm **Test-$\mathcal{P}_t$**. If $\eta < \frac{1}{(4k+14)3^{k+1}}$ then by Lemma 7.12 $g_r \in \mathcal{P}_t$ and, by Lemma 7.13, $\eta_i$ is at least $\Omega(3^{k+1} \cdot d)$, which by the definition of $\varepsilon$ is at least $\Omega(3^{k+1}\varepsilon)$. Hence $\eta_i \geqslant min\left(c(3^{k+1}\varepsilon), \frac{1}{(4k+14)3^{k+1}}\right)$, for some fixed constant $c > 0$. □

**Remark 7.4.** *Theorem 7.2 follows from a similar argument.*

### *7.4.3 Proof of Lemma 7.11*

Observe that the goal of Lemma 7.11 is to show that at any fixed point $y$, if $g_i$ is interpolated from a random hyperplane, then w.h.p. the interpolated value is the most popular vote. To ensure this we show that if $g_i$ is interpolated on two independently random hyperplanes, then the probability that these interpolated values are the same, that is the collision probability, is large. To estimate this collision probability, we show that the difference of the interpolation values can be rewritten as a sum of $T_f^i$ on small number of random hyperplanes. Thus if the test passes often (that is, $T_f^i$ evaluates to zero w.h.p.), then this sum (by a simple union bound) evaluates to zero often, which proves the high collision probability.

The improvement will arise because we will express differences involving $T_f^i(\cdots)$ as a telescoping series to essentially reduce the number of events in the union bound. To do this we will need the following claims. We note that a similar claim for $p = 2$ was proven by expanding the terms on both sides in [1]. However, the latter does not give much insight into the general case i.e., for $\mathbb{F}_p$. We provide proofs that have a much cleaner structure based on the underlying geometric structure, i.e., flats or pseudoflats.

**Claim 7.14.** *For every $l \in \{2, \cdots, k+1\}$, for every $y(= y_1), z, w, b, y_2, \cdots, y_{l-1}, y_{l+1}, \cdots, y_{k+1} \in \mathbb{F}_p^n$, let let*

$$S_f^l(y, z) \stackrel{def}{=} T_f^0(y, y_2, \cdots, y_{l-1}, z, y_{l+1}, \cdots, y_{k+1}, b).$$

*The the following holds:*

$$S_f^l(y, w) - S_f^l(y, z) = \sum_{e \in \mathbb{F}_p^*} \left[S_f^l(y + ew, z) - S_f^l(y + ez, w)\right].$$

*Proof.* Assume $y, z, w$ are independent. If not then both sides are equal to 0 and hence the equality is trivially satisfied. To see why this claim is true for the left hand side, recall the definition of $T_f^0(\cdot)$ and note that the sets of points in the flat generated by $y, y_2, \cdots, y_{l-1}, w$, $y_{l+1}, \cdots, y_{k+1}$ at $b$ and the flat generated by $y, y_2, \cdots, y_{l-1}, z, y_{l+1}, \cdots, y_{k+1}$ at $b$ are the same. A similar argument works for the expression on the right hand side of the equality.

We claim that it is enough to prove the result for $k = 1$ and $b = \mathbf{0}$. A linear transform (or renaming the co-ordinate system appropriately) reduces the case of $k = 1$ and $b \neq \mathbf{0}$ to the case of $k = 1$ and $b = \mathbf{0}$. We now show how to reduce the case of $k > 1$ to the $k = 1$ case. Fix some values $c_2, \cdots, c_{l-1}, c_{l+1}, \cdots, c_{k+1}$ and note that one can write $c_1 y + c_2 y_2 + \cdots c_{l-1} y_{l-1} + c_l w + c_{l+1} y_{l+1} + c_{k+1} y_{k+1} + b$ as $c_1 y + c_l w + b'$, where $b' = \sum_{j \in \{2, \cdots, l-1, l+1, \cdots, k+1\}} c_j y_j + b$. Thus,

$$S_f^l(y, w) = \sum_{(c_2, \cdots, c_{l-1}, c_{l+1}, \cdots, c_{k+1}) \in \mathbb{F}_p}^{k-1} \sum_{(c_1, c_l) \in \mathbb{F}_p^2} f(c_1 y + c_l w + b').$$

One can rewrite the other $S_f^l(\cdot)$ terms similarly. Note that for a fixed vector $(c_2, \cdots, c_{l-1}, c_{l+1}, \cdots, c_{k+1})$, the value of $b'$ is the same. Finally note that the equality in the general case is satisfied if $p^{k-1}$ similar equalities hold in the $k = 1$ case.

Now consider the space $\mathcal{H}$ generated by $y, z$ and $w$ at $\mathbf{0}$. Note that $S_f^l(y, w)$ (with $b = \mathbf{0}$) is just $f \cdot 1_L$, where $1_L$ is the incidence vector of the flat given by the equation $z = 0$. Therefore $1_L$ is equivalent to the polynomial $(1 - z^{p-1})$ over $\mathbb{F}_p$. Similarly $S_f^l(y, z) = f \cdot 1_{L'}$ where $L'$ is given by the polynomial $(1 - w^{p-1})$ over $\mathbb{F}_p$. We use the following polynomial identity (in $\mathbb{F}_p$)

$$w^{p-1} - z^{p-1} = \sum_{e \in \mathbb{F}_p^*} \left[ [1 - (ew + y)^{p-1}] - [1 - (ez + y)^{p-1}] \right] \tag{7.14}$$

Now observe that the polynomial $(1 - (ew + y)^{p-1})$ is the incidence vector of the flat generated by $y - e^{-1} w$ and $z$. Similarly, the polynomial $(1 - (ez + y)^{p-1})$ is the incidence vector of the flat generated by $y - e^{-1} z$ and $w$. Therefore, interpreting the above equation in terms of incidence vectors of flats, Observation 7.4 completes the proof assuming (7.14) is true.

We complete the proof by proving (7.14). Consider the sum: $\sum_{e \in \mathbb{F}_p^*} (ew + y)^{p-1}$. Expanding the terms and rearranging the sums we get $\sum_{j=0}^{p-1} \binom{p-1}{j} w^{p-1-j} y^j \sum_{e \in \mathbb{F}_p^*} e^{p-1-j}$. By Lemma 7.1 the sum evaluates to $(-w^{p-1} - y^{p-1})$. Similarly, $\sum_{e \in \mathbb{F}_p^*} (ez + y)^{p-1} = (-z^{p-1} - y^{p-1})$ which proves (7.14). □

We will also need the following claim.

**Claim 7.15.** *For every $i \in \{1, \cdots, p-2\}$, for every $l \in \{2, \cdots, k+1\}$ and for every $y (= y_1), z, w, b, y_2, \cdots, y_{l-1}, y_{l+1}, \cdots, y_{k+1} \in \mathbb{F}_p^n$, denote*

$$S_f^{i,l}(y, w) \stackrel{def}{=} T_f^i(y, y_2, \cdots, y_{l-1}, w, y_{l+1}, \cdots, y_{k+1}, b).$$

*Then there exists $c_i$ such that*

$$S_f^{i,l}(y, w) - S_f^{i,l}(y, z) = c_i \sum_{e \in \mathbb{F}_p^*} \left[ S_f^{i,l}(y + ew, z) - S_f^{i,l}(y + ez, w) \right].$$

*Proof.* As in the proof of Claim 7.14, we only need to prove the claim for $k = 1$ and $b = \mathbf{0}$. Observe that $S_f^{i,l}(y,z) = f \cdot E_{L_i}$, where $E_{L_i}$ denotes the $(2,i)$-pseudoflat vector of the pseudoflat $L$ generated by $y, z$ at $b$ exponentiated along $y$. Note that the polynomial defining $E_{L_i}$ is just $y^i(w^{p-1}-1)$. Similarly we can identify the other terms with polynomials over $\mathbb{F}_p$. To complete the proof, we need to prove the following identity (which is similar to the one in (7.14)):

$$y^i(w^{p-1} - z^{p-1}) = c_i \sum_{e \in \mathbb{F}_p^*} \left[(y+ew)^i[1-(y-ew)^{p-1}] - (y+ez)^i[1-(y-ez)^{p-1}]\right]. \tag{7.15}$$

where $c_i = 2^i$. Before we prove the identity, note that $(-1)^j\binom{p-1}{j} = 1$ in $\mathbb{F}_p$. This is because for $1 \leqslant m \leqslant j$, $m = (-1)(p-m)$. Therefore $j! = (-1)^j \frac{(p-1)!}{(p-j-1)!}$ holds in $\mathbb{F}_p$. Substitution yields the desired result. Also note that $\sum_{e\in\mathbb{F}_p^*}(y+ew)^i = -y^i$ (expand and apply Lemma 7.1). Now consider the sum

$$\begin{aligned}
\sum_{e\in\mathbb{F}_p^*}(y+ew)^i(y-ew)^{p-1} &= \sum_{e\in\mathbb{F}_p^*}\sum_{\substack{0\leqslant j\leqslant i;\\ 0\leqslant m\leqslant p-1}} (-1)^m\binom{i}{j}\binom{p-1}{m}y^{p-1+i-j-m}w^{j+m}e^{j+m} \\
&= \sum_{\substack{0\leqslant j\leqslant i;\\ 0\leqslant m\leqslant p-1}} (-1)^m\binom{i}{j}\binom{p-1}{m}y^{p-1+i-j-m}w^{j+m}\sum_{e\in\mathbb{F}_p^*}e^{j+m} \\
&= -y^{p-1+i} + (-1)^p\sum_{j=0}^{i}\binom{i}{j}\underbrace{\binom{p-1}{p-1-j}(-1)^j}_{=1}y^iw^{p-1}] \\
&= (-1)[y^i + y^iw^{p-1}2^i] \qquad (7.16)
\end{aligned}$$

Similarly one has $\sum_{e\in\mathbb{F}_p^*}(y+ez)^i(y-ez)^{p-1} = (-1)[y^i + y^iz^{p-1}2^i]$. Substituting and simplifying one gets (7.15). □

Finally, we will also need the following claim.

**Claim 7.16.** *For every $i \in \{1,\cdots,p-2\}$, for every $l \in \{2,\cdots,l+1\}$ and for every $y(=y_l), z, w, b, y_2, \cdots, y_{l-1}, y_{l+1}, \cdots, y_{k+1} \in \mathbb{F}_p^n$, there exists $c_i \in \mathbb{F}_p^*$ such that*

$$\begin{aligned}
S_f^{i,l}(w,y) - S_f^{i,l}(z,y) = \sum_{e\in\mathbb{F}_p^*}\Big[&S_f^{i,l}(y+ew,y-ew) - S_f^{i,l}(w+ey,w-ey) + \\
&S_f^{i,l}(z+ey,z-ey) - S_f^{i,l}(y+ez,y-ez) \\
&+c_i\left[S_f^{i,l}(y+ew,z) - S_f^{i,l}(y+ez,w)\right]\Big]
\end{aligned}$$

*Proof.* As in the proof of Claim 7.15, the proof boils down to proving a polynomial identity over $\mathbb{F}_p$. In particular, we need to prove the following identity over $\mathbb{F}_p$:

$$w^i(1-z^{p-1}) - z^i(1-w^{p-1}) = (w^i-y^i)(1-z^{p-1}) - (z^i-y^i)(1-w^{p-1}) + y^i(w^{p-1}-z^{p-1}).$$

We also use that $\sum_{e\in\mathbb{F}_p^*}(w+ey)^i = -w^i$ and Claim 7.15 to expand the last term. Note that $c_i = 2^i$ as before. □

We need one more simple fact before we can prove Lemma 7.11. For a probability vector $v \in [0,1]^n$, $\left\|v\right\|_\infty = \text{Max}_{i\in[n]}\{v_i\} \geqslant \text{Max}_{i\in[n]}\{v_i\}\cdot(\sum_{i=1}^n v_i) = \sum_{i=1}^n v_i \cdot \text{Max}_{i\in[n]}\{v_i\} \geqslant \sum_{i=1}^n v_i^2 = \left\|v\right\|_2^2$.

**Proof of Lemma 7.11:** We first prove the lemma for $g_0(y)$. We fix $y \in \mathbb{F}_3^n$ and let $\gamma \stackrel{def}{=} \Pr_{y_1,\cdots,y_{k+1}\in\mathbb{F}_3^n}[g_0(y) = f(y) - T_f^0(y-y_1, y_2, \cdots, y_{k+1}, y_1)]$. Recall that we want to lower bound $\gamma$ by $1-(4k+14)\eta_0$. In that direction, we bound a slightly different but related probability. Define

$$\mu = \Pr_{y_1,\cdots,y_{k+1},z_1,\cdots,z_{k+1}\in\mathbb{F}_3^n}[T_f^0(y-y_1,y_2,\cdots,y_{k+1},y_1) = T_f^0(y-z_1,z_2,\cdots,z_{k+1},z_1)]$$

Denote $Y = \langle y_1,\cdots,y_{k+1}\rangle$ and similarly $Z$. Then by the definitions of $\mu$ and $\gamma$ we have, $\gamma \geqslant \mu$. Note that we have

$$\mu = \Pr_{y_1,\cdots,y_{k+1},z_1,\cdots,z_{k+1}\in\mathbb{F}_3^n}[T_f^0(y-y_1,y_2,\cdots,y_{k+1},y_1) - T_f^0(y-z_1,z_2,\cdots,z_{k+1},z_1) = 0].$$

We will now use a hybrid argument. Now, for any choice of $y_1,\cdots,y_{k+1}$ and $z_1,\cdots,z_{k+1}$ we have:

$$\begin{aligned}
&T_f^0(y-y_1,y_2,\cdots,y_{k+1},y_1) - T_f^0(y-z_1,z_2,\cdots,z_{k+1},z_1)\\
&\quad= T_f^0(y-y_1,y_2,\cdots,y_{k+1},y_1) - T_f^0(y-y_1,y_2,\cdots,y_k,z_{k+1},y_1)\\
&\qquad+ T_f^0(y-y_1,y_2,\cdots,y_k,z_{k+1},y_1) - T_f^0(y-y_1,y_2,\cdots,y_{k-1},z_k,z_{k+1},y_1)\\
&\qquad+ T_f^0(y-y_1,\cdots,y_{k-1},z_k,z_{k+1},y_1) - T_f^0(y-y_1,\cdots,y_{k-2},z_{k-1},z_k,z_{k+1},y_1)\\
&\qquad\qquad\vdots\\
&\qquad+ T_f^0(y-y_1,z_2,z_3,\cdots,z_{k+1},y_1) - T_f^0(y-z_1,z_2,\cdots,z_{k+1},y_1)\\
&\qquad+ T_f^0(y-z_1,z_2,z_3,\cdots,z_{k+1},y_1) - T_f^0(y-y_1,z_2,\cdots,z_{k+1},z_1)\\
&\qquad+ T_f^0(y-y_1,z_2,z_3,\cdots,z_{k+1},z_1) - T_f^0(y-z_1,z_2,\cdots,z_{k+1},z_1)
\end{aligned}$$

Consider any pair $T_f^0(y-y_1,y_2,\cdots,y_l,z_{l+1},\cdots,z_{k+1},y_1) - T_f^0(y-y_1,y_2,\cdots,y_{l-1},z_l,\cdots,z_{k+1},y_1)$ that appears in the first $k$ "rows" in the sum above. Note that $T_f^0(y-y_1,y_2,\cdots,y_l,z_{l+1},\cdots,z_{k+1},y_1)$ and $T_f^0(y-y_1,y_2,\cdots,y_{l-1},z_l,\cdots,z_{k+1},y_1)$ differ only in a single parameter. We apply Claim 7.14 and obtain:

$T_f^0(y-y_1,y_2,\cdots,y_l,z_{l+1},\cdots,z_{k+1},y_1) - T_f^0(y-y_1,y_2,\cdots,y_{l-1},z_l,\cdots,z_{k+1},y_1) = T_f^0(y-y_1+y_l,y_2,\cdots,y_{l-1},z_l,\cdots,z_{k+1},y_1) + T_f^0(y-y_1-y_l,y_2,\cdots,y_{l-1},z_l,\cdots,z_{k+1},y_1) - T_f^0(y-y_1+z_l,y_2,\cdots,y_l,z_{l+1},\cdots,z_{k+1},y_1) - T_f^0(y-y_l-z_l,y_2,\cdots,y_l,z_{l+1},\cdots,z_{k+1},y_1)$.

Recall that $y$ is fixed and $y_2,\cdots,y_{k+1},z_2,\cdots,z_{k+1} \in \mathbb{F}_3^n$ are chosen uniformly at random, so all the parameters on the right hand side of the equation are independent and

uniformly distributed. Similarly one can expand the pairs $T_f^0(y-y_1, z_2, z_3, \cdots, z_{k+1}, y_1) - T_f^0(y-z_1, z_2, \cdots, z_{k+1}, y_1)$ and $T_f^0(y-y_1, z_2, z_3, \cdots, z_{k+1}, z_1) - T_f^0(y-z_1, z_2, \cdots, z_{k+1}, z_1)$ into four $T_f^0$ with all parameters being independent and uniformly distributed[3]. Finally notice that the parameters in both $T_f^0(y-z_1, z_2, z_3, \cdots, z_{k+1}, y_1)$ and $T_f^0(y-z_1, z_2, \cdots, z_{k+1}, y_1)$ are independent and uniformly distributed. Further recall that by the definition of $\eta_0$, $\Pr_{r_1,\cdots,r_{k+1}}[T_f^0(r_1, \cdots, r_{k+1}) \neq 0] \leqslant \eta_0$ for independent and uniformly distributed $r_i$s. Thus, by the union bound, we have:

$$\Pr_{y_1,\cdots,y_{k+1},z_1,\cdots,z_{k+1}\in\mathbb{F}_3^n}[T_f^0(y_1, \cdots, y_{k+1}) - T_f^0(z_1, \cdots, z_{k+1}) \neq 0] \leqslant (4k+10)\eta_0. \quad (7.17)$$

Therefore $\gamma \geqslant \mu \geqslant 1-(4k+10)\eta_0$. A similar argument proves the lemma for $g_1(y)$. The only catch is that $T_{f_1}(.)$ is not symmetric– in particular in its first argument. Thus, we use another identity as given in Claim 7.16 to resolve the issue and get four extra terms than in the case of $g_0$, which results in the claimed bound of $(4k+14)\eta_i$. $\square$

**Remark 7.5.** *Analogously, in the case $\mathbb{F}_p$ we have: for every $y \in \mathbb{F}_p^n$, $\Pr_{y_1,y_2,\cdots,y_{k+1}\in\mathbb{F}_p^n}[g_i(y) = f(y) - T_f^i(y-y_1, y_2, \cdots, y_{k+1}, y_1) + f(y)] \geqslant 1-2((p-1)k+6(p-1)+1)\eta_i$.*
*The proof is similar to that of Lemma 7.11 where it can be shown $\mu_i \geqslant 1-2((p-1)k+6(p-1)+1)\eta_i$, for each $\mu_i$ defined for $g_i(y)$.*

### *7.4.4 Proof of Lemma 7.12*

From Theorem 7.1, it suffices to prove that if $\eta_i < \frac{1}{(4k+14)3^{k+1}}$ then $T_{g_i}^i(y_1, \cdots, y_{k+1}, b) = 0$ for every $y_1, \cdots, y_{k+1}, b \in \mathbb{F}_3^n$. Fix the choice of $y_1, \cdots, y_{k+1}, b$. Define $Y = \langle y_1, \cdots, y_{k+1}\rangle$. We will express $T_{g_i}^i(Y, b)$ as the sum of $T_f^i(\cdot)$ with random arguments. We uniformly select $(k+1)^2$ random variables $z_{i,j}$ over $\mathbb{F}_3^n$ for $1 \leqslant i \leqslant k+1$, and $1 \leqslant j \leqslant k+1$. Define $Z_i = \langle z_{i,1}, \cdots, z_{i,k+1}\rangle$. We also select uniformly $(k+1)$ random variables $r_i$ over $\mathbb{F}_3^n$ for $1 \leqslant i \leqslant k+1$. We use $z_{i,j}$ and $r_i$'s to set up the random arguments. Now by Lemma 7.11, for every $I \in \mathbb{F}_3^{k+1}$ (i.e. think of $I$ as an ordered $(k+1)$-tuple over $\{0,1,2\}$), with probability at least $1-(4k+14)\eta_i$ over the choice of $z_{i,j}$ and $r_i$,

$$g_i(I\cdot Y+b) = f(I\cdot Y+b) - T_f^i(I\cdot Y+b-I\cdot Z_1-r_1, I\cdot Z_2+r_2, \cdots, I\cdot Z_{k+1}+r_{k+1}, I\cdot Z_1+r_1), \quad (7.18)$$

where for vectors $X, Y \in \mathbb{F}_3^{k+1}$, $Y\cdot X = \sum_{i=1}^{k+1} Y_i X_i$, holds.

Let $E_1$ be the event that (7.18) holds for all $I \in \mathbb{F}_3^{k+1}$. By the union bound:

$$\Pr[E_1] \geqslant 1 - 3^{k+1}\cdot(4k+14)\eta_i. \quad (7.19)$$

Assume that $E_1$ holds. We now need the following claims. Let $J = \langle J_1, \cdots, J_{k+1}\rangle$ be a $(k+1)$ dimensional vector over $\mathbb{F}_3$, and denote $J' = \langle J_2, \cdots, J_{k+1}\rangle$.

---

[3]Since $T_f^0(\cdot)$ is symmetric in all but its last argument.

**Claim 7.17.** *If (7.18) holds for all $I \in \mathbb{F}_3^{k+1}$, then*

$$\begin{aligned} T^0_{g_0}(Y,b) = & \sum_{0\neq J'\in\mathbb{F}_3^k}\left[-T^0_f(y_1+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,y_{k+1}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},b+\sum_{t=2}^{k+1}J_tr_t)\right] \\ & +\sum_{J'\in\mathbb{F}_3^k}\Bigg[-T^0_f(2y_1-z_{1,1}+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,2y_{k+1}-z_{1,(k+1)}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)}, \\ & \qquad 2b-r_1+\sum_{t=2}^{k+1}J_tr_t) \\ & \qquad +T^0_f\left(z_{1,1}+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,z_{1,k+1}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},r_1+\sum_{t=2}^{k+1}J_tr_t\right)\Bigg] \end{aligned} \tag{7.20}$$

*Proof.*

$$\begin{aligned} T^0_{g_0}(Y,b) &= \sum_{I\in\mathbb{F}_3^{k+1}} g_0(I\cdot Y+b) \\ &= \sum_{I\in\mathbb{F}_3^{k+1}}\big[-T^0_f(I\cdot Y+b-I\cdot Z_1-r_1,I\cdot Z_2+r_2,\cdots,I\cdot Z_{k+1}+r_{k+1}, \\ &\qquad I\cdot Z_1+r_1)+f(I\cdot Y+b)\big] \\ &= -\sum_{I\in\mathbb{F}_3^{k+1}}\Bigg[\Bigg[\sum_{\emptyset\neq J'\in\mathbb{F}_3^k} f(I\cdot Y+b+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\Bigg] \\ &\quad +\Bigg[\sum_{J'\in F_3^k}\Bigg(f(2I\cdot Y+2b-I\cdot Z_1-r_1+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t) \\ &\quad +f(I\cdot Z_1+r_1+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\Bigg)\Bigg]\Bigg] \\ &= -\sum_{0\neq J'\in\mathbb{F}_3^k}\left[\sum_{I\in\mathbb{F}_3^{k+1}} f(I\cdot Y+b+\sum_{t=2}^{k+1}J_tr_t+\sum_{t=2}^{k+1}J_tI\cdot Z_t)\right] \\ &\quad -\sum_{J'\in\mathbb{F}_3^k}\Bigg[\Bigg[\sum_{I\in\mathbb{F}_3^{k+1}} f(2I\cdot Y+2b-I\cdot Z_1-r_1+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\Bigg] \\ &\quad +\Bigg[\sum_{I\in\mathbb{F}_3^{k+1}} f(I\cdot Z_1+r_1+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\Bigg]\Bigg] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{0\neq J'\in\mathbb{F}_3^k}\left[-T_f^0(y_1+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,y_{k+1}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},b+\sum_{t=2}^{k+1}J_tr_t)\right]\\
&+ \sum_{J'\in\mathbb{F}_3^k}\Bigg[-T_f^0(2y_1-z_{1,1}+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,2y_{k+1}-z_{1,(k+1)}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},\\
&\qquad 2b-r_1+\sum_{t=2}^{k+1}J_tr_t)\\
&+ T_f^0(z_{1,1}+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,z_{1,k+1}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},r_1+\sum_{t=2}^{k+1}J_tr_t)\Bigg]
\end{aligned}
$$

□

**Claim 7.18.** *If (7.18) holds for all $I\in\mathbb{F}_3^{k+1}$, then*

$$
\begin{aligned}
T_{g_1}^1(Y,b) = &\sum_{0\neq J'\in\mathbb{F}_3^k}\left[-T_f^1(y_1+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,y_{k+1}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},b+\sum_{t=2}^{k+1}J_tr_t)\right]\\
&+\sum_{J'\in\mathbb{F}_3^k}\Bigg[T_f^1(2y_1-z_{1,1}+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,2y_{k+1}-z_{1,(k+1)}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},\\
&\qquad 2b-r_1+\sum_{t=2}^{k+1}J_tr_t)\Bigg]. \qquad (7.21)
\end{aligned}
$$

*Proof.*

$$
\begin{aligned}
T_{g_1}^1(Y,b) &= \sum_{I\in\mathbb{F}_3^{k+1}} I_1g_1(I\cdot Y+b)\\
&= \sum_{I\in\mathbb{F}_3^{k+1}} I_1\left[-T_f^1(I\cdot Y+b-I\cdot Z_1-r_1,I\cdot Z_2+r_2,\cdots,I\cdot Z_{k+1}+r_{k+1},\right.\\
&\qquad \left. I\cdot Z_1+r_1)+f(I\cdot Y+b)\right]\\
&= -\sum_{I\in\mathbb{F}_3^{k+1}} I_1\Bigg[\Bigg[\sum_{\emptyset\neq J'\in\mathbb{F}_3^k} f(I\cdot Y+b+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\Bigg]\\
&\quad +\Bigg[\sum_{J'\in F_3^k} f(2I\cdot Y+2b-I\cdot Z_1-r_1+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\Bigg]\Bigg]\\
&= -\sum_{0\neq J'\in\mathbb{F}_3^k}\left[\sum_{I\in\mathbb{F}_3^{k+1}} I_1f(I\cdot Y+b+\sum_{t=2}^{k+1}J_tr_t+\sum_{t=2}^{k+1}J_tI\cdot Z_t)\right]
\end{aligned}
$$

$$
\begin{aligned}
& - \sum_{J' \in \mathbb{F}_3^k} \left[ \sum_{I \in \mathbb{F}_3^{k+1}} I_1 f(2I \cdot Y + 2b - I \cdot Z_1 - r_1 + \sum_{t=2}^{k+1} J_t I \cdot Z_t + \sum_{t=2}^{k+1} J_t r_t) \right] \\
= & \sum_{0 \neq J' \in \mathbb{F}_3^k} \left[ -T_f^1(y_1 + \sum_{t=2}^{k+1} J_t z_{t,1}, \cdots, y_{k+1} + \sum_{t=2}^{k+1} J_t z_{t,(k+1)}, b + \sum_{t=2}^{k+1} J_t r_t) \right] \\
+ & \sum_{J' \in \mathbb{F}_3^k} \Bigg[ T_f^1(2y_1 - z_{1,1} + \sum_{t=2}^{k+1} J_t z_{t,1}, \cdots, 2y_{k+1} - z_{1,(k+1)} + \sum_{t=2}^{k+1} J_t z_{t,(k+1)}, \\
& \qquad 2b - r_1 + \sum_{t=2}^{k+1} J_t r_t) \Bigg]
\end{aligned}
$$

□

Let $E_2$ be the event that for every $J' \in \mathbb{F}_3^k, T_f^i(y_1 + \sum_t J_t z_{t,1}, \cdots, y_{k+1} + \sum_t J_t z_{t,(k+1)}, b + \sum_{t=2} k+1 J_t r_t) = 0, T_f^i(2y_1 - z_{1,1} + \sum_{t=2}^{k+1} J_t z_{t,1}, \cdots, 2y_{k+1} - z_{1,k+1} + \sum_{t=2}^{k+1} J_t z_{t,(k+1)}, 2b - r_1 + \sum_{t=2}^{k+1} J_t r_t) = 0$, and $T_f^0(z_{1,1} + \sum_{t=2}^{k+1} J_t z_{t,1}, \cdots, z_{1,k+1} + \sum_{t=2}^{k+1} J_t z_{t,k+1}, r_1 + \sum_{t=2}^{k+1} J_t r_t) = 0$. By the definition of $\eta_i$ and the union bound, we have:

$$\Pr[E_2] \geqslant 1 - 3^{k+1}\eta_i. \tag{7.22}$$

Suppose that $\eta_i \leqslant \frac{1}{(4k+14)3^{k+1}}$ holds. Then by (7.19) and (7.22), the probability that $E_1$ and $E_2$ hold is strictly positive. In other words, there exists a choice of the $z_{i,j}$'s and $r_i$'s for which all summands in either Claim 7.17 or in Claim 7.18, whichever is appropriate, is 0. This implies that $T_{g_i}^i(y_1, \cdots, y_{k+1}, b) = 0$. In other words, if $\eta_i \leqslant \frac{1}{(4k+14)3^{k+1}}$, then $g_i$ belongs to $\mathcal{P}_t$. □

**Remark 7.6.** *Over $\mathbb{F}_p$ we have: if $\eta_i < \frac{1}{2((p-1)k+6(p-1)+1)p^{k+1}}$, then $g_i$ belongs to $\mathcal{P}_t$ (if $k \geqslant 1$).*

*In case of $\mathbb{F}_p$, we can generalize (7.18) in a straightforward manner. Let $E_1'$ denote the event that all such events holds. We can similarly obtain*

$$\Pr[E_1'] \geqslant 1 - p^{k+1} \cdot 2((p-1)k + 6(p-1) + 1)\eta_i. \tag{7.23}$$

**Claim 7.19.** *Assume equivalent of (7.18) holds for all $I \in \mathbb{F}_p^{k+1}$, then*

$$\begin{aligned} T^i_{g_i}(Y,b) &= \sum_{0\neq J'\in\mathbb{F}_p^k}\left[-T^i_f(y_1+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,y_{k+1}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},b+\sum_{t=2}^{k+1}J_tr_t)\right] \\ &+ \sum_{J'\in\mathbb{F}_p^k}\left[\sum_{J_1\in\mathbb{F}_p;J_1\neq 1}J_1^i\left[-T^i_f(J_1y_1-(J_1-1)z_{1,1}+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,J_1y_{k+1}- \right.\right. \\ &\qquad \left.\left.(J_1-1)z_{1,(k+1)}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},J_1b-(J_1-1)r_1+\sum_{t=2}^{k+1}J_tr_t)\right]\right] \end{aligned} \tag{7.24}$$

*Proof.*

$$\begin{aligned} T^i_{g_i}(Y,b) &= \sum_{I\in\mathbb{F}_p^{k+1}} I_1^i g_i(I\cdot Y+b) \\ &= \sum_{I\in\mathbb{F}_p^{k+1}} I_1^i\left[-T^i_f(I\cdot Y+b-I\cdot Z_1-r_1,I\cdot Z_2+r_2,\cdots,I\cdot Z_{k+1}+r_{k+1},\right. \\ &\qquad \left. I\cdot Z_1+r_1)+f(I\cdot Y+b)\right] \\ &= -\sum_{I\in\mathbb{F}_p^{k+1}} I_1^i\left[\left[\sum_{\emptyset\neq J'\in\mathbb{F}_p^k} f(I\cdot Y+b+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\right]\right. \\ &+ \left[\sum_{J_1\in\mathbb{F}_p,J_1\neq 1}J_1^i\left[\sum_{J'\in\mathbb{F}_p^k} f(J_1I\cdot Y+J_1b-(J_1-1)I\cdot Z_1-(J_1-1)r_1\right.\right. \\ &\qquad \left.\left.\left.+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\right]\right] \\ &= -\sum_{0\neq J'\in\mathbb{F}_p^k}\left[\sum_{I\in\mathbb{F}_p^{k+1}} I_1^i f(I\cdot Y+b+\sum_{t=2}^{k+1}J_tr_t+\sum_{t=2}^{k+1}J_tI\cdot Z_t)\right] \\ &- \sum_{J'\in\mathbb{F}_p^k}\left[\sum_{J_1\in\mathbb{F}_p;J_1\neq 1}J_1^i\left[\sum_{I\in\mathbb{F}_p^{k+1}} I_1^i f(J_1I\cdot Y+J_1b-(J_1-1)I\cdot Z_1-(J_1-1)r_1\right.\right. \\ &\qquad \left.\left.\left.+\sum_{t=2}^{k+1}J_tI\cdot Z_t+\sum_{t=2}^{k+1}J_tr_t)\right]\right] \\ &= \sum_{0\neq J'\in\mathbb{F}_p^k}\left[-T^i_f(y_1+\sum_{t=2}^{k+1}J_tz_{t,1},\cdots,y_{k+1}+\sum_{t=2}^{k+1}J_tz_{t,(k+1)},b+\sum_{t=2}^{k+1}J_tr_t)\right] \end{aligned}$$

$$+ \sum_{J' \in \mathbb{F}_p^k} \left[ \sum_{J_1 \in \mathbb{F}_p; J_1 \neq 1} J_1^i \left[ -T_f^i(J_1 y_1 - (J_1 - 1) z_{1,1} + \sum_{t=2}^{k+1} J_t z_{t,1}, \cdots, \quad J_1 y_{k+1} \right. \right.$$
$$\left. \left. -(J_1 - 1) z_{1,(k+1)} + \sum_{t=2}^{k+1} J_t z_{t,(k+1)}, J_1 b - (J_1 - 1) r_1 + \sum_{t=2}^{k+1} J_t r_t \right) \right] \right]$$

□

*Let $E_2'$ be the event analogous to the event $E_2$ in Claim 7.18. Then by the definition of $\eta_i$ and the union bound, we have*

$$\Pr[E_2'] \geqslant 1 - 2p^{k+1}\eta_i. \tag{7.25}$$

*Then if we are given that $\eta_i < \frac{1}{2((p-1)k+6(p-1)+1)p^{k+1}}$, then the probability that $E_1'$ and $E_2'$ hold is strictly positive. Therefore, this implies $T_{g_i}^i(y_1, \cdots, y_{k+1}, b) = 0$.*

### *7.4.5 Proof of Lemma 7.13*

For each $C \in \mathbb{F}_3^{k+1}$, let $X_C$ be the indicator random variable whose value is 1 if and only if $f(C \cdot Y + b) \neq g(C \cdot Y + b)$, where $Y = \langle y_1, \ldots, y_{k+1} \rangle$. Clearly, $\Pr[X_C = 1] = \delta$ for every $C$. It follows that the random variable $X = \sum_C X_C$ which counts the number of points $v$ of the required form in which $f(v) \neq g(v)$ has expectation $\mathbb{E}[X] = 3^{k+1}\delta = \ell \cdot \delta$. It is not difficult to check that the random variables $X_C$ are pairwise independent, since for any two distinct $C_1 = (C_{1,1}, \ldots, C_{i,k+1})$ and $C_2 = (C_{2,1}, \ldots, C_{2,k+1})$, the sums $\sum_{i=1}^{k+1} C_{1,i} y_i + b$ and $\sum_{i=1}^{k+1} C_{2,i} y_i + b$ attain each pair of distinct values in $\mathbb{F}_3^m$ with equal probability when the vectors are chosen randomly and independently. Since $X_C$'s are pairwise independent, $\mathrm{Var}[X] = \sum_C \mathrm{Var}[X_C]$. Since $X_C$'s are boolean random variables, we note

$$\mathrm{Var}[X_C] = \mathbb{E}[X_C^2] - (\mathbb{E}[X_C])^2 = \mathbb{E}[X_C] - (\mathbb{E}[X_C])^2 \leqslant \mathbb{E}[X_C].$$

Thus we obtain $\mathrm{Var}[X] \leqslant \mathbb{E}[X]$, so $\mathbb{E}[X^2] \leqslant \mathbb{E}[X]^2 + \mathbb{E}[X]$. Next we use the following well known inequality which holds for a random variable $X$ taking nonnegative, integer values,

$$\Pr[X > 0] \geqslant \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}.$$

Indeed if $X$ attains value $i$ with probability $p_i$, then we have

$$(\mathbb{E}[X])^2 = \left( \sum_{i>0} i p_i \right)^2 = \left( \sum_{i>0} i \sqrt{p_i} \sqrt{p_i} \right)^2 \leqslant \left( \sum_{i>0} i p_i \right) \left( \sum_{i>0} p_i \right) = \mathbb{E}[X] \cdot \Pr[X > 0],$$

where the inequality follows by the Cauchy-Schwartz inequality. In our case, this implies

$$\Pr[X > 0] \geqslant \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]} \geqslant \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X] + (\mathbb{E}[X])^2} = \frac{\mathbb{E}[X]}{1 + \mathbb{E}[X]}.$$

Therefore,

$$\mathbb{E}[X] \geqslant \Pr[X=1] + 2\Pr[X \geqslant 2] = \Pr[X=1] + 2\left(\frac{\mathbb{E}[X]}{1+\mathbb{E}[X]} - \Pr[X=1]\right)$$
$$= \frac{2\mathbb{E}[X]}{1+\mathbb{E}[X]} - \Pr[X=1].$$

After simplification we obtain,

$$\Pr[X=1] \geqslant \frac{1-\mathbb{E}[X]}{1+\mathbb{E}[X]} \cdot \mathbb{E}[X].$$

The proof is complete by recalling that $\mathbb{E}[X] = \ell \cdot \delta$. □

## *7.5 A Lower Bound and Improved Self-correction*

### *7.5.1 A Lower Bound*

The next theorem is a simple modification of a theorem in [1] and essentially implies that our result is almost optimal.

**Proposition 7.20.** *Let $\mathcal{F}$ be any family of functions $f : \mathbb{F}_p^n \to \mathbb{F}_p$ that corresponds to a linear code $\mathcal{C}$. Let $d$ denote the minimum distance of the code $\mathcal{C}$ and let $\bar{d}$ denote the minimum distance of the dual code of $\mathcal{C}$.*
*Every one-sided testing algorithm for the family $\mathcal{F}$ must perform $\Omega(\bar{d})$ queries, and if the distance parameter $\varepsilon$ is at most $d/p^{n+1}$, then $\Omega(1/\varepsilon)$ is also a lower bound for the necessary number of queries.*

Lemma 7.4 and Proposition 7.20 gives us the following corollary.

**Corollary 7.5.** *Every one-sided tester for testing $\mathcal{P}_t$ with distance parameter $\varepsilon$ must perform $\Omega(max(\frac{1}{\varepsilon}, (1 + ((t+1) \bmod (p-1)))p^{\frac{t+1}{p-1}}))$ queries.*

### *7.5.2 Improved Self-correction*

From Lemmas 7.9, 7.11 and 7.12 the following corollary is immediate:

**Corollary 7.6.** *Consider a function $f : \mathbb{F}_3^n \to \mathbb{F}_3$ that is $\varepsilon$-close to a degree-$t$ polynomial $g : \mathbb{F}_3^n \to \mathbb{F}_3$, where $\varepsilon < \frac{1}{(4k+14)3^{k+1}}$. (Assume $k \geqslant 1$.) Then the function $f$ can be self-corrected. That is, for any given $x \in \mathbb{F}_3^n$, it is possible to obtain the value $g(x)$ with probability at least $1 - 3^{k+1}\varepsilon$ by querying $f$ on $3^{k+1}$ points on $\mathbb{F}_3^n$.*

An analogous result may be obtained for the general case. We, however, improve the above corollary slightly. The above corrector does not allow any error in the $3^{k+1}$ points it queries. We obtain a stronger result by querying on a slightly larger flat $H$, but allowing some errors. Errors are handled by decoding the induced Reed-Muller code on $H$.

**Proposition 7.21.** *Consider a function $f : \mathbb{F}_p^n \to \mathbb{F}_p$ that is $\varepsilon$-close to a degree-$t$ polynomial $g : \mathbb{F}_p^n \to \mathbb{F}_p$. Then the function $f$ can be self-corrected. That is, assume $K > (k+1)$, then for any given $x \in \mathbb{F}_p^n$, the value of $g(x)$ can be obtained with probability at least $1 - \frac{\varepsilon}{(1-\varepsilon \cdot p^{k+1})^2} \cdot p^{-(K-2k-3)}$ with $p^K$ queries to $f$.*

*Proof.* Our goal is to correct the $\mathrm{RM}_p(t,n)$ at the point $x$. Assume $t = (p-1)\cdot k + R$, where $0 \leqslant R \leqslant (p-2)$. Then the relative distance of the code $\delta$ is $(1 - R/p)p^{-k}$. Note that $2p^{-k-1} \leqslant \delta \leqslant p^{-k}$. Recall that the local testability test requires a $(k+1)$-flat, i.e., it tests $\sum_{c_1,\cdots,c_{k+1} \in \mathbb{F}_p} c_1^{p-2-R} f(y_0 + \sum_{i=1}^{k+1} c_i y_i) = 0$, where $y_i \in \mathbb{F}_p^n$.

We choose a slightly larger flat, i.e., a $K$-flat with $K > (k+1)$ to be chosen later. We consider the code restricted to this $K$-flat with point $x$ being the origin. We query $f$ on this $K$-flat. It is known that a majority logic decoding algorithm exists that can decode Reed-Muller codes up to half the minimum distance for any choice of parameters (see [99]). Thus if the number of errors is small we can recover $g(x)$.

Let the relative distance of $f$ from the code be $\varepsilon$ and let $S$ be the set of points where it disagrees with the closest codeword. Let the random $K$-flat be $H = \{x + \sum_{i=1}^{K} t_i u_i | t_i \in \mathbb{F}, u_i \in_R \mathbb{F}_p^n\}$. Let the random variable $Y_{\langle t_1,\cdots,t_K\rangle}$ take the value 1 if $x + \sum_{i=1}^{K} u_i t_i \in S$ and 0 otherwise. Let $D = \mathbb{F}^K \setminus \{0\}$ and $U = \langle u_1, \cdots, u_K \rangle$. Define $Y = \sum_{\langle t_1,\cdots,t_K\rangle \in D} Y_{\langle t_1,\cdots,t_K\rangle}$ and $\ell = (p^K - 1)$. We would like to bound the probability

$$\Pr_U[|Y - \varepsilon\ell| \geqslant (\delta/2 - \varepsilon)\ell].$$

Since $\Pr_U[Y_{t_1,\cdots,t_K} = 1] = \varepsilon$, by linearity we get $\mathbb{E}_U[Y] = \varepsilon\ell$. Let $T = \langle t_1, \cdots, t_K \rangle$. Now

$$\begin{aligned} Var[Y] = & \sum_{T \in \mathbb{F}^K - \{0\}} Var[Y_T] + \sum_{T \neq T'} Cov[Y_T, Y_{T'}] \\ = & \ \ell(\varepsilon - \varepsilon^2) + \sum_{T \neq \lambda T'} Cov[Y_T, Y_{T'}] + \sum_{T = \lambda T'; 1 \neq \lambda \in \mathbb{F}^*} Cov[Y_T, Y_{T'}] \\ \leqslant & \ \ell(\varepsilon - \varepsilon^2) + \ell \cdot (p-2)(\varepsilon - \varepsilon^2) \\ = & \ \ell(\varepsilon - \varepsilon^2)(p-1) \end{aligned}$$

The above follows from the fact that when $T \neq \lambda T'$ then the corresponding events $Y_T$ and $Y_{T'}$ are independent and therefore $Cov[Y_T, Y_{T'}] = 0$. Also, when $Y_T$ and $Y_{T'}$ are dependent then $Cov[Y_T, Y_{T'}] = \mathbb{E}_U[Y_T Y_{T'}] - \mathbb{E}_U[Y_T]\mathbb{E}_U[Y_{T'}] \leqslant \varepsilon - \varepsilon^2$.
Therefore, by Chebyshev's inequality we have (assuming $\varepsilon < p^{-(k+1)}$)

$$\Pr_U[|Y - \varepsilon\ell| \geqslant (\delta/2 - \varepsilon)\ell] \leqslant \frac{\ell\varepsilon(1-\varepsilon)(p-1)}{(\delta/2 - \varepsilon)^2 \ell^2}$$

Now note $(\delta/2 - \varepsilon) \geqslant (p^{-k-1} - \varepsilon) = (1 - \varepsilon \cdot p^{k+1})p^{-k-1}$. We thus have

$$\begin{aligned}\Pr_U[|Y - \varepsilon\ell| \geqslant (\delta/2 - \varepsilon)\ell] &\leqslant \frac{\varepsilon(1-\varepsilon)(p-1)}{(1-\varepsilon\cdot p^{k+1})^2 p^{-2k-2}\ell} \\ &\leqslant \frac{\varepsilon p}{(1-\varepsilon\cdot p^{k+1})^2 p^{-2k-2}(\ell+1)} \\ &= \frac{\varepsilon}{(1-\varepsilon\cdot p^{k+1})^2}\cdot p^{-(K-2k-3)}.\end{aligned}$$

□

### *7.6 Bibliographics Notes*

The results presented in this chapter appear in [72].

As was mentioned earlier, the study of low degree testing (along with self-correction) dates back to the work of Blum, Luby and Rubinfeld ([21]), where an algorithm was required to test whether a given function is linear. The approach in [21] later naturally extended to yield testers for low degree polynomials over fields larger than the total degree. Roughly, the idea is to project the given function on to a random line and then test if the projected univariate polynomial has low degree. Specifically, for a purported degree $t$ function $f : \mathbb{F}_q^n \to \mathbb{F}_q$, the  test works as follows. Pick vectors $y$ and $b$ from $\mathbb{F}_q^n$ (uniformly at random), and distinct $s_1, \cdots, s_{t+1}$ from $\mathbb{F}_q$ arbitrarily. Query the oracle representing $f$ at the $t+1$ points $b + s_i y$ and extrapolate to a degree $t$ polynomial $P_{b,y}$ in one variable $s$. Now test for a random $s \in \mathbb{F}_p$ if

$$P_{b,y}(s) = f(b + sy)$$

(for details see [93],[42]). Similar ideas are also employed to test whether a given function is a low degree polynomial in each of its variable (see [36, 8, 6]).

Alon et al. give a tester over field $\mathbb{F}_2$ for any degree up to the number of inputs to the function (i.e., for any non-trivial degree) [1]. In other words, their work shows that Reed-Muller codes are locally testable. Under the coding theory interpretation, their tester picks a random minimum-weight codeword from the dual code and checks if it is orthogonal to the input vector. It is important to note that these minimum-weight code words generate the Reed-Muller code.

Specifically their test works as follows: given a function $f : \{0,1\}^n \to \{0,1\}$, to test if the given function $f$ has degree at most $t$, pick $(t+1)$-vectors $y_1, \cdots, y_{t+1} \in \{0,1\}^n$ and test if

$$\sum_{\emptyset \neq S \subseteq [t+1]} f(\sum_{i \in S} y_i) = 0.$$

Independent of [72], Kaufman and Ron, generalizing a characterization result of [42], gave a tester for low degree polynomials over general finite fields (see [74]). They show that a given polynomial is of degree at most $t$ *if and only if* the restriction of the polynomial to every affine subspace of suitable dimension is of degree at most $t$. Following this

idea, their tester chooses a random affine subspace of a suitable dimension, computes the polynomial restricted to this subspace, and verifies that the coefficients of the higher degree terms are zero[4]. To obtain constant soundness, the test is repeated many times. An advantage of the approach presented in this chapter is that in one round of the test (over the prime field) we test only one linear constraint, whereas their approach needs to test multiple linear constraints.

A basis of RM consisting of minimum-weight codewords was considered in [28, 29]. We extend their result to obtain a different exact characterization for low-degree polynomials. Furthermore, it seems likely that their exact characterization can be turned into a robust characterization following analysis similar to our robust characterization. However, our basis is cleaner and yields a simpler analysis. We point out that for degree smaller than the field size, the exact characterization obtained from [28, 29] coincides with [21, 93, 42]. This provides an alternate proof to the exact characterization of [42] (for more details, see Remark 7.3 and [42]).

In an attempt to generalize our result to more general fields, we obtain an exact characterization of low degree polynomials over general finite fields [71] (see [86] for more details). This provides an alternate proof to the result of Kaufman and Ron [74] described earlier. Specifically the result says that a given polynomial is of degree at most $t$ *if and only if* the restriction of the polynomial to every affine subspace of dimension $\lceil \frac{t+1}{q-q/p} \rceil$ (and higher) is of degree at most $t$.

Recently Kaufman and Litsyn ([73]) show that the dual of BCH codes are locally testable. They also give a sufficient condition for a code to be locally testable. The condition roughly says that if the number of fixed length codewords in the dual of the union of the code and its $\varepsilon$-far coset is suitably smaller than the same in the dual of the code, then the code is locally testable. Their argument is more combinatorial in nature and needs the knowledge of weight-distribution of the code and thus differs from the self-correction approach used in this work.

[4]Since the coefficients can be written as linear sums of the evaluations of the polynomial, this is equivalent to check several linear constraints