Electronic Colloquium on Computational Complexity, Report No. 11 (2012)

Testers and their Applications

Nader H. Bshouty
Department of Computer Science
Technion, Israel
bshouty@cs.technion.ac.il

February 7, 2012

Abstract

We develop a new notion called *tester of a class $\mathcal{M}$ of functions $f : \mathcal{A} \to \mathcal{C}$* that maps the elements $\boldsymbol{a} \in \mathcal{A}$ in the domain $\mathcal{A}$ of the function to a finite number (the size of the tester) of elements $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_t$ in a smaller sub-domain $\mathcal{B} \subset \mathcal{A}$ where the property $f(\boldsymbol{a}) \neq 0$ is preserved for all $f \in \mathcal{M}$. I.e., for all $f \in \mathcal{M}$ and $\boldsymbol{a} \in \mathcal{A}$ if $f(\boldsymbol{a}) \neq 0$ then $f(\boldsymbol{b}_i) \neq 0$ for some i.

We use tools from elementary algebra and algebraic function fields to construct testers of almost optimal size in deterministic polynomial time in the size of the tester. We then apply testers to deterministically construct new set of objects with some combinatorial and algebraic properties that can be used to derandomize some algorithms.

We show that those new constructions are almost optimal and for many of them meet the union bound of the problem. Constructions include, d-restriction problems, perfect hash, universal sets, cover-free families, separating hash functions, polynomial restriction problems, black box polynomial identity testing for polynomials and circuits over small fields and hitting sets.

Keywords: Combinatorial objects, Derandomization, d-Restriction problems, Perfect hash, Universal sets, Cover-Free families, Separating hash functions, Hitting sets, Polynomial restriction problems.

ISSN 1433-8092

Contents

1 Introduction

1.1 Testers

A *tester* of a class of multivariate polynomial $\mathcal{M}$ over n variables is a set L of maps from a complex (algebraic) structure $\mathcal{A}^n$ to a simple one $\mathcal{B}^n$ that preserve the property $f(\boldsymbol{a}) \neq 0$ for every $f \in \mathcal{M}$, i.e., for all $f \in \mathcal{M}$ and $\boldsymbol{a} \in \mathcal{A}$ if $f(\boldsymbol{a}) \neq 0$ then $f(\ell(\boldsymbol{a})) \neq 0$ for some $\ell \in L$. See a formal definition in section 2 page 14. In this paper we study testers when $\mathcal{A}$, the domain of the functions in $\mathcal{M}$, is a field or a subspace of a field and $\mathcal{B} \subset \mathcal{A}$ is a small subfield. We use tools from elementary algebra and algebraic function fields to construct testers of almost optimal size $|L|$ in polynomial time.

One application of tester is the following: Suppose we need to construct a small set of vectors $S \subset \Sigma^n$ for some alphabet Σ that satisfies some property P. We map Σ into a field $\mathbb{F}$ and find a set of functions $\mathcal{M}_P$ where $S \subset \mathbb{F}^n$ satisfies property P if and only if S is a hitting set for $\mathcal{M}_P$, i.e., for every $f \in \mathcal{M}_P$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$. We then extend $\mathbb{F}$ to a larger field $\mathbb{K}$ (or $\mathbb{F}$-algebra $\mathcal{A}$). Find $S' \subset \mathbb{K}^n$ that is a hitting set for $\mathcal{M}_P$ (which supposed to be easier). Then use tester to change the hitting set $S' \subset \mathbb{K}^n$ over $\mathbb{K}$ to a hitting set $S \subset \mathbb{F}^n$ over $\mathbb{F}$.

In this paper we consider two main classes of multivariate polynomials over finite fields $\mathbb{F}_q$ with q elements. The first class is $\mathcal{P}(\mathbb{F}_q, n, d)$, the class of all multivariate polynomials with n variables and total degree d. The second class is $\mathcal{DML}(\mathbb{F}_q, n, d)$, the class of all multivariate polynomials f with dn variables $x_{i,j}$, $i = 1, \ldots, d$, $j = 1, \ldots, n$ where each monomial in f is of the form $x_{1,i_1} x_{2,i_2} \cdots x_{d,i_d}$.

For the class $\mathcal{P}(\mathbb{F}_q, n, d)$ we give the following testers that map elements from $\mathcal{A} = \mathbb{F}_{q^t}$ to $\mathcal{B} = \mathbb{F}_q$

t	q	Poly time size=$O()$	2^t time size=$O()$	Explicit size=$O()$	Lower Bound size=$\Omega()$
All	$q \geq d(t-1)+1$	dt	dt	dt	dt
I.S.	$q \geq c(d+1)^2$, q P.S.	d^3t	d^2t	dt	dt
All	$q \geq c(d+1)^2$, q P.S.	d^3t	d^2t	d^2t	dt
I.S.	$q \geq c(d+1)$	d^4t	d^3t	d^2t	dt
All	$q \geq c(d+1)$	d^4t	d^3t	d^3t	dt
I.S.	$q \geq d+1$	d^5t	d^4t	d^3t	dt
All	$q \geq d+1$	d^5t	d^4t	d^4t	dt
All	$q < d+1$				∞

In the table I.S. stands for "infinite sequence of integers t" and P.S. stands for "perfect square". The second column in the above table gives bounds on q, the size of the field. The parameter c is any constant greater than 1. Notice that when $q < d+1$ no tester exists for $\mathcal{P}(\mathbb{F}_q, n, d)$. The third column gives the size of testers that can be constructed in polynomial time. The forth column gives the size of the testers that can be constructed in time 2^t. This will be used for fields $\mathbb{F}_{q^r}$ where $r = O(\log t)$.

The fifth column gives the size of the testers that can be explicitly constructed but are not known to be polynomial time constructed. The last column gives lower bounds. For example, the second row and third column in the table says that, for any constant $c > 1$, perfect square $q \geq c(d+1)^2$ there is a tester L for $\mathcal{P}(\mathbb{F}_q, n, d)$ of size $O(d^3 t)$ for infinite sequence of integers t.

We then study testers for $\mathcal{P}(\mathbb{F}_q, n, d)$ that map subspaces $S \subseteq \mathbb{F}_{q^T}$ of size q^t to $\mathbb{F}_q$ for some $T > t$. We were able to give better bounds for the size of such testers that can be constructed in polynomial time. For small t those bounds match the size of the explicit constructions in the above table.

For the class $\mathcal{DML}(\mathbb{F}_q, n, d)$ when $q \geq d+1$ the testers for $\mathcal{P}(\mathbb{F}_q, n, d)$ are also testers for $\mathcal{DML}(\mathbb{F}_q, n, d)$. For $q \leq d$ we give the following testers that map elements from $\mathcal{A} = \mathbb{F}_{q^t}$ to $\mathcal{B} = \mathbb{F}_q$

q	Upper Bound size =	Lower Bound size =
2	$2^{1.66d} \cdot t$	$2^d \cdot t$
3	$2^{1.12d} \cdot t$	$2^{0.58d} \cdot t$
4	$2^{0.87d} \cdot t$	$2^{0.41d} \cdot t$
5	$2^{0.72d} \cdot t$	$2^{0.32d} \cdot t$
7	$2^{0.55d} \cdot t$	$2^{0.22d} \cdot t$
q	$2^{O(\log q/q)d} \cdot t$	$2^{\Omega(1/q)d} \cdot t$

The upper bound and lower bounds are bounds on the size of the testers. The time complexity of constructing the above testers is within $poly(t)$ of their upper bound size.

1.2 d-Restriction Problems

We apply testers to the following problems

d-Restriction Problems: A *d-restriction problem* [53, 5] is a problem of the following form: Given an alphabet Σ of size $|\Sigma| = q$, a length n and a class $\mathcal{M}$ of nonzero functions $f_i : \Sigma^d \to \{0, 1\}$, $i = 1, 2, \ldots, t$. Find a set $A \subseteq \Sigma^n$ of small size such that: For any $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ and $f \in \mathcal{M}$ there is $\boldsymbol{a} \in A$ such that $f(a_{i_1}, \ldots, a_{i_d}) \neq 0$.

The d-restriction problems considered in this paper are perfect hash, universal sets, cover-free families and separating hash functions. Using tester we achieve an asymptotically optimal construction for those problems.

Perfect Hash: For $d \leq q$ we say that the set H of function $h : [n] \to \mathbb{F}_q$ is a (n, q, d)-*perfect hashing* [5] (or (n, d, q)-*splitter* [53]) if for all subsets $S \subseteq [n]$ of size $|S| = d$ there is a *hash function* $h \in H$ such that $h|_S$ is injective (one-to-one) on S, i.e., $|h|_S(S)| = d$. Obviously, perfect hash is a d-restriction problem with the alphabet $\Sigma = \mathbb{F}_q$ where we regard each element in $A \subseteq \mathbb{F}_q^n$ as a function.

In [53, 5] it was shown that there is an (n, d^2, d)-perfect hashing of size $O(d^4 \log d \log n)$ that can be constructed in $poly(n, d)$ time. Wang and Xing [81] used algebraic function fields and gave an explicit[1] (n, d^4, d)-perfect hashing of size $O((d^2/\log d)\log n)$ for infinite sequence of integers n. For any q the only known polynomial time construction is of size $O(d^2 \log d \log n)$, [53, 5]. Blackburn and Wild [23] gave an explicit optimal construction when q is very large compared to d and $\log n$. In this paper we use testers to give a polynomial time construction that is almost optimal for any q and n as described in the following Table

n	q	poly time Size =	Union Bound	Lower Bound
I.S.	$q \geq \frac{c}{4}d^4$	$d^2\frac{\log n}{\log q}$	$d\frac{\log n}{\log q}$	$\frac{\log n}{\log q}$
all	$q \geq \frac{c}{4}d^4$	$d^4\frac{\log n}{\log q}$	$d\frac{\log n}{\log q}$	$\frac{\log n}{\log q}$
I.S.	$q \geq \frac{c}{2}d^2$	$d^4\frac{\log n}{\log d}$	$d\frac{\log n}{\log(2q/(d(d-1)))}$	$\frac{\log n}{\log q}$
all	$q \geq \frac{c}{2}d^2$	$d^6\frac{\log n}{\log d}$	$d\frac{\log n}{\log(2q/(d(d-1)))}$	$\frac{\log n}{\log q}$
I.S.	$q \geq \frac{d(d+1)}{2}+1$	$d^6\frac{\log n}{\log d}$	$d\log n$	$\frac{\log n}{\log q}$
all	$q \geq \frac{d(d+1)}{2}+1$	$d^8\frac{\log n}{\log d}$	$d\log n$	$\frac{\log n}{\log q}$

In the table, I.S. means "for infinite sequence of integers n" and $c > 1$ is any constant. The union bound is a non-constructive bound that can be achieved using probabilistic method by randomly uniformly choosing hash functions [7].

Our results improve all the results in the literature and give the first polynomial time $O(\log n/\log q)$ size construction for a fixed d and any n and q.

Universal sets: Another d-restriction problem is the problem of (n, d)-universal set over an alphabet of size q. This problem is d-restriction problem where $\mathcal{M}$ contains all the nonzero functions $f : \Sigma^d \to \{0, 1\}$. That is, $A \subseteq \mathbb{F}_q^n$ is (n, k)-universal set if for every $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ and $\alpha_1, \ldots, \alpha_d \in \mathbb{F}_q$ there is $\boldsymbol{a} \in A$ such that $a_{i_1} = \alpha_1, a_{i_2} = \alpha_2, \ldots, a_{i_d} = \alpha_d$.

The lower bound $\Omega(q^{d-1}\log n/\log q)$ for the size of (n, d)-universal set over $\mathbb{F}_q$ can be derived from [44]. The union bound gives the upper bound $O(dq^d \log n)$. The best known polynomial time[2] construction for this problem gives a universal set of size $d^{O(\log d/\log q)}q^d \log n$ for $q < d$ and $O(d^5(\log d)^2\ q^d \log n)$, for $q > d$ [53, 2]. For $q < d$ the size of the construction in [53], $d^{O(\log d/\log q)}q^d \log n$, is better than the size we get here. For $q \geq d$ we give a polynomial time construction of (n, d)-universal set of size $O(d^4(q^d/\log q)\log n)$. When $q = \Omega(d^2)$ and perfect square, for infinite number of integers n,

[1] In this paper, when we say "explicit construction" we mean a construction that uses elementary algebra and algebraic function fields in which each step of the construction is indicated. But it is not clear if the construction is polynomial time construction.

[2] In this paper, when we say polynomial time we mean $poly(n) \cdot s$ where s is the size of the construction.

we give a polynomial time construction of size $O(d^2(q^d/\log q)\log n)$ and an explicit construction of size $O(d(q^d/\log q)\log n)$. Although those results slightly improve the existing results from the literature, the latter bound is a surprising result since it exceeds the union bound $O(dq^d \log n)$ achieved by probabilistic method. This shows that our new technique may lead to new combinatorial bounds that exceed the union bounds.

The following Table summarizes the results

n	q	Poly Time Size=	Explicite Size=
I.S.	$q \geq c(d+1)^2$, q P.S.	$d^2 q^d \frac{\log n}{\log q}$	$dq^d \frac{\log n}{\log q}$
all	$q \geq c(d+1)^2$, q P.S.	$d^2 q^d \frac{\log n}{\log q}$	$d^2 q^d \frac{\log n}{\log q}$
I.S.	$q \geq c(d+1)$	$d^3 q^d \frac{\log n}{\log d}$	$d^2 q^d \frac{\log n}{\log d}$
all	$q \geq c(d+1)$	$d^3 q^d \frac{\log n}{\log d}$	$d^3 q^d \frac{\log n}{\log d}$
I.S.	$q \geq d+1$	$d^4 q^d \frac{\log n}{\log d}$	$d^3 q^d \frac{\log n}{\log d}$
all	$q \geq d+1$	$d^4 q^d \frac{\log n}{\log d}$	$d^4 q^d \frac{\log n}{\log d}$
all	$q \leq d+1$	$d^7 q^{(1+c_q/\log q)d} \log n$	$d^5 q^{(1+c_q/\log q)d} \log n$
all	$q = 2$	$d^7 2^{2.66d} \log n$	$d^5 2^{2.66d} \log n$

In the table $c \geq 1$ is any constant, I.S. means "for infinite sequence of integers n", P.S. means "perfect square" and c_q is the constant from Theorem 21.

Cover-Free Families (CFF): Let X be a set with N elements and $\mathcal{B}$ be a set of subsets (blocks) of X. We say that $(X, \mathcal{B})$ is (w, r)-*cover-free family* ((w, r)-CFF), [43], if for any w blocks $B_1, \ldots, B_w \in \mathcal{B}$ and any other r blocks $A_1, \ldots, A_r \in \mathcal{B}$, we have

$$\bigcap_{i=1}^{w} B_i \not\subseteq \bigcup_{j=1}^{r} A_j.$$

The goal is to find (w, r)-CFF with small N. Let $N((w, r), n)$ denotes the minimum number of points in any (w, r)-CFF having n blocks. When $w = 1$, the problem is called *group testing*. It is easy to see that CFF is d-restriction problem.

The problem (w, r)-cover-free family is equivalent to the following problem: An (w, r)-cover-free family is a set $\mathcal{F} \subseteq \{0, 1\}^n$ such that for every $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ where $d = w + r$ and every $J \subset [d]$ of size $|J| = w$ there is $\boldsymbol{a} \in \mathcal{F}$ such that $a_{i_k} = 0$ for all $k \notin J$ and $a_{i_j} = 1$ for all $j \in J$. Then $N((w, r), n)$ is the minimum size of such $\mathcal{F}$.

There are several lower bounds for $N((w, r), n)$. We give the one in [78]

$$N((w, r), n) \geq \Omega\left(\frac{d\binom{d}{w}}{\log\binom{d}{w}} \log n\right).$$

It follows from [79], that for infinite sequence of integers n, a (w,r)-cover free family of size $M = O\left((wr)^{\log^* n}\log n\right)$ can be constructed in polynomial time. In [48], Liu and Shen show that for fixed d (and therefore fixed w,r) there is an explicit construction of size $O(\log n)$ for infinite sequence of integers n. This also follows, for all n and fixed d, from the (n,d)-universal set constructed in [52].

In this paper we show: for any constant $c>1$, the following (w,r)-CFF can be constructed in polynomial time

n	w	Poly time Size=	Union Bound	Lower Bound
I.S	$O(1)$	$\frac{r^{w+2}}{\log r}\log n$	$r^{w+1}\log n$	$\frac{r^{w+1}}{\log r}\log n$
all	$O(1)$	$\frac{r^{w+3}}{\log r}\log n$	$r^{w+1}\log n$	$\frac{r^{w+1}}{\log r}\log n$
I.S.	$o(r)$	$\frac{w^2(ce)^w r^{w+2}}{\log r}\log n$	$\frac{r^{w+1}}{(w/e)^{w-1/2}}\log n$	$\frac{r^{w+1}}{(w/e)^{w+1}\log r}\log n$
all	$o(r)$	$\frac{w^3(ce)^w r^{w+3}}{\log r}\log n$	$\frac{r^{w+1}}{(w/e)^{w-1/2}}\log n$	$\frac{r^{w+1}}{(w/e)^{w+1}\log r}\log n$

To the best of our knowledge, this is the first asymptotically optimal (within $poly(r)$) construction for any d, $w=O(1)$ and n. For $w=o(r)$ our construction is within $r^{o(w)}$ of the lower bound.

Separating Hash Family: Let X and Σ be sets of size n and q, respectively. We call a set $\mathcal{F}$ of functions $f: X\to\Sigma$ an $(M;n,q,\{d_1,d_2,\ldots,d_r\})$ *separating hash family* (SHF), [75, 76], if $|\mathcal{F}|=M$ and for all pairwise disjoint subsets $C_1,C_2,\ldots,C_r\subseteq X$ with $|C_i|=d_i$ for $i=1,2,\ldots,r$, there is at least one function $f\in\mathcal{F}$ such that $f(C_1),f(C_2),\ldots,f(C_r)$ are pairwise disjoint subsets. The minimal M is denoted by $M(n,q,\{d_1,d_2,\ldots,d_r\})$. It is easy to see that this problem is d-restriction problem.

In [22], Bazrafshan and Trund proved that for $D_1=d_1+d_2+\cdots+d_r$,

$$M(n,q,\{d_1,d_2,\ldots,d_r\})=\Omega\left(D_1\frac{\log n}{\log q}\right).$$

In [79], Stinson et. al. proved that an $(M;n,q,\{d_1,d_2\})$ separating hash families of size

$$M=O((d_1d_2)^{\log^* n}\log n)$$

can be constructed in polynomial time for infinite sequence of integers n and $q>d_1d_2$. The same proof gives a polynomial time construction for any separating hash family of size

$$M=O(D_2^{\log^* n}\log n)$$

where

$$D_2=\sum_{1\le i_1<i_2\le r} d_{i_1}d_{i_2}$$

when $q > D_2$.

In [48], Liu and Shen provide an explicit construction of $(M; n, q, \{d_1, d_2\})$ separating hash family using algebraic curves over finite fields. They show that for infinite sequence of integers n there is an explicit $(M; n, q, \{d_1, d_2\})$ separating hash families of size $O(\log n)$ for fixed d_1 and d_2. This also follows from [52], an $(n, d_1 + d_2)$-universal set over $\{0, 1\}$ is a separating hash family of size $O(\log n)$ for fixed d_1 and d_2. In this paper we give a polynomial time construction of an $(M; n, q, \{d_1, d_2\})$ separating hash family of size $M = ((d_1 d_2)^4 \log n / \log q)$ for any $q \geq d_1 d_2(1 + o(1))$ and any n.

We show, for constant $c > 1$ and $q > D_2$, the following $(M; n, q, \{d_1, d_2, \ldots, d_r\})$ separating hash families can be constructed in polynomial time

n	q	poly time Size =	Union Bound	Lower Bound [17]
I.S.	$q \geq cD_2^2$	$D_2 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$
all	$q \geq cD_2^2$	$D_2^2 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$
I.S.	$q \geq cD_2$	$D_2^2 \frac{\log n}{\log d}$	$D_1 \frac{\log n}{\log(q/D_2))}$	$D_1 \frac{\log n}{\log q}$
all	$q \geq cD_2$	$D_2^3 \frac{\log n}{\log d}$	$D_1 \frac{\log n}{\log(q/D_2)}$	$D_1 \frac{\log n}{\log q}$
I.S.	$q \geq D_2 + 1$	$D_2^3 \frac{\log n}{\log q}$	$D_1 \log n$	$D_1 \frac{\log n}{\log q}$
all	$q \geq D_2 + 1$	$D_2^4 \frac{\log n}{\log q}$	$D_1 \log n$	$D_1 \frac{\log n}{\log q}$

We now turn to another application of tester.

1.3 Black Box PIT

The second application of tester is getting asymptotically optimal black box PIT sets for classes of multivariate polynomials over finite fields with substitutions from an extension field of optimal dimension.

The black box Polynomial Identity Testing (PIT) problem is the following: Given an arithmetic circuit C that either identical to the zero function or from a class of circuits $\mathcal{C}$ over a field $\mathbb{F}$, with input variables $x_1, x_2, \ldots, x_n$ and given a *substitution oracle* that for an input $\boldsymbol{a} \in \mathbb{F}^n$ returns $f(\boldsymbol{a})$. Determine whether C computes the identically zero polynomial. We say that $S \subset \mathbb{F}^n$ is a *black box PIT set* for $\mathcal{C}$ if for every $f \in \mathcal{C}$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$.

When the field is finite $\mathbb{F}_q$ many simple classes of circuits, such as circuits that compute monomials, require black box PIT sets of exponential size. Therefore, many papers in the literature allow the substitution oracle to receive assignment $\boldsymbol{a}$ from some extension field $\mathbb{F}_{q^t}$ of $\mathbb{F}_q$. We say that $S \subset \mathbb{F}_{q^t}^n$ is a *black box PIT set over* $\mathbb{F}_{q^t}$ for $\mathcal{C}$ if for every $f \in \mathcal{C}$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$. Our goal will be to minimize the size of the black box PIT set $|S|$ and minimize the dimension of the extension field t.

For this problem we use tester in the following way. First we construct an optimal black box PIT set

S for the class over a large field $\mathbb{F}_{q^T}$. We then use a tester to map the assignments in S to assignments in a smaller field $\mathbb{F}_{q^t}$.

Asymptotically Optimal Black Box PIT Sets for Polynomials: We study the following classes of multivariate polynomials: $\mathcal{P}(\mathbb{F}_q, n)$ is the class of all multivariate polynomials in $\mathbb{F}_q[x_1, \ldots, x_n]$ of variable degree (maximal degree of the variables) at most $q-1$, $\mathcal{P}(\mathbb{F}_q, n, (d, r))$ is the class of all multivariate polynomials in $\mathcal{P}(\mathbb{F}_q, n)$ of degree at most d and variable degree at most $r \leq q-1$, $\mathcal{P}(\mathbb{F}_q, n, d) = \mathcal{P}(\mathbb{F}_q, n, (d, q-1))$, $\mathcal{P}(\mathbb{F}_q, n, s)$ is the class of all multivariate polynomials in $\mathcal{P}(\mathbb{F}_q, n)$ with at most s monomials. Polynomials in the latter class are called in the literature "sparse multivariate polynomials". $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ is the class of all multivariate polynomials in $\mathcal{P}(\mathbb{F}_q, n)$ of degree at most d, variable degree at most r and with at most s monomials.

To the best of our knowledge all the algorithms in the literature that construct black box PIT sets for the above classes are either randomized, deterministic for some fixed extension field $\mathbb{F}_{q^t}$ or obtains non-optimal results in both the extension field dimension and the size of the black box PIT set [34, 24, 80, 39, 45, 9]. In this paper we give polynomial time constructions of black box PIT sets of size that are within $poly(n)$ of the optimal size, Also, the dimensions of the extension fields are optimal.

The following table summarizes some of our main results. See other results in Section 5.

Class	Extension Field $\mathbb{F}_{q^t}$	Lower Bound	Upper Bound	Explicit Construction	Poly Time Construction
$\mathcal{P}(\mathbb{F}_q, n)$	$t \geq \log_q n + 2$	$\frac{q^n}{t}$	$\frac{(\log(qn))\cdot q^n}{t}$	$\frac{n\cdot q^n}{t}$	$\frac{n\cdot q^n}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r)$	$t \geq \log_q n + 2$	$\frac{(r+1)^n}{t}$	$\frac{(\log(rn))\cdot(r+1)^n}{t}$	$\frac{n\cdot(r+1)^n}{t}$	$\frac{n\cdot(r+1)^n}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r))$	$t \geq \log_q(d+1)$	$\frac{R(n,d,r)}{t}$	$\frac{(\log d)\cdot R(n,d,r)}{t}$	$\frac{dn^{d+1}}{t}$	$\frac{dn^{d+1}}{t}$
$\mathcal{P}(\mathbb{F}_q, n, s)$	$t \geq \log_q n + 2$	$\frac{n\cdot s}{t}$	$\frac{(\log n)n\cdot s}{t}$	$\frac{q^5 n^6\cdot s}{t}$	$\frac{n^7 q^{29}(\log^2 q)\cdot s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r, s)$	$t \geq \log_q n + 2$	$\frac{n\log(r+1)}{\log q}\cdot\frac{s}{t}$	$\frac{n(\log n)\log(r+1)}{\log q}\cdot\frac{s}{t}$	$\frac{n^6 r^5\log(r+1)}{\log q}\cdot\frac{s}{t}$	$\frac{n^7 r^5 q^{24}(\log^2 r)\cdot s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$	$t \geq \log_q(d+1)$	$\frac{d\log n}{\log q}\cdot\frac{s}{t}$	$\frac{d(\log d)\log n}{\log q}\cdot\frac{s}{t}$	$\frac{d^6\log n}{\log q}\cdot\frac{s}{t}$	$\frac{q^{24}d^7(\log^2 n)\cdot s}{t}$

Notice that our polynomial time constructions (column 6 in the above table) are optimal in the largest parameters (q^n, $(r+1)^n$, n^d and s in the last three rows of the table). For sparse polynomials $\mathcal{P}(\mathbb{F}_q, n, s)$, $\mathcal{P}(\mathbb{F}_q, n, r, s)$ and $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$, all the results in the literature give black box PIT sets of size that are at least quadratic in the size s. Also, the tight tradeoff with the field dimension $(\cdot/t)$ in each row of the table was not known before. The bound on the dimension of the field extension in the second column is $\log_q(degree) + 1$ and is known to be the best possible dimension (even for randomized algorithms) if one uses Schwartz-Zippel Lemma. Therefore, our constructions are tight in the dimension of the field extension. For the class $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ (the last row of the table), the best known result in the literature [45, 9] used field extension of dimension that depends on the number of the variables n. Our result uses field extension of dimension $\log_q(d+1)$ that is independent of the number of variables. Also notice, that when $q \geq d+1$ no extension field is needed. This will be further studied in [12] to give new

Pseudorandom generators over small fields.

In this paper we also develop a new bound that beats the upper bound of Schwartz-Zippel Lemma (See Lemma 70). This new bound gives a randomized algorithm for polynomial size black box PIT set even over field extension of dimension 2 (See the Table in Lemma 73). Our new bound also significantly improves the upper bounds when the size of the sparse multivariate polynomial is small.

Reduction of Black Box PIT Sets over Large Field to Sets over Small Field: We give several polynomial time reductions of black box PIT sets over large fields to black box PIT sets over small fields. For example, consider a subclass $\mathcal{C} \subseteq \mathcal{P}(\mathbb{F}_q, n, d)$ and t such that $q^t \geq d+1$. We give a polynomial time algorithm that takes a black box PIT set S of size w for $\mathcal{C}$ over an extended field $\mathbb{F}_{q^T}$ and constructs a black box PIT set for $\mathcal{C}$ over $\mathbb{F}_{q^t}$ of size $O(d^5 \cdot wT/t)$.

In particular, we apply this result to the results in [70, 73, 74, 6] and get black box PIT sets over smaller fields as indicated in the following table

Circuit Class	Field size $\geq$	Size of PIT set	New Field size $\geq$	New Size of PIT set	Ref.
$\Sigma\Pi\Sigma(k,d,n)$	dnk^2	$poly(n) \cdot d^k$	$d+1$	$poly(n) \cdot d^{k+5}$	[70]
$\Sigma_r \mathrm{P}_k\mathrm{ROF}$	kn^3	$(kn)^{O(r+\log n)}$	$kn+1$	$(kn)^{O(r+\log n)}$	[73]
ML $\Sigma\Pi\Sigma\Pi(k)$	n^2	$n^{O(k^3)}$	$n+1$	$n^{O(k^3)}$	[74]
R_k ML	n^2	$n^{k^{O(k)}+O(k\log n)}$	$n+1$	$n^{k^{O(k)}+O(k\log n)}$	[6]

The classes in the table are: $\Sigma\Pi\Sigma(k,d,n)$ is the class of depth-3 circuits with n variables, degree d and top fanin k. P_kROF is the class of read once formulas (ROF, each variable appears at most once in the formula) with n variables in which we are allowed to replace each variable x_i with a univariate polynomial $T_i(x_i)$ of degree at most k. $\Sigma_r\mathrm{P}_k$ROF is the sum of r P_kROF formulas. Multilinear (ML) $\Sigma\Pi\Sigma\Pi(k)$ is the class of depth 4 circuits with n variables in which the fan-in of the top Σ gate is a constant k and each multiplication gate computes a multilinear polynomial. R_k ML is the class of multilinear formulas with n variables where each variable appears at most k times in the formula.

The table shows the reduction to a smaller field. For example, in the first row in the table, in [70], Saxena and Seshadhri gave a black box PIT set for $\Sigma\Pi\Sigma(k,d,n)$ over fields of size at least dnk^2. We apply our reduction to give a black box PIT set for $\Sigma\Pi\Sigma(k,d,n)$ over fields of size at least $d+1$. Notice that our field size is independent of n and when the field $\mathbb{F}_q$ satisfies $q \geq d+1$, no extension field is needed.

1.4 Polynomial Restriction Problems

We now give another application of tester. Our new problem is called *polynomial restriction problem* and is a generalization of the d-restriction problem. In the d-restriction problems all the functions

that determine the restrictions on the combinatorial structure depend on the same d variables. In the polynomial restriction problem the functions that determine the restrictions may depend on all the variables but are multivariate polynomials with some bounded parameters (such as degree, variable degree and size). Obviously, d-restriction problems are also polynomial restriction problems with polynomials of degree d. We show that also for those problems several polynomial time constructions are almost optimal. We now give a formal definition

Restriction Problem: A restriction problem is a problem of the following form: Given an alphabet Σ of size $|\Sigma| = q$, an integer n and a class $\mathcal{M}$ of nonzero functions $f : \Sigma^n \to \{0, 1\}$. Find a small set $A \subseteq \Sigma^n$ such that: For every $f \in \mathcal{M}$ there is $\boldsymbol{a} \in A$ such that $f(\boldsymbol{a}) \neq 0$.

We will study restriction problems when $\mathcal{M}$ is a class of multivariate polynomials over $\mathbb{F}_q$.

(s, d)-Sparse (d, r)-Degree Polynomial Restriction Problem: Let $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ be the class of all multivariate polynomials in $\mathbb{F}_q[x_1, \ldots, x_n]$ of degree d and variable degree at most r with at most s monomials of degree d and any number of monomials of degree less than d. We denote $\mathcal{P}(\mathbb{F}_q, n, ((d, q-1), s))$ by $\mathcal{P}(\mathbb{F}_q, n, (d, s))$.

The *(s, d)-sparse (d, r)-degree polynomial problem* over $\mathbb{F}_q$ problem is the following: Given the class $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$. Find a small set $S \subset \mathbb{F}_q^n$ such that for every $f \in \mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$. This problem can be regarded as black box PIT over $\mathbb{F}_q$, hitting set problem or polynomial restriction problem. We will call S a *hitting set* for $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$. We note that this class is not studied in the literature. Also notice that if $f(\boldsymbol{x}) \in \mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ then for any vector of constants $\boldsymbol{\alpha} \in \mathbb{F}_q^n$ we have $f(\boldsymbol{x} + \boldsymbol{\alpha}) \in \mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$. The class $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ is much larger than $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ since we allow any number of monomials of degree less than d in f and therefore the union bound does not give good bounds for this class. We develop new techniques that find lower and upper bounds for the size of a hitting set for this class and apply tester to construct almost optimal size hitting set.

The following table summarizes the results for the size of the hitting sets for $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ when $r \leq p - 1$ where p is the characteristic of the field.

$\mathbb{F}_q, d$	Lower Bound	Upper Bound	Explicit Construction	Poly Time Construction
$q = 2$	$2^d s \log n$	$2^{2d} s \log n$	$2^{2.66d} s \log n$	$2^{2.66d} s \log^2 n$
$q = 2^\ell \leq d$	$\pi_{q,r}^d s \frac{\log n}{\log q}$	$d(2\pi_{q,1})^d s \log n$	$2^{(1+c_q)d} s \log n$	$2^{(1+c_q)d} s \log^2 n$
$q \leq d$	$\pi_{q,r}^d s \frac{\log n}{\log q}$	$d(2\pi_{q,1})^d s \log n$	$2^{(1+c_q)d} d! s \log n$	$2^{(1+c_q)d} d! s \log^2 n$
$q \leq d = O(1)$	$s \log n$	$s \log n$	$s \log n$	$s \log n$
$q \geq d + 1$	$ds \frac{\log n}{\log q}$	$d^2 s \frac{\log n}{\log(q/d)}$	$d^6 s \frac{\log n}{\log q}$	$q^{24} d^7 s \cdot \log^2 n$
$q \geq d + 1 = O(1)$	$s \frac{\log n}{\log q}$	$s \frac{\log n}{\log(q/d)}$	$s \frac{\log n}{\log q}$	$s \frac{\log n}{\log q}$

In the table, the constant $\pi_{q,r}$ satisfies

$$1+\frac{1}{q-1} \leq \pi_{q,r} := \left(\frac{q}{q-r}\right)^{1/r} \leq q^{1/(q-1)} = 1+\frac{\ln q}{q-1}+O\left(\frac{\log^2 q}{q^2}\right) \leq 2.$$

and the constant c_q satisfies

$$c_q = \sum_{i=0}^{\infty} \frac{\log(q^{2^i}+1)}{q^{2^i}} = \frac{\log(q+1)}{q} + O\left(\frac{\log q}{q^2}\right) < 1.66.$$

In the table all the bounds are tight in the parameter s, the number of monomials of degree d. The explicit constructions (see footnote in page 6) are also tight in $\log n$ where in the polynomial time constructions we get $(\log n)^2$. This is because our algorithm requires an element of the field of multiplicative order $n^{O(d)}$. When $d = O(1)$ (rows 4 and 6 in the table) n^d is polynomial and finding such element can be done in polynomial time.

1.5 Organization of this Paper

This paper is organized as follows. In Section 2 we study testers. In Subsection 2.1 we give the definition of tester and the classes of polynomials that will be studied here. In Subsection 2.2 we use elementary algebra to give some basic properties of testers. In Subsection 2.3 we use algebraic function fields to get the first non-trivial tester that reduces $\mathbb{F}_{q^t}$ to $\mathbb{F}_q$ for polynomials of degree $d \leq q-1$. In Subsection 2.4 we give testers that reduce some subspace $S \subset \mathbb{F}_{q^t}$ to $\mathbb{F}_q$. Those testers will have small size. In Subsection 2.5 we study testers that reduce $\mathbb{F}_{q^t}$ to $\mathbb{F}_q$ for a subclass of polynomials of degree d with no restriction on d. In Subsection 2.6 we define symmetric and reducible testers. Those will be used to construct testers in polynomial time. In Subsection 2.7 we give lower bounds for the size of testers. Those bounds will show that our testers are almost asymptotically optimal.

In Section 3 we show that there is a polynomial time algorithm that constructs almost optimal testers that reduce $\mathbb{F}_{q^t}$ to $\mathbb{F}_q$. The way we prove that is as follows. In Subsection 3.3 we show how to reduce the dimension of the field t to $O(\log\log t)$, and then in Subsection 3.4 we show how to exhaustively search for a symmetric tester for the smaller field in polynomial time. In Appendix B we discuss another possible approach for building testers of smaller size. In Subsection 3.5 we give a polynomial time construction for testers from a subspace $S \subset \mathbb{F}_{q^t}$ to $\mathbb{F}_q$. Those have size smaller than the above testers and can be used to get better constructions for some problems. In Subsection 3.6 we show how to construct a tester for a subclass of polynomial of degree d with no restriction on d. Then in Subsection 3.7 we show how to construct a tester from $\mathbb{F}_{q^t}$ to $\mathbb{F}_{q^{t'}}$ for any $t' \leq t$.

In Section 4 we show how to apply tester to d-restriction problems. We study perfect hash in Subsection 4.1, universal sets in Subsection 4.2, cover-free families in Subsection 4.3 and seperating hash family in Subsection 4.4.

In Section 5 we show how to apply tester to black box PIT. In Subsections 5.1–5.2 we define the polynomial classes we study and summarize the main results of this section. In Subsection 5.3 we give some preliminary results. In Subsection 5.4 we show how to reduce a black box PIT set over a large field to a black box PIT set over a smaller field. In Subsection 5.5 we prove new lower and upper bounds for the size of black box PIT sets for several polynomial classes. Then in Subsections 5.6–5.9 we give the polynomial time constructions for the classes. In Subsection 5.10 we consider other black box PIT sets of classes of circuits known from the literature and apply the reduction to them.

In Section 6 we show how to apply tester to polynomial restriction problems. In Subsection 6.1 we give a lower bound for this problem and in Subsection 6.2 we give a nonconstructive upper bound. Then in Subsection 6.3 we give the polynomial time construction. We then finish the paper in Section 7 where we discuss the results and future work.

2 Testers

In this section we define testers and construct explicit testers that will be used in the sequel.

2.1 Definition of Tester

Definition 1. *Let $\mathbb{F}$ be a field and $\mathcal{A}$ and $\mathcal{B}$ be two commutative $\mathbb{F}$-algebras. Let $\mathcal{M} \subseteq \mathbb{F}[x_1, x_2, \ldots, x_n]$ be a class of multivariate polynomials. Let $S \subseteq \mathcal{A}$ and $R \subseteq \mathcal{B}$ be sub-linear spaces over $\mathbb{F}$ and $L = \{\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_\nu\}$ be a set of maps $S^n \to R^n$. We denote by fL the map $S^n \to \mathcal{B}^\nu$ where for $\boldsymbol{a} \in S^n$, $(fL)(\boldsymbol{a}) = (f(\boldsymbol{\ell}_1(\boldsymbol{a})), \ldots, f(\boldsymbol{\ell}_\nu(\boldsymbol{a})))$. We say that L is an $(\mathcal{M}, S, R)$-tester if for every $\boldsymbol{a} \in S^n$ and $f \in \mathcal{M}$ we have (Here $\mathbf{0} = 0^\nu$ is the zero vector of length ν)*

$$(fL)(\boldsymbol{a}) = \mathbf{0} \implies f(\boldsymbol{a}) = 0.$$

The integer $\nu = |L|$ is called the *size of the tester*. The minimal size of an $(\mathcal{M}, S, R)$-tester is denoted by $\nu_R^\circ(\mathcal{M}, S)$. If no L exists then we write $\nu_R^\circ(\mathcal{M}, S) = \infty$. When S and R are known from the context we then just say *tester for* $\mathcal{M}$.

We will also allow $L = \{\ell_1, \ldots, \ell_\nu\}$ to be a set of maps $S \to R$. In that case $\boldsymbol{\ell}_i : S^n \to R^n$ is defined as $\boldsymbol{\ell}_i(\boldsymbol{a}) = (\ell_i(a_1), \ldots, \ell_i(a_n))$ where $\boldsymbol{a} = (a_1, \ldots, a_n) \in S^n$. In such case we call the tester *symmetric tester*.

We say that the tester is *componentwise tester* if $\boldsymbol{\ell}_i(\boldsymbol{a}) = (\ell_{i,1}(a_1), \ldots, \ell_{i,n}(a_n))$ for some $\ell_{i,j} : S \to R$. All the testers that will be considered in this paper are componentwise testers. A componentwise tester is called *reducible* if $\ell_{i,j}(1_\mathcal{A}) = 1_\mathcal{B}$ where $1_\mathcal{A}$ and $1_\mathcal{B}$ are the identities of the algebras $\mathcal{A}$ and $\mathcal{B}$ respectively.

Let $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)$ and $\boldsymbol{y}_i = (y_{i,1}, \ldots, y_{i,n})$ be indeterminates over $\mathbb{F}$ for $i = 1, \ldots, d$. A multilinear polynomial in $\mathbb{F}[\boldsymbol{y}]$ is called (n, d)-*multilinear polynomial* if each monomial of f contains exactly one

variable from each $\boldsymbol{y}_i$, $i = 1, \ldots, d$. Let $\mathcal{DML}(\mathbb{F}, n, d)$ be the class of all (n, d)-multilinear polynomials in $\mathbb{F}[\boldsymbol{y}]$. Notice that $\mathcal{DML}(\mathbb{F}, n, 2)$ is the class of all bilinear forms $\boldsymbol{y}_1^T A \boldsymbol{y}_2$ where $A \in \mathbb{F}^{n \times n}$. We denote $\nu_R(d, S) = \nu_R^\circ(\mathcal{DML}(\mathbb{F}, n, d), S)$. Let $\mathcal{P}(\mathbb{F}, n, d)$ be the class of all multivariate polynomials in $\mathbb{F}[x_1, \cdots, x_n]$ of degree d and $\mathcal{HP}(\mathbb{F}, n, d)$ be the class of all homogeneous polynomials of degree d. We denote $\nu_R^{\mathcal{P}}(d, S) = \nu_R^\circ(\mathcal{P}(\mathbb{F}, n, d), S)$ and $\nu_R^{\mathcal{HP}}(d, S) = \nu_R^\circ(\mathcal{HP}(\mathbb{F}, n, d), S)$. Then

$$\nu_R(d, S) \leq \nu_R^{\mathcal{HP}}(d, S) \leq \nu_R^{\mathcal{P}}(d, S). \tag{1}$$

Note that we omit the parameter n in the definition of ν. We can do that since the bounds we find in this paper are independent of n. When $\mathbb{F}$ is not clear from the context then we write: $\nu_R((d, \mathbb{F}), S)$, $\nu_R^{\mathcal{P}}((d, \mathbb{F}), S)$ and $\nu_R^{\mathcal{HP}}((d, \mathbb{F}), S)$.

In the following subsections we will study bounds for $\nu_{\mathbb{F}_q}$, $\nu_{\mathbb{F}_q}^{\mathcal{P}}$ and $\nu_{\mathbb{F}_q}^{\mathcal{HP}}$ where $\mathbb{F}_q$ is the finite field with q elements.

2.2 Preliminary Results for Testers

In this subsection we prove some elementary results on testers.

The following three Lemmas immediately follows from the definition of tester

Lemma 2. *Let $\mathcal{A}$ and $\mathcal{B}$ be commutative $\mathbb{F}$-algebras and $C \subseteq S \subseteq \mathcal{A}$ and $R \subset \mathcal{B}$ be subspaces over $\mathbb{F}$. Let $\mathcal{M} \subseteq \mathbb{F}[x_1, \ldots, x_n]$. If L is a $(\mathcal{M}, S, R)$-tester then L is a $(\mathcal{M}, C, R)$-tester. In particular, $\nu_R^\circ(\mathcal{M}, C) \leq \nu_R^\circ(\mathcal{M}, S)$.*

Lemma 3. *Let $\mathcal{A}$ and $\mathcal{B}$ be commutative $\mathbb{F}$-algebras and $S \subseteq \mathcal{A}$ and $R \subseteq \mathcal{B}$ be subspaces over $\mathbb{F}$. Let $\mathcal{M} \subset \mathcal{N} \subseteq \mathbb{F}[x_1, \ldots, x_n]$. If L is a $(\mathcal{N}, S, R)$-tester then L is a $(\mathcal{M}, S, R)$-tester. In particular, $\nu_R^\circ(\mathcal{M}, S) \leq \nu_R^\circ(\mathcal{N}, S)$.*

Lemma 4. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be commutative $\mathbb{F}$-algebras and $\mathcal{M} \subseteq \mathbb{F}[x_1, \ldots, x_n]$. Let $S_1 \subseteq \mathcal{A}$, $S_2 \subseteq \mathcal{B}$ and $S_3 \subseteq \mathcal{C}$ be subspaces over $\mathbb{F}$. If L_1 is a $(\mathcal{M}, S_1, S_2)$-tester and L_2 is a $(\mathcal{M}, S_2, S_3)$-tester then $L_2 \circ L_1 = \{\boldsymbol{\ell}_2(\boldsymbol{\ell}_1) \mid \boldsymbol{\ell}_1 \in L_1, \boldsymbol{\ell}_2 \in L_2\}$ is $(\mathcal{M}, S_1, S_3)$-tester. In particular,*

$$\nu_{S_3}^\circ(\mathcal{M}, S_1) \leq \nu_{S_3}^\circ(\mathcal{M}, S_2) \cdot \nu_{S_2}^\circ(\mathcal{M}, S_1).$$

In particular we have

Corollary 5. *Let $\mathbb{K}$ be an extension field of $\mathbb{F}$, $\mathcal{A}$ be a $\mathbb{K}$-algebra and $S \subseteq \mathcal{A}$ be a sublinear space. Let $\mathcal{M} \subseteq \mathbb{F}[x_1, \ldots, x_n]$. Then*

$$\nu_{\mathbb{F}}^\circ(\mathcal{M}, S) \leq \nu_{\mathbb{F}}^\circ(\mathcal{M}, \mathbb{K}) \cdot \nu_{\mathbb{K}}^\circ(\mathcal{M}, S).$$

In particular, for any integers t and m we have

$$\nu_{\mathbb{F}_q}^\circ(\mathcal{M}, \mathbb{F}_{q^{tm}}) \leq \nu_{\mathbb{F}_q}^\circ(\mathcal{M}, \mathbb{F}_{q^t}) \cdot \nu_{\mathbb{F}_{q^t}}^\circ(\mathcal{M}, \mathbb{F}_{q^{tm}})$$

and

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^{tm}}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}(d, \mathbb{F}_{q^{tm}}).$$

We must note here that $\nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}(d, \mathbb{F}_{q^{tm}})$ may be understood as $\nu^{\circ}_{\mathbb{F}_{q^t}}(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^{tm}}) = \nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}((d, \mathbb{F}_q), \mathbb{F}_{q^{tm}})$, which is what we meant in Corollary 5, or as $\nu^{\circ}_{\mathbb{F}_{q^t}}(\mathcal{P}(\mathbb{F}_{q^t}, n, d), \mathbb{F}_{q^{tm}}) = \nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}((d, \mathbb{F}_{q^t}), \mathbb{F}_{q^{tm}})$. The inequality is true for both since, by Lemma 3

$$\nu^{\circ}_{\mathbb{F}_{q^t}}(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^{tm}}) \leq \nu^{\circ}_{\mathbb{F}_{q^t}}(\mathcal{P}(\mathbb{F}_{q^t}, n, d), \mathbb{F}_{q^{tm}}).$$

We now prove

Lemma 6. *Let*

$$\mathcal{M} \subseteq \mathbb{F}[\boldsymbol{x}]\mathbb{F}[\boldsymbol{y}] := \left\{ \sum_{i=1}^{s} h_i(\boldsymbol{x}) g_i(\boldsymbol{y}) \;\middle|\; h_i \in \mathbb{F}[\boldsymbol{x}], g_i \in \mathbb{F}[\boldsymbol{y}], s \in \mathbb{N} \right\}$$

be a class of multivariate polynomials where $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_m)$ are indeterminates. Let $\mathcal{A}$ be a commutative $\mathbb{F}$-algebra, $S \subseteq \mathcal{A}$ sublinear space,

$$\mathcal{M}_x = \left\{ \sum_{i=1}^{s} \lambda_i h_i(\boldsymbol{x}) \;\middle|\; \sum_{i=1}^{s} h_i(\boldsymbol{x}) g_i(\boldsymbol{y}) \in \mathcal{M}, \boldsymbol{\lambda} \in \mathbb{F}^s, s \in \mathbb{N} \right\}$$

and

$$\mathcal{M}_y = \left\{ \sum_{i=1}^{s} \lambda_i g_i(\boldsymbol{y}) \;\middle|\; \sum_{i=1}^{s} h_i(\boldsymbol{x}) g_i(\boldsymbol{y}) \in \mathcal{M}, \boldsymbol{\lambda} \in \mathbb{F}^s, s \in \mathbb{N} \right\}.$$

If L_x is an $(\mathcal{M}_x, S, \mathbb{F})$-tester and L_y is an $(\mathcal{M}_y, S, \mathbb{F})$-tester then $L_x \times L_y$ is a $(\mathcal{M}, S, \mathbb{F})$-tester. That is,

$$\nu^{\circ}_{\mathbb{F}}(\mathcal{M}, S) \leq \nu^{\circ}_{\mathbb{F}}(\mathcal{M}_x, S) \cdot \nu^{\circ}_{\mathbb{F}}(\mathcal{M}_y, S).$$

In particular, for any d_1 and d_2

$$\nu_{\mathbb{F}}(d_1 + d_2, S) \leq \nu_{\mathbb{F}}(d_1, S) \cdot \nu_{\mathbb{F}}(d_2, S).$$

Proof. Suppose $f(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{s} h_i(\boldsymbol{x}) g_i(\boldsymbol{y}) \in \mathcal{M}$, $(\boldsymbol{a}, \boldsymbol{b}) \in S^{n+m}$ and $(f(L_x \times L_y))(\boldsymbol{a}, \boldsymbol{b}) = 0$. Then for every $\boldsymbol{\ell}_x \in L_x$ and $\boldsymbol{\ell}_y \in L_y$ we have $f(\boldsymbol{\ell}_x(\boldsymbol{a}), \boldsymbol{\ell}_y(\boldsymbol{b})) = 0$. Since $f(\boldsymbol{\ell}_x(\boldsymbol{a}), \boldsymbol{y}) \in \mathcal{M}_y$ and L_y is an $(\mathcal{M}_y, S, \mathbb{F})$-tester we have: for every $\boldsymbol{\ell}_x \in L_x$, $f(\boldsymbol{\ell}_x(\boldsymbol{a}), \boldsymbol{b}) = 0$. Let ℓ be any linear transformation in $\mathcal{A}^*$. Then for every $\boldsymbol{\ell}_x \in L_x$

$$\sum_{i=1}^{s} h_i(\boldsymbol{\ell}_x(\boldsymbol{a})) \ell(g_i(\boldsymbol{b})) = \ell(f(\boldsymbol{\ell}_x(\boldsymbol{a}), \boldsymbol{b})) = 0.$$

Since $\sum_{i=1}^{s} h_i(\boldsymbol{x})\ell(g_i(\boldsymbol{b})) \in \mathcal{M}_x$ and L_x is an $(\mathcal{M}_x, S, \mathbb{F})$-tester we have: $\sum_{i=1}^{s} h_i(\boldsymbol{a})\ell(g_i(\boldsymbol{b})) = 0$ for every linear transformation $\ell \in \mathcal{A}^*$. Now let $\{\omega_1, \ldots, \omega_r\} \subset \mathcal{A}$ be a basis for Span $_{\mathbb{F}}\{g_1(\boldsymbol{b}), \ldots, g_s(\boldsymbol{b})\}$ and let ℓ_{ω_i}, $i = 1, 2, \ldots, r$ be linear transformations in $\mathcal{A}^*$ such that $g_i(\boldsymbol{b}) = \sum_{j=1}^{r} \ell_{\omega_j}(g_i(\boldsymbol{b}))\omega_j$. Then

$$\begin{aligned} f(\boldsymbol{a}, \boldsymbol{b}) &= \sum_{i=1}^{s} h_i(\boldsymbol{a}) g_i(\boldsymbol{b}) \\ &= \sum_{i=1}^{s} h_i(\boldsymbol{a}) \sum_{j=1}^{r} \ell_{\omega_j}(g_i(\boldsymbol{b}))\omega_j \\ &= \sum_{j=1}^{r} \omega_j \sum_{i=1}^{s} h_i(\boldsymbol{a})\ell_{\omega_j}(g_i(\boldsymbol{b})) = 0. \end{aligned}$$

□

We now prove two lemmas for $\nu_{\mathbb{F}_q}^{\mathcal{P}}$ and $\nu_{\mathbb{F}_q}^{\mathcal{HP}}$. Let $\mathbb{F}_q[X]_w$ be the linear space of polynomials in $\mathbb{F}_q[X]$ of degree at most w.

Lemma 7. *We have*

1. *If $q \geq d(t-1)$ then*
$$\nu_{\mathbb{F}_q}^{\mathcal{HP}}(d, \mathbb{F}_{q^t}) \leq \nu_{\mathbb{F}_q}^{\mathcal{HP}}(d, \mathbb{F}_q[X]_{t-1}) \leq d(t-1) + 1.$$
2. *If $q \geq d(t-1) + 1$ then*
$$\nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^t}) \leq \nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_q[X]_{t-1}) \leq d(t-1) + 1.$$

Proof. Let $\mathbb{F}_{q^t} = \mathbb{F}_q[\alpha]/(g(\alpha))$ for some irreducible polynomial $g(\alpha) \in \mathbb{F}_q[\alpha]$ of degree t. Every element in $\mathbb{F}_{q^t}$ can be represented as $\omega_0 + \omega_1\alpha + \cdots + \omega_{t-1}\alpha^{t-1}$ where $\omega_i \in \mathbb{F}_q$ for $i = 0, 1, \ldots, t-1$. We first define the map $\ell(\omega_0 + \omega_1\alpha + \cdots + \omega_{t-1}\alpha^{t-1}) = \omega_0 + \omega_1 X + \cdots + \omega_{t-1}X^{t-1}$ where X is indeterminates over $\mathbb{F}_q$. Obviously, for any $f \in \mathbb{F}_q[x_1, \cdots, x_n]$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{F}_{q^t}^n$, if $f(\ell(\beta_1), \ldots, \ell(\beta_n)) = 0$ then $f(\beta_1, \ldots, \beta_n) = 0$. This gives a $(\mathbb{F}_q[x_1, \cdots, x_n], \mathbb{F}_{q^t}, \mathbb{F}_q[X]_{t-1})$-tester of size 1.

For every $f \in \mathcal{P}(\mathbb{F}_q, n, d)$ and $z_1, \ldots, z_n \in \mathbb{F}_q[X]_{t-1}$ we have $f(z_1, \ldots, z_n) \in \mathbb{F}_q[X]_{d(t-1)}$. Therefore, if $q \geq d(t-1)+1$, we can choose $F \subseteq \mathbb{F}_q$ of size $d(t-1)+1$ and if $f(z_1, \ldots, z_n)|_{X=\beta} = 0$ for all $\beta \in F$ then $f(z_1, \ldots, z_n) = 0$. Therefore, $L = \{\ell_\beta \mid \beta \in F\}$ where $\ell_\beta : \mathbb{F}_q[X]_{t-1} \to \mathbb{F}_q$ is defined as $\ell_\beta(z(X)) = z(\beta)$ is a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_q[X]_{t-1}, \mathbb{F}_q)$-tester of size $d(t-1)+1$.

Combining both testers, by Lemma 3 and 4, we get a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $d(t-1)+1$. Therefore, for $q \geq d(t-1)+1$,

$$\nu_{\mathbb{F}_q}^{\mathcal{HP}}(d, \mathbb{F}_{q^t}) \leq \nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^t}) \leq \nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_q[X]_{t-1}) \leq d(t-1)+1.$$

This proves *1* and *2* for $q \geq d(t-1)+1$.

For *1*, when $q = d(t-1)$, consider $f \in \mathcal{HP}(\mathbb{F}_q, n, d)$ and $z_1, \ldots, z_n \in \mathbb{F}_q[X]_{t-1}$. Let $F = \mathbb{F}_q \cup \{\infty\}$ and define for $z \in \mathbb{F}_q[X]_{t-1}$, $\ell_\beta(z) = z(\beta)$ if $\beta \in \mathbb{F}_q$ and $\ell_\infty(z)$ to be the coefficient of X^{t-1} in z. Let $L = \{\ell_\beta \mid \beta \in \mathbb{F}_q \cup \{\infty\}\}$. It is easy to see that the coefficient of $X^{d(t-1)}$ in $f(z_1, \ldots, z_n)$ is $f(\ell_\infty(z_1), \ldots, \ell_\infty(z_n))$. Now if $f(\ell_\infty(z_1), \ldots, \ell_\infty(z_n)) = 0$ then $f(z_1, \ldots, z_n)$ is of degree $d(t-1)-1$ and then if we also have $f(\ell_\beta(z_1), \ldots, \ell_\beta(z_n)) = f(z_1, \ldots, z_n)|_{X=\beta} = 0$ for all $\beta \in \mathbb{F}_q$ then $f(z_1, \ldots, z_n) \equiv 0$. Therefore $L = \{\ell_\beta \mid \beta \in F\}$ is a $(\mathcal{HP}(\mathbb{F}_q, n, d), \mathbb{F}_q[X]_{t-1}, \mathbb{F}_q)$-tester of size $d(t-1)+1$. □

The next result in this subsection shows how to reduce testers for degree d polynomials in $\mathbb{F}_{q^t}$ to testers in $\mathbb{F}_{q^{\log(dt)/\log q}}$. This will be used to construct testers that are almost (within $poly(d)$) optimal in polynomial time.

For any positive integer k, let $N_q(k)$ denotes the number of monic irreducible polynomial of degree k over $\mathbb{F}_q$ [47]. We first prove the following

Lemma 8. *For any finite field $\mathbb{F}_q$ and any integers r and t such that $q^{r-1} \geq dt - d + 1$, we have*

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \left(\frac{dt-d+1}{r} + 1\right) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^r}).$$

Proof. By Lemma 7 and Lemma 4 we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_q[X]_{t-1}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_{q^r}}((d, \mathbb{F}_q), \mathbb{F}_q[X]_{t-1}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^r}). \tag{2}$$

Now it is enough to prove the following lemma. □

Lemma 9. *For any finite field $\mathbb{F}_q$ and any integers r and t such that $q^{r-1} \geq dt - d + 1$, we have*

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^r}}((d, \mathbb{F}_q), \mathbb{F}_q[X]_{t-1}) \leq \frac{dt-d+1}{r} + 1. \tag{3}$$

Proof. It is known that $q^{r-1} < rN_q(r) \leq q^r$, [47]. Let $\mathcal{R}'$ be the set of all monic irreducible polynomials of degree r. Since

$$deg\left(\prod_{p \in \mathcal{R}'} p\right) = rN_q(r) > q^{r-1} \geq dt - d + 1,$$

we can choose $\mathcal{R} \subseteq \mathcal{R}'$ such that

$$dt - d + 1 \leq deg\left(\prod_{p \in \mathcal{R}} p\right) < dt - d + 1 + r.$$

Let $f \in \mathcal{P}(\mathbb{F}_q, n, d)$, $z_1, \ldots, z_n \in \mathbb{F}_q[X]_{t-1}$ and $g(X) = f(z_1, \ldots, z_n) \in \mathbb{F}_q[X]_{dt-d}$. Now $g \equiv 0$ if and only if $g \bmod (\prod_{p \in \mathcal{R}} p) \equiv 0$ if and only if $g \bmod p \equiv 0$ for all $p \in \mathcal{R}$. Now since $g \in \mathbb{F}_q[X]$ then $g \bmod p \equiv 0$ if and only if $g(\beta) = f(z_1(\beta), \ldots, z_n(\beta)) = 0$ for one root $\beta \in \mathbb{F}_{q^r}$ of p. See Theorem 3.33 (ii) in [47].

Therefore,

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^r}}((d, \mathbb{F}_q), \mathbb{F}_q[X]_{t-1}) \leq |\mathcal{R}| \leq \frac{dt - d + 1 + r}{r}.$$

□

We note that a slightly better bound can be proved if $\mathcal{R}'$ is the set of all the irreducible polynomials of degree at most r. Now notice that when $t_1 = t$ and $t_{i+1} = \lceil \log(dt_i)/\log q \rceil + 1$ then $q^{t_{i+1}-1} \geq dt_i - d + 1$, and by Lemma 8 we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^{t_i}}) \leq \left(\frac{dt_i - d + 1}{t_{i+1}} + 1 \right) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}\left(d, \mathbb{F}_{q^{t_{i+1}}}\right) \leq (d+1)\frac{t_i}{t_{i+1}} \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}\left(d, \mathbb{F}_{q^{t_{i+1}}}\right).$$

Therefore, if $q \geq d + 1$ and using Lemma 7 for the last step of the above recurrence, we get

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq d^{\log^* t + O(1)} \cdot t$$

where $\log^* t$ is the minimum integer i in which $t_i \leq 4$ (or any other constant). In the next subsection we prove that for $q \geq d + 1$

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq d^{O(1)} \cdot t.$$

When r divides t then a better bound is proved in the following

Lemma 10. *For any finite field $\mathbb{F}_q$, any integers r and t such that $r|t$ and $q^r \geq d(t/r - 1) + 1$, we have*

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \left(\frac{dt}{r} - d + 1 \right) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}\left(d, \mathbb{F}_{q^r}\right).$$

Proof. By Corollary 5 we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^r}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^r}}((d, \mathbb{F}_q), \mathbb{F}_{(q^r)^{t/r}}).$$

Now by Lemma 7 and 3, if $q^r \geq d(t/r - 1) + 1$ we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^r}}((d, \mathbb{F}_q), \mathbb{F}_{(q^r)^{t/r}}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_{q^r}}(d, \mathbb{F}_{(q^r)^{t/r}}) \leq \frac{dt}{r} - d + 1.$$

□

2.3 Testers for Large Fields

In this subsection we use algebraic function fields to construct explicit testers for large fields. We prove

Theorem 11. *For any $q \geq d+1$ and any t we have*

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq poly(d) \cdot t.$$

In particular, the bound is also true for $\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$ and $\nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$.

We will in fact prove the bound $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^4) \cdot t$. For better bounds when q is large see Corollary 17.

For notations used in this subsection we refer the reader to Appendix A. See also Subsections $1.1 - 1.4$ in [67]. In this subsection we prove Theorem 11 and then in Subsection 3.4 we show how to construct such tester in polynomial time.

We first prove few lemmas using the technique used in [8], Lemma 2.2. (Here we use ℓ for linear maps and l for the dimension of divisors.)

Lemma 12. *Let $F/\mathbb{F}_q$ be a function field, $P_1, \ldots, P_s$ be distinct places of $F/\mathbb{F}_q$ of degree 1 and $D = P_1 + P_2 + \cdots + P_s$. Let G be a divisor of $F/\mathbb{F}_q$ such that $(\text{supp } D) \cap (\text{supp } G) = \emptyset$. Let $L = \{\ell_{P_1}, \ldots, \ell_{P_s}\}$ where $\ell_{P_i} : \mathscr{L}(G) \to \mathbb{F}_q \cup \{\infty\}$ is defined as $\ell_{P_i}(x) = x(P_i)$. Then*

1. *If $s > d \deg(G)$ then any $L' \subseteq L$ where $|L'| = d \deg(G) + 1$ is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathscr{L}(G), \mathbb{F}_q)$-tester.*
2. *If $s > d \deg(G)$ then there is $\bar{L} \subseteq L$ of size $l(dG) \leq d \deg(G) + 1$ such that $\bar{L}$ is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathscr{L}(G), \mathbb{F}_q)$-tester.*
3. *If $s > d \deg(G) \geq 2g - 1$ then there is $\bar{L} \subseteq L$ of size $l(dG) = d \deg(G) - g + 1$ such that $\bar{L}$ is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathscr{L}(G), \mathbb{F}_q)$-tester. In particular,*

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathscr{L}(G)) \leq l(dG) = d \deg(G) - g + 1.$$

Proof. Let $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, d)$ and $\mathcal{M}(\mathscr{L}(G)^n) = \{f(\boldsymbol{x}) \mid f \in \mathcal{M}, \boldsymbol{x} \in \mathscr{L}(G)^n\}$. We first show that ℓ_{P_i} is well defined, i.e., $\ell_{P_i} : \mathscr{L}(G) \to \mathbb{F}_q$. Let $z \in \mathscr{L}(G)$. Then for every i, $v_{P_i}(z) \geq -v_{P_i}(G) = 0$ and therefore by *6* and *7* in Proposition 89, $\ell_{P_i}(z) = z(P_i) \neq \infty$. Also, since P_i is of degree 1 we have $\ell_{P_i}(z) = z(P_i) \in \mathbb{F}_q$.

In the same way the linear function $\ell_{P_i} : \mathscr{L}(dG) \to \mathbb{F}_q$ is well defined. By Proposition 88, for $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathscr{L}(G)^n$ we have $f(\ell_{P_i}(x_1), \ldots, \ell_{P_i}(x_n)) = \ell_{P_i}(f(\boldsymbol{x}))$ and by Proposition 91, $f(\boldsymbol{x}) \in$

$\mathcal{M}(\mathscr{L}(G)^n) \subseteq \mathscr{L}(dG)$. Now if $f(\ell_{P_i}(x_1), \ldots, \ell_{P_i}(x_n)) = 0$ for every $P_i \in L'$ then $\ell_{P_i}(f(\boldsymbol{x})) = 0$ for every $P_i \in L'$ and therefore $f(\boldsymbol{x}) \in \mathrm{Ker}(\ell_{P_i})$ for every $P_i \in L'$. Therefore, it is enough to show that

$$\mathscr{L}(dG) \bigcap \left(\bigcap_{P_i \in L'} \mathrm{Ker}(\ell_{P_i}) \right) = \{0\}.$$

If $z \in \mathscr{L}(dG) \cap \mathrm{Ker}(\ell_{P_i})$ for all $P_i \in L'$ then $\ell_{P_i}(z) = z(P_i) = 0$ and by *7* in Proposition 89, $v_{P_i}(z) \geq 1$. Then, since $(\mathrm{supp}\ D) \cap (\mathrm{supp}\ G) = \emptyset$ we have $z \in \mathscr{L}(dG - D')$ where

$$D' = \sum_{P_i \in L'} P_i.$$

This implies that

$$\mathscr{L}(dG) \cap \bigcap_{P_i \in L'} \mathrm{Ker}(\ell_{P_i}) \subseteq \mathscr{L}(dG - D').$$

Now since $\deg(dG - D') = d \deg G - |L'| = -1 < 0$, by *3* in Proposition 90, we have $\mathscr{L}(dG - D') = \{0\}$. Therefore L' is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathscr{L}(G), \mathbb{F}_q)$-tester.

Obviously, the tester L is symmetric. Since by Proposition 88, $\ell_{P_i}(1) = 1(P_i) = 1$, the tester is also reducible. This completes the proof of *1*.

Now consider the linear map

$$\begin{array}{rcl} T : \mathscr{L}(dG) & \to & \mathbb{F}_q^s \\ f & \mapsto & (f(P_1), f(P_2), \ldots, f(P_s)). \end{array}$$

In *1* we actually have proved that $\ker T = \{0\}$. Therefore, there are $w = \dim \mathscr{L}(dG)$ places $P_{i_1}, \ldots, P_{i_w}$ such that the map

$$\begin{array}{rcl} T' : \mathscr{L}(dG) & \to & \mathbb{F}_q^w \\ f & \mapsto & (f(P_{i_1}), f(P_{i_2}), \ldots, f(P_{i_w})) \end{array}$$

is an isomorphism. This completes the proof of *2*.

Now by Proposition 92, we have $w = \dim \mathscr{L}(dG) = l(dG) = \deg(dG) - g + 1 = d \deg(G) - g + 1$. This proves *3*. □

Lemma 13. *Let $F/\mathbb{F}_q$ be a function field. Let G be a divisor of $F/\mathbb{F}_q$ and Q be a prime divisor of degree $\deg Q = t = l(G)$ such that $v_Q(G) = 0$. If $l(G - Q) = 0$ then the map*

$$\begin{array}{rcl} E : \mathscr{L}(G) & \to & F_Q = \mathbb{F}_{q^t} \\ f & \mapsto & f(Q) \end{array}$$

is isomorphism of vector spaces over $\mathbb{F}_q$ *and* $L = \{E^{-1}\}$ *is a symmetric and reducible* $(\mathbb{F}_q[\boldsymbol{x}], \mathbb{F}_{q^t}, \mathscr{L}(G))$*-tester where* $\boldsymbol{x} = (x_1, \ldots, x_m)$*. In particular,*

$$\nu^{\circ}_{\mathscr{L}(G)}(\mathbb{F}_q[\boldsymbol{x}], \mathbb{F}_{q^t}) = 1.$$

Proof. We first show that the map E is an isomorphism of linear spaces. Let $f \in \mathscr{L}(G)$. Then $v_Q(f) \geq -v_Q(G) = 0$ and therefore by Proposition 89, $f(Q) \neq \infty$ and the map is well defined. Since Q is prime divisor of degree t we have $\dim \mathscr{L}(G) = l(G) = t = [F_Q : \mathbb{F}_q] = \dim F_Q$. Therefore, $F_Q = \mathbb{F}_{q^t}$.

Now we show that $\mathrm{Ker}E = \{0\}$. Let $f \in \mathscr{L}(G)$. If $f \in \mathrm{Ker}E$ then $f(Q) = 0$ and therefore $v_Q(f) > 0$. Since $v_Q(G) = 0$ and $l(G - Q) = 0$ we have $f \in \mathscr{L}(G - Q) = \{0\}$. Therefore E is an isomorphism of vector spaces.

Now suppose $h \in \mathbb{F}_q[\boldsymbol{x}]$ and let $\boldsymbol{E} = (E, E, \ldots, E)$ and $\boldsymbol{E}^{-1} = (E^{-1}, E^{-1}, \ldots, E^{-1})$. Let $\boldsymbol{a} \in \mathbb{F}_{q^t}^m$. If $(hL)(\boldsymbol{a}) = 0$ then $0 = ((hL)(\boldsymbol{a}))(Q) = h(\boldsymbol{E}^{-1}(\boldsymbol{a}))(Q) = h(\boldsymbol{E}(\boldsymbol{E}^{-1}(\boldsymbol{a}))) = h(\boldsymbol{a})$.

Since $E(1) = 1(Q) = 1$ we have $E^{-1}(1) = 1$ and the tester is reducible. □

Ballet in [8] shows that there is a devisor G and a prime devisor Q that satisfy the conditions in Lemma 13. We state the result in the following

Lemma 14. *Let* $F/\mathbb{F}_q$ *be algebraic function field of genus* g *that contains at least* $g+1$ *places of degree* 1*. Let* Q *be a prime divisor of degree* t*. There is a divisor* G *that satisfies the following*

1. $v_Q(G) = 0$
2. $v_P(G) = 0$ *for any prime divisor* P *of degree* 1*.*
3. $l(G) = \deg Q = t$*.*
4. $\deg(G) = t + g - 1$*.*
5. $l(G - Q) = 0$*.*

Proof. The proof follows immediately from Lemma 2.1 and 2.2 in [8]. □

We now use the above three lemmas to prove

Lemma 15. *Let* $F/\mathbb{F}_q$ *be a function field of genus* g *and* $t \geq 3 + 2\log_q(2g+1)$*. If* F *has* $d(t+g-1)+1$ *places of degree* 1 *then*

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq dt + (d-1)(g-1).$$

Proof. First, by Corollary 5.2.10 (c) in [67], if $2g+1 \leq q^{(t-1)/2}(q^{1/2}-1)$ then there is a prime divisor of degree t. Since $t \geq 3 + 2\log_q(2g+1)$ the inequality holds and there is at least one prime divisor of degree t. Let Q be such divisor. Let $P_1, \ldots, P_s$, $s = d(t+g-1)+1$, be distinct places of $F/\mathbb{F}_q$ of degree 1 and $D = P_1 + P_2 + \cdots + P_s$. In Lemma 14 we proved that there is a divisor G of $F/\mathbb{F}_q$ such that $(\text{supp } D) \cap (\text{supp } G) = \emptyset$, $\deg Q = t = l(G)$, $v_Q(G) = 0$, $\deg G = t+g-1$ and $l(G-Q) = 0$.

By Lemmas 13, 12 and Lemma 4 we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathscr{L}(G)) \cdot \nu^{\mathcal{P}}_{\mathscr{L}(G)}(d, \mathbb{F}_{q^t}) \leq d \deg G - g + 1 = dt + (d-1)(g-1).$$

□

We are now ready to give the explicit construction.

A *tower* of function fields over $\mathbb{F}_q$ is a sequence $\mathcal{F} = (F^{(0)}, F^{(1)}, F^{(2)}, \cdots)$ of function fields $F^{(i)}/\mathbb{F}_q$ with $F^{(0)} \subseteq F^{(1)} \subseteq F^{(2)} \subseteq \cdots$ where each extension $F^{(k+1)}/F^{(k)}$ is finite and separable

There are many explicit towers known from the literature. We will use the following $\mathcal{W}_1$ tower defined in [36]. See also [37] Chapter 1.

Lemma 16. *Let x_1 be indeterminate over $\mathbb{F}_{q^2}$ and $F^{(1)} = \mathbb{F}_{q^2}(x_1)$. For $k \geq 1$ let $F^{(k)} = F^{(k-1)}(x_k)$ where*

$$x_k^q + x_k = \frac{x_{k-1}^q}{x_{k-1}^{q-1} + 1}.$$

Let g_k be the genus of $F^{(k)}/\mathbb{F}_{q^2}$ and N_k the number of places in $F^{(k)}/\mathbb{F}_{q^2}$ of degree 1. Then

$$g_k = \begin{cases} q^k - 2q^{k/2} + 1 & \text{if } k = 0 \mod 2 \\ q^k - q^{(k+1)/2} - q^{(k-1)/2} + 1 & \text{if } k = 1 \mod 2 \end{cases} \tag{4}$$

and

$$N_k = \begin{cases} (q^2-q)q^{k-1} + 2q & \text{if } k \geq 3, q = 0 \mod 2 \\ (q^2-q)q^{k-1} + 2q^2 & \text{if } k \geq 3, q = 1 \mod 2 \end{cases} \tag{5}$$

We are now ready to prove Theorem 11.

Proof. We first prove the result for $q \geq 2(d+1)$. Let r be an integer such that $q^{r+1} < 2dt \leq q^{r+2}$. Consider the function field $F^{(r+1)}/\mathbb{F}_{q^2}$ defined in Lemma 16. By Lemma 16, $F^{(r+1)}$ is of genus $g \leq q^{r+1}$ and has $N \geq q^{r+2} - q^{r+1}$ places of degree 1. By Lemma 15, since

$$\begin{aligned} N &\geq q^{r+2} - q^{r+1} = \frac{q^{r+2}}{2} + \left(\frac{q^{r+2}}{2} - q^{r+1}\right) \\ &\geq dt + dq^{r+1} \\ &\geq d(t+g-1) + 1 \end{aligned}$$

we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{q^{2t}}) \leq dt + (d-1)(g-1) \leq d(t + q^{r+1}) \leq d(2d+1)t. \tag{6}$$

Now by Lemma 2 and 7 and Corollary 5 we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^{2t}}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^2}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{q^{2t}}) = d(d+1)(2d+1)t.$$

This proves the result for $q \geq 2(d+1)$.

Notice that thus far we have proved that: for any t and any $q \geq 2(d+1)$ we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq d(d+1)(2d+1)t$. In particular, for any t and any $q^2 \geq 2(d+1)$ we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{q^{2t}}) \leq d(d+1)(2d+1)t. \tag{7}$$

Now for $q \geq d+1$ we have $q^2 \geq 2(d+1)$ and with (7) we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^{2t}}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^2}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{q^{2t}}) = d(d+1)^2(2d+1)t. \tag{8}$$

This proves the result. □

Recall that $\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$. Therefore all the above bounds and the bounds in the following Corollary are also true for $\nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$ and $\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$. We now prove

Corollary 17. *Let $c > 1$ be any constant. We have*

1. *For perfect square q where $q \geq c(d+1)^2$ there is an infinite sequence of integers t such that $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(dt)$.*
2. *For perfect square q where $q \geq c(d+1)^2$ and any integer t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^2t)$.*
3. *For any $q \geq c(d+1)$ there is an infinite sequence of integers t such that $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^2t)$.*
4. *For any $q \geq c(d+1)$ and any integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^3t)$.*
5. *For any $q \geq d+1$ there is an infinite sequence of integers t such that $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^3t)$.*
6. *For any $q \geq d+1$ and any integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^4t)$.*
7. *For any $q \geq d$ and any integers t we have $\nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^4t)$.*
8. *For any $q < d+1$ and any integers $t > 1$ we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = \infty$.*

9. For any $q < d$ and any integers $t > 1$ we have $\nu_{\mathbb{F}_q}^{\mathcal{HP}}(d, \mathbb{F}_{q^t}) = \infty$.

Proof. Let $q \geq c(d+1)$ for some $c > 1$ and r be an integer such that $q^{r+1} < (c/(c-1))dt \leq q^{r+2}$. Consider the function field $F^{(r+1)}/\mathbb{F}_{q^2}$ as in the proof of Theorem 11. Then $F^{(r+1)}/\mathbb{F}_{q^2}$ is of genus $g \leq q^{r+1}$ and has $N \geq q^{r+2} - q^{r+1}$ places of degree 1. Since

$$N \geq q^{r+2} - q^{r+1} \geq q^{r+2} - \left(\frac{q}{c} - d\right) q^{r+1} \geq \frac{c-1}{c} q^{r+2} + dq^{r+1} \geq d(t + q^{r+1}) \geq d(t+g-1)+1,$$

by Lemma 15,

$$\nu_{\mathbb{F}_{q^2}}^{\mathcal{P}}(d, \mathbb{F}_{q^{2t}}) \leq d(t+q^{r+1}) \leq \left(\frac{c}{c-1}d + 1\right) dt. \tag{9}$$

Therefore, for a perfect square $q \geq c(d+1)^2$ we have $\sqrt{q} \geq \sqrt{c}(d+1)$. Then (9) implies

$$\nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^t}) \leq \left(\frac{\sqrt{c}}{\sqrt{c}-1}d + 1\right) dt. \tag{10}$$

This proves *2.*

If $t = \lfloor (c-1)q^{r+2}/(cd) \rfloor$ then $q^{r+1} < (c/(c-1))dt \leq q^{r+2}$ and therefore

$$\begin{aligned} \nu_{\mathbb{F}_{q^2}}^{\mathcal{P}}(d, \mathbb{F}_{q^{2t}}) &\leq d(t+q^{r+1}) \\ &\leq d\left(t + \frac{d}{q}\frac{c(t+1)}{(c-1)}\right) \\ &\leq d\left(t + \frac{(t+1)}{(c-1)}\right) \leq \frac{2c}{c-1}dt. \end{aligned} \tag{11}$$

Therefore, for a perfect square $q \geq c(d+1)^2$ we have $\sqrt{q} \geq \sqrt{c}(d+1)$. Then for $t = \lfloor (\sqrt{c}-1)q^{(r+2)/2}/(\sqrt{c}d) \rfloor$, (11) implies

$$\nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^t}) \leq \frac{2\sqrt{c}}{\sqrt{c}-1}dt. \tag{12}$$

This proves *1.*

Now, for $q \geq c(d+1)$, by Lemma 2, Lemma 7, Corollary 5 and (9), we have

$$\nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^t}) \leq \nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^{2t}}) \leq \nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^2}) \cdot \nu_{\mathbb{F}_{q^2}}^{\mathcal{P}}(d, \mathbb{F}_{q^{2t}}) \leq \left(\frac{c}{c-1}d + 1\right) d(d+1)t. \tag{13}$$

For $t = \lfloor (c-1)q^{r+2}/(cd) \rfloor$, by (11),

$$\nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^t}) \leq \nu_{\mathbb{F}_q}^{\mathcal{P}}(d, \mathbb{F}_{q^2}) \cdot \nu_{\mathbb{F}_{q^2}}^{\mathcal{P}}(d, \mathbb{F}_{q^{2t}}) \leq \frac{2c}{c-1} d(d+1)t. \tag{14}$$

This proves *3* and *4*.

Now by (13) and (14) if $q \geq d+1$ then $q^2 = (d+1)^2 \geq 2(d+1)$ and therefore,

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{(q^2)^t}) \leq (2d+1)d(d+1)t$$

and for $t = \lfloor (q^2)^{r+2}/(2d) \rfloor$,

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{(q^2)^t}) \leq 4d(d+1)t.$$

Therefore

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^{2t}}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^2}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{q^{2t}}) \leq d(d+1)^2(2d+1)t,$$

and for $t = \lfloor (q^2)^{r+2}/(2d) \rfloor$,

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^{2t}}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^2}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{q^{2t}}) \leq 4d(d+1)^2 t.$$

This proves *5* and *6*.

Now by Corollary 5, Lemma 7 and *4*, for $q \geq d$,

$$\nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^2}) \cdot \nu^{\mathcal{HP}}_{\mathbb{F}_{q^2}}(d, \mathbb{F}_{(q^2)^t}) \leq O(d^4)t.$$

This proves *7*.

8 and *9* are proved in Lemma 28. □

We note here that there are other results that can be obtained with other conditions on d and q that are not included in the above lemma. For example, when q is perfect square and $q \geq (d+2)^2$ then $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq O(d^3 t)$. This follows if we choose $c = (d+2)/(d+1)$ in the above proof.

In Subsection 3.4 we show that a tester with the bound in Theorem 11 can be constructed in polynomial time. In Subsection 2.7 we show that the size of the above tester is almost optimal.

Before we leave this subsection we state the following open problems

Open Problems 1. *Let $c > 1$ be any constant. Prove*

1. *For perfect square q where $q \geq c(d+1)^2$ and any integer t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(dt)$.*
   This will implies
2. *For any $q \geq c(d+1)$ and any integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^2 t)$ and*
3. *For any $q \geq d+1$ and any integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^3 t)$.*
   Is the following true?
4. *There is a constant $C > 1$ such that for every $q \geq C(d+1)$ and infinite sequence of integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(dt)$.*

2.4 Testers for Subspaces of Large Fields

In this subsection we study $(\mathcal{M}, S, \mathbb{F}_q)$-testers when $S \subset \mathbb{F}_{q^t}$ is a linear subspace over $\mathbb{F}_q$. In the next section we show that such testers can be easily constructed in polynomial time and, for many applications, are almost as good as the testers in the previous subsection.

We start with the following

Lemma 18. *Let $F/\mathbb{F}_q$ be a function field of genus g. Let G be a divisor of $F/\mathbb{F}_q$ and Q be a prime divisor of degree $\deg Q = t = \deg G + 1$ such that $v_Q(G) = 0$. Then the map*

$$\begin{aligned} E : \mathscr{L}(G) &\to F_Q = \mathbb{F}_{q^t} \\ f &\mapsto f(Q) \end{aligned}$$

is one-to-one linear map and there is a symmetric and reducible $(\mathbb{F}_q[\boldsymbol{x}], S, \mathscr{L}(G))$-tester of size 1 where $S = E(\mathscr{L}(G))$ and $\boldsymbol{x} = (x_1, \ldots, x_m)$.

In particular, $S \subseteq \mathbb{F}_{q^t}$ is a linear subspace over $\mathbb{F}_q$, $|S| = q^{l(G)} \geq q^{t-g}$ and

$$\nu^{\circ}_{\mathscr{L}(G)}(\mathbb{F}_q[\boldsymbol{x}], S) = 1.$$

Proof. We first show that the map E is a one-to-one linear map. Let $f \in \mathscr{L}(G)$. Then $v_Q(f) \geq -v_Q(G) = 0$ and therefore $f(Q) \neq \infty$ and the map is well defined. The fact that it is a linear map follows from Proposition 88.

Now we show that $\mathrm{Ker}E = \{0\}$. Let $f \in \mathscr{L}(G)$. Then $(f) \geq -G$. If $f \in \mathrm{Ker}E$ then $f(Q) = 0$ and therefore $v_Q(f) > 0$. Since $v_Q(G) = 0$ we have $f \in \mathscr{L}(G - Q)$. Since $deg(G - Q) = -1$ by *3* in Proposition 90, $\mathscr{L}(G - Q) = \{0\}$ and therefore $f = 0$. Therefore E is a one-to-one linear map and the map $E_S : \mathscr{L}(G) \to S$ defined as $E_S(f) = E(f)$ is an isomorphism of linear spaces over $\mathbb{F}_q$.

Now suppose $h \in \mathbb{F}_q[\boldsymbol{x}]$ and let $\boldsymbol{E}_S = (E_S, E_S, \ldots, E_S)$ and $\boldsymbol{E}_S^{-1} = (E_S^{-1}, E_S^{-1}, \ldots, E_S^{-1})$. Consider the tester $L = \{\boldsymbol{E}_S^{-1}\}$. Let $\boldsymbol{a} \in S^m$. If $(hL)(\boldsymbol{a}) = 0$ then $0 = ((hL)(\boldsymbol{a}))(Q) = h(\boldsymbol{E}_S^{-1}(\boldsymbol{a}))(Q) = h(\boldsymbol{E}_S(\boldsymbol{E}_S^{-1}(\boldsymbol{a}))) = h(\boldsymbol{a})$.

Since $E_S(1) = 1(Q) = 1$ we have $E_S^{-1}(1) = 1$ and the tester is reducible.

Since E_S is isomorphism we have $|S| = |\mathscr{L}(G)| = q^{l(G)}$ and by the Riemann-Roch Theorem, Proposition 92, we have $l(G) \geq \deg G + 1 - g = t - g$. □

We now prove

Lemma 19. *Let $F/\mathbb{F}_q$ be a function field of genus g and let $t, g > 4$. If F has $d(t + g - 1) + 2$ places of degree 1 then there is a sublinear space $S \subseteq \mathbb{F}_{q^{t+g}}$ of size q^t such that $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S) \leq d(t + g - 1) + 1$.*

Proof. First, by Corollary 5.2.10 (c) in [67], if $2g+1 \leq q^{(t+g-1)/2}(q^{1/2}-1)$ then there is a prime divisor of degree $t+g$. Since $t, g > 4$ the inequality holds and there is at least one prime divisor of degree $t+g$. Let Q be a prime divisor of degree $t+g$.

Let $P_1, \ldots, P_s, P_\infty$, $s = d(t+g-1)+1$, be distinct places of $F/\mathbb{F}_q$ of degree 1 and $D = P_1 + P_2 + \cdots + P_s$ and let $G = (t+g-1)P_\infty$. By Lemma 18 there is a subspace $S \subseteq \mathbb{F}_{q^{t+g}}$ of size $|S| = q^t$ such that $\nu^{\mathcal{P}}_{\mathscr{L}(G)}(d, S) = 1$.

By *1* in Lemma 12 and Lemma 4 we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathscr{L}(G)) \cdot \nu^{\mathcal{P}}_{\mathscr{L}(G)}(d, S) \leq d \deg G + 1 = d(t+g-1)+1.$$

□

Note that in the above lemma we could have used *2* in Lemma 12 and get $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S) \leq dt + (d-1)(g-1)$. This will not effect the asymptotic bounds we get in this paper. Also, we will see in the next subsection that the construction of such tester is easier using this bound.

Now we prove a result similar to Corollary 17

Corollary 20. *Let $c > 1$ be any constant. For every t there is a sublinear space $S_t \subseteq \mathbb{F}_{q^T}$ over $\mathbb{F}_q$ for some $T = O(dt)$ of size $|S_t| = q^t$ such that:*

1. *For perfect square q where $q \geq c(d+1)^2$ and any integer t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S_t) = O(d^2 t)$.*
2. *For any $q \geq c(d+1)$ and any integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S_t) = O(d^3 t)$.*
3. *For any $q \geq d+1$ and any integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S_t) = O(d^4 t)$.*

Let $c > 1$ be any constant. For every t there is a sublinear space $S_t \subseteq \mathbb{F}_{q^T}$ over $\mathbb{F}_q$ for some $T = O(t)$ of size $|S_t| = q^t$ such that:

4. *For perfect square q where $q \geq c(d+1)^2$ there is an infinite sequence of integers t such that $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S_t) = O(dt)$.*
5. *For any $q \geq c(d+1)$ there is an infinite sequence of integers t such that $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S_t) = O(d^2 t)$.*
6. *For any $q \geq d+1$ there is an infinite sequence of integers t such that $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, S_t) = O(d^3 t)$.*

Proof. The proof is very similar to that of Corollary 17. Let $q \geq c(d+1)$ and r be an integer such that $q^{r+1} < (c/(c-1))dt \leq q^{r+2}$. Consider the function field $F^{(r+1)}/\mathbb{F}_{q^2}$ as in the proof of Theorem 11. Then $F^{(r+1)}/\mathbb{F}_{q^2}$ is of genus $g \leq q^{r+1}$ and has $N \geq q^{r+2} - q^{r+1}$ places of degree 1. Since

$$N \geq q^{r+2} - q^{r+1} \geq q^{r+2} - \left(\frac{q}{c} - d\right) q^{r+1} \geq \frac{c-1}{c} q^{r+2} + dq^{r+1} \geq d(t + q^{r+1}) \geq d(t+g-1)+2$$

we have $N \geq d(t+g-1)+2$ and by Lemma 19, there is a sublinear space $S \subseteq \mathbb{F}_{(q^2)^{t+g}}$ of size $(q^2)^t$ such that

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d,S) \leq d(t+q^{r+1}) \leq \left(\frac{c}{c-1}d+1\right)dt. \tag{15}$$

Also $T = t+g \leq t+q^{r+1} = O(dt)$. This proves *1.*

When $t = \lfloor (c-1)q^{r+2}/(cd) \rfloor$ then

$$\begin{aligned}
\nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d,S) &\leq d(t+q^{r+1}) \\
&\leq d\left(t + \frac{d}{q}\frac{c(t+1)}{(c-1)}\right) \\
&\leq d\left(t + \frac{(t+1)}{(c-1)}\right) \leq \frac{2c}{c-1}dt.
\end{aligned}$$

In this case $T = t+g \leq t+q^{r+1} = O(t)$. This proves *4.*

Now for the same $S \subseteq \mathbb{F}_{(q^2)^{t+g}} = \mathbb{F}_{q^{2t+2g}}$ as above, by Lemma 7 and Corollary 5 we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d,S) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d,\mathbb{F}_{q^2}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d,S) = \left(\frac{c}{c-1}d+1\right)d(d+1)t. \tag{16}$$

and for $t = \lfloor (c-1)q^{r+2}/(cd) \rfloor$,

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d,S) \leq \frac{2c}{c-1}d(d+1)t. \tag{17}$$

Since $T = 2t+2g \leq 2t+2q^{r+1} = O(dt)$ in the former and $= O(t)$ in the latter, this proves *2* and *5.*

Now for $q \geq d+1$ we have $q^2 = (d+1)^2 \geq 2(d+1)$ and therefore by (16) and (17), there is $S' \subseteq \mathbb{F}_{(q^2)^{(2t+2g)}} = \mathbb{F}_{q^{(4t+4g)}}$ of size $|S'| = q^t$ such that

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d,S') \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d,\mathbb{F}_{q^2}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^2}}(d,S') = d(d+1)^2(2d+1)t,$$

and for $t = \lfloor q^{r+2}/(2d) \rfloor$,

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d,S') \leq 4d(d+1)^2.$$

Since $T = 4t+4g \leq 4t+4q^{r+1} = O(dt)$ in the former and $= O(t)$ in the latter, this proves *3* and *6.* □

We note here that Corollary 20 is subsumed by Corollary 17 but as we will show in the sequel, its construction requires less time complexity.

2.5 Testers for Small Fields

In this subsection we use some elementary algebra and Theorem 11 to construct testers for small fields. We prove

Theorem 21. *For any $q < d+1$ and t we have*

$$\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq poly(d) \cdot 2^{c_q d} \cdot t$$

where

$$c_q = \sum_{i=0}^{\infty} \frac{\log(q^{2^i}+1)}{q^{2^i}} = O\left(\frac{\log q}{q}\right).$$

In particular we have following values of c_q

q	c_q
2	1.659945821
3	1.116191294
4	0.867464571
5	0.719921672
7	0.548433289

Proof. By Lemma 2, Corollary 5, Lemma 6 and 7 and (1) we have

1. $\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^{t_1}}) \leq \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^{t_1 t_2}})$.
2. $\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^{t_1 t_2}}) \leq \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^{t_1}}) \cdot \nu_{\mathbb{F}_{q^{t_1}}}(d, \mathbb{F}_{q^{t_1 t_2}})$.
3. $\nu_{\mathbb{F}_q}(d_1 + d_2, \mathbb{F}_{q^t}) \leq \nu_{\mathbb{F}_q}(d_1, \mathbb{F}_{q^t}) \cdot \nu_{\mathbb{F}_q}(d_2, \mathbb{F}_{q^t})$.
4. $\nu_{\mathbb{F}_q}(q, \mathbb{F}_{q^2}) \leq q+1$.

Let r be such that $q^{2^{r-1}} < d+1 \leq q^{2^r}$. Then

$$\begin{aligned}
\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) &\leq \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^{t \cdot 2^r}}) \quad \text{By (1)} \\
&\leq \left(\prod_{i=0}^{r-1} \nu_{\mathbb{F}_{q^{2^i}}}(d, \mathbb{F}_{q^{2^{i+1}}}) \right) \nu_{\mathbb{F}_{q^{2^r}}}(d, \mathbb{F}_{(q^{2^r})^t}) \quad \text{By (2)} \\
&\leq \left(\prod_{i=0}^{r-1} \nu_{\mathbb{F}_{q^{2^i}}}(q^{2^i}, \mathbb{F}_{q^{2^{i+1}}})^{\lceil d/q^{2^i} \rceil} \right) \nu_{\mathbb{F}_{q^{2^r}}}(d, \mathbb{F}_{(q^{2^r})^t}) \quad \text{By (3)} \\
&\leq \left(\prod_{i=0}^{r-1} \left(q^{2^i}+1\right)^{\lceil d/q^{2^i} \rceil} \right) \nu_{\mathbb{F}_{q^{2^r}}}(d, \mathbb{F}_{(q^{2^r})^t}) \quad \text{By (4)} \\
&\leq \left(\left(\prod_{i=0}^{r-1} \left(q^{2^i}+1\right) \right) 2^{c_q d} \right) poly(d) \cdot t \quad \text{By Theorem 11} \\
&\leq poly(d) \cdot 2^{c_q d} \cdot t.
\end{aligned}$$

□

We note here that the $poly(d)$ part in the proof is at most $O(d^6)$. We now show how to reduce it to $O(d^5)$. Take r such that $q^{2^{r-1}} < 2(d+1) \leq q^{2^r}$. Then

$$\prod_{i=0}^{r-1} (q^{2^i}+1) = \frac{q^{2^r}-1}{q-1} \leq 4(d+1)^2$$

and by Corollary 17 for $q^{2^r} \geq 2(d+1)$

$$\nu_{\mathbb{F}_{q^{2^r}}}(d, \mathbb{F}_{(q^{2^r})^t}) = O(d^3)t.$$

Therefore

$$\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \left(\left(\prod_{i=0}^{r-1} \left(q^{2^i}+1\right) \right) 2^{c_q d} \right) \nu_{\mathbb{F}_{q^{2^r}}}(d, \mathbb{F}_{(q^{2^r})^t}) \leq O(d^5) 2^{c_q d} \cdot t.$$

In Subsection 3.6 we show that a tester with such size can be constructed in time $2^{c_q d} \cdot poly(t)$. In Subsection 2.7 we give the lower bound $\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = 2^{\Omega(d/q)}t$ for the above tester. This shows that the size of the above tester is almost optimal. We now list some open problems

Open Problems 2.

1. *For $q = 2$ the upper bound for $\nu_{\mathbb{F}_2}(d, \mathbb{F}_{2^t})$ is $2^{1.6599d} \cdot t$ and the lower bound is $2^d \cdot t$ (see Theorem 27). Find better bounds.*

2. *A computer program may exhaustively searches for a better upper bound for $\nu_{\mathbb{F}_2}(d, \mathbb{F}_{2^2})$ for small d. This will lead to a better upper bound for $\nu_{\mathbb{F}_2}(d, \mathbb{F}_{2^t})$ for any d. For example, $\nu_{\mathbb{F}_2}(4, \mathbb{F}_{2^2}) \leq \nu_{\mathbb{F}_2}(2, \mathbb{F}_{2^2})^2 = 9$. If we can prove $\nu_{\mathbb{F}_2}(4, \mathbb{F}_{2^2}) \leq 8$ then $\nu_{\mathbb{F}_2}(d, \mathbb{F}_{2^2}) \leq 8^{\lceil d/4 \rceil}$ and then we get the upper bound $\nu_{\mathbb{F}_2}(d, \mathbb{F}_{2^t}) \leq 2^{1.6174d} \cdot t$.*

2.6 Symmetric and Reducible Testers

In this subsection we give a classification of symmetric and reducible tester. This classification will first help us understand the algebraic structure of symmetric and reducible tester. Then it will show that the problem of deciding, given a set of maps L, whether L is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester is in NP. This will be used in the next subsection to construct a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of almost (within d^2) optimal size in polynomial time.

2.6.1 Definition

We recall the definition of symmetric and reducible tester

Definition 22. *Let $\mathbb{F}$ be a field and $\mathcal{A}$ and $\mathcal{B}$ be two commutative $\mathbb{F}$-algebras. Let $\mathcal{M} \subseteq \mathbb{F}[x_1, x_2, \ldots, x_n]$ be a class of multivariate polynomials. Let $S \subseteq \mathcal{A}$ and $R \subseteq \mathcal{B}$ be sub-linear spaces over $\mathbb{F}$ and $L = \{\ell_1, \ldots, \ell_\nu\}$ be a set of maps $S \to R$. Let $\boldsymbol{\ell}_i : S^n \to R^n$ where for $\boldsymbol{a} \in S^n$,*

$$\boldsymbol{\ell}_i(\boldsymbol{a}) = (\ell_i(a_1), \ell_i(a_2), \ldots, \ell_i(a_n)).$$

We denote by fL the map $S^n \to \mathcal{B}^\nu$ where for $\boldsymbol{a} \in S^n$, $(fL)(\boldsymbol{a}) = (f(\boldsymbol{\ell}_1(\boldsymbol{a})), \ldots, f(\boldsymbol{\ell}_\nu(\boldsymbol{a})))$. We say that L is an symmetric $(\mathcal{M}, S, R)$-tester if for every $\boldsymbol{a} \in S^n$ and $f \in \mathcal{M}$ we have

$$(fL)(\boldsymbol{a}) = 0^\nu \implies f(\boldsymbol{a}) = 0.$$

We say that L is reducible if for all i, $\ell_i(1_\mathcal{B}) = 1_\mathcal{A}$ where $1_\mathcal{A}$ and $1_\mathcal{B}$ are the identities of the algebras $\mathcal{A}$ and $\mathcal{B}$ respectively.

For $q \geq d+1$ we define $\tau(d, q, t)$ the constant for which a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^{\tau(d,q,t)} \cdot t)$ exists. Define $\tau^*(d, q, t)$ the constant for which a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^{\tau^*(d,q,t)} \cdot t)$ exists. Obviously

$$\tau(d, q, t) \leq \tau^*(d, q, t).$$

Since the testers in Lemma 12 and Lemma 13 are symmetric and reducible and Corollary 5 is also true for symmetric and reducible testers, the testers constructed in Theorem 11 and Corollary 17 are symmetric and reducible. Therefore

$$\tau^*(d, q, t) \leq 4 \tag{18}$$

for any $q \geq d + 1$ and any t and

Corollary 23. *We have*

1. *For perfect square q where $q \geq c(d+1)^2$ there is infinite sequence of integers t such that $\tau^*(d,q,t) \leq 1$.*
2. *For perfect square q where $q \geq c(d+1)^2$ and any integer t we have $\tau^*(d,q,t) \leq 2$.*
3. *For any $q \geq c(d+1)$ there is an infinite sequence of integers t such that $\tau^*(d,q,t) \leq 2$.*
4. *For any $q \geq c(d+1)$ and any integers t we have $\tau^*(d,q,t) \leq 3$.*
5. *For any $q \geq d+1$ there is an infinite sequence of integers t such that $\tau^*(d,q,t) \leq 3$.*
6. *For any $q \geq d+1$ and any integers t we have $\tau^*(d,q,t) \leq 4$.*

2.6.2 Classification

In this subsection we give a classification of symmetric and reducible tester

We now prove

Lemma 24. *We have: $L = \{\ell_1, \ldots, \ell_\nu\}$ is a symmetric $(\mathcal{HP}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester for all n if and only if there are $\beta_1, \ldots, \beta_\nu \in \mathbb{F}_{q^t}$ such that for every $a_1, a_2, \ldots, a_d \in \mathbb{F}_{q^t}$*

$$a_1 a_2 \cdots a_d = \sum_{i=1}^{\nu} \beta_i \ell_i(a_1)\ell_i(a_2)\cdots\ell_i(a_d).$$

Proof. ($\Rightarrow$) Suppose L is a symmetric $(\mathcal{HP}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester. For $\boldsymbol{a} = (a_1, a_2, \ldots, a_d) \in \mathbb{F}_{q^t}^d$ define $\lambda_i(\boldsymbol{a}) = \ell_i(a_1)\ell_i(a_2)\cdots\ell_i(a_d)$ and $\Lambda(\boldsymbol{a}) = (\lambda_1(\boldsymbol{a}), \ldots, \lambda_\nu(\boldsymbol{a}))$. Let

$$G = \{\Lambda(\boldsymbol{a}) \mid \boldsymbol{a} \in \mathbb{F}_{q^t}^d\}.$$

Define the binary relation $\phi \subset G \times \mathbb{F}_{q^t}$

$$\phi := \{(\Lambda(\boldsymbol{a}), a_1 a_2 \cdots a_d) \mid \boldsymbol{a} \in \mathbb{F}_{q^t}^d\}.$$

Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{F}_{q^t}^d$. Since $f := x_1 x_2 \cdots x_d - x_{d+1}x_{d+2}\cdots x_{2d} \in \mathcal{HP}(\mathbb{F}_q, n, d)$ we have: $f(\ell_i(a_1), \ldots, \ell_i(a_d), \ell_i(b_1), \ldots, \ell_i(b_d)) = 0$ for all i implies $f(a_1, \ldots, a_d, b_1, \ldots, b_d) = 0$. Thus, if $\Lambda(\boldsymbol{a}) = \Lambda(\boldsymbol{b})$ then $a_1 \cdots a_d = b_1 \cdots b_d$ and therefore the relation ϕ is a function $\phi : G \to \mathbb{F}_{q^t}$ and $\phi(\Lambda(\boldsymbol{a})) = a_1 a_2 \cdots a_d$.

Now, let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r, \boldsymbol{b} \in \mathbb{F}_{q^t}^d$ and suppose $\Lambda(\boldsymbol{b}) = \gamma_1\Lambda(\boldsymbol{a}_1) + \cdots + \gamma_r\Lambda(\boldsymbol{a}_r)$ for some $\gamma_1, \ldots, \gamma_r \in \mathbb{F}_q$. Let

$$g = x_1 \cdots x_d - \gamma_1 x_{d+1} \cdots x_{2d} - \cdots - \gamma_r x_{rd+1} \cdots x_{(r+1)d} \in \mathcal{HP}(\mathbb{F}_q, n, d).$$

Then

$$(gL)(\boldsymbol{b}, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_r) = \Lambda(\boldsymbol{b}) - \gamma_1 \Lambda(\boldsymbol{a}_1) - \cdots - \gamma_r \Lambda(\boldsymbol{a}_r) = 0^\nu$$

and therefore $g(\boldsymbol{b}, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_r) = \phi(\Lambda(\boldsymbol{b})) - \gamma_1 \phi(\Lambda(\boldsymbol{a}_1)) - \cdots - \gamma_r \phi(\Lambda(\boldsymbol{a}_r)) = 0$ and

$$\phi(\Lambda(\boldsymbol{b})) = \gamma_1 \phi(\Lambda(\boldsymbol{a}_1)) + \cdots + \gamma_r \phi(\Lambda(\boldsymbol{a}_r)).$$

Therefore ϕ is linear function restricted on G and there is a natural unique extension of ϕ to a linear function $\hat{\phi} : \text{Span } G \to \mathbb{F}_{q^t}$ where $\hat{\phi}|_G = \phi$. This implies the result.

($\Leftarrow$) Suppose there are $\beta_1, \ldots, \beta_\nu \in \mathbb{F}_{q^t}$ such that for every $a_1, a_2, \ldots, a_d \in \mathbb{F}_{q^t}$

$$a_1 a_2 \cdots a_d = \sum_{i=1}^{\nu} \beta_i \ell_i(a_1) \ell_i(a_2) \cdots \ell_i(a_d).$$

Then for every

$$f = \sum_{\boldsymbol{i} \in I} c_{\boldsymbol{i}} x_{i_1} \cdots x_{i_d} \in \mathcal{HP}(\mathbb{F}_q, n, d)$$

where $I \subseteq [n]^d$ and every $(b_1, \ldots, b_n) \in \mathbb{F}_{q^t}^n$, we have

$$\begin{aligned} f(b_1, \ldots, b_n) &= \sum_{\boldsymbol{i} \in I} c_{\boldsymbol{i}} b_{i_1} \cdots b_{i_d} \\ &= \sum_{\boldsymbol{i} \in I} c_{\boldsymbol{i}} \sum_{j=1}^{\nu} \beta_j \ell_j(b_{i_1}) \ell_j(b_{i_2}) \cdots \ell_j(b_{i_d}) \\ &= \sum_{j=1}^{\nu} \beta_j \sum_{\boldsymbol{i} \in I} c_{\boldsymbol{i}} \ell_j(b_{i_1}) \ell_j(b_{i_2}) \cdots \ell_j(b_{i_d}) \\ &= \sum_{j=1}^{\nu} \beta_j f(\ell_j(b_1), \ldots, \ell_j(b_n)). \end{aligned}$$

Therefore, if $f(\ell_j(b_1), \ldots, \ell_j(b_n)) = 0$ for all $j = 1, 2, \ldots, \nu$ then $f(b_1, \ldots, b_n) = 0$. □

We now prove

Lemma 25. *We have: $L = \{\ell_1, \ldots, \ell_\nu\}$ is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester for all n if and only if there are $\beta_1, \ldots, \beta_\nu \in \mathbb{F}_{q^t}$ such that for every $a_1, a_2, \ldots, a_d \in \mathbb{F}_{q^t}$*

$$a_1 a_2 \cdots a_d = \sum_{i=1}^{\nu} \beta_i \ell_i(a_1) \ell_i(a_2) \cdots \ell_i(a_d) \tag{19}$$

and $\ell_i(1) = 1$ for all i.

Proof. ($\Rightarrow$) If L is a symmetric $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester then it is symmetric $(\mathcal{HP}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester and by Lemma 24 the result follows.

($\Leftarrow$) For any $d' < d$ we have

$$\begin{aligned} a_1 a_2 \cdots a_{d'} &= a_1 a_2 \cdots a_{d'} \cdot 1 \overset{d-d'}{\cdots} 1 \\ &= \sum_{i=1}^{\nu} \beta_i \ell_i(a_1)\ell_i(a_2)\cdots\ell_i(a_{d'}))\ell_i(1) \overset{d-d'}{\cdots} \ell_i(1) \\ &= \sum_{i=1}^{\nu} \beta_i \ell_i(a_1)\ell_i(a_2)\cdots\ell_i(a_{d'}) \end{aligned}$$

Then as in the proof of Lemma 24 we get that for every $f \in \mathcal{P}(\mathbb{F}_q, n, d)$

$$f(b_1, \ldots, b_n) = \sum_{j=1}^{\nu} \beta_j f(\ell_j(b_1), \ldots, \ell_j(b_n))$$

and therefore $L = \{\ell_1, \ldots, \ell_\nu\}$ is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester. □

In the following lemma we show that when the tester is symmetric then we may assume without loss of generality that all ℓ_i are linear functions from the dual vector space $\mathbb{F}^*_{q^t}$.

Lemma 26. *If $L = \{\ell_1, \ldots, \ell_\nu\}$ is a symmetric (and reducible) $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester for all n then there is a symmetric (and reducible) $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester $\tilde{L} = \{\tilde{\ell}_1, \ldots, \tilde{\ell}_\nu\}$ with the same size where each $\tilde{\ell}_i$, $i = 1, \ldots, \nu$ is a linear function in $\mathbb{F}^*_{q^t}$.*

Proof. Since $L = \{\ell_1, \ldots, \ell_\nu\}$ is a symmetric (and reducible) $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester for all n, by Lemma 24 (and Lemma 25) there are $\beta_1, \ldots, \beta_\nu \in \mathbb{F}_{q^t}$ such that for every $a_1, a_2, \ldots, a_d \in \mathbb{F}_{q^t}$

$$a_1 a_2 \cdots a_d = \sum_{i=1}^{\nu} \beta_i \ell_i(a_1)\ell_i(a_2)\cdots\ell_i(a_d) \tag{20}$$

(and $\ell_i(1) = 1$ for all i). Let $W = \{\omega_1, \ldots, \omega_t\}$ be a basis for $\mathbb{F}_{q^t}$ over $\mathbb{F}_q$ where $\omega_1 = 1$. Define the linear functions $\tilde{\ell}_i$ where $\tilde{\ell}_i(\omega_j) = \ell_i(\omega_j)$ for all $i = 1, 2, \ldots, \nu$ and $j = 1, 2, \ldots, t$, and for every element $a = \lambda_1\omega_1 + \cdots + \lambda_t\omega_t \in \mathbb{F}_{q^t}$ where $\lambda_k \in \mathbb{F}_q$, $k = 1, 2, \ldots, t$ we have $\tilde{\ell}_i(a) = \lambda_1\tilde{\ell}_i(\omega_1) + \cdots + \lambda_t\tilde{\ell}_i(\omega_t)$. Obviously $\tilde{\ell}_i \in \mathbb{F}^*_{q^t}$. We now claim that $\tilde{L} = \{\tilde{\ell}_1, \ldots, \tilde{\ell}_\nu\}$ is a symmetric (and reducible) $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester.

Let $a_i = \lambda_{i,1}\omega_1 + \cdots + \lambda_{i,t}\omega_t$ for $i = 1, 2, \ldots, d$. Then

$$\begin{aligned}
a_1 a_2 \cdots a_d &= \prod_{i=1}^{d}\left(\sum_{j=1}^{t}\lambda_{i,j}\omega_j\right) = \sum_{\boldsymbol{k}\in[t]^d}\left[\left(\prod_{i=1}^{d}\lambda_{i,k_i}\right)\left(\prod_{i=1}^{d}\omega_{k_i}\right)\right] \\
&= \sum_{\boldsymbol{k}\in[t]^d}\left[\left(\prod_{i=1}^{d}\lambda_{i,k_i}\right)\left(\sum_{m=1}^{\nu}\beta_m\ell_m(\omega_{k_1})\ell_m(\omega_{k_2})\cdots\ell_m(\omega_{k_d})\right)\right] \\
&= \sum_{\boldsymbol{k}\in[t]^d}\left[\left(\prod_{i=1}^{d}\lambda_{i,k_i}\right)\left(\sum_{m=1}^{\nu}\beta_m\tilde{\ell}_m(\omega_{k_1})\tilde{\ell}_m(\omega_{k_2})\cdots\tilde{\ell}_m(\omega_{k_d})\right)\right] \\
&= \sum_{m=1}^{\nu}\left(\beta_m\sum_{\boldsymbol{k}\in[t]^d}\left[\left(\prod_{i=1}^{d}\lambda_{i,k_i}\right)\left(\tilde{\ell}_m(\omega_{k_1})\tilde{\ell}_m(\omega_{k_2})\cdots\tilde{\ell}_m(\omega_{k_d})\right)\right]\right) \\
&= \sum_{m=1}^{\nu}\left(\beta_m\prod_{i=1}^{d}\left(\sum_{j=1}^{t}\lambda_{i,j}\tilde{\ell}_m(\omega_j)\right)\right) \\
&= \sum_{m=1}^{\nu}\left(\beta_m\prod_{i=1}^{d}\tilde{\ell}_m\left(\sum_{j=1}^{t}\lambda_{i,j}\omega_j\right)\right) \\
&= \sum_{m=1}^{\nu}\beta_m\tilde{\ell}_m(a_1)\tilde{\ell}_m(a_2)\cdots\tilde{\ell}_m(a_d).
\end{aligned}$$

Therefore, by Lemma 24, $\tilde{L} = \{\tilde{\ell}_1, \ldots, \tilde{\ell}_\nu\}$ is a symmetric $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester.

Now if the tester is reducible then since $\omega_1 = 1$ we have $\tilde{\ell}_i(\omega_1) = \ell_i(\omega_1) = 1$. By Lemma 25, $\tilde{L}$ is reducible. □

Testers of the form (19) are called symmetric linear builders [14]. For $d = 2$, symmetric builders and their connection to tensor rank are studied in [72]. We will further study builders in [14] and their connection to d-dimensional tensor rank and give other results.

We now list some open problems

Open Problems 3.

1. *Give a classification of non-symmetric testers.*
2. *Since symmetric testers are equivalent to (symmetric) rank of d-dimensional tensors, many of the results in the theory of bilinear complexity (which is the non-symmetric case when $d = 2$) in [16]*

and in the literature are also true for symmetric testers for any d. We wonder if those results may lead to other applications.

2.7 Lower Bounds

In this subsection we give some lower bounds for the size of testers.

In Theorem 21 we have proved that for any $q < d+1$ and t we have

$$\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \leq \left(1 + \frac{\ln(q+1)}{q} + O\left(\frac{\log q}{q^2}\right)\right)^d \cdot t = 2^{O\left(\frac{\log q}{q}\right)d} \cdot t.$$

We now give the following lower bound

Theorem 27. *For any q, d and t we have*

$$\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \geq \left(1 + \frac{1}{q-1} - \frac{1}{(q-1)q^{t-1}}\right)^{d-1} \cdot t = 2^{\Omega\left(\frac{1}{q}\right)d} \cdot t.$$

Proof. Consider the class of functions

$$\mathcal{M} = \left\{ \left(\prod_{i=1}^{d-1} \sum_{j=1}^{t} \lambda_{i,j} y_{i,j} \right) (y_{d,k_1} - y_{d,k_2}) \;\middle|\; (\lambda_{i,j})_j \in P^t(\mathbb{F}_q) \text{ for all } i = 1, \ldots, d-1, 1 \leq k_1 < k_2 \leq q^t \right\},$$

where $P^t(\mathbb{F}_q)$ is the t-dimensional projective space over $\mathbb{F}_q$. For $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \ldots, \boldsymbol{\lambda}_{d-1}) \in P^t(\mathbb{F}_q)^{d-1}$ we will denote $f_{\boldsymbol{\lambda}} = \prod_{i=1}^{d-1}(\sum_{j=1}^{t} \lambda_{i,j} y_{i,j})$. Let $\mathcal{M}' = \{(y_{d,k_1} - y_{d,k_2}) \mid 1 \leq k_1 < k_2 \leq q^t\}$.

Obviously, $\mathcal{M} \subseteq \mathcal{DML}(\mathbb{F}_q, n, d)$. Let $L = \{\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_\nu\}$ be a $(\mathcal{DML}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester with minimal size $\nu = \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$. Then it is a tester for $\mathcal{M}$. Let α be a primitive root in $\mathbb{F}_{q^t}$ and consider the assignments $\boldsymbol{z}_i = (\alpha^0, \alpha^1, \ldots, \alpha^{t-1}, 0, \ldots, 0) \in \mathbb{F}_{q^t}^{q^t}$ for all $i = 1, 2, \ldots, d-1$ and $\boldsymbol{z}_d = (0, \alpha^0, \alpha^1, \ldots, \alpha^{q^t-2}) \in \mathbb{F}_{q^t}^{q^t}$ and $\boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_d)$. Let $\boldsymbol{c}^{(i)} = \boldsymbol{\ell}_i(\boldsymbol{z}) \in (\mathbb{F}_q^{q^t})^d$ for $i = 1, \ldots, \nu$ and $C = \{\boldsymbol{c}^{(i)} \mid i = 1, 2, \ldots, \nu\}$. Since $f(\boldsymbol{z}) \neq 0$ for all $f \in \mathcal{M}$ and L is a tester for $\mathcal{M}$, for every $f \in \mathcal{M}$ there is $\boldsymbol{c} \in C$ such that $f(\boldsymbol{c}) \neq 0$. That is, C is a hitting set for $\mathcal{M}$.

Notice that if for some $\boldsymbol{c} \in C$ we have $(c_{i,1}, c_{i,2}, \ldots, c_{i,t}) = 0$ for some $i = 1, 2, \ldots, d-1$ then $f(\boldsymbol{c}) = 0$ for all $f \in \mathcal{M}$ and then $C \backslash \{\boldsymbol{c}\}$ is a hitting set for $\mathcal{M}$. Therefore we may assume w.l.o.g that $(c_{i,1}, c_{i,2}, \ldots, c_{i,t}) \neq 0$ for all $\boldsymbol{c} \in C$ and $i = 1, 2, \ldots, d-1$.

Now for every $\boldsymbol{\lambda} \in P^t(\mathbb{F}_q)^{d-1}$ consider the set $C_{\boldsymbol{\lambda}} = \{\boldsymbol{c} \in C \mid f_{\boldsymbol{\lambda}}(\boldsymbol{c}) \neq 0\}$. It is easy to see that $C_{\boldsymbol{\lambda}}$ is a tester for $\mathcal{M}'$. Then for every $1 \leq k_1 < k_2 \leq q^t$ there is $\boldsymbol{c} \in C_{\boldsymbol{\lambda}}$ such that $c_{d,k_1} \neq c_{d,k_2}$. Therefore $|C_{\boldsymbol{\lambda}}| \geq \log q^t / \log q = t$. Now it is easy to see that since $(c_{i,1}, c_{i,2}, \ldots, c_{i,t}) \neq 0$ for all $\boldsymbol{c} \in C$ and

$i = 1, 2, \ldots, d-1$, every $\boldsymbol{c} \in C$ appears in exactly

$$\left(\frac{q^t-1}{q-1} - \frac{q^{t-1}-1}{q-1}\right)^{d-1} = \left(\frac{q^t-q^{t-1}}{q-1}\right)^{d-1}$$

of the $C_{\boldsymbol{\lambda}}$s. Therefore

$$\begin{aligned}
\nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = \nu = |C| &\geq \frac{\sum_{\boldsymbol{\lambda}} |C_{\boldsymbol{\lambda}}|}{\left(\frac{q^t-q^{t-1}}{q-1}\right)^{d-1}} \\
&\geq \frac{|P^t(\mathbb{F}_q)^{d-1}| \cdot t}{\left(\frac{q^t-q^{t-1}}{q-1}\right)^{d-1}} = \frac{\left(\frac{q^t-1}{q-1}\right)^{d-1} \cdot t}{\left(\frac{q^t-q^{t-1}}{q-1}\right)^{d-1}} \\
&= \left(1 + \frac{1}{q-1} - \frac{1}{(q-1)q^{t-1}}\right)^{d-1} t.
\end{aligned}$$

□

For $\nu^{\mathcal{P}}$ and $\nu^{\mathcal{HP}}$, we first prove that there is no tester for $\mathcal{P}(\mathbb{F}_q, n, d)$ when $q \leq d$ and no tester for $\mathcal{HP}(\mathbb{F}_q, n, d)$ when $q \leq d-1$.

Lemma 28. *For $q \leq d$ we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = \infty$ and for $q \leq d-1$ we have $\nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = \infty$.*

Proof. Let $\mathbb{F}_q = \{\alpha_1, \ldots, \alpha_q\}$ and consider the polynomial $f(x) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_q)$. Let $\beta \in \mathbb{F}_{q^t}$ such that $f(\beta) \neq 0$. Since $f(\ell(\beta)) = 0$ for all $\ell : \mathbb{F}_{q^t} \to \mathbb{F}_q$ the first result follows.

For the second result we take $f(x_1, x_2) = x_2(x_1 - \alpha_1 x_2)(x_1 - \alpha_2 x_2)\cdots(x_1 - \alpha_q x_2)$. □

In Corollary 17 we have shown that for $q \geq d+1$ we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(d^{\tau(d,q,t)}t)$ where for any constant $c > 1$

$$\tau(d,q,t) = \begin{cases} 1 & \text{if } \ q \text{ perfect square }, q \geq c(d+1)^2, \text{ I.S. } t \\ 2 & \text{if } \ q \text{ perfect square }, q \geq c(d+1)^2 \\ 2 & \text{if } \ q \geq c(d+1), \text{ I.S. } t \\ 3 & \text{if } \ q \geq c(d+1) \\ 3 & \text{if } \ q \geq d+1, \text{ I.S. } t \\ 4 & \text{if } \ q \geq d+1 \end{cases}$$

where I.S. stands for "infinite sequence of". In the following Theorem we give the lower bound $\tau(d,q,t) \geq 1$, which is tight for perfect square q, $q \geq c(d+1)^2$ and infinite sequence of t.

Theorem 29. *For any $q \geq d+1$ and t we have*

$$\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \geq \nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) \geq dt - d + 1.$$

In particular, $\tau(d, q, t) \geq 1$.

Proof. Consider the set

$$\mathcal{M} = \left\{ \prod_{k=1}^{d} (x_j - x_{i_k}) \;\middle|\; 1 \leq i_1 < i_2 < \cdots < i_d \leq q^t, 1 \leq j \leq q^t, j \notin \{i_1, i_2, \ldots, i_d\} \right\}.$$

Obviously, $\mathcal{M} \subseteq \mathcal{HP}(\mathbb{F}_q, n, d)$. Let $L = \{\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_\nu\}$ be a $(\mathcal{HP}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester with minimal size $\nu = \nu^{\mathcal{HP}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$. By Lemma 3, L is a $(\mathcal{M}, \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester. Let α be a primitive root in $\mathbb{F}_{q^t}$ and consider the assignment $\boldsymbol{z} = (0, \alpha^0, \alpha^1, \ldots, \alpha^{q^t-2}) \in \mathbb{F}_{q^t}^{q^t}$. Let $\boldsymbol{c}_i = \boldsymbol{\ell}_i(\boldsymbol{z})$ for $i = 1, \ldots, \nu$ and $C = \{\boldsymbol{c}_i \mid i = 1, 2, \ldots, \nu\}$. Since $f(\boldsymbol{z}) \neq 0$ for all $f \in \mathcal{M}$ and L is a tester for $\mathcal{M}$, for every $f \in \mathcal{M}$ there is $\boldsymbol{c} \in C$ such that $f(\boldsymbol{c}) \neq 0$. That is, for every $1 \leq j \leq q^t$ and every $1 \leq i_1 < i_2 < \cdots < i_d \leq q^t$ such that $j \notin \{i_1, i_2, \ldots, i_d\}$ there is $\boldsymbol{c} \in C$ such that $c_j \notin \{c_{i_1}, \ldots, c_{i_d}\}$. Such set is called $(\nu; q^t, q, \{1, d\})$-separating hash family [75, 76]. See Subsection 4.4 in this paper. In [22], Bazrafshan and van Trang proved that

$$q^t \leq dq^{\lceil \frac{\nu}{d} \rceil}.$$

See also [17]. Therefore

$$\left\lceil \frac{\nu}{d} \right\rceil \geq t - \frac{\log d}{\log q}.$$

If $q \geq d + 1$ then $\lceil \nu/d \rceil \geq t$ and $\nu \geq dt - d + 1$. □

Our last result in this subsection gives a lower bound for the size of symmetric tester. We note that this result is subsumed by Theorem 29 but uses different algebraic technique that is used in [16]. We prove

Theorem 30. *If* $L = \{\ell_1, \ldots, \ell_\nu\}$ *is a symmetric* $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$*-tester for all* n *then*

$$\nu \geq dt - d + 1.$$

In particular, $\tau^*(d, q, t) \geq 1$.

Proof. We prove the result by induction on d. The case $d = 1$ is trivial. Suppose the lower bound is true for $d - 1$. By Lemma 24 and Lemma 26 there is a symmetric $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester $\tilde{L} = \{\tilde{\ell}_1, \ldots, \tilde{\ell}_\nu\} \subseteq \mathbb{F}_{q^t}^*$ and $\beta_i \in \mathbb{F}_{q^t}$, $i = 1, \ldots, \nu$ such that for every $a_1, a_2, \ldots, a_d \in \mathbb{F}_{q^t}$,

$$a_1 a_2 \cdots a_d = \sum_{i=1}^{\nu} \beta_i \tilde{\ell}_i(a_1) \tilde{\ell}_i(a_2) \cdots \tilde{\ell}_i(a_d).$$

Define the linear function $L : \mathbb{F}_{q^t} \to \mathbb{F}_q^{t-1}$ where $L(a) = (\tilde{\ell}_1(a), \ldots, \tilde{\ell}_{t-1}(a))$. Since by the rank-nullity theorem $\dim \ker L \geq 1$, there is a non-zero element $b \in \mathbb{F}_{q^t} \backslash \{0\}$ such that $L(b) = 0$. Thus $\tilde{\ell}_i(b) = 0$ for $i = 1, \ldots, t-1$. Now for all $a_1, \ldots, a_{d-1} \in \mathbb{F}_{q^t}$,

$$\begin{aligned} a_1 a_2 \cdots a_{d-1} &= (a_1 a_2 \cdots a_{d-1} b) b^{-1} \\ &= \left(\sum_{i=1}^{\nu} \beta_i \tilde{\ell}_i(a_1) \tilde{\ell}_i(a_2) \cdots \tilde{\ell}_i(a_{d-1}) \tilde{\ell}_i(b) \right) b^{-1} \\ &= \sum_{i=t}^{\nu} \gamma_i \tilde{\ell}_i(a_1) \tilde{\ell}_i(a_2) \cdots \tilde{\ell}_i(a_{d-1}) \end{aligned}$$

where $\gamma_i = \beta_i \tilde{\ell}_i(b) b^{-1}$ for $i = t, t+1, \ldots, \nu$. Therefore $L = \{\ell_t, \ldots, \ell_\nu\}$ is a symmetric $(\mathcal{P}(\mathbb{F}_q, n, d-1), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester. By the induction hypothesis we have $\nu - (t-1) \geq (d-1)t - (d-1) + 1$ and therefore $\nu \geq dt - d + 1$. □

We end this subsection with some open problems

Open Problems 4.

1. *For $q < d+1$, close the gap between the lower bound $2^{\Omega\left(\frac{1}{q}\right)d} \cdot t$ in Theorem 27 and the upper bound $2^{O\left(\frac{\log q}{q}\right)d} \cdot t$ in Theorem 21.*

2. *For $q \geq d+1$ the lower bounds in Theorem 29 and Theorem 30 matches the upper bounds in Corollary 17 within at most a factor of $O(1)$ to $O(d^3)$. It is interesting to close those gaps.*

3 Constructing Testers in Polynomial Time

In this subsection we show that testers of almost optimal size can be constructed in polynomial time.

3.1 Time Complexity of Constructing Irreducible Polynomials and $\mathbb{F}_{q^t}$

In some applications the construction of irreducible polynomials of degree n over $\mathbb{F}_q$ and the construction of the field $\mathbb{F}_{q^t}$ is also needed and their complexity must be included in the overall time complexity of the problem.

To construct the field $\mathbb{F}_{q^t}$ one should construct an irreducible polynomial $f(x)$ of degree t in $\mathbb{F}_q[x]$ and then use the representation $\mathbb{F}_{q^t} = \mathbb{F}_q[x]/(f(x))$. For a comprehensive survey on this problem see [64] Chapter 3. See also [4, 29, 61]. We give here the results that will be used in this paper.

Lemma 31. *Let $\mathbb{F}_q$ be a field of characteristic p. There is an algorithm that constructs an irreducible polynomial of degree t with T arithmetic operations in the field $\mathbb{F}_q$ where T is as described in the following table.*

Type	**Field**	**Assumption**	*Time* $= T$	**Poly**
Probabilistic	*Any*	–	$O\left(t^2 \log^{2+\epsilon} t + t \log q \log^{1+\epsilon} t\right)$	$= poly(t, \log q)$
Deterministic	*Any*	–	$O\left(p^{1/2+\epsilon} t^{3+\epsilon} + (\log q)^{2+\epsilon} t^{4+\epsilon}\right)$	$= poly(p, t, \log q)$
Deterministic	*Any*	*ERH*	$O(\log^2 q + t^{4+\epsilon} \log q)$	$= poly(t, \log q)$
Deterministic	$\mathbb{F}_2$	–	$O(t^{3+\epsilon})$	$= poly(t)$

Here ERH stands for the Extended Riemann Hypothesis and ϵ is any small constant.

In some of the applications it is enough to construct an extension field of dimension close to t. The following lemma is proved in [58]

Lemma 32. *There is a deterministic algorithm that constructs an irreducible polynomial of degree d where $t \leq d \leq t \log q$ with $poly(t, \log q)$ arithmetic operations in the field $\mathbb{F}_q$.*

One constraint that follows from using finite fields as an alphabet is that the size of the alphabet must be a power of prime. Shparlinski showed in [63] that

Lemma 33. *for any q large enough one can construct a finite field $\mathbb{F}_Q$ with $Q = q + o(q)$ elements in deterministic time $poly(\log q)$.*

In Lemma 8 one should construct many irreducible polynomials of certain degree. We now prove

Lemma 34. *There is a deterministic algorithm that runs in time $m \cdot poly(t, p, \log q)$ (and $m \cdot poly(t, \log q)$ assuming ERH) and construct m distinct irreducible polynomials of degree t in $\mathbb{F}_q[x]$.*

Proof. By Lemma 31, $\mathbb{F}_{q^t}$ can be constructed in polynomial time. It is known that a normal basis $\{\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{t-1}}\}$ in $\mathbb{F}_{q^t}$ can be constructed in $poly(t, \log q)$ time [54]. For any $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_t) \in \mathbb{F}_q^t$,

$$\beta_{\boldsymbol{\lambda}} = \lambda_1 \alpha + \lambda_2 \alpha^q + \lambda_3 \alpha^q + \cdots + \lambda_{t-1} \alpha^{q^{t-1}}$$

is a root of an irreducible polynomial of degree t if and only if $\beta_{\boldsymbol{\lambda}}, \beta_{\boldsymbol{\lambda}}^q, \beta_{\boldsymbol{\lambda}}^{q^2}, \ldots, \beta_{\boldsymbol{\lambda}}^{q^{t-1}}$ are distinct. It is easy to see that this is true if and only if the vectors

$$\boldsymbol{\lambda}^0 := \boldsymbol{\lambda},\ \boldsymbol{\lambda}^1 := (\lambda_t, \lambda_1, \ldots, \lambda_{t-1}),\ \boldsymbol{\lambda}^2 := (\lambda_{t-1}, \lambda_t, \lambda_1, \ldots, \lambda_{t-2}), \cdots, \boldsymbol{\lambda}^{t-1} := (\lambda_2, \lambda_3, \ldots, \lambda_t, \lambda_1)$$

are distinct. Such $\boldsymbol{\lambda}$ is called a vector of period t.

If we have a vector $\boldsymbol{\lambda}$ of period t then $\beta_{\boldsymbol{\lambda}}$ is a root of irreducible polynomial $f_{\beta_{\boldsymbol{\lambda}}}(x)$ of degree t and $f_{\beta_{\boldsymbol{\lambda}}}(x) \equiv (x - \beta_{\boldsymbol{\lambda}})(x - \beta_{\boldsymbol{\lambda}}^q) \cdots (x - \beta_{\boldsymbol{\lambda}}^{q^{t-1}})$. Notice that $f_{\beta_{\boldsymbol{\lambda}}}(x)$ can be computed in polynomials time

$poly(\log q, t)$. Therefore, it remains to construct m vectors of period t that generates m distinct irreducible polynomial.

Now choose any total order $<$ on $\mathbb{F}_q$ and consider the lexicographic order in $\mathbb{F}_q^t$ with respect to $<$ and consider the sequence of all the elements of $\mathbb{F}_q^t$ with this order. It is easy to see that for any two consecutive elements $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{F}_q^t$ in this sequence there is at least one $\boldsymbol{\lambda}_i$, $i \in \{1, 2\}$ of period t. Also, each irreducible polynomial $f_{\beta_{\boldsymbol{\lambda}}}$ of degree t can be constructed by exactly t elements (i.e., $\boldsymbol{\lambda}^0, \boldsymbol{\lambda}^1, \ldots, \boldsymbol{\lambda}^{t-1}$) in the sequence. This implies that the first $2tm$ elements in this sequence generate at least m distinct irreducible polynomials. □

Throughout this paper, when we say polynomial time or write $poly(t, p, \log q)$ we mean $poly(t, p, \log q)$ without any assumption and $poly(t, \log q)$ assuming ERH.

3.2 Preliminary Results

In this subsection we give some preliminary results

For $q \geq d+1$ we define $\tau_{poly}(d, q, t)$ the constant for which a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^{\tau_{poly}(d,q,t)} \cdot t)$ can be constructed in deterministic polynomial time. Since some of the applications in this paper require testers for fields of logarithmic dimension we also define $\tau_{poly}(d, q, t, r)$ the constant for which a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^r}, \mathbb{F}_q)$-tester of size $O(d^{\tau_{poly}(d,q,t,r)} \cdot r)$ can be constructed in deterministic time $poly(t, p, \log q)$.

In Lemma 7 we have shown that when $q \geq d(t-1)+1$, then a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester can be constructed in time complexity that is equal to the time complexity of constructing the field $\mathbb{F}_{q^t}$. Therefore, such tester can be constructed in polynomial time $poly(t, p, \log q)$. Hence, we may assume that

$$d + 1 \leq q \leq d(t-1).$$

Notice that in this case, $poly(t, p, \log q) = poly(d, q, t) = poly(t)$.

The results in the following lemma follow immediately from Lemma 7, Corollary 5 and the proof of Corollary 17

Lemma 35. *We have*

1. *For $q \geq d(t-1)+1$ we have $\tau_{poly}(d, q, t) = 1$.*
2. *For $q \geq d+1$ we have $\tau_{poly}(d, q, t) \leq \tau_{poly}(d, q^2, t) + 1$.*
3. *For any $c > 1$, $q \geq c(d+1)^2$, q perfect square, any integer r and $t = \lfloor (\sqrt{c}-1) q^{(r+2)/2} / \sqrt{c} d \rfloor$ we have $\tau^*(d, q, t) = 1$.*

We now prove the following lemma. The construction in this lemma is computationally expensive in its own right, but with the parameters we will be using it, it will take time polynomial in the parameters of the main problem.

Lemma 36. *A $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^\tau t)$ can be constructed in poly$(q^{d^\tau t^2})$ time where $\tau = \tau^*(d, q, t)$.*

Proof. By Corollary 23 there is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^\tau t)$. By Lemma 25 we have $L = \{\ell_1, \ldots, \ell_\nu\}$ is a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester if and only if $\ell_i(1) = 1$ for all i and there are $\beta_i \in \mathbb{F}_{q^t}$ such that for every $\boldsymbol{a} \in \mathbb{F}_{q^t}^d$ we have

$$a_1 a_2 \cdots a_d = \sum_{i=1}^{\nu} \beta_i \ell_i(a_1) \ell_i(a_2) \cdots \ell_i(a_d). \tag{21}$$

Then by Lemma 26 we may assume without loss of generality that $\ell_i \in \mathbb{F}_{q^t}^*$ and $\ell_i(1) = 1$ for all $i = 1, 2, \ldots, \nu$.

Now to construct a symmetric and reducible tester we can exhaustively search for one. That is, we can try all possible $\beta_i \in \mathbb{F}_{q^t}$ and $\ell_i \in \mathbb{F}_{q^t}^*$ where $\ell_i(1) = 1$, $i = 1, 2, \ldots, \nu$ and check if (21) is true for all $\boldsymbol{a} \in \mathbb{F}_{q^t}^d$. The number of all possible testers of size ν is at most

$$\binom{|\mathbb{F}_{q^t}|}{\nu} \cdot \binom{|\mathbb{F}_{q^t}^*|}{\nu} = \binom{q^t}{\nu}^2 \leq q^{2\nu t}.$$

Checking each tester takes $|\mathbb{F}_{q^t}^d| = q^{dt}$ substitutions in (21). Therefore, for $\tau = \tau^*(d, q, t)$, with at most $q^{O(d^\tau t^2)} q^{dt} = q^{O(d^\tau t^2)}$ arithmetic operations in the field $\mathbb{F}_{q^t}$ one can find a symmetric and reducible $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester. This implies the result. □

In particular we have

Lemma 37. *For every $q \geq d + 1$ and $\tau = \tau^*(d, q, r)$ if*

$$r \leq \sqrt{\frac{\log t}{d^\tau \log q}}$$

then $\tau_{poly}(d, q, t, r) \leq \tau^*(d, q, r)$.

3.3 Reductions of the Problem

In this subsection we show that a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester can be reduced in deterministic polynomial time to a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^{t'}}, \mathbb{F}_q)$-tester where t' is logarithmic in t. This reduction blows up the size of the tester by at most a factor of d.

Our main technique for constructing testers in polynomial time is to reduce the dimension t of the field in the tester to a small dimension $r = o(t)$ (Lemma 8) and then construct the tester for dimension r in time polynomial in t. Formally,

Lemma 38. *For any $t \leq t'$ and any $r = O(\log_q t')$*

$$r \geq \left\lceil \frac{\log(dt)}{\log q} \right\rceil + 1$$

(or $q^{r-1} \geq dt - d + 1$) we have

$$\tau_{poly}(d, q, t', t) \leq \tau_{poly}\left(d, q, t', r\right) + 1.$$

Proof. We use Lemma 8. Since $r \geq \lceil \log(dt)/\log q \rceil + 1$ we have $q^{r-1} \geq dt - d + 1$. To use the construction in Lemma 8 we need to find $O(dt/r)$ polynomials of degree r. Since $r = O(\log_q t')$ we have $|\mathbb{F}_{q^r}| = poly(t')$. Therefore the field $\mathbb{F}_{q^r}$ can be constructed in $poly(t')$ time and a primitive root of the field and therefore $O(dt/r)$ irreducible polynomials of degree r over $\mathbb{F}_q$ can be constructed in $poly(t')$ time.

If there is a deterministic $poly(t', p, \log q)$ time construction algorithm for a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^r}, \mathbb{F}_q)$-tester of size $O(d^c r)$ then the construction in Lemma 8 is deterministic $poly(t', p, \log q)$ time construction for $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size

$$\left(\frac{dt - d + 1}{r} + 1\right) \cdot O(d^c r) \leq O(d^{c+1} t).$$

This implies the result. □

3.4 Testers for $q \geq d + 1$ in Polynomial Time

In this subsection we prove that for every t, d and $q \geq d+1$ a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^5 \cdot t)$ can be constructed in deterministic polynomial time.

We first prove

Lemma 39. *For every t, constant $c > 1$, perfect square $q \geq c(d+1)^2$ and $d \geq 3$ a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^3 t)$ can be constructed in deterministic polynomial time.*

Proof. By *1* in Lemma 35, for $t \leq d + 2$ we have $q \geq d(t-1) + 1$ and therefore $\tau_{poly}(d, q, t) = 1$. By Lemma 38, for $d + 3 \leq t \leq q^{d+1}/d$ we have

$$\tau_{poly}(d, q, t) \leq \tau_{poly}(d, q, t, d + 2) + 1 \leq 2.$$

Then again by Lemma 38, for $q^{d+1}/d \le t \le q^{dq^d}$ and $r_0 = \lfloor q^{d+1}/d \rfloor - 1$ we have

$$\tau_{poly}(d,q,t) \le \tau_{poly}(d,q,t,r_0) + 1 \le 3.$$

Therefore, we may assume that

$$t \ge t_0 := q^{dq^d}.$$

Let

$$r_1 = \left\lceil \frac{\log dt}{\log q} \right\rceil + 1. \tag{22}$$

By Lemma 38, we have

$$\tau_{poly}(d,q,t) \le \tau_{poly}(d,q,t,r_1) + 1. \tag{23}$$

Let r be an integer such that

$$\frac{(\sqrt{c}-1)q^{(r+2)/2}}{\sqrt{c}d} \le \sqrt{\frac{\log t}{d \log q}} < \frac{(\sqrt{c}-1)q^{(r+3)/2}}{\sqrt{c}d} \tag{24}$$

and

$$r_2 := \left\lfloor \frac{(\sqrt{c}-1)q^{(r+2)/2}}{\sqrt{c}d} \right\rfloor. \tag{25}$$

By *3* in Lemma 35,

$$\tau^*(d,q,r_2) = 1. \tag{26}$$

In Claim 1 below we prove that $q^{r_2-1} \ge dr_1 - d + 1$. By Lemma 38, since $q^{r_2-1} \ge dr_1 - d + 1$ we have

$$\tau_{poly}(d,q,t,r_1) \le \tau_{poly}(d,q,t,r_2) + 1. \tag{27}$$

Finally, by Lemma 37, (24), (25) and (26) we have

$$\tau_{poly}(d,q,t,r_2) \le \tau^*(d,q,r_2) = 1.$$

This with (27) and (23) gives the result. □

It remains to prove

Claim 1. *We have* $q^{r_2-1} \ge dr_1 - d + 1$.

Proof. Since $t \geq t_0 \geq q^{dq^d}$ we have $r_1 \geq \log_q t \geq dq^d$. By (22) we have $r_1 \leq \log_q t + 3$ and therefore $t \geq q^{r_1-3}$. By (24) and (25) we have

$$\begin{aligned} r_2 &\geq \frac{(c-1)q^{(r+2)/2}}{cd} - 1 \geq \frac{1}{q^{1/2}} \frac{(c-1)q^{(r+3)/2}}{cd} - 1 \\ &\geq \sqrt{\frac{\log t}{qd \log q}} - 1 \geq \sqrt{\frac{r_1 - 3}{qd}} - 1. \end{aligned}$$

Therefore

$$q^{r_2-1} \geq q^{\sqrt{\frac{r_1-3}{qd}}-2}$$

and it is enough to prove that for $r_1 \geq dq^d$,

$$q^{\sqrt{\frac{r_1-3}{qd}}-2} \geq dr_1.$$

This is true since $d \geq 3$ and $q \geq (d+1)^2$.

□

We now prove

Theorem 40. *For every n, t, d and $q \geq d+1$, a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^5 t)$ can be constructed in polynomial time.*

Proof. If $q \geq d+1$ then $q^4 \geq 2(d+1)^2$. Therefore by Lemma 39, $\tau_{poly}(d, q^4, t) \leq 3$. Now by *2* in Lemma 35 we have

$$\tau_{poly}(d, q, t) \leq \tau_{poly}(d, q^2, t) + 1 \leq \tau_{poly}(d, q^4, t) + 2 \leq 5.$$

□

By Corollary 23 and the proof of Theorem 40 we also get

Corollary 41. *Let $c > 1$ be a constant and $r = poly(d, \log t)$. We have the following upper bounds for*[3] *$\tau_{poly}(d, q, t, r)$ and $\tau_{poly}(d, q, t)$*

[3]Here and elsewhere we will add to the table other columns so that the reader can compare the result with the upper and lower bounds.

q	t	*Upper B.* $\tau^*(d,q,t)$	*Upper B.* $\tau_{poly}(d,q,t,r)$	*Upper B.* $\tau_{poly}(d,q,t)$	*Lower B.* $\tau(d,q,t)$
$q \geq c(d+1)^2$, *q P.S.*	*I.S.*	1	2	3	1
$q \geq c(d+1)^2$ *q P.S.*	*all*	2	2	3	1
$q \geq c(d+1)$	*I.S.*	2	3	4	1
$q \geq c(d+1)$	*all*	3	3	4	1
$q \geq d+1$	*I.S.*	3	4	5	1
$q \geq d+1$	*all*	4	4	5	1

where I.S. stands for "for infinite sequence of integers t" and P.S. for "perfect square".

3.5 Testers for Subspaces of Fields in Polynomial Time

In this subsection we discuss the explicit construction in Subsection 2.4 of testers for subspaces of fields and what is needed in order to get a polynomial time construction. We then show that testers for subspaces of extension fields of logarithmic dimension can be constructed in polynomial time. Those testers will have better size than the testers in the previous subsection and will close the gaps in the table in Corollary 41 for $\tau_{poly}(d,q,t,r)$ when $r = O(\log t/\log q)$.

Our goal is to construct the tester in Corollary 20 and Lemma 19 in polynomial time. For the construction in Corollary 20 and Lemma 19, we need to construct the $(\mathcal{P}(\mathbb{F}_q,n,d),\mathscr{L}(G),\mathbb{F}_q)$-tester defined in Lemma 12 and the $(\mathbb{F}_q[\boldsymbol{x}],S,\mathscr{L}(G))$-tester defined in Lemma 18. To construct those testers we need

1. To find all the places of degree 1 in $F^{(r+1)}/\mathbb{F}_q$ for the tower $\mathcal{W}_1$ defined in Lemma 16, [36, 37], where $q^{r+1} < (c/(c-1))dt \leq q^{r+2}$. By Lemma 16, the number of such places is less than $q^{r+2} = O(dqt) = poly(t)$.
2. To find a prime divisor Q of $F^{(r+1)}/\mathbb{F}_q$ of degree $t+g$. By Lemma 16, $g \leq q^{r+1} = O(dt)$.
3. To find a basis for $\mathscr{L}((t+g-1)P_\infty)$ for some devisor P_∞ of degree 1.

Notice that

$$r = \Theta\left(\frac{\log t}{\log q}\right).$$

For 1 and 3, it is known from [65] and [69] that all the places of degree 1, a place P_∞ and the basis for $\mathscr{L}((t+g-1)P_\infty)$ can be found in time $poly(t+g) = poly(t)$.

For 2, we need to construct a prime divisor of degree $t+g$. If Q is a prime divisor of degree s over $\mathbb{F}_q$, then there exist conjugate prime divisors $Q_1, Q_2, \ldots, Q_s$ of degree 1 over $\mathbb{F}_{q^s}$ such that $Q = \sum_{i=1}^{s} Q_i$. Therefore we need to find a divisor Q_i of degree one in $F^{(r)}\mathbb{F}_{q^s}$ with s different conjugates. We do not

know whether this can be done in deterministic polynomial time when $s = t + g$ and $r = O(\log t/\log q)$. We will call this problem $\mathbf{PRIME}(s, r)$.

We now prove

Lemma 42. *The problem* $\mathbf{PRIME}(s, r)$ *can be solved in time* $poly(q^{s+r})$.

In particular, $\mathbf{PRIME}(O(\log t/\log q), O(\log t/\log q))$ *can be solved in polynomial time*

Proof. To solve $\mathbf{PRIME}(s, r)$ we can compute all the prime divisors of degree 1 in $F^{(r)}\mathbb{F}_{q^s}$ and its orbits under the Frobenius automorphism of $\mathbb{F}_{q^s}/\mathbb{F}_q$. Each orbit of length s corresponds to a prime divisor of degree s. To find a prime divisor of degree 1 in $F^{(r)}\mathbb{F}_{q^s}$ we solve the system of equation $x_k^q + x_k = x_{k-1}^q/(x_{k-1}^{q-1} + 1)$ for $k = 1, 2, \ldots, r$ and all $x_0 \in \mathbb{F}_{q^s}$. For a fixed x_{k-1} the system $x_k^q + x_k = x_{k-1}^q/(x_{k-1}^{q-1} + 1)$ has at most q solutions for x_k. To find the solutions we exhaustively search for them. This takes $|F_{q^s}| = q^s$ substitutions. Therefore, to find all solutions we need at most q^{2s+r} substitutions and therefore time $poly(q^{s+r})$. □

Obviously, when $s + r = \omega(\log t/\log q)$, the above is not polynomial time construction. Therefore, as we did in the last subsection we need to reduce the dimension of the problem to $\log t/\log q$ and then use Lemma 42. Unfortunately, it is not clear here how to reduce the dimension. Our reduction in Lemma 8 does not seems to be working in this case.

If the dimension of the extension field is $r = O(\log t/\log q)$ then we need to solve $\mathbf{PRIME}(s, r)$ for $s = O(\log(\log t/\log q)/\log q)$ and $r = O(\log t/\log q) + g$ where g is as in the proof of Corollary 20. If $g = O(\log t/\log q)$ then by Lemma 42, $\mathbf{PRIME}(s, r)$ (and therefore item 2) can be solved in deterministic polynomial time. For the cases *4-6* in Corollary 20, $r + g = O(\log t/\log q)$. See the proof of Corollary 20. This implies

Corollary 43. *Let* $c > 1$ *be a constant and* $r = O(\log t/\log q)$. *There is a polynomial time algorithm that constructs a subspace* $S \subseteq \mathbb{F}_{q^R}$ *of size* $|S| = q^r$ *where* $R \leq O(r)$ *and a* $(\mathcal{P}(\mathbb{F}_q, n, d), S, \mathbb{F}_q)$*-tester of size* $O(d^\tau r)$ *where* τ *is as indicated in the following table (the forth column in the table)*

q	t	*Upper B.* $\tau^*(d, q, t)$	*Upper B.* τ	*Lower B.* $\tau(d, q, t)$
$q \geq c(d+1)^2$, *q P.S.*	*I.S.*	1	1	1
$q \geq c(d+1)$	*I.S.*	2	2	1
$q \geq d+1$	*I.S.*	3	3	1

where I.S. stands for "for infinite sequence of integers t" and P.S. for "perfect square".

Notice that $\tau_{poly}(d, q, t, r)$ in Corollary 41 with τ in Corollary 43 give bounds that meet the bounds for $\tau^*(d, q, t)$ when the dimension is $O(\log t/\log q)$.

3.6 Testers for $q < d+1$ in Polynomial Time

In this subsection we study $(\mathcal{DML}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-testers when $q < d+1$. The following result follows immediately from Lemma 31 and the testers constructed in Lemmas 2, 4, 6, 7, Theorem 21 and Theorem 40.

Theorem 44. *For any $q < d+1$ and t a $(\mathcal{DML}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^7 \cdot 2^{c_q d} \cdot t)$ can be constructed in time $2^{c_q d} \cdot poly(t)$ where*

$$c_q = \sum_{i=0}^{\infty} \frac{\log(q^{2^i}+1)}{q^{2^i}} = O\left(\frac{\log q}{q}\right).$$

We denote by $\sigma(q)$ (respectively, $\sigma_{poly}(q)$) the constant for which a $(\mathcal{DML}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester L of size $2^{\sigma(q)\cdot d + o(d)} \cdot t$ exists (respectively, and can be constructed in $2^{\sigma(q)\cdot d} \cdot poly(t)$ time). By Theorem 21 and Theorem 27 we have

$$O\left(\frac{\log q}{q}\right) = \sum_{i=0}^{\infty} \frac{\log(q^{2^i}+1)}{q^{2^i}} = c_q \geq \sigma_{poly}(q) \geq \sigma(q) \geq \log\left(1 + \frac{1}{q-1}\right) = \Omega\left(\frac{1}{q}\right).$$

In particular we have following bounds for $\sigma_{poly}(q)$

q	c_q Upper B.	Lower B.
2	1.659945821	1
3	1.116191294	0.584962501
4	0.867464571	0.415037499
5	0.719921672	0.321928095
7	0.548433289	0.222392421

We now give some open problems

Open Problems 5.

1. *We have shown that a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester of size $O(d^5 t)$ can be constructed in polynomial time. It is easy to see that a tester of size $O(d^4 t)$ can be constructed in quasi-polynomial time. Find a better construction.*

2. *See Appendix B. Prove Conjecture 1.*

3. *Is* **BASIS**(s, r) *solvable in deterministic polynomial time?*

3.7 Tester from any Field to any Field

In this subsection we study $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_{q^{t'}})$-testers when $q^{t'} \geq d+1$. Those testers will be used to reduce the dimension of the extension field used in the black box interpolation and identity testing to a smaller dimension. Such tester can be constructed using the fact that $\mathbb{F}_{q^t}$ can be embedded into the field $\mathbb{F}_{(q^{t'})^t} \cong \mathbb{F}_{(q^t)^{t'}}$ and then the $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{(q^{t'})^t}, \mathbb{F}_{q^{t'}})$-tester constructed in the previous subsections is also a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_{q^{t'}})$-tester. This tester is optimal within a factor of t'. Our goal in this subsection is to find such tester that is optimal within a factor of $poly(d)$.

One approach is to first construct a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_{q^T})$-tester such that $T > dt$ and $t'|T$. Then construct a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^T}, \mathbb{F}_{q^{t'}})$-tester and combine both testers. Another approach, that is used here, is to first construct a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_{q^T})$-tester where $t'|T$ and $T = O(\log(dt)/\log q)$. Then construct a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^T}, \mathbb{F}_{q^{t'}})$-tester and combine both testers.

We prove

Lemma 45. *Let $t' \geq 1$ be an integer such that $q^{t'} \geq d+1$. We have*

1. *If $t' \geq \lceil \log(dt)/\log q \rceil + 1$ then $\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_{q^t}) \leq (d+1)t/t'$ and a tester with such size can be constructed in polynomial time.*
2. *If $t' \geq \lceil \log(d\log(dt)/\log q)/\log q \rceil \geq 2$ then $\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_{q^t}) \leq d(d+1)t/t'$ and a tester with such size can be constructed in polynomial time.*
3. *For any t' we have $\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_{q^t}) \leq O(d^5 t/t')$ and a tester with such size can be constructed in polynomial time.*

Proof. We first prove *1.* As in the proof of Lemma 7, we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_q[X]_{t-1})$$

and by Lemma 9 since $q^{t'-1} \geq dt - d + 1$, we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_q[X]_{t-1}) \leq \frac{dt-d+1}{t'} + 1 \leq (d+1)\frac{t}{t'}.$$

This proves *1.*

We now prove *2.* Let $T = \lceil \log(dt)/\log q \rceil + c$ where $c \geq 1$ is an integer such that $t'|T$. Obviously $c \leq t'$. By (2) we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_{q^t}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_{q^T}}((d, \mathbb{F}_q), \mathbb{F}_q[X]_{t-1}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_{q^T}). \tag{28}$$

By *1*, we have $\nu^{\mathcal{P}}_{\mathbb{F}_{q^T}}((d,\mathbb{F}_q),\mathbb{F}_q[X]_{t-1}) \leq (d+1)t/T$ and since $q^{t'} \geq d(T/t'-1)$, by Lemma 7,

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d,\mathbb{F}_q),\mathbb{F}_{q^T}) = \nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d,\mathbb{F}_q),\mathbb{F}_{(q^{t'})^{T/t'}}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}(d,\mathbb{F}_{(q^{t'})^{T/t'}}) \leq d(T/t').$$

This with (28) implies *2*.

To prove *3*, we use *1*, (28) and Corollary 41. Since $T < poly(d, \log t)$, we get

$$\begin{aligned}
\nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d,\mathbb{F}_q),\mathbb{F}_{q^t}) &\leq \nu^{\mathcal{P}}_{\mathbb{F}_{q^T}}((d,\mathbb{F}_q),\mathbb{F}_q[X]_{t-1}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d,\mathbb{F}_q),\mathbb{F}_{q^T}) \\
&\leq \nu^{\mathcal{P}}_{\mathbb{F}_{q^T}}((d,\mathbb{F}_q),\mathbb{F}_q[X]_{t-1}) \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}(d,\mathbb{F}_{(q^{t'})^{T/t'}}) \\
&\leq (d+1)\frac{t}{T} \cdot O\left(d^4\frac{T}{t'}\right) = O\left(d^5\frac{t}{t'}\right).
\end{aligned}$$

□

4 Applications of Tester for d-Restriction Problems

In this section we give some applications of testers for d-restriction problems. A *d-restriction problem* [53, 5] is a problem of the following form:

Given an alphabet Σ of size $|\Sigma| = q$, an integer n and a class $\mathcal{M}$ of nonzero functions $f_i : \Sigma^d \to \{0,1\}$, $i = 1, 2, \ldots, t$.

Find a small set $A \subseteq \Sigma^n$ such that: For every $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ and $f \in \mathcal{M}$ there is $\boldsymbol{a} \in A$ such that $f(a_{i_1}, \ldots, a_{i_d}) \neq 0$.

We give applications of testers for the following four d-restriction problems: Perfect hash, universal set, cover-free family and separating hash family.

Note. For all the applications in this paper we will give deterministic polynomial time constructions, but in some cases some explicit constructions are also given. We remind the reader that when say an "explicit construction" we mean a construction using elementary algebra and algebraic function fields in which each step of the construction is indicated. But it is not clear whether the construction is polynomial time construction. For a construction that has exponential size, when we say that such construction can be constructed in "polynomial time" we mean in time $s \cdot poly(n)$ where s is the size of the construction.

4.1 Perfect Hash

In this subsection we show that testers can be used for constructing almost optimal perfect hashing in polynomial time.

Let H be a family of functions $h : [n] \to [q]$. We say that H is an (n, q)-*family of perfect hash functions* if for all subset $S \subseteq [n]$ of size $|S| = q$ there is an $h \in H$ such that the restriction of h to S, $h|_S$, is onto, i.e., $h|_S(S) = [q]$. In [53] Naor et. al. gave a deterministic construction of an (n, q)-family of perfect hash functions of size $s = e^q q^{O(\log q)} \log n$ that can be constructed in time $s \cdot poly(n)$.

For $d \leq q$ we say that H is an (n, q, d)-*perfect hashing* [5] (or (n, d, q)-*splitter* [53]) if for all subsets $S \subseteq [n]$ of size $|S| = d$ there is a *hash function* $h \in H$ such that $h|_S$ is injective (one-to-one) on S, i.e., $|h|_S(S)| = d$. Thus (n, q)-family of perfect hash functions is (n, q, q)-perfect hashing.

In [53, 5] it was shown that there are (n, d^2, d)-perfect hashing of size $O(d^4 \log d \log n)$ that can be constructed in $poly(n, d)$ time. Wang and Xing [81] used algebraic function fields and gave an explicit (n, d^4, d)-perfect hashing of size $O((d^2/\log d) \log n)$ for infinite sequence of integers n. Their construction is similar to the construction of $(\mathcal{P}(\mathbb{F}_q, n, d), \mathscr{L}(G), \mathbb{F}_q)$-tester. For any q the only known polynomial time construction is of size $O(d^2 \log d \log n)$, [53, 5]. Blackburn and Wild [23] gave an explicit optimal construction when q is very large compared to d and $\log n$.

Let $N(n, d, q)$ be the size of the smallest (n, q, d)-perfect hashing. Obviously, the problem of finding a small (n, q, d)-*perfect hashing* is a d-restriction problem where $\mathcal{M} = \{f\}$, $f : \Sigma^d \to \{0, 1\}$ and $f(x_1, x_2, \ldots, x_d) = 1$ if and only if $|\{x_1, \ldots, x_d\}| = d$. We will be interested in the case where $d = o(n)$. We now use union bound to give a nonconstructive upper bound

Lemma 46. *Let $q \geq d(d-1)/2 + 1$. Then*

$$N(n, q, d) \leq \frac{\log \binom{n}{d}}{\log\left(\frac{1}{1-g(q,d)}\right)} \leq \frac{d \log n}{\log \frac{2q}{d(d-1)}} \tag{29}$$

where $g(q, d) = (q-1)(q-2)\cdots(q-d+1)/q^d$.

In particular, when $q = \Theta(d^2)$ then $N(n, q, d) = O(d \log n)$ and when $q \geq d^{2+\epsilon}$ for some constant $\epsilon > 0$ then

$$N(n, q, d) \leq \frac{d \log n}{\log q}.$$

Proof. The bound (29) follows from union bound and the fact that

$$g(q, d) = \left(1 - \frac{1}{q}\right)\left(1 - \frac{2}{q}\right)\cdots\left(1 - \frac{d-1}{q}\right) \geq 1 - \frac{d(d-1)}{2q}.$$

When $q \geq d^{2+\epsilon}$ for some constant ϵ we have $N(n, q, d) = O(d \log n / \log q)$ and when $q = O(d)$ and $q > 2d$ then $N(n, q, d) = O(d \log n)$. When $d(d-1)/2 + 1 \leq q \leq 2d$ we have

$$-\ln g(q, d) = -\sum_{i=1}^{d-1} \ln\left(1 - \frac{i}{q}\right) = \sum_{i=1}^{d-1} \sum_{j=1}^{\infty} \frac{i^j}{j q^j} \geq \frac{d(d-1)}{2q} + O(d^3/q^2) \geq \frac{1}{4},$$

and therefore $g(q,d) < 0.78$ and $N(n,q,d) = O(d \log n)$. □

By [50] (see also [32, 41, 42]) we have the following lower bound

Lemma 47. *Let $q \geq d(d-1)/2 + 1$. Then*

$$N(n,q,d) \geq \frac{g(n,d-1)\log(n-d+2)}{g(q,d-1)\log(q-d+2)} = \Omega\left(\frac{\log n}{\log q}\right). \tag{30}$$

We now prove

Lemma 48. *Let $q > d(d-1)/2$ be a power of prime. Let t be an integer such that $q^t \geq n$. Let $S \subseteq \mathbb{F}_{q^T}$ be a sublinear space for some $T \geq t$ where $|S| = q^t$. There is an explicit (n,q,d)-perfect hashing of size*

$$\nu_{\mathbb{F}_q}^{\mathcal{HP}}\left(\frac{d(d-1)}{2}, S\right).$$

In particular, There is an explicit (n,q,d)-perfect hashing of size

$$\nu_{\mathbb{F}_q}^{\mathcal{HP}}\left(\frac{d(d-1)}{2}, \mathbb{F}_{q^t}\right).$$

In particular, for a constant $c > 1$, the following (n,q,d)-perfect hashing can be constructed in polynomial time

n	q	*poly time* *Size =*	*Union Bound*	*Lower Bound*
I.S.	$q \geq \frac{c}{4}d^4$	$d^2\frac{\log n}{\log q}$	$d\frac{\log n}{\log q}$	$\frac{\log n}{\log q}$
all	$q \geq \frac{c}{4}d^4$	$d^4\frac{\log n}{\log q}$	$d\frac{\log n}{\log q}$	$\frac{\log n}{\log q}$
I.S.	$q \geq \frac{c}{2}d^2$	$d^4\frac{\log n}{\log d}$	$d\frac{\log n}{\log(2q/(d(d-1)))}$	$\frac{\log n}{\log q}$
all	$q \geq \frac{c}{2}d^2$	$d^6\frac{\log n}{\log d}$	$d\frac{\log n}{\log(2q/(d(d-1)))}$	$\frac{\log n}{\log q}$
I.S.	$q \geq \frac{d(d+1)}{2} + 1$	$d^6\frac{\log n}{\log d}$	$d\log n$	$\frac{\log n}{\log q}$
all	$q \geq \frac{d(d+1)}{2} + 1$	$d^8\frac{\log n}{\log d}$	$d\log n$	$\frac{\log n}{\log q}$

Proof. Consider the set of functions

$$\mathcal{F} = \{\Delta_{\{i_1,\ldots,i_d\}}(x_1,\ldots,x_n) \mid 1 \leq i_1 < \cdots < i_d \leq n\}$$

in $\mathbb{F}_q[x_1, x_2, \ldots, x_n]$ where

$$\Delta_{\{i_1,\ldots,i_d\}}(x_1,\ldots,x_n) = \prod_{1\leq k<j\leq d}(x_{i_k} - x_{i_j}).$$

Consider n distinct elements $\alpha_1, \alpha_2, \ldots, \alpha_n \in S$ and let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$. Since $g(\boldsymbol{\alpha}) \neq 0$ for every $g \in \mathcal{F}$, and since $\mathcal{F} \subset \mathcal{HP}(\mathbb{F}_q, n, d(d-1)/2)$, using testers, there is a set $B \subseteq \mathbb{F}_q^n$ of size $\nu_{\mathbb{F}_q}^{\mathcal{HP}}(d(d-1)/2, S)$ such that for every $g \in \mathcal{F}$ there is $\boldsymbol{b} \in B$ where $g(\boldsymbol{b}) \neq 0$. This gives an (n, q, d)-perfect hashing.

Now, the results in the table follows from Corollaries 41 and 43. □

When $q > d(d+1)/2$ is not a power of prime number then we can take the nearest prime $q' < q$ and construct an (n, q', d)-perfect hashing that is also (n, q, d)-perfect hashing. It is known that the nearest prime $q' \geq q - \Theta(q^{.525})$, [19], and therefore the result in the above table is also true for any integer $q \geq d(d+1)/2 + O(d^{1.05})$.

4.2 (n, d)-Universal Set

An (n, d)-universal set over an alphabet Σ is a set $\mathcal{F} \subseteq \Sigma^n$ such that for every $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ and every $(\sigma_1, \ldots, \sigma_d) \in \Sigma^d$ there is $\boldsymbol{a} \in \mathcal{F}$ such that $a_{i_j} = \sigma_j$ for all $j = 1, \ldots, d$.

Let $|\Sigma| = q$. Let $U(n, d, q)$ be the size of the smallest (n, d)-universal set over the alphabet Σ. Obviously, finding a small (n, d)-universal set is a d-restriction problem.

The union bound shows that there is an (n, d)-universal set over an alphabet Σ of size

$$U(n, d, q) \leq dq^d \left(\ln \frac{n}{d} + \ln q \right) = O\left(dq^d \log n \right).$$

We first give a better bound when q is a power of prime

Lemma 49. *Let $q > 2$ be a power of prime. We have*

$$U(n, d, q) = O\left(d \frac{q^d}{\log q} \log n \right).$$

Proof. Randomly uniformly choose dr vectors $\boldsymbol{y}_1^{(j)}, \ldots, \boldsymbol{y}_d^{(j)} \in \mathbb{F}_q^n$, $j = 1, \ldots, r$ and take

$$\mathcal{F} = \bigcup_{j=1}^{r} \text{Span } \{\boldsymbol{y}_1^{(j)}, \cdots, \boldsymbol{y}_d^{(j)}\},$$

where Span is the linear span over $\mathbb{F}_q$. The set $\mathcal{F}$ is (n, d)-universal set over an alphabet $\mathbb{F}_q$ if for every $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ there is $j \leq r$ such that the matrix $Y_r = [y_{\ell, i_k}^{(j)}]_{\ell, k}$ is singular. The probability that Y_r is singular is

$$\left(1 - \frac{1}{q^n}\right)\left(1 - \frac{1}{q^{n-1}}\right) \cdots \left(1 - \frac{1}{q}\right) \geq 1 - \sum_{i=1}^{d} \frac{1}{q^i} \geq 1 - \frac{1}{q-1}.$$

Now we use union bound to get the result. □

For $q = 2$, a lower bound of $\Omega(2^d \log n)$ was proved in [44]. For completeness we prove the following lower bound using the techniques used in [44] and [66]

Lemma 50. *We have*

$$U(n, d, q) = \Omega\left(\frac{q^{d-1}}{\log q} \log n\right).$$

Proof. We first show that $t := U(n, 2, q) = \Omega(q \log n / \log q)$. Let $R = \{\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(t)}\}$ be an $(n, 2)$-universal set over $\mathbb{Z}_q$. It is easy to see that for any vector $\boldsymbol{v} \in \mathbb{Z}_q^n$ the set $R + \boldsymbol{v}$ is $(n, 2)$-universal set. For each entry $1 \leq i \leq n$ we choose v_i such that

$$|\{\boldsymbol{r} \in R \mid r_i + v_i = 0\}| \leq \frac{|R|}{q} = \frac{t}{q}.$$

Consider $S = R + \boldsymbol{v} = \{\boldsymbol{s}^{(1)}, \ldots, \boldsymbol{s}^{(t)}\}$ where $\boldsymbol{s}^{(i)} = \boldsymbol{r}^{(i)} + \boldsymbol{v}$. Then the set $W = \{(s_j^{(1)}, \ldots, s_j^{(t)}) \mid j = 1, \ldots, n\}$ contains vectors in $\mathbb{Z}_q^t$ where each vector in W contains at most t/q entries that are zero. For $\boldsymbol{s} \in W$ let $Z(\boldsymbol{s})$ be the set of indices of the entries that are zero in $\boldsymbol{s}$. There are no $\boldsymbol{s}^{(1)}, \boldsymbol{s}^{(2)} \in S$ that satisfy $Z(\boldsymbol{s}^{(1)}) \subset Z(\boldsymbol{s}^{(2)})$ because otherwise the set S is not $(n, 2)$-universal set. Therefore,

$$n = |W| \leq \binom{t}{t/q} \leq (eq)^{t/q}.$$

Therefore, $U(n, 2, q) = t = \Omega(q \log n / \log q)$.

Now let S be an (n, d)-universal set over $\mathbb{Z}_q$. Consider

$$S_{c_1, c_2, \ldots, c_{d-2}} = \{(v_{d-1}, v_d, \ldots, v_n) \mid \boldsymbol{v} \in S, v_i = c_i \text{ for all } i = 1, 2, \ldots, d-2\}.$$

Obviously, each $S_{\boldsymbol{c}}$, $\boldsymbol{c} \in \mathbb{Z}_q^{d-2}$ is $(n - d + 2, 2)$-universal set over $\mathbb{Z}_q$. Therefore

$$U(n, d, q) \geq q^{d-2} U(n - d + 2, 2, q) = \Omega\left(\frac{q^{d-1}}{\log q} \log n\right).$$

□

The best known polynomial time (i.e., $poly(q^d, n)$) construction for this problem gives a universal set of size $d^{O(\log d / \log q)} q^d \log n$ for $q < d$ and $O(d^5 (\log d)^2 q^d \log n)$, for $q > d$ [53, 2]. Here we will show the following

Lemma 51. *For any q, d and an integer r such that $q^r \geq n > q^{r-1}$ there is an explicit (n, d)-universal set over $\Sigma = \mathbb{F}_q$ of size*

$$q^d \cdot \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^r}).$$

In particular, for any constant $c > 1$, the following (n, d)-universal sets over $\Sigma = \mathbb{F}_q$ can be constructed in polynomial time (third column) and explicitly (forth column).

n	q	*Poly Time* *Size=*	*Explicite* *Size=*
I.S.	$q \geq c(d+1)^2$, *q P.S.*	$d^2 q^d \frac{\log n}{\log q}$	$dq^d \frac{\log n}{\log q}$
all	$q \geq c(d+1)^2$, *q P.S.*	$d^2 q^d \frac{\log n}{\log q}$	$d^2 q^d \frac{\log n}{\log q}$
I.S.	$q \geq c(d+1)$	$d^3 q^d \frac{\log n}{\log d}$	$d^2 q^d \frac{\log n}{\log d}$
all	$q \geq c(d+1)$	$d^3 q^d \frac{\log n}{\log d}$	$d^3 q^d \frac{\log n}{\log d}$
I.S.	$q \geq d+1$	$d^4 q^d \frac{\log n}{\log d}$	$d^3 q^d \frac{\log n}{\log d}$
all	$q \geq d+1$	$d^4 q^d \frac{\log n}{\log d}$	$d^4 q^d \frac{\log n}{\log d}$
all	$q \leq d+1$	$d^7 q^{(1+c_q/\log q)d} \log n$	$d^5 q^{(1+c_q/\log q)d} \log n$
all	$q = 2$	$d^7 2^{2.66d} \log n$	$d^5 2^{2.66d} \log n$

Proof. Consider the Reed Solomon $[q^r, q^r - d, d+1]$ code C over $\mathbb{F}_{q^r}$ where $q^r \geq n > q^{r-1}$. Consider the $d \times n$ matrix H_C that is the first n columns of the parity check matrix of C. Consider the set of (n,d)-multilinear polynomials $D_I(\boldsymbol{y}) = \det([y_{j,i_k}]_{j,k})$ where $j,k = 1,\ldots,d$, $I = \{i_1,\ldots,i_d\}$, $1 \leq i_1 < i_2 < \cdots < i_d \leq n$, $\boldsymbol{y} = (\boldsymbol{y}_1,\ldots,\boldsymbol{y}_d)^T$ and $\boldsymbol{y}_i = (y_{i,1},\ldots,y_{i,n})$. Here for notational convenience we regard $\boldsymbol{y}$ as a $d \times n$ matrix of indeterminates. Since every d columns of H_C are linearly independent we have $D_I(H_C) \neq 0$ for all I.

Let $L = \{\boldsymbol{\ell}_1,\ldots,\boldsymbol{\ell}_\nu\}$, $\boldsymbol{\ell}_i : \mathbb{F}_{q^r}^{d\times n} \to \mathbb{F}_q^{d\times n}$, be a $(\mathcal{DML}(\mathbb{F},n,d),\mathbb{F}_{q^r},\mathbb{F}_q)$-tester where $\nu = \nu_{\mathbb{F}_q}(d,\mathbb{F}_{q^r})$. Consider $G_i = \boldsymbol{\ell}_i(H_C) \in \mathbb{F}_q^{d\times n}$, $i = 1,2,\ldots,\nu$. By the definition of tester, since for all I, $D_I(H_C) \neq 0$ we have: for every I there is j such that $D_I(G_j) \neq 0$. That is, for every d distinct indices $I = \{i_1,\ldots,i_d\}$ there is a matrix G_j such that the columns $i_1,\ldots,i_d$ in G_j are linearly independent.

Now consider the set $\mathcal{F} = \cup_i \mathrm{Span}_{\mathbb{F}_q} G_i$ where $\mathrm{Span}_{\mathbb{F}_q} G_i$ is the linear space spanned by the rows of G_i. It is obvious that $\mathcal{F}$ is (n,d)-universal set over $\mathbb{F}_q$ of size $q^d \cdot \nu_{\mathbb{F}_q}(d,\mathbb{F}_{q^r})$.

The results in the first 6 rows of the table follow from the fact that $\nu_{\mathbb{F}_q}(d,\mathbb{F}_{q^r}) \leq \nu^{\mathcal{P}}_{\mathbb{F}_q}(d,\mathbb{F}_{q^r})$, Corollary 17 and Corollary 41. The results in the last two rows of the table follows from Theorem 21 and Theorem 44. □

For $q \geq c(d+1)^2$, perfect square q and infinite sequence of integers n we get an explicit construction of size $O(dq^d \log n/\log q)$, and for all n, a polynomial time construction of size $O(d^2 q^d \log n/\log q)$. Notice that the bound in the explicit construction exceeds the union bound and meets the upper bound in Lemma 49. For constant d all the above bounds meet the bounds in Lemma 49. For $q = 2$ (and small q) the bound $2^{d+O(\log^2 d)} \log n$ in [53] exceeds our bound $d^5 2^{2.66d} \log n$. The advantage of our construction is in that each bit in the construction can be constructed in $poly(d, \log n)$ time where the construction in [53] needs $\omega(poly(d)) \cdot poly(\log n)$ time. This is studied in more details in [12].

4.3 Cover-Free Families

Let X be a set with N elements and let $\mathcal{B}$ be a set of subsets (blocks) of X. We say that $(X, \mathcal{B})$ is (w, r)-*cover-free family* ((w, r)-CFF), [43], if for any w blocks $B_1, \ldots, B_w \in \mathcal{B}$ and any other r blocks $A_1, \ldots, A_r \in \mathcal{B}$, we have

$$\bigcap_{i=1}^{w} B_i \not\subseteq \bigcup_{j=1}^{r} A_j.$$

Let $N((w, r), n)$ denotes the minimum number of points in any (w, r)-CFF having n blocks. When $w = 1$, the problem is called *group testing*. The problem of group testing which was first presented during World War II was presented as follows [26, 51]: Among n soldiers, at most r carry a fatal virus. We would like to blood test the soldiers to detect the infected ones. Testing each one separately will give n tests. To minimize the number of tests we can mix the blood of several soldiers and test the mixture. If the test comes negative then none of the tested soldiers are infected. If the test comes out positive, we know that at least one of them is infected. The problem is to come up with a small number of group test.

This problem is equivalent to $(1, r)$-CFF and is equivalent to finding a small set $\mathcal{F} \subseteq \{0, 1\}^n$ such that for every $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ and every $1 \leq j \leq d$ there is $\boldsymbol{a} \in \mathcal{F}$ such that $a_{i_k} = 0$ for all $k \neq j$ and $a_{i_j} = 1$.

Group testing has the following lower bound [27, 28, 31]

$$N((1, r), n) \geq \Omega\left(\frac{r^2}{\log r} \log n\right). \tag{31}$$

It is known that a group testing of size $O(r^2 \log n)$ can be constructed in polynomial time [26, 55, 40].

The problem (w, r)-cover-free family is equivalent to the following problem: An (w, r)-cover-free family is a set $\mathcal{F} \subseteq \{0, 1\}^n$ such that for every $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ where $d = w + r$ and every $J \subset [d]$ of size $|J| = w$ there is $\boldsymbol{a} \in \mathcal{F}$ such that $a_{i_k} = 0$ for all $k \notin J$ and $a_{i_j} = 1$ for all $j \in J$. Then $N((w, r), n)$ is the minimum size of such $\mathcal{F}$.

There are several lower bounds for $N((w, r), n)$. We give the one in [78]

$$N((w, r), n) \geq \Omega\left(\frac{d\binom{d}{w}}{\log \binom{d}{w}} \log n\right).$$

We first use union bound to show the following

Lemma 52. *For $d = w + r = o(n)$ we have*

$$N((w, r), n) \leq O\left(\sqrt{wrd} \cdot \binom{d}{w} \log n\right).$$

Proof. We choose a random vector $\boldsymbol{a} \in \{0,1\}^n$ where $\mathbf{Pr}[a_i = 1] = w/d$ for all $i = 1, \ldots, n$. For distinct $i_1, \ldots, i_d \in [n]$, let $A_{i_1,\ldots,i_d}$ be the event that $(a_{i_1}, a_{i_2}, \cdots, a_{i_d})$ is of weight w. Then, by Stirling's formula, we have

$$\mathbf{Pr}[A_{i_1,\ldots,i_d}] = \binom{d}{w} \left(\frac{w}{d}\right)^w \left(1 - \frac{w}{d}\right)^{d-w} = O\left(\sqrt{\frac{d}{wr}}\right).$$

Now using union bound the result follows. □

It follows from [79], that for infinite sequence of integers n, a (w,r)-cover free family of size

$$M = O\left((wr)^{\log^* n} \log n\right)$$

can be constructed in polynomial time. For constant d, the (n,d)-universal set over $\Sigma = \{0,1\}$ constructed in [52] of size $M = O(2^{3d} \log n)$ (and in [53] of size $M = 2^{d+O(\log^2 d)} \log n$) is (w,r)-cover free family for any w and r of size $O(\log n)$. See also [48].

We now prove

Lemma 53. *Let t be such that $q^t \geq n$ and $q \geq wr+1$. Let $S \subseteq \mathbb{F}_{q^T}$ be a sublinear space for some $T \geq t$ where $|S| = q^t$. Then*

$$N((w,r),n) \leq N((w,r),q) \cdot \nu_{\mathbb{F}_q}^{\mathcal{HP}}(wr, S).$$

In particular, there is an explicit (w,r)-CFF of size

$$\binom{q}{w} \cdot \nu_{\mathbb{F}_q}^{\mathcal{HP}}(wr, S).$$

In particular, for any constant $c > 1$, the following (w,r)-CFF can be constructed in polynomial time in their sizes

n	w	*Poly time Size=*	*Union Bound*	*Lower Bound*
I.S	$O(1)$	$\frac{r^{w+2}}{\log r} \log n$	$r^{w+1} \log n$	$\frac{r^{w+1}}{\log r} \log n$
all	$O(1)$	$\frac{r^{w+3}}{\log r} \log n$	$r^{w+1} \log n$	$\frac{r^{w+1}}{\log r} \log n$
I.S.	$o(r)$	$\frac{w^2 (ce)^w r^{w+2}}{\log r} \log n$	$\frac{r^{w+1}}{(w/e)^{w-1/2}} \log n$	$\frac{r^{w+1}}{(w/e)^{w+1} \log r} \log n$
all	$o(r)$	$\frac{w^3 (ce)^w r^{w+3}}{\log r} \log n$	$\frac{r^{w+1}}{(w/e)^{w-1/2}} \log n$	$\frac{r^{w+1}}{(w/e)^{w+1} \log r} \log n$

Proof. Consider the set of non-zero functions

$$\mathcal{M} = \{\Delta_{\boldsymbol{i}} \mid \boldsymbol{i} \in [n]^d,\ i_1, i_2, \ldots, i_d \text{ are distinct}\}$$

where

$$\Delta_{\boldsymbol{i}}(x_1,\ldots,x_n)=\prod_{1\le k\le w \text{ and } w<j\le d}(x_{i_k}-x_{i_j}).$$

Consider n distinct elements $\alpha_1,\alpha_2,\ldots,\alpha_n\in S$ and let $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_n)$. Since $g(\boldsymbol{\alpha})\neq 0$ for every $g\in\mathcal{M}$, and since $\mathcal{M}\subset\mathcal{HP}(\mathbb{F}_q,n,wr)$, using testers, there is a set $B\subseteq\mathbb{F}_q^n$ of size $\nu^{\mathcal{HP}}_{\mathbb{F}_q}(wr,S)$ such that for every $g\in\mathcal{M}$ there is $\boldsymbol{b}\in B$ where $g(\boldsymbol{b})\neq 0$.

Let $\mathcal{F}\subseteq\{0,1\}^q$ be a (w,r)-CFF with $|X|=q$ elements of size $N((w,r),q)$. Regard each $f\in\mathcal{F}$ as a function $f:\mathbb{F}_q\to\{0,1\}$. It is easy to see that

$$\{(f(b_1),f(b_2),\ldots,f(b_n))\mid \boldsymbol{b}\in B, f\in\mathcal{F}\}\subseteq\{0,1\}^n$$

is (w,r)-CFF of size $|\mathcal{F}|\cdot|B|=N((w,r),q)\cdot\nu^{\mathcal{HP}}_{\mathbb{F}_q}(wr,S)$.

Now for every subset $R\subseteq\mathbb{F}_q$ define the function $\chi_R:\mathbb{F}_q\to\{0,1\}$ where for $\beta\in\mathbb{F}_q$ we have $\chi_R(\beta)=1$ if $\beta\in R$ and $\chi_R(\beta)=0$ otherwise. Then $\{\chi_R\mid R\subseteq\mathbb{F}_q,|R|=w\}\subseteq\{0,1\}^{\mathbb{F}_q}$ is a (w,r)-CFF with $|\mathbb{F}_q|=q$ elements of size $\binom{q}{w}$. Therefore

$$C=\{(\chi_R(b_1),\chi_R(b_2),\ldots,\chi_R(b_n))\mid \boldsymbol{b}\in B, R\subseteq\mathbb{F}_q,|R|=w\}$$

is (w,r)-CFF of size

$$|C|\le\binom{q}{w}\nu^{\mathcal{HP}}_{\mathbb{F}_q}(wr,S).$$

Now for the results in the table consider a constant $c>1$ and let q be a power of prime such that $q=cwr+o(wr)$. This is possible by [19]. Let $t=\lceil\log n/\log q\rceil$ and let $S\subseteq\mathbb{F}_{q^T}$ where $T=O(t)$ be the linear space defined in Corollary 20. By Corollaries 20, 43 and the above result, for infinite sequence of integers n, there is a (w,r)-CFF of size

$$\binom{q}{w}\cdot\nu^{\mathcal{HP}}_{\mathbb{F}_q}(wr,S)\le\left(\frac{qe}{w}\right)^w(wr)^2t=O\left((cer)^w(wr)^2\frac{\log n}{\log r}\right)=O\left(\frac{w^2(ce)^w r^{w+2}}{\log r}\log n\right)$$

that can be constructed in polynomial time. By Corollary 41 there is a (w,r)-CFF of size

$$\binom{q}{w}\cdot\nu^{\mathcal{HP}}_{\mathbb{F}_q}(wr,\mathbb{F}_{q^t})\le\left(\frac{qe}{w}\right)^w(wr)^3t=O\left(\frac{w^3(ce)^w r^{w+3}}{\log r}\log n\right)$$

that can be constructed in polynomial time. □

In Corollary 17 we have showed that for any constant $c>1$ and $q\ge c(d+1)$, for infinite sequence of integers t we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d,\mathbb{F}_{q^t})=O(d^2t)$. One of the open problems in this paper is whether this bound

can be improved to $O(dt)$. By the proof of Lemma 53, if for some constant $c > 1$ and any $q = c(d+1)$ we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(dt)$ then for $w = O(1)$

$$N((w,r),n) = O\left(\frac{r^{w+1}}{\log r}\log n\right).$$

This bound matches the lower bound and therefore closes the gap between the union bound and the lower bound.

In particular, we will have

$$N((1,r),n) = \Theta\left(\frac{r^2}{\log r}\log n\right)$$

for the group testing problem. We state this in the following

Lemma 54. *If for some constant $c > 1$ and $q = c(d+1)$ we have $\nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O(dt)$ then*

$$N((1,r),n) = \Theta\left(\frac{r^2}{\log r}\log n\right).$$

4.4 Separating Hash Family

Let X and Σ be sets of cardinalities n and q, respectively. We call a set $\mathcal{F}$ of functions $f : X \to \Sigma$ an $(M; n, q, \{d_1, d_2, \ldots, d_r\})$ *separating hash family* (SHF), [75, 76], if $|\mathcal{F}| = M$ and for all pairwise disjoint subsets $C_1, C_2, \ldots, C_r \subseteq X$ with $|C_i| = d_i$ for $i = 1, 2, \ldots, r$, there is at least one function $f \in \mathcal{F}$ such that $f(C_1), f(C_2), \ldots, f(C_r)$ are pairwise disjoint subsets. The goal is to find $(M; n, q, \{d_1, d_2, \ldots, d_r\})$ SHF with small M. The minimal M is denoted by $M(n, q, \{d_1, d_2, \ldots, d_r\})$.

In [22], Bazrafshan and Trund proved that for

$$D_1 = \sum_{i=1}^{r} d_i,$$

$$M(n, q, \{d_1, d_2, \ldots, d_r\}) > (D_1 - 1)\,\frac{\log n - \log(D_1 - 1) - \log q}{\log q} = \Omega\left(D_1 \frac{\log n}{\log q}\right). \tag{32}$$

In [79], Stinson et. al. proved that an $(M; n, q, \{d_1, d_2\})$ separating hash families of size

$$M = O((d_1 d_2)^{\log^* n} \log n)$$

can be constructed in polynomial time for infinite sequence of integers n and $q > d_1 d_2$. The same proof gives a polynomial time construction for any separating hash family of size

$$M = O(D_2^{\log^* n} \log n)$$

where

$$D_2 = \sum_{1 \le i_1 < i_2 \le r} d_{i_1} d_{i_2}$$

when $q > D_2$.

In [48], Liu and Shen provide an explicit constructions of $(M; n, q, \{d_1, d_2\})$ separating hash families using algebraic curves over finite fields. They show that for infinite sequence of integers n there is an explicit $(M; n, q, \{d_1, d_2\})$ separating hash families of size $O(\log n)$ for fixed d_1 and d_2. This also follows from [52], an $(n, d_1 + d_2)$-universal set over two symbols alphabet is a separating hash families of size $O(\log n)$ for fixed d_1 and d_2. Their construction is similar to the construction of the tester defined in Lemma 12. The following lemma gives a polynomial time construction of an $(M; n, q, \{d_1, d_2\})$ separating hash families of size $M = ((d_1 d_2)^4 \log n / \log q)$ for any $q \ge d_1 d_2(1 + o(1))$ and any n.

Lemma 55. *Let $q > D_2$. Let t be an integer such that $q^t \ge n$. Let $S \subseteq \mathbb{F}_{q^T}$ be a sublinear space for some $T \ge t$ where $|S| = q^t$. There is an explicit $(M; n, q, \{d_1, d_2, \ldots, d_r\})$ separating hash family of size*

$$M = \nu_{\mathbb{F}_q}^{\mathcal{HP}}(D_2, S).$$

For $q' \le D_2$ we have

$$M(n, q', \{d_1, d_2, \ldots, d_r\}) \le M(q, q', \{d_1, d_2, \ldots, d_r\}) \cdot \nu_{\mathbb{F}_q}^{\mathcal{HP}}(D_2, S).$$

In particular, for any constant $c > 1$ and $q > D_2$, the following $(M; n, q, \{d_1, d_2, \ldots, d_r\})$ separating hash family can be constructed in polynomial time

n	q	*poly time* $Size =$	*Union Bound*	*Lower Bound [17]*
I.S.	$q \ge c(D_2+1)^2$, *q P.S.*	$D_2 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$
all	$q \ge c(D_2+1)^2$, *q P.S.*	$D_2^2 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$	$D_1 \frac{\log n}{\log q}$
I.S.	$q \ge c(D_2+1)$	$D_2^2 \frac{\log n}{\log d}$	$D_1 \frac{\log n}{\log(q/D_2))}$	$D_1 \frac{\log n}{\log q}$
all	$q \ge c(D_2+1)$	$D_2^3 \frac{\log n}{\log d}$	$D_1 \frac{\log n}{\log(q/D_2)}$	$D_1 \frac{\log n}{\log q}$
I.S.	$q \ge D_2 + 1$	$D_2^3 \frac{\log n}{\log q}$	$D_1 \log n$	$D_1 \frac{\log n}{\log q}$
all	$q \ge D_2 + 1$	$D_2^4 \frac{\log n}{\log q}$	$D_1 \log n$	$D_1 \frac{\log n}{\log q}$

and an $(M; n, r, \{d_1, d_2, \ldots, d_r\})$ separating hash family of size

$$\frac{\binom{cD_2}{d_1\ d_2\ \cdots\ d_r} D_2^3}{\log D_2} \log n,$$

can be constructed in polynomial time.

Proof. Consider the set of functions

$$\mathcal{F} = \{\Delta_{(C_1,\ldots,C_r)}(x_1,\ldots,x_n) \mid C_1,\ldots,C_d \text{ are pairwise disjoint}, |C_i| = d_i\}$$

in $\mathbb{F}_q[x_1, x_2, \ldots, x_n]$ where

$$\Delta_{(C_1,\ldots,C_r)} = \prod_{1\leq k<j\leq r} \prod_{i_1\in C_k, i_2\in C_j} (x_{i_1} - x_{i_2}).$$

The proof then proceeds as the proof of Lemma 48. □

Open Problems 6.

1. *Can one somehow combines the constructions we have here with the techniques used in [53] to get polynomial time constructions for small size alphabet? If we allow quasi-polynomial time for the constructions then we can use Theorem 1 in [53] to get almost optimal size constructions.*
2. *Find polynomial time constructions of almost optimal size for (w, r)-CFF with alphabet of size less than wr and for $(M; n, q, \{d_1, d_2, \ldots, d_r\})$ SHF with $q < D_2$.*

5 Application of Tester for Black Box PIT Sets over Small Field

In this section we show how to reduce a black box PIT set over large field to a black box PIT set over small field. We then apply this to different subsets of multivariate polynomials.

The black box Polynomial Identity Testing (PIT) problem is the following: Given an arithmetic circuit C that either identical to the zero function or from a class of circuits $\mathcal{C}$ over a field $\mathbb{F}$, with input variables $x_1, x_2, \ldots, x_n$ and given a *substitution oracle* that for an input $\boldsymbol{a} \in \mathbb{F}^n$ returns $f(\boldsymbol{a})$. Determine whether C computes the identically zero polynomial. We say that $S \subset \mathbb{F}^n$ is a *black box PIT set* or a *hitting set* for $\mathcal{C}$ if for every $f \in \mathcal{C}$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$.

When the field is finite $\mathbb{F}_q$ many simple classes of circuits, such as circuits that computes monomials, requires black box PIT sets of exponential size. Therefore, many papers in the literature allow the substitution oracle to receive assignment $\boldsymbol{a}$ from some extension field $\mathbb{F}_{q^t}$ of $\mathbb{F}_q$. One problem with that approach is that some functions, such as $x_1^q - x_1$, are identically zero over $\mathbb{F}_q$ but not over any extension field $\mathbb{F}_{q^t}$ of $\mathbb{F}_q$. Therefore, when using an extension field, they assume that each output node in the circuit computes a function of variable degree less than q. That is, the degree of each variable in the function is less than q. In that case, C is identically zero over $\mathbb{F}_q$ if and only if it is identically zero over $\mathbb{F}_{q^t}$ for any t. We say that $S \subset \mathbb{F}_{q^t}^n$ is a *black box PIT set over* $\mathbb{F}_{q^t}$ or a *hitting set over* $\mathbb{F}_{q^t}$ for $\mathcal{C}$ if for every $f \in \mathcal{C}$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$. Our goal will be to minimize the size of the black box PIT set and the dimension of the extension field.

We first start with some definitions

5.1 Sets of Multivariate Polynomials

For a multivariate polynomial f the *total degree* (or just *degree*) of f is the maximum over the sums of the exponents of each multivariate monomial in f and the *variable degree* of f is the maximum over the degree of each variable in f.

In this section we will study the following classes of multivariate polynomials

1. $\mathcal{P}(\mathbb{F}_q, n)$ is the class of all multivariate polynomials in $\mathbb{F}_q[x_1, \ldots, x_n]$ of variable degree at most $q-1$.
2. $\mathcal{P}(\mathbb{F}_q, n, (d, r))$ is the class of all multivariate polynomials in $\mathcal{P}(\mathbb{F}_q, n)$ of degree at most d and variable degree at most r.
3. $\mathcal{P}(\mathbb{F}_q, n, d) = \mathcal{P}(\mathbb{F}_q, n, (d, q-1))$ is the class of all multivariate polynomials in $\mathcal{P}(\mathbb{F}_q, n)$ of degree at most d.
4. $\mathcal{P}(\mathbb{F}_q, n, s)$ is the class of all multivariate polynomials in $\mathcal{P}(\mathbb{F}_q, n)$ with at most s monomials. This class is called in the literature "sparse multivariate polynomials".
5. $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ is the class of all multivariate polynomials in $\mathcal{P}(\mathbb{F}_q, n)$ of degree at most d and variable degree at most r with at most s monomials.

In the same way we define the classes $\mathcal{P}(\mathbb{F}_q, n, r)$, $\mathcal{P}(\mathbb{F}_q, n, d, s)$ and $\mathcal{P}(\mathbb{F}_q, n, r, s)$. Notice that when we write $\mathcal{P}(\mathbb{F}_q, n, 1)$ it is not clear whether $r = 1$, $d = 1$ or $s = 1$. To avoid confusion we will use "=" in the definition of the class to indicate which parameter is meant. For example, $\mathcal{P}(\mathbb{F}_q, n, r = 1)$ is the class of multilinear polynomials where $\mathcal{P}(\mathbb{F}_q, n, s = 1)$ is the class of monomials and $\mathcal{P}(\mathbb{F}_q, n, d = 1)$ is the class of linear functions.

5.2 Main Results

We will study black box PIT sets for the above classes over $\mathbb{F}_{q^t}$. To the best of our knowledge all the algorithms in the literature that construct black box PIT sets for the above classes are either randomized, deterministic for some fixed extension field $\mathbb{F}_{q^t}$ or obtains non-optimal results in both the extension field dimension and the size of the black box PIT set [34, 24, 80, 39, 45, 9]. See more details in the next subsections. In this paper all the results we obtain are within $poly(n)$ of the optimal black box PIT set size and the dimension of the extension field is optimal.

The following table summarizes some of our main results.

Class	Extension Field $\mathbb{F}_{q^t}$	Lower Bound	Upper Bound	Explicit Construction	Poly Time Construction
$\mathcal{P}(\mathbb{F}_q, n)$	$t \geq \log_q n + 2$	$\frac{q^n}{t}$	$\frac{(\log(qn))\cdot q^n}{t}$	$\frac{n\cdot q^n}{t}$	$\frac{n\cdot q^n}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r)$	$t \geq \log_q n + 2$	$\frac{(r+1)^n}{t}$	$\frac{(\log(rn))\cdot(r+1)^n}{t}$	$\frac{n\cdot(r+1)^n}{t}$	$\frac{n\cdot(r+1)^n}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r))$	$t \geq \log_q(d+1)$	$\frac{R(n,d,r)}{t}$	$\frac{(\log d)\cdot R(n,d,r)}{t}$	$\frac{dn^{d+1}}{t}$	$\frac{dn^{d+1}}{t}$
$\mathcal{P}(\mathbb{F}_q, n, s)$	$t \geq \log_q n + 2$	$\frac{n\cdot s}{t}$	$\frac{(\log n)n\cdot s}{t}$	$\frac{q^5 n^6\cdot s}{t}$	$\frac{n^7 q^{29}(\log^2 q)\cdot s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r, s)$	$t \geq \log_q n + 2$	$\frac{n\log(r+1)}{\log q}\cdot\frac{s}{t}$	$\frac{n(\log n)\log(r+1)}{\log q}\cdot\frac{s}{t}$	$\frac{n^6 r^5\log(r+1)}{\log q}\cdot\frac{s}{t}$	$\frac{n^7 r^5 q^{24}(\log^2 r)\cdot s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$	$t \geq \log_q(d+1)$	$\frac{d\log n}{\log q}\cdot\frac{s}{t}$	$\frac{d(\log d)\log n}{\log q}\cdot\frac{s}{t}$	$\frac{d^6\log n}{\log q}\cdot\frac{s}{t}$	$\frac{q^{24}d^7(\log^2 n)\cdot s}{t}$

Notice that our polynomial time constructions (column 6 in the above table) are optimal in the largest parameters (q^n, $(r+1)^n$, n^d and s in the last three rows of the table). For sparse polynomials $\mathcal{P}(\mathbb{F}_q, n, s)$, $\mathcal{P}(\mathbb{F}_q, n, r, s)$ and $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$, all the results in the literature give black box PIT sets of size that are at least quadratic in the size s. Also, the tight tradeoff with the field dimension $(\cdot/t)$ in each row of the table was not known before. The bound on the dimension of the field extension in the second column is $\log_q(degree) + 1$ and is known to be the best possible dimension (even for randomized algorithms) if one uses Schwartz-Zippel Lemma. Therefore, our constructions are tight in the dimension of the field extension. For the class $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ (the last row of the table), the best known result in the literature [45, 9] used field extension of dimension that depends on the number of the variables n. Our result uses field extension of dimension $\log_q(d+1)$ that is independent of the number of variables. Also notice, that when $q \geq d+1$ no extension field is needed. This will be further studied in [12] to give new Pseudorandom generators over small fields.

See other results in the following subsections.

5.3 Preliminary Results

Before we move to the constructions we give some preliminary results that will be needed in this section

5.3.1 A Primitive Root in the Field

In this subsection we give some results from the literature about deterministic algorithms for finding a primitive root (called also primitive element) and an element of large multiplicative order in the field. See more results in [64].

In [62] it is shown that

Lemma 56. *In any field $\mathbb{F}_q$ a primitive root can be found in deterministic time $O(q^{1/4+\epsilon})$ for any constant $\epsilon > 0$.*

In some applications one can also use an element of large order in finite field instead of a primitive root. Von zur Gathen and Shparlinski in [35] showed

Lemma 57. *There is a deterministic algorithm that for every integer n, gives $N = n + o(n)$ and finds an element in $\mathbb{F}_{q^N}$ of multiplicative order*

$$2^{(10n^{1/2}/q^{12})-25}$$

in time poly(n).

In [68], Shoup proved (see also [64], Theorem 2.6)

Lemma 58. *There is a deterministic algorithm that for every integer n, gives $N = n + o(n)$ and finds a primitive root in $\mathbb{F}_{p^N}$ in time*

$$p^{O(n/\log\log n)}.$$

5.3.2 Sidon Sequences

In this subsection we define the Sidon Sequence and prove some results that will be used in the sequel

A sequence of non-negative integers $a_1, a_2, \cdots, a_n$ is called *Sidon B_d sequence* if the sums $a_{i_1} + \cdots + a_{i_d}$ for all $\boldsymbol{i} \in [n]^d$ are distinct up to rearrangements of the summands. The sequence is called *Sidon $B_{\leq d}$ sequence* if the sums $a_{i_1} + \cdots + a_{i_r}$ for all $\boldsymbol{i} \in [n]^d$ and $r \leq d$ are distinct up to rearrangements of the summands. We will study Sidom $B_{\leq d}$ sequences when $d = o(n)$. It is easy to see that $\max_j a_j \geq n^{d-o(d)}$. See more results in [11] and references within.

Several explicit constructions of Sidon B_d sequences were given in the literature [11]. Bose [10] constructed B_2 Sidon sequences by using finite affine geometry. His construction was extended in [15] to the following Sidon B_d sequences. Let q be a prime power and θ a primitive root of $\mathbb{F}_{q^d}$. Define

$$B_d(q, \theta) = \{a \in [q^d - 1] \mid \theta^a - \theta \in \mathbb{F}_q\}.$$

Then $B_d(q, \theta)$ is a Sidon B_d sequence. The construction of this sequence requires discrete log and therefore it is not clear whether it can be done in polynomial time.

We prove

Lemma 59. *A Sidon $B_{\leq d}$ sequence $a_1, a_2, \cdots, a_n$ with*

$$\max_j a_j \leq (n + O(n^{0.525}))^{d+1}$$

can be constructed in deterministic $O(n^{d/2+1.5})$ time.

Proof. Consider the smallest prime number p that is greater than n. By [19], $p \leq n+O(n^{0.525})$. Consider $B' = B_{d+1}(p,\theta)$ for some primitive root $\theta \in \mathbb{F}_{p^{d+1}}$. By Lemma 56, a primitive root for $\mathbb{F}_{p^{d+1}}$ can be found in deterministic $O(p^{(d+1)/4+\epsilon}) = O(n^{(d+1)/4+\epsilon})$ time for any ϵ. To find all $a \in [p^{d+1}-1]$ such that $\theta^a - \theta \in \mathbb{F}_p$ we can use Daniel Shanks baby-step giant-step algorithm for discrete log that runs in time $p^{(d+1)/2+1} = O(n^{d/2+1.5})$. Therefore, the complexity of constructing the Sidon B_{d+1} sequence is $O(n^{d/2+1.5})$.

Since $|B'| = p > n$ we can choose $B = \{a_1, a_2, \ldots, a_n\} \subset B'$ that contains n elements. Since $a_j \in [p^{d+1}-1]$ we have $a_j \leq p^{d+1} \leq (n+O(n^{0.525}))^{d+1}$.

We now show that this sequence is Sidon $B_{\leq d}$ sequence. One way to show this is to show that given a where $a = a_{i_1} + \cdots + a_{i_{d'}}$ for some $d' \leq d$, one can uniquely determines d' and $a_{i_1}, \cdots, a_{i_{d'}}$.

Given $a = a_{i_1} + \cdots + a_{i_{d'}}$. Let $\alpha_i = \theta^{a_i} - \theta \in \mathbb{F}_p$ for $i = 1, 2, \ldots, d' \leq d$. Consider

$$\theta^a = \theta^{a_{i_1}+\cdots+a_{i_{d'}}} = (\theta+\alpha_{i_1})\cdots(\theta+\alpha_{i_{d'}}) = \theta^{d'} + A_{d'-1}\theta^{d'-1} + \cdots + A_0$$

for $A_i \in \mathbb{F}_p$, $i = 0, 1, \ldots, d'-1$. This uniquely determines d' and A_i for $i = 0, 1, \ldots, d'-1$. Since

$$x^{d'} + A_{d'-1}x^{d'-1} + \cdots + A_0 = (x+\alpha_{i_1})\cdots(x+\alpha_{i_{d'}}),$$

by factoring $x^{d'} + A_{d'-1}x^{d'-1} + \cdots + A_0$ over $\mathbb{F}_q$ we get α_{i_j}, $j = 1, \ldots, d'$ which uniquely determine a_{i_j}, $j = 1, \ldots, d'$. □

For $B_{\leq d}$ Sidon sequence that can be constructed in polynomial time we prove

Lemma 60. *A Sidon $B_{\leq d}$ sequence $a_1, a_2, \cdots, a_n$ with*

$$\max_j a_j < (2dn)^{2d}$$

can be constructed in time poly(n).

Proof. Consider the smallest prime number p that is greater than n. By [19], $p \leq n + O(n^{0.525}) \leq 2n$. Consider the Reed-Solomon code $[n, n-2d, 2d+1]$ over $\mathbb{K} = \mathbb{F}_p$ with the code locators $\alpha^0, \alpha^1, \ldots, \alpha^{n-1}$, for some primitive root $\alpha \in \mathbb{F}_p$. Let

$$H = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha^0 & \alpha^1 & \cdots & \alpha^{n-1} \\ (\alpha^0)^2 & (\alpha^1)^2 & \cdots & (\alpha^{n-1})^2 \\ \vdots & \vdots & \vdots & \vdots \\ (\alpha^0)^{2d-1} & (\alpha^1)^{2d-1} & \cdots & (\alpha^{n-1})^{2d-1} \end{bmatrix}.$$

Define a function $\rho : \mathbb{F}_p \to \mathbb{Z}$ as follows. For $\bar{\beta} \in \mathbb{F}_p$ we have $\rho(\bar{\beta}) = \beta \in \mathbb{Z}$ where $\bar{\beta} \equiv \beta \mod p$ and $0 \leq \beta \leq p-1$. Consider the $(2d) \times n$ matrix $\rho(H) = [\rho(H_{i,j})]_{i,j}$. Since every $2d$ columns in H are linearly independent over $\mathbb{F}_p$, every $2d$ columns in $\rho(H)$ are linearly independent over $\mathbb{R}$.

Let

$$(a_1, \cdots, a_n) = (1, (p-1)d+1, ((p-1)d+1)^2, \cdots, ((p-1)d+1)^{2d-1}) \cdot \rho(H).$$

Now, our claim is that the sequence $a_1, a_2, \cdots, a_n$ is Sidon $B_{\leq d}$ sequence with $\max_j a_j \leq (2dn)^{2d}$.

First, since the entries of $\rho(H)$ are in $\{0, 1, \ldots, p-1\} \subset \mathbb{Z}$, we have

$$a_j \leq (p-1) + ((p-1)d+1)(p-1) + \cdots + ((p-1)d+1)^{2d-1}(p-1) \leq (2dn)^{2d}.$$

Now given an integer m that is equal to a sum of $d' \leq d$ elements $a_{i_1} + \cdots + a_{i_{d'}}$. Let $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\} \subset \mathbb{Z}^n$ be the standard basis and define $\boldsymbol{b} = \boldsymbol{e}_{i_1} + \cdots + \boldsymbol{e}_{i_{d'}}$. Then

$$m = (a_1, \ldots, a_n)\boldsymbol{b}^T = (1, (p-1)d+1, ((p-1)d+1)^2, \cdots, ((p-1)d+1)^{2d-1}) \cdot (\rho(H)\boldsymbol{b}^T).$$

Since the entries of $\rho(H)\boldsymbol{b}^T$ are non-negative and less than or equal to $(p-1)d$ writing

$$m = m_0 + m_1((p-1)d+1) + \cdots + m_{2d-1}((p-1)d+1)^{2d-1}$$

in base $(p-1)d+1$ gives

$$(m_0, \ldots, m_{2d-1}) = \rho(H)\boldsymbol{b}^T.$$

Since every $2d$ columns in $\rho(H)$ are linearly independent over $\mathbb{R}$ and $\boldsymbol{b}$ contains at most d nonzero entries $\rho(H)\boldsymbol{b}^T$ uniquely determines $\boldsymbol{b}$. This uniquely determines d' and $a_{i_1}, \ldots, a_{i_{d'}}$. □

5.3.3 The Operator ϕ_d

In this subsection we introduce a new notion that will be used in the sequel.

Let $\mathbb{F}$ be any field. Consider a multivariate polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ of degree d. Let $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)$ where $\boldsymbol{y}_i = (y_{i,1}, \ldots, y_{i,n})$ are new indeterminates for $i = 1, \ldots, d$. Define the operator $\phi_d : \mathbb{F}[\boldsymbol{x}] \to \mathbb{F}[\boldsymbol{y}]$

$$\phi_d f = \sum_{J \subseteq [d]} (-1)^{d-|J|} f\left(\sum_{i \in J} \boldsymbol{y}_i\right), \tag{33}$$

where $\sum_{i \in \emptyset} \boldsymbol{y}_i = 0$. We now show

Lemma 61. *We have*

1. *For a monomial M of degree less than d we have $\phi_d M = 0$*

2. *For a monomial $M_{\boldsymbol{i}} = x_{i_1} \cdots x_{i_d}$ of degree d we have $\phi_d M_{\boldsymbol{i}} = \mathrm{Perm} Y_{M_{\boldsymbol{i}}}$ where* Perm *is the permanent of matrices and*

$$Y_{M_{\boldsymbol{i}}}(\boldsymbol{y}) = \begin{pmatrix} y_{1,i_1} & y_{1,i_2} & \cdots & y_{1,i_d} \\ y_{2,i_1} & y_{2,i_2} & \cdots & y_{2,i_d} \\ \vdots & \vdots & \vdots & \vdots \\ y_{d,i_1} & y_{d,i_2} & \cdots & y_{d,i_d} \end{pmatrix}.$$

3. *For any monomial M of degree at most d we have $\phi_d M = 0$ if and only if M is of degree less than d or one of the variables in M has degree that is greater than or equal to the characteristic of the field p.*

Proof. Let $M = x_{i_1} x_{i_2} \cdots x_{i_r}$. If $r < d$ then

$$\phi_d M = \sum_{J \subseteq [d]} (-1)^{d-|J|} \prod_{k=1}^{r} \sum_{j \in J} y_{j,i_k} = \sum_{\boldsymbol{j} \in [d]^r} \lambda_{\boldsymbol{j}} y_{j_1,i_1} \cdots y_{j_r,i_r}$$

for some $\lambda_{\boldsymbol{j}} \in \mathbb{F}$. For every $\boldsymbol{j} \in [d]^r$ the coefficient $\lambda_{\boldsymbol{j}}$ of $y_{j_1,i_1} \cdots y_{j_r,i_r}$ in $\phi_d M$ is a multiple of (when $i_1, \ldots, i_k$ are distinct then it is exactly)

$$\sum_{\{j_1,j_2,\ldots,j_r\} \subset J \subseteq [d]} (-1)^{d-|J|} = \sum_{i=t}^{d} (-1)^{d-i} \binom{d-t}{i-t} = \sum_{i=0}^{d-t} (-1)^{d-t-i} \binom{d-t}{i} = (1-1)^{d-t} = 0,$$

where $t = |\{j_1, j_2, \ldots, j_r\}| \leq r$. This implies *1*.

For $r = d$, by Ryser's formula [57], result *2* follows.

We now prove *3*. If $M = x_1^{d_t} \cdots x_t^{d_t}$ where $d_1 + \cdots + d_t = d$ then, by *2*, each monomial in $\phi_d M$ appears (and therefore its coefficient is equal to) $d_1! d_2! \cdots d_t!$ times. Therefore $\phi_d M$ is zero if and only if there is i such that $d_i \geq p$. Combining this with *1* the result follows. □

Suppose

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} x_{i_1} \cdots x_{i_d} + g(\boldsymbol{x})$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)$, $g(\boldsymbol{x})$ is a multivariate polynomial of degree less than d, $\boldsymbol{i} = (i_1, i_2, \ldots, i_d)$, $1 \leq i_1 \leq i_2 \leq \cdots \leq i_d \leq n$, $\lambda_{\boldsymbol{i}} \neq 0$ and $I \subset [n]^d$. Since ϕ_d is linear we have

$$(\phi_d f)(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} \mathrm{Perm}\left(Y_{M_{\boldsymbol{i}}}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)\right). \tag{34}$$

5.4 The Reduction from Large Field to Small Field

In this subsection we show how to reduce a black box PIT set over an extension field $\mathbb{F}_{q^t}$ to a black box PIT set over $\mathbb{F}_{q^{t'}}$ where $t' < t$.

We prove

Lemma 62. *If $\mathcal{M} \subseteq \mathbb{F}[x_1, \ldots, x_n]$ and there is a black box PIT set $T \subset S^n$ for $\mathcal{M}$ over a subspace $S \subseteq \mathcal{A}$ for an $\mathbb{F}$-algebra $\mathcal{A}$ then there is a black box PIT set $R \subset \mathbb{F}^n$ for $\mathcal{M}$ of size $|R| = |T| \cdot \nu^{\circ}_{\mathbb{F}}(\mathcal{M}, S)$.*

In particular,

1. *If $\mathcal{M} \subseteq \mathcal{DML}(\mathbb{F}_q, n, d)$ and there is a black box PIT set over an extension field $\mathbb{F}_{q^t}$ for $\mathcal{M}$ of size w then there is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_q$ of size $w \cdot \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$.*
2. *If $\mathcal{M} \subseteq \mathcal{P}(\mathbb{F}_q, n, d)$, $q \geq d+1$ and there is a black box PIT set for $\mathcal{M}$ over an extended field $\mathbb{F}_{q^t}$ of size w then there is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_q$ of size $w \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$.*
3. *If $\mathcal{M} \subseteq \mathcal{P}(\mathbb{F}_q, n, d)$, $q^{t'} \geq d+1$ and there is a black box PIT set for $\mathcal{M}$ over an extended field $\mathbb{F}_{q^t}$ of size w then there is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_{q^{t'}}$ of size $w \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^{t'}}}((d, \mathbb{F}_q), \mathbb{F}_{q^t})$.*

Proof. The results follow immediately from the definition of testers. □

By Corollary 41, Theorem 44 and Lemma 45 we have

Lemma 63. *We have*

1. *Let $\mathcal{M} \subseteq \mathcal{DML}(\mathbb{F}_q, n, d)$. Let S be a black box PIT set for $\mathcal{M}$ over an extended field $\mathbb{F}_{q^t}$ of size w. There is an algorithm that runs in time $w \cdot 2^{c_q d} \cdot poly(t)$ and constructs a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_q$ of size $O(d^7 \cdot 2^{c_q d} \cdot tw)$.*
2. *Let $\mathcal{M} \subseteq \mathcal{P}(\mathbb{F}_q, n, d)$ and $q \geq d+1$. Let S be a black box PIT set for $\mathcal{M}$ over an extended field $\mathbb{F}_{q^t}$ of size w. There is an algorithm that runs in time $w \cdot poly(d, t)$ and constructs a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_q$ of size $O(d^{\tau_{poly}(d,q,t)} \cdot tw) = O(d^5 \cdot tw)$.*
3. *Let $\mathcal{M} \subseteq \mathcal{P}(\mathbb{F}_q, n, d)$ and $q^{t'} \geq d+1$. Let S be a black box PIT set for $\mathcal{M}$ over an extended field $\mathbb{F}_{q^t}$ of size w. There is an algorithm that runs in time $w \cdot poly(d, t)$ and constructs a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_{q^{t'}}$ of size $O(d^5 \cdot wt/t')$.*

We now prove

Lemma 64. *If $\mathcal{M} \subseteq \mathcal{P}(\mathbb{F}_q, n, d, r = p-1) \backslash \mathcal{P}(\mathbb{F}_q, n, d-1, r = p-1)$ where p is the characteristic of the field, $q \leq d$ and there is a black box PIT set for $\phi_d \mathcal{M} = \{\phi_d f \mid f \in \mathcal{M}\}$ over an extension field $\mathbb{F}_{q^t}$ of size w then there is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_q$ of size $2^d w \cdot \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$.*

In particular, let S be a black box PIT set for $\phi_d\mathcal{M}$ over an extended field $\mathbb{F}_{q^t}$ of size w. There is an algorithm that runs in time $w \cdot 2^{(1+c_q)d} \cdot poly(t)$ and constructs a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_q$ of size $O(d^7 2^{(1+c_q)d} \cdot tw)$.

Proof. Let $S \subseteq (\mathbb{F}_{q^t}^n)^d$ be a black box PIT set for $\phi_d\mathcal{M}$ of size w. Let $f \in \mathcal{P}(\mathbb{F}_q, n, d, r = p-1)$ be of degree d and $f \not\equiv 0$. By Lemma 61, $\phi_d f \not\equiv 0$. Since $\phi_d\mathcal{M} \subset \mathcal{DML}(\mathbb{F}_q, n, d)$ and S is a black box PIT set for $\phi_d\mathcal{M}$ over $\mathbb{F}_{q^t}$, by Lemma 62, there is a black box PIT set $R \subseteq (\mathbb{F}_q^n)^d$ for $\phi_d\mathcal{M}$ over $\mathbb{F}_q$ of size $|S| \cdot \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$. Therefore, for every $f \in \mathcal{M}$ there is $\boldsymbol{a} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_d) \in R$ such that $(\phi_d f)(\boldsymbol{a}) \neq 0$. Now, by (33), there is $J \subseteq [d]$ such that $f(\sum_{j \in J} \boldsymbol{a}_j) \neq 0$. Therefore

$$\left\{ \sum_{j \in J} \boldsymbol{a}_j \,\middle|\, J \subset [d], \boldsymbol{a} \in R \right\},$$

is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_q$ of size $2^d w \cdot \nu_{\mathbb{F}_q}(d, \mathbb{F}_{q^t})$.

The other result follows from Theorem 44. □

5.5 Lower and Upper Bounds

In this subsection we give lower and upper bounds for the size of black box PIT sets for classes of polynomials. We first give folklore results that are used in many papers. The folklore lower bound is information theoretic and the upper bound uses Schwartz-Zipple Lemma with the union bound. Then we give a new bound that improves Schwartz-Zipple Lemma and use it with the union bound to get better upper bounds.

5.5.1 Folklore Lower and Upper Bounds

In this subsection we give some folklore lower and upper bounds for the size of a black box PIT set.

The following is a folklore result. We prove it for completeness.

Lemma 65. *Let $\mathcal{M} \subseteq \mathbb{F}_q[x_1, x_2, \ldots, x_n]$ be a class of multivariate polynomials over $\mathbb{F}_q$. Let $\mathcal{M}' \subseteq \mathcal{M}$ such that $\mathcal{M}' - \mathcal{M}' = \{g - g' \mid g \neq g' \text{ and } g, g' \in \mathcal{M}'\} \subseteq \mathcal{M}$. Then any black box PIT $S \subseteq \mathbb{F}_{q^t}^n$ for $\mathcal{M}$ is of size*

$$|S| \geq \frac{\log |\mathcal{M}'|}{t \log q}.$$

Proof. Let $S = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_w\}$ be a black box PIT set for $\mathcal{M}$, and therefore for $\mathcal{M}'$. Let $\mathcal{M}' = \{g_1, \ldots, g_v\}$. Consider the elements $\boldsymbol{g}^{(i)} = (g_i(\boldsymbol{a}_1), \ldots, g_i(\boldsymbol{a}_w)) \in \mathbb{F}_{q^t}^w$ for $i = 1, \ldots, v$. If $|\mathbb{F}_{q^t}^w| < v$, then there is $i \neq j$ such that $\boldsymbol{g}^{(i)} = \boldsymbol{g}^{(j)}$. Then $0 \neq f = g_i - g_j \in \mathcal{M}$ satisfies $f(\boldsymbol{a}_i) = 0$ for all $i = 1, 2, \ldots, w$ and we get a contradiction. Therefore, we must have $|\mathbb{F}_{q^t}|^{|S|} = |\mathbb{F}_{q^t}^w| \geq v = |\mathcal{M}'|$. This implies the result. □

Another folklore result follows from Schwartz-Zippel Lemma, [60, 82, 49], and the union bound

Lemma 66. *Let $\mathcal{M} \subseteq \mathbb{F}_q[x_1, x_2, \ldots, x_n]$ be any class of multivariate polynomials over $\mathbb{F}_q$ of degree d. For t such that $q^t \geq d+1$ and any constant $c > 0$ there exists a black box PIT set $S \subseteq \mathbb{F}_{q^t}^n$ for $\mathcal{M}$ of size*

$$|S| \leq \frac{\log |\mathcal{M}|}{t \log q - \log d} = \begin{cases} \frac{\log |\mathcal{M}|}{t \log q} & \text{if} \quad q^t \geq d^{1+c} \\ \log |\mathcal{M}| & \text{if} \quad q^t \geq (1+c)d \\ d \log |\mathcal{M}| & \text{if} \quad q^t \geq d+1 \end{cases} .$$

Proof. By Schwartz-Zippel Lemma, for any $q^t \geq d+1$ and $f \in \mathcal{M}$

$$\mathbf{Pr}_{\boldsymbol{a} \in U(\mathbb{F}_{q^t}^n)}[f(\boldsymbol{a}) \neq 0] \geq 1 - \frac{d}{q^t}$$

where $U(\mathbb{F}_{q^t}^n)$ is the uniform distribution over $\mathbb{F}_{q^t}^n$. Now the result follows from union bound. □

In particular, we have the following upper and lower bounds[4]

Lemma 67. *For any positive constant c the following are bounds for the size of black box PIT set for $\mathcal{M}$ over $\mathbb{F}_{q^t}$*

[4]Here we assume that $s = o(\text{maximal possible size in the class})$. For example for $\mathcal{P}(\mathbb{F}_q, n, s)$, $s = o(q^n)$.

$\mathcal{M}$	$\lvert\mathbb{F}_{q^t}\rvert \geq$	*Lower Bound* $= \Omega(\ )$	*Upper Bound* $= O(\ )$
$\mathcal{P}(\mathbb{F}_q, n)$	$((q-1)n)^{1+c}$	q^n/t	q^n/t
$\mathcal{P}(\mathbb{F}_q, n)$	$(1+c)(q-1)n$	q^n/t	$\log(qn) \cdot q^n/t$
$\mathcal{P}(\mathbb{F}_q, n)$	$(q-1)n+1$	q^n/t	$(qn\log(qn)) \cdot q^n/t$
$\mathcal{P}(\mathbb{F}_q, n, r)$	$(rn)^{1+c}$	$(r+1)^n/t$	$(r+1)^n/t$
$\mathcal{P}(\mathbb{F}_q, n, r)$	$(1+c)rn$	$(r+1)^n/t$	$(\log rn) \cdot (r+1)^n/t$
$\mathcal{P}(\mathbb{F}_q, n, r)$	$rn+1$	$(r+1)^n/t$	$(rn(\log rn)) \cdot (r+1)^n/t$
$\mathcal{P}(\mathbb{F}_q, n, (d, r))$	d^{1+c}	$R(n,d,r)/t$	$R(n,d,r)/t$
$\mathcal{P}(\mathbb{F}_q, n, (d, r))$	$(1+c)d$	$R(n,d,r)/t$	$\log d \cdot R(n,d,r)/t$
$\mathcal{P}(\mathbb{F}_q, n, (d, r))$	$d+1$	$R(n,d,r)/t$	$(d\log d) \cdot R(n,d,r)/t$
$\mathcal{P}(\mathbb{F}_q, n, s)$	$((q-1)n)^{1+c}$	$n \cdot \frac{s}{t}$	$n \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, s)$	$(1+c)(q-1)n$	$n \cdot \frac{s}{t}$	$(\log n)n \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, s)$	$(q-1)n+1$	$n \cdot \frac{s}{t}$	$(\log n)qn^2 \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r, s)$	$(rn)^{1+c}$	$\frac{n\log(r+1)}{\log q} \cdot \frac{s}{t}$	$\frac{n\log(r+1)}{\log q} \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r, s)$	$(1+c)rn$	$\frac{n\log(r+1)}{\log q} \cdot \frac{s}{t}$	$\frac{n(\log(rn))\log(r+1)}{\log q} \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r, s)$	$rn+1$	$\frac{n\log(r+1)}{\log q} \cdot \frac{s}{t}$	$\frac{n^2 r(\log(rn))\log(r+1)}{\log q} \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$	d^{1+c}	$\frac{d\log n}{\log q} \cdot \frac{s}{t}$	$\frac{d\log n}{\log q} \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$	$(1+c)d$	$\frac{d\log n}{\log q} \cdot \frac{s}{t}$	$\frac{d(\log d)\log n}{\log q} \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$	$d+1$	$\frac{d\log n}{\log q} \cdot \frac{s}{t}$	$\frac{d^2(\log d)\log n}{\log q} \cdot \frac{s}{t}$

where $R(n,d,r)$ is the number of monomials in $\mathbb{F}_q[x_1,\ldots,x_n]$ of degree at most d and variable degree at most r.

Proof. We first use Lemma 65 to prove the lower bounds. For $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n)$ or $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, r)$ we take $\mathcal{M}' = \mathcal{M}$. For $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, s)$ we take $\mathcal{M}' = \mathcal{P}(\mathbb{F}_q, n, \lfloor s/2 \rfloor)$ and for $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, r, s)$ we take $\mathcal{M}' = \mathcal{P}(\mathbb{F}_q, n, r, \lfloor s/2 \rfloor)$.

Consider the class $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, (d, r), s)$. Let $\mathcal{M}'$ be the class of multilinear polynomials in $\mathcal{P}(\mathbb{F}_q, n, (d, 1), \lfloor s/2 \rfloor)$ that contain no monomials of degree less than d. Then $\mathcal{M}' - \mathcal{M}' \subseteq \mathcal{M}$. Therefore any black box set for $\mathcal{M}$ must be of size at least

$$\frac{\log|\mathcal{M}'|}{\log|\mathbb{F}_{q^t}|} = \Omega\left(\frac{ds\log n}{t\log q}\right). \tag{35}$$

For $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, (d, r))$ we take the class of polynomials $\mathcal{M}' \subseteq \mathcal{P}(\mathbb{F}_q, n, (d, r))$ that contain no monomials of degree less than d. We have $\mathcal{M}' - \mathcal{M}' \subseteq \mathcal{M}$ and therefore any black box PIT set for $\mathcal{M}$ is of size

$$\frac{\log|\mathcal{M}'|}{\log|\mathbb{F}_{q^t}|} \geq \frac{R(n,d,r)}{t}.$$

For the upper bounds we use Lemma 66. □

Notice that to use Schwartz-Zippel Lemma the size of the field $|\mathbb{F}_{q^t}| = q^t$ must be at least $d+1$ where d is an upper bound on the degree of the polynomials in the class $\mathcal{M}$. In the next subsection we give a new result that improves Schwartz-Zippel Lemma. This result will first, improve the above upper bound when $q^t \geq d+1$, and second, improve the minimum size of the field to the variable degree r (rather than the total degree $d+1$).

5.5.2 New Non-Constructive Upper Bounds

In this subsection we first give a new non-constructive upper bound for the size of a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r, s)$ over $\mathbb{F}_{q^t}$. Although our analysis also gives new randomized algorithms for black box PIT even over $\mathbb{F}_{q^2}$, we will defer those results to our future works [12, 13, 14]. We then give a new non-constructive upper bound for the size of a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, (d, r))$.

For $f \in \mathcal{P}(\mathbb{F}_q, n, r, s)$, by Schwartz-Zippel Lemma, [60, 82, 49],

$$\mathbf{Pr}_{\boldsymbol{x} \in U(S^n)}[f(\boldsymbol{x}) \neq 0] \geq 1 - \frac{rn}{|S|}$$

where $S \subseteq \mathbb{F}_{q^t}$ and $U(S^n)$ is the uniform distribution over S^n.

The following Lemma improves this bound

Lemma 68. *Let $f \in \mathcal{P}(\mathbb{F}_q, n, r, s)$ be any non-zero polynomial. Then*

$$\mathbf{Pr}_{\boldsymbol{a} \in U(S^n)}[f(\boldsymbol{a}) \neq 0] \geq \left(1 - \frac{r}{|S|}\right)^{\min(n, \lfloor \log s \rfloor)} \geq 1 - \frac{r \min(n, \log s)}{|S|}$$

where $S \subseteq \mathbb{F}_{q^t} \backslash \{0\}$ and $U(S^n)$ is the uniform distribution over S^n.

In particular, for any non-zero polynomial $f \in \mathcal{P}(\mathbb{F}_q, n, s)$ we have

$$\mathbf{Pr}_{\boldsymbol{a} \in U(S^n)}[f(\boldsymbol{a}) \neq 0] \geq \left(1 - \frac{q-1}{|S|}\right)^{\min(n, \lfloor \log s \rfloor)}$$

Proof. We prove the result for any $f \in \mathcal{P}(\mathbb{F}_{q^t}, n, r, s)$. The proof will be by induction on n. For $n = 0$, f is a non-zero constant polynomial and $s = 1$ and the above probability is 1. For $n = 1$, if $s = 1$ then $f(x_1) = \lambda x_1^d$ for some $\lambda \neq 0$ and $r \geq d \geq 0$ and therefore $\mathbf{Pr}_{\boldsymbol{a} \in U(S)}[f(\boldsymbol{a}) \neq 0] = 1$. If $s > 1$ then since $f(x_1)$ is of degree at most r, it has at most r roots in $\mathbb{F}_{q^t} \backslash \{0\}$ and

$$\mathbf{Pr}_{\boldsymbol{a} \in U(S)}[f(\boldsymbol{a}) \neq 0] \geq \left(1 - \frac{r}{|S|}\right).$$

Suppose the lemma is true for $n-1$ and consider $f \in \mathcal{P}(\mathbb{F}_{q^t}, n, r, s)$. We write f in the form $f = p_1(x_n)M_1 + \cdots + p_{s'}(x_n)M_{s'}$ where $p_1, \ldots, p_{s'} \in \mathbb{F}_{q^t}[x_n]$ and $M_1, \ldots, M_{s'} \in \mathbb{F}_{q^t}[x_1, \ldots, x_{n-1}]$ are $s' \leq s$ distinct monomials. Now we have two cases: If for some $i \leq s'$, $p_i(x_n)$ contains one monomial then for any $a_n \in \mathbb{F}_{q^t} \backslash \{0\}$, $f' := f(x_1, \ldots, x_{n-1}, a_n) \not\equiv 0$. This is because $p_i(a_n) \neq 0$ and then f' must contain the monomial M_i. In this case, by the induction hypothesis,

$$\begin{aligned}
\mathbf{Pr}_{\boldsymbol{a} \in U(S^n)}[f(\boldsymbol{a}) \neq 0] &= \sum_{a_n \in S} \frac{1}{|S|} \mathbf{Pr}_{\boldsymbol{a}' \in U(S^{n-1})}[f(\boldsymbol{a}', a_n) \neq 0] \\
&\geq \left(1 - \frac{r}{|S|}\right)^{\min(n-1, \lfloor \log s' \rfloor)} \\
&\geq \left(1 - \frac{r}{|S|}\right)^{\min(n, \lfloor \log s \rfloor)}.
\end{aligned}$$

The other case is when every polynomial $p_i(x_n)$ contains at least two monomials. In that case $s' \leq \lfloor s/2 \rfloor$ and for any $a_n \in \mathbb{F}_{q^t}$ the polynomial $f(x_1, \ldots, x_{n-1}, a_n)$ is of size at most $\lfloor s/2 \rfloor$. Since $\deg(p_1) \leq r$, for $T = \{\beta \in S | p_1(\beta) = 0\}$ we have $|T| \leq r$ and by the induction hypothesis

$$\begin{aligned}
\mathbf{Pr}_{\boldsymbol{a} \in U(S^n)}[f(\boldsymbol{a}) \neq 0] &= \sum_{a_n \in S \backslash T} \frac{1}{|S|} \mathbf{Pr}_{\boldsymbol{a}' \in U(S^{n-1})}[f(\boldsymbol{a}', a_n) \neq 0] \\
&\geq \left(1 - \frac{r}{|S|}\right)\left(1 - \frac{r}{|S|}\right)^{\min(n-1, \lfloor \log s' \rfloor)} \\
&= \left(1 - \frac{r}{|S|}\right)^{\min(n, \lfloor \log s \rfloor)}.
\end{aligned}$$

□

For $f \in \mathcal{P}(\mathbb{F}_q, n, (d, r))$, Schwartz-Zippel Lemma gives

$$\mathbf{Pr}_{\boldsymbol{x} \in U(S)}[f(\boldsymbol{x}) \neq 0] \geq 1 - \frac{d}{|S|}.$$

Now we prove the following better bound

Lemma 69. *Let $f \in \mathcal{P}(\mathbb{F}_q, n, (d, r))$, $r \geq 1$ be any non-zero polynomial. Then*

$$\mathbf{Pr}_{\boldsymbol{x} \in U(S^n)}[f(\boldsymbol{x}) \neq 0] \geq \left(1 - \frac{r}{|S|}\right)^{d/r}$$

where $S \subseteq \mathbb{F}_{q^t}$ and $U(S^n)$ is the uniform distribution over S^n.

Proof. We prove the lemma for $f \in \mathcal{P}(\mathbb{F}_{q^t}, n, (d, r))$, $r \geq 1$ by induction on n. For $n = 0$ the polynomial f is nonzero constant and $\mathbf{Pr}[f \neq 0] = 1$. If $n = 1$ then the polynomial f is in one variable x_1. Then it has at most r roots, $d = r$ and the probability that $f(\boldsymbol{x}) \neq 0$ for $\boldsymbol{x} \in U(S)$ is at least $(|S|-r)/|S| = 1-r/|S|$.

Suppose the result is true for all integers less than or equal to $n-1$. We now prove it for n. Let A be the set of all the elements α in S such that $f|_{x_n=\alpha}$ is identically zero. Since $f \in \mathbb{F}_{q^t}[x_1, x_2, \ldots, x_{n-1}][x_n]$ is of degree r in x_n we have $t := |A| \leq r$. Also

$$f(\boldsymbol{x}) = \left(\prod_{\alpha \in A} (x_n - \alpha) \right) g(\boldsymbol{x})$$

where $g(\boldsymbol{x}) \in \mathbb{F}_{q^t}[x_1, x_2, \ldots, x_n]$ is of degree $d-t$. For $\beta \notin A$ let r_β and $d_\beta \leq d-t$ be the variable degree and total degree of $g|_{x_n=\beta}$, respectively. Since $\phi(x) = (1-x)^{1/x}$ is a monotonically decreasing function in $[0,1]$, $0 \leq r_\beta \leq r \leq q-1$ and $t \leq r \leq q-1$ we have

$$\left(1 - \frac{r_\beta}{|S|}\right)^{1/r_\beta} \geq \left(1 - \frac{r}{|S|}\right)^{1/r} \text{ and } 1 - \frac{t}{|S|} \geq \left(1 - \frac{r}{|S|}\right)^{t/r}$$

and therefore by the induction hypothesis we have

$$\begin{aligned}
\mathbf{Pr}_{\boldsymbol{x} \in U(S^n)}[f(\boldsymbol{x}) \neq 0] &= \mathbf{Pr}[x_n \notin A] \cdot \mathbf{Pr}_{\boldsymbol{x} \in U(S^n)}\left[f(\boldsymbol{x}) \neq 0 \mid x_n \notin A\right] \\
&= \left(1 - \frac{t}{|S|}\right) \sum_{\beta \notin A} \frac{1}{|S|-t} \mathbf{Pr}_{\boldsymbol{x} \in U(S^n)}[f|_{x_n=\beta}(\boldsymbol{x}) \neq 0] \\
&= \left(1 - \frac{t}{|S|}\right) \sum_{\beta \notin A} \frac{1}{|S|-t} \mathbf{Pr}_{\boldsymbol{x} \in U(S^n)}[g|_{x_n=\beta}(\boldsymbol{x}) \neq 0] \\
&\geq \left(1 - \frac{t}{|S|}\right) \sum_{\beta \notin A} \frac{1}{|S|-t} \left(1 - \frac{r_\beta}{|S|}\right)^{d_\beta/r_\beta} \\
&\geq \left(1 - \frac{t}{|S|}\right) \sum_{\beta \notin A} \frac{1}{|S|-t} \left(1 - \frac{r_\beta}{|S|}\right)^{(d-t)/r_\beta} \\
&\geq \left(1 - \frac{t}{|S|}\right) \sum_{\beta \notin A} \frac{1}{|S|-t} \left(1 - \frac{r}{|S|}\right)^{(d-t)/r} \\
&\geq \left(1 - \frac{r}{|S|}\right)^{t/r} \cdot \left(1 - \frac{r}{|S|}\right)^{(d-t)/r} = \left(1 - \frac{r}{|S|}\right)^{d/r}.
\end{aligned}$$

This completes the proof. □

Now by Lemma 68, Lemma 69 and since $d/r \leq n$ we have

Lemma 70. *Let $f \in \mathcal{P}(\mathbb{F}_q, n, (d,r), s)$ be any non-zero polynomial. Then*

$$\mathbf{Pr}_{\boldsymbol{x} \in U(S^n)}[f(\boldsymbol{x}) \neq 0] \geq \left(1 - \frac{r}{|S|}\right)^{\min(d/r, \lfloor \log s \rfloor)}$$

where $S \subseteq \mathbb{F}_{q^t} \backslash \{0\}$ and $U(S^n)$ is the uniform distribution over S^n.

Using the union bound with Lemma 70 we get

Lemma 71. *For $\mathcal{M} \subseteq \mathcal{P}(\mathbb{F}_q, n, (d,r), s)$, if*

$$|\mathcal{M}| \left(1 - \left(1 - \frac{r}{q^t - 1}\right)^{\min(d/r, \lfloor \log s \rfloor)}\right)^m < 1$$

then there is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_{q^t}$ of size m.

In particular, $m = O(M)$ where

$$M = \begin{cases} \left(\frac{q^t-1}{q^t-1-r}\right)^{\min(d/r, \log s)} \log |\mathcal{M}| & \text{if } \frac{q^t-1}{r} = O(1), \frac{d}{r} = \omega(1) \\ \log |\mathcal{M}| & \text{if } \frac{q^t-1}{r} = O(1), \frac{d}{r} = O(1) \\ e^{\frac{\min(d, r \log s)}{q^t - 1}} \log |\mathcal{M}| & \text{if } \frac{q^t-1}{r} = \omega(1), \frac{q^t-1}{\min(d, r \log s)} = o(1) \\ \log |\mathcal{M}| & \text{if } \frac{q^t-1}{r} = \omega(1), \frac{q^t-1}{\min(d, r \log s)} = O(1) \\ \frac{\log |\mathcal{M}|}{t \log q - \log \min(d, r \log s)} & \text{if } \frac{q^t-1}{r} = \omega(1), \frac{q^t-1}{\min(d, r \log s)} = \omega(1) \end{cases}$$

where the O, o and ω are with respect to the parameters s and d.

The above bound improves the upper bounds in Lemma 67 when q^t is close to the degree of f. In particular we have

Lemma 72. *For any positive constant c there is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_{q^t}$ of size $O(m)$ where m is as given in the following table*

$\mathcal{M}$	$\lvert\mathbb{F}_{q^t}\rvert \geq$	*Lower Bound*	*Upper Bound Lemma 67*	*New Upper Bound m*
$\mathcal{P}(\mathbb{F}_q, n)$	cqn	q^n/t	$(qn \log(qn)) \cdot q^n/t$	$(\log(qn)) \cdot q^n/t$
$\mathcal{P}(\mathbb{F}_q, n, r)$	crn	$(r+1)^n/t$	$(rn \log(rn)) \cdot (r+1)^n/t$	$(\log(rn)) \cdot (r+1)^n/t$
$\mathcal{P}(\mathbb{F}_q, n, (d,r))$	cd	$R(n,d,r)/t$	$(d \log d) \cdot R(n,d,r)/t$	$(\log d) \cdot R(n,d,r)/t$
$\mathcal{P}(\mathbb{F}_q, n, s)$	cqn	$n \cdot \frac{s}{t}$	$qn^2 \log n \cdot \frac{s}{t}$	$n \log n \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, r, s)$	crn	$\frac{n \log(r+1)}{\log q} \cdot \frac{s}{t}$	$\frac{n^2 r (\log n) \log(r+1)}{\log q} \cdot \frac{s}{t}$	$\frac{n (\log n) \log(r+1)}{\log q} \cdot \frac{s}{t}$
$\mathcal{P}(\mathbb{F}_q, n, (d,r), s)$	cd	$\frac{d \log n}{\log q} \cdot \frac{s}{t}$	$\frac{d^2 (\log d)(\log n)}{\log q} \cdot \frac{s}{t}$	$\frac{d (\log d) \log n}{\log q} \cdot \frac{s}{t}$

In the next Lemma we show that using our new bounds one can also get an efficient black box PIT set size even if the field extension is 2. This also gives new randomized algorithms over small extension fields that will be discussed in future works [12, 13, 14].

Lemma 73. *There is a black box PIT set for $\mathcal{M}$ over $\mathbb{F}_{q^2}$ of size $O(m)$ where m is as given in the following table*

$\mathcal{M}$	m	*Lower Bound*
$\mathcal{P}(\mathbb{F}_q, n)$	$(\log q)(q+1)^n$	q^n
$\mathcal{P}(\mathbb{F}_q, n, r)$	$(\log q)\left(r+1+\frac{r^2}{q^2}+O\left(\frac{r}{q^2}\right)\right)^n$	$(r+1)^n$
$\mathcal{P}(\mathbb{F}_q, n, r=1)$	$(\log q)\left(2+\frac{2}{q^2-2}\right)^n$	2^n
$\mathcal{P}(\mathbb{F}_2, n, s)$	$n \cdot s^{1.585}$	$n \cdot s$
$\mathcal{P}(\mathbb{F}_q, n, s)$	$n \cdot s^{1+\log\left(\frac{q+1}{q}\right)}$	$n \cdot s$
$\mathcal{P}(\mathbb{F}_q, n, r=1, s)$	$n \cdot s^{1+\log((q^2-1)/(q^2-2))}$ $= n \cdot s^{1+O(1/q^2)}$	$\frac{n \cdot s}{\log q}$
$\mathcal{P}(\mathbb{F}_q, n, r, s)$	$n \cdot s^{1+\log((q^2-1)/(q^2-1-r))}$ $= n \cdot s^{1+O(r/q^2)}$	$\frac{n \log(r+1) \cdot s}{\log q}$
$\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$	$d(\log n) s^{1+\log((q^2-1)/(q^2-1-r))}$ $= d(\log n) \cdot s^{1+O(r/q^2)}$	$\frac{d(\log n) \cdot s}{\log q}$

5.6 Constructive Upper Bound for $\mathcal{P}(\mathbb{F}_q, n)$ and $\mathcal{P}(\mathbb{F}_q, n, r)$

In this subsection we give polynomial time constructions of a black box PIT sets for $\mathcal{P}(\mathbb{F}_q, n)$ and $\mathcal{P}(\mathbb{F}_q, n, r)$ of size that asymptotically match the lower bounds in Lemma 67. We give two constructions. The first is a very simple construction that uses combinatorial Nullstellensatz [1] and is asymptotically tight for small fields and the other one uses testers and is asymptotically tight for large fields. Notice that since $\mathcal{P}(\mathbb{F}_q, n) = \mathcal{P}(\mathbb{F}_q, n, r = q-1)$, it is enough to study the class $\mathcal{P}(\mathbb{F}_q, n, r)$.

The first construction uses the following combinatorial Nullstellensatz

Lemma 74. *([1]) Let $\mathbb{F}$ be a field and f be a multivariate polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ of total degree d. Let $x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$ be a monomial in f of degree d. If $S_1, \cdots, S_n$ are subsets of $\mathbb{F}$ with $|S_i| = r_i + 1$, there is $s_1 \in S_1, s_2 \in S_2, \ldots, s_n \in S_n$ so that $f(s_1, \ldots, s_n) \neq 0$.*

This immediately gives the following construction

Lemma 75. *Let $\mathbb{F}_q$ be any field. Let $S \subseteq \mathbb{F}_{q^t}$ be a set of size $r+1$. Then S^n is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r)$ of size $(r+1)^n$.*

We now use testers to prove the following

Lemma 76. *For any q there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r)$ over $\mathbb{F}_{q^t}$ of size at most*

$$2n\frac{(r+1)^n}{t}$$

that can be constructed in $(r+1)^n poly(t,n)$ time.

In particular, there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n)$ over $\mathbb{F}_{q^t}$ of size at most

$$2n\frac{q^n}{t}$$

that can be constructed in $q^n poly(t,n)$ time.

Proof. If $t \leq 2n$ then the black box PIT set in Lemma 75 is of size $(r+1)^n \leq 2n \cdot (r+1)^n/t$. Therefore we may assume that $t > 2n$.

Consider the linear space $\mathbb{F}_q[x]_{T-1}$ where $T = (r+1)^{n-1}+1$. Since for two distinct monomials M_1 and M_2 with variable degree r we have

$$M_1(x^1, x^{(r+1)}, x^{(r+1)^2} \ldots, x^{(r+1)^{n-1}}) \not\equiv M_2(x^1, x^{(r+1)}, x^{(r+1)^2} \ldots, x^{(r+1)^{n-1}})$$

for $f \in \mathcal{P}(\mathbb{F}_q, n, r)$ we have $f \not\equiv 0$ if and only if $h(x) := f(x^1, x^{(r+1)}, x^{(r+1)^2} \ldots, x^{(r+1)^{n-1}}) \not\equiv 0$. This gives a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r)$ over $\mathbb{F}_q[x]_{T-1}$ of size 1.

By Lemma 62, since $f \in \mathcal{P}(\mathbb{F}_q, n, d = rn)$ and $q^t \geq q^{2n+1} \geq rn+1$ there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r)$ over $\mathbb{F}_{q^t}$ of size

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}\left((rn, \mathbb{F}_q), \mathbb{F}_q[x]_{T-1}\right).$$

By Lemma 9, since

$$q^{t-1} \geq q^{2n} \geq rnT - rn + 1$$

we have

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}\left((rn, \mathbb{F}_q), \mathbb{F}_q[x]_{T}\right) \leq (rn+1)\frac{T-1}{t} = \frac{rn+1}{r+1}\frac{(r+1)^n}{t} \leq 2n\frac{(r+1)^n}{t}.$$

□

5.7 Constructive Upper Bound for $\mathcal{P}(\mathbb{F}_q, n, d)$ and $\mathcal{P}(\mathbb{F}_q, n, (d, r))$

In this subsection we give a black box PIT set for the class $\mathcal{P}(\mathbb{F}_q, n, (d, r))$ for every d, r and q. We give two constructions. The first is a very simple construction that uses combinatorial Nullstellensatz lemma and the second one uses testers. For small d, the size of the black box PIT set asymptotically matches the lower bound in Lemma 65.

By Lemma 67 and 72 we have that for $q^t \geq d+1$, a minimum size black box PIT set S over $\mathbb{F}_{q^t}$ for $\mathcal{P}(\mathbb{F}_q, n, (d, r))$ satisfies

$$\frac{R(n,d,r)}{t} \leq |S| \leq (\log d)\frac{R(n,d,r)}{t}$$

where $R(n,d,r)$ is the number of monomials in $\mathbb{F}_q[x_1, \ldots, x_n]$ of degree at most d and variable degree at most r. The function $R(n,d,r)$ satisfies the following bounds

$$\sum_{i=0}^{d} \binom{n}{i} = R(n,d,1) \leq R(n,d,r) \leq R(n,d,d) = \binom{n+d}{d}.$$

The combinatorial Nullstellensatz lemma gives the following construction

Lemma 77. *Let $S = \{\sigma_0, \sigma_1, \ldots, \sigma_r\} \subseteq \mathbb{F}$ where $|S| = r+1$. The set*

$$S = \{(\sigma_{i_1}, \ldots, \sigma_{i_n}) \mid 0 \leq i_0 \leq d,\ i_j \leq r \text{ for } j = 1, \ldots, n, \quad \text{and} \quad i_0 + i_1 + i_2 + \cdots + i_n = d\}$$

is a black box PIT set for $\mathcal{P}(\mathbb{F}, n, (d, r))$ of size

$$|S| = R(n,d,r).$$

Proof. Let $f \in \mathcal{P}(\mathbb{F}, n, (d, r))$ and let $x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$ be a monomial in f of degree $d' = \deg(f)$. Consider $S_j = \{\sigma_0, \ldots, \sigma_{r_j}\} \subseteq S$ for $j = 1, \ldots, n$. By Lemma 74 there is $\sigma_{i_1} \in S_1, \sigma_{i_2} \in S_2, \ldots, \sigma_{i_n} \in S_n$ so that $f(\sigma_{i_1}, \ldots, \sigma_{i_n}) \neq 0$. Since $\sigma_{i_j} \in S_j$ we have $i_j \leq r_j \leq r$. Also $i_1 + \cdots + i_n \leq r_1 + \cdots + r_n = d'$ and for $i_0 = d - (i_1 + \cdots + i_n)$ we have $i_0 + i_1 + \cdots + i_n = d$ and therefore $(\sigma_{i_1}, \ldots, \sigma_{i_n}) \in S$. □

Our second construction uses testers. We prove

Lemma 78. *For any $q^t \geq d+1$ there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d)$ over $\mathbb{F}_{q^t}$ of size*

$$\frac{d((1+o(1))n)^{d+1}}{t}$$

that can be constructed in time $n^{d+2} \cdot poly(t, d)$.

Proof. If $t \leq dn$ then by Lemma 78 we get the result. Therefore, we may assume that $t > dn$. Let $f \in \mathcal{P}(\mathbb{F}_q, n, d)$ where $q^t \geq d+1$. Then

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} \cdot x_{i_1} \cdots x_{i_d}$$

where $\boldsymbol{x} = (x_0, x_1, \ldots, x_n)$, $x_0 = 1$, $I \subseteq \{(i_1, i_2, \ldots, i_d) \mid 0 \leq i_1 \leq i_2 \leq \cdots \leq i_d \leq n\}$ and $\lambda_{\boldsymbol{i}} \in \mathbb{F}_q$. Note here that we are using x_0 to create monomials of degree less than d.

Consider the Sidon $B_{\leq d}$ sequence $a_1, a_2, \cdots, a_n$ in Lemma 59 where $\max_j a_j \leq N-1 = (n+o(n))^{d+1}$. In Lemma 59 we have shown that this sequence can be constructed in time $n^{d/2+1.5}$. Consider the assignment $\boldsymbol{x}' = (x^{a_0}, x^{a_1}, x^{a_2}, \ldots, x^{a_d})$ where $a_0 = 0$ and x is a new indeterminate. Then

$$f(\boldsymbol{x}') = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} \cdot x^{a_{i_1}+a_{i_2}+\cdots+a_{i_d}}.$$

Notice that $n_{\boldsymbol{i}} := a_{i_1} + a_{i_2} + \cdots + a_{i_d}$, $\boldsymbol{i} \in I$ are distinct. Therefore, $f(\boldsymbol{x}') \neq 0$. That is, $S = \{\boldsymbol{x}'\}$ is a black box PIT set over $\mathbb{F}_q[x]_{N-1}$ of size 1. By Lemma 63 and 9 there is a black box PIT set over $\mathbb{F}_{q^t}$ of size

$$\nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}((d, \mathbb{F}_q), \mathbb{F}_q[x]_{N-1}) \leq \frac{dN-d+1}{t} + 1 \leq \frac{d((1+o(1))n)^{d+1}}{t}.$$

□

5.8 Constructive Upper Bounds for $\mathcal{P}(\mathbb{F}_q, n, s)$ and $\mathcal{P}(\mathbb{F}_q, n, r, s)$

In this subsection we construct black box PIT sets for $\mathcal{P}(\mathbb{F}_q, n, s)$ and $\mathcal{P}(\mathbb{F}_q, n, r, s)$.

In [24], Clausen et. al. showed that any black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, s)$ over $\mathbb{F}_q$ has size at least $\Omega((n/\log s)^{\log s})$. Then they gave a black box PIT set of size $s^{O(\log\log q)}(n/\log s)^{\log s}$ [80]. They also show that there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, s)$ over $\mathbb{F}_{q^t}$ where $t \geq n$ of size $s+1$. Constructing the latter set requires constructing an element of the field with multiplicative order at least q^n. In [34], Grigoriev gave a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, s)$ over $\mathbb{F}_{q^t}$, where $t \geq 2\log_q(ns)+4$, of size $O(qn^2s^3)$. It follows from [45] that there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, s)$ over $\mathbb{F}_{q^t}$ where $t \geq 12\log_q n + O(1)$ of size $(qns)^c$ for some constant $c \geq 2$. All the above results give black box PIT sets for $\mathcal{P}(\mathbb{F}_q, n, s)$ of size that are at least quadratic in s. In this subsection we give a black box PIT set for this class over $\mathbb{F}_{q^t}$, where $t \geq \lceil \log_q n \rceil + 2$, of size that is linear in s/t in deterministic time $poly(n) \cdot s$.

We prove the following.

Lemma 79. *We have*

1. *There is an explicit black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r, s)$ over $\mathbb{F}_{q^t}$ of size*

$$\begin{cases} O\left(\frac{rn^2\log(r+1)}{\log q} \cdot \frac{s}{t}\right) & \text{if} \quad t \geq 2\log_q n + 2 \\ O\left(\frac{r^2n^3\log(r+1)}{\log q} \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \log_q n + O(\log_q \log_q n) \\ O\left(\frac{r^5n^6\log(r+1)}{\log q} \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \log_q n + 2. \end{cases}$$

   In particular, of size $poly(n) \cdot s/t$ for $t \geq \log_q n + 2$.

2. There is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r, s)$ over $\mathbb{F}_{q^t}$ of size

$$\begin{array}{ll} O\left(q^{24} r n^3 \log^2 r \cdot \frac{s}{t}\right) & \text{if} \quad t \geq 2\log_q n + 2 \\ O\left(q^{24} r^2 n^4 \log^2 r \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \log_q n + O(\log_q \log_q n) \\ O\left(q^{24} r^5 n^7 \log^2 r \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \log_q n + 2 \end{array}$$

that can be constructed in deterministic time $s \cdot poly(n)$. In particular, of size $poly(n) \cdot s/t$, for $t \geq \log_q n + 2$, that can be constructed in deterministic time $s \cdot poly(n)$.

Proof. Let $f \in \mathcal{P}(\mathbb{F}_q, n, r, s)$. As in the proof of Lemma 76, since the variable degree of f is at most r, f is not equivalent to 0 if and only if

$$g(x) := f\left(x^1, x^{(r+1)}, x^{(r+1)^2}, \ldots, x^{(r+1)^{n-1}}\right)$$

is not equivalent to 0. Also, $g(x)$ is of degree at most $(r+1)^n - 1$ and contains at most s monomials. Let

$$T = \left\lceil \frac{n \log(r+1)}{\log q} \right\rceil.$$

Let α be an element of multiplicative order at least $(r+1)^n$ in $\mathbb{F}_{q^T}$. Then one of the values $g(\alpha^0), g(\alpha^1), \ldots, g(\alpha^{s-1})$ is not zero. Therefore there exists an explicit black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r, s)$ over $\mathbb{F}_{q^T}$ of size s. By Lemma 62 and Lemma 45, if $q^t \geq rn + 1$, then there is an explicit black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r, s)$ over $\mathbb{F}_{q^t}$ of size

$$s \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}((rn, \mathbb{F}_q), \mathbb{F}_{q^T}) = \begin{cases} O\left(\frac{rn^2 s \log(r+1)}{t \log q}\right) & \text{if} \quad t \geq 2\log_q n + 2 \\ O\left(\frac{r^2 n^3 s \log(r+1)}{t \log q}\right) & \text{if} \quad t \geq \log_q n + O(\log_q \log_q n) \\ O\left(\frac{r^5 n^6 s \log(r+1)}{t \log q}\right) & \text{if} \quad t \geq \log_q n + 2. \end{cases}$$

We now prove *2.* By Lemma 57, for

$$M = \frac{1}{100} q^{24} (n \log(r+1) + 25)^2 = O(q^{24} (\log r)^2 n^2)$$

an element α in $\mathbb{F}_{q^N}$ for some $M \leq N \leq M + o(M)$ of multiplicative order $(r+1)^n$ can be constructed in time $poly(n)$. If $f(x)$ is not identically zero then g is not identically zero and then one of the values $g(\alpha^0), g(\alpha^1), \ldots, g(\alpha^{s-1})$ is not zero. Therefore there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, s)$ over $\mathbb{F}_{q^N}$ of size s that can be constructed in time $s \cdot poly(n)$.

By Lemma 62 and Lemma 45, if $q^t \geq rn+1$, then there is an explicit black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, r, s)$ over $\mathbb{F}_{q^t}$ of size

$$s \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}((rn, \mathbb{F}_q), \mathbb{F}_{q^M}) = \begin{cases} O\left(\frac{q^{24} rn^3 s \log^2 r}{t}\right) & \text{if} \quad t \geq 2\log_q n + 2 \\ O\left(\frac{q^{24} r^2 n^4 s \log^2 r}{t}\right) & \text{if} \quad t \geq \log_q n + O(\log_q \log_q n) \\ O\left(\frac{q^{24} r^5 n^7 s \log^2 r}{t}\right) & \text{if} \quad t \geq \log_q n + 2 \end{cases}$$

that can be constructed in time $s \cdot poly(n)$. □

Using testers with the result in [34], one can get a black box PIT set over $\mathbb{F}_{q^t}$ where $t \geq \lceil \log_q n \rceil + 2$ of size $\tilde{O}(qn^2 s^3)$, which is better than the above size for small s. Since, in this paper, we are seeking black box PIT sets of size that is linear in s (see the note in Section 6) we will not state this result here. For small s, better black box PIT set size and smaller field extension are achieved in [12].

5.9 Constructive Upper Bound for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ and $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$

In this subsection we construct black box PIT sets for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ for every q, s and d.

Klivans and Spielman, [45], gave a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^t}$, where $t \geq 6 \log n + 6 \log d$, of size $(ns)^c$ for some constant $c \geq 2$ that can be constructed in polynomial time $poly(n, s)$. Then Bogdanov [9] gave a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^t}$, where $t \geq 18 \log d + 18 \log \log n$, of size $(s \log n)^c$ for some constant $c \geq 2$ that can be constructed in polynomial time $poly(n, s)$. In this subsection we give a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^t}$, where $t \geq \lceil \log(d+1)/\log q \rceil$, of size $poly(d, \log n) \cdot s/t$ in time $s \cdot poly(n)$. In particular, when $q \geq d+1$ then $t = 1$ and therefore no extension of the field is needed.

We prove

Lemma 80. *We have*

1. *There is an explicit black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^t}$, where $t \geq \lceil \log_q(d+1) \rceil$, of size*

$$\begin{array}{ll} O\left(\frac{d^2 \log n}{\log q} \cdot \frac{s}{t}\right) & \text{if} \quad t \geq 2 \log_q d + \log_q \log_q n + 2 \\ O\left(\frac{d^3 \log n}{\log q} \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \log_q d + \log_q \log_q \log_q n + 2 \\ O\left(\frac{d^6 \log n}{\log q} \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \lceil \log_q(d+1) \rceil . \end{array}$$

2. *There is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^t}$, where $t \geq \lceil \log_q(d+1) \rceil$, of size*

$$\begin{array}{ll} O\left(q^{24} d^3 \log^2 n \cdot \frac{s}{t}\right) & \text{if} \quad t \geq 2 \log_q d + \log_q \log_q n + 2 \\ O\left(q^{24} d^4 \log^2 n \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \log_q d + \log_q \log_q \log_q n + 2 \\ O\left(q^{24} d^7 \log^2 n \cdot \frac{s}{t}\right) & \text{if} \quad t \geq \lceil \log_q(d+1) \rceil \end{array}$$

   that can be constructed in deterministic time $s \cdot poly(n)$.

3. *In particular, when $q \geq d+1$ then there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_q$ of size $poly(q, \log n) \cdot s$ that can be constructed in deterministic time $s \cdot poly(n)$.*

Proof. Let $f \in \mathcal{P}(\mathbb{F}_q, n, d, s)$ and t be integer such that $q^t \geq d+1$. Then

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} \cdot x_{i_1} \cdots x_{i_d}$$

where $\boldsymbol{x} = (x_0, x_1, \ldots, x_n)$, $x_0 = 1$, $I \subseteq \{(i_1, i_2, \ldots, i_d) \mid 0 \leq i_1 \leq i_2 \leq \cdots \leq i_d \leq n\}$, $|I| \leq s$ and $\lambda_{\boldsymbol{i}} \in \mathbb{F}_q$. Note here that we are using x_0 to create monomials of degree less than d.

Consider the Sidon $B_{\leq d}$ sequence $a_1, a_2, \cdots, a_n$ in Lemma 59 where $\max_j a_j \leq N-1 = ((1+o(1))n)^{d+1}$. Consider the field $\mathbb{F}_{q^T}$ where $q^T \geq dN$ and let α be a primitive root of this field. Consider the assignment $\boldsymbol{x}' = (x^{a_0}, x^{a_1}, x^{a_2}, \ldots, x^{a_d})$ where $a_0 = 0$ and x is a new indeterminate. Then

$$h(x) := f(\boldsymbol{x}') = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} \cdot x^{a_{i_1} + a_{i_2} + \cdots + a_{i_d}}.$$

Notice that $n_{\boldsymbol{i}} := a_{i_1} + a_{i_2} + \cdots + a_{i_d}$, $\boldsymbol{i} \in I$ are distinct and $n_{\boldsymbol{i}} < dN$ and therefore if f is not identically zero then $h(x)$ is a nonzero polynomial of degree at most $dN-1$ with at most s monomials. Therefore, one of $h(\alpha^0), h(\alpha^1), \ldots, h(\alpha^{s-1})$ is not zero. That is, $S = \{\boldsymbol{x}'(\alpha^0), \boldsymbol{x}'(\alpha^1), \ldots, \boldsymbol{x}'(\alpha^{s-1})\}$ is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^T}$ of size s. By *3* in Lemma 62 there is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^t}$ of size $s \cdot \nu^{\mathcal{P}}_{\mathbb{F}_{q^t}}(d, \mathbb{F}_{q^T})$ and by Lemma 45, result *1* follows.

To prove *2* we consider the Sidon $B_{\leq d}$ sequence $a_1, a_2, \cdots, a_n$ in Lemma 60 where $\max_j a_j \leq N = (2dn)^{2d}$ that can be constructed in $poly(n)$ time. By Lemma 57, for

$$M = \frac{1}{100} q^{24} (3d \log(dn) + 25)^2 = O(q^{24} d^2 (\log n)^2)$$

an element β in $\mathbb{F}_{q^K}$ for some $M \leq K \leq M + o(M)$ of multiplicative order at least dN can be constructed in time $poly(n)$. As above $S = \{\boldsymbol{x}'(\beta^0), \boldsymbol{x}'(\beta^1), \ldots, \boldsymbol{x}'(\beta^{s-1})\}$ is a black box PIT set for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ over $\mathbb{F}_{q^K}$ of size s. Now the result follows from *3* in Lemma 62 and Lemma 45. □

5.10 Field Reduction of Other Circuit Classes

In this subsection we apply the reduction in Lemma 62 for the black box PIT sets in [70, 73, 74, 6] and get a black box PIT sets over smaller fields. The results are indicated in the following table

Circuit Class	Field size $\geq$	Size of PIT set	New Field size $\geq$	New Size of PIT set	Ref.
$\Sigma\Pi\Sigma(k,d,n)$	dnk^2	$poly(n) \cdot d^k$	$d+1$	$poly(n) \cdot d^{k+5}$	[70]
$\Sigma_r \mathrm{P}_k \mathrm{ROF}$	kn^3	$(kn)^{O(r+\log n)}$	$kn+1$	$(kn)^{O(r+\log n)}$	[73]
ML $\Sigma\Pi\Sigma\Pi(k)$	n^2	$n^{O(k^3)}$	$n+1$	$n^{O(k^3)}$	[74]
R_k ML	n^2	$n^{k^{O(k)}+O(k \log n)}$	$n+1$	$n^{k^{O(k)}+O(k \log n)}$	[6]

The classes in the table are: $\Sigma\Pi\Sigma(k,d,n)$ is the class of depth-3 circuits with n variables, degree d and top fanin k. P_kROF is the class of read once formulas (ROF, each variable appears at most once in the formula) with n variables in which we are allowed to replace each variable x_i with a univariate polynomial $T_i(x_i)$ of degree at most k. $\Sigma_r\mathrm{P}_k$ROF is the sum of r P_kROF formulas. Multilinear (ML) $\Sigma\Pi\Sigma\Pi(k)$ is the class of depth 4 circuits with n variables in which the fan-in of the top Σ gate is a constant k and each multiplication gate computes a multilinear polynomial. R_k ML is the class of multilinear formulas with n variables where each variable appears at most k times in the formula.

The table shows the reduction to a smaller field. For example, in the first row in the table, in [70], Saxena and Seshadhri gave a black box PIT set for $\Sigma\Pi\Sigma(k,d,n)$ over fields of size at least dnk^2. We apply our reduction to give a black box PIT set for $\Sigma\Pi\Sigma(k,d,n)$ over fields of size at least $d+1$. Notice that our field size is independent of n and when the field $\mathbb{F}_q$ satisfies $q \geq d+1$, no extension field is needed.

6 Application of Tester for Polynomial Restriction Problems

A restriction problem is a problem of the following form:

Given an alphabet Σ of size $|\Sigma| = q$, an integer n and a class $\mathcal{M}$ of nonzero functions $f : \Sigma^n \to \{0,1\}$.

Find a small set $S \subseteq \Sigma^n$ such that: For any $f \in \mathcal{M}$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$.

We will study restriction problems when $\mathcal{M}$ is a class of multivariate polynomials over $\mathbb{F}_q$.

We remind the reader that the *total degree* (or just *degree*) of a multivariate polynomial is the maximum over the sums of the exponents of each multivariate monomial and the *variable degree* is the maximum over the degree of each variable. We denote by $\mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$ the class of all multivariate polynomial over $\mathbb{F}_q$ of degree exactly d and variable degree at most $r \leq q-1$ that contains at most s monomials of degree d and any number of monomials of degree less than d. We denote $\mathcal{P}(\mathbb{F}_q, n, ((d,q-1),s))$ by $\mathcal{P}(\mathbb{F}_q, n, (d,s))$. Obviously, $\mathcal{P}(\mathbb{F}_q, n, d, s) \subseteq \mathcal{P}(\mathbb{F}_q, n, (d,s))$ and $\mathcal{P}(\mathbb{F}_q, n, (d,r), s) \subseteq \mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$. Notice that, for example, $f = (x_1+1)(x_2+1)\cdots(x_d+1)$ is in $\mathcal{P}(\mathbb{F}_q, n, (d,1), 2^d)$ and $\mathcal{P}(\mathbb{F}_q, n, ((d,1),1))$ but not in $\mathcal{P}(\mathbb{F}_q, n, (d,1), s)$ for any $s < 2^d$.

In this section we consider the following **s-sparse (d,r)-degree polynomial problem** over $\mathbb{F}_q$: Given the class $\mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$. Find a small set $S \subset \mathbb{F}_q^n$ such that for every $f \in \mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$ there is $\boldsymbol{a} \in S$ such that $f(\boldsymbol{a}) \neq 0$.

This problem can be regarded as black box PIT set over $\mathbb{F}_q$ problem, hitting set problem or polynomial restriction problem. We will call S a *hitting set* for $\mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$.

Note: Throughout this section we will assume

$$d, r, q = o(n). \tag{36}$$

Our technique can also handle other cases that will also be considered here. Although our results are

true for any s, we will assume that $s \gg n$ and therefore we will concentrate on constructions that gives hitting sets of size that is linear in s. In [12] other constructions of hitting sets that have size quadratic in s are also studied.

6.1 Lower Bound for $\mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$

In this subsection we give a lower bound for the size of any hitting set for $\mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$.

We prove

Lemma 81. *For any n, d, r, s and ℓ, where $\ell \leq n^{c_1}$ and $s \leq n^{c_2(1-c_1)\ell}$ for some constants $c_1, c_2 < 1$, any hitting set S for $\mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$ is of size*

$$|S| = \Omega\left(\left(\frac{q}{q-r}\right)^{\lfloor (d-\ell)/r \rfloor} \frac{\ell s \log n}{\log q}\right).$$

In particular, for $s = poly(n)$ there is infinite sequence of integers d such that

$$|S| = \Omega\left(\pi_{q,r}^{d} \cdot \frac{s \log n}{\log q}\right) \tag{37}$$

where

$$\pi_{q,r} = \left(\frac{q}{q-r}\right)^{1/r}$$

and for $s = poly(n)$ and $q = 2$

$$|S| = \Omega(2^d \cdot s \log n).$$

Proof. Let $S = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_t\} \subseteq \mathbb{F}_q^n$ be a hitting set for $\mathcal{P}(\mathbb{F}_q, n, ((d,r),s))$. Let $w = \lfloor (d-\ell)/r \rfloor$. Pick random uniform sets $A_1, A_2, \ldots, A_w \subset \mathbb{F}_q$ each of size r. For a fixed $\boldsymbol{b} \in S$ we have

$$\mathbf{Pr}\left[b_1 \notin A_1, \ldots, b_w \notin A_w\right] = \left(\frac{q-r}{q}\right)^w.$$

Therefore there exist w sets $A_1, \ldots, A_w \in \mathbb{F}_q$ each of size r such that at most $t((q-r)/q)^w$ elements $\boldsymbol{b} \in S$ satisfy $b_i \notin A_i$ for all $i = 1, \ldots, w$. Let $B' \subseteq S$ be those elements and $B = \{(b_{w+1}, \ldots, b_n) \mid \boldsymbol{b} \in B'\}$. Then

$$|B| \leq |B'| \leq \left(\frac{q-r}{q}\right)^w t. \tag{38}$$

Note that the polynomial

$$p(\boldsymbol{x}) = \prod_{i=1}^{w} \prod_{\alpha \in A_i} (x_i - \alpha)$$

is zero on all the elements of $S \backslash B'$.

Let $\mathcal{M}' \subset \mathcal{P}(\mathbb{F}_q, n-w-1, ((\ell, r), \lfloor s/2 \rfloor))$ be a set of all polynomials over $\mathbb{F}_q$ of degree ℓ and variable degree at most r over the $n-w-1$ variables $(x_{w+2}, \ldots, x_n)$ of size at most $\lfloor s/2 \rfloor$ that contain no monomials of degree less than ℓ. By the proof of Lemma 65, if $|\mathbb{F}_q|^{|B|} < |\mathcal{M}'|$ then there are two distinct functions $f, g \in \mathcal{M}'$ such that $f - g$ is equal to zero on all the elements in B. Then, the multivariate polynomial

$$p(\boldsymbol{x}) \cdot x_{w+1}^{d-\ell-wr} \cdot (f(x_{w+2}, \ldots, x_n) - g(x_{w+2}, \ldots, x_n)) \in \mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$$

is non-zero polynomial of degree d and variable degree r and has at most s monomials of degree d and is equal to 0 on all the elements of S. Therefore we must have $|\mathbb{F}_q|^{|B|} \geq |\mathcal{M}'|$.

Now, by (38), we get

$$t \geq \left(\frac{q}{q-r}\right)^w |B| \geq \left(\frac{q}{q-r}\right)^w \frac{\log |\mathcal{M}'|}{\log q} = \Omega\left(\left(\frac{q}{q-r}\right)^{\lfloor (d-\ell)/r \rfloor} \frac{\ell s \log n}{\log q}\right).$$

□

We note here that a better bound can be obtained in (37) if we choose $\ell = r/\log(q/(q-r))$ and $d \bmod r = \ell$. We avoid this since it will give a cumbersome analysis.

We note here that in (37) the constant $\pi_{q,r}$ satisfies

$$1 + \frac{1}{q-1} \leq \pi_{q,r} = \left(\frac{q}{q-r}\right)^{1/r} \leq q^{1/(q-1)} = 1 + \frac{\ln q}{q-1} + O\left(\frac{\log^2 q}{q^2}\right) \leq 2.$$

6.2 Nonconstructive Upper Bound for $\mathcal{P}(\mathbb{F}_q, n, (d, s))$

In this subsection we give a non-constructive upper bound for the size of hitting sets for $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$.

We first prove the following

Lemma 82. *For any non-zero function $f \in \mathcal{P}(\mathbb{F}_q, n, (d, r))$, $r \geq 1$ we have*

$$\mathbf{Pr}_{\boldsymbol{x} \in U(\mathbb{F}_q^n)}[f(\boldsymbol{x}) \neq 0] \geq \pi_{q,r}^{-d}$$

where

$$\pi_{q,r} = \left(\frac{q}{q-r}\right)^{1/r}$$

and $U(\mathbb{F}_q^n)$ is the uniform distribution over $\mathbb{F}_q^n$.

Proof. The result follows from Lemma 69.

□

The union bound with Lemma 82 gives

Lemma 83. *We have*

1. *There is a hitting set for $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ of size*
$$\frac{\log |\mathcal{P}(\mathbb{F}_q, n, (d, r), s)|}{\log\left(\frac{1}{1-\pi_{q,r}^{-d}}\right)}$$
for $q \leq d$ and of size
$$\frac{\log |\mathcal{P}(\mathbb{F}_q, n, (d, r), s)|}{\log(q/d)}$$
for $q \geq d+1$.

2. *In particular, for any $d \leq n^{c_1}$ and $s \leq n^{(1-c_1)c_2 d}$, for some constant $c_1, c_2 < 1$, there is a hitting set for $\mathcal{P}(\mathbb{F}_q, n, (d, r), s)$ of size*
$$O\left(d\pi_{q,r}^d s \log n\right)$$
for $q \leq d$ and of size
$$O\left(\frac{ds\log n}{\log(q/d)}\right) = \begin{cases} O(d^2 s \log n) & q = d + c \\ O(ds \log n) & q = c(d+1) \\ O\left(\frac{d}{\log(q/d)} s \log n\right) & q = \omega(d) \\ O\left(\frac{d}{\log q} s \log n\right) & q \geq d^{1+c} \end{cases}$$
for $q \geq d+1$ and any constant $c > 0$.

3. *In particular, there is a hitting set for $\mathcal{P}(\mathbb{F}_2, n, d, s)$ of size $O(d2^d s \log n)$.*

Since $f \in \mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ can contain any number of monomials of degree less than d, the size of the class $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ is at least $q^{(n/d)^{d-1}}$ and therefore the union bound will not give a good bound. In what follows we use a new technique that gives better bounds

Lemma 84. *We have*

1. *Let $d \leq n^{c_1}$ and $s \leq n^{(1-c_1)c_2 d}$ for some constants $c_1, c_2 < 1$. If $r \leq p-1$, where p is the characteristic of the field, and $q \leq d$ then there is a hitting set for $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ of size*
$$O\left(d\left(2 \cdot \pi_{q,1}\right)^d s \log n\right) = O\left(d\left(2 + \frac{2}{q-1}\right)^d s \log n\right).$$

2. *In particular, there is a hitting set for* $\mathcal{P}(\mathbb{F}_2, n, (d, s))$ *of size* $O(d2^{2d}s\log n)$.

3. *If* $q \geq d+1$ *then for any constant* $c \geq 0$*, there is a hitting set for* $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ *of size*

$$O\left(\frac{d^2 s\log n}{\log(q/d)}\right) = \begin{cases} O(d^3 s\log n) & q = d+c \\ O(d^2 s\log n) & q = c(d+1) \\ O\left(\frac{d^2}{\log(q/d)} s\log n\right) & q = \omega(d) \\ O\left(\frac{d^2}{\log q} s\log n\right) & q \geq d^{1+c}. \end{cases}$$

Proof. We first prove *1* and *2*. Consider $\phi_d\mathcal{M} = \{\phi_d f \mid f \in \mathcal{P}(\mathbb{F}_q, n, (d, s))\} \subseteq \mathcal{P}(\mathbb{F}_q, dn, (d, r = 1))$. First of all, by Lemma 61, for every $f \in \mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$, $r \leq p-1$, $f \not\equiv 0$ if and only if $\phi_d f \not\equiv 0$. Second of all, $\phi_d f$ depends only on the monomials of degree d in f and therefore

$$|\phi_d\mathcal{M}| \leq q^{s\binom{\binom{n+d-1}{d}}{s}}.$$

Now by Lemma 82 and the union bound there is a hitting set $Y \subseteq (\mathbb{F}_q^n)^d$ for $\phi_d\mathcal{M}$ of size

$$\frac{\log|\phi_d\mathcal{M}|}{\log\left(\frac{1}{1-\pi_{q,1}^{-d}}\right)} = O\left(d\pi_{q,1}^d s\log n\right).$$

Now if for some $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_d) \in Y$ and $f \in \mathcal{P}(\mathbb{F}_q, n, (d, s))$ we have $(\phi_d f)(\boldsymbol{y}) \neq 0$ then by (33) $f(\sum_{j\in J} \boldsymbol{y}_j) \neq 0$ for some $J \subseteq [d]$. Therefore,

$$Y' = \left\{ \sum_{j\in J} \boldsymbol{y}_j \;\middle|\; J \subseteq [d], \boldsymbol{y} \in Y \right\}$$

is a hitting set for $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$ of size

$$O\left(d2^d\pi_{q,1}^d s\log n\right).$$

Now we prove *3*. By Lemma 83 there is a hitting set $X \subseteq \mathbb{F}_q^n$ for $\mathcal{P}(\mathbb{F}_q, n, d, s)$ of size $O((ds\log n)/\log(q/d))$. Now choose $d+1$ distinct elements $B = \{\beta_1, \ldots, \beta_{d+1}\} \subseteq \mathbb{F}_q$. This is possible since $q \geq d+1$. Then define the set $S = \{\beta_j \boldsymbol{x}' \mid j = 1, \ldots, d+1, \boldsymbol{x}' \in X\}$. We now show that S is a hitting set for $\mathcal{P}(\mathbb{F}_q, n, (d, s))$.

Let $f(\boldsymbol{x}) \in \mathcal{P}(\mathbb{F}_q, n, (d, s))$ and define $g(\boldsymbol{x}, y) = f(yx_1, yx_2, \ldots, yx_n)$ where y is a new indeterminate. Then $g(\boldsymbol{x}, y) = g_d(\boldsymbol{x})y^d + g_{d-1}(\boldsymbol{x})y^{d-1} + \cdots + g_0(\boldsymbol{x})$ and $g_d(\boldsymbol{x}) \in \mathcal{P}(\mathbb{F}_q, n, d, s)$. Therefore, $g(\boldsymbol{x}', y) \not\equiv 0$ for some $\boldsymbol{x}' \in X$. Since $g(\boldsymbol{x}', y)$ is a polynomial of degree at most d in y we have $g(\boldsymbol{x}', \beta_j) \neq 0$ for some

$j = 1, \ldots, d+1$. Therefore $f(\beta_j \boldsymbol{x}') = g(\boldsymbol{x}', \beta_j) \neq 0$ for some $\boldsymbol{x}' \in X$ and $\beta_j \in B$. Thus, S is a hitting set for $\mathcal{P}(\mathbb{F}_q, n, (d, s))$. Now

$$|S| = |X||B| = O\left(\frac{d^2 s \log n}{\log(q/d)}\right).$$

□

6.3 Constructive Upper Bound for $\mathcal{P}(\mathbb{F}_q, n, (d, s))$

In this subsection we give explicit and polynomial time constructions of hitting sets for $\mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$. Our first result is for $q \geq d + 1$. We prove

Lemma 85. *Let $q \geq d + 1$ and $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, (d, s))$.*

1. *There is an explicit hitting set for $\mathcal{M}$ of size*

$$O\left(d^{\tau+1} s \frac{\log n}{\log q}\right) = \begin{cases} O\left(d^2 s \frac{\log n}{\log q}\right) & \text{if } q \text{ perfect square}, q \geq c(d+1)^2, \text{ I.S. } n \\ O\left(d^3 s \frac{\log n}{\log q}\right) & \text{if } q \text{ perfect square}, q \geq c(d+1)^2 \\ O\left(d^3 s \frac{\log n}{\log q}\right) & \text{if } q \geq c(d+1), \text{ I.S. } n \\ O\left(d^4 s \frac{\log n}{\log q}\right) & \text{if } q \geq c(d+1) \\ O\left(d^4 s \frac{\log n}{\log q}\right) & \text{if } q \geq d+1, \text{ I.S. } n \\ O\left(d^5 s \frac{\log n}{\log q}\right) & \text{if } q \geq d+1 \end{cases}$$

   where $c > 1$ is any constant.

2. *For a constant d there is a hitting set for $\mathcal{M}$ of size*

$$O\left(s \frac{\log n}{\log q}\right)$$

   that can be constructed in polynomial time $poly(n)$.

3. *For $d = O(\log \log \log n)$ and a prime q there is a hitting set for $\mathcal{M}$ of size*

$$O\left(d^{\tau+1} s \frac{\log n}{\log q}\right)$$

   that can be constructed in time $s \cdot poly(n)$.

4. For any d there is a hitting set for $\mathcal{M}$ of size

$$O\left(q^{24}d^7 s \cdot \log^2 n\right)$$

that can be constructed in time $s \cdot poly(n)$.

Proof. Let $f \in \mathcal{P}(\mathbb{F}_q, n, (d, s))$. Consider the Sidon $B_{\geq d}$ sequence in Lemma 60, $a_1, a_2, \cdots, a_n$ where $\max a_j \leq N := (2dn)^{2d}$ that can be constructed in polynomial time. Consider the assignment $\boldsymbol{x}' = (x^{a_1}y, \ldots, x^{a_n}y)$ where x and y are new indeterminates. Then $g(x, y) := f(\boldsymbol{x}') \in \mathbb{F}_q[x, y]$ can be represented as $g(x, y) = g_d(x)y^d + g_{d-1}(x)y^{d-1} + \cdots + g_0(x)$. Since $f \in \mathcal{P}(\mathbb{F}_q, n, (d, s))$, $g_d(x)$ is a nonzero polynomial of degree at most dN with at most s monomials. Consider the field $\mathbb{F}_{q^t}$ such that $q^{t-1} < dN \leq q^t$ and α a primitive root of $\mathbb{F}_{q^t}$. Since g_d contains at most s monomials of degree less than dN, one of the values $g_d(\alpha^0), \ldots, g_d(\alpha^{s-1})$ is not equal to 0. Since $g(x, y)$ is of degree d in y, we have $g(\alpha^i, \alpha^j) \neq 0$ for some $i = 0, \ldots, s-1$ and $j = 0, \ldots, d$. This gives a hitting set for $\mathcal{P}(\mathbb{F}_q, n, (d, s))$ over $\mathbb{F}_{q^t}$ of size $(d+1)s$. Now using a $(\mathcal{P}(\mathbb{F}_q, n, d), \mathbb{F}_{q^t}, \mathbb{F}_q)$-tester and by Corollary 17 we get a hitting set for $\mathcal{P}(\mathbb{F}_q, n, (d, s))$ of size

$$(d+1)s \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^t}) = O\left(d^{\tau(d,q,t)+1} s \frac{\log n}{\log q}\right).$$

This proves *1.*

The bottleneck in the above algorithm is finding a primitive root in $\mathbb{F}_{q^t}$ where $q^t \geq dN$. By Lemma 56, this can be done in polynomial time when d is constant. This implies *2.*

For $d = O(\log \log \log n)$ we use Lemma 58. If q is prime then by Lemma 58 a primitive root in $\mathbb{F}_{q^T}$ where $T = t + o(t)$ can be found in time

$$q^{O(t/\log\log t)} = q^{O(d \log n/(\log q \log\log\log n))} = poly(n).$$

This with Corollaries 41 and 43 implies *3.*

One can also use, instead of a primitive root, an element in $\mathbb{F}_{q^t}$ of multiplicative order at least dN. By Lemma 57 for $m = O(q^{24}d^2 \log^2 n)$ there is a polynomial time algorithm that constructs an element in $\mathbb{F}_{q^M}$ of order at least dN where $M \leq m + o(m)$. This, with Corollary 41, gives a hitting set of size

$$(d+1)s \cdot \nu^{\mathcal{P}}_{\mathbb{F}_q}(d, \mathbb{F}_{q^m}) = O\left(q^{24}d^7 s \cdot \log^2 n\right).$$

This implies *4.* □

Notice that, for any constant $c > 1$ and perfect square $q \geq c(d+1)^2$, our explicit construction meets the non-constructive bound in Lemma 84.

We now study the problem for $q \leq d$. We will first study the case where $q = 2$ and then consider other fields. We prove

Lemma 86. *Let $\mathcal{M} = \mathcal{P}(\mathbb{F}_2, n, (d, s))$. Then*

1. *There is an explicit hitting set for $\mathcal{M}$ of size $O(2^{2.66d} s \log n)$.*
2. *For a constant d there is a hitting set for $\mathcal{M}$ of size $O(s \log n)$ that can be constructed in polynomial time $poly(n)$.*
3. *For $d = O(\log\log\log n)$, there is a hitting set for $\mathcal{M}$ of size $O(2^{2.66d} s \log n)$ that can be constructed in time $s \cdot poly(n)$.*
4. *There is a hitting set for $\mathcal{M}$ of size $O(2^{2.66d} s \log^2 n)$ that can be constructed in time $2^{2.66d} s \cdot poly(n)$.*

Proof. The idea of the proof is the following: First we change the multivariate polynomials in $\mathcal{M}$ to (n, d)-multilinear polynomials in $\phi_d\mathcal{M}$. Then we use Sidon sequence to find a hitting set for $\phi_d\mathcal{M}$ over $\mathbb{F}_{2^r}$ where $r = O(d \log n)$. Then we use a tester to change the hitting set over $\mathbb{F}_{2^r}$ to a hitting set over $\mathbb{F}_2$. Then we use the hitting set for $\phi_d\mathcal{M}$ to get a hitting set for $\mathcal{M}$.

Let $\mathcal{M} = \mathcal{P}(\mathbb{F}_2, n, (d, s))$. Consider $f \in \mathcal{M}$ and suppose

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} x_{i_1} \cdots x_{i_d} + g(\boldsymbol{x})$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)$, $g(\boldsymbol{x})$ is a multivariate polynomial of degree less than d, $\boldsymbol{i} = (i_1, i_2, \ldots, i_d)$, $1 \le i_1 < i_2 < \cdots < i_d \le n$, $\lambda_{\boldsymbol{i}} = 1$, $I \subset [n]^d$ and $|I| \le s$. Let $M_{\boldsymbol{i}} = x_{i_1} \cdots x_{i_d}$. Let $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)$ where $\boldsymbol{y}_i = (y_{i,1}, \ldots, y_{i,n})$ are new indeterminates for $i = 1, \ldots, d$. Recall the operator $\phi_d : \mathbb{F}_2[\boldsymbol{x}] \to \mathbb{F}_2[\boldsymbol{y}]$ in (33). Since the field is of characteristic 2 and since ϕ_d is linear we have

$$(\phi_d f)(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} \mathrm{Perm}\left(Y_{M_{\boldsymbol{i}}}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)\right) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} \det\left(Y_{M_{\boldsymbol{i}}}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)\right). \tag{39}$$

Notice that $\phi_d f$ is (n, d)-multilinear polynomial. Let $\boldsymbol{z} = (z_1, \ldots, z_n)$ be new indeterminates. Then

$$(\phi_d f)(\boldsymbol{z} \star \boldsymbol{y}_1, \ldots, \boldsymbol{z} \star \boldsymbol{y}_d) = \sum_{\boldsymbol{i} \in I} \lambda_{\boldsymbol{i}} z_{i_1} \cdots z_{i_d} \det(Y_{M_{\boldsymbol{i}}}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)),$$

where $\boldsymbol{z} \star \boldsymbol{y}_i = (z_1 y_{i,1}, \ldots, z_n y_{i,n})$.

Let $\mathbb{K} = \mathbb{F}_{2^r}$ where $r = \lceil 4d \log(2dn) \rceil$ and let α be a primitive root in $\mathbb{K}$. We substitute $\boldsymbol{y}' = (\boldsymbol{y}'_1, \ldots, \boldsymbol{y}'_d)$ in $\boldsymbol{y}$ where $y'_{i,j} = \alpha^{ij}$ and get

$$(\phi_d f)(\boldsymbol{z} \star \boldsymbol{y}'_1, \ldots, \boldsymbol{z} \star \boldsymbol{y}'_d) = \sum_{i \in I} \lambda_{\boldsymbol{i}} z_{i_1} \cdots z_{i_d} \det\left([\alpha^{i_k j}]_{k,j}\right) = \sum_{\boldsymbol{i} \in I} \Lambda_{\boldsymbol{i}} z_{i_1} \cdots z_{i_d},$$

where, by Vandermonde, $\Lambda_{\boldsymbol{i}} = \lambda_{\boldsymbol{i}} \det([\alpha^{i_k j}]_{k,j}) \neq 0$ for all $\boldsymbol{i} \in I$. Let $a_1, a_2, \ldots, a_n$ be the Sidon $B_{\leq d}$ sequence that is defined in Lemma 60 where $\max_j a_j \leq (2dn)^{2d}$. We now substitute $\boldsymbol{z} = \boldsymbol{z}' := (w^{a_1}, \ldots, w^{a_n})$ for a new indeterminate w and get

$$t(w) := (\phi_d f)(\boldsymbol{z}' \star \boldsymbol{y}'_1, \ldots, \boldsymbol{z}' \star \boldsymbol{y}'_d) = \sum_{\boldsymbol{i} \in I} \Lambda_{\boldsymbol{i}} w^{a_{i_1} + \cdots + a_{i_d}}.$$

Since $t(w)$ is a polynomial in one variable w of size at most s and degree at most $d(2dn)^{2d} < |\mathbb{K}|$, one of the values $t(1), t(\alpha), \ldots, t(\alpha^{s-1})$ is not zero.

This gives a hitting set for $\phi_d \mathcal{M}$ over $\mathbb{K}$ of size s. By the definition of testers there is a hitting set for $\phi_d \mathcal{M}$ over $\mathbb{F}_2$ of size $s \cdot \nu_{\mathbb{F}_2}(d, \mathbb{F}_{2^r})$. Since each substitution in $\phi_d f$ can be simulated by 2^d substitutions in f and by Theorem 21, there is a hitting set for $\mathcal{M}$ over $\mathbb{F}_2$ of size

$$2^d s \cdot \nu_{\mathbb{F}_2}(d, \mathbb{F}_{2^r}) = 2^{d + c_2 d} d^5 sr \leq d^6 2^{d + c_2 d} s \log n = O(2^{2.66d} s \log n).$$

This proves *1*.

In Lemma 59 we show that a Sidon $B_{\leq d}$ sequence $a_1, a_2, \cdots, a_d$ with $N := \max_j a_j \leq (1 + o(1))^d n^d$ can be constructed in deterministic $O(n^{d/2+1})$ time. We can use this in the above proof. Then we can choose $\mathbb{K} = \mathbb{F}_{2^r}$ where $2^r \geq dN > 2^{r-1}$ and, by Lemma 56, a primitive root in $\mathbb{K}$ can be found in time $(dN)^{1/4} = O(n^{d/2+1})$. This gives a construction of size $O(2^{2.66d} s \log n)$ that can be constructed in time $n^{d/2+1}$. For constant d, we get *2*.

By Lemma 58, for some constant c, there is $M = 2d \log(2dn) + o(2d \log(2dn))$ such that a primitive root in $\mathbb{F}_{2^M}$ can be found in time

$$T = 2^{c \frac{2d \log(2dn)}{\log \log(2d \log(2dn))}}.$$

When $d = O(\log \log \log n)$ we have $T = poly(n)$. This implies *3*.

For the above construction one can also use an element of large multiplicative order in finite field instead of a primitive root. By Lemma 57, for $m = O(d^2 \log^2(dn)))$ one can find an element in $\mathbb{F}_{2^M}$ for some $m \leq M \leq m + o(m)$ of multiplicative order at least $(2dn)^{2d}$ in polynomial time. This gives a polynomial time construction of a hitting set of size $O(2^{2.66d} s \log^2 n)$. □

When applying the above to any other field we encounter two bottlenecks. The first is that in (39), $\mathrm{Perm}\left(Y_{M_{\boldsymbol{i}}}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)\right) = \det\left(Y_{M_{\boldsymbol{i}}}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)\right)$ is true only for multilinear polynomials over fields of characteristic 2. Using permanent instead of determinant for fields of characteristic not equal to 2 will add a factor of $d!$ to the size of the construction. The second bottleneck is that the operator in (33) gives a factor of 2^d to the size of the construction and this, for large field $\mathbb{F}_q$, gives a large gap from the lower bound $(q/(q-1))^d s \log n = 2^{\Omega(d/q)} s \log n$. We therefore get

Lemma 87. *Let $q \leq d$, $r \leq p - 1$ and $\mathcal{M} = \mathcal{P}(\mathbb{F}_q, n, ((d, r), s))$. Then*

1. *There is an explicit hitting set for $\mathcal{M}$ of size $O(2^{(1+c_q)d}d!s\log n)$.*
2. *If $q=2^\ell$ then there is an explicit hitting set for $\mathcal{M}$ of size $O(2^{(1+c_q)d}s\log n)$.*
3. *If d is constant then there is a hitting set for $\mathcal{M}$ of size $O(s\log n)$ that can be constructed in polynomial time $poly(n)$.*
4. *There is a hitting set for $\mathcal{M}$ of size $O(2^{(1+c_q)d}sd!\log^2 n)$ that can be constructed in time $2^{(1+c_q)d}d!s\cdot poly(n)$.*
5. *If $q=2^\ell$ then there is a hitting set for $\mathcal{M}$ of size $O(2^{(1+c_q)d}s\log^2 n)$ that can be constructed in time $2^{(1+c_q)d}s\cdot poly(n)$.*

The table in Figure 1 summarizes the results for $\mathcal{M}=\mathcal{P}(\mathbb{F}_q,n,((d,r),s))$ where $r\le p-1$. See also Open Problems 7.2 and 7.3.

q,d	Lower	Upper	Explicit	Poly Time
$q=2$	$2^d s\log n$	$2^{2d}s\log n$	$2^{2.66d}s\log n$	$2^{2.66d}s\log^2 n$
$q=2^\ell\le d$	$\pi_{q,r}^d s\frac{\log n}{\log q}$	$d2^d\pi_{q,1}^d s\log n$	$2^{(1+c_q)d}s\log n$	$2^{(1+c_q)d}s\log^2 n$
$q\le d$	$\pi_{q,r}^d s\frac{\log n}{\log q}$	$d2^d\pi_{q,1}^d s\log n$	$2^{(1+c_q)d}d!s\log n$	$2^{(1+c_q)d}d!s\log^2 n$
$q\le d=O(1)$	$s\log n$	$s\log n$	$s\log n$	$s\log n$
$q\ge d+1$	$ds\frac{\log n}{\log q}$	$d^2 s\frac{\log n}{\log(q/d)}$	$d^{\tau+1}s\frac{\log n}{\log q}$	$q^{24}d^7 s\cdot\log^2 n$
$q\ge d+1=O(1)$	$s\frac{\log n}{\log q}$	$s\frac{\log n}{\log(q/d)}$	$s\frac{\log n}{\log q}$	$s\frac{\log n}{\log q}$

Figure 1: Bounds for the size of black box PIT set for $\mathcal{P}(\mathbb{F}_q,n,((d,r),s))$.

$$1+\frac{1}{q-1}=\pi_{q,1}\le\pi_{q,r}:=\left(\frac{q}{q-r}\right)^{1/r}\le\pi_{q,q-1}=q^{1/(q-1)}=1+\frac{\ln q}{q-1}+O\left(\frac{\log^2 q}{q^2}\right)\le 2.$$

Open Problems 7.

1. *Find a polynomial time algorithm that for every integer n, constructs a $B_{\le d}$ Sidon sequence $a_1,a_2,\cdots,a_n$ with $\max_j a_j\le n^d$.*
2. *Close the gaps between the lower bounds and upper bounds in the table in Figure 1. Recently, we have developed in [13] a new technique that uses what we will call "general permanent" and "semi-symmetric testers" that narrow those gaps.*
3. *In Lemma 85, for large q, one can use Blaşer et. al. technique, [18], to get a hitting set of size $poly(s,d,\log n)$ in time $poly(s,d,n)$. The size in those constructions are not linear in the size s. This will be studied in [12].*

7 Conclusion and Future Work

In this paper we have developed a new notion called tester and new techniques for using it for different applications. It is a useful technique because it gave many almost optimal constructions that could not be achieved using the previously known techniques. We believe that testers will further help in derandomizing many algorithms especially those that uses algebraic approach. Recently, this research has evolved in different directions. We will discuss some of them that will be studied in more details in [12, 13, 14].

Pseudorandom Generator. In this paper we studied testers that reduce hitting sets over an $\mathbb{F}$-algebra $\mathcal{A}$ to hitting sets over $\mathbb{F}$. In [12] we define $(1-\epsilon)$-tester where $f(\boldsymbol{a}) \neq 0$ implies $\mathbf{Pr}_{\boldsymbol{\ell} \in L}[f(\boldsymbol{\ell}(\boldsymbol{a})) \neq 0] \geq 1-\epsilon$. Then the size of the minimal $(1-\epsilon)$-tester $\nu_{\mathbb{F}_q}(\mathcal{M}, \mathbb{F}_{q^t}, 1-\epsilon)$ is studied. Using $(1-\epsilon)$-testers we give new high density hitting sets and Pseudorandom Generators [9] over small fields.

Randomized Algorithms with Small Number of Random Bits. Hitting sets of high density can reduce the number of random bits in randomized algorithms. We give one example.

To minimize the number of random bits in the randomized black box PIT algorithms for $\mathcal{M}$, one can consider an extension field $\mathbb{K} \supset \mathbb{F}$ such that $\mathbf{Pr}_{\boldsymbol{x} \in \mathbb{K}^n}[f(\boldsymbol{x}) \neq 0] \geq 1-\epsilon$ for all $f \in \mathcal{M}$ and then use $(\mathcal{M}, \mathbb{K}, \mathbb{F})$-tester L to change the point $\boldsymbol{x}$ to a set of points $S_{\boldsymbol{x}} = \{\ell(\boldsymbol{x}) \mid \ell \in L\}$ in $\mathbb{F}^n$. Then for a random uniform $\boldsymbol{x} \in \mathbb{K}^n$ the set $S_{\boldsymbol{x}} \subseteq \mathbb{F}^n$ is a hitting set for f with probability at least $1-\epsilon$. If each element in $\mathbb{K}$ can be represented with k bits then this algorithm uses kn random bits and gives a hitting set with probability at least $1-\epsilon$.

An interesting example is the following: Let $\mathcal{M} = \mathcal{P}(\mathbb{F}_2, n, d = c \log n)$ for constant c. Any deterministic black box PIT set for $\mathcal{M}$ is of size at least $n^{\Omega(\log n)}$. The (folklore) randomize black box PIT algorithm for $\mathcal{M}$ uses $n^{c+1} \log(1/\epsilon)$ random bits, runs in time $n^{c+1} \log(1/\epsilon)$ and with probability at least $1-\epsilon$ gives a hitting point. Using a $(\mathcal{M}, \mathbb{F}_{2^k}, \mathbb{F}_2)$-tester where $k = \log(c(\log n)/\epsilon)$, we get a randomized black box PIT algorithm for $\mathcal{M}$ that uses $O(n(\log n)\log((\log n)/\epsilon))$ random bits, runs in time $n^{2.66c+1}\log((\log n)/\epsilon)$ and with probability at least $1-\epsilon$ gives a hitting point. In [12] we show how to use dense hitting sets in order to reduce the number of random bits to $poly(\log n)$.

Non-Adaptive Learning and Interpolation of Multivariate Polynomial: Testers also give the first adaptive deterministic algorithm for learning the class $C = \mathcal{P}(\mathbb{F}_2, n, d, s)$ in time $poly(2^d, n, s)$ from membership queries (returns the value of the function in an assignment $\boldsymbol{a}$). In particular the algorithm runs in time $poly(n, s)$ for $d = O(\log n)$. Previous algorithms for this class was either randomized [20] or uses equivalence queries [71] (returns a counterexample to any hypothesis h suggested by the learner). To the best of our knowledge this is the first deterministic polynomial time learning algorithm for this class.

Notice that, by *4* in Lemma 86, there is a hitting set S for $\mathcal{P}(\mathbb{F}_2, n, d, s)$ of size $2^{2.66d} s \log^2 n$ that can be constructed in time $poly(2^d, s)$. Now using the polynomial time algorithms in [20, 71] the result follows. In what follows we describe a simple learning algorithm.

The algorithm goes like this. By the definition of hitting set we have $f \equiv 0$ if and only if $f(\boldsymbol{a}) = 0$ for all $\boldsymbol{a} \in S$. Therefore a hitting set is also a test set that tests if the function $f \in C$ is identically zero. We set variables in f to zero as long as the function is not identically zero. When no more variables can be set to zero then the product of those variables not set to zero is a monomial M in f. If all variables are set to zero then the monomial is 1. Assuming, at iteration t, we have found the monomials $M_1, M_2, \ldots, M_{t-1}$ in f. To find a new monomial we recursively run the above on $f + M_1 + M_2 + \cdots + M_{t-1}$. This gives an adaptive deterministic polynomial time algorithm for $\mathcal{P}(\mathbb{F}_2, n, d, s)$.

We can extend the above learning algorithm to an algorithm that learns the monomials of degree d in $\mathcal{M} = \mathcal{P}(\mathbb{F}_2, n, (d, s))$ in time $poly(2^d, n, s)$. The algorithm goes like this. We simulate membership queries to $\phi_d f$ using membership queries to f using (33). Each membership query to $\phi_d f$ can be simulated using 2^d membership queries to f. By the proof of Lemma 86, there is a hitting set for $\phi_d \mathcal{M}$ of size $2^{c_2 d} s \log^2 n$ that can be constructed in time $s \cdot poly(n, 2^d)$. We use this hitting set to test if $\phi_d f \equiv 0$. Notice that $\phi_d f \equiv 0$ if and only if f is of degree less than d. We now run the above algorithm (for the class $\mathcal{P}(\mathbb{F}_2, n, d, s)$) with this test. It is easy to see that at each iteration a monomial of degree d is found.

The above algorithms can also be extended to any field with the complexities described in Table 1.

This solves the open problem of *deterministic* adaptive learning boolean $O(\log n)$-multivariate polynomial (i.e., $C = \mathcal{P}(\mathbb{F}_2, n, O(\log n), s)$) from membership queries in polynomial time. This result was only true for decision trees of depth $O(\log n)$ [20]. An interesting open problem is to find a deterministic polynomial time non-adaptive algorithm for this problem and other interpolation problems. In [14] we define "Builder" which builds $f(\boldsymbol{a})$ from $f(\boldsymbol{\ell}(\boldsymbol{a}))$, $\boldsymbol{\ell} \in L$. We also study the connection of builders to the tensor rank of multidimensional matrices and give a non-adaptive learning algorithm for the above problem.

Hitting Set of Small degree Polynomial over any Field: In Lemma 87 we encountered two bottlenecks. The first is that in (39), $\text{Perm}\left(Y_{M_i}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)\right) = \det\left(Y_{M_i}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)\right)$ is true only for multilinear polynomials over fields of characteristic 2. This adds a factor of $d!$ to the size of the construction. The second bottleneck is that the operator in (33) gives a factor of 2^d to the size of the construction and this, for large field $\mathbb{F}_q$, gives a large gap from the lower bound $(q/(q-1))^d s \log n = 2^{\Omega(d/q)} s \log n$.

In [13] we develop a generalization of Ryser's formula and define a new notion called "general permanent" and use what we will call "semi-symmetric testers" to get tighter bounds.

Locally Explicit and Randomly Uniformly Explicit: All the constructions we have in this paper are deterministic polynomial time constructions. It can also be shown that all the constructions in this paper are polynomial time locally explicit, (i.e., every bit in the construction can be obtained in poly-logarithmic time in its size) and polynomial time randomly uniformly explicit (i.e., a random uniform row in the construction can be constructed in polynomial time in the size of the row with $\lceil \log T \rceil$ random bits where T is the size of the construction). This is studied in more details in [12].

Acknowledgement. I am grateful to Ronny Roth, Joachim von zur Gathen, Igor Shparlinski, Kenneth Shum and Ilya Volkovich for the number of interesting conversations.

References

[1] N. Alon. Combinatorial Nullstellensatz. *Combinatorics, Probability and Computing*, 8(1-2), pp. 7–29. (1999).

[2] N. Alon, J. Bruck, J. Naor, M. Naor, R. M. Roth. Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on Information Theory*, 38(2), pp. 509–516. (1992).

[3] M. Agrawal, N. Kayal, N. Saxena. PRIMES is in P, *Annals of Mathematics*, 160(2), pp. 781–793. (2004).

[4] L. M. Adleman, H. W. Lenstra. Finding irreducible polynomial over finite field. In Proc. 18th Annual ACM Symposium on Theory of Computing, (STOC 86), pp. 350–355. (1986).

[5] N. Alon, D. Moshkovitz, S. Safra. Algorithmic construction of sets for k-restrictions. *ACM Transactions on Algorithms*, 2(2), pp. 153–177. (2006).

[6] M. Anderson, D. van Melkebeek, I. Volkovich. Derandomizing polynomial identity testing for multilinear constant-read formulae. IEEE Conference on Computational Complexity 2011, pp. 273–282. (2011).

[7] N. Alon, J. H. Spencer. The probabilistic method, Wiley, Third Edition, (2008).

[8] S. Ballet. Curves with many points and multiplication complexity in any extension of $\mathbb{F}_q$. *Finite Fields and Their Applications*, 5(4) , pp. 364–377. (1999).

[9] A. Bogdanov. Pseudorandom generators for low degree polynomials. Proceedings of the 37th ACM Symposium on Theory of Computing, (STOC 2005), pp. 21–30. (2005).

[10] R. C. Bose. An affine analogue of Singer's theorem. *Journal of the Indian Mathematical Society,* 6, pp. 1–15. (1942).

[11] K. O'Bryant. A complete annotated bibliography of work related to Sidon sequences. *The Electric Journal of Combinatorics*, (2004).

[12] N. H. Bshouty. $(1-\epsilon)$-Testers and their applications. In Preperation.

[13] N. H. Bshouty. Testers via general permanent and combinatorial nullstellensatz. In Preperation.

[14] N. H. Bshouty. Builders and their Applications. In Preperation.

[15] R. C. Bose, S. Chowla. Theorems in additive theory of numbers. *Comment. Math. Helv.*, 37, pp. 141–147. (1962/1963).

[16] P. Buergisser, M. Clausen, A. Shokrollahi. Algebraic complexity theory. Grundlehren der mathematischen Wissenschaften. Springer Verlag, Heidelberg. (1996).

[17] S. R. Blackburn, T. Etzion, D. R. Stinson, G. M. Zaverucha. A bound on the size of separating hash families. *Journal of Combinatorial Theory, Series A*, 115(7), pp. 1246–1256. (2008).

[18] M. Blaŝer, M. Hardt, R. J. Lipton, N. K. Vishnoi. Deterministically testing sparse polynomial identities of unbounded degree. *Inf. Process. Lett.*, 109(3), pp. 187–192. (2009).

[19] R. C. Baker, G. Harman, J. Pintz. The difference between consecutive primes. II. Proceedings of the London Mathematical Society, 83(3), pp. 532-562. (2001).

[20] N. H. Bshouty, Y. Mansour. Simple learning algorithms for decision trees and multivariate polynomials. Proceedings of the 36th Annual Symposium on Foundations of Computer Science, (FOCS 95), pp. 304–311. (1995).

[21] D. Le Brigand, J. J. Risler. Algorithme de Brill-Noether et codes de Goppa, *Bull. Soc. math. France,* 116, pp. 231–253. (1988).

[22] M. Bazrafshan, T. van Trung. Bounds for separating hash families. *Journal of Combinatorial Theory, Series A*, 118(3), pp. 1129–1135. (2011).

[23] S. R. Blackburn, P. R. Wild. Optimal linear perfect hash families. *Journal of Combinatorial Theory, Series A*, 83(2), pp. 233–250. (1998).

[24] M. Clausen, A. Dress, J. Grabmeier, M. Karpinski. On Zero-testing and interpolation of k-sparse multivariate polynomials over finite fields. *Theoretical Computer Science*, 84(2), pp. 151–164. (1991).

[25] M. Deuring. Lectures on the theory of algebraic functions of one variable. Lecture Notes in Mathematics. Springer-Verlag, 314. (1973).

[26] D. Z. Du, F. K. Hwang. Combinatorial group testing and its applications. Volume 12 of Series on Applied Mathematics. World Scientific, New York, second edition, (2000).

[27] A. G. Dýachkov and V. V. Rykov. Bounds on the length of disjunctive codes. *Problemy Peredachi Inf*, 18(3), pp. 7–13. (1982).

[28] A. G. Dýachkov, V. V. Rykov, A. M. Rashad. Superimposed distance codes. Problems Control Inform. Theory/Problemy Upravlen. Teor. Inform., 18(4), pp. 237–250. (1989).

[29] S. A. Evdokimov. Factoring a solvable polynomial over finite fields and generalized Riemann hypothesis. *Zapiski Nauchn Semin. Leningr. Otdel Matem. Inst. Acad. Sci.*, USSR 176, pp. 104–117, (1989).

[30] W. Fulton. Algebraic curves: An introduction to algebraic geometry. W. A. Benjamin, Inc., New-York, Amesterdam. (1969).

[31] Z. Füredi. On r-cover-free families. *Journal of Combinatorial Theory, Series A*, 73(1), pp. 172–173. (1996).

[32] M. L. Fredman, J. Komlós. On the size of seperating systems and families of perfect hash function, *SIAM J. Algebraic and Discrete Methods*, 5(1), pp. 61–68. (1984).

[33] J. von zur Gathen, J. Gerhard. Modern computer algebra. 2nd edition. University of Bonn, Germany. Cambridge University Press. (2003).

[34] D. Y. Grigoriev, M. Karpinski, M. F. Singer. Fast parralel algorithms for sparse multivariate polynomial interpolation over finite fields. *SIAM J. Comput.*, 19(6), pp. 1059–1063. (1990).

[35] J. von zur Gathen and I. Shparlinski. Orders of Gauss periods in finite fields. *Applicable Algebra in Engineering, Communication and Computing*, 9(1), pp. 15–24. (1998).

[36] A. Garcia, H. Stichtenoth. On the asymptotic behaviour of some towers of function fields over finite fields. *Journal of Number Theory*, 61(2), pp. 248–273. (1996).

[37] A. Garcia, H. Stichtenoth. Topics in geometry, coding theory and cryptography. Algebra and applications. Springer. (2007).

[38] F. Hess. Computing Riemann-Roch spaces in algebraic function fields and related topics. *J. of Symbolic Computation*, 33(4), pp. 425–445. (2002).

[39] M. A. Huang, A. J. Rao. Interpolation of sparse multivariate polynomials over large finite fields with applications. *J. of Algorithms*, 33(2), pp. 204–228. (1999).

[40] P. Indyk, H. Q. Ngo, A. Rudra. Efficiently decodable non-adaptive group testing. In the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 10), pp. 1126–1142. (2010).

[41] J. Körner. Fredman-Komlós bounds and information theory, *SIAM J. Algebraic and Discrete Methods*, 7(4), pp. 560–570. (1986).

[42] J. Körner, K. Marton. New bounds for perfect hashing via information theory. *Europ. J. of Combinatorics*, 9(6), pp. 523–530. (1988).

[43] W. H. Kautz, R. C. Singleton, Nonrandom binary superimposed codes, *IEEE Trans. Inform. Theory*, 10(4), pp. 363-377. (1964).

[44] D. J. Kleitman, J. Spencer. Families of k-independent sets. *Discrete Mathematics*, 6(3), pp. 255–262. (1972).

[45] A. Klivans, D. A. Spielman. Randomness efficient identity testing of multivariate polynomials. In Proceedings on 33rd Annual ACM Symposium on Theory of Computing, (STOC 2001), pp. 216–223. (2001).

[46] K. S. Laursen. The computational complexity of effective construction of geometric Goppa codes. Proceedings of IEEE International Symposium on Information Theory, p. 380. (1997).

[47] R. Lidl, H. Niederreiter. Finite Fields. Encyclopedia of mathematics and its applications. Addison-Wesley Publishing Company. (1984).

[48] L. Liu, H. Shen. Explicit constructions of separating hash families from algebraic curves over finite fields. *Designs, Codes and Cryptography*, 41(2), pp. 221-233. (2006).

[49] D. Moshkovitz. An alternative proof of the Schwartz-Zippel lemma. Electronic Colloquium on Computational Complexity, Report No. 96. (2010).

[50] A. Nilli. Perfect hashing and probability. *Combinatorics, Probability and Computing*, 3(3), pp. 407–409. (1994).

[51] H. Q. Ngo, D. Z. Du. A survey on combinatorial group testing algorithms with applications to DNA library screening. *Theoretical Computer Science*, 55, pp. 171-182. (2000).

[52] J. Naor, M. Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM J. Comput.*, 22(4), pp. 838–856. (1993).

[53] M. Naor, L. J. Schulman, A. Srinivasan. Splitters and near-optimal derandomization, Proc. of the 36th IEEE Symp. on Foundations of Computer Science, (FOCS 95), pp. 182–191. (1995).

[54] A. Poli. A deterministic construction of normal bases with complexity $O(n^3 + n \log n \log\log n \log q)$. *J. Symb. Comp.*, **19**, pp. 305–319. (1995).

[55] E. Porat, A. Rothschild. Explicit non-adaptive combinatorial group testing schemes. *IEEE Transactions on Information Theory*, 57(12), pp. 7982–7989. (2011).

[56] R. M. Roth. Introduction to coding theory. Cambridge University Press, Cambridge, UK. (2006).

[57] H. J. Ryser. Combinatorial mathematics, the carus mathematical monographs No. 14, The Mathematical Association of America. (1963).

[58] C. Saha. A note on irreducible polynomials and identity testing. Manuscript. (2008).

[59] N. Saxena. Progress on polynomial identity testing. Electronic Colloquium on Computational Complexity. Report No. 101. (2009).

[60] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4), pp. 701-717. (1980).

[61] V. Shoup. New algorithms for finding irreducible polynomial over finite field. *Mathematics of Computation*, 54(189), pp. 435–447. (1990).

[62] I. Shparlinski. On finding primitive roots in finite fields. *Theoretical Computer Science*, 157(2), pp. 273–275. (1996).

[63] I. Shparlinski. Finding irreducible and primitive polynomials. *Applicable Algebra in Engineering Communication and Computing*, 4(4), pp. 263–268. (1993).

[64] I. Shparlinski. Finite fields: theory and computation. Mathematics and Its Applications, Vol. 477. (1999).

[65] K. W. Shum. A low-complexity algorithm for the construction of algebraic-geometric codes better than the Gilbert-Varshamov bound. A Dissertation. University of Souuthern California. (2000).

[66] G. Seroussi, N. H. Bshouty. Vector sets for exhaustive testing of logic circuits. *IEEE transaction on information theory*, 34(3), pp. 513–522. (1988).

[67] H. Stichtenoth. Algebraic function fields and codes. Second Edition, Vol. 254, Springer. (2008).

[68] V. Shoup. Finding irreducible and primitive polynomials. *Appl. Algebra in Engin., Commun. and Computing*, 4, pp. 263–268. (1993).

[69] K. W. Shum, I. Aleshnikov, P. V. Kumar, H. Stichtenoth, V.Deolalikar. A low-complexity algorithm for the construction of algebraic-geometric codes better than the Gilbert-Varshamov bound. *IEEE Transactions on Information Theory*, 47(6), pp. 2225–2241. (2001).

[70] N. Saxena, C. Seshadhri. Blackbox identity testing for bounded top fanin depth-3 circuits: the field doesn't matter. Electronic Colloquium on Computational Complexity, Report No. 46. (2011).

[71] R. E. Schapire, L. Sellie. Learning sparse multivariate polynomials over a field with queries and counterexamples. *J. Comput. Syst. Sci.*, 52(2), pp. 201–213. (1996).

[72] G. Seroussi, A. Lempel. On symmetric algorithms for bilinear forms over finite fields. *J. Algorithms*, 5(3), pp. 327–344. (1984).

[73] A. Shpilka, I. Volkovich. Improved polynomial identity testing for read-once formulas. Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization, (APPROX-RANDOM 09), pp. 700–713. (2009).

[74] S. Saraf, I. Volkovich. Black-box identity testing of depth-4 multilinear circuits. Electronic Colloquium on Computational Complexity, Report No. 46. (2011).

[75] D.R. Stinson, T. van Trung, R. Wei. Secure frameproof codes, key distribution patterns, group testing algorithms and related structures, *J. Stat. Planning and Inference*, 86(2), pp. 595-617. (2000).

[76] D.R. Stinson, R. Wei, K. Chen. On generalised separating hash families. *Journal of Combinatorial Theory, Series A*, 115(1), pp. 105–120. (2008).

[77] D. R. Stinson, R. Wei, L. Zhu. New constructions for perfect hash families and related structures using combinatorial designs and codes. *Journal of Combinatorial Designs*, 8(3), pp. 189-200. (2000).

[78] D. R. Stinson, R. Wei, L. Zhu. Some new bounds for cover-free families, *Journal of Combinatorial Theory, Series A*, 90(1), pp. 224-234. (2000).

[79] D. R. Stinson, R. Wei, L. Zhu. New constructions for perfect hash families and related structures using combintorial designs and codes, *J. Combin. Designs.*, 8(3), pp. 189-200. (2000).

[80] K. Werther. The complexity of sparse polynomial interpolation over finite fields. *Applicable Algebra in Engineering, Communication and Computing*, 5(2), pp. 91–103. (1994).

[81] H. Wang and C. P. Xing. Explicit Constructions of perfect hash families from algebraic curves over finite fields. *J. of Combinatorial Theory, Series A*, 93(1), pp. 112–124. (2001).

[82] R. Zippel. Probabilistic algorithms for sparse polynomials. In Proceedings of the International Symposiumon on Symbolic and Algebraic Computation, 72, pp. 216-226. (1979).

8 Appendices

8.1 Appendix A. Algebraic Function Fields

In this appendix we give some notations and basic results from the theory of algebraic function fields in order to be able to follow the proofs in the paper.

An algebraic function field $F/\mathbb{F}_q$ of one variable is an extension field $F \supset \mathbb{F}_q$ such that F is a finite algebraic extension of $\mathbb{F}_q(x)$ for some element $x \in F$ which is transcendental over $\mathbb{F}_q$.

A *valuation ring* $\mathcal{O}$ is a ring $\mathbb{F}_q \subset \mathcal{O} \subset F$ that is not $\mathbb{F}_q$ and not F such that for every $z \in F$ we have that $z \in \mathcal{O}$ or $z^{-1} \in \mathcal{O}$. The ring $\mathcal{O}$ is a local ring (see [67], Proposition 1.1.5). Let $\mathcal{O}^\times$ be the group of units of $\mathcal{O}$. Then $P_\mathcal{O} = \mathcal{O}\backslash\mathcal{O}^\times$, the set of non-units in $\mathcal{O}$, is a unique maximal ideal of $\mathcal{O}$ and is a principle ideal, that is, $P_\mathcal{O} = t\mathcal{O}$ for some $t \in P_\mathcal{O}$ (see [67], Proposition 1.1.5 and Theorem 1.1.6). Also, the element t is unique up to unit element multiplication, that is, if $P_\mathcal{O} = t'\mathcal{O}$ then $t = t'u$ for some $u \in \mathcal{O}^\times$ (follows from Theorem 1.1.13 in [67]). Such t is called *prime* element of P.

A *place* P of a function field is the maximal ideal of some valuation ring $\mathcal{O}$. Given a place P. We have $F\backslash\mathcal{O} = P^{-1} = \{a^{-1} \mid a \in P\}$ and $\mathcal{O} = F\backslash P^{-1}$ ([67], page 4). Therefore a place P uniquely determines the valuation ring $\mathcal{O}$. We denote by $\mathcal{O}_P$ the valuation ring corresponds to the place P. We denote by $\mathbb{P}_F$ the set of all places of $F/\mathbb{F}_q$.

Since P is a maximal ideal of $\mathcal{O}_P$, the *residue class ring* $F_P = \mathcal{O}_P/P$ is a field. For $x \in \mathcal{O}_P$ we define $x(P) \in \mathcal{O}_P/P$ to be the residue class of x modulo P. For $x \in F\backslash\mathcal{O}_P$ we define $x(P) = \infty$. The map $x \mapsto x(P)$ is called the *residue class map*. It is easy to verify that this map satisfies the following.

Proposition 88. *We have*

1. *If $a(P) + b(P)$ is defined then $(a+b)(P) = a(P) + b(P)$.*
2. *If $a(P)b(P)$ is defined then $(ab)(P) = a(P)b(P)$.*
3. *$0(P) = 0$ and $1(P) = 1$.*

Here $\alpha + \infty = \infty + \alpha = \infty$ for every $\alpha \in F_P$ and $\infty + \infty$ is not defined. Also for $\alpha \in F_P\backslash\{0\}$, $\alpha \cdot \infty = \infty \cdot \alpha = \infty \cdot \infty = \infty$ and $\infty \cdot 0$ is not defined. It follows from the above definition that for $\alpha \neq 0, \infty$ we have $\alpha/\infty = 0$ and $\alpha/0 = \infty$. Notice also that

$$P = \{z \in F \mid z(P) = 0\} \quad \text{and} \quad \mathcal{O}_P = \{z \mid z(P) \in \mathbb{F}_q\}.$$

The *degree* of a place P is $\deg P = [F_P : \mathbb{F}_p]$ and is always finite (see [67], Proposition 1.1.15).

To a place $P \in \mathbb{P}_F$ we associate a function $v_P : F \to \mathbb{Z} \cup \{\infty\}$ (or ord_P) as follows: Choose a prime element t of P. Then every $0 \neq z \in F$ has a unique representation $z = t^n u$ with some $u \in \mathcal{O}_P^\times$ and $n \in \mathbb{Z}$ (see [67], Theorem 1.1.6). Then define $v_P(z) = n$. Also define $v_P(0) = \infty$. It can be shown that $v_P(z)$ depends only on P, not on the choice of t (see [67], Theorem 1.1.6). The value $v_P(z)$ (or $\operatorname{ord}_P(z)$) is called the *valuation* (or the *order*) of z.

The place P is called a *zero* of x if $v_P(x) > 0$ and a *pole* of x if $v_P(x) < 0$.

The following are some properties of valuation (see [67], pages 4–6)

Proposition 89. *For $a, b \in F$ we have*

1. $v_P(a) = \infty$ *if and only if* $a = 0$.
2. $v_P(ab) = v_P(a) + v_P(b)$
3. $v_P(a + b) \geq \min(v_P(a), v_P(b))$
4. $v_P(a) = 0$ *for all* $a \in \mathbb{F}_q^\times$.
5. $v_P(x) < 0$ *if and only if* $x \in F \backslash R_P$ *if and only if* $x(P) = \infty$.
6. $v_P(x) = 0$ *if and only if* $x \in R_P^\times$ *if and only if* $x(P) \notin \{0, \infty\}$
7. $v_P(x) > 0$ *if and only if* $x \in P$ *if and only if* $x(P) = 0$.
8. *The prime elements are* $t \in R$ *such that* $v_P(t) = 1$.

The additive free abelian group which generated by $\mathbb{P}_F$ is denoted by $\operatorname{Div}(F)$ and is called the *divisor group* of $F/\mathbb{F}_q$. Each divisor is a formal sum

$$D = \sum_{P \in \mathbb{P}_F} n_P P$$

where $n_P \in \mathbb{Z}$ and all except a finite number of which are not zero. A divisor $D = P$ with $P \in \mathbb{P}_F$ is called a *prime divisor*. The *support* of D is

$$\operatorname{supp} D = \{P \in \mathbb{P}_F \mid n_P \neq 0\}.$$

The *principle divisor* (x) (or $\operatorname{div}(x)$) of $x \in F$ is

$$(x) = \sum_{P \in \mathbb{P}_F} v_P(x) P.$$

We will also use v_P for divisors and denote $v_P(D) = n_P$. A divisor D is called *integral divisor* or *positive* if for every place $P \in \mathbb{P}_F$ we have $v_P(D) \geq 0$. Every divisor D can be written as $D = D_+ - D_-$ where D_+ and D_- are positive divisors and

$$D_+ = \sum_{v_P(D)>0} v_P(x)P, \quad D_- = \sum_{v_P(D)<0} (-v_P(x))P.$$

For the principle divisor (x) we define: $(x)_0 = (x)_+$ and call it the *zero divisor* of x, and $(x)_\infty = (x)_-$ and call it the *pole divisor* of x. Then $(x) = (x)_0 - (x)_\infty$.

The *degree* of the divisor D is defined as

$$\deg D = \sum_{P \in \mathbb{P}_F} v_P(D) \cdot \deg P.$$

All principle divisors have degree 0 (see [67], Theorem 1.4.11).

The set of principle divisors $\mathrm{Princ}(F) = \{(x) \mid 0 \neq x \in F\}$ is called the *group of principle divisors* of $F/\mathbb{F}_q$. The factor group $\mathrm{Cl}(F) = \mathrm{Div}(F)/\mathrm{Princ}(F)$ is called the *divisor class group* of $F/\mathbb{F}_q$. For a divisor $D \in \mathrm{Div}(F)$ we denote by $[D]$ the divisor class of D in $\mathrm{Cl}(F)$. For two divisors D_1 and D_2 we write $D_1 \sim D_2$ if $[D_1] = [D_2]$.

For two divisors D_1 and D_2 we write $D_1 \leq D_2$ if $v_P(D_1) \leq v_P(D_2)$ for every $P \in \mathbb{P}_F$.

For a divisor $D \in \mathrm{Div}(F)$ define the *Riemann-Roch* space associated to D by

$$\mathscr{L}(D) = \{x \in F \mid (x) \geq -D\} \cup \{0\}. \tag{40}$$

Here are some properties (see [67], Lemmas 1.4.6 and 1.4.7)

Proposition 90. *Let $D, D_1, D_2 \in \mathrm{Div}(F)$ and $P \in \mathbb{P}_F$. We have*

1. *$z \in \mathscr{L}(D)$ if and only if for every place P we have $v_P(z) \geq -v_P(D)$.*
2. *$\mathscr{L}(D)$ is a linear space over $\mathbb{F}_q$.*
3. *If $\deg D < 0$ then $\mathscr{L}(D) = \{0\}$ and $l(D) = 0$*
4. *If $D_1 \leq D_2$ then $\mathscr{L}(D_1) \subseteq \mathscr{L}(D_2)$.*
5. *If $P \notin \mathrm{supp}\ D$, $z \in \mathscr{L}(D)$ and $v_P(z) > 0$ then $z \in \mathscr{L}(D - P)$.*

We define the *dimension* of the divisor D as $l(D) = dim\ \mathscr{L}(D)$.

We now prove

Proposition 91. *Let $D \in \mathrm{Div}(F)$. For $f \in \mathcal{P}(\mathbb{F}_q, n, d)$ and $\boldsymbol{z} = (z_1, \ldots, z_n) \in \mathscr{L}(D)^n$ we have $f(\boldsymbol{z}) \in \mathscr{L}(dD)$.*

Proof. Let $M = x_{i_1} \cdots x_{i_{d'}}$ be a monomial in f where $d' \leq d$. Since by Proposition 89, for every place P we have

$$v_P(z_{i_1} \cdots z_{i_{d'}}) = \sum_{j=1}^{d'} v_P(z_{i_j}) \geq -\sum_{j=1}^{d'} v_P(D) = -v_P(d'D) \geq -v_P(dD),$$

$M(\boldsymbol{z}) \in \mathscr{L}(dD)$. Since this is true for every monomial M in f and since $\mathscr{L}(dD)$ is a linear space we have $f(\boldsymbol{z}) \in \mathscr{L}(dD)$. □

The *genus* g or $g(F)$ of $F/\mathbb{F}_q$ is defined by

$$g(F) = \max\{deg\ D - l(D) + 1 \mid D \in \mathrm{Div}(F)\}.$$

Therefore for every divisor D we have $l(D) \geq deg\ D + 1 - g$. A divisor D is called *non-special* if $l(D) = deg\ D + 1 - g$. A divisor W is called *canonical* if $\deg W = 2g - 2$ and $l(W) \geq g$. One of the most important theorem in the theory algebraic function field is the following (see [67], Theorem 1.5.15 and Theorem 1.5.17)

Proposition 92. (Riemann-Roch Theorem) *Let W be a canonical divisor and A be any divisor. Then*

1. *$l(A) = \deg A + 1 - g + l(W - A)$.*
2. *If $\deg A \geq 2g - 1$ then A is non-special and $l(A) = \deg A + 1 - g$.*

8.2 Appendix B. Toward Testers for $q \geq d + 1$ with Better Size

In this appendix we show that if there is a polynomial time algorithm that finds for a given divisor U a basis for $\mathscr{L}(U)$ then a tester with a better size than in Corollary 41 can be constructed in polynomial time.

To find a basis for $\mathscr{L}(U)$ one can use the Brill-Noether Theorem [21] or Hess algorithm [38]. We were unable to find the time complexity of those algorithms. Laursen in [46] shows that the Brill-Noether algorithm runs in polynomial number of "steps" but it is not clear that each step can be performed in polynomial time.

Before we give the main result, we recall the definition of the $\mathcal{W}_1$ tower. Let x_1 be indeterminate over $\mathbb{F}_{q^2}$ and $F^{(1)} = \mathbb{F}_{q^2}(x_1)$. For $k \geq 1$ let $F^{(k)} = F^{(k-1)}(x_k)$ where

$$x_k^q + x_k = \frac{x_{k-1}^q}{x_{k-1}^{q-1} + 1}.$$

We will denote by **BASIS**(s, r) the problem of finding a basis of $\mathscr{L}(U)$ for a divisors U where $\deg(U_-)+\deg(U_+) \le s$ in $F^{(r)}$. We don't know whether this problem can be solved in polynomial time even when $s = O(\log t/\log q)$. So here we can only state the following claim as a conjecture

Conjecture 1. *The* **BASIS**$(O(d\log t/\log q), O(\log(\log t/\log q)/\log q))$ *problem can be solved in polynomial time.*

In this appendix we prove the following

Corollary 93. *Let $c > 2$ be constant and $r = O(\log t/\log q)$. If Conjecture 1 is true then we get the following upper bounds for $\tau_{poly}(d, q, t)$ and $\tau_{poly}(d, q, t, r)$ (columns 4 and 5 in the table)*

q	t	*Upper B.* $\tau^*(d, q, t)$	*Upper B.* $\tau_{poly}(d, q, t, r)$	*Upper B.* $\tau_{poly}(d, q, t)$	*Lower B.* $\tau(d, q, t)$
$q \ge c^2(d+1)^2$, *q P.S.*	*I.S.*	1	1	2	1
$q \ge c^2(d+1)^2$, *q P.S.*	*all*	2	2	3	1
$q \ge c(d+1)$	*I.S.*	2	2	3	1
$q \ge c(d+1)$	*all*	3	3	4	1
$q \ge d+1$	*I.S.*	3	3	4	1
$q \ge d+1$	*all*	4	4	5	1

where I.S. stands for "for infinite sequence of integers t" and P.S. for "perfect square".

Note. In some cases we can replace **BASIS**$(O(d\log t/\log q), O(\log(\log t/\log q)/\log q))$ in Conjecture 1 to **BASIS**$(O(\log t/\log q), O(\log(\log t/\log q)/\log q))$, but we will not discuss this here.

We now give a detailed sketch of the proof. First, as in the proof of Lemma 39, we may assume that $t \ge q^d$.

The idea of the proof is to first use Lemma 38 to reduce the dimension t of the problem to $t' = O(\log t/\log q)$ and then use a construction similar to the construction in Theorem 11 for dimension t'.

By Lemma 38, we have

$$\begin{aligned}\tau_{poly}(d, q, t) &= \tau_{poly}(d, q, t, t)\\ &\le \tau_{poly}\left(d, q, t, \left\lceil\frac{\log(dt)}{\log q}\right\rceil + 1\right) + 1\end{aligned}$$

Our goal now will be to give a construction similar to the construction in Theorem 11 for

$$t' = \left\lceil\frac{\log(dt)}{\log q}\right\rceil + 1.$$

Notice that for the construction in Theorem 11, we need to construct the $(\mathcal{P}(\mathbb{F}_q, n, d), \mathscr{L}(G), \mathbb{F}_q)$-tester defined in Lemma 12 and the $(\mathbb{F}_q[\boldsymbol{x}], \mathbb{F}_{q^{t'}}, \mathscr{L}(G))$-tester defined in Lemma 13. To construct those testers we need

1. To find all the places of degree 1 in $F^{(r)}/\mathbb{F}_{q^2}$ for the tower $\mathcal{W}_1$, [36, 37], where $r = O(\log t'/\log q)$.
2. Find a prime divisor Q of $F^{(r)}/\mathbb{F}_{q^2}$ of degree t'.
3. Find a divisor G of degree $t' + g - 1 = O(dt')$ that satisfies the conditions in Lemma 14.
4. Find a basis for $\mathscr{L}(G)$.

For 1, it is known from [65] and [69] that all the places of degree 1 in this function field can be found in time $poly(d, t)$. For 2, by Lemma 42, a prime divisor of degree t' can be constructed in polynomial time. Item 4 follows from Conjecture 1. For 3, it is not clear whether this can be done in polynomial time. Instead, we use a different result that gives a slightly weaker tester.

We first prove

Lemma 94. *Let $F/\mathbb{F}_q$ be algebraic function field of genus g that contains at least $2g$ places $P_1, P_2, \ldots, P_{2g}$ of degree 1. Let $t \geq g$ be an integer. Let Q and R be a prime divisors of degree t and $t-1$, respectively. There are $1 \leq i_1 < i_2 < \cdots < i_g \leq 2g$ such that*

$$G = R + P_{i_1} + \cdots + P_{i_g}$$

satisfies

1. $v_Q(G) = 0$
2. $v_P(G) = 0$ *for any prime divisor* $P \notin \{P_{i_1}, \ldots, P_{i_g}\}$ *of degree* 1.
3. $l(G) = \deg Q = t$.
4. $\deg(G) = t + g - 1$.
5. $l(G - Q) = 0$.

Proof. *1.* *2.* and *4.* are obvious for any $1 \leq i_1 < i_2 < \cdots < i_g \leq 2g$. Also since $\deg G = t-1+g \geq 2g-1$ by Proposition 92, $l(G) = \deg G + 1 - g = t$. This implies *3.* It remains to prove *5.*

The proof is by induction. Since $\deg(R - Q) = -1$, by *3* in Proposition 90, we have $l(R - Q) = 0$. Suppose for some $k < g$ there are $0 \leq i_1 < i_2 < i_3 < \cdots < i_k \leq 2g$ where $R_k = R + P_{i_1} + \cdots + P_{i_k}$ satisfies $l(R_k - Q) = 0$. Suppose without loss of generality $i_1 = 2g, i_2 = 2g - 1, \ldots, i_k = 2g - k + 1$. Since $2g - k + 1 \geq g + 2$ we have $i_1, \ldots, i_k \in \{g + 2, g + 3, \ldots, 2g\}$.

Consider the divisors $R_k+P_1-Q,\ldots,R_k+P_{g+1}-Q$. Suppose $l(R_k+P_i-Q)\neq 0$ for all $i=1,2,\ldots,g+1$. Then there is $z_i\in\mathscr{L}(R_k+P_i-Q)\backslash\mathscr{L}(R_k-Q)$ for all $i=1,2,\ldots,g+1$. Therefore, by the definition of the Riemann-Roch space (40), it follows that $v_{P_i}(z_i)=-1$ and for all $j\in[g+1]$, $j\neq i$ we have $v_{P_j}(z_i)\geq 0$.

We now show that $z_1,\ldots,z_{g+1}$ are linearly independent over $\mathbb{F}_q$. Suppose there are $\lambda_i\in\mathbb{F}_q$, $i=1,\ldots,g+1$ such that $\lambda_1z_1+\cdots+\lambda_{g+1}z_{g+1}=0$. Let $t_i\in F$ be such that $v_{P_i}(t_i)=1$. The existence of such element follows from the strong approximation theorem, Theorem 1.6.5 in [67]. Then by *2* in Proposition 89, $v_{P_i}(t_iz_i)=0$ and $v_{P_i}(t_iz_j)\geq 1$ for all $j\neq i$. Then by *7* in Proposition 89 and Proposition 88, $0=(\lambda_1t_iz_1+\cdots+\lambda_{g+1}t_iz_{g+1})(P_i)=\lambda_i\cdot(t_iz_i)(P_i)$. Since by *6* in Proposition 89, $(t_iz_i)(P_i)\neq 0$ we get $\lambda_i=0$.

Now let $D=R_k-Q+(g-k)P_1+P_2+P_3+\cdots+P_g+P_{g+1}$. Then $\deg D=2g-1$. Since $D\geq R_k+P_i-Q$, $i=1,\ldots,g+1$ we have $\mathscr{L}(R_k+P_i-Q)\subseteq\mathscr{L}(D)$. Therefore $z_1,\ldots,z_{g+1}\in\mathscr{L}(D)$. Since $z_1,\ldots,z_{g+1}$ are linearly independent over $\mathbb{F}_q$ we have $l(D)\geq g+1$. On the other hand since $\deg D=2g-1$ by *2* in Proposition 92, we have $l(D)=g$. This gives a contradiction. Therefore there exists $P_{i_{k+1}}\in\{P_1,P_2,\ldots,P_{g+1}\}$ such that for $R_{k+1}=R_k+P_{i_{k+1}}$ we have $l(R_{k+1}-Q)=0$. □

Notice now that to construct G we need to compute $\mathscr{L}(R_k+P_i-Q)$ where $\deg(R_k)+\deg(P_i)+\deg(Q)=O(dt')$ and find a prime divisor of degree $t'-1$. By Conjecture 1 and Lemma 42 such divisor can be constructed in polynomial time.

The problem with the divisor G in Lemma 94 is that it uses (and therefore it burns) g places of degree 1 that cannot be used in the tester. See Lemma 12.

Now the proof proceed exactly the same as the proof of Corollary 17 where the only change is that only $N-g\geq q^{r+2}-2q^{r+1}$ places can be used instead of $N\geq q^{r+2}-q^{r+1}$.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de