Sparse affine-invariant linear codes are locally testable

Eli Ben-Sasson* Noga Ron-Zewi † Madhu Sudan‡

April 27, 2012

Abstract

We show that sparse affine-invariant linear properties over arbitrary finite fields are locally testable with a constant number of queries. Given a finite field $\mathbb{F}_q$ and an extension field $\mathbb{F}_{q^n}$, a property is a set of functions mapping $\mathbb{F}_{q^n}$ to $\mathbb{F}_q$. The property is said to be affine-invariant if it is invariant under affine transformations of $\mathbb{F}_{q^n}$, and it is said to be sparse if its size is polynomial in the domain size. Our work completes a line of work initiated by Grigorescu et al. [RANDOM 2009] and followed by Kaufman and Lovett [FOCS 2011]. The latter showed such a result for the case when q was prime. Extending to non-prime cases turns out to be non-trivial and our proof involves some detours into additive combinatorics, as well as a new calculus for building property testers for affine-invariant linear properties.

*Department of Computer Science, Technion, Haifa, Israel and Microsoft Research New-England, Cambridge, MA. eli@cs.technion.ac.il. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 240258.

†Department of Computer Science, Technion, Haifa. nogaz@cs.technion.ac.il. Research was conducted while the author was an intern at Microsoft Research New-England, Cambridge, MA, and supported by the Israel Ministry of Science and Technology.

‡Microsoft Research New-England, Cambridge, MA. madhu@mit.edu

Contents

1 Introduction

This paper investigates property testing in the context of linear, affine-invariant properties and proves that all sparse properties in this class are testable. We describe these notions more precisely below, before explaining the context and motivation for this study.

1.1 The problem and main result

Given finite sets D and R (for domain and range), a *property* of functions mapping D to R is simply given by a subset $\mathcal{F} \subseteq \{D \to R\}$ ($\mathcal{F}$ is the subset of functions that satisfy the property). Property testing investigates the possibility of efficient algorithms that make few queries to an oracle for $f : D \to R$ and accepts $f \in \mathcal{F}$ while rejecting f that is very far from $\mathcal{F}$ with constant probability. Distance here is measured in normalized Hamming distance and so $\delta(f, g) = \frac{1}{|D|} \cdot |\{x | f(x) \neq g(x)\}|$ and $\delta(f, \mathcal{F}) = \min_{g \in \mathcal{F}}\{\delta(f, g)\}$. A property $\mathcal{F}$ is said to be k-locally testable if there exists a tester making at most k queries to a function $f : D \to R$, that accepts $f \in \mathcal{F}$ with probability 1, while rejecting all f with probability at least $\delta(f, \mathcal{F})$.

A large, and very important, class of properties, namely the algebraic ones, are abstracted best by the features of being *linear* and *affine-invariant*. In such settings the range of the property is a (small) finite field $\mathbb{F}_q$ (where $\mathbb{F}_q$ denotes the field of size q) and the domain is a (large) finite extension $\mathbb{F}_{q^n}$. A property $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is *linear* if it is an $\mathbb{F}_q$-vector space, i.e., $\forall f, g \in \mathcal{F}$ and $\alpha \in \mathbb{F}_q$ we have $\alpha f + g \in \mathcal{F}$. The property $\mathcal{F}$ is said to be affine-invariant if it is invariant under affine-transformations of the domain, i.e., $\forall \alpha, \beta \in \mathbb{F}_{q^n}$ with $\alpha \neq 0$, and $\forall f \in \mathcal{F}$ it is the case that $f_{\alpha,\beta}$ given by $f_{\alpha,\beta}(x) = f(\alpha \cdot x + \beta)$ is also in $\mathcal{F}$.

Finally, we say that $\mathcal{F}$ is sparse if it contains only polynomially many functions in its domain size. More precisely, we say that $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is t-(size-)sparse if $|\mathcal{F}| \leq q^{nt}$. Our main theorem shows that all sparse properties are testable with a constant number of queries.

Theorem 1.1 (Main)**.** *For every q and t there exists $k = k_{q,t}$ such that for every n, every t-sparse, linear, affine-invariant property $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is k-locally testable.*

Our work extends prior work of Grigorescu et al. [GKS09] and Kaufman and Lovett [KL11]. The latter, in particular, proved the above theorem when q is prime, leaving open the case of all extensions of prime fields. We describe the relationship to previous work and explain our technical contributions after discussing the motivation for studying affine-invariant linear properties.

1.2 Motivation

Property Testing: The general motivation to understand linear, affine-invariant, properties is that they form the most natural abstraction of some of the most useful class of property tests that have played a role in the construction of locally testable codes and probabilistically checkable proofs. Some central properties that have been utilized in such constructions have been the "linearity" property and the "low-degree" property. Affine-invariant properties abstract such properties in as natural a manner as "graph properties" abstract specific properties such as triangle-freeness or bipartiteness. Given the major role played by algebraic properties, understanding their testability seems as important as understanding testability of, say, graph properties.

Locally testable codes: If the study of affine-invariance is natural in the context of property testing, the restriction to linearity is as natural in the context of error-correcting codes. Most well-studied error-correcting codes are linear and the locally testable ones are usually derived from linear

locally testable properties. We note that the very fact that a property is linear, affine-invariant and locally testable implies that it is an error-correcting code. By the work of Ben-Sasson et al. [BHR05] it is known that all locally testable codes must be what are known as "LDPC codes", where the code is defined by a collection of local constraints. However it is also known, from the work of Ben-Sasson et al. [BGK+05] that in order to be locally testable the LDPC code must have a redundant collection of local constraints. Redundancy among local constraints is a relatively rare phenomenon and imposing some symmetry (such as affine-invariance) is one way of getting such redundancy. Indeed the symmetry offered by affine-invariance is the only setting where (with some additional features) the redundancy is known to lead to testable codes. Thus affine-invariant linear properties lead to some of the most natural and broad classes of locally testable codes.

In spite our relatively good understanding of the structure of affine-invariant linear properties we do not yet have a *characterization* of what makes such properties locally testable, as is the case for graph properties [AFNS06, BCL+06]. The current belief seems to be that a k-query testable property $\mathcal{F} \subset \left\{\mathbb{F}_q^n \to \mathbb{F}_q\right\}$ is a combination of a constant number of "base-properties" where base-properties are of two kinds — "low-degree" properties (also known as Reed-Muller codes of constant degree) and "sparse" ones. (For a detailed description of this belief and its ensuing conjectures see Section 5 in Ben-Sasson et al. [BGM+11a].) But our limited understanding of affine-invariant linear locally testable codes means that we cannot rule out the existence of some other property, neither "low-degree" nor "sparse", that nevertheless is locally testable. And till this work, it was not even known that every combination of "base-properties" does indeed lead to testability. Thus this work finally completes the "easy direction" of the project aiming to characterize affine-invariant linear locally testable properties, and does this by showing that all finite combinations of "base-properties" that are believed to be testable are indeed so. What is still lacking now is a limitation result saying that the remaining classes of affine-invariant linear properties are not testable.

1.3 Comparison with previous work

The task of testing sparse codes was initiated in Kaufman and Litsyn [KL05], and then pursued further in Kaufman and Sudan [KS07b] and most recently by Kopparty and Saraf [KS10]. All the above results show that if a code is sparse *and of very high distance* then it is testable. ([KL05], [KS07b] only deal with binary codes. The results of [KS10] seems to extend to prime-alphabet codes, or even q-ary alphabet case, though the results are not stated so.)

The task of testing sparse affine-invariant linear properties was initiated by [GKS09]. They showed that in some special cases binary sparse affine-invariant linear properties were testable. [KL11] extended the result vastly — they showed that every sparse affine-invariant linear property over a *prime* field $\mathbb{F}_p$ is testable. The main ingredient in the proofs of the above results shows that sparse affine-invariant linear properties satisfy the sufficient condition (high-distance) required in the results mentioned in the previous paragraph. While they also give "nice" tests in the process, this may be viewed as a bonus, but not necessary for testability.

Testing over non-prime finite fields turns out to be more involved for a fundamental reason. Codes over $\mathbb{F}_q$ where $q = p^s$, p is a prime and $s > 1$, have decent distance, but certainly nowhere close to being "excellent" in the sense required in all the previous works. Indeed previous results relied crucially on the fact that every non-zero function from the sparse property in question was roughly balanced (took on every value in the range roughly the same number of times). Such a statement is simply not true in our setting. The reason is not just that $\mathbb{F}_q$ contains $\mathbb{F}_p$ as a subfield, but moreover that $\mathbb{F}_q$ contains many vector spaces over the prime subfield $\mathbb{F}_p$. Indeed for every such subspace V of $\mathbb{F}_q$ it is possible to create sparse properties that contain functions which take on values only from V, and take on every value in V roughly the same number of times. This obstacle

turns out to be sufficient enough to derail the previous proof techniques (which are still useful, but insufficient).

To overcome this obstacle we revisit the structure of affine-invariant linear properties and introduce a simple calculus for building tests for such properties. Our final tests also use some of the algebraic machinery coming from the proofs of the sum-product theorems to build the necessary tests.

1.4 Technical contributions

Previous works on testing affine-invariant linear properties have already shown that it suffices to consider tests that distinguish some "basic" functions. Specifically, if we let $\mathrm{Trace} : \mathbb{F}_{q^n} \to \mathbb{F}_q$ denote the standard trace map given by $\mathrm{Trace}(x) = x + x^q + \cdots + x^{q^{n-1}}$, then it (roughly) suffices to build "tests" that simultaneously accept some good functions, of the form $\mathrm{Trace}(x^d)$ with $d \in G$, while rejecting all bad ones, of the form $\mathrm{Trace}(x^e)$ for $e \in B$. For simplicity think of a "test" of arity k as specified by a tuple $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_{q^n}$ in conjunction with a $\mathbb{F}_q$-linear form $(\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$. The "test" accepts d if $\sum_{i=1}^k \lambda_i \mathrm{Trace}(\alpha_i^d) = 0$ and it rejects e if $\sum_{i=1}^k \lambda_i \mathrm{Trace}(\alpha_i^e) \neq 0$. The sets G and B depend on the property $\mathcal{F}$ being tested. Previous analysis, especially [KL11], picked a random test of constant size that accepted all the good functions and were able to claim that with (overwhelmingly) high probability such a test would reject all bad functions. This claim relied on the fact that all non-zero functions (good/bad) took on each value in the range roughly equally often. This fact is no longer true in our case and translates into an algebraic challenge. For some $d \in G$ and $e \in B$ it is no longer the case that a random "test" that accepts $\mathrm{Trace}(x^d)$ will reject $\mathrm{Trace}(x^e)$ with high probability. A particularly challenging case for us is when $e = p^i d$, where p is the characteristic of the field we are working with. For this specific case, we manage to "handcraft" a test, using some methods from additive combinatorics, that accepts $\mathrm{Trace}(x^d)$ while rejecting $\mathrm{Trace}(x^{p^i d})$. This is the central technical contribution of this work and we give some insight into it next.

If we are so lucky as to have an element α in $\mathbb{F}_{q^n}$ such that $\lambda \triangleq \alpha^d$ is contained in $\mathbb{F}_q$ but not contained in $\mathbb{F}_{p^i}$ then we are in good shape: The "test" that checks whether "$\lambda \cdot f(1) = f(\alpha)$?" accepts $f(x) = \mathrm{Trace}(x^d)$ while rejecting $f(x) = \mathrm{Trace}(x^{p^i d})$. In general we cannot guarantee the existence of such a lucky α. Therefore we consider the set $A = \left\{\alpha^d \mid \alpha \in \mathbb{F}_{q^n}\right\}$ and its ℓ-wise sumset $\ell A = \{a_1 + \ldots + a_\ell \mid a_i \in A\}$. If we could prove that ℓA contains an element $\lambda \in \mathbb{F}_q \setminus \mathbb{F}_{p^i}$ for some constant ℓ (possibly depending on the sparsity of $\mathcal{F}$ and q) we would still be okay. This also seems plausible, since the set A is completely closed under multiplication and so the sum-product estimates [BGK06] show that $|\ell A| \gg |A|$. Thus it is conceivable that the larger set ℓA might contain a nice λ, and if so we would have a constraint of arity roughly ℓ separating $\mathrm{Trace}(x^d)$ from $\mathrm{Trace}(x^{p^i d})$.

Determining the smallest ℓ for which ℓA is closed under addition (for a given d) is well-studied as Waring's problem for finite fields. The best bound, due to Cochrane and Cipra, is roughly of the form $\ell \leq d^{1/\log|A|}$ [CC11] (see [Cip10] for more information). For general d, the parameter ℓ may need to grow with n, however in our case d is restricted (due to the sparsity of $\mathcal{F}$), so the above bound gives constant ℓ. For the sake of presenting a simple and self-contained proof, we provide a solution to a problem that is somewhat more specific than Waring's problem, yet suffices for our purposes and lends more easily to analysis. Based on the simplified analysis of the sum-product theorem in [BIW06], we consider sets A_ℓ of the form $A_\ell = (\ell A - \ell A)/(\ell A - \ell A)$ (i.e., sets containing ratios of two elements each of which is expressible as the difference of two elements of ℓA). We show, with a self-contained elementary proof, that for sufficiently large ℓ the set A_ℓ is closed under

addition, hence contains a $\lambda \in \mathbb{F}_q \setminus \mathbb{F}_{p^i}$. With some additional work we are then able to mimic the "lucky" case above to get a constraint of arity $O(\ell)$ separating $\mathrm{Trace}(x^d)$ from $\mathrm{Trace}(x^{p^i d})$.

Unfortunately, while the handcrafted test manages to settle the toy challenge for a single pair d, e, it fails to build a single test that *simultaneously* accepts all the good functions $\mathrm{Trace}(x^d)$, $d \in G$, while rejecting all the bad functions. In particular, the literature on affine-invariant property testing that reduced testing to distinguishing basic functions seemed to crucially rely on the fact that the tests simultaneously accepted all the functions $\mathrm{Trace}(x^d)$ for $d \in G$. Tests that accept just one of the basic functions seem to be useless in their setting. Indeed we call our tests distinguishing $\mathrm{Trace}(x^d)$ from $\mathrm{Trace}(x^e)$ "pseudo-tests" due to this reason. To use our pseudo-tests, we build a calculus for combining pseudo-tests which allows us to build larger pseudo-tests which combine smaller pseudo-tests to either enlarge the set of good functions being accepted or to enlarge the set of bad functions being rejected. Other than the "handcrafted" pseudo-test mentioned above, we also use the proof method of Kaufman and Lovett to find pseudo-tests distinguishing other pairs of good and bad functions. We then combine them using our calculus till we get a "pseudo-test" which does accept all the good functions, and rejects all the bad functions. At this stage we can now apply the previous works to get a tester for the family $\mathcal{F}$.

1.5 Organization of rest of the paper

In Section 3 we prove our main theorem after recalling in Section 2 the required tools from previous works. In Section 4 we use additive combinatorics to construct a "pseudo-test" for the most challenging case of separating x^d from x^e for $e = p^i d$ (see the discussion in the previous subsection). Section 5 generalizes the main theorem of [KL11] and constructs a "pseudo-test" for separating x^d from x^e for other e's of interest. In Section 6 we introduce our calculus for composing "pseudo-tests". Section 7, though not needed for obtaining our main result, is worth noting. It contains a useful simplification of the constraints used in the study of affine-invariant property testing.

2 Preliminaries

We start by recalling the notions of k-single-orbit characterizability, the degree set and the border set of an affine-invariant linear family and their role in the testing of these properties. All information presented in this section has already appeared in previous works [KS08, GKS08, GKS09, BS11, BGM+11a]. We follow the presentation in [BGM+11a, Sections 2, 3].

2.1 Establishing the k-single-orbit characterization property is sufficient for k-local testability

Our tester for sparse affine-invariant linear properties comes from a structural theorem which shows that every such property has a "single-orbit characterization". To describe this notion we need a couple of definitions.

Definition 2.1 (k-(basic)-constraint, k-characterization)**.** A *k-constraint* $C = (\overline{\alpha}, \{\overline{\lambda}_i\}_{i=1}^r)$ over $\mathbb{F}_{q^n}$ is given by a vector $\overline{\alpha} = (\alpha_1, \ldots, \alpha_k) \in \mathbb{F}_{q^n}^k$ together with r vectors $\overline{\lambda}_i = (\lambda_{i,1}, \ldots \lambda_{i,k}) \in \mathbb{F}_q^k$ for $1 \le i \le r$. We say that the constraint C *accepts* a function $f : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ if $\sum_{j=1}^k \lambda_{i,j} f(\alpha_j) = 0$ for all $1 \le i \le r$. Otherwise we say that C *rejects* f. We say a constraint is *basic* if $r = 1$.

Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a linear property. A *k-characterization* of $\mathcal{F}$ is a collection of k-constraints $C_1, \ldots, C_m$ such that $f \in \mathcal{F}$ if and only if C_j accepts f, for every $j \in \{1, \ldots, m\}$.

It is well-known [BHR05] that every k-locally testable linear property must have a k-characterization. In the case of affine-invariant linear properties some special characterizations are known to lead to k-testability. We describe these special characterizations next.

Definition 2.2 (k-single-orbit characterization (k-s-o-c))**.** Let $C = \left(\overline{\alpha}, \left\{\overline{\lambda}_i\right\}_{i=1}^{r}\right)$ be a k-constraint over $\mathbb{F}_{q^n}$. The *orbit* of C under the set of affine transformations is the following set of k-constraints

$$\{T \circ C\}_T = \left\{ \left((T(\alpha_1), \ldots, T(\alpha_k)), \left\{\overline{\lambda}_i\right\}_{i=1}^{r}\right) \mid T : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n} \text{ is an affine transformation} \right\}.$$

We say that C is a *k-single-orbit characterization* (k-s-o-c) of $\mathcal{F}$ if the orbit of C forms a k-characterization of $\mathcal{F}$.

We say that $\mathcal{F}$ has a basic k-s-o-c if the constraint C above is a basic one. (One of the simplifications proved in this work is that basic single-orbit characterizations are equivalent to general single-orbit characterizations, cf. Section 7.) A theorem due to Kaufman and Sudan [KS08] (see also [KS07a]) says that k-s-o-c implies local testability.

Theorem 2.3 (k-s-o-c implies local testability,[KS07a] Theorem 2.9.)**.** *Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be an affine-invariant linear property. If $\mathcal{F}$ has a k-single-orbit characterization, then $\mathcal{F}$ is also* poly(k)*-locally testable.*

2.2 Degree sets of affine-invariant linear properties

Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a linear affine-invariant property of functions. Note that every member of $\{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ can be written uniquely as a polynomial of degree at most $q^n - 1$ from $\mathbb{F}_{q^n}[x]$. Thus for a function $f : \mathbb{F}_{q^n} \to \mathbb{F}_q$ we define its *support*, denoted $\mathrm{supp}(f)$, to be the set of exponents in the support of the associated polynomial. I.e., $\mathrm{supp}(f) = \{d \in \{0, \ldots, q^n - 1\} | c_d \neq 0\}$ where $f(x) = \sum_d c_d x^d$. The *degree set* of $\mathcal{F}$ is simply the union of the supports of the functions in $\mathcal{F}$:

$$\mathrm{Deg}(\mathcal{F}) = \cup_{f \in \mathcal{F}} \mathrm{supp}(f).$$

Conversely, for a set of degrees $D \subseteq \{0, \ldots, q^n - 1\}$ let

$$\mathrm{Fam}_q(D) = \{f \mid f : \mathbb{F}_{q^n} \to \mathbb{F}_q,\ \mathrm{supp}(f) \subseteq D\}.$$

Affine-invariant linear properties are characterized in terms of their degree-sets (this is stated formally in the next lemma) and these degree-sets have a special structure — they are "closed" under "p-shadows" and "(q,n)-shifts" as explained next.

Define the *p-shadow* of an integer d to be the set of integers whose base-p representation is not larger, point-wise, than the base-p representation of d. More precisely, writing d in base p as $\sum_{i \geq 0} d_i p^i$ we define

$$\mathrm{Shadow}_p(d) = \left\{ \sum_{i \geq 0} e_i p^i \mid e_i \in \{0, 1, \ldots, d_i\}\, \forall i \geq 0 \right\}.$$

It is known from [KS08] that whenever $d \in \mathrm{Deg}(\mathcal{F})$ for some affine-invariant linear property $\mathcal{F}$ then $q \cdot d \mod q^n - 1$ also belongs to $\mathrm{Deg}(\mathcal{F})$. This motivates the following definition of the (q,n)-*shift* of an integer d as

$$\mathrm{Shift}_{q,n}(d) = \begin{cases} \{0\} & d = 0 \\ \left\{c \in \{1, \ldots, q^n - 1\} \mid c \equiv q^i \cdot d \mod q^n - 1 \text{ for some } 0 \leq i \leq n\right\} & 1 \leq d \leq q^n - 1 \end{cases}$$

The reason for treating 0 differently than $q^n - 1$ is that these two exponents induce somewhat different functions, namely $0^0 = 1$ but $0^{q^n-1} = 0$.

The p-shadow of a set of integers D is $\text{Shadow}_p(D) = \bigcup_{d \in D} \text{Shadow}_p(d)$. We say D is *p-shadow closed* if $D = \text{Shadow}_p(D)$. The (q,n)-shift of D is similarly defined and we say D is *(q,n)-shift closed* if $D = \text{Shift}_{q,n}(D)$.

The following lemma is [BGM+11a, Lemma 2.11]. It says that an affine-invariant linear property in $\{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is characterized by its degree-set, and this degree-set is p-shadow and (q,n)-shift closed.

Lemma 2.4 (Characterization of affine-invariant linear properties by degree-sets)**.** *let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be an affine-invariant linear property. Then $\text{Deg}(\mathcal{F})$ is (q,n)-shift-closed, p-shadow-closed, and $\mathcal{F} = \text{Fam}_q(\text{Deg}(\mathcal{F}))$. Conversely, suppose that D is a (q,n)-shift-closed and p-shadow-closed set of degrees. Then $\text{Fam}_q(D)$ is an affine-invariant linear property and $D = \text{Deg}(\text{Fam}_q(D))$.*

Remark 2.5 (The role of p and q in Lemma 2.4). We point out that the characteristic p of the field $\mathbb{F}_q$ and its size q play different roles in the lemma above. The shadow of an integer is with respect to base-p representations, whereas the shift of an integer is computed by taking q-multiples of it.

2.3 The border set of affine-invariant linear properties

The fact that the degree sets of affine-invariant linear properties are p-shadow closed motivates the following definition of the Border introduced in [BGM+11a]. This notion will play a central role in constructing our tester for sparse affine-invariant linear properties.

Definition 2.6 (Border)**.** The *border* of an affine-invariant linear property $\mathcal{F} \subset \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$, where q is a power of a prime p, is the set of degrees e that are "just outside" of $\text{Deg}(\mathcal{F})$, meaning that e is not in $\text{Deg}(\mathcal{F})$ but every element in the p-shadow of e is:

$$\text{Border}(\mathcal{F}) = \{e \in \{0,\ldots,q^n-1\} \mid e \notin \text{Deg}(\mathcal{F}) \text{ but } (\text{Shadow}_p(e) \setminus \{e\}) \subseteq \text{Deg}(\mathcal{F})\}\,.$$

In what follows, we say that a constraint C over $\mathbb{F}_{q^n}$ accepts the degree d if it accepts the function $f(x) = x^d$, otherwise we say that C rejects the degree d. For a set of degrees $D \subseteq \{0,1,\ldots,q^n-1\}$, we say that the constraint C accepts D if it accepts all degrees in D. In our proof of Theorem 1.1 we shall use the following equivalent definition of k-single-orbit characterization via the notion of the border.

Lemma 2.7 (Equivalent definition of k-single-orbit characterizable property via the border, [BGM+11b], Lemma 3.2.)**.** *Let $\mathcal{F}$ be an affine-invariant linear property, and let C be a k-constraint. Then C forms a k-single-orbit characterization of $\mathcal{F}$ if and only if C accepts all degrees in $\text{Deg}(\mathcal{F})$ and rejects all degrees in $\text{Border}(\mathcal{F})$.*

3 Proof of Main Theorem

In this section we prove our main theorem (Theorem 1.1) and along the way explain the main new ingredients and the need for them. Like all previous works on k-local testability of affine-invariant linear properties, our main theorem is obtained from showing the existence of the k-s-o-c property.

Theorem 3.1 (Sparse affine-invariant linear properties have a k-single-orbit characterization)**.** *For every q that is a power of a prime p and every integer t there exists an integer $k = k(t,q)$ such that the following holds. If $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is a t-sparse linear affine-invariant property then $\mathcal{F}$ has a k-single-orbit characterization.*

Proof of Main Theorem 1.1. Follows immediately from Theorem 3.1 and Theorem 2.3. □

3.1 Pseudo-tests suffice for local testability

Our single-orbit characterizations are obtained by introducing a notion that we call a "pseudo-test", which we define below.

Definition 3.2 (Pseudo-test)**.** For disjoint sets $D, B \subseteq \{0, \ldots, q^n - 1\}$, and a k-constraint $C = (\overline{\alpha}, \{\overline{\lambda}_i\}_{i=1}^r)$, we say that C is a *k-pseudo-test* separating D from B if C accepts all degrees in D and rejects all degrees in B. We say that C is a basic-pseudo-test if it is a basic-constraint. (When the arity k is clear from context we will often drop it, calling it simply a "pseudo-test".)

As such the pseudo-test above need not satisfy any semantic properties. While the test itself accepts every function in the span of the monomials $\{x^d | d \in D\}$ it clearly accepts a vast number of other functions (since it is a single deterministic test and hence accepts a subspace of dimension $q^n - r$). So it is far from being sound. Our intent is to use the orbit of the pseudo-test as the test, but then this orbit is now not complete! It may not accept x^d with probability 1, even for $d \in D$. Thus pseudo-tests seem to be completely irrelevant to the task at hand.

However as we note below in the next corollary, in some circumstances they do work well as tests. Furthermore, somewhat surprisingly it is possible to take two pseudo-tests each of which is incomplete, or unsound, and combine them to get something that is complete and sound. Indeed the value of the pseudo-tests are that they can be composed together nicely. Some of the basic steps are given later on in Propositions 6.2 and 6.4 and the resulting "broad" Composition Lemma 3.10 says that it is possible to construct ("nice") pseudo-tests "piecemeal" from (not so nice but) simpler pseudo-tests.

The relation between pseudo-tests and single-orbit characterizability is given by the following Corollary which is an immediate consequence of Lemma 2.7 and the definition of a pseudo-test.

Corollary 3.3 (Equivalent definition of k-single-orbit characterizable property via pseudo-tests)**.** *Let $\mathcal{F}$ be an affine-invariant linear property, and let C be a k-constraint. Then C forms a k-single-orbit characterization of $\mathcal{F}$ if and only if $\mathcal{F}$ is a pseudo-test separating* $\mathrm{Deg}(\mathcal{F})$ *from* $\mathrm{Border}(\mathcal{F})$.

When applying the above corollary it will be useful for us to use the following simple lemma which says that a constraint over $\mathbb{F}_{q^n}$ accepts a degree d if and only if it accepts all degrees in its (q, n)-shift.

Lemma 3.4. *Let d be a degree in $\{0, 1, \ldots, q^n - 1\}$, and let C be a k-constraint over $\mathbb{F}_{q^n}$. Then C accepts d if and only if it accepts all degrees in* $\mathrm{Shift}_{q,n}(d)$.

Proof. Let $d' \in \mathrm{Shift}_{q,n}(d)$ be such that $d' \equiv d \cdot q^\ell \mod q^n - 1$. Then for all $1 \leq i \leq r$ we have that

$$\left(\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^d\right)^{q^\ell} = \sum_{j=1}^{k} \lambda_{i,j}^{q^\ell} \alpha_j^{d \cdot q^\ell} = \sum_{j=1}^{k} \lambda_{i,j} \alpha_j^{d'},$$

where the first equality is due to the fact that raising to the power q^ℓ is a linear operation over $\mathbb{F}_{q^n}$, while the second equality is due to the fact that $\lambda_{i,j} \in \mathbb{F}_q$, and hence $\lambda_{i,j}^{q^\ell} = \lambda_{i,j}$. Thus we have that $\sum_{j=1}^k \lambda_{i,j} \alpha_j^d = 0$ if and only if $\sum_{j=1}^k \lambda_{i,j} \alpha_j^{d'} = 0$. □

Given the notion of pseudo-test we can state the main technical theorem whose proof occupies the rest of this paper. In what follows we say D' is a *(q, n)-shift representative set* for a (q, n)-shift closed set D if $\mathrm{Shift}_{q,n}(D') = D$.

Theorem 3.5 (Main Technical — Sparse affine-invariant linear properties have a k-pseudo-test)**.** *For every q that is a power of a prime p and every integer t there exists an integer $k = k(t,q)$ such that the following holds. Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a t-sparse affine-invariant linear property and let D', B' be (q,n)-representative sets of* $\mathrm{Deg}(\mathcal{F})$, $\mathrm{Border}(\mathcal{F})$ *respectively. Then there exists a k-pseudo-test that separates D' from B'.*

Proof of Theorem 3.1. Follows immediately from Theorem 3.5, Corollary 3.3 and Lemma 3.4. □

3.2 Overview of the proof of Main Technical Theorem 3.5

Fix (q,n)-shift representative sets D', B' for $\mathrm{Deg}(\mathcal{F})$ and $\mathrm{Border}(\mathcal{F})$ respectively. We construct a pseudo-test that separates D' from B' in three steps as follows.

1. Cover $D' \times B'$ by a constant number of product sets
$$D' \times B' = D'_0 \times B'_0 \cup \ldots \cup D'_\ell \times B'_\ell \tag{1}$$
where the constant ℓ depends only on t and q, and, crucially, is independent of n.
2. For each $i = 1, \ldots, \ell$ construct a k'-pseudo-test that separates D'_i from B'_i, where k' does not depend on n (it, too, depends on q and t).
3. Show that all ℓ of the k'-pseudo-tests can be "composed" to derive a single k-pseudo-test that separates D' from B' with $k = k(k', t, \ell)$. This separates D' from B' by a pseudo-test of size that depends only on q and t and is independent of n and thereby proves Theorem 3.5.

We now elaborate on each of the steps. The second step will be broken up into two sub-steps because there are two very different kinds of pair-sets that we need to consider, and each requires its own set of tools.

3.3 Covering the (q,n)-shift representative sets

First we define the cover of $D' \times B'$ by set-pairs and then bound the number of set-pairs in our cover in Lemma 3.7. (Inspection reveals that our cover is actually a partition of $D' \times B'$ but the rest of our proofs only need the weaker assumption of a cover.)

Definition 3.6 (Cover)**.** Given D', B' that are (q,n)-shift representative sets of $\mathrm{Deg}(\mathcal{F})$ and $\mathrm{Border}(\mathcal{F})$ respectively, where $q = p^s$ for a prime p, partition B' into
$$B_0 = B' \setminus \mathrm{Shift}_{p,sn}(D'); \quad B_1 = B' \cap \mathrm{Shift}_{p,sn}(D'). \tag{2}$$
Set $D'_0 = D'$ and $B'_0 = B_0$. Order the pairs in $D' \times B_1$ arbitrarily as $\{(d_1, b_1), \ldots, (d_\ell, b_\ell)\}$ where $\ell = |D'| \cdot |B_1|$ and let $D'_i = \{d_i\}$ and $B'_i = \{b_i\}$ for all $i = 1, \ldots, \ell$.

Notice that although elements of $\mathrm{Border}(\mathcal{F})$ do not belong to $\mathrm{Deg}(\mathcal{F})$ (cf. Definition 2.6), they can potentially belong to $\mathrm{Shift}_{p,sn}(\mathrm{Deg}(\mathcal{F}))$, so the set B_1 can indeed be nonempty.

Inspection reveals that the above set of pairs in Definition 3.6 covers $D' \times B'$. The following lemma bounds the number of pairs by bounding $|D'| \cdot |B_1|$. The second part of the lemma will be used soon and since its proof relies on the first part we find it convenient to include it here. To state the second part we define the *p-weight* $\mathrm{wt}_p(d)$ of an integer d as the sum of digits of the base-p representation of d. Formally, if $d = \sum_{i \geq 0} d_i p^i$ then $\mathrm{wt}_p(d) = \sum_{i \geq 0} d_i$.

Lemma 3.7 (t-sparse properties have sparse representative sets)**.** *Suppose that $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is a t-sparse affine-invariant linear property. Then the following holds:*

1. *There exist (q,n)-shift representative sets D', B' for $\mathrm{Deg}(\mathcal{F})$, $\mathrm{Border}(\mathcal{F})$ respectively such that $|D'| \leq 2t+1$, and assuming $q = p^s$ where p is a prime, $B_1 = \mathrm{Border}(\mathcal{F}) \cap \mathrm{Shift}_{p,sn}(D')$ is of size at most $s(2t+1)$.*

2. *All integers in $\mathrm{Deg}(\mathcal{F})$ have p-weight at most $2t$ and those of $\mathrm{Border}(\mathcal{F})$ have p-weight at most $2t+1$.*

Proof. Our starting point is Lemma 2.15 from [BGM+11a]. It says that if $\mathcal{F}$ is t-sparse then it has a (q,n)-shift representative set D' of size at most $2t+1$. The border could be potentially of much larger size but if we restrict our attention only to the elements that lie in $\mathrm{Shift}_{p,sn}(D')$, then they can be represented by a set B_1 of size at most $s|D'|$ because for each nonzero $d \in D'$ the (q,n)-shifts of $d, dp, \ldots, dp^{s-1}$ cover the (p,sn)-shift of d.

To prove the second part we claim that $\mathrm{Deg}(\mathcal{F})$ contains integers of p-weight at most $2t$. By definition, this will immediately imply (cf. Definition 2.6) that the p-weight of every element of $\mathrm{Border}(\mathcal{F})$ is at most $2t+1$. To see that $\mathrm{Deg}(\mathcal{F})$ cannot contain an integer of p-weight greater than $2t$ notice that Lemma 2.4 implies that if an integer of p-weight r belongs to $\mathrm{Deg}(\mathcal{F})$ then there are integers of p-weight r' in $\mathrm{Deg}(\mathcal{F})$ for every $r' = 0, 1, \ldots, r-1$. Since the (q,n)-shift of an integer d contains only integers of the same p-weight as d, this implies that $|D'| > r$. The assumption $|D'| \leq 2t+1$ therefore shows that no integer in $\mathrm{Deg}(\mathcal{F})$ has p-weight greater than $2t$ as claimed and this completes our proof. □

3.4 Separating a pair of sets with disjoint p-shifts

We now turn to the task of separating individual pairs of sets from our cover given in Definition 3.6. We start by showing a pseudo-test which separates a pair of sets D, B such that B does not contain any p-shift of a degree in D. This pseudo-test will be used for separating the sets D'_0 from B'_0 and in addition for separating all pairs (d_i, b_i) such that the degree b_i does not belong to a p-shift of the degree d_i. Our proof method uses the work of Kauffman and Lovett [KL11]. Stated using our language of pseudo-tests, they proved that for every t-sparse affine-invariant linear property $\mathcal{F} \subseteq \{\mathbb{F}_{p^n} \to \mathbb{F}_p\}$ over a *prime* field $\mathbb{F}_p$ there exists a $k(t)$-pseudo-test that separates D' from B', where D', B' are (p,n)-shift representative sets of $\mathrm{Deg}(\mathcal{F})$, $\mathrm{Border}(\mathcal{F})$ respectively. And by Corollary 3.3 and Lemma 3.4. this readily implies $\mathcal{F}$ is also $k(t)$-single-orbit characterizatable. We observe that the proof method of [KL11] actually gives the following more general pseudo-test.

Lemma 3.8 (Separation of distinct (p,n)-shifts)**.** *For every t, w and prime p there exists $k = k(t,w)$ such that the following holds for sufficiently large n: Let $D, B \subseteq \{0, \ldots, p^n - 1\}$ such that $|D| \leq t$, B does not contain any (p,n)-shift of a degree in D and in addition $\mathrm{wt}_p(d) \leq w$ for every degree $d \in D \cup B$. Then there exists a single (basic) k-pseudo-test C that separates D from B.*

We prove this lemma in Section 5. To see that the result of [KL11] is a special case of it note that if $\mathcal{F} \subseteq \{\mathbb{F}_{p^n} \to \mathbb{F}_p\}$ is a t-sparse affine-invariant linear property over a prime field $\mathbb{F}_p$ then Lemma 3.7 implies that $\mathrm{Deg}(\mathcal{F})$ has a (p,n)-shift representative set D' of size at most $2t+1$. Moreover, Part 2 of the same Lemma implies that if B' is a (p,n)-shift representative set of $\mathrm{Border}(\mathcal{F})$ then $\mathrm{wt}_p(d) \leq 2t+1$ for every $d \in D' \cup B'$. Finally, note that the fact that $\mathcal{F}$ is an affine-invariant linear property over $\mathbb{F}_p$ implies that it is (p,n)-shift-closed and hence B' does not contain any (p,n)-shift of a degree in D'.

3.5 Separating a pair of degrees in the same p-shift

Lemma 3.8 only gives a pseudo-test which separate pairs of degrees that belong to different (p,n)-shifts. As explained above, this suffices in order to prove single-orbit characterizability of affine-invariant linear properties over a prime field $\mathbb{F}_p$ since the degree sets of such properties are (p,n)-shift closed. However, in the case of non-prime fields of size $q=p^s$ affine-invariant linear properties are not necessarily (p,sn)-shift closed, and thus we need to be able to separate also pairs of degrees that belong to the same (p,sn)-shift. The following lemma covers this case.

Lemma 3.9 (Separation of two degrees in the same (p,sn)-shift)**.** *Let $q=p^s$ for a prime p, and let $d\in\{0,\ldots,q^n-1\}$, $b\in \mathrm{Shift}_{p,sn}(d)\setminus \mathrm{Shift}_{q,n}(d)$ be a pair of degrees of p-weight at most w. Then there exists a single (basic) k-pseudo-test C that separates $\{d\}$ from $\{b\}$ for $k=4\cdot 8^{4^{w+1}}$.*

As mentioned earlier this is the case where we have to design the pseudo-tests explicitly. We prove this lemma, using machinery that comes from the proofs of the sum-product theorem, in Section 4 .

3.6 A Calculus for composing pseudo-tests

So far we have managed to find a separating pseudo-test for each pair of sets in our cover of $D'\times B'$ given in Definition 3.6. In order to obtain a *single* pseudo-test that separates all of D' from all of B' and thereby prove Theorem 3.5 we introduce a natural calculus for composing pseudo-tests that separate distinct pairs of degree-sets. Suppose C_1 is a k_1-pseudo-test that separates D_1 from B_1 and C_2 is a k_2-pseudo-test that separates D_2 from B_2. One of the basic operations in our calculus takes the "union" of C_1 and C_2 and gives a (k_1+k_2)-pseudo-test that separates $D_1\cap D_2$ from $B_1\cup B_2$ (cf. Proposition 6.2). The second operation takes the "tensor" of C_1 and C_2 and gives a $(k_1\cdot k_2)$-pseudo-test which separates $D_1\cup D_2$ from $B_1\cap B_2$ (cf. Proposition 6.4). The combination of the two operations yields the following result that allows us to combine many different pseudo-tests into one.

Lemma 3.10 (Composition of pseudo-tests)**.** *For every k', t and ℓ, there exists $k=k(k',t,\ell)$ such that the following holds. Let $D,B\subseteq\{0,\ldots,q^n-1\}$ be disjoint sets with $|D|\leq t$ and let $D\times B=D_1\times B_1\cup\ldots\cup D_\ell\times B_\ell$ be a cover of $D\times B$. Suppose that for all $i=1,\ldots,\ell$ there exists a k'-pseudo-test C_i which separates D_i from B_i. Then there exists a k-pseudo-test C that separates D from B.*

The proof of the lemma, along with a detailed description of the calculus of "unions" and "tensors" that underlie it, appears in Section 6.

3.7 Completing the proof of Theorem 3.5

We are now in the position to complete the formal proof of the main technical theorem.

Proof of Main Technical Theorem 3.5. Let $q=p^s$ for a prime p and let $\mathcal{F}\subseteq\{\mathbb{F}_{q^n}\to\mathbb{F}_q\}$ be a t-sparse affine-invariant linear property. Let B',D' be the (q,n)-shift representative sets for $\mathrm{Deg}(\mathcal{F}),\mathrm{Border}(\mathcal{F})$ respectively guaranteed by Lemma 3.7. By the lemma we have $|D'|\leq 2t+1$ and $|B'\cap\mathrm{Shift}_{p,sn}(D')|\leq s(2t+1)$ and all integers in D' have p-weight at most $2t$ and those of B' have p-weight at most $2t+1$. Cover $D'\times B'$ as in Definition 3.6 and notice the number of sets in this cover is bounded by $1+|D'|\cdot|B_1|\leq s(2t+1)^2+1$.

Apply Lemma 3.8 to conclude that D'_0 can be separated from B'_0 by a k_1-pseudo-test C_0 for k_1 that depends only on t. Apply Lemma 3.8 also to each pair $D'_i\times B'_i, i=1,\ldots,\ell$ which satisfy

$\text{Shift}_{p,sn}(d_i) \cap \text{Shift}_{p,sn}(b_i) = \emptyset$ to obtain a k_1-pseudo-test C_i that separates D'_i from B'_i. Finally, apply Lemma 3.9 to each pair $D'_i \times B'_i, i = 1, \ldots, \ell$ that satisfy $\text{Shift}_{p,sn}(d_i) \cap \text{Shift}_{p,sn}(b_i) \neq \emptyset$ to obtain a k_2-pseudo-test C_i that separates D'_i from B'_i where k_2 depends only on t.

Apply the composition Lemma 3.10 to $C_0, \ldots, C_\ell$ and, recalling ℓ is bounded by a polynomial in s, t and $|D'| \leq 2t + 1$ we conclude the existence of a k-pseudo-test C that separates D' from B' where k depends only on t and q. This completes the proof of Theorem 3.5. □

4 Separating pairs of degrees in the same p-shift — Proof of Lemma 3.9

In this section we prove Lemma 3.9, showing for every pair of degrees $d \in \{0, \ldots, q^n - 1\}$, $b \in \text{Shift}_{p,sn}(d) \setminus \text{Shift}_{q,n}(d)$ of constant p-weight the existence of a k-constraint which separates d from b.

The proof idea of Lemma 3.9 is the following. Suppose we wish to find a constraint which separates the degree d from the degree $d \cdot p^i \in \text{Shift}_{p,sn}(d) \setminus \text{Shift}_{q,n}(d)$. Let $A = \left\{\alpha^d | \alpha \in \mathbb{F}_{q^n}\right\}$, and assume for simplicity that $A \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i}) \neq \emptyset$. Then in this case there exists a constraint $C = (\overline{\alpha}, \overline{\lambda})$ of arity 2 which separates d from $d \cdot p^i$: Let $\gamma^d \in A \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i})$ and $\overline{\alpha} = (\alpha_1, \alpha_2) = (1, \gamma) \in \mathbb{F}_{q^n}^2$, and let $\overline{\lambda} = (\lambda_1, \lambda_2) = (-1, \gamma^{-d}) \in \mathbb{F}_q^2$ (the fact that $\gamma^{-d} \triangleq 1/\gamma^d$ is in $\mathbb{F}_q$ follows from our assumption that $\gamma^d \in \mathbb{F}_q$). Then for the degree d we have that

$$\lambda_1 \alpha_1^d + \lambda_2 \alpha_2^d = -1 + \gamma^{-d} \gamma^d = 0, \tag{3}$$

and hence C accepts d. On the other hand, for the degree $d \cdot p^i$ we have that

$$\lambda_1 \alpha_1^{d \cdot p^i} + \lambda_2 \alpha_2^{d \cdot p^i} = -1 + \gamma^{-d} \gamma^{d \cdot p^i}. \tag{4}$$

Note that (4) equals zero if and only if $(\gamma^d)^{p^i} = \gamma^d$. But by assumption $\gamma^d \notin \mathbb{F}_{p^i}$, and hence $(\gamma^d)^{p^i} \neq \gamma^d$ which implies in turn that (4) is non-zero.

However, our assumption that $A \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i}) \neq \emptyset$ was too optimistic. To resolve this we resort to the closure $\mathbb{F}(A)$ of A in $\mathbb{F}_{q^n}$, defined as the smallest subfield of $\mathbb{F}_{q^n}$ containing A. Note that $\mathbb{F}_p \subseteq \mathbb{F}(A) \subseteq \mathbb{F}_{q^n}$. We first prove (in Lemma 4.1) that $\mathbb{F}(A) \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i}) \neq \emptyset$. Then in Lemma 4.2 we prove, using machinery developed for the proof of a version of the sum-product theorem from [BIW06], that if $d \leq q^{(1-\epsilon)n}$ then every element $\gamma \in \mathbb{F}(A)$ can be written as $\gamma = \frac{\gamma_1}{\gamma_2}$ where both γ_1 and γ_2 are the sum of a constant number of elements in A (this constant depends only on ϵ). This gives in turn the desired constraint $C = (\overline{\alpha}, \overline{\lambda})$ which separates d from $d \cdot p^i$.

We start by claiming that $\mathbb{F}(A)$ contains an element in $\mathbb{F}_q \setminus \mathbb{F}_{p^i}$.

Lemma 4.1. *Let $q = p^s$ for some prime p, and let $d \in \{0, \ldots, q^n - 1\}$ be such that $d \cdot p^i \notin \text{Shift}_{q,n}(d)$. Let $A = \left\{\alpha^d | \alpha \in \mathbb{F}_{q^n}\right\}$, and let $\mathbb{F}(A)$ be the smallest subfield of $\mathbb{F}_{q^n}$ containing A. Then $\mathbb{F}(A) \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i}) \neq \emptyset$.*

Proof. Let $\mathbb{F}(A) = \mathbb{F}_{p^m}$, and suppose by way of contradiction that $\mathbb{F}_{p^m} \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i}) = \emptyset$. Then $\mathbb{F}_{p^m} \cap \mathbb{F}_q \subseteq \mathbb{F}_{p^i}$. In order to arrive at a contradiction, we will show that $d \cdot p^i$ is a (q, n)-shift of d contradicting our assumption.

Let $r = \gcd(s, m)$. Then we have $\mathbb{F}_{p^m} \cap \mathbb{F}_q = \mathbb{F}_{p^r} \subseteq \mathbb{F}_{p^i}$ and hence r divides i. Our first observation is that since $r = \gcd(s, m)$ there exists a pair of integers t, ℓ such that $tm + \ell s = r$. Let $t' = \frac{it}{r}$, $\ell' = \frac{i\ell}{r}$. Since r divides i we have that t', ℓ' are integers and

$$t'm + \ell' s = \frac{it}{r} m + \frac{i\ell}{r} s = \frac{i}{r}(tm + \ell s) = i. \tag{5}$$

Our second observation is that since $\mathbb{F}(A) = \mathbb{F}_{p^m}$ then for every $\alpha \in \mathbb{F}_{q^n}$ we have that $(\alpha^d)^{p^m} = \alpha^d$ and hence the polynomial $x^{d \cdot p^m} - x^d$ is identically zero over $\mathbb{F}_{q^n}$. This implies in turn that

$$d \cdot p^m \equiv d \mod q^n - 1. \tag{6}$$

From (5) and (6) we have

$$\begin{aligned} d \cdot p^i &\equiv d \cdot p^{t'm+\ell' s} \mod q^n - 1 \text{ (From (5))} \\ &\equiv d \cdot p^{\ell' s} \mod q^n - 1 \text{ (From (6))} \\ &\equiv d \cdot q^{\ell'} \mod q^n - 1 \text{ (Since } q = p^s) \end{aligned}$$

Thus we have that $d \cdot p^i$ is a (q, n)-shift of d — a contradiction. □

We now prove that if d is not too large then every element $\gamma \in \mathbb{F}(A)$ can be written in the form $\gamma = \frac{\gamma_1}{\gamma_2}$ where both γ_1 and γ_2 are sums of a constant number of elements in A.

Lemma 4.2. *Let $d \in \{0, 1, \ldots, q^n - 1\}$ be a degree which satisfies $d \leq q^{(1-\epsilon)n}$, let $A = \{\alpha^d | \alpha \in \mathbb{F}_{q^n}\}$, and let $\mathbb{F}(A)$ be the smallest subfield of $\mathbb{F}_{q^n}$ containing A. Then $\mathbb{F}(A) \subseteq \frac{\ell A - \ell A}{\ell A - \ell A}$ for $\ell = 8^{4^{\lceil 1/\epsilon \rceil}}$.*

Using [CC11, Theorem 1.2] the bound on ℓ above can be improved to exponential in $1/\epsilon$, i.e., $\ell = 2^{O(1/\epsilon)}$. For the sake of presenting a simple self-contained proof, we prove the above lemma using the following theorem from [BIW06] which was proved there as a step towards a simplified version of the sum-product theorem of [BKT04]. For a set A and $\odot$ an arithmetic operation in $\{+, -, \div, \times\}$ let $A \odot A = \{a \odot a' | a, a' \in A\}$.

Theorem 4.3 ([BIW06], Claim A.4.)**.** *Let $\mathbb{F}$ be a finite field, and let $A \subseteq \mathbb{F}$ and $k \in \mathbb{N}$ (with $k \geq 2$) be such that $|\mathbb{F}|^{1/k} < |A| \leq |\mathbb{F}|^{1/(k-1)}$. Then $|\frac{A-A}{A-A}| \geq |\mathbb{F}|^{1/(k-1)}$.*

Proof of Lemma 4.2. Apply Theorem 4.3 iteratively. Set $A_0 := A$ and for $i = 1, 2, 3, \ldots$ let $A'_i = \frac{A_{i-1} - A_{i-1}}{A_{i-1} - A_{i-1}}$, and $A_i = A'_i + A'_i \cdot A'_i$. The proof consists of two main steps. In the first step we will argue that there exists $t \leq \lceil 1/\epsilon \rceil + 1$ for which $\mathbb{F}(A) \subseteq A_t$. In the second step we will prove by induction on i that $A_i \subseteq \frac{8^{4^{i-1}}A - 8^{4^{i-1}}A}{8^{4^{i-1}}A - 8^{4^{i-1}}A}$ for every $i = 1, 2, 3, \ldots$. Lemma 4.2 follows from the $i = t$ case.

We start by showing the existence of $t \leq \lceil 1/\epsilon \rceil + 1$ for which $\mathbb{F}(A) \subseteq A_t$. To see this note first that since $d \leq q^{(1-\epsilon)n}$, for every $\beta \in \mathbb{F}_{q^n}$ there are at most $q^{(1-\epsilon)n}$ solutions in x to the equation $x^d = \beta$. Thus $|A_0| \geq \frac{q^n}{q^{(1-\epsilon)n}} = q^{\epsilon n}$. Choose $k = \lceil \frac{1}{\epsilon} \rceil + 1$, and note that $\epsilon > \frac{1}{k}$. Theorem 4.3 implies that $|A'_1| = |\frac{A_0 - A_0}{A_0 - A_0}| \geq (q^n)^{1/(k-1)}$. If A'_1 is a field then we are done since it can be verified that $A \subseteq A'_1$ (since $0 \in A$) and hence $\mathbb{F}(A) \subseteq A'_1$ from the minimality of $\mathbb{F}(A)$. Otherwise we have that $|A_1| = |A'_1 + A'_1 \cdot A'_1|$ is strictly greater than $(q^n)^{1/(k-1)}$, and thus we can apply Theorem 4.3 again to the set A_1. Continuing this process iteratively we have that at the i-th step either A'_i is a field and hence $\mathbb{F}(A) \subseteq A'_i \subseteq A_i$ or that $|A_i| > |A'_i| \geq (q^n)^{1/(k-i)}$. Since $A_i \subseteq \mathbb{F}_{q^n}$ for all i, this process must terminate after at most $k = \lceil \frac{1}{\epsilon} \rceil + 1$ steps, and thus we have that $\mathbb{F}(A) \subseteq A_t$ for $t \leq \lceil \frac{1}{\epsilon} \rceil + 1$.

Next we show by induction on i that $A_i \subseteq \frac{8^{4^{i-1}}A - 8^{4^{i-1}}A}{8^{4^{i-1}}A - 8^{4^{i-1}}A}$ for every $i = 1, 2, 3, \ldots$. This will imply in turn that

$$\mathbb{F}(A) \subseteq A_t \subseteq \frac{8^{4^{\lceil 1/\epsilon \rceil}}A - 8^{4^{\lceil 1/\epsilon \rceil}}A}{8^{4^{\lceil 1/\epsilon \rceil}}A - 8^{4^{\lceil 1/\epsilon \rceil}}A}$$

Base case — $i = 1$. Noting that A is closed under multiplication, in this case we have that

$$A_1 = A'_1 + A'_1 \cdot A'_1 = \frac{A-A}{A-A} + \frac{A-A}{A-A} \cdot \frac{A-A}{A-A} \subseteq \frac{A-A}{A-A} + \frac{2A-2A}{2A-2A} \subseteq \frac{8A-8A}{4A-4A} \subseteq \frac{8A-8A}{8A-8A}$$

Induction step. Suppose that the claim holds for index i and we will prove that it holds for index $i+1$ as well.

$$\begin{aligned}
A_{i+1} &= A'_{i+1} + A'_{i+1} \cdot A'_{i+1} = \frac{A_i - A_i}{A_i - A_i} + \frac{A_i - A_i}{A_i - A_i} \cdot \frac{A_i - A_i}{A_i - A_i} \\
&\subseteq \frac{8^{4^{i-1}}A - 8^{4^{i-1}}A}{8^{4^{i-1}}A - 8^{4^{i-1}}A} + \frac{8^{4^{i-1}}A - 8^{4^{i-1}}A}{8^{4^{i-1}}A - 8^{4^{i-1}}A} \cdot \frac{8^{4^{i-1}}A - 8^{4^{i-1}}A}{8^{4^{i-1}}A - 8^{4^{i-1}}A} \text{ (Induction hypothesis)} \\
&\subseteq \frac{8^{4^{i-1}}A - 8^{4^{i-1}}A}{8^{4^{i-1}}A - 8^{4^{i-1}}A} + \frac{2 \cdot 8^{2 \cdot 4^{i-1}}A - 2 \cdot 8^{2 \cdot 4^{i-1}}A}{2 \cdot 8^{2 \cdot 4^{i-1}}A - 2 \cdot 8^{2 \cdot 4^{i-1}}A} \\
&\subseteq \frac{8 \cdot 8^{3 \cdot 4^{i-1}}A - 8 \cdot 8^{3 \cdot 4^{i-1}}A}{4 \cdot 8^{3 \cdot 4^{i-1}}A - 4 \cdot 8^{3 \cdot 4^{i-1}}A} \\
&\subseteq \frac{8^{4^i}A - 8^{4^i}A}{8^{4^i}A - 8^{4^i}A}
\end{aligned}$$

□

In our proof of Lemma 3.9 we would like to apply Lemma 4.2 to the degree d. In order to apply this lemma we need d to be small. However, all we know about d is that it has small p-weight and this does not guarantee that d is small. In order to deal with this we shall first prove that since d has a small p-weight, it has a degree d' in its (q,n)-shift that is small. We will then show a constraint over $\mathbb{F}_{q^n}$ which separates d' from $d' \cdot p^i$. From Lemma 3.4 this will also imply that the constraint C separates d from $d \cdot p^i$. The following lemma says that every degree of small q-weight has a degree in its (q,n)-shift that is small.

Lemma 4.4. *For every degree $d \in \{0, 1, \ldots, q^n - 1\}$ there exists a degree $d' \in \mathrm{Shift}_{q,n}(d)$ such that $d' \leq q^{(1-1/\mathrm{wt}_q(d))n+1}$*

Proof. Let $t = \mathrm{wt}_q(d)$, and let $d = \sum_{i=0}^{n-1} d_i q^i$ be the representation of d in base-q. Since $\mathrm{wt}_q(d) \leq t$ the pigeonhole principle implies that d has at least $\frac{n-t}{t} = \frac{n}{t} - 1$ consecutive digits $d_j, d_{j+1 \mod n}, \ldots, d_{j+\frac{n}{t}-2 \mod n}$ which equal zero. Let $d' \in \mathrm{Shift}_{q,n}(d)$ be such that $d' \equiv d \cdot q^{n+1-n/t-j} \mod q^n - 1$. Then d' satisfies that all indices $d'_{n-\frac{n}{t}+1}, \ldots, d'_{n-1}$ equal zero, where $d' = \sum_{i=0}^{n-1} d'_i q^i$ is the representation of d' in base-q. But this implies in turn that $d' \leq \sum_{i=0}^{n(1-1/t)} (q-1) q^i \leq q^{n(1-1/t)+1}$. □

We now proceed to the proof of Lemma 3.9

Proof of Lemma 3.9. Suppose that $b \in \mathrm{Shift}_{q,n}(d)$ such that $b \equiv d \cdot p^i \mod q^n - 1$. Since $\mathrm{wt}_q(d) \leq \mathrm{wt}_p(d) \leq w$, from Lemma 4.4 we have that there exists a degree $d' \in \mathrm{Shift}_{q,n}(d)$ such that $d' \leq q^{(1-1/w)n+1} \leq q^{(1-1/(w+1))n}$ for sufficiently large n. Let $b' \in \mathrm{Shift}_{q,n}(d')$ such that $b' \equiv d' \cdot p^i \mod q^n - 1$ and note that $b' \in \mathrm{Shift}_{p,sn}(d') \setminus \mathrm{Shift}_{q,n}(d')$. From Lemma 3.4 it suffices to show a k-constraint which separates d' from b'.

Let $A = \left\{\alpha^{d'} | \alpha \in \mathbb{F}_{q^n}\right\}$, and let $\mathbb{F}(A)$ be the smallest subfield of $\mathbb{F}_{q^n}$ containing A. From Lemma 4.1 we have that $\mathbb{F}(A) \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i}) \neq \emptyset$, let $\gamma \in \mathbb{F}(A) \cap (\mathbb{F}_q \setminus \mathbb{F}_{p^i})$. From Lemma 4.2 and since

$d' \leq q^{(1-1/(w+1))n}$ we have that $\mathbb{F}(A) \subseteq \frac{\ell A - \ell A}{\ell A - \ell A}$ for $\ell = 8^{4^{w+1}}$, and thus $\gamma \in \frac{\ell A - \ell A}{\ell A - \ell A}$. In particular there exist $\beta_1, \ldots, \beta_{4\ell} \in \mathbb{F}_{q^n}$ such that

$$\gamma = \frac{(\beta_1^{d'} + \ldots + \beta_\ell^{d'}) - (\beta_{\ell+1}^{d'} + \ldots + \beta_{2\ell}^{d'})}{(\beta_{2\ell+1}^{d'} + \ldots + \beta_{3\ell}^{d'}) - (\beta_{3\ell+1}^{d'} + \ldots + \beta_{4\ell}^{d'})} \tag{7}$$

The constraint $C = (\overline{\alpha}, \overline{\lambda})$ will be the (4ℓ)-constraint defined by $\overline{\alpha} = (\alpha_1, \ldots, \alpha_{4\ell}) \in \mathbb{F}_{q^n}^{4\ell}$ and $\overline{\lambda} = (\lambda_1, \ldots \lambda_{4\ell}) \in \mathbb{F}_q^{4\ell}$, where $\alpha_i = \beta_i$ for all $1 \leq i \leq 4\ell$, and

$$\lambda_i = \begin{cases} -1 & 1 \leq i \leq \ell \\ 1 & \ell+1 \leq i \leq 2\ell \\ \gamma & 2\ell+1 \leq i \leq 3\ell \\ -\gamma & 3\ell+1 \leq i \leq 4\ell \end{cases}.$$

It remains to show that the constraint C accepts d' and rejects b'. For the degree d' we have from (7) that

$$\sum_{i=1}^{4\ell} \lambda_i \alpha_i^{d'} = -\left(\sum_{i=1}^{\ell} \beta_i^{d'} - \sum_{i=\ell+1}^{2\ell} \beta_i^{d'}\right) + \gamma\left(\sum_{i=2\ell+1}^{3\ell} \beta_i^{d'} - \sum_{i=3\ell+1}^{4\ell} \beta_i^{d'}\right) = 0$$

On the other hand, for the degree b' we have

$$\begin{aligned}
\sum_{i=1}^{4\ell} \lambda_i \alpha_i^{b'} &= \sum_{i=1}^{4\ell} \lambda_i \alpha_i^{d' \cdot p^i} \text{ (Since } b' \in \mathrm{Shift}_{q,n}(d') \text{ such that } b' \equiv d' \cdot p^i \mod q^n - 1) \\
&= -\left(\sum_{i=1}^{\ell} \beta_i^{d' \cdot p^i} - \sum_{i=\ell+1}^{2\ell} \beta_i^{d' \cdot p^i}\right) + \gamma\left(\sum_{i=2\ell+1}^{3\ell} \beta_i^{d' \cdot p^i} - \sum_{i=3\ell+1}^{4\ell} \beta_i^{d' \cdot p^i}\right) \\
&= -\left(\sum_{i=1}^{\ell} \beta_i^{d'} - \sum_{i=\ell+1}^{2\ell} \beta_i^{d'}\right)^{p^i} + \gamma\left(\sum_{i=2\ell+1}^{3\ell} \beta_i^{d'} - \sum_{i=3\ell+1}^{4\ell} \beta_i^{d'}\right)^{p^i} \text{ (Since the mapping } y \mapsto y^{p^i} \text{ is linear)} \\
&= \left(\sum_{i=1}^{\ell} \beta_i^{d'} - \sum_{i=\ell+1}^{2\ell} \beta_i^{d'}\right)^{p^i} (-1 + \gamma \cdot \gamma^{-p^i}) \text{ (From (7))}
\end{aligned}$$

To see that the above equation is non-zero note that $\gamma \notin \mathbb{F}_{p^i}$ and hence $\gamma^{p^i} \neq \gamma$. This implies in turn that $-1+\gamma \cdot \gamma^{-p^i} \neq 0$. Also, since $\gamma \neq 0$, from (7) we have that $\left(\sum_{i=1}^{\ell} \beta_i^{d'} - \sum_{i=\ell+1}^{2\ell} \beta_i^{d'}\right)^{p^i} \neq 0$. Hence the above equation is non-zero which concludes the proof of the lemma. □

5 Separating a pair of sets with disjoint p-shifts — Proof of Lemma 3.8

In this section we prove Lemma 3.8 which shows the existence of a k-constraint which separates degree sets with disjoint (p,n)-shifts. We will prove Lemma 3.8 using probabilistic arguments following the proof method of [KL11]. More precisely, we will show that if we choose $\overline{\lambda} \in (\mathbb{F}_p^*)^k$, $\overline{\alpha} \in (\mathbb{F}_{p^n})^k$ uniformly and independently at random for sufficiently large k, then with sufficiently high probability the constraint $C = (\overline{\alpha}, \overline{\lambda})$ will accept the set D and will reject the set B. In order to prove this we first compute bounds on the probability that such a random constraint accepts all degrees in an arbitrary degree set $D \subseteq \{0, 1, \ldots, p^n - 1\}$.

Lemma 5.1. *Let p be a prime, and let $D \subseteq \{0, 1, \ldots, p^n - 1\}$ be such that D has a (p, n)-shift representative set of size t, and $\mathrm{wt}_p(d) \leq w$ for every $d \in D$. Let $\overline{\lambda} \in (\mathbb{F}_p^*)^k$, $\overline{\alpha} \in (\mathbb{F}_{p^n})^k$ be chosen uniformly and independently at random. Then if $0 \notin D$,*

$$\left| \Pr_{\overline{\lambda} \in (\mathbb{F}_p^*)^k,\, \overline{\alpha} \in (\mathbb{F}_{p^n})^k} [C = (\overline{\alpha}, \overline{\lambda}) \text{ accepts } D] - p^{-|\mathrm{Shift}_{p,n}(D)|} \right| \leq |\mathbb{F}_{p^n}|^{-k/(2w^2 2^w t)},$$

while if $0 \in D$,

$$\left| \Pr_{\overline{\lambda} \in (\mathbb{F}_p^*)^k,\, \overline{\alpha} \in (\mathbb{F}_{p^n})^k} [C = (\overline{\alpha}, \overline{\lambda}) \text{ accepts } D] - \left(1 + (-1)^k \left(\frac{1}{p-1}\right)^{k-1}\right) p^{-|\mathrm{Shift}_{p,n}(D)|} \right| \leq |\mathbb{F}_{p^n}|^{-k/(2w^2 2^w t)}.$$

The proof of the above lemma appears in Section 5.1. Before proving this lemma, we present the proof of Lemma 3.8 based on it.

Proof of Lemma 3.8. Suppose first that $0 \notin B$, we will deal with the case in which $0 \in B$ later. Choose $k = 2w^2 2^w (t+1) + 2$ (note that k is even), and let $\overline{\lambda} \in (\mathbb{F}_p^*)^k$, $\overline{\alpha} \in (\mathbb{F}_{p^n})^k$ be chosen uniformly and independently at random. Denote by $P(D, B)$ the probability that the random constraint $C = (\overline{\alpha}, \overline{\lambda})$ accepts all degrees in D and rejects all degrees in B. Our goal will be to show that $P(D, B)$ is at least $(p-1)^{-k} p^{-nk}$, the probability assigned under the uniform distribution to a fixed pair of vectors $\overline{\lambda} \in (\mathbb{F}_p^*)^k$, $\overline{\alpha} \in (\mathbb{F}_{p^n})^k$. This will imply the existence of a k-constraint $C = (\overline{\alpha}, \overline{\lambda})$ which accepts all degrees in D and rejects all degrees in B. We compute a lower bound on $P(D, B)$ using Lemma 5.1.

$$\begin{aligned}
P(D,B) &\geq \Pr_{\overline{\lambda} \in (\mathbb{F}_p^*)^k,\, \overline{\alpha} \in (\mathbb{F}_{p^n})^k} [C = (\overline{\alpha}, \overline{\lambda}) \text{ accepts } D] - \sum_{b \in B} \Pr_{\overline{\lambda} \in (\mathbb{F}_p^*)^k,\, \overline{\alpha} \in (\mathbb{F}_{p^n})^k} [C = (\overline{\alpha}, \overline{\lambda}) \text{ accepts } D \cup \{b\}\,] \\
&\geq p^{-|\mathrm{Shift}_{p,n}(D)|} - p^{-nk/(2w^2 2^w t)} - \\
&\quad \sum_{b \in B} \left(p^{-|\mathrm{Shift}_{p,n}(D \cup \{b\})|} + p^{-nk/(2w^2 2^w (t+1))} \right) \quad \text{(from Lemma 5.1 and since } k \text{ is even)} \\
&= p^{-|\mathrm{Shift}_{p,n}(D)|} - p^{-nk/(2w^2 2^w t)} - \\
&\quad \sum_{b \in B} \left(p^{-(|\mathrm{Shift}_{p,n}(D)| + |\mathrm{Shift}_{p,n}(b)|)} + p^{-nk/(2w^2 2^w (t+1))} \right) \quad (b \notin \mathrm{Shift}_{p,n}(D)) \\
&\geq p^{-|\mathrm{Shift}_{p,n}(D)|} \left(1 - \sum_{b \in B} p^{-|\mathrm{Shift}_{p,n}(b)|}\right) - (|B| + 1) p^{-nk/(2w^2 2^w (t+1))} \qquad (8)
\end{aligned}$$

In order to bound the above expression we bound the sizes of $\mathrm{Shift}_{p,n}(D)$, $\mathrm{Shift}_{p,n}(b)$ and B. The size of $\mathrm{Shift}_{p,n}(D)$ can be bounded easily noting that $|D| \leq t$ and $|\mathrm{Shift}_{p,n}(d)| \leq n$ for every $d \in D$ which yields

$$|\mathrm{Shift}_{p,n}(D)| \leq tn. \qquad (9)$$

Next we bound $\mathrm{Shift}_{p,n}(b)$ from below for $b \in B$. Let $b = \sum_{i=0}^{n-1} b_i p^i$ be the base-p representation of b. From our assumption that $0 \notin B$ we have that $b \neq 0$ and hence $b_i \neq 0$ for some $0 \leq i \leq n-1$. Suppose that $|\mathrm{Shift}_{p,n}(b)| = r$, note that this implies that r divides n. But this implies that $p^{jr} \cdot b \equiv b \mod p^n - 1$ for every integer $0 \leq j \leq \frac{n}{r}$ which implies in turn that $b_i \equiv b_{i+jr} \mod n$ for every integer $0 \leq j \leq \frac{n}{r}$. Thus we have that $\mathrm{wt}_p(b) \geq \frac{n}{r}$ which gives

$$|\mathrm{Shift}_{p,n}(b)| \geq \frac{n}{w}. \qquad (10)$$

Finally, we bound the size of B. Since all degrees in B have p-weight at most w, we can bound the size of B for sufficiently large n by

$$|B| \leq \sum_{i=0}^{w} \binom{n+w-1}{w-1} \leq (w+1) \cdot (n+w-1)^{w-1}. \tag{11}$$

Plugging Equations (9), (10) and (11) into (8) we obtain

$$\begin{aligned} P(D,B) &\geq p^{-tn}\big(1-(w+1)\cdot(n+w-1)^{w-1}\cdot p^{-n/w}\big) - \big((w+1)\cdot(n+w-1)^{w-1}+1\big)p^{-nk/(2w^2 2^w(t+1))} \\ &\geq \frac{1}{2}p^{-tn} \quad \text{(due to our choice of } k = 2w^2 2^w(t+1)+2 \text{ and for sufficiently large } n) \\ &\geq (p-1)^{-k}p^{-nk} \end{aligned}$$

It remains to deal with the case in which $0 \in B$. The same calculations as above show the existence of a k-constraint C which is a pseudo-test separating D from $B \setminus \{0\}$. Note also that the 1-constraint $C' = (\alpha, \lambda)$ defined by $\alpha = 0$, $\lambda = 1$ rejects the degree 0 and accepts all degrees in D and thus forms a pseudo-test separating D from $\{0\}$. Thus the union constraint $C'' = C' \cup C$ (cf. Definition 6.1) forms a $(k+1)$-constraint which is a pseudo-test separating D from B.

□

5.1 Proof of Lemma 5.1

In order to prove this lemma we first show that the task of computing the probability that a random constraint satisfies a degree set D can be reduced to the task of computing the expected bias of the trace of sparse polynomials supported on degrees in D. We will then use a special bound on the distribution of the image of sparse polynomials from [KL11] in order to compute this latter expectation.

Recall that the trace operator over $\mathbb{F}_{q^n}$ is the function $\mathrm{Trace}_{q^n \to q} : \mathbb{F}_{q^n} \to \mathbb{F}_q$ defined as $\mathrm{Trace}_{q^n \to q}(x) = \sum_{i=0}^{n-1} x^{q^i}$. The following are well-known facts regarding the trace function that we shall use for the proof of Lemma 5.1.

Fact 5.2. *The trace operator is F_q-linear, i.e. for $\alpha, \beta \in \mathbb{F}_{q^n}$ and $\gamma \in \mathbb{F}_q$, $\mathrm{Trace}_{q^n \to q}(\alpha+\beta) = \mathrm{Trace}_{q^n \to q}(\alpha) + \mathrm{Trace}_{q^n \to q}(\beta)$ and $\mathrm{Trace}_{q^n \to q}(\gamma\alpha) = \gamma \mathrm{Trace}_{q^n \to q}(\alpha)$. Moreover, it is a q^{n-1}-to-1 map, i.e., for every $\alpha \in \mathbb{F}_q$, $|\mathrm{Trace}^{-1}_{q^n \to q}(\alpha)| = q^{n-1}$.*

Fact 5.3 (Trace of linear functions is unbiased)**.** *Let p be a prime. Then for every $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}_{p^n}$ not all zero we have that*

$$\mathbb{E}_{x_1, x_2, \ldots, x_n \in \mathbb{F}_{p^n}}\left[\omega_p^{\mathrm{Trace}_{p^n \to p}(\sum_{i=1}^n \alpha_i x_i)}\right] = 0,$$

where $\omega_p = e^{2\pi i/p}$ is the complex p-root of unity.

For a degree set $D \subseteq \{0, 1, \ldots, p^n - 1\}$ and a vector $\overline{\beta} = (\beta_d)_{d \in D} \in (\mathbb{F}_{p^n})^{|D|}$ let $f_{\overline{\beta}}(x) = \sum_{d \in D} \beta_d x^d$. For a function $g : \mathbb{F}_{p^n} \to \mathbb{F}_p$, let $\mathrm{Bias}_p(g) = \frac{\Pr_{\alpha \in \mathbb{F}_{p^n}}[g(\alpha)=0] - 1/p}{1-1/p}$.

Lemma 5.4. *Let p be a prime, and let $D \subseteq \{0, 1, \ldots, p^n - 1\}$. Let $\overline{\lambda} \in (\mathbb{F}_p^*)^k$, $\overline{\alpha} \in (\mathbb{F}_{p^n})^k$ be chosen uniformly and independently at random. Then*

$$\Pr_{\overline{\lambda} \in (\mathbb{F}_p^*)^k,\, \overline{\alpha} \in (\mathbb{F}_{p^n})^k}[C = (\overline{\alpha}, \overline{\lambda}) \text{ accepts } D] = \mathbb{E}_{\overline{\beta} \in (\mathbb{F}_{p^n})^{|D|}}\left[\mathrm{Bias}_p(\mathrm{Trace}_{p^n \to p}(f_{\overline{\beta}}))\right]^k.$$

Proof. Our first observation is that

$$\Pr_{\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}[C=(\overline{\alpha},\overline{\lambda})\text{ accepts }D]=\mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|},\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}\left[\omega_p^{\mathrm{Trace}_{p^n\to p}(\sum_{d\in D}\beta_d(\sum_{i=1}^k\lambda_i\alpha_i^d))}\right].$$

To see this let $\mu(\overline{\beta},\overline{\alpha},\overline{\lambda})=\omega_p^{\mathrm{Trace}_{p^n\to p}(\sum_{d\in D}\beta_d(\sum_{i=1}^k\lambda_i\alpha_i^d))}$. Note that for every $\overline{\alpha},\overline{\lambda}$ such that $C=(\overline{\alpha},\overline{\lambda})$ accepts D we have that $\sum_{i=1}^k\lambda_i\alpha_i^d=0$ for all $d\in D$, and hence $\mu(\overline{\beta},\overline{\alpha},\overline{\lambda})=1$ for all $\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|}$. In particular, in this case $\mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|}}\mu(\overline{\beta},\overline{\alpha},\overline{\lambda})=1$. On the other hand, for every $\overline{\alpha},\overline{\lambda}$ such that $C=(\overline{\alpha},\overline{\lambda})$ does not accept D there exists $d\in D$ such that $\sum_{i=1}^k\lambda_i\alpha_i^d\neq 0$, and hence $\mathrm{Trace}_{p^n\to p}(\sum_{d\in D}\beta_d(\sum_{i=1}^k\lambda_i\alpha_i^d))$ is distributed uniformly over $\mathbb{F}_p$ when $\overline{\beta}$ is distributed uniformly over $(\mathbb{F}_{p^n})^{|D|}$. This implies in turn that $\mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|}}\mu(\overline{\beta},\overline{\alpha},\overline{\lambda})=0$.

Thus we have

$$\begin{aligned}
\Pr_{\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}[C=(\overline{\alpha},\overline{\lambda})\text{ accepts }D] &= \mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|},\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}\left[\omega_p^{\mathrm{Trace}_{p^n\to p}(\sum_{d\in D}\beta_d(\sum_{i=1}^k\lambda_i\alpha_i^d))}\right]\\
&= \mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|},\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}\left[\omega_p^{\sum_{i=1}^k\lambda_i\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_i))}\right]\\
&= \mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|},\lambda_0\in\mathbb{F}_p^*,\,\alpha_0\in\mathbb{F}_{p^n}}\left[\omega_p^{\lambda_0\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_0))}\right]^k\\
&= \mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|}}[\mathrm{Bias}_p(\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}))]^k
\end{aligned}$$

To see the last equality let $\rho(\lambda_0,\alpha_0)=\omega_p^{\lambda_0\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_0))}$. Note that for every $\alpha_0\in\mathbb{F}_{p^n}$ such that $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_0))=0$ we have that $\rho(\lambda_0,\alpha_0)=1$ for all $\lambda_0\in\mathbb{F}_p$. In particular, in this case $\mathbb{E}_{\lambda_0\in\mathbb{F}_p}\rho(\lambda_0,\alpha_0)=1$. On the other hand, for every $\alpha_0\in\mathbb{F}_{p^n}$ such that $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_0))\neq 0$ we have that $\lambda_0\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_0))$ is distributed uniformly over $\mathbb{F}_p$ when λ_0 is distributed uniformly over $\mathbb{F}_p$ and hence $\mathbb{E}_{\lambda_0\in\mathbb{F}_p}\rho(\lambda_0,\alpha_0)=0$. This implies in turn that

$$\mathbb{E}_{\lambda_0\in\mathbb{F}_p,\,\alpha_0\in\mathbb{F}_{p^n}}\rho(\lambda_0,\alpha_0)=\Pr_{\alpha_0\in\mathbb{F}_{p^n}}[\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_0))=0],$$

and hence

$$\mathbb{E}_{\lambda_0\in\mathbb{F}_p^*,\,\alpha_0\in\mathbb{F}_{p^n}}\rho(\lambda_0,\alpha_0)=\frac{\Pr_{\alpha_0\in\mathbb{F}_{p^n}}[\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}(\alpha_0))=0]-1/p}{1-1/p}=\mathrm{Bias}_p(\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}})).$$

□

In order to compute the expectation in the right-hand side of Lemma 5.4 we shall use the following special bound on the distribution of the image of sparse polynomials from [KL11] (see also [KL10]).

Theorem 5.5 (Bounds on the image distribution of sparse polynomials, [KL10], Theorem 1.3.)**.** *Let p be a prime, and let $f(x)$ be a univariate polynomial over $\mathbb{F}_{p^n}$. Suppose that $f(x)$ is the sum of at most t monomials over $\mathbb{F}_{p^n}$, each of degree if p-weight at most w with respect to p. Then either* $\mathrm{Trace}_{p^n\to p}(f(x))$ *is constant for every $x\in\mathbb{F}_{p^n}$, or*

$$\left|\mathbb{E}_{x\in\mathbb{F}_{p^n}}\left[\omega_p^{Trace_{p^n\to p}(f(x))}\right]\right|\leq|\mathbb{F}_{p^n}|^{-1/(2w^2 2^w t)}.$$

Note that the above bound applies only if $\mathrm{Trace}_{p^n\to p}(f(x))$ is non-constant, and hence in order to compute the expectation in the right-hand side of Lemma 5.4 we shall also need to compute the probability, over all vectors $\overline{\beta} \in (\mathbb{F}_{p^n})^{|D|}$, that the function $f_{\overline{\beta}}(x) = \sum_{d\in D} \beta_d x^d$ is the constant function. This is done in the following lemma.

Lemma 5.6. *Let p be a prime, and let D be a subset of $\{0, 1, \ldots, p^n - 1\}$ which satisfies that $\mathrm{Shift}_{p,n}(d) \cap \mathrm{Shift}_{p,n}(d') = \emptyset$ for every distinct $d, d' \in D$. Then*

$$\Pr_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|}}[\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \text{ is the zero function}] = p^{-|\mathrm{Shift}_{p,n}(D)|}.$$

In addition, for every $c \in \mathbb{F}_p^*$,

$$\Pr_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|}}[\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \equiv c] = \begin{cases} p^{-|\mathrm{Shift}_{p,n}(D)|}, & 0 \in D \\ 0, & \text{otherwise} \end{cases}$$

Proof. Our first observation is that $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}})$ is the zero function if and only if the polynomial $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \mod x^{p^n} - x$ is the zero polynomial. Note that

$$\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) = \mathrm{Trace}_{p^n\to p}\left(\sum_{d\in D} \beta_d x^d\right) = \sum_{d\in D} \mathrm{Trace}_{p^n\to p}\left(\beta_d x^d\right).$$

Our second observation is that since all degrees in D lie in distinct (p, n)-shifts then all monomials in $\mathrm{Trace}_{p^n\to p}\left(\beta_d x^d\right) \mod x^{p^n} - x$ are distinct from all monomials in $\mathrm{Trace}_{p^n\to p}\left(\beta_{d'} x^{d'}\right) \mod x^{p^n} - x$ for all distinct $d, d' \in D$.

One last observation is that if $d \in D$ satisfies $|\mathrm{Shift}_{p,n}(d)| = \ell$ then all monomials in $\mathrm{Trace}_{p^n\to p}\left(\beta_d x^d\right) \mod x^{p^n} - x$ are of the form $x^{d\cdot p^i \mod p^n-1}$, $0 \le i \le \ell - 1$, where the coefficient of the monomial $x^{d\cdot p^i \mod p^n-1}$ is $(\mathrm{Trace}_{p^n\to p^\ell}(\beta_d))^{p^i}$. Note that $(\mathrm{Trace}_{p^n\to p^\ell}(\beta_d))^{p^i}$ equals zero if and only if $(\mathrm{Trace}_{p^n\to p^\ell}(\beta_d))^{p^j}$ equals zero, and that the number of elements $\beta_d \in \mathbb{F}_{p^n}$ which satisfy $(\mathrm{Trace}_{p^n\to p^\ell}(\beta_d))^{p^i} = 0$ is exactly $p^{n-\ell}$. Thus the probability that a random element $\beta_d \in \mathbb{F}_{p^n}$ satisfies that $\mathrm{Trace}_{p^n\to p}\left(\beta_d x^d\right) \mod x^{p^n} - x$ is the zero polynomial is exactly $p^{-\ell}$.

Concluding, we have that for every $d \in D$ the probability that a random element $\beta_d \in \mathbb{F}_{p^n}$ satisfies that $\mathrm{Trace}_{p^n\to p}\left(\beta_d x^d\right) \mod x^{p^n} - x$ is the zero polynomial equals $p^{-|\mathrm{Shift}_{p,n}(d)|}$. Since all polynomials of the form $\mathrm{Trace}_{p^n\to p}\left(\beta_d x^d\right) \mod x^{p^n} - x$ for $d \in D$ have distinct monomials, this implies in turn that

$$\Pr_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|D|}}[\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \text{ is the zero function}] = \prod_{d\in D} p^{-|\mathrm{Shift}_{p,n}(d)|} = p^{-\sum_{d\in D}|\mathrm{Shift}_{p,n}(d)|} = p^{-|\mathrm{Shift}_{p,n}(D)|}.$$

The second part of the lemma can be proved similarly by observing that $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \equiv c$ if and only if the polynomial $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \mod x^{p^n} - x$ equals c, and that this happens if and only if the coefficient of the monomial 1 in $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \mod x^{p^n} - x$ equals c while all other coefficients equal 0.

□

We are now ready for the proof of Lemma 5.1.

Proof of Lemma 5.1. Let S be a set of (p,n)-shift representative set of D, $|S| = t$. Lemma 3.4 implies that the constraint C accepts D if and only if it accepts S, and thus it suffices to prove the lemma for the set S. From Lemma 5.4 we have that

$$\Pr_{\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}[C=(\overline{\alpha},\overline{\lambda}) \text{ accepts } S] = \mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|S|}}[\mathrm{Bias}_p(\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}))]^k.$$

Suppose first that $0 \notin D$. Note that $f_{\overline{\beta}}$ is a univariate polynomial over $\mathbb{F}_{p^n}$ which is the sum of at most t monomials, each of degree of p-weight at most w. Thus from Theorem 5.5 we have that the function $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}})$ is either the zero function, in which case its bias equals 1, or it is a non-constant function, in which case it satisfies

$$|\mathbb{E}_{x\in\mathbb{F}_{p^n}}\left[\omega_p^{Trace_{p^n\to p}(f_{\overline{\beta}})}\right]| \le |\mathbb{F}_{p^n}|^{-1/(2w^2 2^w t)},$$

and in particular

$$|\mathrm{Bias}_p(\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}))| \le |\mathbb{F}_{p^n}|^{-1/(2w^2 2^w t)}.$$

Moreover, from Lemma 5.6 we have that the probability over all $\overline{\beta}\in(\mathbb{F}_{p^n})^{|S|}$ that $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}})$ is the zero polynomial is $p^{-|\mathrm{Shift}_{p,n}(S)|} = p^{-|\mathrm{Shift}_{p,n}(D)|}$. Concluding, we have that

$$\left|\Pr_{\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}[C=(\overline{\alpha},\overline{\lambda}) \text{ accepts } D] - p^{-|\mathrm{Shift}_{p,n}(D)|}\right| =$$

$$\left|\mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|S|}}[\mathrm{Bias}_p(\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}))]^k - p^{-|\mathrm{Shift}_{p,n}(D)|}\right| \le |\mathbb{F}_{p^n}|^{-k/(2w^2 2^w t)}$$

Next suppose that $0 \in D$. From Theorem 5.5 we have that the function $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}})$ is either the zero function, in which case its bias equals 1, or a constant non-zero function, in which case its bias equals $-\frac{1}{p-1}$, or it satisfies

$$|\mathrm{Bias}_p(\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}))| \le |\mathbb{F}_{p^n}|^{-1/(2w^2 2^w t)}.$$

Moreover, from Lemma 5.6 we have that for every $c\in\mathbb{F}_p$ the probability over all $\overline{\beta}\in(\mathbb{F}_{p^n})^{|S|}$ that $\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}) \equiv c$ is $p^{-|\mathrm{Shift}_{p,n}(S)|} = p^{-|\mathrm{Shift}_{p,n}(D)|}$. Concluding, we have that

$$\left|\Pr_{\overline{\lambda}\in(\mathbb{F}_p^*)^k,\,\overline{\alpha}\in(\mathbb{F}_{p^n})^k}[C=(\overline{\alpha},\overline{\lambda}) \text{ accepts } D] - p^{-|\mathrm{Shift}_{p,n}(D)|} - (p-1)\left(-\frac{1}{p-1}\right)^k p^{-|\mathrm{Shift}_{p,n}(D)|}\right| =$$

$$\left|\mathbb{E}_{\overline{\beta}\in(\mathbb{F}_{p^n})^{|S|}}[\mathrm{Bias}_p(\mathrm{Trace}_{p^n\to p}(f_{\overline{\beta}}))]^k - \left(1+(-1)^k\left(\frac{1}{p-1}\right)^{k-1}\right)p^{-|\mathrm{Shift}_{p,n}(D)|}\right| \le |\mathbb{F}_{p^n}|^{-k/(2w^2 2^w t)}$$

□

6 A calculus for composing pseudo-tests — Proof of Lemma 3.10

Below we introduce two operations on constraints, each operation combines two constraints to create a single constraint (of larger arity). We describe the property of these combination operators in the language of pseudo-tests. For a pair of vectors $\gamma = (\gamma_1,\ldots,\gamma_k)$ and $\gamma' = (\gamma'_1,\ldots,\gamma'_k)$ let $\gamma\circ\gamma' = (\gamma_1,\ldots,\gamma_k,\gamma'_1,\ldots,\gamma'_k)$ denote their concatenation. Let 0_k denote the all-zeros vector of length k.

Definition 6.1 (Union of constraints)**.** Let $C_1 = \left(\overline{\alpha}^{(1)}, \left\{\overline{\lambda}_i^{(1)}\right\}_{i=1}^{r_1}\right)$ be a k_1-constraint and let $C_2 = \left(\overline{\alpha}^{(2)}, \left\{\overline{\lambda}_i^{(2)}\right\}_{i=1}^{r_2}\right)$ be a k_2-constraint. Their union $C = C_1 \cup C_2$ is the $(k_1 + k_2)$-constraint $C = \left(\overline{\alpha}, \left\{\overline{\lambda}_i'^{(1)}\right\}_{i=1}^{r_1} \cup \left\{\overline{\lambda}_i'^{(2)}\right\}_{i=1}^{r_2}\right)$ defined by

$$\overline{\alpha} = \overline{\alpha}^{(1)} \circ \overline{\alpha}^{(2)},$$

$$\overline{\lambda}_i'^{(1)} = \overline{\lambda}_i^{(1)} \circ 0_{k_2} \text{ for all } 1 \le i \le r_1,$$

and

$$\overline{\lambda}_i'^{(2)} = 0_{k_1} \circ \overline{\lambda}_i^{(2)} \text{ for all } 1 \le i \le r_2.$$

Proposition 6.2. *Let C_1 be a pseudo-test separating D_1 from B_1 and let C_2 be a pseudo-test separating D_2 from B_2. Then $C_1 \cup C_2$ is a pseudo-test separating $D_1 \cap D_2$ from $B_1 \cup B_2$.*

Proof. For every degree d,

$$\sum_{j=1}^{k_1+k_2} \lambda_{i,j}'^{(1)} \alpha_j^d = \sum_{j=1}^{k_1} \lambda_{i,j}^{(1)} \left(\alpha_j^{(1)}\right)^d \text{ for all } 1 \le i \le r_1,$$

and similarly

$$\sum_{j=1}^{k_1+k_2} \lambda_{i,j}'^{(2)} \alpha_j^d = \sum_{j=1}^{k_2} \lambda_{i,j}^{(2)} \left(\alpha_j^{(2)}\right)^d \text{ for all } 1 \le i \le r_2.$$

Thus the constraint C accepts the degree d if and only if both C_1 and C_2 accept the degree d. Hence C accepts all degrees in $D_1 \cap D_2$ and rejects all degrees in $B_1 \cup B_2$.

□

Definition 6.3 (Tensor of constraints)**.** Let $C_1 = \left(\overline{\alpha}^{(1)}, \left\{\overline{\lambda}_i^{(1)}\right\}_{i=1}^{r_1}\right)$ be a k_1-constraint and let $C_2 = \left(\overline{\alpha}^{(2)}, \left\{\overline{\lambda}_i^{(2)}\right\}_{i=1}^{r_2}\right)$ be a k_2-constraint. Their tensor product $C = C_1 \otimes C_2$ is the $(k_1 \cdot k_2)$-constraint $C = \left(\overline{\alpha}, \left\{\overline{\lambda}_{(i_1,i_2)}\right\}_{i_1=1,i_2=1}^{r_1,r_2}\right)$, where $\overline{\alpha} \in (\mathbb{F}_{q^n})^{k_1 \times k_2}$ and $\overline{\lambda}_{(i_1,i_2)} \in (\mathbb{F}_q)^{k_1 \times k_2}$ for all $1 \le i \le r_1$, $1 \le i \le r_2$, and are defined as follows:

$$\alpha_{(j_1,j_2)} = \alpha_{j_1}^{(1)} \cdot \alpha_{j_2}^{(2)},$$

and

$$\lambda_{(i_1,i_2),(j_1,j_2)} = \lambda_{i_1,j_1}^{(1)} \cdot \lambda_{i_2,j_2}^{(2)}$$

for all $1 \le j_1 \le k_1$, $1 \le j_2 \le k_2$, $1 \le i_1 \le r_1$, $1 \le i_2 \le r_2$.

Proposition 6.4. *Let C_1 be a pseudo-test separating D_1 from B_1 and let C_2 be a pseudo-test separating D_2 from B_2. Then $C_1 \otimes C_2$ is a pseudo-test separating $D_1 \cup D_2$ from $B_1 \cap B_2$.*

Proof. For every degree d, and for every $1 \leq i_1 \leq r_1$, $1 \leq i_2 \leq r_2$ we have that

$$\begin{aligned}\sum_{j_1=1}^{k_1}\sum_{j_2=1}^{k_2}\lambda_{(i_1,i_2),(j_1,j_2)}(\alpha_{(j_1,j_2)})^d &= \sum_{j_1=1}^{k_1}\sum_{j_2=1}^{k_2}\lambda^{(1)}_{i_1,j_1}\cdot\lambda^{(2)}_{i_2,j_2}\cdot\left(\alpha^{(1)}_{j_1}\right)^d\cdot\left(\alpha^{(2)}_{j_2}\right)^d \\ &= \left(\sum_{j_1=1}^{k_1}\lambda^{(1)}_{i_1,j_1}\left(\alpha^{(1)}_{j_1}\right)^d\right)\cdot\left(\sum_{j_2=1}^{k_2}\lambda^{(2)}_{i_2,j_2}\left(\alpha^{(2)}_{j_2}\right)^d\right).\end{aligned}$$

Thus the constraint C accepts the degree d if and only if at least one of the constraints C_1, C_2 accepts d. Hence C accepts all degrees in $D_1 \cup D_2$ and rejects all degrees in $B_1 \cap B_2$.

□

The above, simple constructions, result in the Composition Lemma 3.10, which allows to decompose a testing problem into simpler ones.

Proof of Lemma 3.10. Fix $d \in D$ and let $S \subseteq [\ell]$ be the set $\{j \in [\ell] | \exists b \in B \text{ s.t. } C_j \text{ is a pseudo-test separating } \{d\} \text{ from } \{b\}\}$. Let C_d be the union of all constraints indexed by indices in S. Then C_d is at most a $(k' \cdot \ell)$-constraint and by Proposition 6.2 above C_d is a pseudo-test separating $\{d\}$ from B.

Now let C be the tensor of all constraints C_d for $d \in D$. The constraint C is of size at most $(k' \cdot \ell)^t$ and by Proposition 6.4, C is a pseudo-test separating D from B . The lemma thus holds for $k = (k'\ell)^t$. □

7 Equivalence of basic and general single-orbit characterizations

Recall that an affine-invariant linear property $\mathcal{F}$ is k-single-orbit characterizable if there exists a k-constraint $C = (\overline{\alpha}, \{\overline{\lambda}_i\}_{i=1}^r)$ which forms a k-single-orbit characterization of $\mathcal{F}$. We say that $\mathcal{F}$ has a *basic* k-single-orbit characterization if $r = 1$.

A natural question is whether every affine-invariant linear property which has a single-orbit characterization with $r > 1$ and of small locality also has a basic single-orbit characterization of small locality. In this section we answer this question in the affirmative by showing that every k-single-orbit characterization of an affine-invariant linear property $\mathcal{F}$ can be transformed into a basic k'-single-orbit characterization of $\mathcal{F}$, where k' depends only on k.

Theorem 7.1 (Equivalence of basic and general single-orbit characterizations)**.** *For every integer k there exists an integer $n_0 = n_0(k)$ such that the following holds for all $n \geq n_0$. If an affine-invariant linear property $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ has a k-single-orbit characterization then it also has a basic k^2-single-orbit characterization.*

For the proof of the above theorem we shall use the following theorem from [BS11] (see also [BS10]) which gives a bound on the q-weight of degrees in the degree set of affine-invariant linear properties which are accepted by a k-constraint.

Theorem 7.2 (Weight-degree of affine-invariant linear properties accepted by a k-constraint, [BS10], Theorem 2.9.)**.** *Suppose that $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is an affine-invariant linear property, where $q = p^s$ for a prime p. Suppose furthermore that there exists a k-constraint which accepts all functions $f \in \mathcal{F}$. Then $\mathrm{wt}_q(d) \leq (k-1)q/p$ for every degree $d \in \mathrm{Deg}(\mathcal{F})$.*

We shall also use the following well-known Schwartz-Zippel lemma that bounds the number of zeros of multivariate polynomials.

Lemma 7.3 (Schwartz-Zippel, [Sch80, Zip79]). *Let $p \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial of degree d over a finite field $\mathbb{F}$. Then*

$$\Pr_{x_1,x_2,\ldots,x_n}[p(x_1, x_2, \ldots, x_n) = 0] \leq d/|\mathbb{F}|$$

Proof of Theorem 7.1. Since $\mathcal{F}$ is k-single-orbit characterizable there exists a k-constraint $C = \left(\overline{\alpha}, \left\{\overline{\lambda}_i\right\}_{i=1}^{r}\right)$ that forms a k-single-orbit characterization of $\mathcal{F}$. Without loss of generality we may assume that all vectors $\overline{\lambda}_1, \overline{\lambda}_2, \ldots, \overline{\lambda}_r$ are linearly independent and hence $r \leq k$.

We will show that for all sufficiently large n there exists a choice of elements $\gamma_1, \gamma_2, \ldots, \gamma_r \in \mathbb{F}_{q^n}$ such that the $(k \cdot r)$-basic-constraint $C' = (\overline{\alpha}, \overline{\lambda})$, $\overline{\alpha} \in (\mathbb{F}_{q^n})^{r \times k}$, $\overline{\lambda} \in (\mathbb{F}_q)^{r \times k}$, defined by $\alpha_{i,j} = \alpha_j \cdot \gamma_i$ and $\lambda_{i,j} = \lambda_{i,j}$ for all $1 \leq i \leq r$, $1 \leq j \leq k$ forms a single-orbit characterization of $\mathcal{F}$. Note that the fact that $r \leq k$ implies that C' is indeed a k^2-constraint.

Let B' be a (q, n)-shift representative set for Border$(\mathcal{F})$ which contains the minimal degree from each (q, n)-shift in Border$(\mathcal{F})$. From Lemmas 2.7 and 3.4 it suffices to show that C' accepts all degrees in Deg$(\mathcal{F})$ and rejects all degrees in B'.

For every degree $d \in \{0, 1, \ldots, q^n - 1\}$ let $P_d(x_1, x_2, \ldots, x_r)$ be the polynomial in the variables $x_1, x_2, \ldots, x_r$, defined as follows.

$$P_d(x_1, x_2, \ldots, x_r) = \sum_{i=1}^{r} \sum_{j=1}^{k} \lambda_{i,j} \cdot \left(\alpha_j \cdot x_i\right)^d = \sum_{i=1}^{r} x_i^d \left(\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^d \right)$$

Note that for every choice of elements $\gamma_1, \gamma_2, \ldots, \gamma_r \in \mathbb{F}_{q^n}$, the constraint C' accepts the degree d if and only if $P_d(\gamma_1, \gamma_2, \ldots, \gamma_r) = 0$. Thus we need to show that there exists a choice of elements $\gamma_1, \gamma_2, \ldots, \gamma_r \in \mathbb{F}_{q^n}$ such that $P_d(\gamma_1, \gamma_2, \ldots, \gamma_r) = 0$ for every degree $d \in Deg(\mathcal{F})$ and $P_b(\gamma_1, \gamma_2 \ \ldots, \gamma_r) \neq 0$ for every degree $b \in B'$.

We first observe that $P_d(\gamma_1, \ldots, \gamma_r) = 0$ for every degree $d \in$ Deg$(\mathcal{F})$ and for every choice of elements $\gamma_1, \ldots, \gamma_r \in \mathbb{F}_{q^n}$. To see this recall that the constraint C accepts d and therefore $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^d = 0$ for all $1 \leq i \leq r$.

It remains to show that there exists a choice of elements $\gamma_1, \ldots, \gamma_r \in \mathbb{F}_{q^n}$ such that $P_b(\gamma_1, \ldots, \gamma_r)$ is non-zero for all $b \in B'$. In order to show this we shall bound the probability that random elements $\gamma_1, \ldots, \gamma_r \in \mathbb{F}_{q^n}$ satisfy $P_b(\gamma_1, \ldots, \gamma_r) = 0$ for some $b \in B'$. By union bound,

$$\begin{aligned} \Pr_{\gamma_1,\ldots,\gamma_r \in \mathbb{F}_{q^n}}[P_b(\gamma_1, \ldots, \gamma_r) = 0 \text{ for some } b \in B'] &\leq \sum_{b \in B'} \Pr_{\gamma_1,\ldots,\gamma_r \in \mathbb{F}_{q^n}}[P_b(\gamma_1, \ldots \gamma_r) = 0] \\ &\leq |B'| \cdot \max_{b \in B'} \Pr_{\gamma_1,\ldots,\gamma_r \in \mathbb{F}_{q^n}}[P_b(\gamma_1, \ldots \gamma_r) = 0] \quad (12) \end{aligned}$$

In what follows we show an upper bound on the size of B' and on the probability $\Pr_{\gamma_1,\ldots,\gamma_r \in \mathbb{F}_{q^n}}[P_b(\gamma_1, \ldots \gamma_r) = 0]$ for $b \in B'$ based on Lemmas 7.2 and 7.3.

We start with bounding the probability $\Pr_{\gamma_1,\ldots,\gamma_r \in \mathbb{F}_{q^n}}[P_b(\gamma_1, \ldots \gamma_r) = 0]$ for $b \in B'$. Since C is a k-constraint which accepts all degrees in Deg$(\mathcal{F})$, from Theorem 7.2 we have that $\mathrm{wt}_q(d) \leq (k-1)q/p$ for every $d \in$ Deg$(\mathcal{F})$. From the definition of the border this implies that $\mathrm{wt}_q(b) \leq (k-1)q/p + 1 \leq kq/p$ for all $b \in$ Border$(\mathcal{F})$. Since B' contains the minimal degree from each (q, n)-shift in Border$(\mathcal{F})$, Lemma 4.4 implies that $b \leq q^{(1-p/(qk))n+1}$ for every $b \in B'$. Thus we

have that for every $b \in B'$ the degree of the polynomial $P_b(x_1, \ldots, x_r)$ is at most $q^{(1-p/(qk))n+1}$, and hence the Schwartz-Zippel Lemma (Lemma 7.3) implies that

$$\Pr_{\gamma_1,\ldots,\gamma_r \in \mathbb{F}_{q^n}}[P_b(\gamma_1, \ldots \gamma_r) = 0] \leq \frac{q^{(1-p/(qk))n+1}}{q^n}. \tag{13}$$

Next we bound the size of B'. This can be done by noticing that the fact that all degrees in Border($\mathcal{F}$) are of q-weight at most kq/p implies that

$$|B'| \leq |\text{Border}(\mathcal{F})| \leq \sum_{i=0}^{kq/p} \binom{n}{i} q^i \leq \frac{kq}{p} q^{kq/p} n^{kq/p} + 1. \tag{14}$$

Plugging (13) and (14) into (12) we obtain

$$\Pr_{\gamma_1,\ldots,\gamma_r \in \mathbb{F}_{q^n}}[P_b(\gamma_1, \ldots, \gamma_r) = 0 \text{ for some } b \in B'] \leq \left(\frac{kq}{p} q^{kq/p} n^{kq/p} + 1\right) \frac{q^{(1-p/(qk))n+1}}{q^n}.$$

This implies in turn that for sufficiently large n there exists a choice of elements $\gamma_1, \ldots, \gamma_r \in \mathbb{F}_{q^n}$ such that $P_b(\gamma_1, \ldots, \gamma_r)$ is non-zero for all $b \in B'$ which concludes the proof of the theorem.

□

References

[AFNS06] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: it's all about regularity. In Jon M. Kleinberg, editor, *STOC*, pages 251–260. ACM, 2006.

[BCL+06] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, Balázs Szegedy, and Katalin Vesztergombi. Graph limits and parameter testing. In Jon M. Kleinberg, editor, *STOC*, pages 261–270. ACM, 2006.

[BGK+05] Eli Ben-Sasson, Venkatesan Guruswami, Tali Kaufman, Madhu Sudan, and Michael Viderman. Locally testable codes require redundant testers. *SICOMP: SIAM Journal on Computing*, 39(7):3230–3247, 2005.

[BGK06] J. Bourgain, A. A. Glibichuk, and S. V. Konyagin. Estimates for the number of sums and products and for exponential sums in fields of prime order. *Journal of the London Mathematical Society*, 73(2):380–398, 2006.

[BGM+11a] Eli Ben-Sasson, Elena Grigorescu, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. On sums of locally testable affine invariant properties. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques, volume 6845 of LNCS*, pages 400–411, 2011.

[BGM+11b] Eli Ben-Sasson, Elena Grigorescu, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. On sums of locally testable affine invariant properties. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:79, 2011.

[BHR05] Eli Ben-Sasson, Prahladh Harsha, and Sofya Raskhodnikova. Some 3CNF properties are hard to test. *SICOMP: SIAM Journal on Computing*, 35, 2005.

[BIW06] Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. *SICOMP: SIAM Journal on Computing*, 36, 2006.

[BKT04] J. Bourgain, N. Katz, and T. Tao. A sum-product extimate in finite fields, and applications. *GAFA*, 14:27–57, 2004.

[BS10] Eli Ben-Sasson and Madhu Sudan. Limits on the rate of locally testable affine-invariant codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:108, 2010.

[BS11] Eli Ben-Sasson and Madhu Sudan. Limits on the rate of locally testable affine-invariant codes. In Leslie Ann Goldberg, Klaus Jansen, R. Ravi, and José D. P. Rolim, editors, *APPROX-RANDOM*, volume 6845 of *Lecture Notes in Computer Science*, pages 412–423. Springer, 2011.

[CC11] Todd Cochrane and James Cipra. Sum-product estimates applied to Waring's problem over finite fields. *INTEGERS*, 11, 2011.

[Cip10] James Arthur Cipra. *WARING'S NUMBER IN FINITE FIELDS*. PhD thesis, KANSAS STATE UNIVERSITY, Manhattan, Kansas, USA, 2010.

[GKS08] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. 2-transitivity is insufficient for local testability. In *IEEE Conference on Computational Complexity*, pages 259–267. IEEE Computer Society, 2008.

[GKS09] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. Succint representation of codes with applications to testing. In *Proceedings of RANDOM-APPROX 2009*, volume 5687 of *Lecture Notes in Computer Science*, pages 534–547. Springer, 2009.

[KL05] Tali Kaufman and Simon Litsyn. Almost orthogonal linear codes are locally testable. In *FOCS*, pages 317–326. IEEE Computer Society, 2005.

[KL10] Tali Kaufman and Shachar Lovett. Testing of exponentially large codes, by a new extension to Weil bound for character sums. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:65, 2010.

[KL11] Tali Kaufman and Shachar Lovett. New extension of the weil bound for character sums with applications to coding. In *The 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2011)*, 2011.

[KS07a] Tali Kaufman and Madhu Sudan. Algebraic property testing: The role of invariance. *Electronic Colloquium on Computational Complexity (ECCC)*, 14(111), 2007.

[KS07b] Tali Kaufman and Madhu Sudan. Sparse random linear codes are locally decodable and testable. In *FOCS*, pages 590–600. IEEE Computer Society, 2007.

[KS08] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In Cynthia Dwork, editor, *STOC*, pages 403–412. ACM, 2008.

[KS10] Swastik Kopparty and Shubhangi Saraf. Local list-decoding and testing of random linear codes from high error. In ACM, editor, *Proceedings of the 2010 ACM International Symposium on Theory of Computing: June 5–8, 2010, Cambridge, MA, USA*, pages 417–426, pub-ACM:adr, 2010. ACM Press.

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, 1980.

[Zip79] R. Zippel. Probabilistic algorithms for sparse polynomials. In *ISSAC '79: Proc. Int'l. Symp. on Symbolic and Algebraic Computation,* Lecture Notes in Computer Science, Vol. 72. Springer-Verlag, 1979. Zippel discusses probabilistic methods for testing polynomial identities and properties of systems of polynomials.