Electronic Colloquium on Computational Complexity, Report No. 79 (2011)

On Sums of Locally Testable Affine Invariant Properties

Eli Ben-Sasson* Elena Grigorescu† Ghid Maatouk‡ Amir Shpilka§
Madhu Sudan¶

Abstract

Affine-invariant properties are an abstract class of properties that generalize some central algebraic ones, such as linearity and low-degree-ness, that have been studied extensively in the context of property testing. Affine invariant properties consider functions mapping a big field $\mathbb{F}_{q^n}$ to the subfield $\mathbb{F}_q$ and include all properties that form an $\mathbb{F}_q$-vector space and are invariant under affine transformations of the domain. Almost all the known locally testable affine-invariant properties have so-called "single-orbit characterizations" — namely they are specified by a single local constraint on the property, and the "orbit" of this constraint, i.e., translations of this constraint induced by affine-invariance. Single-orbit characterizations by a local constraint are also known to imply local testability. Despite this prominent role in local testing for affine-invariant properties, single-orbit characterizations are not well-understood.

In this work we show that properties with single-orbit characterizations are closed under "summation". Such a closure does not follow easily from definitions, and our proof uses some of the rich developing theory of affine-invariant properties. To complement this result, we also show that the property of being an n-variate low-degree polynomial over $\mathbb{F}_q$ has a single-orbit characterization (even when the domain is viewed as $\mathbb{F}_{q^n}$ and so has very few affine transformations). This allows us to exploit known results on the single-orbit characterizability of "sparse" affine-invariant properties to show the following: The sum of any sparse affine-invariant property (properties satisfied by $q^{O(n)}$-functions) with the set of degree d multivariate polynomials over $\mathbb{F}_q$ has a single-orbit characterization (and is hence locally testable) when q is prime. Our result leads to the broadest known family of locally testable affine-invariant properties and gives rise to some intriguing questions/conjectures attempting to classify all locally testable affine-invariant properties.

Keywords: Property testing, Symmetries, Direct sums, Error-correcting codes

*Department of Computer Science, Technion — Israel Institute of Technology, Haifa, Israel, 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 and from the US-Israel Binational Science Foundation under grant number 2006104.

†Georgia Tech, Atlanta, GA, elena@cc.gatech.edu. Supported in part by NSF award CCR-0829672 and NSF award 1019343 to the Computing Research Association for the Computing Innovation Fellowship Program.

‡School of Computer and Communications Sciences, EPFL, Switzerland, ghid.maatouk@epfl.ch. Part of this work was conducted while the author was an intern at Microsoft Research. It is supported in part by Grant 228021-ECCSciEng of the European Research Council.

§Faculty of Computer Science, Technion — Israel Institute of Technology, Haifa, Israel and Microsoft Research, Cambridge, MA, shpilka@cs.technion.ac.il. This research was partially supported by the Israel Science Foundation (grant number 339/10).

¶Microsoft Research New England, Cambridge, Massachusetts, USA, madhu@mit.edu.

ISSN 1433-8092

1 Introduction

Given finite sets D and R, let $\{D \to R\}$ denote the set of functions mapping D to R. A *property* $\mathcal{F}$ of functions mapping D to R is simply given by a set $\mathcal{F} \subseteq \{D \to R\}$. The goal of property testing [RS96, GGR98] is to design "query efficient" tests for various properties. Specifically, a (k, ϵ, δ)-tester for $\mathcal{F}$ is a probabilistic oracle algorithm that, given oracle access to a function $f : D \to R$, makes k-queries to f and accepts $f \in \mathcal{F}$ with probability one, while rejecting f that is δ-far from $\mathcal{F}$ with probability at least ϵ. Here, distance is measured by normalized Hamming distance: $\delta(f, g) = |\{x \in D \mid f(x) \neq g(x)\}|/|D|$ denotes the distance between f and g, and $\delta(f, \mathcal{F}) = \min_{g \in \mathcal{F}}\{\delta(f, g)\}$. f is said to be δ-far from $\mathcal{F}$ if $\delta(f, \mathcal{F}) > \delta$ and δ-close otherwise. To minimize notation we say $\mathcal{F}$ is k-locally testable if for every $\delta > 0$ there exists $\epsilon = \epsilon(k, \delta) > 0$ such that $\mathcal{F}$ is (k, ϵ, δ)-locally testable. Our interest is in families of properties that are k-locally testable for some constant k.

In this work we consider testing of "affine-invariant (linear) properties". The domain and range of such properties are fields. Let $\mathbb{F}_q$ denote the field of size q and let $\mathbb{F}_q^*$ denote the non-zero elements of $\mathbb{F}_q$. We consider properties $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ (so q is a prime power and n is a positive integer). $\mathcal{F}$ is *linear* if for every $f, g \in \mathcal{F}$ and $\alpha \in \mathbb{F}_q$, the function $\alpha \cdot f + g$ belongs to $\mathcal{F}$, where $(\alpha \cdot f + g)(x) = \alpha \cdot f(x) + g(x)$. A function $A : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ is *affine* if there exist $\alpha, \beta \in \mathbb{F}_{q^n}$ such that $A(x) = \alpha x + \beta$. We say A is an affine permutation if A is affine and bijective. Note this is equivalent to saying $A(x) = \alpha x + \beta$ for some $\alpha \in \mathbb{F}_{q^n}^*$ and $\beta \in \mathbb{F}_{q^n}$. A property $\mathcal{F}$ is said to be *affine-invariant* if for $f \in \mathcal{F}$ and every affine permutation $A : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$, the function $f \circ A$ is also in $\mathcal{F}$, where $(f \circ A)(x) = f(A(x))$.[1]

The main contribution of this work is to describe a new class of affine-invariant properties that are locally testable. We show that a broad class of locally testable affine-invariant properties (one that includes most known ones) is closed under "sums". But before presenting our results, we motivate the study of affine-invariant properties briefly.

Motivation: The study of affine-invariance was originally motivated in [KS08] by its connections to locally testable codes and to property testing (cf. the recent survey [Sud10]). Indeed, many "base-constructions" of locally testable codes — crucially used in constructing probabilistically checkable proofs [AS98, ALM+98] — are algebraic in nature and come from families of low-degree polynomials. This motivates the search for the minimal algebraic requirements sufficient to obtain families of locally testable codes, and affine-invariance offers a rich and interesting framework in which to study abstract properties shared by low-degree functions and other algebraic locally testable properties. In this respect, the study of affine-invariant property testing is similar to the study of graph property testing initiated by [GGR98]. Graph-property testing abstracts and unifies properties such as k-colorability and triangle-free-ness, by focussing only on the invariance induced by being a "graph property" (i.e., the property should remain invariant under renaming of

[1] In all previous works starting with [KS08], affine-invariance was defined as invariance with respect to all afine functions, and not only with respect to affine permutations. However, the latter class is more aesthetic. For instance it forms a group under composition — indeed, a well-studied one. Hence in this paper, we define affine-invariance as invariance with respect to the group of affine-permutations. Fortunately, the class of properties does not change despite the mild change in the definition and we prove this equivalence in Proposition 2.4 allowing us to incorporate all previously known results about such properties.

the vertices). Affine-invariant testing similarly attempts to abstract and unify algebraic properties such as being linear or of low-degree or a BCH codeword by focussing only on the invariance of the property (and the linearity of the code/property). The study of graph property testing however is much further advanced and has culminated in a complete combinatorial characterization of locally-testable properties [AFNS06, BCL+06]. Testing of affine-invariant properties lacks such a characterization and indeed it is not yet clear what shape such a characterization might take.

An additional reason to study affine-invariant properties is because they correspond to *locally correctable codes*. An error correcting code of blocklength n is said to be locally correctable if it has an associated "local corrector". Given an adversarially corrupted codeword $w \in \mathbb{F}_q^n$ and index $i \in \{1, \ldots, n\}$ the (randomized) local corrector makes a *constant* number (hence it is called "local") of queries to entries of w and outputs, with high probability, the ith entry of the "correct" codeword w' — closest in Hamming distance to w. Linear codes that are locally correctable are easily seen to be locally decodable codes as defined by [KT00] and can be used to construct databases that support private information retrieval [CGKS98] (in general though, local correctability is a stronger property than local decodability, see e.g. [BIW10, BDWY10]) . It can be verified that affine-invariant locally testable codes are in fact locally correctable [KS08] hence our results imply new families of locally correctable (and decodable) codes.

We note that affine invariant properties as studied here are somewhat different from general (non-linear) linear-invariant properties studied in [BCSX09, Sha10, KSV08, BGS10]. In their setting they consider properties whose domain is a large vector space over a constant sized field, and the property is invariant under linear-transformations of this domain, but they do not require the property to form a vector space. This difference in emphasis leads to very different concerns and tools in the two settings. We remark also that invariance in property testing can and has been studied in the context of other invariance groups. See [Sud10] and references therein for a broader coverage. Here we restrict ourselves to (linear) affine-invariant properties.

Known Testable Properties: Previous works have shown local testability results for two broad categories of affine-invariant properties: (1) Reed-Muller properties, and (2) Sparse properties. In this section we give a brief description. A more full description is given following our definitions in Section 2.4.

In our language, Reed-Muller properties are obtained by equating the sets $\mathbb{F}_{q^n}$ and $\mathbb{F}_q^n$ with an $\mathbb{F}_q$-linear bijection. This allows us to view $\mathbb{F}_q$-linear subspaces of $\{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ as linear subspaces of $\{\mathbb{F}_q^n \to \mathbb{F}_q\}$ where the latter is the set of n-variate functions over $\mathbb{F}_q$. The *q-ary Reed-Muller property of weight degree w* is given by the set of functions that are n-variate polynomials of degree at most w in this view.[2] The testing result here shows that the Reed-Muller property with parameter w over $\mathbb{F}_q$ is testable with $q^{O(w/q)}$ queries [KR06] (see also [AKK+05, JPRZ09]), independent of n.

Sparse properties are less structured ones. We note that there are two definitions in the literature, one for general linear properties, which we refer to as *size-sparsity*, and a related (but different) one for affine-invariant linear properties, which we simply call *sparsity*. Both are parameterized by an integer t. A property $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is *t-size-sparse* if $|\mathcal{F}| \leq q^{nt}$. We won't give the formal technical definition of t-sparsity here (see Section 2.4 for the formal definition), but it is related and in particular t-sparse properties are also t-size-sparse. t-size sparsity has been studied in

[2] The reason for choosing the name weight degree will become clear in Section 2.4.

[KL05, KS07, KS10] and the results show (over different fields, and under different conditions) that t-sparse properties that have very good "distance" (no two functions in the property are too close to each other) are testable. t-sparsity was studied in [GKS09, KL10] and here the testing results do not need to assume high-distance, but rather prove that it is a consequence of affine-invariance. The main theorem here, due to [KL10] shows that for every prime q and integer t there exists k, such that for every n every t-sparse $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is k-locally testable (see Proposition 2.24 for a formal statement).

Aside from the classes above, the known testable properties are less "explicit" and are derived from the concept of single-orbit characterizations, described next.

Single-orbit characterizations: Local tests of linear properties work by picking k query locations $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_{q^n}$ (non-adaptively) and then verifying that $f(\alpha_1), \ldots, f(\alpha_k)$ satisfy some given constraint (we formalize this later a bit more carefully). If a property is affine-invariant, it should be equally effective to query $A(\alpha_1), \ldots, A(\alpha_k)$ for some affine permutation A, and then test to see if the function values at these points also satisfy the given constraint. The collection of tests so obtained (by trying out all As) is referred to as the *orbit* of the constraint at $\alpha_1, \ldots, \alpha_k$. If the only functions that satisfy all these constraints are the functions in $\mathcal{F}$, then we say that $\mathcal{F}$ has a *single orbit characterization*.[3]

Single-orbit characterizations seem to be playing a central role in testing of affine-invariant properties. On the one hand, it is known that every k-single-orbit characterized property is k-locally testable [KS08] and some non-single-orbit characterized properties are known to be not locally-testable even though they can be characterized by a collection of k-local constraints [BMSS10]. On the other hand, most known locally testable properties also seem to have some "single-orbit" property. Sparse codes over prime fields were shown to be single-orbit characterized in [KL10] (see also [GKS09]). The Reed-Muller property has the single orbit property over the (large) group of affine transformations over the vector space $\mathbb{F}_q^n$ by natural considerations. (This will be insufficient for our purposes and so we will strengthen it to get a single-orbit characterization over the field $\mathbb{F}_{q^n}$ in this work.)

Remaining cases of known locally testable codes are obtained in one of two ways: (1) By lifting: This is an operation introduced in [BMSS10]. Here we start with a single-orbit property over some field $\mathbb{F}_{q^n}$ and then "lift" this property to one over an extension field $\mathbb{F}_{q^{nm}}$ (in a manner we will describe later). (2) By taking intersections: The intersection of testable properties is always testable. The lifts turn out to be single-orbit characterized by definition, and the intersection of a constant number of single-orbit characterized properties also turns out to be single-orbit characterized essentially by definition.

1.1 Main Result

In this work we extend the class of properties over $\mathbb{F}_{q^n}$ that have single orbit characterizations.

[3] We note again that in previous works, single-orbit characterization referred to the orbit of a constraint under all affine transformations, and not just affine permutations. But since the class of properties invariant under affine transformations equals the class of properties invariant under affine permutations, the two notions are equivalent. See Corollary 2.5 following Proposition 2.4.

Before presenting our result we formally define the notions of constraints, characterization and single-orbit characterization.

Definition 1.1 (k-single orbit characterization)**.** *Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a linear family of functions. A* k-local constraint *C for $\mathcal{F}$ is a set of points $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_{q^n}$, together with a subspace $V \subsetneq \mathbb{F}_q^k$ such that $\langle f(\alpha_1), \ldots, f(\alpha_k)\rangle \in V$ for every $f \in \mathcal{F}$. Denote C by $C = (\alpha_1, \ldots, \alpha_k; V)$.*

We say that a collection of local constraints $C_1, \ldots, C_t$ characterizes *$\mathcal{F}$ if it holds that a function f satisfies $C_1, \ldots, C_t$ if and only if $f \in \mathcal{F}$. If each C_i is a k-local constraint then we say that the C_i's form a* k-local characterization of $\mathcal{F}$.

Finally, let $C = (\alpha_1, \ldots, \alpha_k; V)$ be a k-local constraint. The ***orbit*** *of C under the set of affine permutations is the set of constraints $\{T \circ C\}_T = \{(T(\alpha_1), \ldots, T(\alpha_k); V)\}_T$, for all affine permutations T. If C is such that the orbit of C forms a k-local characterization of $\mathcal{F}$, then we say that $\mathcal{F}$ has a* k-single orbit characterization.

Remark 1.2. *We note that in previous works the orbits are taken under the set of general affine transformations and not just permutations, but as shown in Corollary 2.5, the notion of single-orbit remains the same under the two notions.*

Remark 1.3. *In later parts of this paper we will need to be a bit more explicit about how the vector space $V \subseteq \mathbb{F}^k$ is described. We will describe it by a collection of t (independent) linear constraints, where t is the codimension of V. Specifically, V will be represented by $\{\bar{\lambda}_i\}_{i=1}^t$ where $\bar{\lambda}_i = (\lambda_{i,1}, \ldots, \lambda_{i,k}) \in \mathbb{F}_q^k$, and $V = \{(\beta_1, \ldots, \beta_k) \in \mathbb{F}_q^k \mid \sum_j \lambda_{i,j}\beta_j = 0 \ \forall i \in \{1, \ldots, t\}\}$. Thus a constraint C will be given by $(\bar{\alpha}; \{\bar{\lambda}_i\}_{i=1}^t)$ where $\bar{\alpha} = (\alpha_1, \ldots, \alpha_k) \in \mathbb{F}_{q^n}^k$ and $\bar{\lambda}_i = (\lambda_{i,1}, \ldots, \lambda_{i,k}) \in \mathbb{F}_q^k$. Note that t is always between 1 and k since $V \subsetneq \mathbb{F}_q^k$.*

Our first result considers the sum of affine invariant properties. For properties $\mathcal{F}_1, \mathcal{F}_2 \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ their *sum* is $\mathcal{F}_1 + \mathcal{F}_2 = \{f_1 + f_2 \mid f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2\}$. For general linear properties $\mathcal{F}_1 + \mathcal{F}_2$ is also linear, but the testability of $\mathcal{F}_1, \mathcal{F}_2$ does not imply their sum is locally testable. Indeed it may be the case that $\mathcal{F}_1 + \mathcal{F}_2$ satisfies no local constraints. Sums of affine-invariant properties behave more nicely. It is straightforward to see the the sum of affine-invariant properties is affine-invariant. More interestingly, it is also possible to show (relatively easily) that if for every $i \in \{1, 2\}$, $\mathcal{F}_i$ satisfies a k_i-local constraint, then $\mathcal{F}_1 + \mathcal{F}_2$ satisfies a $k_1 \cdot k_2$-local constraint. However this does not seem to imply local-testability. Here we focus on single-orbit characterized properties and prove their sum is single-orbit characterized.

Theorem 1.4. *For every q, k_1, k_2, there exists $\kappa = \kappa(k_1, k_2, q)$ such that for every n, if $\mathcal{F}_1, \mathcal{F}_2 \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ are affine-invariant properties with $\mathcal{F}_i$ having a k_i-single orbit characterization, then $\mathcal{F}_1 + \mathcal{F}_2$ has a κ-single orbit characterization. Specifically, if $n \geq n_0 = 10k^2 \log k + 10$, where $k = \max\{k_1, k_2\}$, then we can set $\kappa = k_1 \cdot k_2$, else $\kappa = q^{n_0}$ works.*

While the theorem sounds simple, its proof (in Section 3) requires a fair bit of the theory of affine-invariant codes, and in particular relies on the upper bounds on the dimension of locally testable affine invariant codes shown recently in [BS10].

To apply the theorem above to get new families of single-orbit characterized properties, we need good base properties. However, the two families mentioned earlier, sparse properties and Reed-Muller properties were not known to have the single-orbit property over the same group. Reed-Muller properties were known to have the single-orbit property over the group of affine permutations

over $\mathbb{F}_q^n$, while sparse properties are invariant only over $\mathbb{F}_{q^n}$. (And there is no point using the theorem above to prove that the sum of two sparse families is single-orbit — this is already known since the sum of sparse families is also sparse!) To remedy this situation we show that the Reed-Muller property is actually single orbit over the group of affine permutations over $\mathbb{F}_{q^n}$.

Theorem 1.5 (Reed-Muller codes have local single-orbit property)**.** *Let $q = p^s$ be a prime power. Let w, n be integers such that $w + 1 < \sqrt{\frac{n}{\log_q(3ns)}}$. Denote $w + 1 = r(p-1) + \ell$, where $0 \leq \ell < p - 1$. Then, the q-ary Reed-Muller family of weight degree w, $\mathrm{RM}_q(w, n)$, has a k-single orbit characterization for $k = p^r \cdot (\ell + 1)$. In particular, for every w, q there exists a $k = k(w, q)$ such that the q-ary Reed-Muller family of weight degree w has a k-single orbit characterization.*

Indeed an immediate consequence of the two theorems above is that the sum of Reed-Muller and sparse properties over prime fields are locally testable.

Corollary 1.6. *For integers t, d and prime p, there exists a $k = k(t, d, p)$ such that for every n and every pair of properties $\mathcal{F}_1, \mathcal{F}_2 \in \{\mathbb{F}_{p^n} \to \mathbb{F}_p\}$, where $\mathcal{F}_1$ is the p-ary Reed-Muller property of order d, and $\mathcal{F}_2$ is t-sparse, the property $\mathcal{F}_1 + \mathcal{F}_2$ has a k-single orbit characterization, and is hence k-locally testable.*

The corollary above describes the broadest known class of testable properties when n and q are prime. When n is not prime, the earlier-mentioned notion of lifting leads to other locally testable binary properties, and then intersection also leads to further richness. Finally, when q is not a prime, then the single-orbit property seems less explored. In fact it is not known whether sparse properties have single orbit characterizations. (The results of [KL10] also do not explicitly cover the case of non-prime fields.) In general the range of the properties seems to figure centrally in our understanding of affine-invariant property testing. Binary properties seem much better understood than properties over other prime fields, e.g., Theorem 4.1 gives optimal single-orbit characterization of the Reed-Muller property in this case for all n (and not just the case of big n). In turn, properties over prime fields, as mentioned above, are still much better understood than properties over general fields.

In addition to the new results enumerated above, this paper makes some pedagogical contributions to the study of affine-invariant properties. Among other things, we note that notation has been variable in the past (with the undesirable consequence of the same phrase being used to describe different concepts, or sometime the same concept without a formally proven equivalence). Making such notation consistent requires some work, such as proving the equivalence of invariance under affine-transformation with invariance under affine-permutations (and many other such minor, but annoying, subtleties). We do all this in Section 2, which hopefully provides a comprehensive collection of useful properties of affine-invariant families. The second main pedagogical contribution of this work is to suggest possible questions for future work (with some conjectures) that may lead to a classification of locally testable affine-invariant properties. We do this in Section 5, where we first collect all the known testability results, and use this to discuss potential classifications.

Organization of this paper. The paper is organized as follows. In Section 2 we give the basic definitions regarding affine-invariant families and discuss important notions concerning affine invariant properties. Some of these are used in later sections, while others are provided simply to

unify "folklore" knowledge. Section 3 contains the proof of Theorem 1.4 showing that the sum of single-orbit characterized properties is single-orbit characterized. In Section 4 we give the proof of Theorem 1.5 (single-orbit characterization of Reed-Muller families) and also include a strengthening of it for the binary field. Finally, in Section 5 we discuss questions and directions related to the program of characterizing all affine invariant properties.

2 The structure of affine-invariant properties

In this section we introduce some of the basic, and by now well-studied, structural aspects of affine-invariant properties. We start with some basic terminology and standard facts that we use in Section 2.1. We then give, in Section 2.2, the main structural features of affine-invariant properties, which characterize them in terms of classes of families of polynomials. In Section 2.3 we give some basic features of sparse affine-invariant families. Finally, in Section 2.4 we describe some of the known testability results for affine-invariant families.

2.1 Preliminaries

In what follows $\mathbb{F}_q$ will denote the field of q elements of characteristic p, where $q = p^s$ for some integer s. Let $d = \sum_i d_i p^i$ be the base p representation of an integer d. The **weight** (or p-weight) of d is defined as $\mathrm{wt}(d) = \sum_i d_i$. I.e. it is the sum of coefficients in the p-ary representation of d. A non-negative integer $e = \sum_i e_i p^i$ is said to be in the p-shadow of d (or simply in the **shadow** of d), denoted $e \leq_p d$, if $e_i \leq d_i$ for all i. We denote $a \equiv_k b$ whenever a is equal to b modulo k. As we will be studying polynomials modulo identities of the form $x^q - x \equiv_p 0$ it will be convenient to define the following variant of the modular operation. Let a and k be integers. We define $a \bmod^* k$ as

$$a \bmod^* k = \begin{cases} 0 & a = 0 \\ b & \text{where } 1 \leq b \leq k \text{ is such that } b \equiv_k a \end{cases}$$

We also say that $a \equiv b \pmod^* k$ if $a \bmod^* k = b \bmod^* k$. Note that the only difference between mod and mod* is that mod* does not send nonzero multiples of k to zero but rather to k. It is now clear that $x^a \equiv_q x^{a \bmod^* q-1}$.

We will use the following well known theorem of Lucas (for a short proof see, e.g., Theorem 2.5 of [CST11]).

Theorem 2.1 (Lucas' theorem)**.** *In the notations above,* $\binom{d}{e} \equiv_p \prod_i \binom{d_i}{e_i}$*, where* $\binom{d_i}{e_i} = 0$ *if* $d_i < e_i$*.*

In particular, $\binom{d}{e} \neq 0$ if and only if $e \leq_p d$. We will often use the simple fact that for any $x, y \in \mathbb{F}_q$ it holds that

$$(x+y)^d = \sum_{e \leq_p d} \binom{d}{e} x^e y^{d-e}.$$

The trace operator over $\mathbb{F}_{q^n}$ is a function $\mathrm{Trace} : \mathbb{F}_{q^n} \to \mathbb{F}_q$ defined as $\mathrm{Trace}(x) = \sum_{i=0}^{n-1} x^{q^i}$. The following properties of Trace are well-known.

Proposition 2.2. *The trace operator is linear, i.e., for $\alpha, \beta \in \mathbb{F}_{q^n}$ and $\gamma \in \mathbb{F}_q$, $\mathrm{Trace}(\alpha+\beta) = \mathrm{Trace}(\alpha) + \mathrm{Trace}(\beta)$ and $\mathrm{Trace}(\gamma\alpha) = \gamma\mathrm{Trace}(\alpha)$. Moreover, it is a q^{n-1}-to-1 map, i.e., for every $\alpha \in \mathbb{F}_q$, $|\mathrm{Trace}^{-1}(\alpha)| = q^{n-1}$.*

Finally, the following standard distance property of low-degree polynomials will be of use to us.

Lemma 2.3 (Schwartz-Zippel [Sch80, Zip79], see e.g., [MR95, Theorem 7.2])**.** *Let $p \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial of total degree d. Then*

$$\Pr_{x \in \mathbb{F}^n}[p(x) = 0] \leq d/|\mathbb{F}|.$$

2.2 Degree sets of affine-invariant properties

The class of properties that we consider are characterized by their algebraic properties. To describe such properties we need to introduce several notions from the works of [KS08, GKS08, GKS09, BS10, BMSS10]. We stress that while we reproduce the notions here, some of the terminology here is different (and we believe more appropriate) and we take care to ensure the results here apply to the class of properties invariant under the group of affine permutations as opposed to the semi-group/monoid of all affine transformations.

We view functions $f : \mathbb{F}_{q^n} \to \mathbb{F}_q$ as functions from $\mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ whose image just happens to be contained in $\mathbb{F}_q \subseteq \mathbb{F}_{q^n}$. This allows us to view f as (the evaluation of) a univariate polynomial of degree $q^n - 1$.

Let $f(x) = \sum_{d=0}^{q^n-1} c_d x^d$. The *support* of f, denoted $\mathrm{supp}(f)$, is the set $\mathrm{supp}(f) = \{d \mid c_d \neq 0\}$.

As mentioned earlier, previous works starting with [KS08] defined affine-invariance as invariance with respect to the class of general affine functions mapping $\mathbb{F}_{q^n}$ to itself. Group-theoretically the nicer invariance to study would be invariance under the *group* of *invertible* affine-transformations mapping $\mathbb{F}_{q^n}$ to itself. The following proposition shows that these two notions lead to the same definitions.

Proposition 2.4. *If $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is invariant under all affine permutations, then it is also invariant under all affine transformations.*

Proof. Notice all we need to prove is that if $\mathcal{F}$ contains a non-zero function, then it contains every constant function. (Formally, we wish to show that if $f \in \mathcal{F}$ and π is a non-invertible affine map, then $f \circ \pi \in \mathcal{F}$. But $\pi(x) = ax + b$ is a non-invertible affine map if and only if $a = 0$, in which case the function $f \circ \pi$ is simply the constant function with value $f(b)$ everywhere. So it suffices to show every constant (and in particular $f(b)$) is in $\mathcal{F}$.)

Let $f(x) = \sum_{d=0}^{q^n-1} c_d x^d \in \mathcal{F}$ be a non-zero function. First we note that we can assume $c_0 \neq 0$. To do so we note that $f_\alpha(x) = f(x+\alpha) \in \mathcal{F}$ for every $\alpha \in \mathbb{F}_{q^n}$. The constant coefficient of f_α is simply $f(\alpha)$, which is non-zero for some α. We can fix such an α and henceforth work with f_α instead of f.

Next we note that the function $\sum_{\beta \in \mathbb{F}_{q^n}^*} f(\beta x)$ is also in $\mathcal{F}$ (by linearity and affine-permutation-invariance). To understand this function, we first study the quantity $\sum_{\beta \in \mathbb{F}_{q^n}^*} \beta^d$ for $d \in \{1, \ldots, q^n -$

2}. Using the fact that we can write $\mathbb{F}_{q^n}^* = \{\omega^0, \omega, \omega^2, \ldots, \omega^{q^n-2}\}$ for some $\omega \in \mathbb{F}_{q^n}^*$ we get that

$$\sum_{\beta \in \mathbb{F}_{q^n}^*} \beta^d = \sum_{i=0}^{q^n-2} (\omega^i)^d = \sum_{i=0}^{q^n-2} (\omega^d)^i = ((\omega^d)^{q^n-1} - 1)/(\omega^d - 1) = 0.$$

This now yields

$$\begin{aligned} \sum_{\beta \in \mathbb{F}_{q^n}^*} f(\beta x) &= \sum_{\beta \in \mathbb{F}_{q^n}^*} \sum_{d=0}^{q^n-1} c_d \beta^d x^d \\ &= \sum_{d=0}^{q^n-1} c_d x^d \sum_{\beta \in \mathbb{F}_{q^n}^*} \beta^d \\ &= -(c_0 + c_{q^n-1} x^{q^n-1}), \end{aligned}$$

where the final equality uses the fact that $\sum_{\beta \in \mathbb{F}_{q^n}^*} \beta^d = 0$ if $d \notin \{0, q^n - 1\}$ and -1 otherwise.

Let us abbreviate $c_{q^n-1} = \gamma$. Now, if $\gamma = 0$, then we are done since we have the non-zero constant function $-c_0 \in \mathcal{F}$ and thus all constant functions are in $\mathcal{F}$ (by linearity of $\mathcal{F}$).

On the other hand, if $\gamma \neq 0$, then note that $f(x) = -c_0$ if $x = 0$ and $-c_0 - \gamma$ for all $x \neq 0$. (It is almost a constant function.) Now consider $g(x) = \sum_{\beta \in \mathbb{F}_{q^n}} f(x - \beta)$. For every x, we have $g(x) = -((q^n - 1) \cdot \gamma + q^n \cdot c_0) = \gamma$, which is a non-zero constant function which is in $\mathcal{F}$ (since every $f(x - \beta) \in \mathcal{F}$). Thus in this case also we have that every constant function is in $\mathcal{F}$ as desired. □

We note in particular that the notion of single-orbit characterization also does not change under the two notions of invariance classes.

Corollary 2.5. *Let $C = (\alpha_1, \ldots, \alpha_k; V)$ be a single-orbit characterization of $\mathcal{F}$ under affine transformations. Then C is also a single-orbit characterization of $\mathcal{F}$ under affine permutations.*

Proof. Let $\mathcal{G}$ be the set of functions that satisfy every constraint in the affine-permutation orbit of C. Then $\mathcal{G}$ is a $\mathbb{F}_q$-vector space and closed under affine-permutations, and hence, by Proposition 2.4 it is closed under affine-transformations. Thus it satisfies the orbit of C under all affine transformations as well and so must equal $\mathcal{F}$. □

The following definition captures an important feature of the structure of affine invariant families.

Definition 2.6 ($\mathrm{Deg}(\mathcal{F})$)**.** *Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a family of functions. The **degree set** of $\mathcal{F}$, denoted $\mathrm{Deg}(\mathcal{F})$, is the set of degrees of monomials that appear in some polynomial in $\mathcal{F}$. Formally,*

$$\mathrm{Deg}(\mathcal{F}) = \{d \mid \exists f \in \mathcal{F} \text{ such that } d \in \mathrm{supp}(f)\}.$$

To better understand affine-invariance we need to describe some basic properties of the degree sets (the ones that are known to lead to local testability). We do so in the next two definitions.

Definition 2.7 (Shift(d), Shift(D), shift-closed, shift-representatives, Fam(D))**.** *Let d be an integer in $\{0,\ldots,q^n-1\}$. The* ***shift*** *of d is defined as the set of degrees obtained when taking all q powers of x^d. Formally,*

$$\text{Shift}_{q,n}(d) = \{q^i \cdot d \bmod^* q^n - 1 \mid \forall 0 \le i \le n\}.$$

Recall that $q^i \cdot d \bmod^* q^n - 1$ is the integer d' such that if $d \neq 0$ then $d' \equiv q^i d \mod (q^n-1)$ and $1 \le d' \le q^n - 1$, and if $d = 0$ then $d' = 0$.

In what follows, we will always be considering degrees in the support of functions from $\mathbb{F}_{q^n}$ to $\mathbb{F}_q$, so that we drop the subscripts. This notion can easily be extended to that of a set of degrees. For a set $D \subseteq \{0,\ldots,q^n-1\}$, the shift of D is defined as

$$\text{Shift}(D) = \bigcup_{d \in D} \text{Shift}(d).$$

A set D is said to be ***shift-closed*** *if*

$$\text{Shift}(D) = D.$$

For a shift-closed D, a set $S \subseteq D$ is said to be a set of shift-representatives *of D if*

$$\text{Shift}(S) = D \text{ and } \text{Shift}(d) \cap \text{Shift}(d') = \emptyset \text{ for } d, d' \in S.$$

(In other words S contains one element from each "shift" class in D; by convention we assume each element of S is the smallest amongst its shifts.)[4] *Finally, for a shift-closed D, we define*

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

Remark 2.8. *In [BMSS10] and in many other previous works, the set* Shift(D) *is usually called the* orbit *of D, and denoted* Orbit(D)*. However, we feel that it is better to use the notion of* shift *as we already use "orbit" to denote orbits under the action of the affine group and we do not wish to mix the two different group actions.*

Another important ingredient that we will use is the *shadow* of a degree.

Definition 2.9 (Shadow, Shadow-closed set)**.** *For a non-negative integer d, the* ***shadow*** *of d is the set*

$$\text{Shadow}(d) = \{e \mid e \le_p d\}.$$

The shadow of a set S of non-negative integers is simply the union of the shadows of its elements, i.e.,

$$\text{Shadow}(S) = \bigcup_{d \in S} \text{Shadow}(d).$$

We will be interested in ***shadow-closed*** *sets playing the role of degree sets of functions. A set S of non-negative integers is* ***shadow-closed*** *if* Shadow$(S) = S$.

For a general (linear) family $\mathcal{F}$, the support of $\mathcal{F}$ does not give much useful information about $\mathcal{F}$. However, for affine invariant families, this set completely describes the family. Furthermore, sets of degrees that are closed under shifts and under shadows completely characterize affine-invariant properties. The following lemma gives an equivalent definition of Fam(D).

[4] As $d' \in \text{Shift}(d)$ if and only if $d \in \text{Shift}(d')$, such S always exists.

Lemma 2.10. *Let $D \subseteq \{0, \ldots, q^n - 1\}$ be p-shadow closed and q-shift closed. Then* $\mathrm{Fam}(D) = \{f : \mathbb{F}_{q^n} \to \mathbb{F}_q \mid \mathrm{supp}(f) \subseteq D\}$.

Proof. Let $S_1 = \{\mathrm{Trace}(f) \mid f \in \mathbb{F}_{q^n}[x], \mathrm{supp}(f) \subseteq D\}$ be the standard definition of $\mathrm{Fam}(D)$. Let $S_2 = \{f : \mathbb{F}_{q^n} \to \mathbb{F}_q \mid \mathrm{supp}(f) \subseteq D\}$ be the purpotedly equivalent set.

To see $S_1 \subseteq S_2$, we note that the image of Trace is always in $\mathbb{F}_q$ and furthermore $\mathrm{supp}(\mathrm{Trace}(f)) \subseteq \mathrm{Shift}(\mathrm{supp}(f))$. Thus if $\mathrm{supp}(f) \subseteq D$ and D is shift-closed, then for every $g = \mathrm{Trace}(f) \in S_1$, we have $g : \mathbb{F}_{q^n} \to \mathbb{F}_q$ and $\mathrm{supp}(g) \subseteq \mathrm{Shift}(D) = D$ and so $g \in S_2$.

We now prove $S_2 \subseteq S_1$. Let $f(x) = \sum_{d \in D} c_d x^d \in S_2$. Fix $d \in S$ and let $b = |\mathrm{Shift}(d)|$. Further, let $f^{(d)}(x) = \sum_{i=0}^{b-1} c_{q^i d} x^{q^i d}$. We will show below that for every d, $f^{(d)}(x) \in S_1$ and by linearity it follows that $f(x) \in S_1$. (In particular $f = \sum_{d \in S} f^{(d)}$ where S is a shift-representative of D.)

Focussing on $f^{(d)}$, using the fact that $f^{q^i} = f$, we see that $c_{q^i d} = (c_d)^{q^i}$. So if $b = n$, then we are set since $f^{(d)}(x) = \mathrm{Trace}(c_d x^d) \in S_1$. In the general case where $b \neq n$, let $\mathbb{K} = \mathbb{F}_{q^n}$, $\mathbb{L} = \mathbb{F}_{q^b}$ and $\mathbb{F} = \mathbb{F}_q$. Let $\mathrm{Trace}_{\mathbb{K} \to \mathbb{L}}$ denote the trace map from $\mathbb{K}$ to $\mathbb{L}$ given by $\mathrm{Trace}_{\mathbb{K} \to \mathbb{L}}(x) = x + x^{q^b} + x^{q^{2b}} + \cdots + x^{q^{n-b}}$, and let $\mathrm{Trace}_{\mathbb{L} \to \mathbb{F}}(x) = x + x^q + x^{q^2} + \cdots + x^{q^{b-1}}$. Then we note that $f^{(d)}(x) = \mathrm{Trace}_{\mathbb{L} \to \mathbb{F}}(c_d x^d)$, and $c_d \in \mathbb{L}$ (since $c_d = c_{dq^b \pmod{q^n - 1}} = (c_d)^{q^b}$). Now let $\beta \in \mathbb{K}$ be such that $\mathrm{Trace}_{\mathbb{K} \to \mathbb{L}}(\beta) = c_d$. (Such a β exists for every $c_d \in \mathbb{L}$ since the trace is a surjective map.) Finally, we note that the traces compose naturally, i.e., $\mathrm{Trace}(x) = \mathrm{Trace}_{\mathbb{L} \to \mathbb{F}}(\mathrm{Trace}_{\mathbb{K} \to \mathbb{L}}(x))$ and so

$$\begin{aligned} \mathrm{Trace}(\beta x^d) &= \mathrm{Trace}_{\mathbb{L} \to \mathbb{F}}(\mathrm{Trace}_{\mathbb{K} \to \mathbb{L}}(\beta x^d)) \\ &= \mathrm{Trace}_{\mathbb{L} \to \mathbb{F}}(x^d \mathrm{Trace}_{\mathbb{K} \to \mathbb{L}}(\beta)) \\ &= \mathrm{Trace}_{\mathbb{L} \to \mathbb{F}}(c_d x^d) \\ &= f^{(d)}(x) \end{aligned}$$

Thus we have $f^{(d)}(x) = \mathrm{Trace}(\beta x^d) \in S_1$ as desired. □

Our next lemma repeats in different forms in the literature [KS08, GKS08, GKS09, BS10]. Specifically, it is Lemma 3.5 in [BMSS10]. Note that while Lemma 3.5 in [BMSS10] uses the notion of invariance under general affine transformations, the lemma applies also to families invariant under affine permutations due to Proposition 2.4.

Lemma 2.11 (Closed degree sets specify affine-invariant properties)**.** *Let $\mathcal{F}$ be a linear and affine-invariant family. Then* $\mathrm{Deg}(\mathcal{F})$ *is shadow-closed and shift-closed, and* $\mathcal{F} = \mathrm{Fam}(\mathrm{Deg}(\mathcal{F}))$. *Conversely, if D is shadow-closed and shift-closed then D is the degree set of some affine invariant family. More specifically,* $\mathrm{Fam}(D)$ *is affine-invariant and* $D = \mathrm{Deg}(\mathrm{Fam}(D))$.

Remark 2.12. *We note that the interplay between q and p can be a bit confusing so we remind the reader the following facts: $q = p^s$ is a prime power. We will be working over the field $\mathbb{F}_{q^n}$. The weight of a degree d is defined with respect to its p-ary representation and so is the shadow of d. On the other hand, the shift of d is defined with respect to the Frobenius map over $\mathbb{F}_q$. Traces are also computed over $\mathbb{F}_q$ and affine-invariance is of course with respect to the group of affine transformations $x \mapsto ax + b$, acting on $\mathbb{F}_{q^n}$, where $a \in \mathbb{F}_{q^n}^*$ and $b \in \mathbb{F}_{q^n}$.*

2.3 The size of affine-invariant properties

The literature on testing of linear properties has two seemingly related, but distinct, notions of sparsity. Both notions lead to testability (under some conditions). Here we define both notions and explicitly compare them.

Definition 2.13 (Size-sparsity, Sparsity)**.** *A family $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is said to be* t-size-sparse *if $|\mathcal{F}| \leq q^{nt}$. An affine-invariant family $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is said to be* t-sparse *if there exists a set $S \subseteq \{0, \ldots, q^n - 1\}$ with $|S| \leq t$ such that $\mathrm{Shift}(S) = \mathrm{Deg}(\mathcal{F})$.*

Thus while size-sparsity is a general combinatorial notion, sparsity is a more algebraic notion. In Section 2.4 we will explain how these notions relate to property testing, but first we try to relate the size of affine-invariant families to their algebraic structure. Our first lemma gives an exact description of the size in terms of the size of the degree sets.

Lemma 2.14. *Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be an affine-invariant property and let $D = \mathrm{Deg}(\mathcal{F})$. Then $|\mathcal{F}| = q^{|D|}$.*

Proof. Let S be a set of shift-representatives of D, clearly $|D| = \sum_{d \in S} |\mathrm{Shift}(d)|$. For $d \in S$, let $C(d) = \{\alpha \in \mathbb{F}_{q^n} \mid \exists f \in \mathrm{Fam}(D), f(x) = \sum_{e \in D} c_e x^e \text{ with } c_d = \alpha\}$, denote the set of admissible coefficients of x^d in $\mathrm{Fam}(D)$ (where we use the equivalent definition of $\mathrm{Fam}(D)$ from Lemma 2.10).

As noted in the proof of Lemma 2.10, we have that $\alpha \in C(d)$ only if $\alpha \in \mathbb{F}_{q^{|\mathrm{Shift}(d)|}}$. Conversely, we also have $\sum_{i=0}^{|\mathrm{Shift}(d)|-1} (\alpha x^d)^{q^i} \in \mathrm{Fam}(D)$ for every $\alpha \in \mathbb{F}_{q^{|\mathrm{Shift}(d)|}}$. So $C(d) = \mathbb{F}_{q^{|\mathrm{Shift}(d)|}}$ and in particular $|C(d)| = q^{|\mathrm{Shift}(d)|}$. We now note that $|\mathcal{F}| = \prod_{d \in S} |C(d)|$. Indeed, if we pick, for every $d \in S$, $\alpha_d \in C(d)$ then $\sum_{d \in S} \sum_{i=0}^{|\mathrm{Shift}(d)-1|} (\alpha_d x^d)^{q^i} \in \mathrm{Fam}(D)$ and because the shifts of different of $d \in S$ are disjoint, we have that the coefficient of x^d is α_d. Putting these together, we have $|\mathcal{F}| = \prod_{d \in S} |C(d)| = \prod_{d \in S} q^{|\mathrm{Shift}(d)|} = q^{\sum_{d \in S} |\mathrm{Shift}(d)|} = q^{|D|}$ as claimed. □

For every element $d \in \{0, \ldots, q^n - 1\}$, the size of $\mathrm{Shift}(d)$ is at most n. For "typical" elements this is also an equality. (We won't define typical, but an insistent reader could use the equality as a definition of being "typical".) Thus for typical degree sets $D = \mathrm{Deg}(\mathcal{F})$ with a set S of shift-representatives we have $|D| = n|S|$ and then the notions of t-sparsity and t-size-sparsity are equivalent.

However, the above is not true for all families. For instance if one considers the case where $n = 2m$ and $\mathcal{F} = \mathrm{Fam}(\mathrm{Shift}(S))$ for $S = \{0, 1, 1 + q^m\}$ then $\mathcal{F}$ is 3-sparse but $|\mathrm{Shift}(S)| = |\mathrm{Shift}(0)| + |\mathrm{Shift}(1)| + |\mathrm{Shift}(1 + q^m)| = 1 + n + m$, which gives $|\mathcal{F}| = q^{1+3m}$ and so its size-sparsity is some real number which is less than 2.[5]

The following lemma however shows that for sufficiently large n ($n > 6$), sparsity and size-sparsity are related to within a factor of 2 (and so always within a factor of 6).

Lemma 2.15. *For all $n > 6$ and q the following holds: Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be an affine-invariant property. If $\mathcal{F}$ is t-sparse then it is also t-size-sparse. Conversely if $\mathcal{F}$ is t-size-sparse then it is also $(2t + 1)$-sparse.*

[5] We note that [KL10, Claim 4.7] do seem to claim the two notions of sparsity are equivalent, but there appears to be a gap in their proof. Our example clearly shows the notions are not equivalent.

To prove the lemma above we use a more technical lemma below (Lemma 2.16) about integers and their shadows.

Proof. Let $D = \mathrm{Deg}(\mathcal{F})$. Let S be a set of shift representatives of D. We will show $q^{n\cdot(|S|-1)/2} \leq |\mathcal{F}| \leq q^{n|S|}$ and this immediately implies the lemma

The upper bound follows easily from Lemma 2.14 since for every d we have $|\mathrm{Shift}(d)| \leq n$ and so $|D| \leq n \cdot |S|$ and thus $|\mathcal{F}| = q^{|D|} \leq q^{n\cdot|S|}$.

To see the lower bound, we first note that since $\mathrm{Shift}(d) \cap \mathrm{Shift}(d') = \emptyset$ for $d, d' \in S$, we have $|D| = |\cup_{d\in S} \mathrm{Shift}(d)| = \sum_{d\in S} |\mathrm{Shift}(d)|$.

Let $T \subseteq S$ be the set $T = \{d \in S - \{0\} \mid |\mathrm{Shift}(d)| \neq n\}$. Notice T does not contain any integers of weight 1 because these integers have a n different shifts. Let $\psi : T \to S$ be any map that satisfies $\psi(d) \leq_p d$ and $\mathrm{wt}(\psi(d)) = \mathrm{wt}(d) - 1$. Then by Lemma 2.16 ψ is an injective map whose image is in $S - T - \{0\}$. Thus we $|T| \leq (|S| - 1)/2$.

We can use this to bound $|D|$ as follows. We have $|D| = \sum_{d\in S} |\mathrm{Shift}(d)| \geq \sum_{d\in S-T-\{0\}} |\mathrm{Shift}(d)| = n(|S| - |T| - 1) \geq n(|S| - 1)/2$. Using Lemma 2.14 we now have $|\mathcal{F}| = q^{|D|} \geq q^{n(|S|-1)/2}$. □

Lemma 2.16. *For every $n > 6$ and every $q = p^s$: If $d \in \{1, \ldots, q^n - 1\}$ is such that $|\mathrm{Shift}(d)| < n$, then for every $e \leq_p d$ such that $\mathrm{wt}(e) = \mathrm{wt}(d) - 1$, the following are true:*

1. *$|\mathrm{Shift}(e)| = n$.*
2. *$d' = d$ is the unique integer satisfying $\mathrm{wt}(d') = \mathrm{wt}(e) + 1$, $e \leq_p d'$, and $|\mathrm{Shift}(d')| < n$.*

Proof. The proof of this lemma uses Lemma 2.17 stated and proved below.

Notice first that we can shift d and e jointly so we can assume without loss of generality $d = \sum_{i=0}^{n-1} d_i p^i$ with $d_1 > 0$, and $e = d - 1$. Let $a = |\mathrm{Shift}(d)|$ and so a satisfies $q^a d \equiv d \,(\mathrm{mod}^*\, q^n - 1)$. Suppose further that $|\mathrm{Shift}(e)| = b < n$. Then we have $q^b(d-1) \equiv d - 1 \,(\mathrm{mod}^*\, q^n - 1)$.

Playing with the above we see that

$$q^{a+b}d \equiv q^b d = q^b(d-1) + q^b \equiv d - 1 + q^b \,(\mathrm{mod}^*\, q^n - 1).$$

But we also have

$$q^{a+b}d = q^{a+b}(d-1) + q^{a+b} \equiv q^a(d-1) + q^{a+b} \equiv d - q^a + q^{a+b} \,(\mathrm{mod}^*\, q^n - 1).$$

Putting the two together we get $q^{a+b} - q^a - q^b + 1 \equiv 0 \,(\mathrm{mod}^*\, q^n - 1)$. Recall that $q = p^s$. Hence, setting $\tilde{a} = as$, $\tilde{b} = bs$, $\tilde{n} = ns$ and rewriting the above we have $(p^{\tilde{a}} - 1)(p^{\tilde{b}} - 1) \equiv 0 \,(\mathrm{mod}^*\, p^{\tilde{n}} - 1)$. This implies that $(p^{\tilde{a}} - 1)(p^{\tilde{b}} - 1) \equiv 0 \,(\mathrm{mod}\ \ p^{\tilde{n}} - 1)$, which by Lemma 2.17 (setting $k = 2$) can only occur if $\tilde{a}$ or $\tilde{b}$ is a multiple of $\tilde{n}$. This yields Part (1), and we move to Part (2).

Consider $d' = d - 1 + p^c$, for some $0 < c < ns$. Let $a = |\mathrm{Shift}(d)|$ and $b = |\mathrm{Shift}(d')|$ and suppose $a, b < n$. Then we have $q^a d \equiv d \,(\mathrm{mod}^*\, q^n - 1)$ and $q^b d' \equiv d' \,(\mathrm{mod}^*\, q^n - 1)$.

Again we play with the expression $q^{a+b}d$ and notice

$$q^{a+b}d \equiv q^b d = q^b(d - 1 + p^c) + q^b - q^b p^c \equiv d - 1 + p^c + q^b - q^b p^c \,(\mathrm{mod}^*\, q^n - 1).$$

On the other hand, we also have

$$\begin{aligned} q^{a+b}d &= q^{a+b}(d-1+p^c) + q^{a+b} - q^{a+b}p^c \equiv q^a(d-1+p^c) + q^{a+b} - q^{a+b}p^c \\ &\equiv d - q^a + q^a p^c + q^{a+b} - q^{a+b}p^c \,(\mathrm{mod}^*\, q^n - 1). \end{aligned}$$

Putting the two together, we get $(q^a-1)\cdot(q^b-1)\cdot(p^c-1) \equiv 0\,(\mathrm{mod}^*\, q^n-1)$. Setting $\tilde{a} = as$, $\tilde{b} = bs$, $\tilde{c} = c$ and $\tilde{n} = ns$ and rewriting the above (and further taking mod instead of mod*) we have $(p^{\tilde{a}}-1)(p^{\tilde{b}}-1)(p^{\tilde{c}}-1) \equiv 0\,(\mathrm{mod}\ p^{\tilde{n}}-1)$, which by Lemma 2.17 (with $k=3$) can only occur if $\tilde{a}$ or $\tilde{b}$ or $\tilde{c}$ is a multiple of $\tilde{n}$, giving the desired contradiction. □

Finally we use the following lemma to finish the proof of Lemma 2.16. We are grateful to Shripad Garge for permission to include his elegant proof below.

Lemma 2.17 ([Gar11])**.** *Let p be a prime, $n > 2$, and $a_1, \ldots, a_k$ be such that $\prod_{i=1}^{k}(p^{a_i}-1) = 0\,(\mathrm{mod}\ p^n-1)$. Then there exists an i such that n divides a_i, unless $n = 6$ and $p = 2$.*

Proof. Without loss of generality we can assume that the a_i's are at most n (since we can replace each a_i by its residue modulo n). The lemma now follows immediately from [Art55, Corollary 2] which shows that there is always a prime r which divides p^n-1 but not p^j-1 for any $j < n$ provided $n > 2$ (unless $n = 6$ and $p = 2$). Thus, in such case r (and hence p^n-1) cannot divide $\prod_{i=1}^{k}(p^{a_i}-1)$, unless $a_i = n$ for some i. The lemma follows. □

2.4 Known Testable Families

Here we describe known testable families more formally. We start by stating a general theorem due to Kaufman and Sudan [KS08] which gives a general sufficient criterion, but does not give any testable properties explicitly.

Theorem 2.18 ([KS08, Theorem 2.9])**.** *If $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ has a k-single orbit characterization, then $\mathcal{F}$ is k-locally testable.*

The theorem of [KS08] is more general than the above. It applies also to families $\mathcal{F} \subseteq \{\mathbb{F}_{q^n}^m \to \mathbb{F}_q\}$. Affine invariance over $\mathbb{F}_{q^n}^m$ implies affine-invariance over $\mathbb{F}_{q^{nm}}$ (under appropriate correspondence between the two domains) but single-orbit characterization over $\mathbb{F}_{q^n}^m$ does not transfer, so one should take care with this. We point out that [KS08] need only "formal characterizations" a notion more general than single-orbit characterizations, but subsequent works have worked with the more restricted notion and indeed the restricted notion seems to suffice to capture all known testable affine-invariant properties.

Moving on, we now turn to two classes of explicit families of affine-invariant properties that are locally testable. While in both cases the testability results were not originally expressed as a consequence of Theorem 2.18, it turns out that both can be derived this way also, as we elaborate below.

Reed-Muller Codes. It is a well known fact that Reed-Muller (RM) codes are locally testable [KR06] (see also [AKK+05, JPRZ09]) and we later prove (Theorem 1.5) that they in fact have a single orbit characterization. Interestingly, there are two different ways to think of Reed-Muller codes as an affine invariant family. The first and the more commonly used one is to view RM codes of degree d as n-variate polynomials of degree d over $\mathbb{F}_q$. This family is invariant under any affine transformation on $(\mathbb{F}_q)^n$. A second less common view of Reed-Muller codes of degree d is as *univariate* polynomials over $\mathbb{F}_{q^n}$ of *weight degree* d. In this case the group acting on the codewords is the group of affine permutations $x \to ax + b$ where $a \in \mathbb{F}_{q^n}^*$ and $b \in \mathbb{F}_{q^n}$. It is not hard to see that the second group is contained in the first group (by embedding $\mathbb{F}_{q^n} \hookrightarrow \mathrm{GL}_n(\mathbb{F}_q)$, the set of invertible matrices over $\mathbb{F}_q$). However, as the second testable family that we have is affine invariant with respect to the smaller group, we will have to give up on some of the structure that Reed-Muller codes poses over $\mathbb{F}_q$ and study their behavior as a family of weight degree d univariate polynomials over $\mathbb{F}_{q^n}$. For that reason it is important for us to prove Theorem 1.5 as it will allow us to get new families of testable (and even single orbit characterizable) properties using Theorem 1.4. We now give the two formal definitions of Reed-Muller codes and show their equivalence.

Let $\mathrm{RM}_q(w,n) \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be the family of *Reed-Muller codes of weight degree* w, defined as

$$\mathrm{RM}_q(w,n) = \{f : \mathbb{F}_{q^n} \to \mathbb{F}_q \mid \forall d \in \mathrm{supp}(f), \mathrm{wt}(d) \leq w\}. \tag{1}$$

Note that $RM_q(w,n)$ is both linear and affine-invariant.

The following theorem proves the equivalence with the more standard definition of RM codes. We include a proof below for completeness, even though the theorem is folklore and has been used before in the related literature (in particular [BS10]). Let $\phi : \mathbb{F}_{q^n} \mapsto \mathbb{F}_q^n$ be any bijective $\mathbb{F}_q$-linear map. Denote the standard definition of Reed-Muller codes by

$$\widetilde{\mathrm{RM}}_q(w,n) = \{f \circ \phi \mid f = \sum f_{d_1,\ldots,d_n} x_1^{d_1} \cdots x_n^{d_n},\ f_{d_1,\ldots,d_n} \in \mathbb{F}_q,\ \sum d_i \leq w\}.$$

Theorem 2.19. *Let* $\mathrm{RM}_q(w,n)$ *be as defined in (1). Then* $\widetilde{\mathrm{RM}}_q(w,n) = \mathrm{RM}_q(w,n)$.

For this we will rely on the following three claims. The proofs of the first two are easy so we omit them. (The first is easy and the second is easy given the first.)

Claim 2.20. *Both* $\widetilde{\mathrm{RM}}_q(w,n)$ *and* $\mathrm{RM}_q(w,n)$ *are* $\mathbb{F}_q$*-linear spaces.*

For two families of functions $\mathcal{F}_1, \mathcal{F}_2$ we denote $\mathcal{F}_1 \odot \mathcal{F}_2$ to be

$$\mathcal{F}_1 \odot \mathcal{F}_2 = \mathrm{span}_{\mathbb{F}_q}\{f_1 \cdot f_2 \mid f_1 \in \mathcal{F}_1 \text{ and } f_2 \in \mathcal{F}_2\}.$$

Claim 2.21. *For every* $w > 1$ *we have that* $\widetilde{\mathrm{RM}}_q(w,n) = \widetilde{\mathrm{RM}}_q(w-1,n) \odot \widetilde{\mathrm{RM}}_q(1,n)$. *Similarly,* $\mathrm{RM}_q(w,n) = \mathrm{RM}_q(w-1,n) \odot \mathrm{RM}_q(1,n)$.

Note that the proof of this claim is immediate given Claim 2.20.

Claim 2.22. $\widetilde{\mathrm{RM}}_q(1,n) = \mathrm{RM}_q(1,n)$.

Proof. Note that both are $\mathbb{F}_q$ linear spaces of dimension exactly $n+1$. Indeed, for $\widetilde{\mathrm{RM}}$ this is obvious and for RM it follows from the fact that if $f(x) = c + \sum_{i=0}^{n-1} c_i x^{q^i}$ maps $\mathbb{F}_{q^n}$ to $\mathbb{F}_q$ then using $f^q = f$ we get that $c_i = c_0^{q^i}$ and that $c \in \mathbb{F}_q$. Hence, the dimension of $\mathrm{RM}_q(1,n)$ is exactly $n+1$ (it is determined by $(c, c_0) \in \mathbb{F}_q \times \mathbb{F}_{q^n}$). As $\mathbb{F}_{q^n}$ is an $\mathbb{F}_q$ vector space of dimension n, the linear space of affine functions $\{g : \mathbb{F}_{q^n} \to \mathbb{F}_q\}$ has dimension exactly $n+1$. Hence, by comparing dimensions we see that both $\widetilde{\mathrm{RM}}_q(1,n)$ and $\mathrm{RM}_q(1,n)$ are equal to the space of affine functions $\{g : \mathbb{F}_{q^n} \to \mathbb{F}_q\}$. □

We can now give the proof of Theorem 2.19.

Proof of Theorem 2.19. Given that Claim 2.21 holds, it suffices to prove the theorem for $w = 1$ which is done in Claim 2.22. □

Sparse Families. The second class of locally testable affine-invariant properties are what are termed "sparse properties". This phrase tends to refer to two different (but seemingly related) notions and we describe them both below along with the associated testability results.

The first refers to testability due to size-sparsity. These results require an additional condition that the family of codes be of very high-distance. To describe the results, we need the notion of *distance* and the even stronger notion of *bias* of a family. Recall $\delta(f,g)$ denotes the (normalized Hamming) distance between functions f and g, i.e., $\delta(f,g) = q^{-n} \cdot |\{\alpha \in \mathbb{F}_{q^n} \mid f(\alpha) \neq g(\alpha)\}|$ and the distance of a family $\mathcal{F}$ is the quantity $\delta(\mathcal{F}) = \min_{f \neq g \in \mathcal{F}}\{\delta(f,g)\}$. We say that $\mathcal{F}$ is ϵ-biased if for all distinct $f, g \in \mathcal{F}$, we have $\frac{1}{q} + \epsilon \leq \delta(f,g) \leq 1 - \frac{1}{q} - \epsilon$.

Proposition 2.23 ([KS07], see also [KL05, KS10])**.** *For every $\gamma > 0$ and $t < \infty$ there exists a k such that for every n the following holds: If $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ is t-size-sparse and has distance $\delta(\mathcal{F}) \geq \frac{1}{2} - 2^{-\gamma n}$, then $\mathcal{F}$ is k-locally testable.*

The results of [KL05, KS10] only hold for the case of families that have bias $2^{-\gamma n}$ (as opposed to the distance condition stated above). However the proof of [KS10] may potentially extend to the case of functions mapping to any field $\mathbb{F}_q$ (as opposed to only $\mathbb{F}_2$). (If true, this would lead to some testability results for affine-invariant properties not covered by the following theorem.)

In the case of affine-invariant properties one can drop the condition on distance (since this turns out to follow from affine-invariance [GKS09, KL10]). Furthermore, now testability can be attributed to single-orbit characterizations.

Proposition 2.24 ([KL10], see also [GKS09])**.** *For every prime number q and integer t there exists $k = k(t,q)$ such that the following holds. For every n and every affine-invariant $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$, if $\mathcal{F}$ is t-sparse then $\mathcal{F}$ has a k-single orbit characterization (and is hence k-locally testable).*

3 Sums of Affine-Invariant Properties

In this section we prove Theorem 1.4. The main idea behind the proof is that instead of looking at the sets of degrees of a locally characterizable family $\mathcal{F}$, we look at the *border* set of degrees. These are the integers that do not themselves belong to $\mathrm{Deg}(\mathcal{F})$ but every integer in their shadow is in $\mathrm{Deg}(\mathcal{F})$.

Definition 3.1 (Border of a family)**.** *Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a family of functions. The **border** of $\mathcal{F}$ is the set of degrees given by*

$$\text{Border}(\mathcal{F}) = \{d \notin \text{Deg}(\mathcal{F}) \mid \forall e <_p d,\, e \in \text{Deg}(\mathcal{F})\}.$$

The following lemma gives several equivalent definitions to being a k-single orbit characterizable family. The lemma can be seen as an extension of Lemma 3.6 in [BMSS10].

Lemma 3.2. *[Equivalent definitions of k-single orbit characterizable family]*
Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a linear affine-invariant family. The following are equivalent:

1. *$(\bar{\alpha}; \{\bar{\lambda}_i\}_{i=1}^t)$ is a k-single orbit characterization of $\mathcal{F}$, where $\bar{\alpha} = (\alpha_1, \ldots, \alpha_k) \in \mathbb{F}_{q^n}^k$ and $\bar{\lambda}_i = (\lambda_{i,1}, \ldots, \lambda_{i,k}) \in \mathbb{F}_q^k$.*
2. *For all d,*
$$d \in \text{Deg}(\mathcal{F}) \;\Leftrightarrow\; \forall i \sum_{j=1}^{k} \lambda_{i,j}(\alpha_j x + y)^d \equiv 0.$$
In other words, the RHS is the zero polynomial, for every i.
3. *For all d,*
$$d \in \text{Deg}(\mathcal{F}) \;\Leftrightarrow\; \forall e \leq_p d,\, \forall i \sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e = 0.$$
4. *For all $d \in \text{Deg}(\mathcal{F}) \cup \text{Border}(\mathcal{F})$,*
$$d \in \text{Deg}(\mathcal{F}) \;\Leftrightarrow\; \forall i \sum_{j=1}^{k} \lambda_{i,j}\alpha_j^d = 0.$$

We break the proof down into four intermediary lemmas. Clearly it is enough to prove Lemmas 3.3, 3.4, 3.5 and 3.6 in order to prove Lemma 3.2.

Lemma 3.3. *Conditions 2 and 3 in Lemma 3.2 are equivalent.*

Proof. For this, it is enough to prove that for any d and for all $1 \leq i \leq t$,

$$\sum_{j=1}^{k} \lambda_{i,j}(\alpha_j x + y)^d = 0 \;\Leftrightarrow\; \forall e \leq_p d,\, \sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e = 0.$$

Fix d and i. By the discussion following Lucas' theorem (Theorem 2.1) we get

$$\begin{aligned}\sum_{j=1}^{k} \lambda_{i,j}(\alpha_j x + y)^d &= \sum_{j=1}^{k} \lambda_{i,j} \sum_{e \leq_p d} \binom{d}{e} \alpha_j^e x^e y^{d-e} \\ &= \sum_{e \leq_p d} \binom{d}{e} \left(\sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e \right) x^e y^{d-e}.\end{aligned}$$

Since $\binom{d}{e} \neq 0$ when $e \leq_p d$, this equality of polynomials tells us that

$$\sum_{j=1}^{k} \lambda_{i,j}(\alpha_j x + y)^d = 0 \;\Leftrightarrow\; \sum_{e \leq_p d} \binom{d}{e} \left(\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e \right) x^e y^{d-e} = 0 \;\Leftrightarrow\; \forall e \leq_p d, \sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e = 0,$$

as required. □

Lemma 3.4. *Conditions 3 and 4 in Lemma 3.2 are equivalent.*

Proof. We start by proving that Condition 3 implies Condition 4. For this, we assume Condition 3 and prove that for $d \in \mathrm{Deg}(\mathcal{F})$, $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^d = 0$ and for $d \in \mathrm{Border}(\mathcal{F})$, $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^d \neq 0$. First note that if $d \in \mathrm{Deg}(\mathcal{F})$, then any e in the shadow of d satisfies $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e = 0$. In particular, $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^d = 0$. Now, consider $d \in \mathrm{Border}(\mathcal{F})$. Since $d \notin \mathrm{Deg}(\mathcal{F})$, by Condition 3 we know that there exists an e in the shadow of d such that $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e \neq 0$. But d is the only element in the shadow of itself that does not belong to $\mathrm{Deg}(\mathcal{F})$, so that all elements e in the strict shadow of d satisfy $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e = 0$. Thus d must be such that $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^d \neq 0$.

We now prove that Condition 4 implies Condition 3. For this we assume Condition 4 and first prove that for $d \in \mathrm{Deg}(\mathcal{F})$, all elements e in the shadow of d satisfy $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e = 0$. Recall that $\mathcal{F}$ is an affine-invariant family, so that $\mathrm{Deg}(\mathcal{F})$ is shadow-closed. Thus, for $d \in \mathrm{Deg}(\mathcal{F})$, all elements e in the shadow of d also belong to $\mathrm{Deg}(\mathcal{F})$, and by Condition 4 must all satisfy $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e = 0$. Now we prove that for any $d \notin \mathrm{Deg}(\mathcal{F})$, there exists an element e in its shadow such that $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e \neq 0$. We claim that for any $d \notin \mathrm{Deg}(\mathcal{F})$, there exists an element e in its shadow that belongs to $\mathrm{Border}(\mathcal{F})$. Using Condition 4, this completes the proof. To see that the claim is true, take $d \notin \mathrm{Deg}(\mathcal{F})$ and consider a chain of elements in its shadow $d \geq_q e_1 \geq_q \cdots \geq_q e_\ell$, where e_ℓ is a minimal element not in $\mathrm{Deg}(\mathcal{F})$, i.e., all the strict shadow of e_ℓ is contained in $\mathrm{Deg}(\mathcal{F})$. Then either $e_\ell \in \mathrm{Border}(\mathcal{F})$, or $e_\ell = 0$. But for nontrivial $\mathcal{F}$, $\mathrm{Deg}(\mathcal{F})$ always contains 0, so that it must be the case that $e_\ell \in \mathrm{Border}(\mathcal{F})$. □

Lemma 3.5. *Condition 2 implies Condition 1, in Lemma 3.2.*

Proof. Assume $\mathcal{F}$ and $(\bar{\alpha}; \{\bar{\lambda}_i\}_{i=1}^{t})$ are such that Condition 2 holds. We want to show that for any $f : \mathbb{F}_{q^n} \to \mathbb{F}_q$,

$$f \in \mathcal{F} \;\Leftrightarrow\; \forall i : \sum_{j=1}^{k} \lambda_{i,j} f(\alpha_j a + b) = 0 \quad \forall a \in \mathbb{F}_{q^n}^*, b \in \mathbb{F}_{q^n}.$$

Let $f(x) = \sum_d f_d x^d$ be in $\mathcal{F}$. Consider a degree d in the support of f. Since $d \in \mathrm{Deg}(\mathcal{F})$ it follows that

$$\forall i \sum_{j=1}^{k} \lambda_{i,j}(\alpha_j x + y)^d = 0.$$

Fix $a \in \mathbb{F}_{q^n}^*, b \in \mathbb{F}_{q^n}$. We have that

$$\sum_{j=1}^{k} \lambda_{i,j} f(\alpha_j a + b) = \sum_{j=1}^{k} \lambda_{i,j} \sum_d f_d (\alpha_j a + b)^d = \sum_d f_d \left(\sum_{j=1}^{k} \lambda_{i,j} (\alpha_j a + b)^d \right) = \sum_d f_d \cdot 0 = 0.$$

To prove the other side of the equivalence, suppose f is such that for every $1 \leq i \leq t$,

$$\sum_{j=1}^{k} \lambda_{i,j} f(\alpha_j a + b) = 0 \quad \forall a \in \mathbb{F}_{q^n}^*, b \in \mathbb{F}_{q^n}. \tag{2}$$

Define the family $\mathcal{F}'$ as the smallest affine-invariant linear family containing f, that is,

$$\mathcal{F}' = \left\{ \sum_{a,b} \gamma_{ab} f(xa + b) \mid a \in \mathbb{F}_{q^n}^*, b \in \mathbb{F}_{q^n}, \gamma_{ab} \in \mathbb{F}_q \right\}.$$

By linearity and Equation (2), for any $g \in \mathcal{F}'$ and $1 \leq i \leq t$ we have that

$$\sum_{j=1}^{k} \lambda_{i,j} g(\alpha_j a + b) = 0 \quad \forall a, b.$$

Now, every degree d in the support of f satisfy $d \in \mathrm{Deg}(\mathcal{F}')$. As $\mathcal{F}'$ is affine-invariant, every $e \leq_p d$ also belongs to $\mathrm{Deg}(\mathcal{F}')$. Thus, for all $e \leq_p d$,

$$\mathrm{Trace}(\beta x^e) \in \mathcal{F}' \quad \forall \beta \in \mathbb{F}_{q^n}.$$

Indeed, this follows immediately from Lemma 2.11 and the fact that $\mathcal{F}'$ is linear affine invariant.

Thus, for all such e and for any β, $\mathrm{Trace}(\beta x^e)$ satisfies

$$0 = \sum_{j=1}^{k} \lambda_{i,j} \mathrm{Trace}(\beta(\alpha_j a + b)^e) = \mathrm{Trace}\left(\beta \sum_{j=1}^{k} \lambda_{i,j} (\alpha_j a + b)^e \right) \quad \forall a \in \mathbb{F}_{q^n}^*, b \in \mathbb{F}_{q^n},$$

and in particular, letting $a = 1$ and $b = 0$, we get $\mathrm{Trace}(\beta \sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e) = 0$. However, this holds for all β if and only if $\sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e = 0$.

Hence, for every degree d in the support of f and $1 \leq i \leq t$, it holds that

$$\forall e \leq_p d, \sum_{j=1}^{k} \lambda_{i,j} \alpha_j^e = 0.$$

By Condition 3 (which by Lemma 3.3 is equivalent to Condition 2), this implies that $d \in \mathrm{Deg}(\mathcal{F})$ for every d in the support of f. Lemmas 2.10 and 2.11 now imply that $f \in \mathcal{F}$. □

Lemma 3.6. *Condition 1 in Lemma 3.2 implies Condition 3.*

Proof. Assume $\mathcal{F}$ and $(\bar{\alpha}; \{\bar{\lambda}_i\}_{i=1}^t)$ are such that

$$f \in \mathcal{F} \;\Leftrightarrow\; \forall i : \sum_{j=1}^{k} \lambda_{i,j} f(\alpha_j a + b) = 0 \quad \forall a \in \mathbb{F}_{q^n}^*, b \in \mathbb{F}_{q^n}.$$

We want to show that for any d,

$$d \in \mathrm{Deg}(\mathcal{F}) \;\Leftrightarrow\; \forall i \text{ and } \forall e \leq_p d, \sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e = 0.$$

First take $d \in \mathrm{Deg}(\mathcal{F})$. Recall that $\mathrm{Deg}(\mathcal{F})$ is shadow-closed by affine-invariance of $\mathcal{F}$, so that all $e \leq_p d$ belong to $\mathrm{Deg}(\mathcal{F})$, and for all such e, $\mathrm{Trace}(\beta x^e) \in \mathcal{F}$ for every $\beta \in \mathbb{F}_{q^n}$. By Condition 1, we thus have, for all $e \leq_p d$ and $1 \leq i \leq t$,

$$0 = \sum_{j=1}^{k} \lambda_{i,j}\mathrm{Trace}\left(\beta(\alpha_j a + b)^e\right) = \mathrm{Trace}\left(\beta \sum_{j=1}^{k} \lambda_{i,j}(\alpha_j a + b)^e\right) \quad \forall a \in \mathbb{F}_{q^n}^*, b \in \mathbb{F}_{q^n}, \beta \in \mathbb{F}_{q^n}.$$

In particular,

$$\mathrm{Trace}\left(\beta \sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e\right) = 0, \;\forall \beta.$$

But this is true if and only if $\sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e = 0$.

Conversely, assume that d is such that for every i and every $e \leq_p d$ it holds that $\sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e = 0$. Noting that

$$\sum_{j=1}^{k} \lambda_{i,j}\mathrm{Trace}\left((\alpha_j a + b)^d\right) = \mathrm{Trace}\left(\sum_{e \leq_p d} \binom{d}{e} \left(\sum_{j=1}^{k} \lambda_{i,j}\alpha_j^e\right) a^e b^{d-e}\right),$$

we see that $\sum_{j=1}^{k} \lambda_{i,j}\mathrm{Trace}\left((\alpha_j a + b)^d\right) = 0$ for every $a \in \mathbb{F}_{q^n}^*$, $b \in \mathbb{F}_{q^n}$, so that $\mathrm{Trace}(x^d) \in \mathcal{F}$ and $d \in \mathrm{Deg}(\mathcal{F})$. □

3.1 Proof of Main Theorem

Now we have all the required definitions and tools to prove our main theorem.

Proof of Theorem 1.4. Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be linear, affine-invariant families of functions from $\mathbb{F}_{q^n}$ to $\mathbb{F}_q$, such that $\mathcal{F}_1$ has a k_1-single orbit characterization and $\mathcal{F}_2$ has a k_2-single orbit characterization. We denote $\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2$. Let us also denote $k = \max\{k_1, k_2\}$. By Condition 2 in Lemma 3.2, we know that there exist $(\bar{\alpha}^{(1)}; \{\bar{\lambda}_i^{(1)}\}_{i=1}^{t_1})$, where $\bar{\alpha}^{(1)} = (\alpha_1^{(1)}, \ldots, \alpha_{k_1}^{(1)}) \in \mathbb{F}_{q^n}^{k_1}$ and $\bar{\lambda}_i^{(1)} = (\lambda_{i,1}^{(1)}, \ldots, \lambda_{i,k_1}^{(1)}) \in \mathbb{F}_q^{k_1}$, such that for every possible degree d,

$$d \in \mathrm{Deg}(\mathcal{F}_1) \quad \Leftrightarrow \quad \forall 1 \leq i \leq t_1 \quad \sum_{j=1}^{k} \lambda_{ij}^{(1)}(\alpha_j^{(1)} x + y)^d = 0.$$

Similarly, there exist $(\bar{\alpha}^{(2)}; \{\bar{\lambda}_i^{(2)}\}_{i=1}^{t_2})$, where $\bar{\alpha}^{(2)} = (\alpha_1^{(2)}, \ldots, \alpha_{k_2}^{(2)}) \in \mathbb{F}_{q^n}^{k_2}$ and $\bar{\lambda}_i^{(2)} = (\lambda_{i,1}^{(2)}, \ldots, \lambda_{i,k_2}^{(2)}) \in \mathbb{F}_q^{k_2}$, such that for every degree d,

$$d \in \mathrm{Deg}(\mathcal{F}_2) \quad \Leftrightarrow \quad \forall 1 \le i \le t_2 \quad \sum_{j=1}^{k} \lambda_{ij}^{(2)} (\alpha_j^{(2)} x + y)^d = 0.$$

The proof will follow from a counting argument. We will use the fact that any shift of any local constraint defines another viable local constraint. We will then count the number of elements in $\mathrm{Border}(\mathcal{F})$ and show that each of them can satisfy only a certain number of constraints. Then, we will use the abundance of k-local constraints for $\mathcal{F}$ to prove that we can take the orbit of one of them to get a single orbit characterization of $\mathcal{F}$. The counting will be algebraic and will rely on the ability to view a 'bad' constraint as a root of a certain polynomial of a not too high degree. The Schwartz-Zippel lemma (Lemma 2.3) then guarantees the the polynomial does not have too many roots and hence most constraints are 'good'.

For any degree d, $1 \le i_1 \le t_1$ and $1 \le i_2 \le t_2$, consider the formal polynomial

$$\begin{aligned} p_{i_1,i_2,d}(x_1, x_2, y_1, y_2) &= \sum_{j_1=1}^{k_1} \sum_{j_2=1}^{k_2} \lambda_{i_1,j_1}^{(1)} \lambda_{i_2,j_2}^{(2)} (\alpha_{j_1}^{(1)} x_1 + y_1)^d (\alpha_{j_2}^{(2)} x_2 + y_2)^d \\ &= \left(\sum_{j_1=1}^{k_1} \lambda_{i_1,j_1}^{(1)} (\alpha_{j_1}^{(1)} x_1 + y_1)^d \right) \left(\sum_{j_2=1}^{k_2} \lambda_{i_2,j_2}^{(2)} (\alpha_{j_2}^{(2)} x_2 + y_2)^d \right). \end{aligned}$$

Note that $\mathrm{Deg}(\mathcal{F}) = \mathrm{Deg}(\mathcal{F}_1) \cup \mathrm{Deg}(\mathcal{F}_2)$, so that for any d, it is straightforward to see that

$$d \in \mathrm{Deg}(\mathcal{F}) \quad \Leftrightarrow \quad \forall i_1, i_2 \quad p_{i_1,i_2,d}(x_1, x_2, y_1, y_2) = 0. \tag{3}$$

For any assignment $(a_1, a_2, b_1, b_2) \in \mathbb{F}_{q^n}^{*2} \times \mathbb{F}_{q^n}^2$ of the variables (x_1, x_2, y_1, y_2), we know from equation (3) that the set of points $\{\gamma_{i_1,i_2,j_1,j_2} = (\alpha_{i_1,j_1}^{(1)} a_1 + b_1)(\alpha_{i_2,j_2}^{(2)} a_2 + b_2)\}_{i_1,i_2,j_1,j_2}$ and scalars $\{\lambda_{i_1,j_1}^{(1)} \lambda_{i_2,j_2}^{(2)}\}_{i_1,i_2,j_1,j_2}$ satisfy

$$\forall d \in \mathrm{Deg}(\mathcal{F}),\ p_{i_1,i_2,d}(a_1, a_2, b_1, b_2) = \sum_{j_1,j_2} \lambda_{i_1,j_1}^{(1)} \lambda_{i_2,j_2}^{(2)} \gamma_{i_1,i_2,j_1,j_2}^d = 0.$$

By Condition 4 in Lemma 3.2, to give a κ-single orbit characterization of $\mathcal{F}$ it is enough to find an assignment $(a_1, a_2, b_1, b_2) \in \mathbb{F}_{q^n}^{*2} \times \mathbb{F}_{q^n}^2$ satisfying

$$\forall d \in \mathrm{Border}(\mathcal{F}),\ p_{i_1,i_2,d}(a_1, a_2, b_1, b_2) = \sum_{j_1,j_2} \lambda_{i_1,j_1}^{(1)} \lambda_{i_2,j_2}^{(2)} \gamma_{i_1,i_2,j_1,j_2}^d \neq 0.$$

For any (i_1, i_2, d), define the event $\mathrm{BAD}_{i_1,i_2,d}$, over $(a_1, a_2, b_1, b_2) \in \mathbb{F}_{q^n}^{*2} \times \mathbb{F}_{q^n}^2$, as the event that $p_{i_1,i_2,d}(a_1, a_2, b_1, b_2) = 0$ and let

$$\mathrm{BAD} = \bigcup_{d \in \mathrm{Border}(\mathcal{F})} \bigcup_{i_1=1}^{t_1} \bigcup_{i_2=1}^{t_2} \mathrm{BAD}_{i_1,i_2,d}.$$

We would like to upper bound Pr[BAD], where the probability is taken over a uniformly random choice of $(a_1, a_2, b_1, b_2) \in \mathbb{F}_{q^n}^{*2} \times \mathbb{F}_{q^n}^2$. By the union bound,

$$\begin{aligned}\Pr[\text{BAD}] &\leq \sum_{\substack{d \in \text{Border}(\mathcal{F}) \\ 1 \leq i_1 \leq t_1, 1 \leq i_2 \leq t_2}} \Pr[\text{BAD}_{i_1,i_2,d}] \\ &\leq |\text{Border}(\mathcal{F})| \cdot t_1 \cdot t_2 \cdot \max_{\substack{d \in \text{Border}(\mathcal{F}) \\ i_1,i_2}} \Pr[\text{BAD}_{i_1,i_2,d}] \\ &\leq |\text{Border}(\mathcal{F})| \cdot k^2 \cdot \max_{\substack{d \in \text{Border}(\mathcal{F}) \\ i_1,i_2}} \Pr[\text{BAD}_{i_1,i_2,d}],\end{aligned} \tag{4}$$

where we used the simple observation that $t_1 \leq k_1 \leq k$ and $t_2 \leq k_2 \leq k$. The rest of the proof follows from three claims.

Claim 3.7. *For all $d \in \text{Border}(\mathcal{F})$, $\text{wt}(d) \leq k$.*

Proof. The proof follows from (our) Lemma 3.5 and from Theorem 3.6 in [BS10]. □

Claim 3.8. *For all i_1, i_2 and $d' \in \text{Shift}(d)$, $BAD_{i_1,i_2,d} \Leftrightarrow BAD_{i_1,i_2,d'}$.*

Proof. Let d' be a shift of d, so that $d' \equiv q^\ell \cdot d \,(\text{mod}^*\, q^n - 1)$ for some ℓ. $\text{BAD}_{i_1,i_2,d}$ is the event that $p_{i_1,i_2,d}(a_1, a_2, b_1, b_2) = \sum_{j_1,j_2} \lambda^{(1)}_{i_1,j_1} \lambda^{(2)}_{i_2,j_2} \gamma^d_{i_1,i_2,j_1,j_2} = 0$. But if $\text{BAD}_{i_1,i_2,d}$ holds, then we have that $p_{i_1,i_2,d}(a_1, a_2, b_1, b_2)^{q^\ell} = 0$, so that

$$\sum_{j_1,j_2} (\lambda^{(1)}_{i_1,j_1})^{q^\ell} (\lambda^{(2)}_{i_2,j_2})^{q^\ell} \gamma^{q^\ell d}_{i_1,i_2,j_1,j_2} = \sum_{j_1,j_2} \lambda^{(1)}_{i_1,j_1} \lambda^{(2)}_{i_2,j_2} \gamma^{d'}_{i_1,i_2,j_1,j_2} = 0,$$

which holds since $\lambda^{(1)}_{i_1,j_1}$ and $\lambda^{(2)}_{i_2,j_2}$ are in the base field, and since $\gamma^{q^\ell d}_{i_1,i_2,j_1,j_2} = \gamma^{q^\ell d \,(\text{mod}^*\, q^n - 1)}_{i_1,i_2,j_1,j_2} = \gamma^{d'}_{i_1,i_2,j_1,j_2}$. But this is exactly the event $\text{BAD}_{i_1,i_2,d'}$, so that $\text{BAD}_{i_1,i_2,d} \Rightarrow \text{BAD}_{i_1,i_2,d'}$. The converse is proven similarly by noting that $d = q^{n-\ell} \cdot d' \,(\text{mod}^*\, q^n - 1)$. □

Claim 3.9. *For any d of weight $\text{wt}(d) = w$, there exists $d' \in \text{Shift}(d)$ satisfying*

$$d' < q \cdot q^{n(1-\frac{1}{w})}.$$

Proof. The proof of the claim relies on the following fact: in a string of (p-ary) digits of length m and weight w, there exists a substring consisting of at least $(\lceil m/w \rceil - 1)$ consecutive zeroes. Now, if we represent d in base q then it also has weight at most k. We can now shift this q-ary representation of d by looking at some shift $q^i d \,(\text{mod}^*\, q^n - 1)$ such that the last $(\lceil n/w \rceil - 1)$ q-ary digits in its representation are zeros. Letting $d' = q^i d \,(\text{mod}^*\, q^n - 1)$ the claim follows. □

By Claim 3.7, we can upper-bound $|\text{Border}(\mathcal{F})|$ by the number of degrees of weight at most k. If we denote $q = p^s$ then a rough upper bound is $\binom{ns+k}{k}$. To find an upper bound for $\max_{\substack{d \in \text{Border}(\mathcal{F}) \\ i_1,i_2}} \Pr[\text{BAD}_{i_1,i_2,d}]$, first note that for any $d \in \text{Border}(\mathcal{F})$, the polynomial $p_{i_1,i_2,d}(x_1, x_2, y_1, y_2)$ is not identically zero, by equation (3). We can thus upper-bound,

by the Schwartz-Zippel lemma (Lemma 2.3), the probability that (a_1, a_2, b_1, b_2) is a root of $p_{i_1,i_2,d}(x_1, x_2, y_1, y_2)$ as

$$\Pr[\mathrm{BAD}_{i_1,i_2,d}] \leq \frac{\deg(p_{i_1,i_2,d})}{q^n - 1} \leq \frac{2d}{q^n - 1}.$$

Claim 3.8 allows us to upper-bound $\Pr[\mathrm{BAD}_{i_1,i_2,d}]$, for any d, by the best upper bound Schwartz-Zippel gives us for any d' in the orbit of d. In particular,

$$\Pr[\mathrm{BAD}_{i_1,i_2,d}] \leq \min_{d' \in \mathrm{Shift}(d)} \frac{2d'}{q^n - 1} < \frac{2qq^{-\frac{n}{wt(d)}}}{1 - q^{-n}},$$

where the second inequality follows from Claim 3.9. Finally, using Claim 3.7, we have that for any $d \in \mathrm{Border}(\mathcal{F})$,

$$\Pr[\mathrm{BAD}_{i_1,i_2,d}] \leq \frac{2qq^{-\frac{n}{k}}}{1 - q^{-n}},$$

since $\mathrm{wt}(d) \leq k$. We can finally upper bound $\Pr[\mathrm{BAD}]$, assuming $n \geq k$, as

$$\Pr[\mathrm{BAD}] \leq \frac{2qq^{-\frac{n}{k}}}{1 - q^{-n}} \binom{ns + k}{k} \leq \frac{2qq^{-\frac{n}{k}}}{1 - q^{-n}} \left(\frac{e(sn + k)}{k}\right)^k,$$

where e is the natural logarithm. Easy calculation show that for $n \geq 10k^2 \log_q(k) + 10$ it holds that $\Pr[\mathrm{BAD}] < 1$.

Therefore, for $n \geq 10k^2 \log_q(k) + 10$, there exists $a_1, a_2 \in \mathbb{F}_{q^n}^*$ and $b_1, b_2 \in \mathbb{F}_{q^n}$ such that the set of points $\{\gamma_{i_1,i_2,j_1,j_2} = (\lambda^{(1)}_{i_1,j_1} a_1 + b_1)(\lambda^{(2)}_{i_2,j_2} a_2 + b_2)\}_{i_1,i_2,j_1,j_2}$ and scalars $\{\lambda^{(1)}_{i_1,j_1} \lambda^{(2)}_{i_2,j_2}\}_{i_1,i_2,j_1,j_2}$ form a k_1k_2-single orbit characterization for $\mathcal{F}$.

When $n < 10k^2 \log_q(k) + 10$, it is clear that any property can be characterized by reading q^n locations. In other words, the property has a q^n-single orbit local characterization. □

4 Single orbit characterization of Reed-Muller codes

In this section we prove Theorem 1.5. For the sake of readability we restate it.

Theorem 1.5. *Let $q = p^s$ be a prime power. Let w, n be integers such that $w + 1 < \sqrt{\frac{n}{\log_q(3ns)}}$. Denote $w + 1 = r(p - 1) + \ell$, where $0 \leq \ell < p - 1$. Then, the q-ary Reed-Muller family of weight degree w, $\mathrm{RM}_q(w, n)$, has a k-single orbit characterization for $k = p^r \cdot (\ell + 1)$. In particular, for every w, q there exists a $k = k(w, q)$ such that the q-ary Reed-Muller family of weight degree w has a k-single orbit characterization.*

For the case of the binary field we can prove that no matter what w is, $\mathrm{RM}_2(w, n)$ is single orbit characterized with locality 2^{w+1}. Notice that the locality is the same as the guarantee in Theorem 1.5 and the only difference is that w is not restricted.

Theorem 4.1. *For all positive integers w, n, the binary Reed-Muller family of weight degree w $\mathrm{RM}_2(w, n)$ is 2^{w+1}-single orbit.*

We first prove Theorem 1.5. The proof of Theorem 4.1 is given in Section 4.1.

For the proof we will need the following notations and facts from [GSL10].

Given a polynomial f, we define the first derivative along y, denoted $f_{(y,1)}$, as

$$f_{(y,1)}(x) = f(x+y) - f(x) \ .$$

We define the ℓ^{th} derivative along y for $\ell \geq 1$ inductively as

$$f_{(y,\ell)}(x) = f_{(y,\ell-1)}(x+y) - f_{(y,\ell-1)}(x).$$

It is easy to verify that

$$f_{(y,\ell)}(x) = \sum_{0 \leq j \leq \ell} (-1)^{\ell-j} \binom{\ell}{j} f(x+jy) \ . \tag{5}$$

We define multiple derivatives in multiple directions, which we denote by $f_{(y_1,\ell_1),\ldots,(y_{r+1},\ell_{r+1})}(x)$. To derive a formula for those derivatives we define the following quantity for all ℓ, c

$$\mu(\ell, c) = \sum_{0 \leq j \leq \ell} (-1)^{\ell-j} \binom{\ell}{j} j^c \ .$$

The following combinatorial identities are well-known (a proof can be found in [GSL10]).

Fact 4.2 (Fact 3.2 in [GSL10])**.** *Let $\ell \leq p-1$. Then*

$$\begin{aligned} &\mu(\ell, c) = 0 \text{ for } c \in \{0, \ldots, \ell-1\} \ , \\ &\mu(\ell, \ell) \not\equiv 0 \mod p \ . \end{aligned}$$

Hence, for $\ell \leq p-1$,

$$\begin{aligned} x^d_{(y,\ell)} &= \sum_{j=0}^{\ell} (-1)^{\ell-j} \binom{\ell}{j} (x+jy)^d \\ &= \sum_{j=0}^{\ell} (-1)^{\ell-j} \binom{\ell}{j} \sum_{e \leq d} \binom{d}{e} x^{d-e} (jy)^e \\ &=^{(*)} \sum_{e \leq d} \binom{d}{e} x^{d-e} y^e \sum_{j=0}^{\ell} (-1)^{\ell-j} \binom{\ell}{j} j^{\mathrm{wt}(e)} \\ &= \sum_{e \leq d} \binom{d}{e} x^{d-e} y^e \mu(\ell, \mathrm{wt}(e)) \\ &=^{(\dagger)} \sum_{\substack{e \leq d \\ \mathrm{wt}(e) \geq \ell}} \binom{d}{e} x^{d-e} y^e \mu(\ell, \mathrm{wt}(e)) \end{aligned}$$

where in $(*)$ we use the fact that for $0 \leq j < p$, $j^e = j^{\mathrm{wt}(e)}$ (since $j^{p^i} = j$) and in $(\dagger)$ we use $\mu(\ell, \mathrm{wt}(e)) = 0$ for $\mathrm{wt}(e) < \ell \leq p-1$. Thus, differentiating ℓ times along y reduces the weight of the degree of x by at least ℓ, as one would expect.

By repeating this calculation, we can compute an expression for derivatives in multiple "directions". Given vectors $d, e_1, \ldots, e_{r+1}$ we use the notation $\binom{d}{e_1,\ldots,e_{r+1}}$ for the product of multinomials $\binom{d}{e_1}\binom{d-e_1}{e_2}\ldots$. We have

$$x^d_{(y_1,\ell_1),\ldots,(y_{r+1},\ell_{r+1})} = \sum_{e_1+\cdots+e_{r+1}\le d} \binom{d}{e_1,\ldots,e_{r+1}} x^{d-(e_1+\cdots+e_{r+1})} \cdot \prod_{j=1}^{r+1} \mu(\ell_j, \mathrm{wt}(e_j)) y_j^{e_j} = \sum_{\mathrm{wt}(e_1)\ge\ell_1,\ldots,\mathrm{wt}(e_{r+1})\ge\ell_{r+1}} \binom{d}{e_1,\ldots,e_{r+1}} x^{d-(e_1+\cdots+e_{r+1})} \cdot \prod_{j=1}^{r+1} \mu(\ell_j, \mathrm{wt}(e_j)) y_j^{e_j} \,. \tag{6}$$

The following lemma is immediate given Equation (6).

Lemma 4.3 (Lemma 3.5 in [GSL10])**.** *Let*

$$\mathrm{wt}(d) = r(p-1)+\ell \quad \textit{where} \quad \ell \le p-1 \,,$$
$$\ell_1 = \cdots = \ell_r = p-1 \text{ and } \ell_{r+1} = \ell \,.$$

Then $(x^d)_{(y_1,\ell_1),\ldots,(y_{r+1},\ell_{r+1})}$ *is a non-zero polynomial in* $y_1, \ldots, y_{r+1}$.

Proof. By Equation (6), the only terms contributing to $(x^d)_{(y_1,\ell_1),\ldots,(y_{r+1},\ell_{r+1})}$, must come from terms such that

$$\mathrm{wt}(e_1) = \cdots = \mathrm{wt}(e_r) = p-1 \quad \text{and} \quad \mathrm{wt}(e_{r+1}) = \ell \,,$$

as otherwise $d - (e_1 + \cdots + e_{r+1}) < 0$. In other words, for our choice of $\ell_1, \ldots, \ell_{r+1}$ (recalling $\mathrm{wt}(d) = r(p-1)+\ell$) we have that

$$x^d_{(y_1,\ell_1),\ldots,(y_{r+1},\ell_{r+1})} = (\mu(p-1,p-1)^r \cdot \mu(\ell,\ell)) \cdot \sum_{\substack{\mathrm{wt}(e_1)=p-1,\ldots,\mathrm{wt}(e_r)=p-1 \\ \mathrm{wt}(e_{r+1})=\ell}} \binom{d}{e_1,\ldots,e_{r+1}} \prod_{j=1}^{r+1} y_j^{e_j} \,. \tag{7}$$

Furthermore, for any such choice of the e_i, the coefficient of $\prod_{j=1}^{r} y_j^{e_j}$ is nonzero. This follows from Fact 4.2 and Lucas' theorem, noticing that two different vectors $(e_1, \ldots, e_{r+1})$ and $(e'_1, \ldots, e'_{r+1})$ give rise to different monomials. □

We can now give the proof of Theorem 1.5. The idea is that we can find directions such that taking derivatives in those directions (of the right order) gives a nonzero value for any monomial x^d, where $\mathrm{wt}(d) = w+1$, and a zero value for any monomial of smaller weight.

Proof of Theorem 1.5. Since the degree set of $\mathrm{RM}_q(w,n)$ is $S = \{d \mid \mathrm{wt}(d) \le w\}$ it follows that $\mathrm{Border}(S) = \{d \mid \mathrm{wt}(d) = w+1\}$. By Lemma 3.2 it suffices to consider degrees $d \in \mathrm{Border}(S)$. So consider $\mathrm{Shift}(d)$ for an arbitrary $d \in \mathrm{Border}(S)$. By Claim 3.9, we can assume w.l.o.g. that d is such that $d \le q \cdot q^{n(1-\frac{1}{w+1})}$. From Lemma 4.3, we get by the Schwartz-Zippel lemma (Lemma 2.3) that if we pick the y_j at random, then the probability that $(x^d)_{(y_1,p-1),\ldots,(y_r,p-1),(y_{r+1},\ell)} = 0$ is at

most d/q^n. Thus, the probability is at most $q \cdot q^{-\frac{n}{w+1}}$. As there are at most[6] $\binom{ns+w}{w+1}$ monomials of weight $w+1$ (recall $q=p^s$) it follows that if $\binom{ns+w}{w+1}q^{-\frac{n}{w+1}} < 1$ then there is a choice of y_j for which $(x^d)_{(y_1,p-1),\ldots,(y_r,p-1),(y_{r+1},\ell)} \neq 0$, for any $\mathrm{wt}(d) = w+1$. Using the estimate $\binom{a}{b} \leq (ea/b)^b$, where here e is the natural logarithm, it follows that if $w+1 < \sqrt{n/\log_q(3ns)}$ then this probability is indeed smaller than 1.

By the definition of the derivative, it is also clear that taking a derivative in directions $(y_1, p-1), \ldots, (y_r, p-1), (y_{r+1}, \ell)$ and evaluating at, say $x=0$, amounts to evaluating x^d on $p^r \cdot (\ell+1)$ different points, where $w+1 = r(p-1)+\ell$ $(0 \leq \ell < p-1)$. Indeed, using the fact that modulo p it holds that $\binom{p-1}{j} = (-1)^j$, we conclude from (5) that

$$\begin{aligned}
&(x^d)_{(y_1,p-1),\ldots,(y_r,p-1),(y_{r+1},\ell)}(0) \\
&= \sum_{\substack{0\leq j_1,\ldots,j_r\leq p-1 \\ 0\leq j_{r+1}\leq \ell}} \left(\prod_{i=1}^{r}(-1)^{(p-1)-j_i}\binom{p-1}{j_i}\right)\cdot\left((-1)^{\ell-j_{r+1}}\binom{\ell}{j_{r+1}}\right)\cdot\left(\sum_{i=1}^{r+1} j_i y_i\right)^d \\
&= \sum_{\substack{0\leq j_1,\ldots,j_r\leq p-1 \\ 0\leq j_{r+1}\leq \ell}} \left(\prod_{i=1}^{r}(-1)^{(p-1)-j_i}(-1)^{j_i}\right)\cdot\left((-1)^{(p-1)-j_{r+1}}\binom{\ell}{j_{r+1}}\right)\cdot\left(\sum_{i=1}^{r+1} j_i y_i\right)^d \\
&= (-1)^{(r+1)(p-1)} \cdot \sum_{\substack{0\leq j_1,\ldots,j_r\leq p-1 \\ 0\leq j_{r+1}\leq \ell}} (-1)^{j_{r+1}}\binom{\ell}{j_{r+1}}\cdot\left(\sum_{i=1}^{r+1} j_i y_i\right)^d .
\end{aligned}$$

It is also clear that if $\mathrm{wt}(d) \leq w$, then $(x^d)_{(y_1,p-1),\ldots,(y_r,p-1),(y_{r+1},\ell)} = 0$ (e.g., by (6), since in this case, $d < e_1 + \ldots + e_{r+1}$).

As a conclusion we get that the Reed-Muller code for degree w is $p^r \cdot (\ell+1)$-single orbit characterized, as required.

□

4.1 Binary Reed-Muller is Single-Orbit

Here we prove that binary Reed-Muller codes are single-orbit characterizable, without restricting the degree w. The proof is similar in nature to the proof of Theorem 1.5 only now we give a better analysis of (7) and notice that it is a determinant which makes the work of finding the good directions easier.

Proof of Theorem 4.1. When $p=2$ notice that since $w+1 = r(p-1)+\ell$, for $0 \leq \ell < p-1$ it follows that $\ell = 0$ and $w+1 = r$. Let $d \in S$ satisfy $\mathrm{wt}(d) = w+1 = r$. Then (7) above reduces to

$$x^d_{(y_1,\ell_1),\ldots,(y_{r+1},\ell_{r+1})} = \sum_{\substack{\forall i\ \mathrm{wt}(e_i)=1 \\ e_1+\ldots+e_{w+1}\leq_2 d}} \prod_{j=1}^{w+1} y_j^{e_j} . \tag{8}$$

[6] Here we lose a lot. E.g., if $p=2$ then there are at most $\binom{n}{w+1}$ monomials of weight $w+1$.

Let $d = d_1 + \ldots + d_{w+1}$ where $d_i = 2^{j_i}$ for some $0 \leq j_i \leq n-1$. Then notice that

$$x^d_{(y_1,\ell_1),\ldots,(y_{r+1},\ell_{r+1})} = \sum_{\substack{\forall i \, \mathrm{wt}(e_i)=1 \\ e_1+\ldots+e_{w+1} \leq_2 d}} \prod_{j=1}^{w+1} y_j^{e_j} = \det \begin{vmatrix} y_1^{d_1} & y_1^{d_2} & \ldots & y_1^{d_{w+1}} \\ y_2^{d_1} & y_2^{d_2} & \ldots & y_2^{d_{w+1}} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & y_i^{d_j} & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ y_{w+1}^{d_1} & \ldots & \ldots & y_{w+1}^{d_{w+1}}. \end{vmatrix} \tag{9}$$

We now show how to pick $y_1, \ldots, y_k$ so that the determinant above is nonzero. Let $y_i = y^i$ for some $y \in \mathbb{F}_{2^n}$. This implies that the matrix above is a Vandermonde matrix and its determinant is nonzero whenever $y^{d_i} \neq y^{d_j}$ for all $i, j \in [w+1]$. Notice that if $i \leq j$ then $y^{2^i} = y^{2^j}$ iff $y^{2^i(2^{j-i}-1)} = 1$ iff y belongs to a subfield $\mathbb{F}_{2^{j-i}}$ of $\mathbb{F}_{2^n}$. Choosing y to be a primitive element[7] of $\mathbb{F}_{2^n}$ ensures that it does not belong to any subfield and hence the determinant in Equation (9) is nonzero for every $d = 2^{j_1} + \ldots + 2^{j_{w+1}} \in \mathrm{Border}(S)$, where $0 \leq j_i \leq n-1$.

This concludes the proof. □

5 Consequences, Questions and Conjectures

Our work further highlights the role played by single-orbit characterizations in the testing of affine-invariant properties. This feature is common (e.g. Reed-Muller property is single-orbit over the smaller group) and also useful (sums of single-orbit characterized properties also have this feature).

At the moment almost all known locally-testable affine-invariant properties are known to be single-orbit characterized. The only exception is the case of sparse properties where the range is not a prime field. This leads to the following question, which we hope can be resolved affirmatively (soon).

Question 5.1. *For every q and t, does there exists a constant $k = k(q,t)$ such that every t-sparse property $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is k-single orbit characterized?*

Assuming an affirmative answer to the questions above, we get a "concise" description of all known testable properties.

5.1 Known locally testable properties

As mentioned earlier, the known "basic" single-orbit characterized affine-invariant families are the Reed-Muller families and sparse families. Three "operations" are also now known that preserve "single-orbit characterizations" and hence local testability of these basic families: (1) Sums of two families, (2) Intersections of two families, and (3) Lift of a single family. Below we define this lifting operator.

[7] y is a primitive element of $\mathbb{F}_{2^n}$ if $\mathbb{F}^*_{2^n} = \{1, y, y^2, \ldots, y^{2^n-2}\}$

Definition 5.2 (Lifted code [BMSS10])**.** *Let $\mathbb{K} \supsetneq \mathbb{L} \supsetneq \mathbb{F}_q$ be finite fields with $q = p^s$. For $D \subseteq \{0, \ldots, |\mathbb{L}| - 1\}$ we define the* lift *of D from $\mathbb{L}$ to $\mathbb{K}$ to be the set of integers*

$$\mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(D) = \left\{ d' \in \{0, \ldots, |\mathbb{K}| - 1\} \;\middle|\; \left(\mathrm{shadow}_p(d') \,(mod^* |\mathbb{L}| - 1)\right) \subseteq D \right\}.$$

For an affine-invariant family $\mathcal{F} \subseteq \{\mathbb{L} \to \mathbb{F}_q\}$ with degree set $D = \mathrm{Deg}(\mathcal{F})$, let $\mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(\mathcal{F})$ be the affine-invariant family with degree set $\mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(D)$, i.e.,

$$\mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(\mathcal{F}) = \{f : \mathbb{K} \to \mathbb{F}_q \mid \mathrm{supp}(f) \subseteq \mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(D)\} = \mathrm{Fam}(\mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(D)).$$

The following proposition follows easily from the definitions

Proposition 5.3 ([BMSS10])**.** *Lifts of single orbit characterized families are also single-orbit characterized. Specifically, if $\mathbb{F}_q \subseteq \mathbb{L} \subseteq \mathbb{K}$ and $(\bar{\alpha}, \{\bar{\lambda_i}\}_{i=1}^t)$ is a k-single orbit characterization of $\mathcal{F} \subseteq \{\mathbb{L} \to \mathbb{F}_q\}$ then $(\bar{\alpha}, \{\bar{\lambda_i}\}_{i=1}^t)$ is also k-single orbit characterization of $\mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(\mathcal{F})$.*

Given the operations above, it is easy to see that one can compose a finite number of basic single-orbit characterized families using a "formula" whose operations are sum, intersection and lifts. We define this concept below.

Definition 5.4 (Formula, size)**.** *A formula Φ of size s, degree d, sparsity t producing a family $\mathcal{F} \subseteq \{\mathbb{K} \to \mathbb{F}_q\}$, denoted $(s, d, t, \mathbb{K}, \mathbb{F})$-formula, is given by the following inductive definition:*

1. *A formula Φ of size 1, is given by $\mathcal{F} \subseteq \{\mathbb{K} \to \mathbb{F}_q\}$ where $\mathcal{F}$ is either a Reed-Muller family of order d, or a t-sparse family.*
2. *A formula of size s is obtained by one of the following operations:*
   (a) *Picking $\mathbb{L}$ such that $\mathbb{F}_q \subseteq \mathbb{L} \subseteq \mathbb{K}$ and letting $\Phi = \mathrm{lift}_{\mathbb{L} \nearrow \mathbb{K}}(\Phi_1)$ where Φ_1 is a $(s - 1, t, d, \mathbb{L}, \mathbb{F})$ formula.*
   (b) *Picking s_1, s_2 such that $s_1 + s_2 + 1 = s$ and letting $\Phi = \Phi_1 \cap \Phi_2$ where Φ_i is an $(s_i, t, d, \mathbb{K}, \mathbb{F})$ formula.*
   (c) *Picking s_1, s_2 such that $s_1 + s_2 + 1 = s$ and letting $\Phi = \Phi_1 + \Phi_2$ where Φ_i is an $(s_i, t, d, \mathbb{K}, \mathbb{F})$ formula.*

The following theorem summarizes the state of knowledge of single-orbit characterized families.

Theorem 5.5. *For every s, t, d, q there exists a $k = k(s, t, d, q)$ such that for every n, every $(s, t, d, \mathbb{F}_{q^n}, \mathbb{F}_q)$-formula produces a k-single orbit characterized family, for prime q.*

Note that the caveat that q is prime can be dropped if we have an affirmative answer to Question 5.1.

5.2 Conjectures and questions

We start with the most obvious question.

Question 5.6. *Is the following statement true? For every k, q there exist s, t, d such that for every n, if $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is a k-locally testable affine-invariant family then $\mathcal{F}$ is given by an $(s, t, d, \mathbb{F}_{q^n}, \mathbb{F}_q)$-formula.*

At the moment our understanding of affine-invariance with respect to its local testability is so far that it is too optimistic to conjecture an affirmative answer to this question. All we can say is that an affirmative answer is not yet ruled out.

The nature of the question seems to become much simpler if we disallow lifts, by insisting that n is prime (then we get no fields $\mathbb{L}$ strictly between $\mathbb{F}_q$ and $\mathbb{F}_{q^n}$). In this setting, intersections become uninteresting and lead to a much tamer question.

Question 5.7. *Is the following statement true? For every k, q there exist t, d such that for every prime n, if $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is a k-locally testable affine-invariant family then $\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2$ where $\mathcal{F}_1 = \mathrm{RM}_q(d', n)$ and $\mathcal{F}_2$ is t'-sparse, for some $d' \leq d$ and $t' \leq t$.*

This question remains quite challenging even when we restrict to the case where $q = 2$ (where our state of understanding does seem somewhat better), and even when we restrict our families to be contained in $\mathrm{RM}_2(2, n)$.

Conjecture 5.8. *For every k there exists a t such that the following holds for every prime n: If $\mathcal{F} \subsetneq \mathrm{RM}_2(2, n)$ is k-locally testable then $\mathcal{F}$ is t-sparse.*

Attempting to prove the conjecture above leads to some interesting questions about the rank of certain Vandermonde like matrices that seem interesting in their own right. We state the conjecture below. We don't prove the connection to the conjecture above, but claim that an affirmative answer to the following implies an affirmative answer to the above.

Conjecture 5.9. *For every k, there exists a t such that for every prime n and every sequence $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_{2^n}$ of elements that are $\mathbb{F}_2$-linearly independent, and every sequence of t distinct elements $e_1, \ldots, e_t \in \{0, \ldots, n-1\}$, the $k \times t$ matrix $M = [M_{ij}]_{ij}$ with $M_{ij} = \alpha_i^{2^{e_j}}$ has rank exactly k.*

Finally a couple of questions which relate to the structure of locally-testable codes (an affirmative answer to both is implied by an affirmative answer to Question 5.6).

Question 5.10. *For every k, q does there exist a $\tilde{k}$ such that for every n, if $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is k-locally testable, then $\mathcal{F}$ has a $\tilde{k}$-single orbit characterization?*

Question 5.11. *For every k, q does there exist a $\tilde{k}$ such that for every n, if $\mathcal{F}_1, \mathcal{F}_2 \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ are k-locally testable, then $\mathcal{F}_1 + \mathcal{F}_2$ is $\tilde{k}$-locally testable?*

Acknowledgments

We would like to thank Shripad Garge for permission to include his proof of Lemma 2.17, and Neeraj Kayal and Dipendra Prasad for directing us to Shripad.

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 *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 251–260. ACM, 2006.

[AKK+05] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing Reed-Muller codes. *IEEE Transactions on Information Theory*, 51(11):4032–4039, 2005.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, May 1998.

[Art55] Emil Artin. The orders of the linear groups. *Communications on Pure and Applied Mathematics*, 8(3):355–365, August 1955.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, January 1998.

[BCL+06] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, Balázs Szegedy, and Katalin Vesztergombi. Graph limits and parameter testing. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 261–270. ACM, 2006.

[BCSX09] Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie. Testing linear-invariant non-linear properties. In *Proceedings of the 26th International Symposium on Theoretical Aspects of Computer Science, STACS 2009, February 26-28, 2009, Freiburg, Germany*, pages 135–146, 2009. Full version at http://www.eccc.uni-trier.de/report/2008/088/.

[BDWY10] Boaz Barak, Zeev Dvir, Avi Wigderson, and Amir Yehudayoff. Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:149, 2010. To appear in STOC 2011.

[BGS10] Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira. A unified framework for testing linear-invariant properties. In *Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 478–487. IEEE Computer Society, 2010.

[BIW10] Omer Barkol, Yuval Ishai, and Enav Weinreb. On locally decodable codes, self-correctable codes, and t-private PIR. *Algorithmica*, 58:831–859, 2010.

[BMSS10] Eli Ben-Sasson, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. Symmetric LDPC codes are not necessarily locally testable. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:199, 2010. To appear in CCC 2011.

[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.

[CGKS98] Benny Chor, Oded Goldreich, Eyal Kushilevitz, and Madhu Sudan. Private information retrieval. *Journal of the ACM*, 45(6):965–981, 1998.

[CST11] Gil Cohen, Amir Shpilka, and Avishay Tal. On the degree of univariate polynomials over the integers. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:2, 2011.

[Gar11] Shripad Garge. Personal Communication, April 2011.

[GGR98] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, 45(4):653–750, 1998.

[GKS08] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. 2-transitivity is insufficient for local testability. In *Proceedings of the 23rd Annual IEEE Conference on Computational Complexity, CCC 2008, 23-26 June 2008, College Park, Maryland, USA*, pages 259–267. IEEE Computer Society, 2008.

[GKS09] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. Succinct representation of codes with applications to testing. In Irit Dinur, Klaus Jansen, Joseph Naor, and José D. P. Rolim, editors, *APPROX-RANDOM*, volume 5687 of *Lecture Notes in Computer Science*, pages 534–547. Springer, 2009.

[GSL10] Parikshit Gopalan, Amir Shpilka, and Shachar Lovett. The complexity of Boolean functions in different characteristics. *Computational Complexity*, 19(2):235–263, 2010.

[JPRZ09] Charanjit S. Jutla, Anindya C. Patthak, Atri Rudra, and David Zuckerman. Testing low-degree polynomials over prime fields. *Random Struct. Algorithms*, 35(2):163–193, 2009.

[KL05] Tali Kaufman and Simon Litsyn. Almost orthogonal linear codes are locally testable. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2005), 23-25 October 2005, Pittsburgh, PA, USA*, 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.

[KR06] T. Kaufman and D. Ron. Testing polynomials over general fields. *SIAM J. on Computing*, 36(3):779–802, 2006.

[KS07] Tali Kaufman and Madhu Sudan. Sparse random linear codes are locally decodable and testable. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA*, pages 590–600. IEEE Computer Society, 2007.

[KS08] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In Cynthia Dwork, editor, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 403–412. ACM, 2008.

[KS10] Swastik Kopparty and Shubhangi Saraf. Local list-decoding and testing of random linear codes from high error. In Leonard J. Schulman, editor, *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 417–426. ACM, 2010.

[KSV08] Daniel Král', Oriol Serra, and Lluís Vena. A removal lemma for systems of linear equations over finite fields. *Israel Journal of Mathematics (to appear)*, 2008. Preprint available at http://arxiv.org/abs/0809.1846.

[KT00] Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, STOC '00, pages 80–86, New York, NY, USA, 2000. ACM.

[MR95] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[RS96] Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM J. on Computing*, 25(2):252–271, 1996.

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

[Sha10] Asaf Shapira. A proof of Green's conjecture regarding the removal properties of sets of linear equations. *Journal of the London Math Society*, 81(2):355–373, February 2010.

[Sud10] Madhu Sudan. Invariance in property testing. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:51, 2010.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and algebraic computation*, pages 216–226. 1979.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de