Algebraic Property Testing: The Role of Invariance

Tali Kaufman* Madhu Sudan†

November 2, 2007

Abstract

We argue that the symmetries of a property being tested play a central role in property testing. We support this assertion in the context of algebraic functions, by examining properties of functions mapping a vector space $\mathbb{K}^n$ over a field $\mathbb{K}$ to a subfield $\mathbb{F}$. We consider $\mathbb{F}$-linear properties that are invariant under linear transformations of the domain and prove that an $O(1)$-local "characterization" is a necessary and sufficient condition for $O(1)$-local testability when $|\mathbb{K}| = O(1)$. (A local characterization of a property is a definition of a property in terms of local constraints satisfied by functions exhibiting a property.) For the subclass of properties that are invariant under affine transformations of the domain, we prove that the existence of a *single* $O(1)$-local constraint implies $O(1)$-local testability. These results generalize and extend the class of algebraic properties, most notably linearity and low-degree-ness, that were previously known to be testable. In particular, the extensions include properties satisfied by functions of degree linear in n that turn out to be $O(1)$-locally testable.

Our results are proved by introducing a new notion that we term "formal characterizations". Roughly this corresponds to characterizations that are given by a single local constraint and its permutations under linear transformations of the domain. Our main testing result shows that local formal characterizations essentially imply local testability. We then investigate properties that are linear-invariant and attempt to understand their local formal characterizability. Our results here give coarse upper and lower bounds on the locality of constraints and characterizations for linear-invariant properties in terms of some structural parameters of the property we introduce. The lower bounds rule out any characterization, while the upper bounds give formal characterizations. Combining the two gives a test for all linear-invariant properties with local characterizations.

We believe that invariance of properties is a very interesting notion to study in the context of property testing, in general and merits a systematic study. In particular, the class of linear-invariant and affine-invariant properties exhibits a rich variety among algebraic properties and offer better intuition about algebraic properties than the more limited class of low-degree functions.

*IAS, Princeton. kaufmant@mit.edu.

†MIT CSAIL. madhu@mit.edu. Research supported in part by NSF Award CCR 0514915.

1 Introduction

Property testing considers the task of testing efficiently, by random sampling, if a function mapping a finite domain to a finite range "essentially" satisfies a given property. The property to be tested can be specified by the family of functions $\mathcal{F}$ that possess the property. A property $\mathcal{F}$ is *k-locally testable* if there exists a randomized test that queries the value of a function f on k inputs and accepts $f \in \mathcal{F}$ and rejects $f \notin \mathcal{F}$ with probability lower bounded by a quantity proportional to the distance of f from $\mathcal{F}$. Proximity of functions is measured in terms of its relative Hamming distance $\delta(f,g) = \Pr_x[f(x) \neq g(x)]$ when x is chosen uniformly from the finite domain. A function f is δ-close to $\mathcal{F}$ if there exists a $g \in \mathcal{F}$ such that $\delta(f,g) \leq \delta$ and δ-far otherwise.

The study of property testing emerged in the wake of the linearity test of Blum, Luby, and Rubinfeld [4] and was defined formally in Rubinfeld and Sudan [19]. The first substantial investigation of property testing occurred in Goldreich, Goldwasser, and Ron [10] who focussed on the testing of properties of combinatorial objects, in particular of graphs. Subsequent works have lead to major strides in the testing of graph properties culminating with the works of Alon et al. and Borgs et al. [1, 6]. The testing of algebraic properties has also seen significant progress since [4, 19] including testing of functions satisfying functional equations [18], and testing of various algebraic properties leading to error-correcting codes e.g. testing of Reed-Muller codes [2], generalized Reed-Muller codes [16, 13], dual-BCH codes [15]. On the negative side, the works of Bogdanov, Obata, and Trevisan [5] and Ben-Sasson, Harsha,and Raskhodnikova [3] give properties that are not locally testable.

In the light of this progress it is natural to ask: What are the essential features that make a property testable. In the context of graph-property testing (in the "dense-graph" model) this question is answered by the works of [1, 6], who show that a certain feature that they term "regularity" is necessary and sufficient for testing graph properties. In the algebraic setting, a similar understanding of properties that lead to local testability is lacking. In this paper we take some steps to remedy this.

Invariance and Property Testing: Our approach to (algebraic) property testing is to attribute testability to some "invariance" features exhibited by the property. Invariance features of a family $\mathcal{F}$, especially under permutations of the domain, seems naturally linked to property testing. For example, let us consider the test for "majority" (the property $\mathcal{F}$ consisting of all functions $f : \{1,\ldots,N\} \to \{0,1\}$ that take the value 1 at least $N/2$ times). This test is considered uninteresting and we propose a formal explanation. This test actually uses the symmetry of the property $\mathcal{F}$, and the symmetry required is the full group of permutations over the domain. Indeed the test easily extends to any other "symmetric" property $\mathcal{F}$ of Boolean functions, which has the feature that if $f \in \mathcal{F}$ and π is a permutation on the domain, the $f \circ \pi(x) = f(\pi(x))$ is also in $\mathcal{F}$. A formal reason to declare the test "obvious" may be that the group of invariances needed in $\mathcal{F}$ is so large (qualitatively).

Graph property testing similarly revolves around symmetries. This setting consider functions $A : \{1,\ldots,n\} \times \{1,\ldots,n\} \to \{0,1\}$, and properties that are invariant under permutations that permute rows and columns simultaneously. The groups of symmetries thus is somewhat smaller ($(\sqrt{N})!$ as opposed to $N!$, where $N = n^2$ is the domain size). But now one needs some more features (monotonicity/heredity) to get property testers [1, 6]. Despite this natural link between property testing and invariances, this link does not seem to have been explicit in prior literature. We make it explicit here. We remark that in independent work, Goldreich and Sheffet [11], also make this notion explicit, and use it to understand the randomness complexity needs of property testing.

In this paper we explore invariances of an algebraic kind. To do so, we consider functions mapping an n-dimensional vector space over a finite field $\mathbb{K}$ to a subfield $\mathbb{F}$ of $\mathbb{K}$. Among such functions the families $\mathcal{F}$

we consider satisfy two properties:

1. They are $\mathbb{K}$-*linear invariant* (or simply *linear invariant*), i.e., for every function $f \in \mathcal{F}$, and linear map $L : \mathbb{K}^n \to \mathbb{K}^n$ (i.e., a function that satisfies $\alpha L(\mathbf{x}) + \beta L(\mathbf{y}) = L(\alpha\mathbf{x} + \beta\mathbf{y})$ for every $\alpha, \beta \in \mathbb{K}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$), it is the case that $f \circ L$, given by $(f \circ L)(\mathbf{x}) = f(L(\mathbf{x}))$, is also in $\mathcal{F}$. If such a closure holds for all affine maps L from $\mathbb{K}^n$ to $\mathbb{K}^n$, then the property $\mathcal{F}$ is said to be *affine-invariant*.

2. They are $\mathbb{F}$-*linear* (or simply *linear*), i.e., for every pair of functions $f, g \in \mathcal{F}$ and $\alpha, \beta \in \mathbb{F}$ it is the case that the function $\alpha f + \beta g$ is also in $\mathcal{F}$. This is the property that typically leads to linear codes over the alphabet $\mathbb{F}$.

In the algebraic context, linear-invariance over the domain seems to be a natural class of invariances (though not necessarily the only class) to consider, and may be viewed as analogous to the choice of working with "graph-properties". The linearity of the family $\mathcal{F}$ (when viewed as a vector space over the range) is an additional property we impose to derive some testability results (analogous to the role played by heredity/monotonicity in graph property testing).

For simplicity we suppress the use of the phrase "$\mathbb{F}$-linear" in this paper, and use the term linear-invariant (affine-invariant) family to reflect families which are both linear-invariant (resp. affine-invariant) and linear. (We stress that this is merely a notational choice. It maybe quite interesting to study non-linear properties that are linear-invariant also, but we don't do so here.)

The resulting collection of families unify most previously considered in algebraic settings. They include the class of linear functions, low-degree polynomials (and thus generalized Reed-Muller codes), as well as the dual-BCH codes. But they also include other families such as homogenous polynomials of any given degree and linearized polynomials. They satisfy nice closure properties e.g., if $\mathcal{F}_1$ and $\mathcal{F}_2$ are linear-invariant, then so are $\mathcal{F}_1 \cap \mathcal{F}_2$ and $\mathcal{F}_1 + \mathcal{F}_2$, the family that consists of the sum of functions from $\mathcal{F}_1$ and $\mathcal{F}_2$. Finally, we remark that the group of symmetries required by linear-invariance is relatively tiny, and only quasipolynomial in the domain size, compared to the exponential sizes relied upon in the symmetric properties as well as in graph properties.

Our principal results are to show necessary and sufficient conditions for testing linear-invariant families mapping $\mathbb{K}^n$ to $\mathbb{F}$. The results hold for all choices of $\mathbb{K}$ and $\mathbb{F}$ as $n \to \infty$, but are specially strong when $|\mathbb{K}| = O(1)$. We describe our results, and approach, below.

Constraints, Characterizations, Formal Characterizations, and Testing: To understand necessary conditions for local testability, we start by recalling the some basic notions in this context, namely those of "constraints" and "characterizations".

We say that a family $\mathcal{F}$ satisfies a *constraint* $C = (x_1, \ldots, x_k; S)$ where $x_1, \ldots, x_k \in \mathbb{K}^n$ and $S \subsetneq \mathbb{F}^k$ if every member $f \in \mathcal{F}$ satisfies $\langle f(x_1), \ldots, f(x_k) \rangle \in S$. We refer to this constraint as a k-local constraint. In order for a property to be k-locally testable, with one-sided error, it must be the case that functions in the family satisfy some k-local "constraint" (since every rejected function must be rejected with a proof of non-membership in the family). Local constraints also essential for a family of functions to be self-correctible and indeed it turns out that all function families we analyze are self-correctible.

Testable properties where every non-member is rejected with positive probability (as required by our definition of a local test) actually need to show even more structure. Specifically, it must be that there is some set of local constraints that completely *characterize* the family, i.e., $f \in \mathcal{F}$ if and only if it satisfies every one of the given set of k-local constraints. (See Definition 2.1 for a formal definition.) In this paper we will consider all function families that are linear invariant and have a local characterization and show that they are testable.

To derive this result we examine the source of the local characterizability of a family. Local characterizability of a family requires that a family be specified by *several* local constraints. In examining the features that lead to property testing it is natural to ask for an explanation for this abundance of local constraints. One way to explain them is via the invariance features of the family. If a family satisfies one local constraint, then every "permutation" of the domain that preserves membership in the family yields a potentially new local constraint. In our case, thus the abundance of constraints can be explained by the linear invariance of the family. Every linear transformation of a constraint, leads to another valid constraint, and together this set can be quite large. Motivated by this, we introduce the notion of a *formal characterization*, which requires that the family be specified by a *single* constraint and its "orbit", i.e., all the other constraints obtained by linear transformations of the given one, characterize the family. (The actual definition allows a slightly broader class of characterizations, see Definition 2.3.) Modulo the formal definitions of these objects, we can state our first theorem informally as follows:

Main Theorem 1 (Informal): *If a family $\mathcal{F}$ is linear-invariant and has a k-local formal characterization, which satisfies some additional restrictions, then it is k-locally testable.* (See Theorem 2.9 for a formal statement.)

The requirement that a single constraint and its orbit characterize a family may seem overly restrictive, but known characterizations of most algebraic functions including those from [4, 19, 2, 16, 13] are actually formal and satisfy the (thus far unspecified) additional restrictions (see Proposition 2.7). As a result Theorem 2.9 already subsumes many of the algebraic testing results. Moreover, as discussed later in this section, the proof is actually somewhat simpler and unifies the different proofs presented in the literature for the different cases.

Our other main results show that the above theorem actually gives testers for all linear-invariant families provided the family is locally characterizable, a clear necessary condition. For the special case of affine-invariant families, we show that the existence of a *single* local constraint suffices to establish testability. Again we describe these theorems informally below.

Main Theorem 2 (Informal): *If a family $\mathcal{F}$ is affine-invariant and has a k-local constraint, then it has a $k^{\mathrm{poly}(|\mathbb{K}|)}$-local formal characterization which satisfies the additional restrictions mentioned in Main Theorem 1 (Informal). Hence $\mathcal{F}$ is $k^{\mathrm{poly}(|\mathbb{K}|)}$-locally testable.* (See Theorem 2.10 for a formal statement.)

Thus when $|\mathbb{K}| = O(1)$, the above pins down the local testability to with polynomial factors. Moving to the case of linear-invariant families, here we do get local formal characterizations, but they do not satisfy the additional restrictions described in Theorem 2.9. However, we still manage to use the theorem to give a local test for all such families.

Main Theorem 3 (Informal): *If a family $\mathcal{F}$ is linear-invariant and has a k-local characterization, then it has a $k^{\mathrm{poly}(|\mathbb{K}|)}$-local formal characterization (which need not satisfy the additional restrictions mentioned in Main Theorem 1 (Informal)). Furthermore, $\mathcal{F}$ is $k^{\mathrm{poly}(|\mathbb{K}|)}$-locally testable.* (See Theorem 2.11 for a formal statement.)

Significance of results: The significance of the results depend on the "novelty" of the class of properties that are linear-invariant, and have local constraints or characterizations. At first look it may appear that linear-invariance is just a rephrasing of the notion of being low-degree polynomials[1]. Indeed we even prove

[1] We remark that it is not possible to deny that every property from $\mathbb{K}^n$ to $\mathbb{F}$ is a property of "polynomials", since every function is from $\mathbb{K}^n$ to $\mathbb{F}$ is a polynomial. However this is no more interesting than saying that the function family is $|\mathbb{K}|^n$-locally testable! What we claim here, and show later in the paper, is that the class of properties showing linear-invariance is not just polynomials of a given upper bound on the degree.

that when $\mathbb{K} = \mathbb{F} = \mathbb{Z}_p$ is a prime field then the only *affine-invariant* families are polynomials of a given bound on their degree. However each restriction, $\mathbb{K} = \mathbb{F}$, $\mathbb{F} = \mathbb{Z}_p$ and the affine-invariance of $\mathcal{F}$ (as opposed to mere linear-invariance), when relaxed leads to a broader set of properties.

For instance, when $\mathbb{K} = \mathbb{F}$ and $\mathbb{F}$ is not a prime field, then the class of "linearized polynomials" lead to an interesting collection of "high-degree" polynomials that are affine-invariant, but testable with much greater locality than their degree would suggest. (Linearized polynomials over the field $\mathbb{F}$ of cardinality p^s for prime p and $s > 1$ are functions of the form $\sum_{i=0}^{s-1} c_i x^{p^i}$.) In Theorem 7.1 we give a generalization of this result to multivariate polynomials and p-degree greater than 1, giving a moderately broad class of functions that are very locally testable using Theorem 2.9.

Moving to the case where $\mathbb{K} \neq \mathbb{F}$, a priori it is not even clear that it is good to think of them as polynomials over $\mathbb{K}$ (though as noted earlier, every function from $\mathbb{K}^n$ to $\mathbb{K}$, and hence from $\mathbb{K}^n$ to $\mathbb{F}$, is a polynomial with coefficients from $\mathbb{K}$). Every non-constant function takes on a constant value $1/|\mathbb{F}|$ fraction of the times and so must be a very high degree polynomial over $\mathbb{K}$ (of degree at least $|K|/|F|$). Yet they can be locally testable with $O(1)$ locality, again suggesting that the "degree" of polynomials in the set is not a good way to measure their testability. This class of functions are interesting in that they capture the "dual-BCH" codes studied (in the context of property testing) by Kaufman and Litsyn [15]. In this paper, we give some basic structural results about such functions (see Section 6) which allows us to get some weak, but general, results about testing multivariate versions of such functions.

The strongest contrast from low-degree polynomials however comes when studying linear-invariant (as opposed to affine-invariant) families. In the previous cases, it was the structure within the field $\mathbb{K}$ that played a central role in differentiating the properties under consideration from the class of low-degree polynomials. While this distinction led to some nice examples, the "coarseness" of our general results (Informal Theorems 2 and 3 above) is weak to capture this distinction. In the case of linear-invariant families, homogenous polynomials start to play a special role and this role is quantitatively much more significant. For example consider the set of n-variate polynomials over $\mathbb{Z}_3$ supported on monomials of odd degree or monomials of degree at most 10. It can be verified that this a linear-invariant family. On the one hand this set includes polynomials of degree upto $2n - 1$, and indeed the supporting set of monomials has cardinality at least 2^n. However, it turns out that this family is testable with $O(1)$-locality independent of n (and this follows from Lemma 6.17 that is used to prove the Informal Theorem 3 above)! Indeed Lemma 6.17 gives a broad generalization of this example to a rich collection of non-low-degree polynomials that are locally testable.

We remark that linear-invariance also leads to other rich effects. As mentioned above, the class of homogenous polynomials of degree d is linear-invariant and $O(d)$-locally testable. Also if $\mathcal{F}_1$ and $\mathcal{F}_2$ are linear-invariant, then so is $\mathcal{F}_1 + \mathcal{F}_2$. It follows, again from Lemma 6.17, that if both are locally testable then so is $\mathcal{F}_1 + \mathcal{F}_2$.

In summary, we assert that the class of linear-invariant properties mapping $\mathbb{K}^n$ to $\mathbb{F}$ form a rich enhancement of the class of low-degree polynomials and our results here show how to extend some of the property testing results to the enhanced collection of properties.

Techniques: Our techniques belong into three different categories.

Unification of previous testing results by Tensor product of codes. Our testing result (Informal Theorem 1) unifies, simplifies and generalizes the proof of the robustness result from several prior works [4, 19, 2, 16, 13]. The later works in this sequence built on the proof structure developed in [4], but then needed to find new ways to address the many variants of a common technical problem that arose in all the proofs. Our insight in this work is to notice that all these problems were hovering around the concept of "tensor products" of linear spaces (or codes). By extracting this element explicitly (see proofs of Lemmas 3.1 and 3.3) we are able to find a single proof (not much more complicated than the first) that simultaneously solves all the

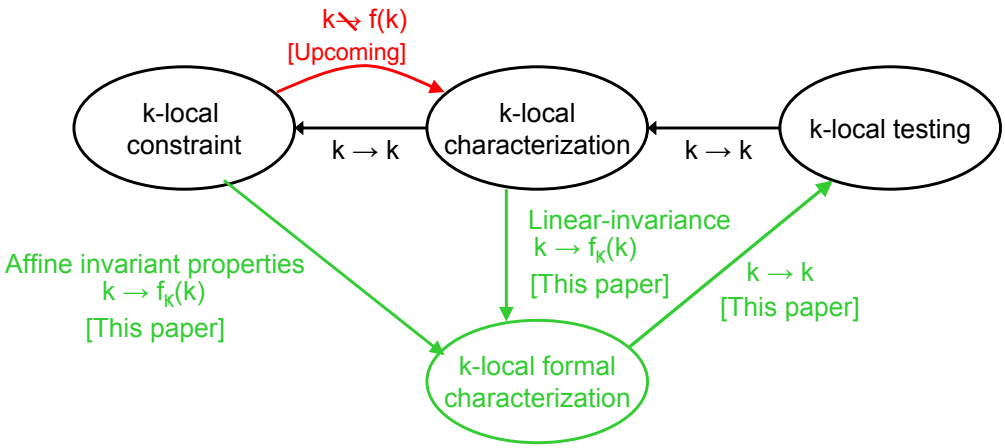

Figure 1: Informal summary of the notions and results in this paper.

problems. We remark that this proof does not specialize to any of the previous proofs, not even in the case of [4]. Previous proofs were more "efficient" in terms of the tradeoff between the rejection probability of the test and the distance from the family $\mathcal{F}$. By sacrificing this efficiency we are able to unearth some of the underlying reasons for why testing works. Given the central role of linearity and low-degree testing in complexity theory, we hope that the additional understanding will be of technical benefit in the future.

Structural theorems for linear invariant families. Our structural theorems about linear-invariant families (Informal Theorems 2 and 3) are based on a careful analysis of polynomials mapping $\mathbb{K}^n$ to $\mathbb{F}$. Recalling that every function from $\mathbb{K}^n$ to $\mathbb{F}$ can be viewed as a n-variate polynomial over $\mathbb{K}$, we ask questions of the form, what does a linear invariant family $\mathcal{F}$ containing a single function (polynomial) f look like? We present some very simple but broadly useful lemmas in this context, which we describe first for the simple case when $\mathbb{K} = \mathbb{F}$. We give a "monomial extraction lemma", Lemma 4.2, which shows that every monomial appearing in the support of f is also in $\mathcal{F}$ (where we view the monomial also as a function from $\mathbb{F}^n$ to $\mathbb{F}$). For example, any linear-invariant family containing the polynomial $x^2 + xy^2 + y^4$ also contains the function xy^2. This turns our attention to linear invariant families $\mathcal{F}$ that contain some given monomial m. We show a "monomial spreading lemma", Lemma 4.6, which describes many other monomials that should be contained in $\mathcal{F}$ as well. For example a family containing the monomial x^2y^3 over a field of characteristic greater than 5 also contains the monomials x^5 and xy^4 etc. We show a similar (more general) variant for affine-invariant families also. These lemmas, though simple, forge the path for a better understanding of linear-invariant and affine-invariant families. In particular they say that these families are completely characterized by the monomials in the families. In the case of affine-invariant families, the maximum degree of the monomials in the family forms a good, though crude, bound on the locality of the characterization/tests of the family, and this leads to the Informal Theorem 2 above.

For linear-invariant families however, the degree turns out to be the wrong measure to estimate the locality of characterizations or tests. Instead we introduce a new parameter that we call the *linear-invariance degree* of a family. For example, for the earlier-mentioned example of the family mapping $\mathbb{Z}_3^n$ to $\mathbb{Z}_3$ supported on all monomials of odd degree and on other monomials of degree upto 10, the linear-invariance degree turns out to be 10. We show that this invariance degree bounds, again crudely, the locality of the characterization/tests of any family and this leads to the Informal Theorem 3 above, in the case of $\mathbb{K} = \mathbb{F}$.

Systematic study of functions from a field $\mathbb{K}$ to a subfield $\mathbb{F}$. Finally we extend the results to the case of

function families mapping $\mathbb{K}^n$ to some subfield $\mathbb{F}$ of $\mathbb{K}$. Thus, our work provides the *first* systematic study of testability of functions from a field to its subfield. In this case we describe a basis for functions mapping $\mathbb{K}^n$ to $\mathbb{F}$, which itself seems somewhat new. This basis generalizes in a common way the well-studied "trace" and "norm" functions, both of which map $\mathbb{K}$ to $\mathbb{F}$. These functions, that we refer to as "Traces of monomials", satisfy similar properties to the monomials in the simpler case of functions from $\mathbb{F}^n$ to $\mathbb{F}$. Viewed as a polynomial over $\mathbb{K}$, if a function f has a support on a monomial m, then the trace of the monomial m is itself a function in any linear-invariant family containing f. Furthermore, the presence of one monomial implies the presence of many others in the family, leading to upper and lower bounds on the characterizations/tests of the family.

Conclusions, the Alon et al. Conjecture and Future Work: Our work attempts to highlight on the role of invariance in property testing. We remark that despite the obvious relationship of this notion to property testing, it has not been highlighted before. The only prior mentions seem to be in the works of Alon et al. [2], and in Goldreich and Sheffet [11].

Our work highlights linear-invariance as a central theme in algebraic property testing. Our results show that this notion yields a wide class of properties that have local property tests. These results are strong when the underlying field $\mathbb{K}$ is small. However when $\mathbb{K}$ is large, the characterization results (in particular, Theorem 2.10) becomes quite weak, even for affine-invariant families. In particular, in the case of the dual-BCH codes (which consider functions mapping $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$), our characterizations are completely trivial, while these codes do have very efficient tests [15]. One way to improve our results would be if Theorem 2.10 could be improved to have no dependence on t. This however is not possible, as shown in upcoming joint work with Grigorescu [12]. Specifically they exhibit a family of affine-invariant functions mapping $\mathbb{F}_{2^t}$ to $\mathbb{F}_2$ that have 8-local constraints, but no $o(t)$-local characterizations. Thus some dependence on $\mathbb{K}$ is necessary in translating constraints to characterizations.

Our work provides the first systematic study of testing functions from a field to its subfield. This setting is different than the well studied case of functions from a field to itself. This difference is best illustrated by the following example

- For affine invariant function family of the form $\mathbb{F}_2^n \to \mathbb{F}_2$ we have : a local constraint imply local characterization and local testability.
- For affine invariant function family of the form $\mathbb{F}_{2^n} \to \mathbb{F}_2$ we might have (by the work of [12]) a local constraint, but *no* local characterization! , and hence *no* local testing!

Moreover, our work suggests a method to construct new locally testable codes by picking the dual code to be a code spanned by an orbit of a short local constraint (orbit under the group of linear transformations).

In general, we feel that the class of linear-invariant functions offer a rich variety of properties, sufficiently wide to test out conjectures about the nature of testable properties. For instance, Alon et al. [2] had conjectured that linear codes of large distance, that have a small weight codeword in the dual, and have a "2-transitive invariant group" are locally testable. When applied to codes derived from affine-invariant function families, their conjecture implies that every affine-invariant family from $\mathbb{K}^n \to \mathbb{F}$ with a k-local constraint, must have an $f_{\mathbb{F}}(k)$-local test and in particular, an $f_{\mathbb{F}}(k)$-local characterization. The aforementioned result [12] refutes this conjecture of [2] by considering affine-invariant families. However, our work (Theorem 2.10) shows that a weak version of the [2] conjecture does hold, within the class of linear-invariant codes, by giving an $f_{\mathbb{K}}(k)$-local algebraic characterization and test.

This leaves the possibility that every locally characterized code with a "2-transitive invariant group" may be locally testable. Again we feel that this question can and should be examined in the context of affine-invariant families. In general, we feel that for every missing arrow, or qualitatively weak one, in Figure 1 poses an interesting open question that we hope will be investigated in future work.

This work put in focus object of the following form: $\mathbb{F}$-linear subspaces that are invariant under permutations of a group G. In this work the group G is the group of linear transformations of the domain. In a future work one may try to understand invariance under different groups in the following sense.

- Does k-local formal characterization imply local-testing also when the group of invariances is different than the group of linear transformations?
- Given a linear subspace that is invariant under permutations of a group G, when it is the case that k-local formal characterization exists (i.e. when there exists one short orbit that span the dual space)?

Organization of this paper:

In Section 2 we introduce some basic definitions needed to present our main results and we provide formal statements of our main results. Then in Section 3 we prove our main result on testing linear-invariant families. Section 7 presents an example of some families that possess very local characterizations and thus local tests. The remaining sections undertake the analysis of locality of characterizations in general linear-invariant families. Section 4 describes some basic structural properties, in particular on the role of monomials in functions mapping $\mathbb{F}^n$ to $\mathbb{F}$. Section 5 turns these results into bounds on the locality of the characterizations and tests for affine-invariant and linear-invariant families mapping $\mathbb{F}^n$ to $\mathbb{F}$. Section 6 extends the results of the previous two sections to the case of functions mapping $\mathbb{K}^n$ to $\mathbb{F}$.

2 Definitions and Statement of Results

We start with some common notation we use. We use $\mathbb{Z}$ to refer to the integers. We use $[n]$ to denote the set $\{1, \ldots, n\}$. Throughout we work with finite fields $\mathbb{F}$ of cardinality $q = p^s$ and $\mathbb{K}$ of cardinality $Q = q^t$. $\mathbb{F}^*$ and $\mathbb{K}^*$ will denote the non-zero elements of the fields. For an integer vector $\mathbf{d} = \langle d_1, \ldots, d_n \rangle$ with $0 \leq d_i < Q$ and $c \in \mathbb{K}^*$, we let $c \cdot \mathbf{x}^{\mathbf{d}}$ denote the monomial $c \cdot \prod_{i=1}^{n} x_i^{d_i}$. We use $\mathbb{K}[\mathbf{x}]$ to denote polynomials in $\mathbf{x}$ with coefficients from $\mathbb{K}$. We use $\mathcal{L}$ to denote the space of linear functions from $\mathbb{K}^n \to \mathbb{K}^n$ and $\mathcal{A}$ to denote the set of affine functions.

2.1 Robust local tests

We start with the formal definitions of constraints, characterizations and formal characterizations.

Definition 2.1 (k-local constraint/characterization) *A k-local constraint C is given by k points $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{K}^n$ and a set $S \subsetneq \mathbb{F}^k$. We say that a family $\mathcal{F}$ satisfies a* k-local constraint $C = (\mathbf{x}_1, \ldots, \mathbf{x}_k; S)$ *if $\langle f(\mathbf{x}_1), \ldots, f(\mathbf{x}_k) \rangle \in S$ for every $f \in \mathcal{F}$. We say that a family $\mathcal{F}$ has a* k-local characterization *if there exists a collection $\mathcal{C}$ of k-local constraints such that $f \in \mathcal{F}$ if and only if f satisfies all constraints $C \in \mathcal{C}$.*

When the property being tested is $\mathbb{F}$-linear, it is well-known [3] that the set S might as well be an $\mathbb{F}$-linear proper subspace of $\mathbb{F}^k$. In what follows we often use the letter V to denote such a subspace (instead of S).

We now introduce the notion of a k-local formal characterization. We start with a strong and elegant definition, though we will soon switch to a slightly weaker (but more cumbersome) definition that is easier to work with. The strong definition formalizes characterizations derived from linear, or affine, translations of a *single* k-local constraint.

Definition 2.2 (Strong Formal Characterization) *A family of functions $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$ has a* strong k-local formal characterization *it there exists a constraint $C = (\mathbf{x}_1, \ldots, \mathbf{x}_k; V \subseteq \mathbb{F}^k)$ such that $f \in \mathcal{F}$ if and only if for every linear function $L : \mathbb{K}^n \to \mathbb{K}^n$ it is the case that $\langle f(L(\mathbf{x}_1)), \ldots, f(L(\mathbf{x}_k))\rangle \in V$.*

Characterizations such as the above are common in property testing. For instance the class of linear functions from $\mathbb{Z}_p^n$ to $\mathbb{Z}_p$, for prime p and $n \geq 2$ can be described by the constraint $C = (\mathbf{a}, \mathbf{b}, \mathbf{a}+\mathbf{b}; V)$ where $\mathbf{a} = \langle 1, 0, \ldots, 0\rangle$, $\mathbf{b} = \langle 0, 1, 0, \ldots, 0\rangle$, and $V = \{\langle \alpha, \beta, \alpha+\beta\rangle | \alpha, \beta \in \mathbb{Z}_p\}$. Similarly, the class of degree d polynomials mapping $\mathbb{Z}_p^n$ to $\mathbb{Z}_p$, for $d \leq p$ and $n \geq 2$ can be described by the constraint $C = (\mathbf{a}, \mathbf{a}+\mathbf{b}, \mathbf{a}+2\mathbf{b}, \ldots, \mathbf{a}+(d+1)\mathbf{b}; V_d)$ where $V_d = \{\langle \alpha_0, \ldots, \alpha_{d+1}\rangle \in \mathbb{F}^{d+2} | \sum_{i=0}^{d+1} (-1)^i \binom{d+1}{i} \alpha_i = 0\}$. More complex expressions can be found for functions mapping polynomials over any (esp. a non-prime) field to itself. However all these definitions do restrict n to be at least 2, which is somewhat artificial. Also for technical reasons we will use a "dual" (and weaker) notion of a "formal" constraint.

In the above version, a formal characterization may be viewed as being given by a collection of constraints: one for every linear map from $\mathbb{K}^n$ to $\mathbb{K}^n$. In the "dual" version below, we will consider a collection of constraints which are parametrized by a constant number of variables taking values in $\mathbb{K}^n$. The "variables" of a constraint, i.e., locations examined by the constraint, are linear functions of the parameters. As usual the constraint requires that the vector of function values at the specified locations come from the set S.

Definition 2.3 ((Weak) k-local formal characterization) *A family $\mathcal{F}$ has a (weak) k-local formal characterization if there exists an integer m; k linear functions $\ell_1, \ldots, \ell_k : (\mathbb{K})^m \to \mathbb{K}$; and a linear subspace $V \subset \mathbb{F}^k$ such that $f \in \mathcal{F}$ if and only if for every $y_1, \ldots, y_m \in \mathbb{K}^n$, we have $\langle f(x_1), \ldots, f(x_k)\rangle \in V$, where $x_i = \ell_i(y_1, \ldots, y_m)$. (Here we interpret the linear function ℓ_i as a map from $(\mathbb{K}^n)^m \to \mathbb{K}^n$ in the natural way.)*

The following proposition establishes a fairly close connection between strong and weak formal characterizations.

Proposition 2.4 *A family $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$ has a weak k-local formal characterization if is it has a strong k-local formal characterization. If $n \geq k$ then the converse also holds.*

Proof: Let $C = (\mathbf{x}_1, \ldots, \mathbf{x}_k; V)$ give a strong formal characterization of $\mathcal{F}$. Renumber $\mathbf{x}_1, \ldots, \mathbf{x}_k$ so that the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are linearly independent and $\mathbf{x}_j = \sum_{i=1}^m \lambda_{ij}\mathbf{x}_i$ for $j \in \{m+1, \ldots, k\}$. Now let $\ell_1, \ldots, \ell_k : \mathbb{K}^m \to \mathbb{K}$ be defined as $\ell_j(z_1, \ldots, z_m) = z_j$ if $j \leq m$ and $\ell_j(z_1, \ldots, z_m) = \sum_{i=1}^m \lambda_{ij} z_i$ for $j \in \{m+1, \ldots, k\}$. Then it can be easily seen that $\ell_1, \ldots, \ell_k$ and V give a weak formal characterization of $\mathcal{F}$.

In the other direction, suppose $\ell_1, \ldots, \ell_k : \mathbb{K}^m \to \mathbb{K}$ and V give a weak formal characterization of $\mathcal{F}$. Let $\alpha_1, \ldots, \alpha_m \in \mathbb{K}^n$ be linearly independent vectors in $\mathbb{K}^n$. (Note such a collection exist since $m \leq k \leq n$.) Let $\mathbf{x}_1, \ldots, \mathbf{x}_k$ be given by $\mathbf{x}_j = \ell_j(\alpha_1, \ldots, \alpha_m)$. Then it can be verified that the constraint $(\mathbf{x}_1, \ldots, \mathbf{x}_k; V)$ gives a strong formal characterization of $\mathcal{F}$. ∎

Henceforth whenever we refer to formal characterizations, we mean weak ones. The formal version of the Informal Theorems 1, 2, and 3 rely on some restricted classes of formal characterizations that we specify below.

Definition 2.5 (2-ary Independent and Affine Formal Characterizations) *A k-local formal characterization $(\ell_1, \ldots, \ell_k; V)$ is* 2-ary independent *if ℓ_1 and ℓ_j are linearly independent for every $j \in \{2, \ldots, k\}$. If all the ℓ_i's are of the form $y_1 + \tilde{\ell}_i(y_2, \ldots, y_m)$, where $\tilde{\ell}_i$'s are non-zero, then we say that the characterization is an* affine *characterization. (Note that every affine characterization is also 2-ary independent.)*

In the propositions below, we mention some general results on the existence of formal local characterizations. The first gives a general transformation, which may be quite weak for large $\mathbb{K}$, but is quite useful for small $\mathbb{K}$. The second summarizes known (quite strong) characterizations in our terms. Both proofs are omitted.

Proposition 2.6 *For every $\mathbb{K}$ there exists a function $g = g_{\mathbb{K}} : \mathbb{Z} \to \mathbb{Z}$ such that if $\mathcal{F}$ has a k-local characterization, then it has a $g(k)$-local formal characterization.*

Proposition 2.7 (Follows from [7, 16]) *The set $\mathcal{F}_{n,d,\mathbb{F}}$ of n-variate polynomials of degree at most d over $\mathbb{F}$ (so here $\mathbb{K} = \mathbb{F}$) of cardinality $q = p^s$, have a $d + 2$-local formal characterization, if $d \leq q - q/p$, and a $q^{\lceil d/(q(1-1/p)) \rceil}$-local formal characterzation if $d \geq d(1 - 1/p)$. In both cases, the formal characterizations are affine.*

A much wider class of properties (other than just the class of low-degree polynomials) have local characterizations. We discuss this in detail shortly, but first we describe a natural test for properties with local formal characterizations.

Definition 2.8 (Linear-invariant test) *For family $\mathcal{F}$ that has a formal local characterization given by $(\ell_1, \ldots, \ell_k; V)$, the linear-invariant test is defined to be: "Pick $x_1, \ldots, x_m \in \mathbb{K}^n$ at random and accept if and only if $\langle f(y_1), \ldots, f(y_k) \rangle \in V$, where $y_i = \ell_i(x_1, \ldots, x_m)$."*

We can now state our main theorem, which formalizes the Informal Theorem 1 of Section 1, for testing linear-invariant families with local formal characterization.

Theorem 2.9 *If $\mathcal{F}$ is a (linear invariant) family of functions mapping $\mathbb{K}^n$ to $\mathbb{F}$, with a 2-ary independent k-local formal characterization, then it is k-locally testable. Specifically, the linear-invariant test accepts all members of $\mathcal{F}$, while a function f that is δ-far from $\mathcal{F}$ is rejected with probability* $\min\left\{\frac{\delta}{2}, \frac{1}{(2k+1)(k-1)}\right\}$.

We prove this Theorem in Section 3. In particular, note that in all cases the rejection probability is independent of n and $\mathbb{K}$. So if $k = O(1)$, then the rejection probability is $\Omega(\delta)$.

For well-known linear-invariant families such as linear functions [4], and Reed-Muller codes [19, 2, 16, 13], the theorem above produces local tests with the same locality as in the previous works, though the rejection probability may be slightly smaller in our case. The rest of this section describes property tests that we can derive that are not already captured by previous results.

To do so we study invariance properties of functions mapping $\mathbb{K}^n$ to $\mathbb{F}$. All functions from $\mathbb{K}^n$ to $\mathbb{F}$ are polynomials. So the principal questions we study here are: "Which subsets of polynomials are linear (or affine) invariant?" and "Which of these families have k-local formal characterizations?"

We differentiate our results into two categories: those for affine-invariant families and those for linear-invariant families. In both cases, as argued earlier there is a rich variety of function families that are not "merely" low-degree polynomials. However in the case of affine-invariant families, the maximum degree

of functions in the family does give a crude bound on the locality of characterizations and tests for the family. On the one hand families that contain even a single high-degree function cannot satisfy any local constraint; and on the other hand families with only low-degree functions have local formal characterizations (see Lemmas 6.9 and 6.14). For affine-invariant families, the characterizations can be converted to affine-invariant, and hence 2-ary independent ones, one can now apply Theorem 2.9 to get a testing result as well. This leads us to the following theorem, which formalizes Informal Theorem 2.

Theorem 2.10 . *For fields $\mathbb{F} \subseteq \mathbb{K}$ with $|\mathbb{F}| = q$ and $|\mathbb{K}| = Q$, let $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$ be an affine-invariant family with a k-local constraint. Then $\mathcal{F}$ has a $k' = (Q^2k)^{Q^2}$-local formal affine characterization. Furthermore $\mathcal{F}$ is k'-locally testable where the test accepts members of $\mathcal{F}$ with probability* 1 *and rejects functions that are δ-far with probability* $\min\left\{\frac{\delta}{2}, \frac{1}{(2k'+1)(k'+1)}\right\}$.

Theorem 2.10 is proved in Section 6, though the simpler case where $\mathbb{K} = \mathbb{F}$ is proved in Section 5.

The gap between the upper and lower bounds is the above theorem is quite weak. Partly this is because the degree of the polynomial in a family is only a weak estimator of the locality of characterizations. In Section 7 we give an example of a family mapping $\mathbb{F}^n$ to $\mathbb{F}$ where the degree is larger than the locality of the characterization by a factor of about q/p. This example is interesting in its own right in that it shows some of the ways in which affine-invariant families differ from families of low-degree polynomials.

In the case of linear-invariant families, the degree is no longer even a crude estimator of the locality of characterizations. In Section 5 we introduce the notion of the linear-invariance degree of a family and use this parameter in Sections 5 and 6 to derive upper bounds on the locality of formal characterizations, while also deriving lower bounds on the locality of (any) characterization (see Lemmas 6.8 and 6.16). These characterizations, unfortunately, are not 2-ary independent. However we manage to reduce the testing of linear-invariant families to some related families that do have 2-ary independent characterizations. This allows us to use Theorem 2.9, in a slightly more involved way, to get local tests for linear-invariant families as well. The following theorem, which formalizes Informal Theorem 3, summarizes this investigation.

Theorem 2.11 . *For fields $\mathbb{F} \subseteq \mathbb{K}$ with $|\mathbb{F}| = q$ and $|\mathbb{K}| = Q$, let $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$ be an linear-invariant family with a k-local characterization. Then $\mathcal{F}$ has a $k' = (Q^2k)^{Q^2}$-local formal characterization. Furthermore $\mathcal{F}$ is k_0-locally testable, for $k_0 = 2Qk'$ where the test accepts members of $\mathcal{F}$ with probability* 1 *and rejects functions that are δ-far with probability* $\min\left\{\frac{\delta}{2}, \frac{Q^2}{(2k_0+Q)(k_0+Q)}\right\}$.

Again, Theorem 2.11 is proved in Section 6, though the simpler case where $\mathbb{K} = \mathbb{F}$ is proved in Section 5.

Part I

3 Local Testing from Local Formal Characterizations

In this section we wish to prove Theorem 2.9 which asserts that a linear-invariant family $\mathcal{F}$ with a 2-ary independent k-local formal characterization is k-locally testable, by the linear-invariant test for $\mathcal{F}$. We restate the theorem below.

Theorem 2.9 (restated) *If $\mathcal{F}$ is a (linear invariant) family of functions mapping $\mathbb{K}^n$ to $\mathbb{F}$, with a 2-ary independent k-local formal characterization, then it is k-locally testable. Specifically, the linear-invariant test accepts all members of $\mathcal{F}$, while a function f that is δ-far from $\mathcal{F}$ is rejected with probability* $\min\left\{\frac{\delta}{2}, \frac{1}{(2k+1)(k-1)}\right\}$.

In particular, the theorem implies that every affine-invariant family $\mathcal{F}$ with a k-local formal characterization is testable.

Recall the linear-invariant test picks $x_1, \ldots, x_m \in \mathbb{K}^n$ at random and accepts if and only if $\langle f(y_1), \ldots, f(y_k)\rangle \in V$, where $y_i = \ell_i(x_1, \ldots, x_m)$ for $i \in [k]$.

Let $\epsilon(f)$ denote the probability that the linear-invariant test rejects a function f. It is clear that if $f \in \mathcal{F}$ then $\epsilon(f) = 0$. So to prove Theorem 2.9 for the case of 2-ary independent formal characterizations, it suffices to show that if $\epsilon(f) < \frac{1}{(2k+1)(k-1)}$, then $\delta(f, \mathcal{F}) \leq 2\epsilon(f)$.

We start by making some notational simplifications. For $i \in [k]$ and $j \in [m]$, let $c_{ij} \in \mathbb{K}$ be such that $\ell_i(x_1, \ldots, x_m) = \sum_{j=1}^{m} c_{ij} x_j$. Without loss of generality, we assume that the first m linear functions simply project on to the first m coordinates; i.e., $\ell_i(x_1, \ldots, x_m) = x_i$ for $i \in [m]$. (This can be achieved by a linear transformation of the variable $x_1, \ldots, x_m$ and by permuting the ℓ_i's.) Furthermore, we assume the remaining coordinates are linearly independent of x_1 and so for every $i \neq 1$, the vector $\langle c_{i2}, \ldots, c_{im}\rangle \neq 0$.

Fix a function f with $\epsilon(f) < 1/((2k+1)(k-1))$. As in [4], we now describe a function $g : \mathbb{K}^n \to \mathbb{F}$ that is close to f, that will turn out to be a member of $\mathcal{F}$. For any choice of values $\alpha_2, \ldots, \alpha_k \in \mathbb{F}$ notice that there is at most one $\alpha \in \mathbb{F}$ such that $\langle \alpha, \alpha_2, \ldots, \alpha_k\rangle \in V$. Define $\textsc{Decode}(\alpha_2, \ldots, \alpha_k)$ to be this α if it exists (and a special symbol $\perp$ denoting error otherwise). For $x \in \mathbb{K}^n$ and let $\mathrm{Sc}^f(x; x_2, \ldots, x_m) = \textsc{Decode}(f(y_2), \ldots, f(y_k))$ where $y_i = \ell_i(x, x_2, \ldots, x_m)$. Note that $\epsilon(f)$ equals the probability that $f(x) \neq \mathrm{Sc}^f(x; x_2, \ldots, x_m)$, when $x, x_2, \ldots, x_m$ are chosen uniformly and independently from $\mathbb{K}^n$. In particular $f(x) = \mathrm{Sc}^f(x; x_2, \ldots, x_m)$ for every $x, x_2, \ldots, x_m$ if and only if $f \in \mathcal{F}$.

Finally, we are ready to define the function g, which we claim to be the function close to f that is in $\mathcal{F}$. For $x \in \mathbb{K}^n$, let $g(x) = \mathrm{plurality}_{\alpha \in (\mathbb{K}^n)^{m-1}} \mathrm{Sc}^f(x, \alpha)$.

We now follow the same sequence of steps as in [4]. It is straightforward to show that f is close to g and we do so in Lemma 3.2. But before we do so, we move to the crucial step, which is to prove that the plurality above is really an overwhelming majority for every x. We show this first in Lemma 3.1. Finally, a proof similar to that of Lemma 3.1 shows that g must be a member of $\mathcal{F}$ and we do so in Lemma 3.3. Theorem 2.9 follows easily from these lemmas.

Lemma 3.1 *For every $x \in \mathbb{K}^n$, $\Pr_{\mathbf{y},\mathbf{z}}[(\mathrm{Sc}^f(x, \mathbf{y}) \neq \mathrm{Sc}^f(x, \mathbf{z}))] \leq 2(k-1)\epsilon(f)$. Hence, for every $x \in \mathbb{K}^n$, $\Pr_{\mathbf{y}}[g(x) \neq \mathrm{Sc}^f(x, \mathbf{y})] \leq 2(k-1)\epsilon(f)$.*

Proof: Let $\epsilon = \epsilon(f)$. We build two $k \times k$ matrices M, N with $M_{ij} \in \mathbb{K}^n$ and $N_{ij} \in \mathbb{F}$ and use properties of these matrices to prove the lemma.

For $i, j \in [m]$ pick $\gamma_{ij} \in \mathbb{K}^n$ as follows. Let $\gamma_{11} = x$, $\gamma_{1j} = y_j$, $\gamma_{i1} = z_i$, and γ_{ij} be chosen independently and uniformly at random from $\mathbb{K}^n$ otherwise. (Note every γ_{ij} except γ_{11} is thus drawn uniformly at random from $\mathbb{K}^n$.) Now for $i \in [k]$ and $j \in [m]$, let $M_{ij} = \ell_i(\gamma_{1j}, \ldots, \gamma_{mj})$. (In particular, we have $M_{ij} = \gamma_{ij}$ for $i, j \in [m]$.) Finally for $i \in [k]$ and $j \in [k]$, let $M_{ij} = \ell_j(M_{i1}, \ldots, M_{im})$. The second matrix N_{ij} is defined to be $f(M_{ij})$ except when $i = j = 1$, in which case we define $N_{11} = \mathrm{Sc}^f(x, \mathbf{y})$.

Below we show that all the rows of N are codewords of V (with high probability), and that all the columns except possibly the first are also codewords of V. This allows us to conclude that the first column is also a codeword of V and this in turn yields the lemma.

We start by examining the properties of M and N. We claim that every row and every column of M corresponds to the queries of a potential test by our tester. We start with the rows. Fix $i \in [k]$ and note that the entries of the ith row correspond to queries of the test with randomness $M_{i1}, \ldots, M_{im}$ (corresponding to queries of the test "Does $f(M_{i1}) = \mathrm{Sc}^f(M_{i1}; M_{i2}, \ldots, M_{im})$?"). Notice further that for $i \neq 1$ the values $M_{i1}, \ldots, M_{im}$ are drawn uniformly and independently at random from $\mathbb{K}^n$ (independent of x). To see this, suppose $c_{ij} \neq 0$ for some $j \in \{2, \ldots, m\}$. Then note that there is a one to one correspondence between $\langle \gamma_{j1}, \ldots, \gamma_{jm} \rangle$ and $\langle M_{i1}, \ldots, M_{im} \rangle$ for any fixed choice of $\{\gamma_{ik}\}_{i \neq j,k}$. Thus choosing $\langle \gamma_{j1}, \ldots, \gamma_{jm} \rangle$ uniformly at random makes $\langle M_{i1}, \ldots, M_{im} \rangle$ uniform over $(\mathbb{K}^n)^m$ independent of $\gamma_{11} = x$. We conclude that the probability that $f(M_{i1}) \neq \mathrm{Sc}^f(M_{i1}; M_{i2}, \ldots, M_{im})$ is at most ϵ. In other words, the probability that the ith row of N is *not* a codeword of V is at most ϵ for $i \neq 1$.

Next we move to the columns of M and N. Note that the construction of M was asymmetric in that every row was defined to form a "query" pattern of our test. However, we note that the same matrix could have been defined by constructing the first m rows first, and then defining each column to be a "query pattern" of the test. To see this recall that $\ell_i(x_1, \ldots, x_m) = \sum_{j=1}^m c_{ij} x_j$. Thus we have

$$\begin{aligned}
M_{ij} &= \ell_j(M_{i1}, \ldots, M_{im}) \\
&= \sum_{j'=1}^{m} c_{jj'} M_{ij'} \\
&= \sum_{j'=1}^{m} c_{jj'} \sum_{i'=1}^{m} c_{ii'} M_{i'j'} \\
&= \sum_{i'=1}^{m} c_{ii'} \sum_{j'=1}^{m} c_{jj'} M_{i'j'} \\
&= \sum_{i'=1}^{m} c_{ii'} M_{i'j} \\
&= \ell_i(M_{1j}, \ldots, M_{mj}).
\end{aligned}$$

By a similar argument to the previous paragraph we now have that the probability that the jth column of N is not a codeword is at most ϵ for $j \neq 1$.

Thus, by the union bound, we have that with probability at most $2(k-1)\epsilon$ there exists a row (other than the first) or a column (other than the first) such that N restricted to the row or the column is not a codeword of V. We now use this to show that the first row of N and the first column of N are also codewords of V. Here we use the properties of tensor products of codes. Recall that the tensor product of V with itself, denoted $V \otimes V$ is the code consisting of all $k \times k$ matrices over $\mathbb{F}$ all of whose rows are codewords of V and all of whose columns are codewords of V. It is well known that if V has distance d then its tensor product with

itself has the following "erasure-correcting" property: Given the projection of any matrix $B = A|_{S\times T}$ to a subset S of the rows and a subset T of the columns with $|S|, |T| \geq k - d + 1$, B can be extended to a (unique) codeword A of $V \otimes V$ if and only if for every row $s \in S$, the sth row of B is consistent with (the projection to T of) some codeword of V, and for every column $t \in T$, the tth column of B is consistent with (the projection to S of) some codeword of V.

In our case, the code V has distance at least 2 and we know the projection of N onto all columns except the first and all rows except the first are consistent with V. Thus the extension to N to a codeword of $V \otimes V$ is unique and this is the unique value which satisfies $N_{11} = \text{DECODE}(N_{12}, \ldots, N_{1k}) = \text{DECODE}(N_{21}, \ldots, N_{k1})$. We conclude that with probability at least $1-2(k-1)\epsilon$, we have $\Pr_{\mathbf{y},\mathbf{z}}[\text{Sc}^f(x, \mathbf{y}) \neq \text{Sc}^f(x, \mathbf{z})] \leq 2(k-1)\epsilon(f)$.

The consequence to g follows from the fact when drawing samples from a distribution, the probability of a collision is no more than the probability of the most likely element. ∎

We now revert to the task of proving that f is close to g and that g is a member of the family $\mathcal{F}$. We start with the former task which we show in exactly the same way as in [4, 19].

Lemma 3.2 $\delta(f, g) \leq 2\epsilon(f)$.

Proof: Let $B = \{x \in \mathbb{K}^n | \Pr_\alpha[f(x) \neq \text{Sc}^f(x, \alpha)] \geq \frac{1}{2}\}$. Notice that $\epsilon(f) \geq \frac{1}{2}\Pr_x[x \in B]$. On the other hand, if $x \notin B$, then $f(x) = \text{plurality}_\alpha[\text{Sc}^f(x, \alpha)]$. The lemma follows. ∎

Next we show that the proof technique of Lemma 3.1 can be adapted to prove also that $g \in \mathcal{F}$. This modification is similar to those in the early papers [4, 19].

Lemma 3.3 *Let f be a function with $\epsilon(f) < \frac{1}{(2k+1)(k-1)}$ and let g be its self-corrected version. Then $g \in \mathcal{F}$.*

Proof: It suffices to show that for every $x_1, \ldots, x_m \in \mathbb{K}^n$ the vector $\langle g(y_1), \ldots, g(y_k)\rangle \in V$, where $y_i = \ell_i(x_1, \ldots, x_m)$. Fix such a sequence $x_1, \ldots, x_m \in \mathbb{K}^n$ and let $y_i = \ell_i(x_1, \ldots, x_m)$. for $i \in [k]$. As in the proof of Lemma 3.1, we will construct a matrix $M \in (\mathbb{K}^n)^{k\times k}$ whose first row will be $y_1, \ldots, y_k$. We will then define a related matrix N and show that all rows of N, except possibly the first, and all columns are codewords of V. We will then conclude that its first row must be a codeword of V and this will imply the lemma.

For $i, j \in [m]$, pick γ_{ij} as follows. $\gamma_{1j} = x_j$ and γ_{ij} is drawn uniformly and independently from $\mathbb{K}^n$ for all other i, j pairs. For $i' \in [k]$ and $j \in [m]$, define $M_{i'j} = \ell_{i'}(\gamma_{1j}, \ldots, \gamma_{mj})$. Finally, for $i', j' \in [k]$, define $M_{i'j'} = \ell_{j'}(M_{i'1}, \ldots, M_{i'm})$. Now let $N_{ij} = g(M_{ij})$ if $i = 1$ and $f(M_{ij})$ otherwise.

As in the proof of Lemma 3.1 we have that all the rows of M except the first represent the queries of a random test, and in particular the queried points are independent of $y_1, \ldots, y_k$. Thus we have that the probability that the i'th row of N is not a codeword of V is at most ϵ, for $i' \neq 1$.

Next we turn to the columns of N. Note that once again we have $M_{ij} = \ell_i(M_{1j}, \ldots, M_{mj})$. Now for every j, the jth column of M represents the queries of a random test through y_j. Thus we have that the probability that the jth column of N is not a codeword of V is given by the probability of the event $g(y_j) \neq \text{Sc}^f(y_j; M_{2j}, \ldots, M_{mj})$ and by Lemma 3.1 the probability of this event is at most $2(k-1)\epsilon$.

Taking the union of all the "bad events" and deducting them, we have that with probability at least $1-(2k+1)(k-1)\epsilon$ we have that all the rows of N except the first, and all the columns of N are codewords of V. We conclude (as in the proof of Lemma 3.1) that the first row of N, i.e., the vector $\langle g(y_1), \ldots, g(y_k)\rangle$ is a

codeword of V. Since $1-(2k+1)(k-1)\epsilon(f) > 0$, we have with positive probability $\langle g(y_1), \ldots, g(y_k)\rangle \in V$. But $y_1, \ldots, y_k$ were chosen deterministically and so the probability of this event is either zero or one, yielding that this event must happen with probability one. ∎

Finally, we can prove our main testing theorem, namely that locally (formally) characterized function families are locally testable.

Proof of Theorem 2.9: From Lemma 3.2, we have $\delta(f,g) \leq 2\epsilon(f)$. and by Lemma 3.3, we have $g \in \mathcal{F}$ and so $\delta(f) \leq 2\epsilon(f)$. ∎

Part II

4 Structure of Affine-/Linear-Invariant Families

In this section we aim to study structural properties of linear-invariant and affine-invariant families of functions mapping $\mathbb{F}^n$ to $\mathbb{F}$. (Later we extend the study to functions from $\mathbb{K}^n$ to $\mathbb{F}$ in Section 6.) We start by proving two basic lemmas that describe some of the members of any linear-invariant (or affine-invariant) family $\mathcal{F}$ containing a given function f. The first of these shows that in a linear-invariant family $\mathcal{F}$ containing f, every monomial in the support of f, when viewed as a function, is also in the family $\mathcal{F}$. The second lemma illustrates that given a single monomial in some linear-invariant family $\mathcal{F}$, one can deduce the presence of many other monomials in $\mathcal{F}$. In fact, over a prime field we show that *all* monomials with degree less than or equal to that of the starting monomial are in the family, if the family is affine-invariant. Together, these lemmas lead to a good understanding of the behavior of linear-invariant families and help the study of (formal) local characterizations in Section 5.

Before launching into the section we first introduce some notation and definitions that apply generally to functions mapping $\mathbb{K}^n \to \mathbb{F}$.

We use $\{\mathbb{K}^n \to \mathbb{F}\}$ to denote the set of all functions mapping $\mathbb{K}^n$ to $\mathbb{F}$.

Definition 4.1 *For a set of functions $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$, $\text{SPAN}_{\mathbb{F}}(\mathcal{F}) = \{\sum_{i=1}^{\ell} \alpha_i \cdot f_i | \ell \in \mathbb{Z}^+, \alpha_i \in \mathbb{F}, f_i \in \mathcal{F}\}$ denotes the linear span (over $\mathbb{F}$) of $\mathcal{F}$. For a family of functions $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$ we let the linear span of $\mathcal{F}$, denoted $\text{L-SPAN}_{\mathbb{F}}(\mathcal{F})$, be the smallest linear-invariant family of functions containing $\mathcal{F}$. Finally, the affine span of $\mathcal{F}$, denoted $\text{A-SPAN}_{\mathbb{F}}(\mathcal{F})$ is the smallest affine-invariant family containing $\mathcal{F}$.*

When the range $\mathbb{F}$ is clear from the context we suppress the subscript and refer to $\text{SPAN}_{\mathbb{F}}(\mathcal{F})$ as simply $\text{SPAN}(\mathcal{F})$. Note that $\text{L-SPAN}(\mathcal{F})$ can be written as $\text{SPAN}(\{f(L(\mathbf{x}))|f \in \mathcal{F} \text{ and } L : \mathbb{K}^n \to \mathbb{K}^n \text{ is a linear function}\})$. Similarly, $\text{A-SPAN}(\mathcal{F})$ can be written as $\text{SPAN}(\{f(A(\mathbf{x}))|f \in \mathcal{F} \text{ and } A : \mathbb{K}^n \to \mathbb{K}^n \text{ is an affine function}\})$.

We will be switching back and forth between functions and polynomials. Specifically, given an n-dimensional vector space $\mathbb{K}^n$, we will associate n variables $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ with the space. Given a function $f : \mathbb{K}^n \to \mathbb{F}$ we will often use it exchangeably to represent the unique polynomial in $p_f \in \mathbb{K}[\mathbf{x}]$ (with coefficients in $\mathbb{K}$) whose degree in each variable is at most $|\mathbb{K}| - 1$, and which evaluates to the function f on every point in $\mathbb{K}^n$. In particular, below we will be thinking of monomials in $\mathbb{F}[\mathbf{x}]$ as functions from $\mathbb{F}^n \to \mathbb{F}$.

4.1 Extracting Monomials in Linear-Invariant Families

For a polynomial $f = \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{x}^{\mathbf{d}}$, we refer to the support of f to be the set of monomials $c_{\mathbf{d}} \mathbf{x}^{\mathbf{d}}$ with $c_d \neq 0$. For a monomial $m = \mathbf{x}^{\mathbf{d}}$, we denote the degree of the monomial by $deg(m) = \sum_{i=1}^{n} d_i$. Our first lemma asserts that in a linear-invariant family mapping $\mathbb{F}^n$ to $\mathbb{F}$, every monomial in the support of a function in the family also belongs to the family.

Lemma 4.2 *[Monomial extraction lemma] For every function $f : \mathbb{F}^n \to \mathbb{F}$, every monomial in the support of f is contained in $\text{L-SPAN}(f)$.*

Proof: We prove the lemma by proving the following claim about univariate polynomials, and then using induction on the number of variables.

Claim 4.3 *Let $f(\mathbf{x}) = \sum_{k=0}^{q-1} f_i \cdot x_n^i$ where $f_i \in \mathbb{F}[x_1, \ldots, x_{n-1}]$ for $i \in \{0, \ldots, q-1\}$ and $q = |\mathbb{F}|$. Then for every such $k \in \{0, \ldots, q-1\}$, $f_k \cdot x_n^k \in$ L-SPAN(f).*

Proof: Note that for $k = 0$, $f_0 = f(x_1, \ldots, x_{n-1}, 0)$ which is obviously in L-SPAN(f). So fix $k \in \{1, \ldots, q-1\}$. Now let $\tilde{f} = f - f_0 = \sum_{i=1}^{q-1} f_i \cdot x_n^i$. Note that since $\tilde{f} \in$ L-SPAN(f), it suffices to show that $f_k \cdot x_n^k \in$ L-SPAN$(\tilde{f})$. Fix a primitive element $\alpha \in \mathbb{F} - \{0\}$. We claim that $f_k \cdot x_n^k = -\sum_{j=1}^{q-1} \alpha^{-kj} \tilde{f}(x_1, \ldots, x_{n-1}, \alpha^j x_n)$, which immediately implies $f_k \cdot x_n^k \in$ L-SPAN$(\tilde{f})$. To verify the claim, we work on the RHS:

$$\begin{aligned} -\sum_{j=1}^{q-1} \alpha^{-kj} \tilde{f}(x_1, \ldots, x_{n-1}, \alpha^j x_n) &= -\sum_{j=1}^{q-1} \alpha^{-kj} \sum_{i=1}^{q-1} f_i \cdot (\alpha^j x_n)^i \\ &= -\sum_{i=1}^{q-1} f_i \cdot x_n^i \cdot \left(\sum_{j=1}^{q-1} \alpha^{(i-k)j} \right) \end{aligned}$$

But the inside term is of the form $\sum_{j=1}^{q-1} \beta^j$ which is zero for every $\beta \in \mathbb{F} - \{1\}$, and equals -1 when $\beta = 1$. We conclude that the above expression simplifies to $f_k \cdot x_n^k$. This concludes the proof of Claim 4.3. ∎

We now conclude the proof of the lemma with a simple inductive argument. Let $f = \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{x}^{\mathbf{d}}$. Fix a vector $\mathbf{e}$ such that $c_{\mathbf{e}} \neq 0$. We will show that $c_{\mathbf{e}} \mathbf{x}^{\mathbf{e}}$ is in L-SPAN(f). To do so let

$$h_i = \sum_{d_1, \ldots, d_i} c_{d_1, \ldots, d_i, e_{i+1}, \ldots, e_n} \prod_{j=1}^{i} x_j^{d_j} \prod_{j=i+1}^{n} x_j^{e_j}.$$

Note that $h_n = f$ and $h_0 = c_{\mathbf{e}} \mathbf{x}^{\mathbf{e}}$ is the monomial of interest to us. From Claim 4.3 we get that for every i, h_i is in L-SPAN(h_{i+1}) and thus in L-SPAN(f). For $i = 0$, this yields the lemma. ∎

4.2 The spread of monomials in linear-/affine-invariant families

The main lemma is a general lemma that asserts that the presence of a single monomial in a family implies the presence of other monomials, with "smaller" degrees in a somewhat technical sense. We follow the lemma up with a corollary that describes some of the ways in which the lemma will be used later. Before presenting the lemma we present a simple useful proposition.

Proposition 4.4 *Let $\mathbf{x}$ and $\mathbf{z}$ be disjoint sets of variables. If a monomial $m = \mathbf{x}^{\mathbf{d}}$ has the monomial $m' = \mathbf{x}^{\mathbf{e}} \cdot \mathbf{z}^{\mathbf{e}'}$ it its linear (affine) span, then the monomial $m \cdot \mathbf{z}^{\mathbf{f}}$ has the monomial $m' \cdot \mathbf{z}^{\mathbf{e}'+\mathbf{f}}$ it its linear (resp. affine) span.*

Proof: We prove the proposition for the case of affine spans. The linear case is similar.

Let n denote the dimension of $\mathbf{x}$ and n' denote the dimension of $\mathbf{z}$. By the fact that m' is in the affine span of m we get that $m' = \mathbf{x}^{\mathbf{e}} \cdot \mathbf{z}^{\mathbf{e}'} = \sum_{i=1}^{\ell} c_i (A_i \mathbf{x} + \mathbf{b}_i)^{\mathbf{d}}$ for some finite sequence $\{(c_i, A_i, \mathbf{b}_i)\}_{i=1}^{\ell}$ with $c_i \in \mathbb{F}$, $A_i \in \mathbb{F}^{(n+n') \times n}$ and $\mathbf{b}_i \in \mathbb{F}^{n+n'}$. For every $i \in [\ell]$, let $A_i' \in \mathbb{F}^{(n+n') \times (n+n')}$ be given by

$$A_i' = \left(A_i \,\middle|\, \begin{matrix} 0 \\ I_{n'} \end{matrix} \right),$$

and let $\mathbf{b}'_i = \mathbf{b}_i$. Note that

$$\sum_{i=1}^{\ell} c_i A'_i(\mathbf{x}, \mathbf{z} + \mathbf{b}'_i)^{\langle \mathbf{d},\mathbf{f} \rangle} = \sum_{i=1}^{\ell} c_i (A_i\mathbf{x} + b_i)^{\mathbf{d}} \cdot \mathbf{z}^{\mathbf{f}} = \mathbf{x}^{\mathbf{e}} \cdot \mathbf{z}^{\mathbf{e}'} \cdot \mathbf{z}^{\mathbf{f}}.$$

Thus we have that the monomial $\mathbf{x}^{\mathbf{d}} \cdot \mathbf{z}^{\mathbf{f}}$ has the monomial $\mathbf{x}^{\mathbf{e}} \cdot \mathbf{z}^{\mathbf{e}'+\mathbf{f}}$ it its affine span. ∎

Proposition 4.5 *Let $m, m' \in \mathbb{F}[\mathbf{x}] \subseteq \mathbb{F}[\mathbf{x}, y]$ be such that $m' \in$ A-SPAN(m). Then $y^{\deg(m)-\deg(m')} \cdot m' \in$ L-SPAN(m).*

Proof: Let $d = \deg(m)$ and $d' = \deg(m')$. Let $m' = \sum_{i=1}^{\ell} c_i m(A_i\mathbf{x} + b_i)$. Then $\sum_{i=1}^{\ell} c_i m(A_i\mathbf{x} + b_i y)$ is a homogenous polynomial $f(\mathbf{x}, y)$ of degree d. Furthermore $f(\mathbf{x}, 1) = m'$. If follows that $f(\mathbf{x}, y) = y^{d-d'} \cdot m'$. To see this, let $f(\mathbf{x}, y) = \sum_{i=0}^{d} f_i(\mathbf{x}) y^i$ where f_i is a homogenous polynomial of degree $d - i$. Then $f(\mathbf{x}, 1) = \sum_{i=0}^{d} f_i(\mathbf{x})$. Note that if $f_i(\mathbf{x}) \neq 0$, then there are no cancellations from the any of the other $f_j(\mathbf{x})$'s since these polynomials have disjoint support. Thus it follows that $f_i(\mathbf{x}) = 0$ for $i \neq d - d'$ and $f_{d-d'}(\mathbf{x}) = m'$, thus yielding the proposition. ∎

We now present the main lemma of this section. To motivate the lemma, we first give an example. Consider the linear span of the monomial $x^5 \in \mathbb{F}[x, y]$. If the characteristic p of $\mathbb{F}$ is greater that 5 (or if $p = 3$), then L-SPAN(x^5) = SPAN$(\{x^5, x^4y, x^3y^2, x^2y^3, xy^4, y^5\})$. On the other hand, if $\mathbb{F}$ is of characteristic 5, the L-SPAN(x^5) = SPAN$(\{x^5, y^5\})$. If $\mathbb{F}$ is of characteristic 2, then L-SPAN(x^5) = SPAN$(\{x^5, x^4y, xy^4, y^5\})$. The lemma below attempts to capture some of this diversity.

Lemma 4.6 (Monomial Spread Lemma) *Let $\mathbf{d} = \langle d_1, \ldots, d_n \rangle \in \{0, \ldots, q-1\}^n$ and $\mathbf{e} = \langle e_1, \ldots, e_n \rangle \in \{0, \ldots, q-1\}^n$. For $i \in [n]$ and $j \in \{0, \ldots, s-1\}$ let d_{ij} and e_{ij} be the unique integers from $\{0, \ldots, p-1\}$ such that $d_i = \sum_{j=0}^{s-1} d_{ij} p^j$ and $e_i = \sum_{j=0}^{s-1} e_{ij} p^j$. Let m be the monomial $\mathbf{x}^{\mathbf{d}}$ and let $m' = \mathbf{x}^{\mathbf{e}}$. If for every $j \in \{0, \ldots, s-1\}$ it is the case that $\sum_{i=1}^{n} e_{ij} \leq \sum_{i=1}^{n} d_{ij}$, then the following hold:*

1. *$m' \in$ A-SPAN(m).*
2. *$y^{f-\deg(m')+\deg(m)} \cdot m' \in$ L-SPAN$(y^f \cdot m)$ for every non-negative f.*

Proof: We only prove Part (2). The affine case follows by setting $y = 1$ in the proof below. Alternately, one can make the general observation that if a monomial $y^a m'$ is contained in L-SPAN(m) for $m, m' \in \mathbb{F}[\mathbf{x}]$, then m' is contained in A-SPAN(m). Applying this observation to the conclusion from Part (2) of the lemma (with $f = 0$) yields Part (1).

We start with a simple claim that deals with the special case of the span of bivariate monomials. The lemma then follows by a simple induction using this claim.

Claim 4.7 *Let $k \in \{0, \ldots, q-1\}$ and $k_0, \ldots, k_{s-1} \in \{0, \ldots, p-1\}$ be such that $k = \sum_{j=0}^{s-1} k_j p^j$. Let ℓ be a non-negative integer and let $j_0 \in \{0, \ldots, s-1\}$ be such that $k_{j_0} > 0$. Then, the monomial $y^{\ell+p^{j_0}} \cdot x^{k-p^{j_0}}$ is contained in* L-SPAN $\left(y^\ell \cdot x^k\right)$.

Proof: Let $M(x, y) = y^\ell \cdot x^k$. We show below that $M(x + y, y)$ has the monomial $M(x, y) \cdot (y/x)^{p^{j_0}}$ in its support. The claim then follows by Lemma 4.2.

For a monomial $y^\ell \prod_{j=1}^{s} x^{a_j \cdot p^{j-1}}$, we say that $\sum_j a_j$ is its pseudo-degree. Then note that $M(x+y,y) = M(x,y) + \sum_{j\in[s]} k_j \cdot M(x,y)\cdot (y/x)^{p_{j-1}} + M'(x,y)$ where $M'(x,y)$ is a polynomial of pseudo-degree less than $(\sum_{j=1}^{s} k_j) - 1$. Thus the coefficient of $M(x,y)(y/x)^{p_{j_0-1}}$ in $M(x+y,y)$ is exactly k_{j_0} which is non-zero if $k_{j_0} \neq 0$. ∎

We now move to the proof of the lemma. We prove this lemma by induction on $\sum_{i,j} |d_{ij} - e_{ij}|$. We consider two cases:

CASE 1: $\exists j$ s.t. $\sum_i d_{ij} > \sum_i e_{ij}$: Let $\ell \in \{0,\ldots,s-1\}$ be such that $\sum_i d_{i\ell} > \sum_i e_{i\ell}$ and let $k \in [n]$ be such that $d_{k\ell} > e_{k\ell}$. Let $\tilde{d}_{ij}$ be given as follows: $\tilde{d}_{k\ell} = d_{k\ell} - 1$ and $\tilde{d}_{ij} = d_{ij}$ otherwise. Let $\tilde{d}_i = \sum_{j=0}^{s-1} \tilde{d}_{ij} p^j$. By Claim 4.7 we have that the monomial $y^{f+p^\ell} \cdot x_k^{\tilde{d}_k}$ is in the linear span of $y^f x_k^{d_k}$ and so (using Proposition 4.4) $y^{f+p^{\ell-1}} \prod_i x_i^{\tilde{d}_i}$ is in the linear span of $y^f \prod_i x_i^{d_i}$. By induction we also have that $y^{f+\deg(m)-\deg(m')}\mathbf{x}^{\mathbf{e}}$ is in the linear span of $y^{f+p^{j-1}}\mathbf{x}^{\mathbf{d}}$. Putting the two together we get the lemma in this case.

CASE 2: $\exists k_1, k_2, \ell$ s.t. $d_{k_1\ell} > e_{k_1\ell}$ and $d_{k_2\ell} < e_{k_2\ell}$: Now define $\tilde{d}$ to be $\tilde{d}_{k_1\ell} = d_{k_1\ell} - 1$, $\tilde{d}_{k_2\ell} = d_{k_2\ell} + 1$ and $\tilde{d}_{ij} = d_{ij}$ otherwise. Again, let $\tilde{d}_i = \sum_{j=0}^{s-1} \tilde{d}_{ij} p^j$. Applying Claim 4.7 to $x = x_{k_1}$ and $y = x_{k_2}$, we now have that the monomial $x_{k_1}^{\tilde{d}_{k_1}} \cdot x_{k_2}^{\tilde{d}_{k_2}}$ is contained in the linear span of $x_{k_1}^{d_{k_1}} \cdot x_{k_2}^{d_{k_2}}$. It follows (using Proposition 4.4) that $y^f \cdot \prod_{i=1}^{n} x_i^{\tilde{d}_i}$ is in the linear span of $y^f \mathbf{x}^{\mathbf{d}}$. Again, by induction, it also follows that $y^{f+\deg(m)-\deg(m')} \cdot \mathbf{x}^{\mathbf{e}'}$ is in the linear span of $y^f \mathbf{x}^{\tilde{\mathbf{d}}}$ We conclude that $y^{f+\deg(m)-\deg(m')} \cdot m'$ is in the linear span of $y^f \cdot m$. This yields the lemma statement for this case.

The lemma now follows since the two cases above are exhaustive. ∎

In the following corollary we describe some of the special cases that are used in later sections.

Corollary 4.8 *The following statements are true:*

1. *If $e_1,\ldots,e_n$ are non-negative integers such that $e_{n-1} + e_n < p$ then the monomial $x_1^{e_1}\cdots x_n^{e_n}$ is in the linear span of the monomial $x_1^{e_1}\cdots x_{n-2}^{e_{n-2}} \cdot x_{n-1}^{e_{n-1}+e_n}$.*

2. *If $q/p \leq d < q$ and f is an arbitrary integer then the monomial $x^{q/p}y^{f+d-q/p}$ is in the linear span of $x^d y^f$. and $x^{q/p}$ is in the affine span of x^d.*

3. *If $d_1 + \cdots + d_n \geq q/p$ and $f \geq 0$, then the monomial $y^{e+f}x_1^{q/p}$ is in the linear span of $y^f x_1^{d_1}\cdots x_n^{d_n}$ for $e = d_1 + \cdots d_n - q/p$, and $x_1^{q/p}$ is in the affine span of $x_1^{d_1}\cdots x_n^{d_n}$.*

Proof: We prove only the containments in the linear span. The affine part can be obtained by setting $y = 1$ in the proofs.

1. Part 1 is obtained as follows. Let $d_i = e_i$ for $i \in [n-2]$ and $d_{n-1} = e_{n-1} + e_n$ and $d_n = 0$. Let d_{ij}s be the unique integers such that $d_i = \sum_{j=1}^{s} d_{ij} p^{j-1}$, and let e_{ij} be defined analogously from e_i. Finally let $f = f' = 0$. Then it is clear that $e_{ij} = d_{ij}$ except possibly when $i \in \{n-1, n\}$ and $j = 1$. In these cases we have $d_{n-1,1} = e_{n-1} + e_n$ and $e_{n-1,1} = e_{n-1}$ and $e_{n,1} = e_n$, which also satisfies $d_{n-1,1} + d_{n,1} \leq e_{n-1,1} + e_{n,1}$. Also note that $f' = f + \sum_{j=1}^{s} p^{j-1} \sum_{i=1}^{n} (d_{ij} - e_{ij}) = 0$. Thus, by Lemma 4.6, we have that $x_1^{e_1}\cdots x_n^{e_n}$ is in the linear span of the monomial $x_1^{e_1}\cdots x_{n-2}^{e_{n-2}} \cdot x_{n-1}^{e_{n-1}+e_n}$.

2. For Part 2, we use Lemma 4.6 with d_{1j} such that $d = \sum_{j=1}^{s} d_{1j} p^{j-1}$ and $e_{1j} = 0$ except when $j = s$ in which case $e_{1s} = 1$. Since $d \geq q/p$ we have $d_{1s} \geq 1 = e_{1s}$, satisfying the condition of the corollary.

3. For Part 3, assume $d_1 \geq d_2 \geq \cdots \geq d_n$. Let $k \geq 1$ be the smallest index such that $\frac{q}{p} \leq \sum_{i=1}^{k} d_i < q$. (Note that such a k exists since either $q/p \leq d_1 < q$ in which case $k = 1$, or $d_i < q/p$ for every i and so for the first k such that $\sum_{i=1}^{k} d_i \geq q/p$, this sum is also less than $2q/p$.) Let $e = \sum_{i=1}^{k} d_i$. Consider the linear transformation that sets $x_1, \cdots, x_k$ to x_1 and $x_{k+1}, \cdots, x_n$ to y. This shows that the monomial $m_1 = {x_1}^e y^{\sum_{i=1}^{n} d_n - e + f}$ is in L-SPAN$(y^f x_1^{d_1} \cdots x_n^{d_n})$. Applying, Part 2 to m_1 we get this part.

∎

Lemma 4.9 *Let $m \in \mathbb{F}[\mathbf{x}, y]$ be a monomial of degree d. Let $\ell = \lfloor d/q \rfloor$. Then $\prod_{i=1}^{\ell} x_i^{q/p}$ is contained in* A-SPAN(m)*, Furthermore, $\{y^{d_1} \cdot m' | m' \in$ A-SPAN$(\prod_{i=1}^{\ell} x_i^{q/p}), d_1 + \deg(m') \equiv d \mod (q-1)\}$ is contained in* L-SPAN(m).

Proof: We prove only the containment for the assertion about A-SPAN(m) and the containment in the linear span follows using Proposition 4.5.
Let $m = x_1^{d_1} \cdots x_n^{d_n}$ and let $d_1 \geq d_2 \geq \cdots \geq d_n$. Partition the variables x_i into blocks where the total degree of the variables within each block (except at most one) is at least q/p and less than q. The number of full blocks (ones of total degree at least q/p) is at least ℓ. Inductively, with ℓ applications of Part 3 of Corollary 4.8 (and using Proposition 4.4), we get that $\prod_{i=1}^{\ell} x_i^{\frac{q}{p}}$ is in A-SPAN$(\mathbf{x^d})$. ∎

We also prove a characterization of affine-invariant families over prime fields, showing that a family of functions over a prime field is affine-invariant if and only it forms a "Generalized Reed-Muller code".

Corollary 4.10 *$\mathcal{F}$ is an affine invariant family mapping $\mathbb{F}_p^n \to \mathbb{F}_p$ if and only if there exists an integer d such that $\mathcal{F}$ is the family of all polynomials over $\mathbb{F}_p$ in n-variables of degree at most d.*

Proof: It is obvious that the set of degree d polynomials form an affine-invariant family, giving one direction. For the other direction, let d be the maximum degree of any polynomial in $\mathcal{F}$, and let m be the monomial of degree d in the support of this polynomial. Then by the Monomial extraction lemma $m \in \mathcal{F}$. Furthermore, using the affine part of Lemma 4.6 (with $s = 1$) we see that every monomial $\mathbf{x^e}$ of degree at most d is contained in $\mathcal{F}$. We conclude that every polynomial of degree d is in $\mathcal{F}$. Finally, $\mathcal{F}$ contains no other functions (since the highest degree of any polynomial in $\mathcal{F}$ is d). We conclude that $\mathcal{F}$ is the set of polynomials of degree d, as asserted. ∎

5 Bounding the Locality of Characterization for Aff/Lin

In this section we prove Theorems 2.10 and 2.11 for the special case when $\mathbb{K} = \mathbb{F}$. In the process we give upper and lower bounds on the locality of formal characterizations of affine-invariant and linear-invariant families, in terms of the degree patterns of the monomials in their support.
Our (upper bounds on) characterizations are obtained by considering the values of a given function on some small dimensional subspace and verifying that these values agree with the values of some function in the family. Keeping this in mind, we define the restriction of a function family to a smaller dimension.

Definition 5.1 (Projections of function families) *For positive integers ℓ and n, and for a linear-invariant family of functions $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$, the ℓ-dimensional restriction (extension) of $\mathcal{F}$, denoted $\mathcal{F}|_\ell$ is the family $\mathcal{F}|_\ell = \{f \circ L | f \in \mathcal{F}, L : \mathbb{K}^\ell \to \mathbb{K}^n$ linear $\}$.*

Note that we don't insist that $\ell \leq n$ and indeed the definition above makes sense also in this case. However in all our usage below, we think of $\ell \leq n$.

For affine-invariant families our characterizations depend simply on the maximum degree of functions in the family. For linear invariant functions this is no longer true. For instance, the family of functions supported on all monomials in $x_1, \ldots, x_n$ of degree $3 \mod 4$ over $\mathbb{F}_5$ has a 2-local characterization even though it contains polynomials of degree $\Omega(n)$. For linear-invariant families, the characterizations depend on a more refined parameter that we define next.

Definition 5.2 *For a linear invariant family $\mathcal{F}$ properly contained in $\{\mathbb{F}^n \to \mathbb{F}\}$, let $d_{\text{lin}}(\mathcal{F})$, the* linear-invariance degree *of $\mathcal{F}$, be the largest integer d such that $\mathcal{F}$ contains a monomial m_1 of degree d, while there also exists a monomial $m_2 \notin \mathcal{F}$ of degree d' for some $d' > 0$ with $d' \equiv d(\mod q-1)$.*

5.1 Upper bounds on locality of characterizations

The next lemma is the crux of our characterizations for linear-invariant as well as affine-invariant families.

Lemma 5.3 . *Let $\mathcal{F} \subseteq \{\mathbb{F}^n \to \mathbb{F}\}$ be a linear-invariant family of linear-invariance degree $d_{\text{lin}}(\mathcal{F}) = d$. Suppose $f : \mathbb{F}^n \to \mathbb{F}$ is not in $\mathcal{F}$. Then, if $n \geq 1 + \left(\frac{2}{p} \cdot (d+q)\right)$, then there exists a linear function $L : \mathbb{F}^{n-1} \to \mathbb{F}^n$ such that $f \circ L \notin \mathcal{F}|_{n-1}$.*

Proof: Let $m = x_1^{d_1} \cdots x_n^{d_n}$ be a monomial of maximal degree in the support of f that is not contained in $\mathcal{F}$. We show that there is a linear map $L : \mathbb{F}^{n-1} \to \mathbb{F}^n$ such that $m \circ L$ is not in $\mathcal{F}|_{n-1}$. We consider two cases:

Case 1: *There exist distinct indices i, j such that $d_i + d_j < p$*: Without loss of generality assume $i = n-1$ and $j = n$. Let m' be the monomial $m' = x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^{d_{n-1}+d_n}$. First note by the Monomial Spread Lemma (in particular, by Part 1 of Corollary 4.8) that $m \in \text{L-SPAN}(m')$. So $m' \notin \mathcal{F}$ and hence $m' \notin \mathcal{F}|_{n-1}$. We claim that for some choice of $\alpha, \beta \in \mathbb{F}$, the map $L_{\alpha,\beta}(x_1, \ldots, x_{n-1}) = \langle x_1, \ldots, x_{n-2}, \alpha x_{n-1}, \beta x_{n-1} \rangle$ leaves the monomial m' with non-zero support in $f \circ L_{\alpha,\beta}$, which would suffice to prove the lemma (in this case).

To see this, let c_i be the coefficient of the monomial $x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^i x_n^{d_{n-1}+d_n-i}$ in f. Let $h(x,y) = \sum_{i=0}^{d_{n-1}+d_n} c_i x^i y^{d_{n-1}+d_n-i}$. It can be verified that the coefficient of m' in $f \circ L_{\alpha,\beta}$ is exactly $h(\alpha, \beta)$. Furthermore, $h(\alpha, \beta)$ is a non-zero polynomial since the coefficient $c_{d_{n-1}}$ is the coefficient of m in f which is non-zero. Thus there must exist α, β such that $h(\alpha, \beta) \neq 0$ and this yields the claim.

Case 2: *For every pair of distinct i, j, $d_i + d_j \geq p$.* Let e denote the degree of m. For every t, we have $d_{2t-1} + d_{2t} \geq p$, and so e, the total degree of m, is at least $p\lfloor n/2 \rfloor \geq d + q$.

We first note that no monomial m' of degree e or $e-(q-1)$ is in $\mathcal{F}$. Otherwise the linear-invariance degree of $\mathcal{F}$ would be the degree of m'. For example, if m' has degree $e-(q-1) > d$, then m' satisfies the role of the monomial m_1 in the definition of the linear-invariance degree and m of degree $e = e-(q-1)(\mod q-1)$ satisfies the role of m_2 in the definition of linear-invariance degree thereby yielding $d_{\text{lin}}(\mathcal{F}) = e-(q-1) > d$. So we conclude m' can not be in $\mathcal{F}$.

But now consider the linear map $L_{\alpha,\beta}$ as in the previous case, i.e., $L_{\alpha,\beta}(x_1,\ldots,x_{n-1}) = \langle x_1,\ldots,x_{n-2},\alpha x_{n-1},\beta x_{n-1}\rangle$. Let m' be the monomial $x_1^{d_1}\cdots x_{n-2}^{d_{n-2}}\cdot x_{n-1}^{f}$ where $f = d_{n-1}+d_n$ if $d_{n-1}+d_n < q$ and $f = d_{n-1}+d_n-(q-1)$ otherwise. The degree of m' is thus e or $e-(q-1)$. We claim that for some α,β, the coefficient of m' is non-zero in $f\circ L_{\alpha,\beta}$ and this will yield the lemma in this case.

To verify the claim, note that the coefficient of m' in $f\circ L_{\alpha,\beta}$ is exactly $h'(\alpha,\beta)$ where $h'(x,y) = \sum_{i=0}^{\min\{q-1,d_n+d_{n-1}\}} c_i x^i y^{d_n+d_{n-1}-i}$ and c_i is the coefficient of the monomial $x_1^{d_1}\cdots x_{n-2}^{d_n-2}\cdot x_{n-1}^{i}x_n^{d_n+d_{n-1}-i}$ in f. Again, we have $h'(x,y)$ is not identically zero (the coefficient $c_{d_{n-1}}$ is non-zero) and so there exists α,β such that $h'(\alpha,\beta)\neq 0$. ∎

We can now give a characterization for linear-invariant families.

Lemma 5.4 *Let $\mathcal{F}$ be a linear invariant family, properly contained in $\{\mathbb{F}^n\to\mathbb{F}\}$, of linear-invariance degree $d_{\max}$. Then $\mathcal{F}$ has a q^ℓ-local formal characterization for $\ell = \frac{2(d+q)}{p}$.*

Proof: We claim that the characterization is simply the one that $f\in\mathcal{F}$ if and only if $f\circ L$ is in $\mathcal{F}|_\ell$ for every linear map $L:\mathbb{F}^\ell\to\mathbb{F}^n$.

It is easy to verify that (if this indeed characterizes the family correctly) this is a q^ℓ-local *formal* characterization. We analyze the correctness below.

In one direction, it is obvious that every $f\in\mathcal{F}$ and $L:\mathbb{F}^\ell\to\mathbb{F}^n$ satisfies $f\circ L\in\mathcal{F}|_\ell$. The other direction is a simple induction based on Lemma 5.3. Fix $f\notin\mathcal{F}$. Let m be any integer between $\ell+1$ and n. Assume by induction on $n-m$ that there a linear map $L_m:\mathbb{F}^m\to\mathbb{F}^n$ such that $f\circ L_m\notin\mathcal{F}|_m$. Now we prove that there is a map $L_{m-1}:\mathbb{F}^{m-1}\to\mathbb{F}^n$ such that $f\circ L_{m-1}\notin\mathcal{F}|_{m-1}$. Since $\mathcal{F}$ is linear-invariant, so is $\mathcal{F}|_m$. Also the linear-invariance degree of $\mathcal{F}|_m$ is at most $d_{\max}$. By Lemma 5.3 there is a linear map $L:\mathbb{F}^{m-1}\to\mathbb{F}^m$ such that $(f\circ L_m)\circ L\notin(\mathcal{F}|_m)|_{m-1} = \mathcal{F}|_{m-1}$. Thus $f\circ L_{m-1}\notin\mathcal{F}|_{m-1}$ for $L_{m-1} = L_m\circ L$. We conclude that the linear map $L_\ell:\mathbb{F}^\ell\to\mathbb{F}^n$ derived from setting $m=\ell+1$, satisfies $f\circ L_\ell\notin\mathcal{F}|_\ell$. ∎

Immediately, we also get a characterization for affine-invariant families (since every affine invariant family with polynomials of degree at most $d_{\max}$ is also a linear-invariant family of linear-invariance degree at most $d_{\max}$).

Lemma 5.5 *Let $\mathcal{F}$ be a proper subset of $\{\mathbb{F}^n\to\mathbb{F}\}$ and let $d_{\max}$ denote the maximum degree of any function in $\mathcal{F}$. Then $\mathcal{F}$ has a q^ℓ-local formal characterization for $\ell\le\frac{2(d+q)}{p}$.*

5.2 Lower bounds on locality of characterizations for affine-invariant families

We now turn to proving lower bounds on the locality of constraints (and thus characterizations) in affine-invariant families. The lower bound is eventually derived from the study of Generalized Reed-Muller codes where it is known that the family of polynomials of degree d has no $q^{\lfloor d/q\rfloor}$-local characterizations. Specifically we have:

Lemma 5.6 ([14, 8]) *$\mathcal{F} = \text{A-SPAN}(\prod_{i=1}^d x_i)$ has no $q^{\lfloor d/q\rfloor}$-local constraints.*

Proof: Note that every monomial $\mathbf{x}^{\mathbf{e}}$ of degree at most d is contained in $\mathcal{F}$. Suppose $\mathbf{e} = \langle e_1,\ldots,e_n\rangle$ then we can substitute x_j for e_j variables in $\prod_i x_i$ for every j, and substitute 1 for the remaining variables to get an affine transormation that transforms $\prod_{i=1}^d x_i$ to $\mathbf{x}^{\mathbf{e}}$. Thus the family $\mathcal{F}_1$ of d-variate polynomials of degree at most d is contained in $\mathcal{F}$.

We can now invoke well-known results from the study of the "Generalized Reed-Muller codes", in particular [14, Theorem 5] (see also [8, Theorems 2.2.1 and 2.6.2]), which state that the class of degree d polynomials in d variables have no constraints of locality $q^{\lfloor d/q \rfloor}$. In other words $\mathcal{F}_1$ has no $q^{\lfloor d/q \rfloor}$-local constraints. Using Claim 5.8 we get that $\mathcal{F}$ also has no $q^{\lfloor \frac{d}{q} \rfloor}$-local constraints. ∎

We are now ready to prove lower bounds on the locality of constraints in affine-invariant families.

Lemma 5.7 *Let $\mathcal{F}$ be an affine invariant family properly contained in $\{\mathbb{F}^n \to \mathbb{F}\}$ containing a polynomial of degree d. Then $\mathcal{F}$ has no q^ℓ-local constraints for $\ell \leq (d - q^2)/q^2$.*

Proof: Before proving the lower bound, we provide generic conditions under which the absence of local constraints in one family of functions imply the absence of local constraints in another family.

Claim 5.8 *Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be non-trivial families of functions from $\mathbb{F}^n \to \mathbb{F}$. Suppose $\mathcal{F}_1$ has no k-local constraints. Then, if there exists a function $g : \mathbb{F}^n \to \mathbb{F}^n$ such that for every $f \in \mathcal{F}_1$ it is the case that $f \circ g \in \mathcal{F}_2$, then $\mathcal{F}_2$ also has no k-local constraints. In particular, if $\mathcal{F}_1 \subseteq \mathcal{F}_2$ then $\mathcal{F}_2$ has no k-local constraints.*

Proof: Suppose $\mathcal{F}_2$ has a k-local constraint of the form $\langle x_1, \ldots, x_k; S \rangle$ where $x_i \in \mathbb{F}^n$ and S is a proper subset of $\mathbb{F}^k$. (I.e., $f \in \mathcal{F}_2$ implies $\langle f(x_1), \ldots, f(x_k) \rangle \in S$ for every $f \in \mathcal{F}_2$.) Then we can use g to translate this into the constraint $\langle g(x_1), \ldots, g(x_k); S \rangle$ for $\mathcal{F}_1$ (since $\langle h(g(x_1)), \ldots, h(g(x_k)) \rangle \in S$ for every $h \in \mathcal{F}_1$), which would be a contradiction.
In particular, if $\mathcal{F}_1 \subseteq \mathcal{F}_2$, then using the identity function $g(x) = x$, we get that $\mathcal{F}_2$ has no k-local constraints. ∎

We now apply Claim 5.8 to the conclusion of Lemma 5.6 to derive a lower bound on the constraints of a family of functions that is slightly more convenient for us to work with.

Claim 5.9 *The family $\mathcal{F} = \text{A-SPAN}(\prod_{i=1}^d x_i^{\frac{q}{p}})$ has no $q^{\lfloor d/q \rfloor}$-local constraints.*

Proof: By Lemma 5.6 we have that $\mathcal{F}_1 = \text{A-SPAN}(\prod_{i=1}^d x_i)$ has no $q^{\lfloor d/q \rfloor}$-local constraints. Let $g(x_1, \ldots, x_n) = \langle x_1^{q/p}, \ldots, x_n^{q/p} \rangle$. Note that $g^{-1}(x_1, \ldots, x_n) = \langle x_1^p, \ldots, x_n^p \rangle$. Note that for every $n \times n$ matrix A and vector $b \in \mathbb{F}^n$, we have $Ag(x) + b = g(g^{-1}(A)x + g^{-1}(b))$ (where $g^{-1}(A)$ simply applies g^{-1} to every column of A). This implies that every $f \in \mathcal{F}_1 = \text{A-SPAN}(\prod_i x_i)$ satisfies $f \circ g \in \mathcal{F}$. So we can apply Claim 5.8 to conclude that $\mathcal{F}$ also has no $q^{\lfloor d/q \rfloor}$-local constraints. ∎

We ready to prove Lemma 5.7. Recall that we are given a family $\mathcal{F}$ with some monomial, say m_1, of degree d. By Lemma 4.9 the monomial m_1 has $\mathcal{F}_1 = \prod_{i=1}^\ell x_i^{q/p}$ in its affine span for $\ell = \lfloor d/q \rfloor$. By Claim 5.9, we have that $\mathcal{F}_1$ has no $q^{\lfloor \ell/q \rfloor}$-local constraints. Since $\mathcal{F} \supset \mathcal{F}_1$, we can now apply Claim 5.8 again (with the identity function g) to conclude that $\mathcal{F}$ has no $q^{\lfloor \ell/q \rfloor}$-local constraints either. The lemma follows using the fact that $\lfloor \lfloor d/q \rfloor / q \rfloor \geq (d - q^2)/q^2$. ∎

5.3 Lower bounds for Linear-Invariant Families

In this section we provide lower bounds on the locality of characterizations of linear-invariant families, based on their "linear-invariance degree" (see Definition 5.2). As shown in Section 5.1, this parameter also yields upper bounds and thus together we find that this parameter governs (in some weak sense, since the bounds are far apart) the locality of characterizations for linear-invariant families.

In order to understand the locality of characterizations, we introduce the notion of a constraint on a family $\mathcal{F}_1$ relative to a family $\mathcal{F}_2$.

Definition 5.10 *For families $\mathcal{F}_1, \mathcal{F}_2 \subseteq \{\mathbb{F}^n \to \mathbb{F}\}$, with $\mathcal{F}_1 \subsetneq \mathcal{F}_2$, we say that a constraint C is a constraint on $\mathcal{F}_1$* relative to *$\mathcal{F}_2$ if every function $f \in \mathcal{F}_1$ satisfies C and there exists a function $g \in \mathcal{F}_2$ that does not satisfy C.*

The following straightforward fact explains the relevance of constraints relative to other families when analyzing characterizations.

Proposition 5.11 *If $C_1, \ldots, C_m$ form a characterization of $\mathcal{F}_1$, then for every family $\mathcal{F}_2 \supsetneq \mathcal{F}_1$, there exists an index j such that the constraint C_j is a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$.*

In what follows, we will consider a family $\mathcal{F}$ of linear invariance degree d. We will construct families $\mathcal{F}_1$ and $\mathcal{F}_2$ related to d such that $\mathcal{F}_1$ has no constraints of small locality relative to $\mathcal{F}_2$. We will then use reductions to tranfer this result to showing that $\mathcal{F}$ has no constraints of small locality relative to some family $\mathcal{F}_3$ which will yield a lower bound on the locality of its characterizations.

Throughout this section we will consider functions from $\mathbb{F}^{n+1} \to \mathbb{F}$, and we will associate them with polynomials from $\mathbb{F}[\mathbf{x}, y]$ where $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ is a collection of n variables.

For a set of functions $\mathcal{G} \subseteq \{\mathbb{F}^m \to \mathbb{F}\}$, let $\textsc{Span}(\mathcal{G})$ denote the span of the functions in $\mathcal{G}$, i.e., $\textsc{Span}(\mathcal{G}) = \{\sum_{i=1}^{t} \alpha_i g_i | \alpha_i \in \mathbb{F}, g_i \in \mathcal{G}\}$.

Lemma 5.12 *Let d and ℓ be positive integers and let $\mathcal{F}_1 = \textsc{Span}(\{y^{d_1} \cdot m | m \in \text{A-}\textsc{span}(\prod_{i=1}^{\ell} x_i^{q/p}), d_1 + \deg(m) \equiv d \mod (q-1), 1 \leq d_1 \leq q-1\})$. Let $\mathcal{F}_2 = \textsc{Span}\{m | m$ monomial in $\mathbb{F}[\mathbf{x}, y], \deg(m) \equiv d \mod (q-1), \deg_y(m) \geq 1\}$, be the collection of all polynomials supported on monomials of degree $d \mod (q-1)$, with positive degree in y. Then $\mathcal{F}_1$ has no constraints of locality $q^{\lfloor \ell/q \rfloor}$ relative to $\mathcal{F}_2$.*

Proof: Note by the definitions of $\mathcal{F}_1$ and $\mathcal{F}_2$ that $\mathcal{F}_1 \subseteq \mathcal{F}_2$. (In particular the degree in y of every monomial in the support of $\mathcal{F}_1$ is positive.) If $\mathcal{F}_1 = \mathcal{F}_2$ then the claim is trivial since there can be no function in $\mathcal{F}_2 - \mathcal{F}_1$ and so none violating any given constraint. So assume $\mathcal{F}_1 \subsetneq \mathcal{F}_2$.

Let $C = (\mathbf{z}_1, \ldots, \mathbf{z}_k; S)$, where S, be a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$. We will show that $k \geq q^{\ell/q}$. To show this we will map (most points of) $\mathbb{F}^{n+1}$ to $\mathbb{F}^n$ in a way that maps homogenous polynomials of positive degree in y to generic polynomials over $\mathbf{x}$.

For a point $\mathbf{z} = \langle x_1, \ldots, x_n, y \rangle \in \mathbb{F}^{n+1}$, let $\pi(\mathbf{z}) = \langle x_1/y, \ldots, x_n/y \rangle$ if $y \neq 0$ and some special symbol $\perp$ if $y = 0$.

Note that for any function $f \in \mathcal{F}_2$ and point $\mathbf{z} \in \mathbb{F}^{n+1}$, $f(\mathbf{z}) = 0$ if $\pi(\mathbf{z}) = \perp$. Further, note that if $\pi(\mathbf{z}_1) = \pi(\mathbf{z}_2)$ then there exists a $\lambda \in \mathbb{F} - \{0\}$ such that $\mathbf{z}_2 = \lambda \mathbf{z}_1$ and $f(\mathbf{z}_2) = \lambda^d f(\mathbf{z}_1)$. We use these observations to "simplify" the constraint C while maintaining the property that it remains a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$.

First note that we can assume w.l.o.g. that $\pi(\mathbf{z}_i) \neq \perp$ for every $i \in [k]$. To see this, suppose $\pi(\mathbf{z}_k) = \perp$. Then it can be verified that the constraint $C' = (\mathbf{z}_1, \ldots, \mathbf{z}_{k-1}; S')$ is a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$, where $S' = \{\langle \alpha_1, \ldots, \alpha_{k-1}\rangle | \langle \alpha_1, \ldots, \alpha_{k-1}, 0\rangle \in S\}$. (We omit the simple verification steps.)

Next, we also note that we can assume that $\pi(\mathbf{z}_i)$'s are all distinct for distinct $i \in [k]$. Again to see this, suppose $\pi(\mathbf{z}_{k-1}) = \pi(\mathbf{z}_k)$. Then it must be that $\mathbf{z}_k = \lambda \mathbf{z}_{k-1}$. Once again it can be verified that the constraint $C' = (\mathbf{z}_1, \ldots, \mathbf{z}_{k-1}; S')$ is a constraint on $\mathcal{F}_1'$ relative to $\mathcal{F}_2$, where $S' = \{\langle \alpha_1, \ldots, \alpha_{k-1}\rangle | \langle \alpha_1, \ldots, \alpha_{k-1}, \lambda^d \alpha_{k-1}\rangle \in S\}$. (Again, we omit the simple verification steps.)

Note that in the "simplification" process above, we may have potentially lost the property that $S \neq \mathbb{F}^k$. But we note that the fact that C is a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$ implies $S \neq \mathbb{F}^k$ as follows: Suppose C is violated by some function $g \in \mathcal{F}_2$. Then we have that $\langle g(\mathbf{z}_1), \ldots, g(\mathbf{z}_k)\rangle \notin S$ and so $S \neq \mathbb{F}^k$.

So we may now assume that the $\pi(\mathbf{z}_i)$'s are distinct elements of $\mathbb{F}^n$ and that $S \neq \mathbb{F}^k$. Let $\mathbf{z}_i = \langle x_{i1}, \ldots, x_{in}, y_i\rangle$. Now consider the constraint $C' = (\pi(\mathbf{z}_1), \ldots, \pi(\mathbf{z}_k), S')$, for $S' = \{\langle \alpha_1/y_1^d, \ldots, \alpha_k/y_k^d\rangle | \langle \alpha_1, \ldots, \alpha_k\rangle \in S\}$. Since $S \neq \mathbb{F}^k$, we also have $S' \neq \mathbb{F}^k$. We claim that C' is a k-local constraint on the family A-SPAN$(\prod_{i=1}^{\ell} x_i^{q/p})$. To verify this claim, we need to show that every $f \in$ A-SPAN$(\prod_{i=1}^{\ell} x_i^{q/p})$ satisfies $\langle f(\pi(\mathbf{z}_i)), \ldots, f(\pi(\mathbf{z}_k)\rangle \in S'$.

Consider the following map from polynomials in $\mathbb{F}[\mathbf{x}]$ to $\mathcal{F}_2$, where a monomial $m \in \mathbb{F}[\mathbf{x}]$ is mapped to the monomial $\hat{m} = m \cdot y^i$ where $i \in [q-1]$ is chosen so that $\deg(m) + i = d \mod (q-1)$. This map can be extended linearly to every polynomial $\mathbb{F}[\mathbf{x}]$ mapping the polynomial p to $\hat{p}$. Note that since functionally $y^{j(q-1)+i} = y^i$ we can w.l.o.g. think of the monomial $\hat{m}$ as having degree $\geq d$. In particular for monomials from A-SPAN$(\prod_{i=1}^{\ell} x_i^{q/p})$ the corresponding monomial has degree exactly d. Thus, for any function $f \in$ A-SPAN$(\prod_{i=1}^{\ell} x_i^{q/p})$, the corresponding function $\hat{f} \in \mathcal{F}_1$. Thus we have that $\hat{f}$ satisfies the constraint C, i.e., $\langle \hat{f}(\mathbf{z}_1), \ldots, \hat{f}(\mathbf{z}_k)\rangle \in S$. By the definition of π and $\hat{f}$, we have that $\hat{f}(\mathbf{z}_i) = y_i^d \cdot f(\pi(\mathbf{z}_i))$. Thus $\langle y_1^d \cdot f(\pi(\mathbf{z}_1)), \ldots, y_k^d \cdot f(\pi(\mathbf{z}_k))\rangle \in S$ and so Thus $\langle f(\pi(\mathbf{z}_1)), \ldots, f(\pi(\mathbf{z}_k))\rangle \in S'$.

Thus C' is a non-trivial constraint on A-SPAN$(\prod_{i=1}^{\ell} x_i^{q/p})$ and so, by Claim 5.9 $k > q^{\lfloor \ell/q \rfloor}$. ∎

Lemma 5.13 *Let $\mathcal{F} \subsetneq \{\mathbb{F}^{n+1} \to \mathbb{F}\}$ be a family of linear invariance degree d. Then $\mathcal{F}$ has no characterizations of locality $q^{(d-q^2)/q^2}$.*

Proof: Let $m \in \mathcal{F} \subset \mathbb{F}[\mathbf{x}, y]$ be a monomial of degree d. Let $m' \in \mathbb{F}[\mathbf{x}, y]$ be a monomial of degree $d' \equiv d \mod (q-1)$ such that $m' \notin \mathcal{F}$. (Such monomials exists, by the definition of linear-invariance degree.) Assume without loss of generality that $\deg_y(m') > 0$ (since we could rename variables to achieve this).

Let $\ell = \lfloor d/q \rfloor$. Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be as in Lemma 5.12, so that $\mathcal{F}_1$ has no constraints of locality $q^{\lfloor \ell/q \rfloor}$ relative to $\mathcal{F}_2$. Note first that by Lemma 4.9 we have that $\mathcal{F}_1$ is contained in L-SPAN$(m) \subseteq \mathcal{F}$.

Let $\mathcal{F}_3 = \mathcal{F} + \mathcal{F}_2$ consist of all functions $\{\alpha f + \beta g | f \in \mathcal{F}, g \in \mathcal{F}_2, \alpha, \beta \in \mathbb{F}\}$.

Note that $\mathcal{F} \subsetneq \mathcal{F}_3$. The containment is by definition, while the propriety of the containment follows from the fact that $m' \in \mathcal{F}_3 - \mathcal{F}$.

We now claim that $\mathcal{F}$ has no $q^{\lfloor \ell/q \rfloor}$-local constraints relative to $\mathcal{F}_3$ and this (combined with Proposition 5.11) yields the lemma.

Suppose $C = (\mathbf{z}_1, \ldots, \mathbf{z}_k; S)$ is a constraint on $\mathcal{F}$ relative to $\mathcal{F}_3$. Without loss of generality, we can assume that S is a $\mathbb{F}$-linear subspace of $\mathbb{F}^k$ (since $\mathcal{F}$ is a linear subspace) [3]. On the one hand, since $\mathcal{F}_1 \subseteq \mathcal{F}$ we have that C is also a constraint on $\mathcal{F}_1$. We now claim that C is actually a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$. Now let $h = \alpha f + \beta g \in \mathcal{F}_3$ not satisfy C, where $f \in \mathcal{F}$ and $g \in \mathcal{F}_2$. Let $v_f = \langle f(\mathbf{z}_1), \ldots, f(\mathbf{z}_k)\rangle$, $v_g = \langle g(\mathbf{z}_1), \ldots, g(\mathbf{z}_k)\rangle$, and $v_h = \langle h(\mathbf{z}_1), \ldots, h(\mathbf{z}_k)\rangle$. Then we have $v_h = \alpha v_f + \beta v_g$. On the one hand,

we have $v_f \in S$ (since $f \in \mathcal{F}$) and on the other, we have $v_h \notin S$. Since S is a linear subspace, it must be that $v_g \notin S$. Thus $g \in \mathcal{F}_2$ violates C and so C is a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$. By Claim 5.9 we have $k > q^{\lfloor \ell/q \rfloor}$. ∎

5.4 Testing Linear Invariant Families

The formal characterization described in Section 5.1 can immediately be turned into an affine invariant characterization for affine-invariant families. Coupled with Theorem 2.9 this leads immediately to a tester for affine-invariant families. However the characterization does not immediately lead to a tester for linear-invariant families, since these characterizations are not necessarily 2-ary independent. In this section we fix this gap.

We start with a definition that isolates a seemingly problematic subclass of linear-invariant families, where the characterizations are necessarily not 2-ary independent.

Definition 5.14 *A linear invariant family $\mathcal{F} \subseteq \{\mathbb{F}^n \to \mathbb{F}\}$ is said to be* projective *if, for every pair of monomials $\mathbf{x}^{\mathbf{d}}$ and $\mathbf{x}^{\mathbf{e}}$ with $\sum_{i=1}^n d_i \equiv \sum_{i=1}^n e_i \mod (q-1)$, it is the case that $\mathbf{x}^{\mathbf{d}}$ is in the support of $\mathcal{F}$ if and only if $\mathbf{x}^{\mathbf{e}}$ is in the support of $\mathcal{F}$.*

Projective families have a very simple local formal characterization, which is unfortunately not 2-ary independent, as described below.

Proposition 5.15 *A family $\mathcal{F}$ is projective if and only if there exists a set of monomials $S \subseteq \{x^0, x^1, \ldots, x^{q-1}\}$ on a single variable x such that the following holds: $f \in \mathcal{F}$ if and only if for every 1-dimensional linear function $L : \mathbb{F} \to \mathbb{F}^n$, the support of $f \circ L$ is contained in S.*

Proof: Let D be the set of degrees of monomials in the support of $\mathcal{F}$ reduced modulo $q-1$ (i.e., to the set $\{1, \ldots, q-1\}$, except if the monomial $\mathbf{x}^{\mathbf{0}}$ is in the support of $\mathcal{F}$, in which case we include 0 in the set D). Let $S = \{x^i | i \in D\}$.

On the one hand, it is clear that that if $f \in \mathcal{F}$ then $f \circ L$ has its support in S for every linear function $L : \mathbb{F} \to \mathbb{F}^n$. For the reverse direction, we reason as in the proof of Case 2 of Lemma 5.3. Let f be a polynomial not in $\mathcal{F}$ and let m be a monomial of maximal degree in the support of f that is not in $\mathcal{F}$. Suppose the degree of m is d. By the definition of projective families, we have that $d \mod (q-1) \notin D$. We first note that there is a linear function $L_n : \mathbb{F}^{n-1} \to \mathbb{F}^n$ such that $f \circ L_n$ has a monomial in its support of degree d or $d-(q-1)$. In either case the degree of this monomial (modulo $(q-1)$) is not in D. We continue this way to find a sequence of linear functions $L_i : \mathbb{F}^{i-1} \to \mathbb{F}^i$ such that for $L = L_n \circ \cdots \circ L_2$ it is the case that $f \circ L$ has a monomial in its support of degree not in D. ∎

Even though projective families do not have a 2-ary independent linear characterization, they turn out to have a simple local test: Namely pick a random line $L : \mathbb{F} \to \mathbb{F}^n$ and verify $f \circ L$ has its support in S. We won't prove the correctness of this test right now (it will follow from the general case). Instead we turn to showing that every linear invariant family can be written as the sum of a nice family (with a 2-ary independent formal characterization) and a projective family and this ends up leading to a test.

Lemma 5.16 *Let $\mathcal{F}$ be a linear-invariant family of linear invariance degree d. Then there exists a linear-invariant family $\mathcal{F}_1$ containing polynomials of degree at most d, and a projective family $\mathcal{F}_2$ such that $\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2$. Furthermore given an oracle to a function $f : \mathbb{F}^n \to \mathbb{F}$ one can construct an oracle for a function $g : \mathbb{F}^n \to \mathbb{F}$ where the oracle for g makes q oracle calls to f, such that $g \in \mathcal{F}_1$ if $f \in \mathcal{F}$ and $\delta(f, \mathcal{F}) \leq \delta(g, \mathcal{F}_1)$.*

Proof: The first part follows from the definition of linear invariance degree. Let $D \subseteq \{1, \ldots, q-1\}$ be given by $D = \{i | \exists$ monomial $m \in \mathcal{F}$ with $\deg(m) > d$ and $\deg(m) \equiv i \mod (q-1)\}$. Let $\mathcal{F}_1$ be the span of the set of monomials in $\mathcal{F}$ of degree at most d. Let $\mathcal{F}_2$ be the span of set of the monomials m' of degree $\deg(m') \equiv i \mod (q-1)$ for some $i \in D$. By the definition of linear-invariance degree, we have that every monomial m' of degree $\deg(m') \equiv i \mod (q-1)$ for some $i \in D$ is contained in $\mathcal{F}$ and so $\mathcal{F}_2 \subseteq \mathcal{F}_2$. It thus follows that $\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2$.

For the second part, we define $g : \mathbb{F}^n \to \mathbb{F}$ as follows. Given $\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle \in \mathbb{F}^n$, Let $f_\alpha(t) = f(t \cdot \alpha)$. Further, let $c_0, \ldots . c_{q-1}$ be such that $f_\alpha(t) = \sum_{i=0}^{q-1} c_i t^i$. Finally let $\tilde{f}_\alpha(t) = \sum_{i \notin D} c_i t^i$. We define $g(\alpha) = \tilde{f}_\alpha(1)$. Note by the definition of g that computing g at any point only requires q oracle calls to the oracle for f.

We claim that $f - g \in \mathcal{F}_2$. This is verified by noting that for linear functions $L : \mathbb{F} \to \mathbb{F}^n$, the function $(f-g) \circ L$ has all of its support on monomials with degree in D. (For instance if $L(t) = t \cdot \alpha$, then $(f-g) \circ L = f_\alpha(t) - \tilde{f}_\alpha(t) = \sum_{i \in D} c_i t^i$.) By Proposition 5.15, it follows that $f - g \in \mathcal{F}_2$.

It is immediate that $\delta(f, \mathcal{F}) = \delta(g, \mathcal{F}) \leq \delta(g, \mathcal{F}_1)$. To see that if $f \in \mathcal{F}$ then $g \in \mathcal{F}_1$, note that $g \circ L$ always has its monomials from $\{0, \ldots, q-1\} - D$. Applying Proposition 5.15, we find that $g \in \mathcal{F}'$ where $\mathcal{F}'$ is the projective space consisting of the span of monomials whose degree, modulo $q-1$, is in $\{0, \ldots, q-1\} - D$. But since $g \in \mathcal{F}$ and the only monomials in $\mathcal{F}$ whose degree modulo $q-1$ is not in D, are those of degree at most d, we conclude that g is of degree at most d and hence $g \in \mathcal{F}_1$. ∎

Finally we use a simple proposition that can be used to give 2-ary independent localy characterizations for family $\mathcal{F}_1$ above.

Proposition 5.17 *Let $\mathcal{F} \subseteq \mathcal{F}'$ have a k_1-local formal characterization. Furthermore suppose $\mathcal{F}'$ has a 2-ary independent k_2-local formal characterization. Then $\mathcal{F}$ has a $k_1 + k_2$-local 2-ary independent formal characterization.*

Proof: Let $m, \ell_1, \ldots, \ell_{k_1}, V$ describe the characterization of $\mathcal{F}$ i.e., $f \in \mathcal{F}$ iff for every $x_1, \ldots, x_m \in \mathbb{K}^n$ it is the case that $\langle f(y_1), \ldots, f(y_{k_1}) \rangle \in V$ for $y_i = \ell_i(x_1, \ldots, x_m)$.

Similarly let $m', \ell'_1, \ldots, \ell'_{k_2}, V'$ denote the characterization of $\mathcal{F}'$.

Then we claim that the characterization $m' + m, \tilde{\ell}'_1, \ldots, \tilde{\ell}'_{k_2}, \tilde{\ell}_1, \ldots, \tilde{\ell}_{k_1}, \tilde{V}$ forms a 2-ary independent characterization of $\mathcal{F}$, where

- $\tilde{\ell}'_i(z_1, \ldots, z_{m'}, x_1, \ldots, x_m) = \ell'_i(z_1, \ldots, z_{m'})$,
- $\tilde{\ell}_i(z_1, \ldots, z_{m'}, x_1, \ldots, x_m) = \ell_i(x_1, \ldots, x_m)$,
- and $\langle a_1, \ldots, a_{k_2}, b_1, \ldots, b_{k_1} \rangle \in \tilde{V}$ if and only if $\langle a_1, \ldots, a_{k_2} \rangle \in V'$ and $\langle b_1, \ldots, b_{k_1} \rangle \in V$.

The claim is immediate: On the one hand, if $f \notin \mathcal{F}$ then there must exist $x_1, \ldots, x_m$ such that $\langle f(y_1) \ldots, f(y_{k_1}) \rangle \notin V$ and thus for every $z_1, \ldots, z_{m'}$ $\langle f(y'_1), \ldots, f(y'_{k_2}), f(y_1), \ldots, f(y_{k_1}) \rangle \notin \tilde{V}$, where $y_i = \ell_i(x_1, \ldots, x_m)$ and $y'_i = \ell'_i(z_1, \ldots, z_{m'})$. On the other hand if $f \in \mathcal{F}$ then f is also in $\mathcal{F}'$ and so for every $x_1, \ldots, x_m, z_1, \ldots, z_{m'}$ we have $\langle f(y'_1), \ldots, f(y'_{k_2}), f(y_1), \ldots, f(y_{k_1}) \rangle \in \tilde{V}$. Finally, it is straightforward to verify that $\tilde{\ell}'_1$ is linearly independent of all the other linear functions: it is independent of $\tilde{\ell}'_i$ by th 2-ary independence of the characterization of $\mathcal{F}'$; and it is independent of $\tilde{\ell}_i$ since it operates on a disjoint set of formal variables. ∎

Putting all the ingredients together we get:

Lemma 5.18 *Let $\mathcal{F} \subseteq \{\mathbb{F}^n \to \mathbb{F}\}$ be a linear-invariant family of linear-invariance degree d. Then it is $k' = 2q \cdot q^{2(d+q)/p}$-locally testable. Specifically, there is k'-query local test that accepts members of $\mathcal{F}$ with probability 1 and rejects functions that are δ-far from $\mathcal{F}$ with probability $\min\left\{\frac{\delta}{2}, \frac{q^2}{(2k'+q)(k'+q)}\right\}$.*

Proof: By Lemma 5.16 there exists a linear-invariant family $\mathcal{F}_1$ of polynomials of degree at most d and a projective family $\mathcal{F}_2$ such that $\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2$.
Since the linear invariance degree of $\mathcal{F}_1$ is at most d, it has, by Lemma 5.4 a $q^{2(d+q)/p}$-local formal characterization. Since $\mathcal{F}_1$ is contained in the family of degree d polynomials it also has a 2-ary independent (in fact, affine) $q^{2(d+q)/p}$-local formal constraint (now using the fact that any constraint on the family of degree d polynomials is a constraint on $\mathcal{F}_1$ and using Lemma 5.4 again to see that the family of degree d polynomials has an affine invariant $q^{2(d+q)/p}$-local constraint). Using Proposition 5.17, we conclude that $\mathcal{F}_1$ has a $2q^{2(d+q)/p}$-local 2-ary independent formal characterization. By Theorem 2.9, we have that $\mathcal{F}_1$ has a $k_1 = 2q^{2(d+q)/p}$-local test that accepts members of $\mathcal{F}_1$ and rejects a member that is δ-far with probability $\min\left\{\frac{\delta}{2}, \frac{1}{(2k_1+1)(k_1+1)}\right\}$.
We now describe the test for membership in $\mathcal{F}$. Given oracle access to a function f, we invoke Lemma 5.16 to get oracle access to the function g such that $\delta(f, \mathcal{F}) \leq \delta(g, \mathcal{F}_1)$ and such that $f \in \mathcal{F}$ implies $g \in \mathcal{F}_1$. We test if $g \in \mathcal{F}_1$ using the test for $\mathcal{F}_1$ from the previous paragraph. This test makes $q \cdot k_1$ queries into the oracle for f (to simulate the k_1 queries to g). If $f \in \mathcal{F}$ then $g \in \mathcal{F}_1$ and this test accepts with probability 1. If f is δ-far from $\mathcal{F}$, then g is also δ-far from $\mathcal{F}_1$ and so the test rejects with probability $\min\left\{\frac{\delta}{2}, \frac{1}{(2k_1+1)(k_1+1)}\right\}$.
The lemma follows using $k' = qk_1$. ∎

5.5 Summarizing: Constraints, Characterizations and Tests

The lemmas proved in the earlier parts of this section combine to prove Theorems 2.10 and Theorems 2.11 for the special case when $\mathbb{K} = \mathbb{F}$. Specifically, we get that affine invariant families have local formal characterizations and local tests if and only if they have a single local constraints. For linear invariant families we get the same conclusion under the stronger hypothesis that they have a local characterization. For the sake of completeness we include a formal statement and proof below.

Theorem 5.19 *If $\mathcal{F} \subseteq \{\mathbb{F}^n \to \mathbb{F}\}$ is an affine-invariant family with a k-local constraint, then it has a $k' = (q^2k)^{q^2}$-local formal affine characterization, where $q = |\mathbb{F}|$. Furthermore $\mathcal{F}$ is k'-locally testable where the test accepts members of $\mathcal{F}$ with probability 1 and rejects functions that are δ-far with probability $\min\left\{\frac{\delta}{2}, \frac{1}{(2k'+1)(k'+1)}\right\}$.*

Proof: By Lemma 5.7 we have that if $\mathcal{F}$ has a k-local constraint then every function of $\mathcal{F}$ has degree $d < q^2 + q^2 \log_q k$. Now, from Lemma 5.5 we have that if every function in $\mathcal{F}$ is a polynomial of degree at most d, then $\mathcal{F}$ has a $k' = q^{2(d+q)/p}$ local formal characterization. Combining the two bounds with some crude manipulations, we get that $k' \leq (q^2k)^{q^2}$. Since every formal characterization of an affine invariant family can be converted into an affine formal characterization, and hence a 2-ary independent formal characterization, with the same locality, we can now apply Theorem 2.9 to conclude that $\mathcal{F}$ is k'-locally testable. ∎

Similarly, by combining Lemmas 5.13, 5.4, and 5.18, we also get an analogous theorem for linear-invariant families where the hypothesis of k-local constraint is replaced by the hypothesis of a k-local characterization, and the parameter of interest in the proof is now the linear-invariance degree of $\mathcal{F}$.

Theorem 5.20 *If $\mathcal{F} \subseteq \{\mathbb{F}^n \to \mathbb{F}\}$ is a linear-invariant family with a k-local characterization then it has a $k' = (q^2k)^{q^2}$-local formal characterization for $q = |\mathbb{F}|$. Furthermore $\mathcal{F}$ is k_0-locally testable for $k_0 = 2qk'$ where the test accepts members of $\mathcal{F}$ with probability 1 and rejects functions that are δ-far with probability $\min\left\{\frac{\delta}{2}, \frac{q^2}{(2k_0+q)(k_0+q)}\right\}$.*

Proof: By Lemma 5.13, we have that the linear-invariance degree of $\mathcal{F}$ is at most $d < q^2 + q^2 \log_q k$. By Lemma 5.4, we have that $\mathcal{F}$ has a $k' = q^{2(d+q)/p}$-local formal characterization. It follows that $k' \leq (q^2k)^{q^2}$. Finally, by Lemma 5.18, we have that $\mathcal{F}$ is $k_0 = 2qk'$-locally testable. ∎

6 Function Families over Extension Fields

In this section we now generalize our study of function families to the case of general fields $\mathbb{K}$ and $\mathbb{F}$ with $\mathbb{K} \supseteq \mathbb{F}$. We extend the results from Sections 4 and 5 to this setting. Throughout the section we let $q = |\mathbb{F}| = p^s$ and $Q = |\mathbb{K}| = q^t$ (though sometimes we will repeat this fact, for redundancy).

We start by describing a basis for functions from $\mathbb{K}^n$ to $\mathbb{F}$ that extends the role played by monomials in the case of functions from $\mathbb{F}^n$ to $\mathbb{F}$. Two well-known functions mapping $\mathbb{K}$ to $\mathbb{F}$ are the "Trace" function, which we will denote $\mathrm{Trace}_0(\cdot)$, and the "Norm" function. The standard Trace function is given by $\mathrm{Trace}_0(x) = x + x^q + \cdots + x^{q^{t-1}}$. The Norm function $N(x)$ is given by $N(x) = x^{1+q+\cdots+q^{t-1}}$. We wish to find a "basis" of all functions that map from $\mathbb{K}^n$ to $\mathbb{F}$, we need a family which generalizes both these families, hopefully in a nice algebraic way. We describe such a generalization below. We refer to the functions we work with as the 'Traces of monomials". (We are not aware of previous use of this family.)

Definition 6.1 *For a vector $\mathbf{d} = \langle d_1, \ldots, d_n \rangle$ of non-negative integers, let $b(\mathbf{d})$ denote the smallest positive integer b such that $d_i \cdot q^b \equiv d_i \mod (Q-1)$ for every $i \in [n]$. Note that $b \leq t$. We say that $c \in \mathbb{K}$ is $\mathbf{d}$-admissible if $c^{q^{b(\mathbf{d})}} = c$. For a vector $\mathbf{d}$ and $\mathbf{d}$-admissible coefficient $c \in K$, the Trace of the monomial $m = c \cdot \mathbf{x}^{\mathbf{d}}$, denoted $\mathrm{Trace}(m)$, is the polynomial $m + m^q + \cdots + m^{q^b-1}$ for $b = b(\mathbf{d})$.*

In what follows it is critical that we do not confuse the monomial Trace function $\mathrm{Trace}(m) : \mathbb{K}^n \to \mathbb{F}$ from the function $\mathrm{Trace}_0 \circ m : \mathbb{K}^n \to \mathbb{F}$. Whereas the latter is more commonly studied, it is the former that is central to this section. For example, over $\mathbb{K} = \mathbb{F}_{16}$ and $\mathbb{F} = \mathbb{F}_2$, $\mathrm{Trace}(x^3) = x^3 + x^6 + x^{12} + x^9$, $\mathrm{Trace}(y^5) = y^5 + y^{10}$, and $\mathrm{Trace}(x^3y^5) = x^3y^5 + x^6y^{10} + x^{12}y^5 + x^9y^{10}$.

In the definition above, we were careful with the coefficients of the monomials in the argument of the Trace function. This is important since the function $\mathrm{Trace}(\alpha x^d)$ could be linearly independent (over $\mathbb{F}$) of the function $\mathrm{Trace}(\beta x^d)$. However, for admissible coefficients, $\mathrm{Trace}(\alpha x^d)$ and $\mathrm{Trace}(\beta x^d)$ generate the same linear span, as we show below. (This proposition simplifies our life later, by letting us ignore the coefficients of the monomials in our basis functions.)

Proposition 6.2 *For a vector $\mathbf{d} = \langle d_1, \ldots, d_n \rangle$ of non-negative integers and $\mathbf{d}$-admissible coefficients $\alpha, \beta \in \mathbb{K}^*$, it is the case that $\mathrm{Trace}(\alpha \mathbf{x}^{\mathbf{d}}) \in \text{L-SPAN}(\mathrm{Trace}(\beta \mathbf{x}^{\mathbf{d}}))$.*

Proof: Let $b = b(\mathbf{d})$. Note that admissibility of α, β implies that they are contained in the field $\mathbb{L} = \mathbb{F}_{q^b}$ (since $\alpha^{q^b} = \alpha$ and $\beta^{q^b} = \beta$). Let S denote the set of coefficients $S = \{\gamma | \ \mathrm{Trace}(\gamma \mathbf{x}^{\mathbf{d}}) \in \text{L-SPAN}(\mathrm{Trace}(\beta \mathbf{x}^{\mathbf{d}}))$. We will prove the proposition by proving $S = \mathbb{L}$.

First note that since the monomial-Trace function is additive, and the linear span of any set is closed under addition, we get that S is closed under addition.

Now we turn to its multiplicative properties. To do so, we need to understand $b = b\mathbf{d}$. Notice that $q^t - 1$ divides $d_i \cdot (q^b - 1)$ for every i. So if we let $e = \gcd(d_1, \ldots, d_n)$, we have that $q^t - 1$ divides $e \cdot (q^b - 1)$ and furthermore b is the smallest positive integer that has this property. Let ω be a primitive $(q^t - 1)$th root of unity in $\mathbb{K}$. Then, by the fact that b is the smallest integer such that $q^t - 1$ divides $e \cdot (q^b - 1)$, we have that $\mathbb{L}$ is the smallest subfield of $\mathbb{K}$ that contains ω^e. We claim that if $\tau \in S$, then $\tau \cdot \omega^e \in S$. To see this, let $a_1, \ldots, a_n$ be integers such that $e = \sum_{i=1}^{n} a_i d_i$. Then note that $\mathrm{Trace}(\tau(\omega^{a_1} x_1)^{d_1} \cdots (\omega^{a_n} x_n)^{d_n}) \in$ L-SPAN$(\mathrm{Trace}(\tau \mathbf{x}^{\mathbf{d}}))$ and so we have $\tau\omega^e \in S$. We thus conclude that S is closed under addition, and under multiplication by ω^e. It follows that $S = \mathbb{L}$. ∎

Thus from now on, whenever we refer to monomials, we may ignore the leading coefficient, since any admissible coefficient is equivalent to the coefficient 1. The central nature of the trace of monomials is explained by the following proposition.

Proposition 6.3 *Every function $f : \mathbb{K}^n \to \mathbb{F}$ can be described by a set of monomials $\mathcal{M}$ such that $f(\mathbf{x}) = \sum_{m \in \mathcal{M}} \mathrm{Trace}(m)$.*

Proof: Let $f(\mathbf{x}) = \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{x}^{\mathbf{d}}$. We prove the lemma by induction on the size of the support of f. Let $\mathbf{e}$ be a vector such that $c_{\mathbf{e}} \neq 0$. Then we note that $c_{(q \cdot \mathbf{e}) \mod (Q-1)} = c_{\mathbf{e}}^q$. This is so since $f(\mathbf{x})^q = f(\mathbf{x})$ (since $f(\mathbf{x}) \in \mathbb{F}$. Furthermore, $f(\mathbf{x})^q = (\sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{x}^{\mathbf{d}})^q = \sum_{\mathbf{d}} c_{\mathbf{d}}^q \mathbf{x}^{q\mathbf{d} \mod (Q-1)}$. By considering the coefficient of $c_{q \cdot \mathbf{e} \mod (Q-1)}$ we get $c_{(q \cdot \mathbf{e}) \mod (Q-1)} = c_{\mathbf{e}}^q$. Note further that since $q^{b(\mathbf{e})}\mathbf{e} \equiv \mathbf{e} \mod (Q-1)$ it follows that $c_{\mathbf{e}}^{q^{b(\mathbf{e})}} = c_{\mathbf{e}}$ and so $c_{\mathbf{e}}$ is $\mathbf{e}$-admissible. It follows that if we subtract $\mathrm{Trace}(c_{\mathbf{e}} \mathbf{x}^{\mathbf{e}})$ from $f(\mathbf{x})$ we get a function on a smaller support. We conclude that f can be decomposed into a sum of traces of monomials. ∎

In what follows, we start by giving an extraction lemma for linear-invariant families of function mapping $\mathbb{K}^n$ to $\mathbb{F}$, which shows that the trace of any monomial that is in the support of a function in the family is also in the family. We then use this, along with standard monomial "spread" properties to give upper bounds (see Section 6.2) and lower bounds (Sections 6.3 and 6.4) on the constraints and characterizations of affine-invariant and linear-invariant families mapping $\mathbb{K}^n$ to $\mathbb{F}$. In Section 6.5 we use the characterizations to build a tester for the linear-invariant case. The resulting theorems are summarized in Section 6.6.

6.1 Extracting Traces of Monomials

For a set of functions $S \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$, recall the notions of SPAN(S) = SPAN$_{\mathbb{F}}(S)$ and L-SPAN(S) and A-SPAN(S) (see Definition 4.1 in Section 4 for the formal definitions). These notions will be used in this and subsequent sections.

Lemma 6.4 (Trace of Monomial Extraction Lemma) *Let $f : \mathbb{K}^n \to \mathbb{F}$. Then for every monomial m in the support of f, we have* $\mathrm{Trace}(m) \in$ L-SPAN(f).

Proof: Let m be a monomial in the support of f. Let $m = c\mathbf{x}^{\mathbf{e}}$ where $\mathbf{e} = \langle e_1, \cdots, e_n \rangle$ and $c \in \mathbb{K}$ is $\mathbf{e}$-admissible. Let $b = b(\mathbf{e})$, so that $\mathrm{Trace}(m) = m + m^q + \cdots + m^{q^{b-1}}$. We wish to show $\mathrm{Trace}(m) \in$ L-SPAN(f).

We first claim that we can assume w.l.o.g. that $e_i \neq 0$ for every $i \in [n]$. If not, and suppose $e_n = 0$, then we are done by induction on the number of variables, since $g(x_1, \ldots, x_{n-1}) \stackrel{\triangle}{=} f(x_1, \ldots, x_{n-1}, 0)$ has the feature that m is in the support of g and g depends on fewer variables. So $\mathrm{Trace}(m) \in$ L-SPAN(g) and $g \in$ L-SPAN(f) yielding $\mathrm{Trace}(m) \in$ L-SPAN(f).

Next we claim that we can assume that for every $i \in [n]$ and for every monomial $c_{\mathbf{d}}\mathbf{x}^{\mathbf{d}}$ in the support of f, it is the case that $d_i \neq 0$. We prove this by induction on i. Assume the statement is true for $j \in [i]$ (i.e., $d_j \neq 0$ for every $j \in [i]$ and every monomial $\mathbf{x}^{\mathbf{d}}$ with non-zero coefficient in f). Now consider the function $\tilde{f}(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) - f(x_1, \ldots, x_i, 0, x_{i+2}, \ldots, x_n)$. $\tilde{f}$ now has no support on monomials of the form $\mathbf{x}^{\mathbf{d}}$ with $d_j = 0$ for any $j \in [i+1]$. But m is still in the support of $\tilde{f}$ and $\tilde{f} \in$ L-SPAN(f). So proving $\mathrm{Trace}(m) \in$ L-SPAN$(\tilde{f})$ suffices to prove $\mathrm{Trace}(m) \in$ L-SPAN(f).

Finally we get to the real case: We now have a monomial $m = c\mathbf{x}^{\mathbf{e}}$ in the support of f. For every monomial $c_{\mathbf{d}}\mathbf{x}^{\mathbf{d}}$ in the support of f and every $i \in [n]$ we have $d_i \neq 0$. We'd like to show $\mathrm{Trace}(m) \in$ L-SPAN(f).

Let $f = \sum_{\mathbf{d}} c_{\mathbf{d}}\mathbf{x}^{\mathbf{d}}$. Let $\mathbb{K}^* = \mathbb{K} - \{0\}$. Consider the following expression.

$$g(\mathbf{x}) = \sum_{s=0}^{b-1} \sum_{\langle \alpha_1, \ldots, \alpha_n \rangle \in (\mathbb{K}^*)^n} (\alpha_1)^{-e_1 \cdot q^s} \cdots (\alpha_n)^{-e_n \cdot q^s} f(\alpha_1^{-1}x_1, , \cdots, \alpha_n^{-1}x_n).$$

We claim that $g(\mathbf{x}) \in$ L-SPAN(f) and that $g(\mathbf{x}) = (-1)^n \cdot \mathrm{Trace}(c_{\mathbf{e}}\mathbf{x}^{\mathbf{e}})$ thereby showing that $\mathrm{Trace}(m) \in$ L-SPAN(f).

For the first part, it is obvious that $g(\mathbf{x})$ is in L-SPAN$_{\mathbb{K}}(f)$, but this is not what we want. We need to show that $g(\mathbf{x}) \in$ L-SPAN$_{\mathbb{F}}(f)$. To see this we use the property of the monomial $m_0 = \mathbf{x}^{\mathbf{e}}$. Note that $\mathrm{Trace}(m_0) = m_0 + m_0^q + \cdots + m_0^{q^{b-1}}$ since $b = b(\mathbf{e})$ is independent of c. Note that

$$g(\mathbf{x}) = \sum_{\alpha \in (\mathbb{K}^*)^n} \mathrm{Trace}(m_0(\alpha)) \cdot f(\alpha_1^{-1}x_1, \alpha_2^{-1}x_2, \cdots, \alpha_n^{-1}x_n).$$

Since $\mathrm{Trace}(m_0)$ maps $\mathbb{K}^n$ to $\mathbb{F}$, we have that the expression for g forms an $\mathbb{F}$-linear combination of f applied to $\mathbb{K}$-linear transforms of the vector $\mathbf{x}$. By definition of L-SPAN we have $g \in$ L-SPAN$_{\mathbb{F}}(f)$.

Next to see that $g(\mathbf{x}) = (-1)^n \mathrm{Trace}(c_{\mathbf{e}}\mathbf{x}^{\mathbf{e}})$, we write $g(\mathbf{x}) = \sum_{s=0}^{b-1} g_s(\mathbf{x})$, where $g_s(\mathbf{x}) = \sum_{\alpha \in (\mathbb{K}^*)^n} m_0(\alpha)^{q^s} \cdot f(\alpha_1^{-1}x_1, \cdots, \alpha_n^{-1}x_n)$. We claim that $g_s(\mathbf{x}) = (-1)^n c_{q^s \cdot \mathbf{e}} m_0(\mathbf{x})^{q^s} = c_{\mathbf{e}}^{q^s} m_0(\mathbf{x})^{q^s}$ and this implies $g(\mathbf{x}) = \mathrm{Trace}(m)$. But then the identity $g_s(\mathbf{x}) = (-1)^n c_{q^s \cdot \mathbf{e}} m_0(\mathbf{x})^{q^s}$, follows easily from the Fourier Transform. Specifically:

$$\begin{aligned}
g_s(\mathbf{x}) &= \sum_{\alpha \in (\mathbb{K}^*)^n} \alpha^{q^s \cdot \mathbf{e}} \cdot f(\alpha_1^{-1}x_1, \cdots, \alpha_n^{-1}x_n) \\
&= \sum_{\alpha \in (\mathbb{K}^*)^n} \alpha^{q^s \cdot \mathbf{e}} \cdot \sum_{\mathbf{d}} c_{\mathbf{d}} \alpha^{-\mathbf{d}} \mathbf{x}^{\mathbf{d}} \\
&= \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{x}^{\mathbf{d}} \sum_{\alpha \in (\mathbb{K}^*)^n} \alpha^{q^s \cdot \mathbf{e} - \mathbf{d}} \\
&= \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{x}^{\mathbf{d}} \prod_{i=1}^{n} \left(\sum_{\alpha_i \in \mathbb{K}^*} \alpha_i^{(q^s \cdot e_i - d_i)} \right).
\end{aligned}$$

Now the summation $\sum_{\alpha_i \in \mathbb{K}^*} \alpha_i^{(q^s \cdot e_i - d_i)}$ equals -1 if $q^s e_i = d_i$ and 0 otherwise. So the final quantity above equals $(-1)^n c_{q^s \mathbf{e}} \mathbf{x}^{q^s \mathbf{e}}$ as desired. ∎

6.2 Upper bounds on the characterizations of linear-invariant families

In this section we give characterizations, and thus upper bounds on the locality of characterizations, of affine-invariant and linear-invariant families mapping $\mathbb{K}^n$ to $\mathbb{F}$. As in the case when $\mathbb{K} = \mathbb{F}$, the affine-invariance locality is a function of the degree of the highest degree polynomial contained in the family under consideration. In the linear-invariant case we need to extend the notion of the linear-invariance degree and we do so below.

Definition 6.5 *For a linear invariant family $\mathcal{F}$ properly contained in $\{\mathbb{K}^n \to \mathbb{F}\}$, let $d_{\text{lin}}(\mathcal{F})$, the* linear-invariance degree *of $\mathcal{F}$, be the largest integer d such that $\mathcal{F}$ contains a monomial m_1 of degree d in its support, while there also exists a monomial m_2 that is not in the support of $\mathcal{F}$, whose degree is d' for some $d' > 0$ with $d' \equiv d \mod (Q-1)$, where $Q = |\mathbb{K}|$.*

To get an upper bound, we first give a simple monomial spread lemma for functions from $\mathbb{K}^n$ to $\mathbb{F}$.

Lemma 6.6 *Let $m = \mathbf{x}^{\mathbf{d}}$ be a monomial with $d_{n-1} < p$ and $d_n = 0$. For $i \in \{0, \ldots, d_{n-1}\}$, let $\tilde{m}$ be the monomial $x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^{i} \cdot x_n^{d_{n-1}-i}$. Then the function* $\text{Trace}(\tilde{m}) \in \text{L-SPAN}(\text{Trace}(m))$.

Proof: Let $f(\mathbf{x}) = \text{Trace}(m(\mathbf{x}))$. Note that since $d_{n-1} < p \leq q$, we have that $q^{t-1}d_{n-1} < q^t$ and so $b(\mathbf{d})$ must equal t. Thus $\text{Trace}(m(\mathbf{x})) = \text{Trace}_0(m(\mathbf{x}))$. So we need to show $\text{Trace}(\tilde{m}) \in \text{L-SPAN}(\text{Trace}_0(m))$. Now consider $\tilde{f}(\mathbf{x}) = f(x_1, \ldots, x_{n-2}, x_{n-1}+x_n, 0) \in \text{L-SPAN}(f)$. We have $\tilde{f}(\mathbf{x}) = \sum_{\ell=0}^{t-1} \sum_{i=0}^{d_{n-1}} \binom{d_{n-1}}{i} (x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^{i} \cdot x_n^{d_{n-1}-i})^{q^\ell}$. We note that the coefficient of $\tilde{m}$ in this expression is exactly $\binom{d_{n-1}}{i}$ which is non-zero. It follows from the monomial extraction lemma that $\tilde{m}$ is in $\text{L-SPAN}(\tilde{f}) \subseteq \text{L-SPAN}(f)$. ∎

The following lemma now shows that one can project non-members of a family $\mathcal{F}$ to smaller dimensional subspaces while preserving non-membership in $\mathcal{F}$.

Lemma 6.7 *Let $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$ be a linear-invariant family of linear-invariance degree $d_{\text{lin}}(\mathcal{F}) = d$. Suppose $f : \mathbb{K}^n \to \mathbb{F}$ is not in $\mathcal{F}$. Then, if $n \geq 1 + \left(\frac{2(d+Q)}{p}\right)$, then there exists a linear function $L : \mathbb{K}^{n-1} \to \mathbb{K}^n$ such that $f \circ L \notin \mathcal{F}|_{n-1}$.*

Proof: The proof is exactly the same as that of Lemma 5.3 with notational changes. We include it below for completeness.

Let $m = x_1^{d_1} \cdots x_n^{d_n}$ be a monomial of maximal degree in the support of f that is not contained in the support of $\mathcal{F}$. We show that there is a linear map $L : \mathbb{K}^{n-1} \to \mathbb{K}^n$ such that $m \circ L$ is not in the support of $\mathcal{F}|_{n-1}$. We consider two cases:

Case 1: *There exist distinct indices i, j such that $d_i + d_j < p$*: Without loss of generality assume $i = n-1$ and $j = n$. Note first that for the the monomial $m' = x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^{d_{n-1}+d_n}$, we have, by Lemma 6.6, $\text{Trace}(m') \notin \mathcal{F}$ (and hence $\text{Trace}(m') \notin \mathcal{F}|_{n-1}$). We claim that for some choice of $\alpha, \beta \in \mathbb{F}$, the map $L_{\alpha,\beta}(x_1, \ldots, x_{n-1}) = \langle x_1, \ldots, x_{n-2}, \alpha x_{n-1}, \beta x_{n-1} \rangle$ leaves the monomial m' with non-zero support in $f \circ L_{\alpha,\beta}$, which would suffice to prove the lemma (in this case).

To see this, let c_i be the coefficient of the monomial $x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^{i} x_n^{d_{n-1}+d_n-i}$ in f. Let $h(x, y) = \sum_{i=0}^{d_{n-1}+d_n} c_i x^i y^{d_{n-1}+d_n-i}$. It can be verified that the coefficient of m' in $f \circ L_{\alpha,\beta}$ is exactly $h(\alpha, \beta)$. Furthermore, $h(\alpha, \beta)$ is a non-zero polynomial since the coefficient $c_{d_{n-1}}$ is the coefficient of m in f which is non-zero. Thus there must exist $\alpha, \beta \in \mathbb{K}$ such that $h(\alpha, \beta) \neq 0$ and this yields the claim.

Case 2: *For every pair of distinct i, j, $d_i + d_j \geq p$.* Let e denote the degree of m. For every t, we have $d_{2t-1} + d_{2t} \geq p$, and so e, the total degree of m, is at least $p\lfloor n/2 \rfloor \geq d + Q$.

We first note that no monomial m' of degree e or $e-(Q-1)$ is in $\mathcal{F}$. Otherwise the linear-invariance degree of $\mathcal{F}$ would be the degree of m'. For example, if m' has degree $e-(Q-1) > d$, then m' satisfies the role of the monomial m_1 in the definition of the linear-invariance degree and m of degree $e = e-(Q-1)(\mod Q-1)$ satisfies the role of m_2 in the definition of linear-invariance degree thereby yielding $d_{\text{lin}}(\mathcal{F}) = e-(Q-1) > d$. So we conclude m' can not be in $\mathcal{F}$.

But now consider the linear map $L_{\alpha,\beta}$ as in the previous case, i.e., $L_{\alpha,\beta}(x_1, \ldots, x_{n-1}) = \langle x_1, \ldots, x_{n-2}, \alpha x_{n-1}, \beta x_{n-1} \rangle$. Let m' be the monomial $x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^{f}$ where $f = d_{n-1} + d_n$ if $d_{n-1} + d_n < Q$ and $f = d_{n-1} + d_n - (Q-1)$ otherwise. The degree of m' is thus e or $e-(Q-1)$. We claim that for some α, β, the coefficient of m' is non-zero in $f \circ L_{\alpha,\beta}$ and this will yield the lemma in this case.

To verify the claim, note that the coefficient of m' in $f \circ L_{\alpha,\beta}$ is exactly $h'(\alpha, \beta)$ where $h'(x, y) = \sum_{i=0}^{\min\{Q-1, d_n+d_{n-1}\}} c_i x^i y^{d_n+d_{n-1}-i}$ and c_i is the coefficient of the monomial $x_1^{d_1} \cdots x_{n-2}^{d_{n-2}} \cdot x_{n-1}^{i} x_n^{d_n+d_{n-1}-i}$ in f. Again, we have $h'(x, y)$ is not identically zero (the coefficient $c_{d_{n-1}}$ is non-zero) and so there exists α, β such that $h'(\alpha, \beta) \neq 0$. ∎

We are now ready to give the characterization for linear-invariant families.

Lemma 6.8 *Let $\mathcal{F}$ be a linear invariant family, properly contained in $\{\mathbb{K}^n \to \mathbb{F}\}$, of linear-invariance degree d. Then $\mathcal{F}$ has a $(Q)^\ell$-local formal characterization for $\ell = \frac{2(d+Q)}{p}$.*

Proof: We claim that the characterization is simply the that $f \in \mathcal{F}$ if and only if $f \circ L$ is in $\mathcal{F}|_\ell$ for every linear map $L : \mathbb{K}^\ell \to \mathbb{K}^n$.

It is again easy to verify that, if correct, this is indeed a q^ℓ-local *formal* characterization. We analyze the correctness below.

In one direction, it is obvious that every $f \in \mathcal{F}$ and $L : \mathbb{K}^\ell \to \mathbb{K}^n$ satisfies $f \circ L \in \mathcal{F}|_\ell$. The other direction is a simple induction based on Lemma 5.3. Fix $f \notin \mathcal{F}$. Let m be any integer between $\ell + 1$ and n. Assume by induction on $n - m$ that there a linear map $L_m : \mathbb{K}^m \to \mathbb{K}^n$ such that $f \circ L_m \notin \mathcal{F}|_m$. Now we prove that there is a map $L_{m-1} : \mathbb{K}^{m-1} \to \mathbb{K}^n$ such that $f \circ L_{m-1} \notin \mathcal{F}|_{m-1}$. Since $\mathcal{F}$ is linear-invariant, so is $\mathcal{F}|_m$. Also the linear-invariance degree of $\mathcal{F}|_m$ is at most $d_{\max}$. By Lemma 5.3 there is a linear map $L : \mathbb{K}^{m-1} \to \mathbb{K}^m$ such that $(f \circ L_m) \circ L \notin (\mathcal{F}|_m)|_{m-1} = \mathcal{F}|_{m-1}$. Thus $f \circ L_{m-1} \notin \mathcal{F}|_{m-1}$ for $L_{m-1} = L_m \circ L$. We conclude that the linear map $L_\ell : \mathbb{K}^\ell \to \mathbb{K}^n$ derived from setting $m = \ell + 1$, satisfies $f \circ L_\ell \notin \mathcal{F}|_\ell$. ∎

Again, using the fact that the maximum total degree of a polynomial in the family $\mathcal{F}$ is an upper bound on the linear-invariance degree of $\mathcal{F}$, we also get the following corollary for affine-invariant families.

Lemma 6.9 *Let $\mathcal{F}$ be a proper subset of $\{\mathbb{K}^n \to \mathbb{F}\}$ and let d denote the maximum degree of any function in $\mathcal{F}$. Then $\mathcal{F}$ has a Q^ℓ-local formal characterization for $\ell \leq \frac{2(d+Q)}{p}$.*

6.3 Lower bounds on the locality of constraints for Affine Invariant Families

We now move to lower bounds on the locality of constraints for affine-invariant families. Our starting point is Claim 5.9 which shows that the family A-SPAN$_{\mathbb{K}}(\prod_{i=1}^{\ell} x_i^{Q/p} \subseteq \{\mathbb{K}^n \to \mathbb{K}\}$ has no $Q^{\lfloor \ell/Q \rfloor}$-local

constraints. We start with a simple proposition that turns this into a lower bound on a family mapping $\mathbb{K}^n$ to $\mathbb{F}$. First we recall some definitions.

Recall that $\mathrm{Trace}_0 : \mathbb{K} \to \mathbb{F}$ denotes the standard trace function given by $\mathrm{Trace}_0(x) = \sum i = 0^{t-1} x^{q^i}$. We extend this to functions and function families as follows. For $f : \mathbb{K}^n \to \mathbb{K}$, $\mathrm{Trace}_0(f)$ is the function that maps $\mathbf{x} \in \mathbb{K}^n$ to $\mathrm{Trace}_0(f(\mathbf{x}))$. For $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{K}\}$, let $\mathrm{Trace}_0(\mathcal{F}) = \{\mathrm{Trace}_0(f) | f \in \mathcal{F}\}$.

We also extend the notion of the trace of monomials to functions and function families. We say that a monomial $c\mathbf{x}^{\mathbf{e}}$ is admissible if c is $\mathbf{e}$-admissible. For a function $f : \mathbb{K}^n \to \mathbb{K}$ which is the sum of admissible monomials, we define $\mathrm{Trace}(f)$ to be the sum of the traces of the monomials in its support. Finally, for a function family $\mathcal{F}$, we let $\mathrm{Trace}(\mathcal{F}) = \{\mathrm{Trace}(f) | f \in \mathcal{F}\}$.

The following proposition relates $\mathrm{Trace}_0(\mathcal{F})$ to $\mathrm{Trace}(\mathcal{F})$.

Proposition 6.10 *For a linear-invariant family $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{K}\}$, $\mathrm{Trace}_0(\mathcal{F}) \subseteq \mathrm{Trace}(\mathcal{F})$.*

Proof: Note that it suffices to show that $\mathrm{Trace}_0(m) \in \mathrm{Trace}(\mathcal{F})$ for every monomial $m \in \mathcal{F}$. Let $m = c\mathbf{x}^{\mathbf{d}}$ and let $b = b(\mathbf{d})$. Note that $\mathrm{Trace}_0(m) = \mathrm{Trace}(m')$ where $m' = c'\mathbf{x}^{\mathbf{d}}$ and $c' = c + c^{q^b} + c^{q^{2b}} + \cdots + c^{q^{t-b}}$. In particular note that $(c')^{q^b} = c'$ and so c' is $\mathbf{d}$-admissible. Also note that m' is an admissible monomial and a member of $\mathcal{F}$ and so $\mathrm{Trace}(m) \in \mathrm{Trace}(\mathcal{F})$. We thus conclude that $\mathrm{Trace}_0(m) \in \mathrm{Trace}(\mathcal{F})$. The proposition follows. ∎

The next proposition shows that a lower bound on the locality of (relative) constraints for a family $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{K}\}$ also yields a lower bound for $\mathrm{Trace}_0(\mathcal{F})$.

Proposition 6.11 *Let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \{\mathbb{K}^n \to \mathbb{K}\}$. If $\mathcal{F}_1$ has no k-local constraints relative to $\mathcal{F}_2$, then $\mathrm{Trace}_0(\mathcal{F}_1)$ has no k-local constraints relative to $\mathrm{Trace}_0(\mathcal{F}_2)$.*

Proof: Let $C = (\mathbf{x}_1, \ldots, \mathbf{x}_k, S)$ with $S \subsetneq \mathbb{F}^k$ be a constraint on $\mathrm{Trace}_0(\mathcal{F}_1)$ relative to $\mathrm{Trace}_0(\mathcal{F}_2)$. Then we claim that $C' = (\mathbf{x}_1, \ldots, \mathbf{x}_k, S')$, where $S' = \{\langle \alpha_1, \ldots, \alpha_k \rangle \in \mathbb{K}^k \mid \langle \mathrm{Trace}_0(\alpha_1), \ldots, \mathrm{Trace}_0(\alpha_k) \rangle \in S\}$, is a constraint on $\mathcal{F}_1$ relative to $\mathcal{F}_2$. We omit the straightforward verification steps. ∎

The above propositions immediately give a family of affine invariant functions with no constraints of small locality.

Lemma 6.12 *For every ℓ, the family* $\text{A-SPAN}(\mathrm{Trace}(\prod_{i=1}^{\ell} x_i^{Q/p}))$ *has no constraints of locality $Q^{\lfloor \ell/Q \rfloor}$.*

Proof: Follows immediately by combining Claim 5.9 with Propositions 6.10 and 6.11. ∎

We now turn to the task of showing that a family with some high degree monomial also contains other high degree monomials. We don't provide a very general lemma, but rather one that is sufficient for our purposes.

Lemma 6.13 *For every vector $\mathbf{d} = \langle d_1, \ldots, d_n \rangle$ of non-negative integers and index $i \in [n]$, the monomial* $\mathrm{Trace}(\mathbf{x}^{\mathbf{e}}) \in \text{A-SPAN}(\mathrm{Trace}(\mathbf{x}^{\mathbf{d}}))$, *where $\mathbf{e} = \langle e_1, \ldots, e_n \rangle$ is given by $e_i = Q/p$, and $e_j = d_j$ for $j \in [n] - \{i\}$, provided $d_i \geq Q/p$.*

Proof: For notational simplicity we assume $i = 1$.

Let $b = b(\mathbf{d})$. Note that since $e_1 = Q/p$, we have that the smallest integer b' such that $e_1^{b'} \equiv e_1 \mod (Q-1)$ is t and thus $\mathrm{Trace}(\mathbf{x}^{\mathbf{e}}) = \mathrm{Trace}_0(\mathbf{x}^{\mathbf{e}})$. Our goal is thus to show that $\mathrm{Trace}_0(\mathbf{x}^{\mathbf{e}})$ is in the affine span of $\mathrm{Trace}(\mathbf{x}^{\mathbf{d}})$. If $d_1 = Q/p$, then this is trivial, and so assume $d_1 > Q/p$.

We claim that, for some $\gamma \in \mathbb{K}$, the monomial $\mathbf{x}^{\mathbf{e}}$ has a non-zero coefficient in the polynomial $g(\mathbf{x}) = \mathrm{Trace}(\langle x_1 + \gamma, x_2, \ldots, x_n \rangle^{\mathbf{d}})$. Assuming this claim, using the fact that $g(\mathbf{x}) \in \text{A-SPAN}(\mathrm{Trace}(\mathbf{x}^{\mathbf{d}}))$ and that every monomial m in the support of $g(\mathbf{x})$ is in A-SPAN(g), we get that for some admissible $\beta \in \mathbb{K}^*$, $\mathrm{Trace}(\beta \cdot \mathbf{x}^{\mathbf{e}})$ is in A-SPAN$(\mathrm{Trace}(\mathbf{x}^{\mathbf{d}}))$. Combining with Proposition 6.2 we see that we can drop the coefficient β and get that $\mathrm{Trace}(\mathbf{x}^{\mathbf{e}})$ is also in A-SPAN$(\mathrm{Trace}(\mathbf{x}^{\mathbf{d}}))$. It thus suffices to prove the claim.
Note that

$$\begin{aligned} g(\mathbf{x}) &= \mathrm{Trace}\left(\left(\sum_{i=0}^{d_1} \binom{d_1}{i} \gamma^{d_1 - 1} x_1^i \right) \cdot \prod_{j=2}^{n} x_j^{d_j} \right) \\ &= \sum_{\ell=0}^{b-1} \left(\sum_{i=0}^{d_1} \binom{d_1}{i}^{q^\ell} \gamma^{q^\ell \cdot (d_1 - i)} x_1^{i \cdot q^\ell} \right) \cdot \prod_{j=2}^{n} x_j^{d_j \cdot q^\ell} \end{aligned}$$

To determine the coefficient of $\mathbf{x}^{\mathbf{e}}$ in the above expression, let $S = \{(i,\ell) | 0 \le i < b, i \cdot q^\ell = Q/p \mod (Q-1)$ and $d_j \cdot q^\ell = d_j \mod (Q-1), \forall j \in \{2, \ldots, n\}\}$. Then the coefficient of $\mathbf{x}^{\mathbf{e}}$ in $g(\mathbf{x})$ is $\sum_{(i,\ell) \in S} \binom{d_1}{i} \cdot \gamma^{q^\ell \cdot (d_1 - i)} = \gamma^{-Q/p} \cdot \sum_{(i,\ell) \in S} \binom{d_1}{i} \cdot \gamma^{q^\ell \cdot d_1}$. This coefficient is itself a polynomial in γ and we prove that it is a non-zero polynomial. To see this we focus on the coefficient of γ^{d_1}. Note that, by the definition of the index b (and $\mathrm{Trace}(\mathbf{x}^{\mathbf{d}})$), the only index ℓ for which $q^\ell \cdot d_1 = d_1 (\mod Q-1)$ (and $q^\ell \cdot d_j = d_j (\mod Q-1)$ for all other j's) is $\ell = 0$. Furthermore, the only i for which $i \cdot q^0 = Q/p$ is Q/p. Thus the pair $(Q/p, 0)$ is the unique pair in S that contributes to the coefficient of γ_1^d in the expression above and this coefficient is $\binom{d_1}{Q/p}$ which can be verified to be non-zero. Thus the coefficient of $\mathbf{x}^{\mathbf{e}}$ is a non-zero polynomial in γ and thus there exists a γ for which this coefficient is non-zero. This proves the claim, and hence the lemma. ∎

The two lemmas above can be combined to derive a lower bound on the locality of constraints for any affine-invariant family containing any high-degree polynomial, as shown next.

Lemma 6.14 *If an affine-invariant family $\mathcal{F} \subsetneq \{\mathbb{K}^n \to \mathbb{F}\}$ contains a polynomial of degree d, then it has no constraints of locality $Q^{(d-Q^2)/Q^2}$.*

Proof: Fix a monomial m such that $\mathrm{Trace}(m)$ is in $\mathcal{F}$ and the degree of m is d. Partition the variables in $\mathbf{x}$ so that the degree of m in each block,except at most one, is between Q/p and $Q-1$ (again this can be done by putting variables of degree greater than Q/p into blocks of their own, and greedily packing the remaining variables into blocks till a block size exceeds Q/p). The number of blocks is thus strictly greater than d/Q. Now replace all variables in block i by the variable x_i to get a new monomial $m' = \mathbf{x}^{\mathbf{d}}$ such that $\mathrm{Trace}(m') \in \mathcal{F}$ and the degree of at least Q/p variables in m' is at least Q/p. Applying Lemma 6.13 to these variables in turn shows that $\mathrm{Trace}(\prod_{i=1}^{\ell} x_i^{Q/p})$ is contained in $\mathcal{F}$ for $\ell = \lfloor d/Q \rfloor$. Applying Lemma 6.12 we conclude that $\mathcal{F}$ has no $Q^{(d-Q^2)/Q^2}$-local constraints. ∎

6.4 Lower Bound in the Linear Invariant Case

We now give a lower bound for the the case of linear-invariant families. We do so by reducing to the lower bound for functions from $\mathbb{K}^n \to \mathbb{K}$.

Lemma 6.15 *Let $\mathbf{d} \in (\mathbb{Z}^+)^n$, $i \in [n]$, and let f be a non-negative integer, such that $d_i \ge Q/p$. Let $\mathbf{e}$ be given by $e_j = d_j$, except when $j = i$ in which case $e_i = Q/p$. Then $\mathbf{x}^{\mathbf{e}} \cdot y^{f + d_i - Q/p} \in \text{L-SPAN}(\mathbf{x}^{\mathbf{d}} \cdot y^f)$.*

Proof: Similar to the proof of Lemma 6.13. ∎

Lemma 6.16 . *Let $\mathcal{F} \subsetneq \{\mathbb{K}^n \to \mathbb{F}\}$ be a linear invariant family of linear invariance degree d. Then $\mathcal{F}$ does not have a $Q^{(d-Q^2)/Q^2}$-local characterization.*

Proof: We proceed as in the proof of Lemma 5.13. Let $m \in \mathbb{K}[\mathbf{x}]$ be a monomial of degree d such that $\mathrm{Trace}(m) \in \mathcal{F}$. Let m' be a monomial of degree $d' \equiv d \mod (Q-1)$ such that $\mathrm{Trace}(m') \notin \mathcal{F}$. Let x_1 be a variable of positive degree in m'. Then note that the monomial $\tilde{m} = m'y/x_1 \in \mathbb{K}[\mathbf{x}, y]$ also satisfies $\mathrm{Trace}(\tilde{m}) \notin \mathcal{F}$ (since $\tilde{m}$ is in the linear span of $\tilde{m}$). Since the degree in y of $\tilde{m}$ is 1, we have that $\mathrm{Trace}(\tilde{m}) = \mathrm{Trace}_0(\tilde{m})$. We now use m and $\tilde{m}$ to get a lower bound on the characterization of $\mathcal{F}$.

Let $\ell = \lfloor d/Q \rfloor$. Let $\mathcal{F}_1 = \textsc{Span}(\{y^{d_1} \cdot m_1 | m_1 \text{ monomial in } \in \text{A-}\textsc{span}(\prod_{i=1}^{\ell} x_i^{q/p}), d_1 + \deg(m_1) \equiv d \mod (q-1), 1 \le d_1 \le q-1\})$. Let $\mathcal{F}_2 = \textsc{Span}\{m_2 | m_2 \text{ monomial in } \mathbb{F}[\mathbf{x}, y], \deg(m_2) \equiv d \mod (q-1), \deg_y(m_2) \ge 1\}$, be the collection of all polynomials supported on monomials of degree $d \mod (q-1)$, with positive degree in y. Recall, by Lemma 5.12, that $\mathcal{F}_1$ has no constraints of locality $Q^{\lfloor \ell/Q \rfloor}$ relative to $\mathcal{F}_2$. By Propositions 6.10 and 6.11, we also have that $\mathrm{Trace}(\mathcal{F}_1)$ has no constraints of locality $Q^{\lfloor \ell/Q \rfloor}$ relative to $\mathrm{Trace}_0(\mathcal{F}_2)$. Furthermore, since $\mathrm{Trace}_0(\tilde{m}) \in \mathrm{Trace}_0(\mathcal{F}_2)$, we have that $\mathrm{Trace}_0(\mathcal{F}_2)$ is not contained $\mathcal{F}$. Thus it suffices to show that $\mathrm{Trace}(\mathcal{F}_1)$ is contained in $\mathcal{F}$.

For this part, we proceed as in the proof of Lemma 6.14. We collect the variables of $\mathbf{x}$ in blocks with each block having degree between Q/p and Q in m. By identifying the variables within a block with copies of a single variable, we get a monomial m_1 of degree between Q/p and Q in at least ℓ variables such that $\mathrm{Trace}(m_1) \in \mathcal{F}$. Repeatedly applying Lemma 6.13 to it, we get that for every monomial $m_2 \in \mathrm{Trace}(\mathcal{F}_1)$, $\mathrm{Trace}(m_2) \in \mathcal{F}$, and thus $\mathrm{Trace}(\mathcal{F}_1) \subseteq \mathcal{F}$.

It follows that $\mathcal{F}$ has no constraints of locality $Q^{\lfloor \ell/Q \rfloor}$ relative to $\mathcal{F} + \mathrm{Trace}_0(\tilde{m})$ and hence does not have a $Q^{\lfloor \ell/Q \rfloor}$-local characterization. The lemma follows by noting that $\lfloor \ell/Q \rfloor = \lfloor q/Q^2 \rfloor \ge (d - Q^2)/Q^2$. ∎

6.5 Testing for linear invariant families

We conclude, as in Section 5.4, by giving a testing theorem for linear-invariant families. Again we remark that the test does not follow immediately from the characterization results, since the characterization are not necessarily 2-ary independent.

However, it follows directly from the results of Section 5.4, and the characterization of Section 6.2, that every linear invariant family of linear invariance degree d is $Q^{2(d+Q)/p}$-locally testable. Specifically we note that:

- Definition 5.14 of "projective" families is still applicable to families mapping $\mathbb{K}^n \to \mathbb{F}$, being subsets of $\{\mathbb{K}^n \to \mathbb{K}\}$.
- Proposition 5.15 characterizing projective families still applies.
- Lemma 5.16 giving a decomposition of every linear invariant family $\mathcal{F}$ into the sum of a family $\mathcal{F}_1$ of bounded degree and a projective family $\mathcal{F}_2$, along with a local reduction to compute a function g whose distance from $\mathcal{F}_1$ estimates the distance of f from $\mathcal{F}$, also still applies.
- The family $\mathcal{F}_1$ derived in the previous step does have a 2-ary independent local formal characterization and thus a local test.

Putting that above observations together, as in the proof of Lemma 5.18 we get:

Lemma 6.17 *Let $\mathcal{F} \subseteq \{\mathbb{K}^n \to \mathbb{F}\}$ be a linear-invariant family with a k-local characterization. Then has a $k' = 2Q \cdot (Q^2 k)^{Q^2}$-local test that accepts members of $\mathcal{F}$ with probability 1, while rejecting δ-far members with probability at least $\min\left\{\frac{\delta}{2}, \frac{Q^2}{(2k'+Q)(k'+Q)}\right\}$.*

6.6 Putting the results together

Combining Lemmas 6.14 and 6.9 and Theorem 2.9 we get a proof of Theorem 2.10.
Similarly, by combining Lemmas 6.16, 6.8, and 6.17, we get a proof of Theorem 2.11.

Part III

7 A non-trivial local formal characterization

In this section we consider a set of polynomials, of potentially very large degrees, which is affine invariant. We show that this set has a formal characterization of very small locality.

Let $\mathbb{F}_q^d[x_1, \ldots, x_n]$ denote the space of polynomials of degree at most d in n variables over $\mathbb{F}_q$. For $q = p^s$, let $\mathbb{F}_q^{\mathrm{char},d}[x_1, \ldots, x_n]$ be the space of functions

$$\{f \in \mathbb{F}_q[x_1, \ldots, x_n] | \exists g \in \mathbb{F}_q^d[z_{11}, \ldots, z_{sn}] \text{ s.t. } f(x_1, \ldots, x_n) = g(x_1, x_1^p, \ldots, x_1^{p^{s-1}}, x_2, \ldots, x_n^{p^{s-1}})\}.$$

When $d = 1$, and g is homogenous, then we get the class of linearized polynomials. Extending this terrible nomenclature, we refer to elements of $\mathbb{F}_q^{\mathrm{char},d}[\mathbf{x}]$ as *d-ized polynomials*. We claim below that the property of being a d-ized polynomial is locally testable with $(d+2)$-local tests.

Theorem 7.1 *For $q = p^m$ with p being prime and for an integer $d \leq p-2$, the family of d-ized polynomials has a $(d+2)$-local formal affine characterization. Specifically a function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ is a d-ized polynomial if and only if $\forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$, $\sum_{i=0}^{d+1} \alpha_{i,d} f(\mathbf{x} + i\mathbf{y}) = 0$, where $\alpha_{i,d} = (-1)^i \binom{d+1}{i}$.*

We remark that the degree of a d-ized polynomial over p^m may be as high as $d \cdot p^{s-1}$ and so the characterization can be quite local even when the polynomial has high degree. To prove Theorem 7.1 we use heavily the characterization from [19, 9] that for a prime field $\mathbb{F}_p$, a function $g : \mathbb{F}_p^n \to \mathbb{F}_p$ is a degree d polynomial if and only if $\sum_{i=0}^{d+1} \alpha_i g(\mathbf{x} + i\mathbf{y}) = 0$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_p^n$.

To translate results about $\mathbb{F}_p$ to results about $\mathbb{F}_q$ for $q = p^m$, we use the following correspondence from $\mathbb{F}_q$ to $\mathbb{F}_p^m$ using linearized polynomials. (For this part we also use the fact that $\mathbb{F}_p$ is contained in $\mathbb{F}_q$, given by the solutions of the equation $x^p - x = 0$.)

Proposition 7.2 *There exist maps $b : \mathbb{F}_q \to \mathbb{F}_p^m$ and $b^{-1} : \mathbb{F}_p^m \to \mathbb{F}_q$ satisfying:*

- *For every $\beta \in \mathbb{F}_q$, $\beta = b^{-1}(b(\beta))$.*
- *$b = \langle b_1, \ldots, b_m \rangle$, where $b_i : \mathbb{F}_q \to \mathbb{F}_q$ is a linearized polynomial (i.e., a polynomial of the form $b_i(x) = \sum_{j=0}^{m-1} c_{ij} x^{p^j}$) with its image being $\mathbb{F}_p$. In particular, the b_i's are $\mathbb{F}_p$-linear maps.*
- *b^{-1} is an $\mathbb{F}_p$-linear map. In particular, $b^{-1}(0) = 0$.*

We extend the maps b and b^{-1} to apply to vectors in $\mathbb{F}_q^n$ and $\mathbb{F}_q^{mn}$ using the extension $b(x_1, \ldots, x_n) = \langle b(x_1), \ldots, b(x_n) \rangle$, and b^{-1} being its inverse. Using these maps we can create an alternate characterization of the d-ized polynomials.

Lemma 7.3 *$f : \mathbb{F}_q^m \to \mathbb{F}_q$ is a d-ized polynomial if and only if there exists polynomials $g_1, \ldots, g_m : \mathbb{F}_p^{mn} \to \mathbb{F}_p$ of degree at most d such that $f(\mathbf{x}) = b^{-1}(g_1(b(\mathbf{x})), \ldots, g_m(b(\mathbf{x})))$.*

Proof: We start with the forward direction, i.e., Assume such $g_1, \ldots, g_m$ exist and prove that f is d-ized in such a case.

For both parts it is useful to see a "more" explicit version of b^{-1}. Since b^{-1} is linear, we have that there exists $w_1, \ldots, w_m \in \mathbb{F}_q$ such that $b^{-1}(\alpha_1, \ldots, \alpha_m) = \sum_{i=1}^{m} \alpha_i w_i$ (for all $\alpha_1, \ldots, \alpha_m \in \mathbb{F}_p$). Using this we see that $f(\mathbf{x}) = b^{-1}(g_1(b(\mathbf{x})), \ldots, g_m(b(\mathbf{x}))) = \sum_{i=1}^{m} g_i(b(\mathbf{x}))w_i$. But $b(\mathbf{x})$ is just a collection of mn linear forms in $x_1, \ldots, x_1^{p^{m-1}}, \ldots, x_n^{p^{m-1}}$, and g_i is a degree d polynomial in its arguments, and so their composition is degree d-ized polynomial in $\mathbf{x}$. Thus we get each $g_i(b(\mathbf{x}))$ is a d-ized polynomial in $\mathbf{x}$ and so f, which is a linear combination of such polynomials, is also a d-ized polynomial.

For the other direction, assume f is d-ized. So $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_1^{p^{m-1}}, \ldots, x_n^{p^{m-1}})$ for some degree d polynomial g. Note that

$$\begin{aligned}
g_i(\mathbf{y}) &= b_i(f(b^{-1}(\mathbf{y}))) \\
&= b_i(f(\sum_j y_{1j}w_j, \ldots, \sum_j y_{nj}w_j)) \\
&= b_i\left(g\left(\left(\sum_j y_{1j}w_j\right), \ldots, \left(\sum_j y_{1j}w_j\right)^{p^{m-1}}, \ldots, \left(\sum_j y_{nj}w_j\right)^{p^{m-1}}\right)\right) \\
&= b_i(\tilde{g}(y_1, \ldots, y_{mn}))
\end{aligned}$$

for some degree d polynomial in mn variables with coefficients from $\mathbb{F}_q$. (For the last step we use the fact that $y_i^p = y_i$ when $y_i \in \mathbb{F}_p$.) Finally we use the fact that we are only interested in the evaluations of $\tilde{g}$ over elements of $\mathbb{F}_p^{mn}$. Note that $b_i(\tilde{g})$ has the same degree as $\tilde{g}$ in this case, since b_i is $\mathbb{F}_p$-linear (and so for a monomial of the form $c \cdot \prod y_{jk}^{e_{jk}}$, we have $b_i(c \cdot \prod y_{jk}^{e_{jk}}) = b_i(c) \cdot \prod y_{jk}^{e_{jk}}$). ∎

We are now ready to prove Theorem 7.1.

Proof: We prove the forward direction first. Suppose $f \in \mathbb{F}_q^{\mathrm{char},d}[\mathbf{x}]$. We wish to show that for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$, $\sum_{i=0}^{d+1} \alpha_i f(\mathbf{x} + i\mathbf{y}) = 0$. By Lemma 7.3 we have that there exist degree d polynomials $g_1, \ldots, g_m : \mathbb{F}_p^{mn} \to \mathbb{F}_p$ such that $f(\mathbf{x}) = b^{-1}(g_1(b(\mathbf{x})), \ldots, g_m(b(\mathbf{x})))$. So we have

$$\begin{aligned}
&\sum_{i=0}^{d+1} \alpha_i f(\mathbf{x} + i\mathbf{y}) \\
&= \sum_{i=0}^{d+1} \alpha_i b^{-1}(g_1(b(\mathbf{x} + i\mathbf{y})), \ldots, g_m(b(\mathbf{x} + i\mathbf{y}))) \\
&= b^{-1}\left(\sum_{i=0}^{d+1} \alpha_i g_1(b(\mathbf{x} + i\mathbf{y})), \ldots, \sum_{i=0}^{d+1} \alpha_i g_m(b(\mathbf{x} + i\mathbf{y}))\right) \quad \text{(By the linearity of } b^{-1}\text{)} \\
&= b^{-1}\left(\sum_{i=0}^{d+1} \alpha_i g_1(b(\mathbf{x}) + ib(\mathbf{y})), \ldots, \sum_{i=0}^{d+1} \alpha_i g_m(b(\mathbf{x}) + ib(\mathbf{y}))\right) \quad \text{(By the linearity of } b\text{)} \\
&= b^{-1}(0, \ldots, 0) \quad \text{(By [19])} \\
&= 0 \quad \text{(By linearity of } b^{-1}\text{)}
\end{aligned}$$

Now for the reverse direction, suppose $f : \mathbb{F}_q^n \to \mathbb{F}_q$ satisfies $\sum_{i=0}^{d+1} \alpha_i f(\mathbf{x} + i\mathbf{y}) = 0$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$. Consider the function $g_i : \mathbb{F}_p^{mn} \to \mathbb{F}_p$ given by $g_i(\mathbf{z} = \langle z_{11}, \ldots, z_{mn} \rangle) = b_i(f(b^{-1}(\mathbf{z})))$. We now note that for any pair $\mathbf{u}, \mathbf{v} \in \mathbb{F}_p^{mn}$, we have

$$\begin{aligned}
\sum_{j=0}^{d+1} \alpha_j g_i(\mathbf{u} + j\mathbf{v}) &= \sum_{j=0}^{d+1} \alpha_j b_i(f(b^{-1}(\mathbf{u} + j\mathbf{v}))) \\
&= b_i \left(\sum_{j=0}^{d+1} \alpha_j f(b^{-1}(\mathbf{u} + j\mathbf{v})) \right) \\
&= b_i \left(\sum_{j=0}^{d+1} \alpha_j f(b^{-1}(\mathbf{u}) + jb^{-1}(\mathbf{v})) \right) \\
&= b_i(0) \\
&= 0
\end{aligned}$$

We conclude that g_i is a degree d polynomial for every $i \in [m]$. But now since $f(\mathbf{x})b^{-1}(g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$, we conclude that f must also be a d-ized polynomial. ∎

Acknowledgments

We would like to thank Oded Goldreich, Elena Grigorescu, Swastik Kopparty, Alex Samorodnitsky, and Avi Wigderson for many valuable discussions.

References

[1] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: it's all about regularity. In Kleinberg [17], pages 251–260.

[2] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing low-degree polynomials over GF(2). In *Proceedings of the 7th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM 2003), Lecture Notes in Computer Science, vol. 2764*, pages 188–199, New York, 2003. Springer.

[3] Eli Ben-Sasson, Prahladh Harsha, and Sofya Raskhodnikova. Some 3CNF properties are hard to test. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 345–354, New York, NY, USA, 2003. ACM Press.

[4] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences*, 47(3):549–595, 1993.

[5] Andrej Bogdanov, Kenji Obata, and Luca Trevisan. A lower bound for testing 3-colorability in bounded-degree graphs. In *FOCS*, pages 93–102. IEEE Computer Society, 2002.

[6] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, Balázs Szegedy, and Katalin Vesztergombi. Graph limits and parameter testing. In Kleinberg [17], pages 261–270.

[7] Stephen D. Cohen. Functions and polynomials in vector spaces. *Archiv der Mathematik*, 48(5):409–419, May 1987.

[8] P. Delsarte, J.M. Goethals, and F.J. MacWilliams. On generalized Reed-Muller codes and their relatives. *Information and Control*, 16(5):403–442, 1970.

[9] Katalin Friedl and Madhu Sudan. Some improvements to total degree tests. In *Proceedings of the 3rd Annual Israel Symposium on Theory of Computing and Systems*, pages 190–198, Washington, DC, USA, 4-6 January 1995. IEEE Computer Society. Corrected version available online at http://theory.csail.mit.edu/~madhu/papers/friedl.ps.

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

[11] Oded Goldreich and Or Sheffet. On the randomness complexity of property testing. In Moses Charikar, Klaus Jansen, Omer Reingold, and José D. P. Rolim, editors, *APPROX-RANDOM*, volume 4627 of *Lecture Notes in Computer Science*, pages 509–524. Springer, 2007.

[12] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. 2-transitivity is insufficient for local testability. In *CCC 2008: Proceedings of the 23rd IEEE Conference on Computational Complexity*, page (to appear). IEEE Computer Society, June 23-26th 2008.

[13] Charanjit S. Jutla, Anindya C. Patthak, Atri Rudra, and David Zuckerman. Testing low-degree polynomials over prime fields. In *FOCS '04: Proceedings of the Forty-Fifth Annual IEEE Symposium on Foundations of Computer Science*, pages 423–432. IEEE Computer Society, 2004.

[14] T. Kasami, S. Lin, and W. W. Peterson. New generalization of the Reed-Muller codes - Part I: Primitive codes. *IEEE Transactions on Information Theory*, 14:189–199, 1968.

[15] Tali Kaufman and Simon Litsyn. Almost orthogonal linear codes are locally testable. In *Proceedings of the Forty-sixth Annual Symposium on Foundations of Computer Science*, pages 317–326, 2005.

[16] Tali Kaufman and Dana Ron. Testing polynomials over general fields. In *Proceedings of the Forty-fifthth Annual Symposium on Foundations of Computer Science*, pages 413–422, 2004.

[17] Jon M. Kleinberg, editor. *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*. ACM, 2006.

[18] Ronitt Rubinfeld. Robust functional equations and their applications to program testing. *SIAM Journal on Computing*, 28(6):1972–1997, 1999.

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