Symmetries in Algebraic Property Testing

by

Elena Grigorescu

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2010

Author .....
Department of Electrical Engineering and Computer Science
August 16, 2010

Certified by .....
Madhu Sudan
Fujitsu Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by .....
Professor Terry P. Orlando
Chair, Department Committee on Graduate Students

Symmetries in Algebraic Property Testing

by

Elena Grigorescu

Submitted to the Department of Electrical Engineering and Computer Science
on August 16, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Computer Science and Engineering

Abstract

Modern computational tasks often involve large amounts of data, and efficiency is a very desirable feature of such algorithms. *Local algorithms* are especially attractive, since they can imply global properties by only inspecting a small window into the data. In Property Testing, a local algorithm should perform the task of distinguishing objects satisfying a given property from objects that require many modifications in order to satisfy the property.

A special place in Property Testing is held by algebraic properties: they are some of the first properties to be tested, and have been heavily used in the PCP and LTC literature. We focus on conditions under which algebraic properties are testable, following the general goal of providing a more unified treatment of these properties. In particular, we explore the notion of symmetry in relation to testing, a direction initiated by Kaufman and Sudan. We investigate the interplay between local testing, symmetry and dual structure in linear codes, by showing both positive and negative results.

On the negative side, we exhibit a counterexample to a conjecture proposed by Alon, Kaufman, Krivelevich, Litsyn, and Ron aimed at providing general sufficient conditions for testing. We show that a single codeword of small weight in the dual family together with the property of being invariant under a 2-transitive group of permutations do not necessarily imply testing.

On the positive side, we exhibit a large class of codes whose duals possess a strong structural property ('the single orbit property'). Namely, they can be specified by a single codeword of small weight and the group of invariances of the code. Hence we show that sparsity and invariance under the affine group of permutations are sufficient conditions for a notion of very structured testing. These findings also reveal a new characterization of the extensively studied BCH codes. As a by-product, we obtain a more explicit description of structured tests for the special family of BCH codes of design distance 5.

Thesis Supervisor: Madhu Sudan
Title: Fujitsu Professor of Electrical Engineering and Computer Science

Acknowledgments

I have been extremely fortunate to be Madhu's advisee – clearly! I am indebted to him for his constant support, encouragements and guidance, for his openness and availability, and for introducing me to so many beautiful research directions. It goes without saying that I have learned a lot about doing research from Madhu, and what I really like the most is the importance of asking daring questions; and being always optimistic about the proximity to a solution; and making collaboration enjoyable, and inclusive; and more. Thanks Madhu for all these, and I wish that we will collaborate again.

I loved working on the topics of this thesis with Tali. Her friendliness and kind advice have been really helpful and motivational.

I am thankful to my committee members Ronitt Rubinfeld and Piotr Indyk for agreeing to read my thesis and for being so accommodating with setting up my defense date. I have also benefited from Ronitt's support and collaboration starting early in my graduate program. Even though we haven't worked together on research, Piotr's charisma and witty sense of humor have always been refreshing.

Yet again, I have been lucky to collaborate with some wonderful colleagues on topics that have not been included in this thesis, and I am grateful to each of them: (in random order) Sofya Raskhodnikova, Madhav Jha, Victor Chen, Kyomin Jung, Arnab Bhattacharyya, Ronald de Wolf, Ronitt Rubinfeld, Irit Dinur, Swastik Kopparty, Ning Xie, David Woodruff, Asaf Shapira, Jakob Nordström.

At different stages in my graduate student life it has been great to have around, share thoughts, plans, advice and support with my colleagues and friends from the lab: (in no particular order) Tasos Sidiropoulos, Eddie Nikolova, Victor Chen, Alex Andoni, Brendan Juba, Angelina Lee, Shubhangi Saraf, Swastik Kopparty, Sergey Yekhanin, David Woodruff, Ben Rossman, Prasant Gopal, Jakob Nordström, Arnab Bhattacharyya, Kevin Matulef, Krzysztof Onak, Kyomin Jung, Ning Xie, Mihai Patraşcu, Petar Maymounkov.

Joanne Hanley's readiness to help with all the administrative issues has made things so much easier for me all these years, and Be Blackburn's sweets corner has always been a nice procrastination destination.

I am also grateful to many people who had a positive impact on my academic development before coming to MIT. To mention only a few of them, I thank Joe Gallian for his trust and guidance during his Research Experience for Undergraduates program, where I got a first taste of research. My undergraduate Math and CS professors from Bard College: Lauren Rose, Bob McGrail, Ethan Bloch, Rebecca Thomas and Sven Anderson encouraged me in all my endeavors. I am also extremely indebted to my Math and Physics professors from high-school: Constantin Grigoriu, Dorel Haralamb, and Camelia Neţa, whose impeccable teaching styles and interest for scientific rigor were exemplary.

I thank all my friends for being such a great and spirited company, for inventing distractions from research and for keeping me in touch with the world outside the lab.

Mom, Dad, Miha and Tanveer, your role in my life is beyond words!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Decision problems occur ubiquitously in computation, and while very often they might be hard to solve exactly, many models have been proposed to approach reasonably efficient relaxations. In Property Testing, the model of our focus, the goal is to distinguish between objects that belong to a class, and objects that differ in many locations from each object in the class. A tester algorithm in this model is allowed to accept objects that do not belong to the class, but which are close to some object in the class. The algorithm should use randomness and run in sub-linear time, hence it should make a correct decision after reading only a restricted, but judiciously chosen portion of the input. Such so-called *local* algorithms are desirable in settings where the decision problem might not be known to be efficiently computable, but also even when it admits polynomial (super-linear) time algorithms.

Surprisingly, many natural questions have been shown to admit local algorithms with good error parameters. A domain where local algorithms are both desirable and computationally feasible is in protecting information from errors. In practice for example, data storage devices use redundant encoding in order to facilitate recovery from errors. However, when the amount of error is too large, recovery could become too time consuming, and in fact unaffordable. Local algorithms to test membership in error correcting codes might be employed in these scenarios in order to quickly decide what data is recoverable.

The focus of this thesis is on testing membership in error correcting codes. Error correcting codes can be viewed as a special subclass of algebraic families of functions. The ultimate driving goal of this research is to understand features that distinguish between algebraic families for which membership can be tested in a time-efficient manner, and those which provably require non-local algorithms. This direction is strongly motivated by intimate connections with Locally Testable Codes and Probabilistically Checkable Proofs, and has received wide attention in recent years.

1.1 Property testing and algebraic property testing

A brief history Blum, Luby and Rubinfeld [33] initiated the field of Property Testing by proposing a tester for the class of linear functions. The immediate follow-up works of Rubinfeld and Sudan [89], Babai, Fortnow and Lund [15], and Babai, Fortnow, Levin and Szegedy [14], considered testing polynomials of higher degrees in various settings of parameters and domains. Algebraic property testing rapidly took off ever since, and has so far found myriads of applications across theoretical computer science. These results were instrumental in the proofs of MIP=NEXP [15] and in the highly acclaimed PCP Theorem of Arora, Lund, Motwani, Sudan and Szegedy [12]. Testing graph properties has also been well-studied. Graph property testing was introduced by Goldreich, Goldwasser and Ron in [53] who considered properties such as colorability, bipartiteness or connectedness in the dense graph model. The recent results of Alon et al. [5] and Borgs et al. [35] completely characterized testable properties in this model.

Property testing Formally, a tester for a property $\mathcal{P}$ is a randomized algorithm which can access a given object given as a black-box via *queries*. Based on the answers to the queries, the algorithm makes an accept/reject decision which satisfies the following conditions: if the object belongs to $\mathcal{P}$ it should accept; otherwise, if the

object is *far* from $\mathcal{P}$ it should reject with high probability over the randomness of the algorithm. $\mathcal{P}$ is thought of as a collection of properties, i.e. $\mathcal{P} = \cup_n \mathcal{P}_n$ for $n \to \infty$, where the objects in $\mathcal{P}_n$ have size n. The tester is *local* if the number of queries it performs does not depend on the size of the input object, that is, the number of queries is independent of n. The notion of distance to $\mathcal{P}$ is captured by metrics dependent on the descriptions of the objects belonging to the property. When the property refers to sets of graphs, distance is measured in the number of edges to be added to or removed from the input graph in order to build a graph in the target collection. In this thesis we concentrate on *algebraic properties*, namely collections of functions $f : D \to R$, mapping a vector space D over a field, into a subfield R. In this case the notion of distance is given by the Hamming metric, which is the number of places the function needs to be modified in order to obtain a function from the target family. A tester could be *adaptive* if the choice of the queries it makes depends on the answers it had received; it is *non-adaptive* otherwise. The definition above describes a *single sided* tester; when the test may also err on objects in the family it is called *double sided.*

Locally Testable Codes *Locally Testable Codes* (LTCs) form a special class of algebraic testable properties. They are error correcting codes for which membership can be tested with a small number of queries. Error correcting codes are usually described by sets of vectors, called *codewords*, such that the Hamming distance between pairwise vectors is somewhat large. Their design parameters are the *block length* N and *alphabet* Σ. An alternate view of a code $\mathcal{C} \subseteq \Sigma^N$ is as collections of functions in $\{D \to R\}$, where $|D| = N$ and $R = \Sigma$ (i.e. if $c \in \mathcal{C}$ the function $c(\alpha) = c_\alpha$ $\forall \alpha \in D$.) The set of linear functions, and the set of low degree polynomials are some basic examples of LTCs. Building LTCs with good parameters has been a direction of active research [57, 22, 24, 44, 84], bearing strong connections with similar trends in the study of PCPs. In this thesis we focus on a second intriguing direction in the area, namely on understanding what features of codes/algebraic properties could reveal global structure from an average local structure. The current trends in the

study of LTCs have been recently surveyed in [21].

1.2 Structure and symmetry in testing

A starting example In coding theoretic terms, testing whether a function is linear [33] corresponds to testing membership in the Hadamard code, formally denoted by

$$H = \{f_a : \mathbb{F}_2^n \to \mathbb{F}_2 \mid f_a(x) = \sum_{i=0}^{n-1} a_i x_i, a \in \mathbb{F}_2^n\}.$$

To describe a tester for this family one uses the evident fact that for any $\alpha, \beta \in \mathbb{F}_2^n$ every linear function $f \in H$ satisfies the *constraint* $f(\alpha) + f(\beta) + f(\alpha + \beta) = 0$. A tester for linearity simply picks uniformly at random $\alpha, \beta \in \mathbb{F}_2^n$ and queries f at α, β and $\alpha + \beta$. It accepts if and only if $f(\alpha) + f(\beta) + f(\alpha + \beta) = 0$, and rejects otherwise.

An alternate way to represent this test is as a binary vector of *weight* 3, indexed by elements of $\mathbb{F}_2^n$ and supported on $\alpha, \beta, \alpha + \beta$. This vector has the property that it has inner product 0 with each codewords of H. This view will become useful in analyzing future testers.

Consider now the set of all binary vectors whose inner product is 0 with each codeword in H. As we will see later on, each such vector could represent *a test* for the Hadamard code, and its support size is called the *locality* of the test. This set of vectors forms a vector space H'. Moreover, this set *characterizes* the Hadamard code, in the sense that no other function $g : \mathbb{F}_2^n \to \mathbb{F}_2$, has inner product 0 with each function in H'. Formally, H' is the *dual* of the Hadamard code, i.e. the Hamming code.

Linearity and Duality The example above illustrates the concepts we will be working with in this thesis. We focus on *linear* families of functions[1], namely families for which if $f, g \in \mathcal{P}$ then $f + g \in \mathcal{P}$. Viewed as a collection of vectors, a linear family is just a vector space. Notice that the Hadamard and Hamming codes are

[1]Not to be confused with families of linear functions.

linear codes. Each linear family can be associated with a unique dual family, which is the vector space dual to it. In other words, the dual of $\mathcal{P} \subseteq \{\mathbb{F}_2^n \to \mathbb{F}_2\}$, denoted $\mathcal{P}^\perp$, is $\{g : \mathbb{F}_2^n \to \mathbb{F}_2, \ \sum_x f(x)g(x) = 0, \quad \forall f \in \mathcal{P}\}$. The dual family is of central interest in testing linear properties, since essentially, every test (even in the adaptive or double-sided setting) must belong to it [23]. This point has motivated the need for a better understanding of the *structural features* of the dual family that are relevant to testing. Our work delves into this connection with the goal of identifying necessary and sufficient conditions for local testability.

Symmetry in algebraic properties The structural features of a family can be studied from the perspective of the set of symmetries that the family exhibits. The initial systematic study by Kaufman and Sudan [72] on the role of symmetries in algebraic property testing has sparked a wave of great interest in this connection [60, 61, 74, 25, 54, 31, 30].

A group of *symmetries or invariances* acting on a family $\mathcal{P} \subseteq \{f : D \to R\}$ is a set of functions $\pi : D \to D$ such that $f \in \mathcal{P}$ if and only if $f \circ \pi \in \mathcal{P}$. When dealing with codes, it is most common to only consider functions π which are permutations, and hence $f \circ \pi$ is a permuted codeword. The largest group of symmetries acting on a family is called the *automorphism* group. [2]

Many common families of algebraic properties exhibit large automorphism groups. Invariance under *linear transformations* of the domain is a most commonly encountered symmetry. The Hadamard code, and Reed Muller codes are invariant under the linear group $\{g : \mathbb{F}_2^n \to \mathbb{F}_2^n \mid g(x) = Ax, A \in \mathbb{F}_2^{n \times n}\}$ (a.k.a $\mathrm{GL}(n, 2)$).

Invariance under *affine transformations* occurs also commonly. For example, Reed Muller codes are invariant under the affine group $\{g : \mathbb{F}_2^n \to \mathbb{F}_2^n \mid g(x) = Ax + b, A \in \mathbb{F}_2^{n \times n}, b \in \mathbb{F}_2^n\}$ (a.k.a $\mathrm{AGL}(n, 2)$)[3].

[2] In the coding literature, the set of all permutation that keep the code invariant is called the *permutation* group. The *automorphism* group of a code includes, besides permutations, transformations that multiply each element of a function by a non-zero element of the field. In this thesis we only focus on binary functions, and thus the two groups coincide. In a few places we slightly abuse its common usage by calling it the automorphism group when we only mean permutation group.

[3] Strictly speaking, $\mathrm{AGL}(n, 2)$ and $\mathrm{GL}(n, 2)$ only refer to nonsingular transformations A. We are

In this thesis we concentrate on linear and affine groups of symmetries and analyze the testability of codes that feature these invariances.

1.3 A few motivating questions

To summarize our introductory exposition, in this work we investigate the relations among (1) bases of low weight vectors, (2) affine/linear invariance and (3) testability. In this section we propose a few basic questions tackling the interplay between these notions. We describe our results in more detail in Section 1.4.

Codes with/without bases of low weight Ben-Sasson et al. [23] formalized the fact that any tester (even adaptive or double-sided ones) for membership in a linear family $\mathcal{F}$ can be reduced to picking $g \in \mathcal{F}^{\perp}$. That is, the tester queries a given function f at locations in the support of g. This result motivated understanding the testability of families whose duals are generated by bases of low weight functions.

Bases of low weight functions have been very relevant in testing many common families. Returning to our starting example of the Hadamard code, one can prove that the Hamming code is generated by the weight 3 codewords, i.e by the minimum weight codewords. Similarly, bases of low weight functions were used in testing Reed Muller codes [7, 63] and dual-BCH codes [67].

However, the existence of a low weight arbitrary basis is not sufficient for testing [23]. Random Low Density Parity Check (LDPC) codes, although generated by weight 3 codewords, require a large number of queries in order to test for membership. Vaguely speaking, this phenomenon is caused by the fact that large collections of low weight functions cannot combine into (sum up to) another low weight function, which would be a necessary condition.

Understanding when a family can be characterized by low weight functions is an important step in understanding its testability. This task could be highly non-trivial:

abusing notation slightly.

given an arbitrary basis possibly of large weight vectors is there a small weight basis for it? What if the family is represented by collections of functions, or say polynomials? What if the family exhibits some large group of symmetries? We note that techniques to analyze this type of questions are rare in the literature: a successful approach was made in [23], by means of analyzing properties of expander graphs.

This leads us to a first question that motivates out work.

Question 1 *Can one exhibit explicit families that are invariant under a large group of symmetries, contain low weight functions but cannot be characterized by low weight functions? That is, can one exhibit such a family for which any basis must contain a large weight vector?*

The family we exhibit for an answer to this question will also be relevant to showing our negative testing results. In particular, we prove that the dual of an affine subcode of Reed Muller codes of order 2 cannot have a low-weight basis (See Chapter 4.) This result has applications to a conjecture of Alon et al. [7] (the AKKLR conjecture), which attempted to unify specific testing approaches in the literature. Namely, the conjecture stated that families that are '2-transitive' and admit small constraints should be testable by tests with small locality.

The AKKLR [7] conjecture has been so far an important source of sufficient conditions for testability in the literature, motivating the major contributions that the work of Kaufman and Sudan [72] has brought in this direction.

The single orbit characterization in common codes Kaufman and Sudan [72] realized that the symmetries of a family can lead to more structured bases and furthermore to structured tests. In particular, they focused on families that can be completely characterized by a single function and its set of transformations under a group of symmetries. This property is denoted by *single orbit characterization* under a group of symmetries of the family.

Again, the Hadamard code provides an example of families that admit single orbit characterization. Indeed, every codeword of the Hamming code can be obtained as a linear combination of the set of permutations $\mathrm{AGL}(n, 2)$ of the codeword supported on $\langle e_1, e_2, e_1 + e_2 \rangle$ (here e_1 and e_2 are the standard basis vectors in $\mathbb{F}_2^n$.)

This property similarly holds for Reed Muller codes. For different settings of field size versus degree, a single orbit characterization exists, but its description might differ with the setting. For example, for Reed Muller codes of order d in fields of characteristic 2 (that is, $\mathrm{RM}(d) = \{f : \mathbb{F}_2^n \to \mathbb{F}_2 | \deg(f) \leq d\}$,) a single orbit characterization of the dual is given by a function supported at $\{\mathrm{Span}(\alpha_0, \alpha_2, \ldots, \alpha_d) + \beta\}$, for some $\alpha_0, \ldots, \alpha_d, \beta \in \mathbb{F}_2^n$. For degree $d < |\mathbb{K}|$ and $\mathrm{RM}(d) = \{f : \mathbb{K}^m \to \mathbb{K}, \deg(f) \leq d\}$ a single orbit generator for the dual is supported at $\{1, w, \ldots, w^{d+1}\}$, where w is a primitive element of $\mathbb{K}$.

The single orbit property is to some extent expected under such large groups of symmetries (i.e. $\mathrm{AGL}(2, n)$): one should able to find a basis of dimension, say $O(2^n)$, among 2^{n^2} vectors. A more intriguing question is whether such single orbit bases could be found for smaller groups of invariances, which leads to a 2nd sequence of questions that we propose.

Question 2 *Do Reed Muller codes have the single orbit property under a smaller group of symmetries? Also do BCH codes (which are also extensively studied codes) have the single orbit characterization property? If so, what is the smallest group under which this property holds for BCH codes?*

We show that Reed Muller codes (in characteristic 2) have the single orbit property under even smaller groups of invariances, namely under affine groups over a domain $\mathbb{F}_{2^n}$ (a.k.a. $\mathrm{AGL}(1, 2^n)$) (see Chapter 5). Furthermore, we show that this property also holds for BCH codes. In some cases, BCH codes have this property under an even smaller group of invariances, i.e. under linear transformations of $\mathbb{F}_{2^n}$ (a.k.a. $\mathrm{GL}(1, 2^n)$) (see Chapter 6).

Structured testing in general settings The reason why the single orbit characterization is a relevant feature is due to the fact that it immediately leads to a notion of *structured testing.* A structured tester simply picks a random permutation from a group of invariances and computes its composition with a low weight generator of the single orbit. Kaufman and Sudan [72] showed that having the single orbit property under affine invariant transformations implies structured testing. Structured tests are nice since as soon as a single low weight generator is known, the rest of the tests are immediately explicitly specified by a group of symmetries. This observation prompts us to another set of questions of broad interest.

Question 3 *What general families of functions admit structured tests under the affine group? What general families have the single orbit property under other groups?*

To answer these questions we show some general families that have the single orbit characterization under affine and linear invariances (see Chapter 6.)

We proceed with a more detailed account of our results.

1.4 Our results

We focus on families $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ that are invariant under affine/cyclic transformations of $\mathbb{F}_{2^n}$ and show positive and negative testing results.

1.4.1 A counterexample to the AKKLR conjecture

As mentioned, our first result provides an answer to Question 1 discussed above and a counterexample to the AKKLR conjecture (in Chapter 4.) We provide an explicit family $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ that is invariant under a group of 2-transitive transformations, but that cannot be tested with a small number of queries. This counterexample is a subcode of Reed Muller codes of order 2, and in addition, it is affine invariant. We argue that any small weight vector in $\mathcal{F}^\perp$ must belong to the

dual of the Reed Muller code of order 2. Hence any basis for the dual of $\mathcal{F}$ must have a large weight codeword. For an illustration of the broader context of this example see Figure 1.4.3.

2-transitivity Informally, a family of functions is 2-transitive if any two coordinates look the same as any other two. This notion of symmetry is related to that of pairwise independence, which in turn plays a crucial role in the analysis of self-correctors for linear properties. To be more precise, a self-corrector for a function f (which is assumed to be correct on most inputs) computes the value of f at each location x, with high probability, from the value of the function at a few other places. Self-correctors and testers for linear properties have very similar features. While local testers use functions of small weight in the dual to test for membership, self-correctors use the same dual-functions in order to correct corrupted locations in the given function. A common analysis argument on the success probability of self correctors (e.g. [33, 7, 70]) relies on the fact that $f(x)$ can be computed from values $f(x_i)$ such that x and x_i are almost pairwise independent. The fact that 2-transitive codes with small dual distance are correctable was formalized in [73]. These apparent interconnections between 2-transitivity, correctability and testability prompted Alon et al. [7] to propose 2-transitivity as an indicator for local testability. Finally, we note that the most natural 2-transitive groups are the affine groups, and in fact families that are 2-transitive but not affine invariant are non-trivial to construct.

Supporting evidence The proposed sufficient conditions were an initial attempt at a more unified theory of the features that enable testability in Hadamard, Reed Muller, and BCH codes.

A confirmation of the conjecture in a broader context was exhibited in [72]. Their results show that the existence of a low weight dual function together with invariance under the affine group implies the existence of a special basis (the single orbit characterization), which in turn implies testability.

As additional evidence from the negative side, the conjecture does not hold for random LDPC codes. Clearly such codes do not have large groups of symmetries, and in particular they are not invariant under a 2-transitive group. As shown by Ben-Sasson et al. [23] random LDPC codes require a large number of queries in order to test for membership, even though their dual may contain small weight functions (in fact they may contain a basis of small weight functions.)

Implications Disproving the conjecture is a step toward a better understanding of the structural properties that enable testing in algebraic settings. Our results here can be stated as saying that for families $\mathcal{F} \subseteq \{\mathbb{K}^m \to \mathbb{F}\}$ where $|\mathbb{K}| \to \infty$ (in our case $m = 1$,) affine invariance and the existence of a small weight dual function does not imply local testing. This contrasts with the case when $\mathbb{K}$ is of fixed size [72]. Even though the conjecture is false in general, it has the merit of identifying symmetry (2-transitivity,) as a possible indicator of testability. This view has inspired positive results in the same vein and has been expanded in subsequent works, including this thesis.

1.4.2 Explicit structured testing in some BCH codes

Chapters 5 and Chapter 6 give partial answers to Question 2 described above. In Chapter 5 we start our study of the single orbit characterization in BCH codes, by presenting an explicit structured basis for the restricted family of BCH codes of design distance 5. We also show that Reed Muller codes can be generated by an explicit single orbit under the group $\mathrm{AGL}(1, 2^n)$. Furthermore, the results of Chapter 6 imply that general BCH codes can be generated by a single low weight codeword under the affine group $\mathrm{AGL}(1, 2^n)$, for some settings of the field size. Moreover, BCH codes admit low weight single orbit generators even for slightly smaller groups, namely under the linear (cyclic) group $\mathrm{GL}(1, 2^n)$.

Explicit succinct representation As far as we are aware, explicit single orbit generators of low weight have only been known for families such as the Hadamard code and Reed-Muller codes. For these codes the explicit description is to some extent obvious: the tests are supported on affine subspaces, or on evenly spaced points on a line ([33, 89, 7, 63, 69]).

We only show explicit single orbit generators of low weight in BCH codes of design distance 5 (see Chapter 5). The novelty of this result lies in the fact that the support of a single orbit generator under $\mathrm{AGL}(1, 2^n)$ can be described by a fixed set of carefully chosen univariate polynomials. This is a somewhat surprising uniform description of such codes, as n grows. While this result is a modest contribution, we believe that the question of finding fully explicit generators for BCH codes of general design distance could open an interesting direction of further investigation.

Previous works Counting arguments using MacWilliams identities can show that BCH codes of small design distance must contain small weight codewords. While BCH codes are well-studied, only recently Kaufman and Litsyn [68] showed that these codes have a basis of almost smallest weight codewords. Such a basis was however arbitrary and unstructured and would require $O(2^n)$ bits to specify. Our result gives a basis that require only $O(n)$ bits to specify (i.e. the support of the generator of the single orbit characterization.)

1.4.3 Sufficient conditions for structured testing

In Chapter 6 we attempt to answer Question 3 above and provide some more general sufficient conditions for the existence of the single orbit property. We show that duals of families $\mathcal{F} \subseteq \{F_{2^n} \to \mathbb{F}_2\}$ that are invariant under affine transformations of $\mathbb{F}_{2^n}$ and which contain a small number of codewords (namely, they are *sparse*,) must contain a small weight function that generates a basis (under the action of the affine group). See Figure 1.4.3 for an illustration of these families in a broader hierarchy. Hence, we exhibit a general class of affine invariant families that admit structured testing.

We note that we can only show our result in some restricted settings of n, which are implied by the number theoretic machinery that we make use of.

We also consider families that are cyclic invariant. Here we also show that duals of sparse families that are invariant under the cyclic group must have a low weight single orbit generator. Since a set of cyclic permutations is about a factor of 2^n smaller than a set of affine permutation of a given word, this is a somewhat stronger result. However, the settings of n for which it holds are more restrictive than before.

Settings of n Our results for affine invariant families hold when n is prime, a condition that ensures that $\mathbb{F}_{2^n}$ does not have non-trivial subfields, as required by Bourgain's number theoretic results that we employ. For the cyclic case, we additionally need that $2^n - 1$ does not have large divisors. This condition is satisfied in particular when $2^n - 1$ is a Mersenne prime. We believe that our results should hold regardless of these restrictions, however our approach could lead to more general settings only if the number theoretic tools we use can be generalized.

1.5 Organization and credits

Credits This work has been written in collaboration with Tali Kaufman and Madhu Sudan. Chapter 4 appeared in [60]; Chapter 6 appeared in [61]; Chapter 5 has not appeared in published form.

Organization In Chapter 2 we introduce some basic definitions that we will use throughout the thesis. In Chapter 3 we give polynomial descriptions of families that are affine and cyclic invariant. We continue with presenting a counterexample to the AKKLR conjecture in Chapter 4. We start our study of the single orbit property by considering explicit tests for the common codes Hadamard, Reed Muller and some BCH in Chapter 5. In Chapter 6 we present our general conditions for succinct representation, and finally we describe some open problems in Chapter 7.

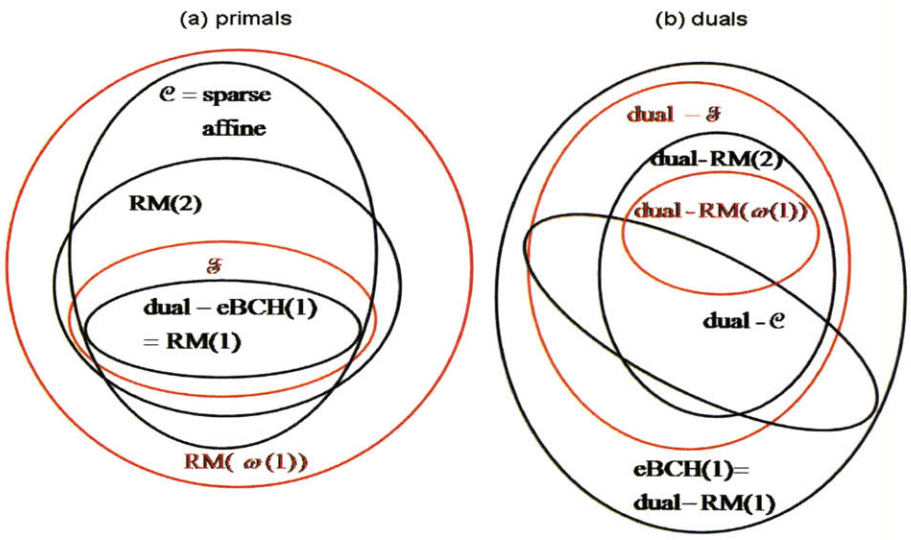

Figure 1-1: Relations among the families considered in this work. In (a): the red/black sets represent non-testable/testable families, respectively. In (b): diagram of the duals of families in (a). The red families also correspond to vector spaces that do not have a basis of low weight vectors. $\mathcal{F}$ represents the counterexample to the AKKLR conjecture from Chapter 4. $\mathcal{C}$ is a sparse, affine invariant family discussed in Chapter 6.

1.6 Bibliographic notes

1.6.1 Testing various properties

The test for linearity [33] generated numerous works on improved analysis [20, 75, 62, 92, 95], and on further generalizations to higher degree polynomials. The initial results [89, 15, 14] dealt with field sizes larger than the degree. More recently, low degree polynomials have also been considered over small characteristic by Alon et al. [7], and for other field sizes smaller than the degree, in works of Kaufman and Ron [69] and Jutla et al. [63]. Many other results demonstrated improved parameters for low degree tests in various settings [13, 48, 50, 90]. In graph properties, [53] also opened up the way for a long list of notable works [2, 4, 9, 8, 34, 66, 55, 5, 35] that considered different models and a wide range of related questions.

Another domain of interest has been in testing boolean functions. Testing dictatorships, juntas, monotonicity, sparsity are just a few examples in the vast literature in

the area [86, 32, 43, 82, 58, 52, 47, 1]. Property testing has also been intensely investigated in the setting of testing distributions. Some examples of properties considered there are uniformity, statistical distance of pairs of distributions or entropy [18, 17, 19, 97, 3]. For detailed surveys on developments on these aspects, see [51, 87, 45, 78, 88].

1.6.2 Symmetries and testing

Assorted results Symmetry has been invoked rather implicitly in testing results before the work of Kaufman and Sudan [72].

Graph testing in the dense model is one area where necessary and sufficient conditions are well understood by now [5, 35]. An implicit feature of classes of graphs is invariance under vertex relabeling. Their group of symmetries could be in fact much larger. For example, the group of symmetries of the class of bipatite graphs includes invariances under vertex removal or edge removal. In general, a graph property is testable if and only if it is 'regularly reducible' - a notion that abstracts invariance under a large group of actions on these graphs. This observation has motivated the search for similar groups of symmetries in algebraic settings that could enable testing applications.

Properties that are symmetric under data relabellings have been also considered in testing distributions. In [97] Valiant considers such questions as testing the entropy of distribution or testing closeness between distributions. Symmetry was studied in cyclic codes by Babai et al. [16] to show that no good cyclic codes are testable. Also Goldreich et al. [56] studied symmetric properties and showed bounds on the randomness complexity of testing these families.

Kaufman and Sudan's results Symmetry has been singled out explicitly as a common characteristic of known algebraic testable families only recently by Kaufman and Sudan in [72]. There they focus on general algebraic families of function $f : \mathbb{K}^m \to \mathbb{F}$, where $\mathbb{K}, \mathbb{F}$ are finite fields and $\mathbb{F}$ is a subfield of $\mathbb{K}$. Any such function is

in fact a special type of m-variate polynomial over $\mathbb{K}$, which takes values only in $\mathbb{F}$. Affine/linear invariance restricts the monomials that may occur in these families, and understanding monomial structure eventually leads to pinning down characteristics of the dual generators of the family.

The general question of focus is how does the existence of a low weight function relate to the existence of a low weight basis (characterization,) and furthermore to testing in linear or in affine invariant families?

They show a few main results in this direction.

1. First, they prove that families $\mathcal{F} \subseteq \{\mathbb{K}^m \to \mathbb{F}\}$ whose duals admit a weight k single orbit under affine/linear transformations of the domain $\mathbb{K}^m$ can be tested by tests of locality roughly k.

2. They then relate the existence of a low weight function to the existence of a low weight single orbit. In families that are invariant under affine transformations of the domain $\mathbb{K}^m$, the existence of a function of weight k implies the existence of a single orbit of locality at most $g(k, |\mathbb{K}|) = (k \cdot |\mathbb{K}|^2)^{|\mathbb{K}|^2}$. Notice that this quantity is independent of m, and it is constant when both k and $|\mathbb{K}|$ are constant. This dependency was slightly improved by Lin et al. [79] to a quantity that is still exponential in $|\mathbb{K}|$.

3. They also relate the existence of an arbitrary low weight basis (characterization) to the existence of a low weight single orbit. In families that are invariant under linear transformations of the domain $\mathbb{K}^m$ a k-weight basis implies a $g(k, |\mathbb{K}|)$-single orbit (under linear transformations).

This perspective unravels a unified view of case specific analyses for testing Hadamard and Reed Muller codes [33, 89, 7, 63, 69], since these families do admit single orbit generators of small weight. It also motivates our search for other families whose testability is owed to the property that a low-weight single orbit generates the dual.

Chapter 2

Preliminaries

We start with some standard notation. $[N]$ denotes the set $\{1, 2, \ldots, N\}$. A finite field of $q = p^t$ elements is denoted by $\mathbb{F}_q$, where p is the prime characteristic of the field; $\mathbb{F}_q^* = \mathbb{F}_q - \{0\}$. A *primitive element* of a field $\mathbb{F}_q$, denoted w, is such that $\mathbb{F}_q = \{0, 1, w, w^2, \ldots, w^{q-2}\}$. The inner product of two vectors $x, y \in \mathbb{F}^n$ is denoted by $\langle x, y \rangle = \sum_{i=1}^{n-1} x_i y_i$. The inner product between two functions $f, g : D \to \mathbb{F}_2$ is $\langle f, g \rangle = \sum_{x \in D} f(x)g(x)$. For a non-negative integer s, let $\mathrm{bwt}(s)$ be the binary weight of s (i.e., if $s = \sum s_i 2^i$ then $\mathrm{bwt}(s) = \sum s_i$.)

2.1 Algebraic properties

Algebraic properties An *algebraic property* is a collection of functions $\{f : D \to R\}$, where D and R are finite of infinite domains and ranges. Most often D and R are vector spaces over finite fields. In particular, $D = \mathbb{K}^{\ell_1}$ and $R = \mathbb{F}^{\ell_2}$ where $\mathbb{K}$ is some finite extension of $\mathbb{F}$. As it is the case in Property Testing, we study families of algebraic properties indexed by $n \to \infty$, namely families $\mathcal{F}_n \subseteq \{f : \mathbb{K}^{\ell_1(n)} \to \mathbb{F}^{\ell_2(n)}\}$. We will often drop the subscript n whenever the family is clear from the context.

In this thesis we focus on binary functions ($R = \mathbb{F}_2$) and on domains of size 2^n or $2^n - 1$. This representation will also be useful at describing error-correcting codes, our focus throughout this work.

Functions vs vectors A function $f : D \to \mathbb{F}_2$ can be represented as a vector in the following sense. Consider a fixed order of the elements in D, say $e_1, e_2, \ldots, e_{|D|}$ and represent f by its evaluation table at these points, i.e. by the vector $\langle f(e_1), f(e_2), \ldots, f(e_{|D|})\rangle$. Similarly, given a binary vector of length N we view it as the evaluation table of a function $f : D \to \mathbb{F}_2$, with $|D| = N$. Throughout this work we will alternate between these two equivalent descriptions of codes and families of functions without further comment.

Linearity A property $\mathcal{F} \subseteq \{f : D \to \mathbb{F}_2\}$ is *linear* if it satisfies the condition that, if $f, g \in \mathcal{F}$ then $f + g \in \mathcal{F}$. A linear property is essentially a vector space over $\mathbb{F}_2$. A *basis* for $\mathcal{F}$ is a set of functions (vectors) that generate it.

Dual of a linear property One can associate with every linear property $\mathcal{F}$ another property, denoted the *dual* of $\mathcal{F}$ defined as follows:

$$\mathcal{F}^{\perp} = \{g : D \to \mathbb{F}_2 | \ \langle f, g\rangle = \sum_{x \in D} f(x)g(x) = 0 \text{ for all } f \in \mathcal{F}\}.$$

Duals of linear properties are extremely useful objects in the context of property testing. One is interested in particular in understanding their structure in terms of functions of small weight.

Weight The *weight* of a function $f : D \to \mathbb{F}_2$ denoted $\mathrm{Wt}(f) = |\{x \in D | f(x) \neq 0\}|$. The *relative weight* is defined as $\mathrm{wt}(f) = \frac{\mathrm{Wt}(f)}{|D|}$.

Hamming distance The *Hamming distance* between $f, g \in \mathcal{F}$ is defined as $\mathrm{Wt}(f-g)$. Similarly, the relative Hamming distance between f, g is $\mathrm{wt}(f, g)$. For a function $g : D \to \mathbb{F}_2$ and property $\mathcal{F} = \{f | f : D \to \mathbb{F}_2\}$ the *distance* from f to $\mathcal{F}$ is $\mathrm{Dist}(f, \mathcal{F}) = \min_{g \in \mathcal{F}} \mathrm{Wt}(f - g)$, and the relative distance from f to $\mathcal{F}$ is $\mathrm{dist}(f, \mathcal{F}) = \min_{g \in \mathcal{F}} \mathrm{wt}(f - g)$. We will most often use the notion of relative distance.

2.2 Error-correcting codes

Typically, an error-correcting code $\mathcal{C}$ over an *alphabet* Σ is defined as the image of a function C that encodes a *message* $m \in \Sigma^K$ into a *codeword* $C(m) \in \Sigma^N$. The alphabet Σ is usually a finite field $\mathbb{F}$, or a vector space over $\mathbb{F}$.

In particular, one can now define the *distance* of $\mathcal{C}$ by $\Delta(\mathcal{C}) = \min_{c_1,c_2 \in \mathcal{C}} \mathrm{Wt}(c_1 - c_2)$ and its *relative distance* by $\delta(\mathcal{C}) = \min_{c_1,c_2 \in \mathcal{C}} \mathrm{wt}(c_1 - c_2)$. The dual of a linear code $\mathcal{C}$ was also implicitly introduced in the previous section as $\mathcal{C}^\perp = \{y |\ \langle y, c \rangle = 0\ \forall c \in \mathcal{C}\}$.

2.3 Property testing and locally testable codes

We begin by defining 2-sided local testers, and then a 1-sided strong testers. The former testers are used in our negative results, while the latter in the positive sufficient testing conditions.

Definition 4 (k-local tester) *For integer k and reals $\epsilon_2 > \epsilon_1 \geq 0$ and $\delta > 0$, a* $(k, \epsilon_1, \epsilon_2, \delta)$-local test *for a property $\mathcal{F}$ is a probabilistic algorithm that, given oracle access to a function $f \in \mathcal{F}$, queries f on k locations (probabilistically, possibly adaptively), and* accepts *$f \in \mathcal{F}$ with probability at least $1 - \epsilon_1$, while* accepting *functions f that are δ-far from $\mathcal{F}$ with probability at most $1 - \epsilon_2$. Property $\mathcal{F}$ is called* $(k, \epsilon_1, \epsilon_2, \delta)$-locally testable *if it has a $(k, \epsilon_1, \epsilon_2, \delta)$-local test.*

Given an ensemble of families $\mathcal{F}' = \{\mathcal{F}_n\}_n$, we say $\mathcal{F}'$ is k-locally testable *if there exist $0 \leq \epsilon_1 < \epsilon_2$ and $\delta > 0$ such that for every n, $\mathcal{F}_n$ is $(k, \epsilon_1 + o(1), \epsilon_2 - o(1), \delta)$-locally testable (where the $o(1)$ term goes to zero as $n \to \infty$).*

We will often drop the subscript n when this is clear from the context.

If both $\epsilon_1 > 0$ and $\epsilon_2 > 0$ then the tester is *double-sided*, else it is *single sided.* A single-sided tester is *perfect* if it accepts every function $f \in \mathcal{F}$ with probability 1.

A tester is called *adaptive* if the queries it makes are based on answers to previous queries. Otherwise, namely when it can send all its queries at once, it is called *non-adaptive.*

Definition 5 (strong k-local tester [57]) *For integer k and real $\alpha > 0$, a* (k, α)-strong local tester *for a property $\mathcal{F}$ is a probabilistic algorithm that, given oracle access to a function $f \in \mathcal{F}$, queries f on k locations (probabilistically, possibly adaptively), and* accepts *$f \in \mathcal{F}$ with probability at least* 1, *while* accepting *functions f that are δ-far from $\mathcal{F}$ with probability at most $1 - \alpha \cdot \delta(f, \mathcal{F})$.*

$\mathcal{F}$ is said to be strongly locally testable if there exist $k < \infty$ and $\alpha > 0$ such that $\mathcal{F}$ is (k, α)-locally testable.

Accordingly, we can define weak and strong Locally Testable Codes.

Definition 6 (Locally Testable Code-weak version) *An error-correcting code $\mathcal{C}$ is* $(k, \epsilon_1, \epsilon_2, \delta)$-locally testable *for some integer k and reals $\epsilon_2 > \epsilon_1 \geq 0$ and $\delta > 0$, if there exists a $(k, \epsilon_1, \epsilon_2, \delta)$-local tester for the property $\mathcal{C}$.*

Definition 7 (Locally Testable Code- strong version) *An error-correcting code $\mathcal{C}$ is* (k, α)-locally testable *for some integer k and real $\alpha \geq 0$, if there exists a (k, α)-local tester for the property $\mathcal{C}$.*

2.4 Invariance

Let $\mathcal{F} \subseteq \{f : D \to \mathbb{F}_2\}$ be a family of binary functions and let $G \subseteq \{\pi : D \to D\}$ be a set of transformations of the domain. A transformation $\pi : D \to D$ is a *permutation* if it is a bijection. We say that $\mathcal{F}$ is *invariant* under G if for every $f \in \mathcal{F}$ and every $\pi \in G$ it is the case that $f \circ \pi \in \mathcal{F}$, where for every $x \in D \quad f \circ \pi(x) = f(\pi(x))$.

Automorphism group The *automorphism group* of family/code $\mathcal{F}$, denoted $\mathrm{Aut}(\mathcal{F})$, is the group of all transformations $\pi : [D] \to [D]$ such that if $c \in \mathcal{C}$ then $c \circ \pi \in \mathcal{C}$.

We are interested in families that are invariant under some well-studied groups (i.e., whose invariant groups contain some well-studied groups).

In particular we look at invariance under linear and affine groups, defined over a domain of size p^n. If $e|n$ then $\mathbb{F}_{p^n} \simeq \mathbb{F}_{p^e}^{n/e}$ under addition, and one can consider a series of nested linear/affine groups by viewing $\mathbb{F}_{p^n}$ as a vector space over $\mathbb{F}_{p^e}$ for all such e.

Most generally, the linear group acting on a domain of size p^n when the domain is seen as a vector space of dimension n/e over the subfield $\mathbb{F}_{p^e}$ is

$$\mathrm{GL}(n/e, p^e) = \{\pi : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n} | \pi(x) = \sum_{i=0}^{n/e-1} a_i x^{p^{ei}}, \forall a_i \in \mathbb{F}_{p^n}\},$$

(see for e.g. [28].) A linear function $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ $\pi(x) = \sum_{i=0}^{n/e-1} a_i x^{p^{ei}}$ is a permutation if and only if the a_i's involved above are linearly independent over $\mathbb{F}_{p^e}$ (For a formal proof see Chapters 4 and 7 of [80].)

Similarly,

$$\mathrm{AGL}(n/e, p^e) = \{\pi : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n} | \pi(x) = \sum_{i=0}^{n/e-1} a_i x^{p^{ei}} + b, \forall a_i, b \in \mathbb{F}_{p^n}\},$$

is the affine group acting on a domain of size p^n when the domain is seen as a vector space of dimension n/e over the subfield $\mathbb{F}_{p^e}$.

These definitions are equivalent to the typical definitions of GL and AGL involving matrix transformations. We choose these descriptions since they provide a concise representation of all the linear/affine groups over a fixed domain size p^n.

Observe that if $e_1 e_2 | n$ then

$$\mathrm{GL}(1, p^n) \subseteq \mathrm{GL}(n/(e_1 e_2), p^{e_1 e_2}) \subseteq \mathrm{GL}(n/e_1, p^{e_1}) \subseteq \mathrm{GL}(n, e)$$

and similarly

$$\mathrm{AGL}(1, p^n) \subseteq \mathrm{AGL}(n/(e_1e_2), 2^{e_1e_2}) \subseteq \mathrm{AGL}(n/e_1, 2^{e_1}) \subseteq \mathrm{AGL}(n, p).$$

In this work we focus on algebraic families that are invariant under the smallest such linear/affine subgroups for $p = 2$, namely on $\mathrm{GL}(1, 2^n)$ and $\mathrm{AGL}(1, 2^n)$, respectively.

In fact, the linear families that are invariant under the largest linear/affine group over a domain of size 2^n (i.e. $\mathrm{GL}(n, 2)$ and $\mathrm{AGL}(n, 2)$) are well studied families of codes, namely variants of Reed Muller codes [42, 72].

Definition 8 (Affine invariance) *A function $\pi : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is an* affine permutation *if there exist $\alpha \in \mathbb{F}_{2^n}^*$ and $\beta \in \mathbb{F}_{2^n}$ such that $\pi(x) = \alpha x + b$. A code $\mathcal{C} \subseteq \mathbb{F}_2^N$ is said to be* affine invariant *if the automorphism group of $\mathcal{C}$ contains the affine group* $\mathrm{AGL}(1, 2^n) = \{\pi(x) = \alpha x + b, \alpha \in \mathbb{F}_{2^n}^*, \beta \in \mathbb{F}_{2^n}\}$.

Linear invariance under the group $\mathrm{GL}(1, 2^n)$ is sometimes called cyclic invariance, due to the fact that the codes invariant under this group have a cyclic structure. We stick with this nomenclature hereafter.

Definition 9 (Cyclic/linear invariance) *A function $\pi : \mathbb{F}_{2^n}^* \to \mathbb{F}_{2^n}^*$ is a* cyclic permutation *if it is of the form $\pi(x) = \alpha x$ for $\alpha \in \mathbb{F}_{2^n}^*$.* [1] *A code $\mathcal{C} \subseteq \mathbb{F}_2^{N-1}$ is said to be* cyclic invariant *(or simply cyclic) if the automorphism group of $\mathcal{C}$ contains the linear (cyclic) group* $\mathrm{GL}(1, 2^n) = \{\pi(x) = \alpha x, \alpha \in \mathbb{F}_{2^n}^*.\}$.

[1] Note that this is a permutation of $\mathbb{F}_{2^n}^*$ if the elements of $\mathbb{F}_{2^n}^*$ are enumerated as $\langle \omega, \omega^2, \ldots, \omega^{N-1} \rangle$ where ω is a primitive element of $\mathbb{F}_{2^n}^*$.

Chapter 3

Description of Cyclic/Affine Invariant Linear Families

Cyclic/affine-invariant families admit nice representations as sets of univariate polynomials. This description will be useful in our results and we start by making this connection explicit. In the last part of this chapter we relate this description to the classical representation of cyclic codes, as ideals in univariate rings of polynomials.

Notation Let $N = 2^n$ and we view elements $c \in \mathbb{F}_2^N$ as functions $c : \mathbb{F}_N \to \mathbb{F}_2$, and similarly we view elements $c \in \mathbb{F}_2^{N-1}$ as functions $c : \mathbb{F}_N^* \to \mathbb{F}_2$. For $d \in \{1, \ldots, N-2\}$, let

$$\text{orb}(d) = \{d, 2d \pmod{N-1}, 4d \pmod{N-1}, \ldots, 2^{n-1}d \pmod{N-1}\}.$$

Also for a set $D \subseteq [N]$ let $\text{orb}(D) = \cup_{d \in D} \text{orb}(d)$. Notice that $2^i d = 2^j d \pmod{N-1}$ iff $2^j(2^{i-j}-1)d = 0 \pmod{2^n-1}$, and recall that $2^\ell - 1 \mid 2^n - 1$ iff $\ell \mid n$. Therefore, whenever n is prime $|\text{orb}(d)| = n$ for all d, and otherwise there exists d such that $|\text{orb}(d)| < n$. Let $|\text{orb}(d)| = \ell_d$. Let min-orb(d) denote the smallest integer in orb(d), and let

$$\mathcal{D} = \{\text{min-orb}(d) \mid d \in \{1, \ldots, N-2\}\} \cup \{N-1\}.$$

For $D \subseteq \mathcal{D}$ let

$$P_{N,D} = \{\alpha_0 + \sum_{d \in D} \alpha_d x^d \mid \alpha_d \in \mathbb{F}_N, \alpha_0, \alpha_{N-1} \in \{0,1\}\}, \text{ and}$$

$$P_{N-1,D} = \{\sum_{d \in D} \alpha_d x^d \mid \alpha_d \in \mathbb{F}_N, \alpha_{N-1} \in \{0,1\}\}.$$

Trace functions In general, for $\mathbb{K} = \mathbb{F}_{p^{t\ell}}$ and $\mathbb{F} = \mathbb{F}_{p^t}$ (with p prime) the *trace function* of $\mathbb{K}$ over $\mathbb{F}$ is $\text{Trace}_{\mathbb{K}/\mathbb{F}} : \mathbb{K} \to \mathbb{F}$

$$\text{Trace}_{\mathbb{K}/\mathbb{F}}(x) = x^{p^t} + x^{p^{2t}} + \ldots + x^{p^{(\ell-1)t}}.$$

In this work we only consider $p = 2$ and $\mathbb{F} = \mathbb{F}_2$, and from here on we will drop the subscript in the trace to refer to the binary $\text{Trace} : \mathbb{F}_{2^n} \to \mathbb{F}_2$ as $\text{Trace}(x) = x + x^2 + x^{2^2} + \ldots + x^{2^{n-1}}$.

In addition to the full trace, in this chapter we also make use of the *truncated trace* defined for each positive degree d as $\text{Trace}_d : \mathbb{F}_N \to \mathbb{F}_N$ by

$$\text{Trace}_d(x) = \sum_{i=0}^{|\text{orb}(d)|-1} x^{2^i}.$$

Properties of the Trace functions

1. The Trace functions are linear, i.e. $\text{Trace}(\alpha+\beta) = \text{Trace}(\alpha) + \text{Trace}(\beta)\ \forall \alpha, \beta \in \mathbb{F}_N$, and $\text{Trace}_d(\alpha + \beta) = \text{Trace}_d(\alpha) + \text{Trace}_d(\beta)\ \forall \alpha, \beta \in \mathbb{F}_N$.

2. Another useful property is that $\text{Trace}(\alpha) = \text{Trace}(\alpha^2)$ for all $\alpha \in \mathbb{F}_{2^n}$, which implies $\text{Trace}(\alpha x^2) = \text{Trace}(\beta x)$ for $\beta = \alpha^{2^{n-1}}$, and all $x \in \mathbb{F}_{2^n}$.

3. While Trace_d could take values outside $\mathbb{F}_2$, for $c \in \mathbb{F}_{2^{\ell_d}}$ the function $\text{Trace}_d(cx^d)$ is a map $\mathbb{F}_N \to \mathbb{F}_2$. Indeed, since $c^{2^{\ell_d}} = c$ and $d2^{\ell_d} = d \pmod{N-1}$, one can check that $(\text{Trace}_d(cx^d))^2 = (\sum_{i=0}^{\ell_d-1}(cx^d)^{2^i})^2 = c^2x^{2d} + c^4x^{4d} + \ldots + c^{2^{\ell_d}}x^{d2^{\ell_d}} = \sum_{i=0}^{\ell_d-1}(cx^d)^{2^i} = \text{Trace}_d(cx^d)$ for all $x \in \mathbb{F}_N$.

4. For $c \in \mathbb{F}_{2^{\ell_d}}$ it is the case that $\text{Trace}(cx^d) = \frac{n}{\ell_d} \cdot \text{Trace}_d(cx^d)$. Therefore,

$$\text{Trace}(cx^d) = \begin{cases} 0, & \text{if } \frac{n}{\ell_d} \text{ is even} \\ \text{Trace}_d(cx^d) & \text{if } \frac{n}{\ell_d} \text{ is odd} \end{cases}$$

The following lemma is the main property of the truncated trace that we exploit in the following section.

Lemma 10 *For every $d \in \mathcal{D}$ and $\beta \in \mathbb{F}_{2^{\ell_d}}$ there exists $\alpha \in \mathbb{F}_N$ such that* $\text{Trace}_d(\beta x^d) = \text{Trace}(\alpha x^d)$.

Proof: Let $L = 2^{\ell_d}$ and notice that for each $\alpha \in \mathbb{F}_N$

$$\begin{aligned} \text{Trace}(\alpha x^d) &= \sum_{i=0}^{n-1} (\alpha x^d)^{2^i} \\ &= \alpha x^d + \alpha^2 x^{2d} + \ldots + \alpha^{2^{\ell_d-1}} x^{d\, 2^{\ell_d-1}} + \alpha^{2^{\ell_d}} x^d + \ldots + \alpha^{2^{n-1}} x^{d\, 2^{\ell_d-1}} \\ &= \sum_{i=0}^{\ell_d-1} ((\alpha + \alpha^L + \alpha^{L^2} + \ldots + \alpha^{2^{L^{(n/\ell_d-1)}}}) x^d)^{2^i} \\ &= \text{Trace}_d(\text{Trace}_{\mathbb{F}_N/\mathbb{F}_L}(\alpha)\ x^d). \end{aligned}$$

It is a well-known fact that the map $\text{Trace}_{\mathbb{F}_N/\mathbb{F}_L} : \mathbb{F}_N \to \mathbb{F}_L$ is surjective. Indeed, the polynomial $\text{Trace}_{\mathbb{F}_N/\mathbb{F}_L}(x)$ has degree $2^{n-\ell_d}$ and it defines a linear map $\mathbb{F}_N \to \mathbb{F}_L$. Therefore its kernel has size at most $2^{n-\ell_d}$ and hence each element in $\mathbb{F}_L$ could be the image of a cosets of size at most $2^{n-\ell_d}$. Therefore, each element of $\mathbb{F}_L$ is the image of some element in $\mathbb{F}_N$.

Finally, this implies that for every $\beta \in \mathbb{F}_L$, there exists $\alpha \in \mathbb{F}_N$ such that $\text{Trace}_{\mathbb{F}_N/\mathbb{F}_L}(\alpha) = \beta$. It now follows that $\text{Trace}_d(\beta x^d)$ =$\text{Trace}_d(\text{Trace}_{\mathbb{F}_N/\mathbb{F}_L}(\alpha)x^d)$ =$\text{Trace}(\alpha x^d)$, which concludes the proof.

∎

3.1 Cyclic-invariant families

In this section we will show the following formal description of cyclic-invariant families.

Proposition 11 *[Description of Cyclic-Invariant Families] For every cyclic invariant family* $\mathcal{C} \subseteq \{\mathbb{F}_N^* \to \mathbb{F}_2\}$ *there exists a (unique) set* $D \subseteq \mathcal{D}$ *such that* $\mathcal{C} = \{\text{Trace}(p) | p \in P_{N-1,D}\}$. *Conversely, every arbitrary set* $D \subseteq \mathcal{D}$ *corresponds to a cyclic invariant family* $\mathcal{C} = \{\text{Trace}(p) | p \in P_{N,D}\}$. *We say that the set* D *uniquely* describes *the family* $\mathcal{C}$.

We begin by describing binary functions defined over $\mathbb{F}_{2^n}$: they are polynomials with monomial degrees closed under multiplication by 2.

Lemma 12 *For every word* $w : \mathbb{F}_N \to \mathbb{F}_2$ *(respectively* $w : \mathbb{F}_N^* \to \mathbb{F}_2$*) there is a unique polynomial* $p \in P_{N,\mathcal{D}}$ *(respectively* $p \in P_{N-1,\mathcal{D}}$*) such that* $w(x) = \text{Trace}(p(x))$.

Proof: Every function $w : \mathbb{F}_N \to \mathbb{F}_N$ is a polynomial, and thus we can write $w(x)$ uniquely as $\sum_{i=0}^{N-1} c_i x^i$ for some coefficients $c_i \in \mathbb{F}_N$. The condition that $w(\alpha) \in \{0,1\}$ for every $\alpha \in \mathbb{F}_N$, yields some constraints on c_i. In particular we have $w(\alpha)^2 = w(\alpha)$ for every $\alpha \in \mathbb{F}_N$ and so $w(x)^2 = w(x) \pmod{x^N - x}$. But $w(x)^2 = \sum_{i=0}^{N-1} c_i^2 x^{2i}$ and so, equating coefficients we have, $c_0^2 = c_0$, $c_{N-1}^2 = c_{N-1}$ (thus $c_0, c_{N-1} \in \mathbb{F}_2$), and moreover $c_{2i \pmod{N-1}} = c_i^2$ for every $i \in \{1, \ldots, N-2\}$.

In particular, for $d \in \mathcal{D} - \{N-1\}$ we must have $c_{d\ 2^{\ell_d} \pmod{N-1}} x^{d \cdot 2^{\ell_d}} = c_d x^d$ and thus $c_d = c_{d2^{\ell_d} \pmod{N-1}} = c_d^{2^{\ell_d}}$, implying $c_d \in \mathbb{F}_{2^{\ell_d}}$ for all $d \in \mathcal{D} - \{N-1\}$.

Thus,

$$
\begin{aligned}
w(x) &= \sum_{i=0}^{N-1} c_i x^i && (3.1)\\
&= c_0 x^0 + c_{N-1} x^{N-1} + \sum_{d \in \mathcal{D}-\{N-1\}} \mathrm{Trace}_d(c_d x^d) && (3.2)\\
&= \mathrm{Trace}(c_0 x^0) + \mathrm{Trace}(c_{N-1} x^{N-1}) + \sum_{d \in \mathcal{D}-\{N-1\}} \mathrm{Trace}_d(c_d x^d) && (3.3)\\
&= \mathrm{Trace}(c_0 x^0) + \mathrm{Trace}(c_{N-1} x^{N-1}) + \sum_{d \in \mathcal{D}-\{N-1\}} \mathrm{Trace}(c'_d x^d) && (3.4)\\
&= \mathrm{Trace}\Big(\sum_{d \in \mathcal{D} \cup \{0\}\}} c'_d x^d \Big), && (3.5)
\end{aligned}
$$

for some $c'_d \in \mathbb{F}_N$ $\forall d \in \mathcal{D} - \{N-1\}$ and $c'_0 = c_0$, $c'_{N-1} = c_{N-1}$. Equation 3.2 follows from writing the set $\{0, \ldots, N-1\}$ (the set of degrees of x) as $\{0, N-1\} \cup (\cup_{d \in \mathcal{D}-\{N-1\}} \mathrm{orb}(d))$, where the sets $\mathrm{orb}(d)$ are disjoint. Equation 3.3 follows by noticing that $c_0 x_0$ and $c_{N-1} x^{N-1}$ belong to $\mathbb{F}_2$ for all $x \in \mathbb{F}_N$, and thus $\mathrm{Trace}(c_i x_i) = c_i x_i$ ($i \in \{0, N-1\}$). Equation 3.4 is implied by Lemma 10, and finally equation 3.5 follows by the linearity of the trace function.

This concludes the proof for the case of functions mapping $\mathbb{F}_N$ to $\mathbb{F}_2$. For the case of functions $w : \mathbb{F}_N^* \to \mathbb{F}_2$, the proof is similar except we start by writing w uniquely as $\sum_{i=1}^{N-1} c_i x^i$ (and so x^{N-1} plays the role of the constant function 1). ∎

Lemma 13 *Suppose $\mathcal{C} \subseteq \{\mathbb{F}_N^* \to \mathbb{F}_2\}$ is a cyclic invariant code containing the word $w = \mathrm{Trace}(p(x))$ for $p \in P_{N-1,\mathcal{D}}$. Then, for every monomial x^e in the support of p, the function $\mathrm{Trace}(x^e)$ is in $\mathcal{C}$. Moreover, if $e \neq N-1$ then for every $\beta \in \mathbb{F}_N$, $\mathrm{Trace}(\beta x^e) \in \mathcal{C}$.*

Proof: The proof is essentially from [72]. Since their proof is a bit more complex (and considers a more general class of functions and non-prime n), we include the proof in our setting for completeness.

Let $p(x) = \sum_{d \in \mathcal{D}} c_d x^d$, where $c_{N-1} \in \{0, 1\}$ and let $w(x) = \mathrm{Trace}(p(x))$. Fix e in the support of p. We first consider the case $e \neq N-1$. We wish to show that

Trace(βx^e) is in $\mathcal{C}$ for every $\beta \in \mathbb{F}_N$. Note that for every $\alpha \in \mathbb{F}_N^*$, $w(\alpha x)$ is in $\mathcal{C}$ (by the cyclic invariance). Furthermore, the function $\sum_{\alpha \in \mathbb{F}_N^*} \text{Trace}(\alpha^{-e}) w(\alpha x)$ is also in $\mathcal{C}$ (by linearity). But as we show below this term is simply Trace$(c_e x^e)$.

$$\begin{aligned}
\sum_{\alpha \in \mathbb{F}_N^*} \text{Trace}(\alpha^{-e}) w(\alpha x) &= \sum_{\alpha \in \mathbb{F}_N^*} \text{Trace}(\alpha^{-e}) \text{Trace}(p(\alpha x)) \\
&= \sum_{\alpha \in \mathbb{F}_N^*} \left(\sum_{j=0}^{n-1} \alpha^{-e \cdot 2^j} \right) \left(\sum_{i=0}^{n-1} \sum_{d \in \mathcal{D}} c_d^{2^i} \alpha^{d \cdot 2^i} x^{d \cdot 2^i} \right) \\
&= \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} \sum_{d \in \mathcal{D}} c_d^{2^i} x^{d \cdot 2^i} \sum_{\alpha \in \mathbb{F}_N^*} \alpha^{d \cdot 2^i - e \cdot 2^j}
\end{aligned}$$

Recall that $\sum_{\alpha \in \mathbb{F}_N^*} \alpha^t$ is 0 if $t \neq 0 (\text{mod} N-1)$ and 1 if $t = 0$. So we conclude that the innermost sum is non-zero only if $d \cdot 2^i = e \cdot 2^j (\text{mod} N-1)$ which in turn happens only when $d = e$ and $j = i$ (since both $d, e \in \mathcal{D} - \{N-1\}$). We conclude $\sum_{\alpha \in \mathbb{F}_N^*} \text{Trace}(\alpha^{-e}) w(\alpha x) = \sum_{i=0}^{n-1} c_e^{2^i} x^{e \cdot 2^i} = \text{Trace}(c_e x^e)$.

Finally, we need to show that Trace(βx^e) is also in $\mathcal{C}$. To see this, consider the set $S \subseteq \mathbb{F}_N$ defined as $S = \{\gamma | \text{Trace}(c_e \gamma x^e) \in \mathcal{C}\}$. We know S is non-empty (since $1 \in S$), S is closed under addition, and if $\beta \in S$, then so is $\beta \cdot \zeta^e$ for every $\zeta \in \mathbb{F}_N$. Thus, in particular, S contains the set $T = \{p(\omega^e) | p \in \mathbb{F}_2[x]\}$ where ω is the multiplicative generator of $\mathbb{F}_N^*$. T is again closed under addition and also under multiplication and so is a subfield of $\mathbb{F}_N$. Finally it includes ω^e as an element and so $T = \mathbb{F}_N$ (the only strict subfield of $\mathbb{F}_N$ is $\mathbb{F}_2$ which does not contain ω^e for $e \in \mathcal{D}$). We thus conclude that both S and T equal $\mathbb{F}_N$ and so for every $\beta \in \mathbb{F}_N$, Trace$(\beta x_e) \in \mathcal{C}$.

It now only remains to consider the case $e = N-1$. By hypothesis $c_{N-1} = 1$ in this case. Thus we consider the simpler function $\sum_{\alpha \in \mathbb{F}_N^*} w(\alpha x)$ which is also in $\mathcal{C}$. It can be argued as above that this function equals $c_{N-1} x^{N-1} = x^{N-1} = \text{Trace}(x^{N-1})$. This concludes the analysis. ∎

Observation 14 *Lemma 13 also holds for affine invariant families. Indeed, we only used the facts that $\mathcal{C}$ is linear, and that $w(\alpha x)$ is in $\mathcal{C}$ for every $\alpha \in \mathbb{F}_N$.*

Proof of Proposition 11: Let D be the set of all integers $d \in \mathcal{D}$ such that there is some polynomial $p \in P_{N-1,\mathcal{D}}$ with positive support on the monomial x^d such that $\text{Trace}(p) \in \mathcal{C}$. By Lemma 13 we have that every function $\text{Trace}(\beta x^d) \in \mathcal{C}$ for every $\beta \in \mathbb{F}_N$, if $d \notin \{0, N-1\}$.

Conversely, any arbitrary set of degrees $D \in \mathcal{D}$ determines the code $\mathcal{C} = \{\text{Trace}(p(x)) \mid p \in P_{N-1,D}\}$. It is easy to notice that this code is cyclic invariant, since $\text{Trace}(p(\alpha x)) \in \mathcal{C}, \ \forall \alpha \in \mathbb{F}_N$.

∎

3.2 Affine-invariant families

We can now explicitly describe affine invariant families. These families are also characterized by a set of degrees, but this set has an additional closure property.

Shadow of degree For non-negative integers d and e we say e is in the *shadow* of d, denoted $e \prec d$, if in the binary representations $d = \sum_i d_i 2^i$ and $e = \sum_i e_i 2^i$ with $d_i, e_i \in \{0,1\}$, it is the case that $e_i \leq d_i$ for every i. The *shadow* of a positive integer d, denote Δd, is the set $\{e | e \prec d\}$. The shadow of a set of non-negative integers D, denoted ΔD, is the set $\{e | e \prec d \text{ for some } d \in D\}$. For $D \subseteq \mathcal{D}$ we say that D is *shadow-closed* if $(\Delta D) \cap \mathcal{D} = D$.

We will also need the following widely known fact.

Fact 15 (Lucas identity) *For* $0 \leq d, e \leq 2^n - 2$ *and* $d = \sum_{i=0}^{n-1} d_i 2^i, \ e = \sum_{i=0}^{n-1} e_i 2^i$

$$\binom{d}{e} \pmod 2 = \binom{d_0}{e_0}\binom{d_1}{e_1}\ldots\binom{d_{n-1}}{e_{n-1}}.$$

Hence

$$\binom{d}{e} \pmod 2 = \begin{cases} 1 & \text{if } e \prec d \\ 0 & \text{if } e \not\prec d. \end{cases}$$

In this section we will show the following structural statement.

Proposition 16 *[Description of Affine Invariant Families] For every affine invariant family* $\mathcal{C} \subseteq \{\mathbb{F}_N \to \mathbb{F}_2\}$ *there exists a (unique) shadow-closed set* $D \subseteq \mathcal{D}$ *such that* $\mathcal{C} = \{\text{Trace}(p) | p \in P_{N,D}\}$. *Conversely, every shadow-closed set* $D \subseteq \mathcal{D}$ *corresponds to an affine invariant family* $\mathcal{C} = \{\text{Trace}(p) | p \in P_{N,D}\}$. *We say that the set* D *uniquely* describes *the family* $\mathcal{C}$.

We start with following lemma.

Lemma 17 *If* $\mathcal{C}$ *is an affine-invariant code,* $\text{Trace}(x^d) \in \mathcal{C}$ *and* $e \prec d$ *then* $\text{Trace}(x^e) \in \mathcal{C}$.

Proof: Since $\text{Trace}(x^d) \in \mathcal{C}$ and $\mathcal{C}$ is affine invariant, then $\text{Trace}((x+1)^d) \in \mathcal{C}$. But by Lucas' formula $(x+1)^d = \sum_{e \leq d} \binom{d}{e} x^e = \sum_{e \prec d} x^e$. Therefore, $\text{Trace}(\sum_{e \prec d} x^e) \in \mathcal{C}$ and by Observation 14 $\text{Trace}(x^e) \in \mathcal{C}$. ∎

Proof of Proposition 16: The first part of the proof is similar to the proof of the cyclic case. Let D be the set of all integers $d \in \mathcal{D}$ such that there is some polynomial $p \in P_{N,\mathcal{D}}$ with positive support on the monomial x^d such that $\text{Trace}(p) \in \mathcal{C}$. By Observation 14 we have that every function $\text{Trace}(\beta x^d) \in \mathcal{C}$ for every $\beta \in \mathbb{F}_N$, if $d \notin \{0, N-1\}$. Furthermore since $\text{Trace}((x+1)^d)$ is also in $\mathcal{C}$, it follows that the constant function 1 is also in $\mathcal{C}$. We conclude that the traces of all the polynomials in $P_{N,D}$ are in $\mathcal{C}$.

By Lemma 17 it follows that for all $e \in \Delta D$ $\text{Trace}(x^e) \in \mathcal{C}$. Thus it must be that $D = \Delta D$.

Conversely, we verify that for any shadow-closed set $D \subset \mathcal{D}$, the family

$$\mathcal{C} = \{\text{Trace}(p(x)) \mid p \in P_{N,D}\}$$

is affine invariant. Indeed, for $p(x) = \sum_{d\in D} p_d x^d$ and $\mathrm{Trace}(p(x)) \in \mathcal{C}$, it follows that

$$\begin{aligned}
\mathrm{Trace}(p(ax+b)) &= \mathrm{Trace}\left(\sum_{d\in D} p_d(ax+b)^i\right) = \mathrm{Trace}\left(\sum_{d\in D} p_d \sum_{0\le e\le d} \binom{d}{e} a^e\ b^{d-e} x^e\right) \\
&= \mathrm{Trace}\left(\sum_{d\in D} p_d \sum_{0\prec e\prec d} a^e\ b^{d-e} x^e\right) \\
&= \mathrm{Trace}\left(\sum_{e}\Big(\sum_{d\in D\ :\ e\prec d} p_d a^e\ b^{d-e}\Big)x^e\right) \\
&= \sum_{e} \mathrm{Trace}\left(\sum_{d\in D\ :\ e\prec d} p_d a^e\ b^{d-e} x^e\right).
\end{aligned}$$

In the above we used Lucas' formula from Fact 15 and the linearity of the trace function. Now recall that $D = \Delta D \cap \mathcal{D}$ and therefore the terms $\mathrm{Trace}(\alpha x^e)$, with $e \in D$ belong to $\mathcal{C}$. Finally, examining the remaining terms, these are monomials of degrees e such that $e \prec d$ for some $d \in D$ and $e \notin D$. Therefore, $e \in \mathrm{orb}(e')$ for some $e' \in D$, and $e = e'2^k$ for some k. But $\mathrm{Trace}(\alpha x^e) = \mathrm{Trace}(\alpha x^{2^k e'}) = \mathrm{Trace}(\alpha^{2^{n-k}} x^{2^n e'}) = \mathrm{Trace}(\alpha^{2^{n-k}} x^{e'})$ which belongs to $\mathcal{C}$.

∎

3.3 Cyclic/affine codes as polynomial ideals

We mention that a common description of cyclic/affine invariant codes is as polynomial ideals. While we do not use this description, we comment on these alternate definitions here in order to make the connection between our results and the literature on these well-studied codes. For more detail see for example [81].

Alternate description of cyclic codes A vector in $v \in \mathbb{F}_2^{N-1}$ can be abstracted as a polynomial $p_v(x) \in \mathbb{F}_2[x]/\langle x^{N-1} - 1\rangle$, where $p_v(x) = \sum_{i=0}^{N-2} v_i x^i$. A cyclic code $\mathcal{C}$ can be represented as an ideal in $\mathbb{F}_2[x]/\langle x^{N-1} - 1\rangle$ generated by a polynomial generator $g(x)$. That is, a polynomial p_v represents a codeword if and only if $p_v =$

$gh \bmod x^{N-1} - 1$, for some $h \in \mathbb{F}_2[x]$. The roots of $g(x)$ in $\mathbb{F}_N$ uniquely describe the code, and bear a one-to-one relationship to the degrees D that characterize the dual code $\mathcal{C}^\perp$ in the sense described in Proposition 11. Namely, the roots of g are *exactly* the field elements w^d, where w is a fixed primitive element of $\mathbb{F}_N$ and $d \in \text{orb}(D)$.

Alternate description of affine invariant codes A vector $v \in \mathbb{F}_2^N$ can be abstracted as a polynomial $p_v(x) \in \mathbb{F}_2[x]/\langle x^N - x\rangle$, where $p_v(x) = \sum_{i=0}^{N-1} v_i x^i$. An affine invariant code $\mathcal{C}$ is the extension of a cyclic code by a parity bit. It is also the ideal generated by a polynomial $a \in \mathbb{F}_2[x]/\langle x^N - x\rangle$ whose roots are in a one-to-one correspondence to the set D that characterizes the affine code $\mathcal{C}^\perp$. Namely, the roots of $a(x)$ are $\{0\}$ and the elements w^d, where $d \in \text{orb}(D)$. The cyclic code whose extension is the affine invariant code characterized by the set of degrees D has the same set of roots, except possibly 0.

3.4 Bibliographic notes

Affine codes (and cyclic codes whose extensions are affine invariant) received a lot of attention in the 1970s and we mention here a few notable results.

Limitations to building good codes In [64] Kasami shows that affine invariant codes for which the distance is linear with respect to growing block length must have vanishing rate. McEliece [83] however prove that there are arbitrarily long non-linear good codes (i.e. codes of constant rate and relative distance).

Automorphism groups and structure of affine invariant codes Structural properties of affine invariant codes were initially described in [65]. Delsarte [42] studied codes that are invariant under subgroups of $\text{AGL}(n, p)$. These families include in particular affine codes of length p^n (since these are invariant under $\text{AGL}(1, p^n)$.) He proved that the only linear codes over an alphabet $\mathbb{F}_p$ invariant under $\text{AGL}(n, p)$ are the Reed Muller codes (this is also proved in [72]). Affine invariant codes could

have larger automorphism groups than just $\mathrm{AGL}(1, p^n)$. Berger [27] showed that the automorphism groups of affine invariant codes are in fact subgroups of $\mathrm{AGL}(p, n)$. More recently Berger and Charpin [28] completely characterized the automorphism groups of many affine invariant families, by explicitly describing their elements as special types of polynomials. Interestingly, they were able to fully characterize the automorphism group of BCH codes, and identified exceptions to the common case when the automorphism group is just $\mathrm{AGL}(1, p^n)$.

Chapter 4

2-Transitivity is Insufficient for Local Testability

In this section we start our study of sufficient conditions for testing, by presenting a negative result. We show that families that are invariant under a 2-transitive group and whose duals contain a small-weight function are not necessarily testable. This disproves a conjecture made by Alon et al. in [7] (denoted here as the AKKLR conjecture.) We start with the definition of 2-transitivity.

Definition 18 (2-Transitivity) *A group G of permutations mapping D to D is* 2-transitive *if for every $x, x', y, y' \in D$ such that $x \neq y$ and $x' \neq y'$, there exists $\pi \in G$ such that $\pi(x) = x'$ and $\pi(y) = y'$.*

Abusing notation slightly, we say that $\mathcal{F}$ is 2-transitive if $\mathrm{Aut}(\mathcal{F})$ *is 2-transitive.*

By now, it has become folklore that having a low-weight function in the dual code is a trivial necessary condition for testability. Indeed, the following result of Ben-Sasson et al. [23] formalizes the fact that any test for a linear property reduces to picking dual functions and checking a linear condition. If the dual family contains only large weight function, no local test exists.

Theorem 19 ([23, Theorem 3.3]) *Let $\mathcal{F} = \{\mathcal{F}_n\}_n$ be a linear property that is k-locally testable. Then $\mathcal{F}$ is k-locally testable by a non-adaptive, perfect tester. Specifically, if $\mathcal{F}_n$ is $(k, \epsilon_1, \epsilon_2, \delta)$-locally testable, then $\mathcal{F}_n$ is $(k, 0, \epsilon_2 - \epsilon_1, \delta)$-locally testable by a non-adaptive, perfect tester. Moreover, if f is the given word, the test checks if $\langle f, v \rangle = 0$ for $v \in \mathcal{F}^\perp$, where v has support of size at most k.*

Theorem 19 will be very useful in presenting our counterexample to the AKKLR conjecture. It implies that in order to rule out any test for linear properties (even adaptive or 2-sided error tests) it is enough to rule out non-adaptive, perfect testers.

4.1 The conjecture

We now formally state the AKKLR-conjecture.

Conjecture 20 ([7]) *For every $d \in \mathbb{N}$, there exists $k = k(d) < \infty$ such that the following holds: Let $\mathcal{F} = \{\mathcal{F}_n\}_n$ be an ensemble of properties such that for every n,*

1. *$\mathcal{F}_n^\perp$ has a non-zero function of weight at most d, and*
2. *$\mathcal{F}_n$ is 2-transitive.*

Then $\mathcal{F}$ is k-locally testable.

We disprove this conjecture in the following formal statement.

Theorem 21 *For every $k \leq \infty$, there is an ensemble of domains $\{D_n\}_n$ and an ensemble of properties $\mathcal{F} = \{\mathcal{F}_n\}_n$ such that the following hold:*

1. *For every n, $\mathcal{F}_n^\perp$ has a non-zero function of weight at most 8.*
2. *For every n, $\mathcal{F}_n$ is 2-transitive.*
3. *$\mathcal{F}$ is not k-locally testable.*

As pointed out earlier, we plan to prove this theorem by ruling out a restrictive class of tests that are non-adaptive and perfect and then using Theorem 19. However to use that theorem we need to ensure that our property is linear. The following theorem gives the more technical result that we show.

Theorem 22 *For every $k \leq \infty$, there is an ensemble of domains $\{D_n\}_n$ and an ensemble of properties $\mathcal{F} = \{\mathcal{F}_n\}_n$ such that the following hold:*

1. *$\mathcal{F}$ is linear.*
2. *For every n, $\mathcal{F}_n^{\perp}$ has a non-zero function of weight at most 8.*
3. *For every n, $\mathcal{F}_n$ is 2-transitive.*
4. *$\mathcal{F}$ is not k-locally testable by a non-adaptive, perfect tester.*

Note that Theorem 21 follows immediately by combining Theorem 22 and Theorem 19. In the rest of this chapter we focus on Theorem 22.

4.2 The counterexample, basic properties, and proof ideas

Our counterexample family comes from a the broad class of affine invariant properties introduced by Kaufman and Sudan [72]. Indeed, as we prove next, every affine-invariant family is 2-transitive. This observation reveals a large collection of 2-transitive families that can be further explored in the context of the conjecture.

Proposition 23 *For every field $\mathbb{K}$ and integer n, the set of affine permutations from $\mathbb{K}^n \to \mathbb{K}^n$ is 2-transitive.*

Proof: It suffices to prove that for every $x_1, x_2, y_1, y_2 \in \mathbb{K}^n$ with $x_1 \neq x_2$ and $y_1 \neq y_2$, there exists an affine permutation $A : \mathbb{K}^n \to \mathbb{K}^n$ such that $A(x_1) = y_1$ and $A(x_2) = y_2$. Let A be given by $A(x) = Mx + b$ where $M \in \mathbb{K}^{n \times n}$ and $b \in \mathbb{K}^n$. The condition that it be a permutation implies M should be non-singular, and satisfy $M(x_1 - x_2) = y_1 - y_2$, while $b = y_1 - Mx_1$. It is easy to see that a non-singular M satisfying $M(x_1 - x_2) = y_1 - y_2$ exists. ∎

Let $\mathbb{K} = \mathbb{F}_{2^n}$ and $N = 2^n$.

Recall from Chapter 3 the definition of the set $\mathcal{D}_n = \{\text{min-orb}(d) \mid d \in \{1, \ldots, N-2\}\} \cup \{N-1\}$. Let $D_{k,n}$ be the set of degrees

$$D_{k,n} = \{2^i + 1 | 0 \leq i \leq k\} \cup \{1\}, \text{ and } D'_{k,n} = \mathcal{D} \cap D_{k,n}.$$

Recall also that $P_{N,D'_{k,n}}$ is the set of polynomials supported on the monomials with degrees in $D'_{k,n}$ and coefficients in $\mathbb{F}_N$.

The counterexample The family that we focus on in this chapter is

$$\mathcal{F}^*_{k,n} = \left\{ \text{Trace}(p(x)) \mid p(x) \in P_{N,D'_{k,n}} \right\}.$$

Proposition 24 *$\mathcal{F}^*_{k,n}$ is linear, and affine-invariant.*

Proof: The proof is immediate from the description of affine-invariant families shown in Chapter 3. ∎

As we will prove later, $\mathcal{F}^*_{k,n}$ is also a strict subfamily of the common Reed-Muller codes of order 2 (RM$(2, n)$), for $k \leq \lfloor n/2 \rfloor$. The dual of $\mathcal{F}^*_{k,n}$ therefore contains the dual of RM$(2, n)$. Since RM$(2, n)$ has small weight codewords (in fact codewords of weight 8) the small dual distance of $\mathcal{F}^*_{k,n}$ comes for free.

The main insight of the proof is the fact that a small number of queries cannot distinguish $\mathcal{F}^*_{k,n}$ from RM$(2, n)$. In other words, a codeword of RM$(2, n)$ is accepted with probability 1 by tests that only employ a number of queries $t < k - 1$. Since the

distance between codewords of $\mathrm{RM}(2,n)$ is large (i.e. $\mathrm{RM}(2,n)$ is a code of constant distance), it follows that there is a word w such that $w \in \mathrm{RM}(2,n) - \mathcal{F}^*_{k,n}$ that is accepted by any t-query test, a contradiction. Hence, for every k there exists $\mathcal{F}^*_{k,n}$, which, even though it does contain functions of low weight, it is not testable with k queries. As a consequence, for $k = \omega(1)$ (think of t as, say $\approx \log n$), $\mathcal{F}^*_{k,n}$ cannot be tested with tests of constant locality.

We establish a basic property of our counterexample family, which will be useful in arguing that $\mathcal{F}^*_{k,n}$ is distinct from $\mathrm{RM}(2,n)$.

Lemma 25 *For every $t < n-1$, $\mathcal{F}^*_{t,n} \subseteq \mathcal{F}^*_{t+1,n}$. If $t < \lfloor n/2 \rfloor$ then $\mathcal{F}^*_{t,n} \subsetneq \mathcal{F}^*_{t+1,n}$.*

Proof: The proof of the first containment follows from the definition. The second part can be derived from, for instance, [81, Chapter 9, Theorem 7]. For the sake of completeness we include a proof here.

We claim that for distinct $1 \leq i, j < n/2$, the functions $\mathrm{Tr}(x^{2^i+1})$ and $\mathrm{Tr}(x^{2^j+1})$ have disjoint support, when viewed as polynomials of degree at most $2^n - 1$. This suffices, since it implies that the function $\mathrm{Tr}(x^{2^t+1}) \notin \mathcal{F}^*_{t-1,n}$. We prove the claim below.

Note that the function $\mathrm{Tr}(x^{2^i+1})$ has support on the monomials x^d for $d = 2^{i+\ell} + 2^\ell (\mathrm{mod} 2^n - 1)$ and similarly $\mathrm{Tr}(x^{2^j+1})$ is supported by the monomials x^d for $d = 2^{j+m} + 2^m (\mathrm{mod} 2^n - 1)$ (here we use the phrase mod non-conventionally to refer to the unique integer in $[2^n - 1]$ from the equivalence class). Suppose for contradiction that $2^{i+\ell} + 2^\ell = 2^{j+m} + 2^m (\mathrm{mod} 2^n - 1)$. Then, by multiplying both sides by $2^{s-\ell}$ and reducing modulo $2^n - 1$, we see that we have $2^i + 1 = 2^{j+m'} + 2^{m'} (\mathrm{mod} 2^n - 1)$ (where $m' = m - \ell$). Now we consider two cases: If $m' \leq n/2$, then the unique integer between 1 and $2^n - 1$ equal to $2^{j+m'} + 2^{m'} (\mathrm{mod} 2^n - 1)$ is $2^{j+m'} + 2^{m'}$. But then $2^{j+m'} + 2^{m'} \neq 2^i + 1$ unless $m' = 0$ and $i = j$ (violating distinctness of i and j). In the other case, if $m' > n/2$, then the unique integer in $[2^n - 1]$ equal to $2^{m'} + 2^{j+m'} > 2^{n/2} > 2^i + 1$. So again the modular equivalence can not hold. This proves the claim, and thus the lemma. ∎

Linearized polynomials On a technical level, the main proof uses properties of so-called *linearized polynomials* and employs simple algebraic arguments.

A linearized polynomial of degree 2^d is a mapping $L : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ defined by

$$L(x) = \sum_{i=0}^{d} l_i x^{2^i}.$$

The Trace function and the Trace_d function are particular examples of linearized polynomials. Notice that $L(0) = 0$ and if $\alpha, \beta \in \mathbb{F}_N$ are roots of L, then $\alpha + \beta$ is a root of L. It follows that the kernel of L is a subspace of dimension at most d, an observation that will be useful in the proof.

4.3 Proof of main theorem

4.3.1 Reed-Muller of Order d families

As already discussed, the counterexample family defined above is included in RM codes of order 2. This is not immediately obvious from the usual definition of RM codes as low degree polynomials.

Definition 26

$$\text{RM}(d,n) = \left\{ f : \mathbb{F}_2^n \to \mathbb{F}_2 | f = \sum a_{d_1,d_2,\ldots,d_n} x_1^{d_1} x_2^{d_2} \ldots x_n^{d_n}, a_i \in \mathbb{F}_2, \text{ with } \sum d_i \leq d \right\}.$$

Notice that it is enough to consider $d_i \in \{0,1\}$, since over $\mathbb{F}_2$ $x^i = x$ for $i \geq 2$. In this section we give an alternative definition and first show how that is equivalent to Definition 26.

Definition 27

$$\mathcal{C}(d,n) = \left\{ f : \mathbb{F}_{2^n} \to \mathbb{F}_2 | f(x) = \text{Trace}(\sum b_i x^{d_i}), b_i \in F_{2^n}, \text{ with } \text{bwt}(d_i) \leq d. \right\}$$

Lemma 28 *Definition 26 and 27 describe the same family.*

Proof: Let w be a primitive element of $\mathbb{F}_{2^n}$ and consider the bijection $\pi : \mathbb{F}_2^n \to \mathbb{F}_{2^n}$ given by $\pi(x_1, \ldots, x_n) = \sum_{i=0}^{n} x_i w^i$.

To show that $\mathrm{RM}(d,n) \subseteq \mathcal{C}(d,n)$ it is enough to show that every monomial $x_1^{d_1} x_2^{d_2} \ldots x_n^{d_n}$ can be written as a univariate polynomial from $\mathcal{C}(d,n)$, and then use the linearity property of the two families.

We first show that for every i, there exists $\alpha_i \in \mathbb{F}_{2^n}$ such that $\mathrm{Trace}(\alpha_i x) = x_i$, where $x = \pi(x_1, \ldots, x_n)$. Indeed, since $\mathbb{F}_2^n \simeq \mathbb{F}_{2^n}$, there exists a bijection between the set of linear transformations mapping $\mathbb{F}_2^n \to \mathbb{F}_2$ (i.e. $\mathcal{L}_1 = \{L_A(x) = A \cdot x,\ A \in \mathbb{F}_2^{n \times n}\}$) to the set of linear functions mapping $\mathbb{F}_{2^n} \to \mathbb{F}_2$ (i.e. $\mathcal{L}_2 = \{l_\alpha(x) = \mathrm{Trace}(\alpha x), \alpha \in \mathbb{F}_{2^n}\}$.) The map $(x_1, \ldots, x_n) \mapsto x_i$ is a linear transformation in $\mathcal{L}_1$, and hence it follows that there exists $\alpha_i \in \mathbb{F}_{2^n}$ such that $x_i = \mathrm{Trace}(\alpha_i x)$ and so $\mathrm{RM}(1,n) \subseteq \mathcal{C}(1,n)$.

Using this observation, we next show that the polynomial $x_1^{d_1} x_2^{d_2}$ can be expressed as a polynomial of the form $\mathrm{Trace}(p(x))$ (with $x = \pi(x_1, x_2, \ldots, x_n)$) such that each monomial degree in p has binary weight at most 2.

$$
\begin{aligned}
x_1^{d_1} x_2^{d_2} &= \mathrm{Trace}(\alpha_1 x)\mathrm{Trace}(\alpha_2 x) = \Big(\sum_{i=0}^{n-1} (\alpha_1 x)^{2^i}\Big)\Big(\sum_{j=0}^{n-1} (\alpha_2 x)^{2^j}\Big) = \sum_{i,j} \alpha_1^{2^i} \alpha_2^{2^j} x^{2^i + 2^j} \\
&= \sum_{i=0}^{n-1} \sum_{j \geq i} (\alpha_1^{2^i} \alpha_2^{2^j} + \alpha_1^{2^j} \alpha_2^{2^i}) x^{2^i + 2^j} = \sum_{i} \sum_{j \geq i} (\alpha_1 \alpha_2^{2^{j-i}} + \alpha_1^{2^{j-i}} \alpha_2)^{2^i} x^{(1 + 2^{j-i}) 2^i} \\
&= \sum_{d=0}^{n-1} \sum_{i} \left((\alpha_1 \alpha_2^{2^d} + \alpha_1^{2^d} \alpha_2) x^{1+2^d} \right)^{2^i} = \sum_{d=0}^{n-1} \mathrm{Trace}((\alpha_1 \alpha_2^{2^d} + \alpha_1^{2^d} \alpha_2) x^{1+2^d}) \\
&= \mathrm{Trace} \sum_{d=0}^{n-1} (\alpha_1 \alpha_2^{2^d} + \alpha_1^{2^d} \alpha_2) x^{1+2^d}.
\end{aligned}
$$

Hence, $\mathrm{RM}(2,n) \subseteq \mathcal{C}(2,n)$. A simple inductive argument implies $x_1^{d_1} \ldots x_n^{d_n} \in \mathcal{C}(d,n)$, for all d.

To show that $\mathcal{C}(d,n) \subseteq \mathrm{RM}(d,n)$ it is enough to prove that $\mathrm{Trace}(\alpha x^{\sum_{i=0}^{n-1} d_i 2^i})$, with $\sum d_i \leq d$, can be expressed as a multivariate polynomial from $\mathrm{RM}(d,n)$. Let $L = \{i | d_i \neq 0\} = \{l_1, \ldots, l_{|L|}\}$, thus $|L| \leq d$. Let $(x_1, \ldots, x_n) = \pi^{-1}(x)$ and let $I = \{i | x_i \neq$

$0\}$.

It follows that

$$\begin{aligned}
\text{Trace}(\alpha x^{\sum_{i=0}^{n-1} d_i 2^i}) &= \sum_{j=0}^{n-1} \alpha^{2^j} \left(\prod_{l \in L} (\sum_{i \in I} x_i w^i)^{2^l} \right)^{2^j} \\
&= \sum_{j=0}^{n-1} \alpha^{2^j} \left(\sum_{i_1,\ldots,i_{|L|} \in I} x_{i_1} \ldots x_{i_{|L|}} w^{\sum_{t=1}^{|L|} i_t 2^{l_t}} \right)^{2^j} \\
&= \sum_{i_1,i_2,\ldots,i_{|L|} \in I} \left(x_{i_1} \cdots x_{i_{|L|}} \right) \sum_{j=0}^{n-1} \left(\alpha \sum_{i_1,\ldots,i_{|L|} \in I} w^{\sum_{t=1}^{|L|} i_t 2^{l_t}} \right)^{2^j} \\
&= \sum_{i_1,i_2,\ldots,i_{|L|} \in I} \left(x_{i_1} \cdots x_{i_{|L|}} \right) \text{Trace}(p(w)) \\
&\in \text{RM}(d,n),
\end{aligned}$$

where $p(x) = \alpha \sum_{i_1,\ldots,i_{|L|} \in I} x^{\sum_{t=1}^{|L|} i_t 2^{l_t}}$, and recall that $x_i \in \mathbb{F}_2, \ \forall i$. ∎

The following proposition is an immediate consequence of Lemma 28

Proposition 29 *For every $t < n$, $\mathcal{F}^*_{t,n} \subseteq \text{RM}(2,n)$.*

We will further need the following notation. For points $x_0, x_1, \ldots, x_\ell \in \mathbb{F}_{2^n}$, define $A(x_0; x_1, \ldots, x_\ell)$ to be the affine subspace generated by $x_1, \ldots, x_\ell$ through x_0. I.e., $A(x_0; x_1, \ldots, x_\ell) = \{x_0 + \sum_{i=1}^{\ell} a_i x_i | a_1, \ldots, a_\ell \in \mathbb{F}_2\}$.

We now note the fact that $\text{RM}(2,n)$ has weight 8 functions in its dual.

Proposition 30 *For $n \geq 3$, $\text{RM}(2,n)^\perp$ contains weight 8 functions.*

Proof: We will show that all $f \in \mathcal{C}(2,n) = \text{RM}(2,n)$ satisfy the '$\text{RM}(2,n)$' constraint $\sum_{z \in A(x_0;x_1,x_2,x_3)} f(z) = 0$ for every $x_0, x_1, x_2, x_3 \in \mathbb{F}_{2^n}$.

Using the linearity of the Trace function ($\text{Trace}(x+y) = \text{Trace}(x) + \text{Trace}(y)$) we note that it suffices to show that every $f \in \{\text{Trace}(\beta), \text{Trace}(\beta_0 x), \text{Trace}(\beta_1 x^{2^1+1}, \ldots, \text{Trace}(\beta_k x^{2^k+1})\}$ satisfies the above constraint, for all $0 \leq k \leq n-1$.

For $f = \text{Trace}(\beta)$ and $f = \text{Trace}(\beta_0 x)$ this is straightforward, since $f(x+y) = f(x)+f(y)$ and so the $\sum_{z \in A(x_0;x_1,x_2,x_3)} f(z) = 8f(x_0)+4f(x_1)+4f(x_2)+4f(x_3) = 0$ (since we are performing the arithmetic modulo 2).

Now consider $\text{Trace}(\beta x^{2^i+1})$. We will show that $\sum_{z \in A(x_0;x_1,\ldots,x_3)} z^{2^i+1} = 0$. It then follows that $\sum_z \text{Trace}(\beta z^{2^i+1}) = \text{Trace}(\beta(\sum_z z^{2^i+1})) = \text{Trace}(0) = 0$. Note further that $(x+y)^{2^i+1} = x^{2^i+1} + y^{2^i+1} + x^{2^i}y + y^{2^i}x$. Using this expansion we have:

$$\begin{aligned}
&\sum_{z \in A(x_0;x_1,\ldots,x_3)} z^{2^i+1} \\
&= \sum_{w \in A(x_0;x_1,x_2)} w^{2^i+1} + (w+x_3)^{2^i+1} \\
&= \sum_{w \in A(x_0;x_1,x_2)} (w x_3^{2^i} + w^{2^i} x_3 + x_3^{2^i+1}) \\
&= x_3^{2^i} \sum_{w \in A(x_0;x_1,x_2)} w + x_3 \sum_{w \in A(x_0;x_1,x_2)} w^{2^i} + 0 \\
&= x_3^{2^i}(4x_0 + 2x_1 + 2x_2) + x_3(4x_0^{2^i} + 2x_1^{2^i} + 2x_2^{2^i}) \\
&= 0
\end{aligned}$$

∎

We will also need to show that the codewords of $\text{RM}(2,n)$ are far away from each other.

Proposition 31 *For every $f \neq g \in \text{RM}(2,n)$, $\delta(f,g) \geq 1/7$.*

Proof: Consider any function $f \in \text{RM}(2,n)$ and let h be such that $\delta(f,h) < 1/14$. We claim that h uniquely specifies f. In particular, the algorithm: Pick x_1, x_2, x_3 at random and output $\sum_{z \in A(x;x_1,x_2,x_3)} h(z)$, outputs $f(x)$ with probability at least $1 - 7\delta(f,h) > 1/2$ and thus defines f uniquely.

We thus conclude that there can not exist $f, g \in \text{RM}(2,n)$ such that $\delta(f,g) < 1/7$.

∎

4.3.2 Key lemma

Finally we move to the main lemma of the paper. The goal of this section is to prove the following lemma.

Lemma 32 (Main Lemma) *Suppose* $g \in (\mathcal{F}^*_{k,n})^\perp$ *has weight* $t \leq k$. *Then* $g \in \mathrm{RM}(2,n)^\perp$.

To prove this lemma we first state three useful sub-lemmas, which yield the main lemma easily. We prove the sub-lemmas later.

The sub-lemmas refer to a positive integer m and the set $U = \{(i,j)|0 \leq i < j \leq m \text{ or } i = j = 0\}$. Note that $|U| = 1 + \binom{m+1}{2}$. We also use b_0 to denote the zero of $\mathbb{F}_{2^n}$.

Lemma 33 *Let* $a_1, \ldots, a_t \in \mathbb{F}_{2^n}$ *be such that* $\sum_{i=1}^t f(a_i) = 0$ *for every* $f \in \mathcal{F}^*_{k,n}$. *Further, suppose there exists* $g \in \mathrm{RM}(2,n)$ *such that* $\sum_{i=1}^t g(a_i) \neq 0$. *Then there exists* $m \leq t$, $\mathbb{F}_2$*-linearly independent elements* $b_1, \ldots, b_m \in \mathbb{F}_{2^n}$, *and a non-zero vector* $\langle \lambda_{ij} \rangle_{(i,j)\in U} \in \mathbb{F}_2^{|U|}$ *such that* $\sum_{(i,j)\in U} \lambda_{ij} f(b_i + b_j) = 0$, *for every* $f \in \mathcal{F}^*_{k,n}$.

Lemma 34 *Suppose* $b_1, \ldots, b_m \in \mathbb{F}_{2^n}$ *are* $\mathbb{F}_2$*-linearly independent elements, and* $\langle \lambda_{ij} \rangle_{(i,j)\in U} \in \mathbb{F}_2^{|U|}$ *is a non-zero vector such that* $\sum_{(i,j)\in U} \lambda_{ij} f(b_i + b_j) = 0$ *for every* $f \in \mathcal{F}^*_{k,n}$. *Then there exists a non-empty set* $E \subseteq \{(i,j)|1 \leq i < j \leq m\}$ *such that for every* $d \in [k]$ *it is the case that* $\sum_{(i,j)\in E} \left(b_i^{2^d} b_j + b_j^{2^d} b_i \right) = 0$.

Finally we show that the conclusion of the previous lemma implies that $m > k + 1$.

Lemma 35 *Suppose* $b_1, \ldots, b_m \in \mathbb{F}_{2^n}$ *are* $\mathbb{F}_2$*-linearly independent elements and suppose* $E \subseteq \{(i,j)|1 \leq i < j \leq m\}$ *is a non-empty set such that for every* $d \in [k]$, $\sum_{(i,j)\in E} \left(b_i^{2^d} b_j + b_j^{2^d} b_i \right) = 0$. *Then* $m > k + 1$.

We first show that Lemma 32 follows from the three sublemmas.

Proof: (of Lemma 32) Let $h \in (\mathcal{F}^*_{k,n})^\perp$ and suppose $h \notin \mathrm{RM}(2,n)^\perp$. We wish to show $t > k$. (We actually show $t > k+1$, but we state the weaker bound for notational simplicity.)

Let $a_1, \ldots, a_t \in \mathbb{F}_{2^n}$ be the points such that $h(a_i) = 1$. By definition of $(\mathcal{F}^*_{k,n})^\perp$ we have that $0 = \sum_{x \in \mathbb{F}_{2^n}} f(x)h(x) = \sum_{i=1}^t f(a_i)$. Since $h \notin \mathrm{RM}(2,n)^\perp$, there must exist a function $g \in \mathrm{RM}(2,n)$ such that $\sum_{i=1}^t g(a_i) \neq 0$. Using Lemma 33 we get that there exist $m \leq t$, linearly independent points $b_1, \ldots, b_m \in \mathbb{F}_{2^n}$, and a non-zero vector $\langle \lambda_{ij} \rangle_{(i,j) \in U} \in \mathbb{F}_2^{|U|}$ such that $\sum_{(i,j) \in U} \lambda_{ij} f(b_i + b_j) = 0$ for every $f \in \mathcal{F}^*_{k,n}$, where $b_0 = 0$. Applying Lemma 34 we get that there exists a non-empty set $E \subseteq \{(i,j) | 1 \leq i < j \leq m\}$ such that for every $d \in [k]$ we have $\sum_{(i,j) \in E} \left(b_i^{2^d} b_j + b_j^{2^d} b_i \right) = 0$. Applying Lemma 35 we then get that $m > k$ and thus $t \geq m > k$ as desired. ∎

We now turn to proving the three sub-lemmas. Again the crucial result here is Lemma 35 and the other two are just to pin the problem down.

Proof: (of Lemma 33) Let $b_1, \ldots, b_m$ be the largest linearly independent subset of points among $a_1, \ldots, a_t$ and let $g \in \mathrm{RM}(2,n)$ be the function satisfying $\sum_{i=1}^t g(a_i) \neq 0$.

We first claim that for every function $f \in \mathcal{F}^*_{k,n}$ at least one of the following must hold: (1) $f(0) \neq g(0)$, or (2) there exists $i \in [m]$ such that $f(b_i) \neq g(b_i)$, or (3) there exist $(i,j) \in [m] \times [m]$ such that $f(b_i + b_j) \neq g(b_i + b_j)$. To see this claim, assume otherwise, for some $f \in \mathcal{F}^*_{k,n}$. Note that we can prove, by induction on the size of the set S, that for every set $S \subseteq [m]$ we have $f(\sum_{i \in S} b_i) = g(\sum_{i \in S} b_i)$. Indeed, this is obviously true for $|S| \leq 2$. Now consider a set $S = T \cup \{i,j\}$ where $i, j \notin T$. Let

$b = \sum_{\ell \in T} b_\ell$. Now note that

$$\begin{aligned} & f(b + b_i + b_j) \\ &= f(0) + f(b) + f(b_i) + f(b_j) + f(b + b_i) \\ &\quad + f(b + b_j) + f(b_i + b_j) \\ &= g(0) + g(b) + g(b_i) + g(b_j) + g(b + b_i) \\ &\quad + g(b + b_j) + g(b_i + b_j) \\ &= g(b + b_i + b_j), \end{aligned}$$

where the first and third inequalities follow from the fact that both $f, g \in \mathrm{RM}(2, n)$ while the middle equality is by induction. But then, we have that f and g agree on the entire subspace, which contradicts the fact that $\sum_{i=1}^{t} f(a_i) \neq \sum_{i=1}^{t} g(a_i)$. Hence our claim must be true.

Consider the set $V = \{\langle f(b_i + b_j)\rangle_{(i,j)\in U} | f \in \mathcal{F}^*_{k,n}\}$. V is a linear subspace of $\mathbb{F}_2^{|U|}$ since $\mathcal{F}^*_{k,n}$ is a linear subspace; but $V \neq \mathbb{F}_2^{|U|}$ (since in particular $\langle g(b_i + b_j)\rangle_{(i,j)\in U} \notin V$. Thus there must be a non-trivial constraint $\langle \lambda_{ij}\rangle_{(i,j)\in U}$ such that every vector $x \in V$ satisfies $\sum_{(i,j)\in U} \lambda_{ij} x_{ij} = 0$. This yields the lemma. ∎

Proof: (of Lemma 34) We use the basis functions to establish this lemma. Let $b_0, b_1, \ldots, b_m$ and $\langle \lambda_{ij}\rangle_{i,j}$ be as given.

This proof also relies on the linearity of the the Trace function, and the additional fact that $\mathrm{Trace}(ax) = 0$ for every $x \in \mathbb{F}_{2^n}$ if and only if $a = 0$. (This is easily seen since $\mathrm{Trace}(ax)$ is a non-zero polynomial of degree 2^{n-1} in x, if $a \neq 0$.)

First consider the constant function $1 = \mathrm{Trace}(\beta)$ for some $\beta \in \mathbb{F}_{2^n}$. Since $\mathrm{Trace}(\beta) \in \mathcal{F}^*_{k,n}$ we have $\sum_{i,j} \lambda_{ij} = \sum_{i,j} \lambda_{ij} \mathrm{Trace}(\beta) = 0$, and thus $\lambda_{00} = \sum_{(i,j)\in U-(0,0)} \lambda_{ij}$.

Next we consider the functions $\mathrm{Trace}(\beta_0 x) \in \mathcal{F}^*_{k,n}$. We have $0 = \sum_{i,j} \lambda_{ij} \mathrm{Trace}(\beta_0(b_i + b_j)) = \mathrm{Trace}\left(\beta_0 \sum_{i,j} \lambda_{ij}(b_i + b_j)\right)$. Using the aforementioned property of the Trace function, we have that the above identity holds for every $\beta_0 \in \mathbb{F}_{2^n}$ only if $\sum_{i,j} \lambda_{i,j}(b_i + b_j) = 0$. Let $\tau_i = \sum_{j<i} \lambda_{ji} + \sum_{j>i} \lambda_{ij}$. (For simplicity of notation below, we will

assume $\lambda_{ij} = \lambda_{ji}$.) Then we have $0 = \sum_{i,j} \lambda_{ij}(b_i + b_j) = \sum_{i=0}^{m} \tau_i b_i = \sum_{i=1}^{m} \tau_i b_i$ (where the last equality follows from $b_0 = 0$). But $b_1, \ldots, b_m$ are linearly independent over $\mathbb{F}_2$ and $\tau_i, \lambda_{ij} \in \mathbb{F}_2$, so the only way $\sum_{i=1}^{m} \tau_i b_i = 0$ is if $\tau_i = 0$ for every i. Thus we get $\lambda_{0i} = \sum_{j \neq 0} \lambda_{ji}$ for every $i \in [m]$

Finally we consider $\text{Trace}(\beta_d x^{2^d+1}) \in \mathcal{F}^*_{k,n}$ for $d \in [k]$. We have

$$0 = \sum_{i,j} \lambda_{ij} \text{Trace}\left(\beta_d (b_i + b_j)^{2^d+1}\right) = \text{Trace}\left(\beta_d \sum_{i,j} \lambda_{ij} (b_i + b_j)^{2^d+1}\right).$$

Again, we have that the above identity holds for every $\beta_d \in \mathbb{F}_{2^n}$ only if $\sum_{i,j} \lambda_{i,j}(b_i + b_j)^{2^d+1} = 0$. Expanding $(x+y)^{2^d+1}$ as $x^{2^d+1} + y^{2^d+1} + x^{2^d}y + xy^{2^d}$, we get

$$\begin{aligned} 0 &= \sum_{i,j} \lambda_{ij} \left(b_i^{2^d+1} + b_j^{2^d+1} + b_i^{2^d} b_j + b_i b_j^{2^d}\right) \\ &= \sum_{i=1}^{m} \tau_i b_i^{2^d+1} + \sum_{1 \leq i < j \leq m} \lambda_{ij} (b_i^{2^d} b_j + b_i b_j^{2^d}) \\ &= \sum_{(i,j) \in E} (b_i^{2^d} b_j + b_i b_j^{2^d}), \end{aligned}$$

where $E = \{(i,j) | 1 \leq i < j \leq m \text{ s.t. } \lambda_{ij} \neq 0\}$ as required for the lemma statement. The only remaining issue is to show that $E \neq \emptyset$.

We claim that if $E = \emptyset$ we have $\lambda_{ij} = 0$ for every i, j. For $i, j \geq 1$ this follows from the definition of E. For $i \neq 0$ and $j = 0$ this follows from the identity above that $\lambda_{0i} = \sum_{j \neq 0} \lambda_{ji} = 0$. For $i = j = 0$, we also have $\lambda_{00} = \sum_{(i,j) \in U - (0,0)} \lambda_{ij} = 0$. But this contradicts the hypothesis that $\langle \lambda_{ij} \rangle \neq 0$, and so we conclude $E \neq \emptyset$. ∎

Proof: (of Lemma 35) This is the crux of our analysis and uses a mix of linear and polynomial algebra arguments. Assume for contradiction that $m \leq k + 1$.

Recall we are given that for every $d \in [k]$ $\sum_{(i,j) \in E} (b_i^{2^d} b_j + b_i b_j 2^d) = 0$. Note further that we also trivially have this condition for $d = 0$, since $\sum_{(i,j) \in E} (b_i^{2^d} b_j + b_i b_j 2^d) = \sum_{(i,j) \in E} (b_i b_j + b_i b_j) = \sum_{(i,j) \in E} 0$.

For $i \in [m]$, let $\rho_i = \sum_{\{j|(i,j) \text{ or } (j,i) \in E\}} b_j$. Then we can rewrite $\sum_{(i,j)\in E}(b_i^{2^d} b_j + b_i b_j 2^d)$ as $\sum_{i=1}^m \rho_i b_i^{2^d}$ and so we have, for every $d \in \{0, 1, \ldots, k\}$ as $\sum_{i=1}^m \rho_i b_i^{2^d} = 0$.

Consider the $m \times m$ matrix $A = (a_{ij})$ with $a_{ij} = b_j^{2^{i-1}}$. Then the previous paragraph implies that $A \cdot \rho = 0$ for the column vector $\rho = \langle \rho_1, \ldots, \rho_m \rangle$. (In particular, we have that the ith entry of $A \cdot \rho$ equals $\sum_{j=1}^m b_j^{2^{i-1}} \rho_j$ which is 0 for every $i \in \{1, \ldots, k+1\} \supseteq \{1, \ldots, m\}$.)

Next we note that $\rho \neq 0$. This is true since for at least one $i \in [m]$ the summation $\sum_{\{j|(i,j) \text{ or } (j,i) \in E\}} b_j$ sums over a non-empty set of indices j (since $E \neq \emptyset$). But now the linear independence of $b_1, \ldots, b_m$ over $\mathbb{F}_2$ implies that the summation, and hence ρ_i, is non-zero.

We conclude that the matrix A is singular. We now use this fact to infer that A has a non-zero vector in its left kernel, i.e., there exists a non-zero row vector $\lambda = \langle \lambda_1, \ldots, \lambda_m \rangle$ such that $\lambda A = 0$. But now consider the polynomial $\Lambda(x) = \sum_{i=1}^m \lambda_i x^{2^{i-1}}$. Using this notation, we have $\lambda A = \langle \Lambda(b_1), \ldots, \Lambda(b_m) \rangle$. Thus the condition $\lambda A = 0$ implies that $\Lambda(b_j) = 0$ for every $j \in \{1, \ldots, m\}$.

But now, we have that $\Lambda(x)$ is a non-zero polynomial (since λ is a non-zero vector), of degree at most 2^{m-1}. Furthermore Λ is a linearized polynomial and satisfies $\Lambda(x+y) = \Lambda(x) + \Lambda(y)$. This implies that $\Lambda(b_S) = 0$ for every $S \subseteq [m]$, where $b_S = \sum_{i \in S} b_i$. The linear independence of $b_1, \ldots, b_m$ furthermore implies that the b_S's are all distinct and thus we get that Λ is a non-zero polynomial of degree at most 2^{m-1} with 2^m distinct roots, yielding the desired contradiction. ∎

4.3.3 Putting it together

We now use the main lemma of the previous subsection to claim that membership in $\mathcal{F}^*_{k,n}$ is not testable with a strong k-local test (i.e. non-adaptive, one sided error). This part is more or less standard and follows, for instance, from the methods in [23]. We include the full details for completeness.

We first summarize our arguments from the previous section in a slightly more convenient form.

Lemma 36 *Fix* $a_1, \ldots, a_t \in \mathbb{F}_{2^n}$. *For* $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ *let* $\pi(f) = \pi_{a_1,\ldots,a_t}(f) = \langle f(a_1), \ldots, f(a_t) \rangle$ *be the projection of* f *to* $a_1, \ldots, a_t$. *Let* $V \subseteq \mathbb{F}_2^t$ *be the set* $V = \{\pi(f) | f \in \mathcal{F}^*_{k,n}\}$, *and let* $W = \{\pi(f) | f \in \mathrm{RM}(2,n)\}$. *If* $t \leq k$, *then* $V = W$.

Proof: We first note that V and W are linear subspaces of $\mathbb{F}_2^t$. This follows from the fact that $\mathcal{F}^*_{k,n}$ and $\mathrm{RM}(2,n)$ are linear spaces. Since $\mathcal{F}^*_{k,n} \subsetneq \mathrm{RM}(2,n)$, it also follows that $V \subseteq W$. Suppose $V \neq W$. Then it follows, by linear algebra, that there exist vectors $u, w \in \mathbb{F}_2^t$ such that $u \cdot v = 0$ for every $v \in V$, $u \cdot w \neq 0$ and $w \in W$. Since $w \in W$ there exists $h \in \mathrm{RM}(2,n)$ such that $w = \pi(h)$. Let $a'_1, \ldots, a'_{t'}$ be the subsequence of $a_1, \ldots, a_t$ corresponding to indices i such that $u_i \neq 0$. Then we have $\sum_{i=1}^{t'} h(a'_i) = 1$ while $\sum_{i=1}^{t'} f(a'_i) = 0$ for every $f \in \mathcal{F}^*_{k,n}$. By Lemma 32 we have $t \geq t' > k$.

∎

We can now prove Theorem 22.

Proof: **(of Theorem 22)** For every n, the domain $D_n = \mathbb{F}_{2^n}$. For every n, the family of functions we work with is $\mathcal{F}_n = \mathcal{F}^*_{k,n}$.

First note, by Proposition 30 that for every n, $\mathcal{F}_n$ has a non-zero function in its dual of weight 8. Next, by Proposition 24 we also have that $\mathcal{F}_n$ is affine invariant and thus (by Proposition 23) 2-transitive. It remains to show that $\mathcal{F}$ is not k-locally testable. Assume $\mathcal{F}$ is t-locally testable, i.e., for all sufficiently large n there is a one-sided error, non-adaptive, tester $T = T_n$ that accepts every member of $\mathcal{F}_n$ while rejecting all functions at distance at least, say, $1/7$ from $\mathcal{F}_n$ with positive probability. We argue below that this can not happen if $t \leq k$ and $n > 2k + 1$.

Suppose $t \leq k$. Fix the coins of T to some string R and let $a_1, \ldots, a_t \in \mathbb{F}_{2^n}$ be the queries of the tester T on random string R. Let π, V and W be as in the statement

of Lemma 36. Since the tester makes one-sided error, it follows that it must accept every pattern in V (i.e., accepts every function f such that $\pi(f) \in V$). By Lemma 36 we have $V = W$ and so the tester accepts every element of $\mathrm{RM}(2,n)$ also on random string R. Thus we get that every element of $\mathrm{RM}(2,n)$ is accepted with probability one by the tester T. Since $\mathrm{RM}(2,n) \neq \mathcal{F}^*_{k,n}$ for $k < \lfloor n/2 \rfloor$ (Lemma 25) there exists a function $h \in \mathrm{RM}(2,n) - \mathcal{F}^*_{k,n}$ that is accepted with probability one. Furthermore, by the distance of $\mathrm{RM}(2,n)$ (Proposition 31) and the fact that $\mathcal{F}^*_{k,n} \subseteq \mathrm{RM}(2,n)$, we have that $\delta(h, \mathcal{F}^*_{k,n}) \geq 1/7$. We conclude that the tester T accepts functions at distance $1/7$ from $\mathcal{F}^*_{k,n}$ with probability one, violating the requirement above. ∎

4.4 Discussion

As mentioned in the introduction, our results here show the existence of families $\mathcal{F}^*_{k,n}$ whose duals contain weight 8 functions but any basis must contain functions of weight $> k-2$. This is enough to show that for growing k these families are not testable (hence, they are not weakly testable in the sense of Definition 6.)

A stronger variant of the AKKLR conjecture considers 2-transitive families that are spanned by low weight functions. In this case one can also consider the strong testability notion from Definition 7. In that notion any codeword that does not belong to the code is rejected with some non-zero probability. Note that in order for the test to reject codewords that are very close to the code it must be the case that the code is spanned by low-weight codewords.

We now formally state a stronger variant of the AKKLR-conjecture which considers families whose duals have a basis of low weight rather than just small distance (See also [21].)

Conjecture 37 (variant of AKKLR) *For every $d \in \mathbb{N}$, there exists $k = k(d) < \infty$ such that the following holds: Let $\mathcal{F} = \{\mathcal{F}_n\}_n$ be an ensemble of properties such that for every n,*

1. $\mathcal{F}_n^{\perp}$ *is spanned by functions of weight at most* d, *and*

2. $\mathcal{F}_n$ *is* 2*-transitive.*

Then $\mathcal{F}$ *is* k*-locally testable.*

We remark that this variant remains open and it leads to similar conjectures for other groups of symmetries (such as affine, cyclic, or less-explicit abelian or non-abelian groups.)

Chapter 5

Explicit Structured Testing in Common Algebraic Families

In this chapter we focus on *explicit* and *succinct* representations of common algebraic families, two properties which imply nice testing applications, namely *structured* testing. We show the existence of explicit structured tests for BCH codes of design distance 5 (i.e. BCH$(2, n)$, for any n) and the novelty of this result lies in the fact that the support of the test can be described by a fixed set of carefully chosen univariate polynomials.

As described in the introduction, explicitness and succinctness are motivated by the recent results of Kaufman and Sudan [72], who identified 'the single orbit property' as a source of sufficient conditions for testing. A family exhibiting the single orbit property can be generated as a vector space by one function of small weight and all its translates under permutations in its automorphism group.

An immediate consequence of the single orbit property is that it implies a succinct description of the family, in the following sense. To specify a vector space one needs to specify a set of basis vectors, which could amount to specifying $dN \log N$ bits (where N is the length of a vector, and d is the dimension of the vector space). However, when the vector space has a single orbit generator of small weight, then it can be specified by the support of this generator, i.e. by $O(\log N)$ bits.

Let us discuss explicitness and succinctness for Reed-Muller codes of order 1, i.e. the family of n-variate polynomials over $\mathbb{F}_2$ of total degree at most 1 (A similar argument was shown in the introduction for the Hadamard code, which is linear invariant under $\mathrm{GL}(n,2)$ but not affine invariant under $\mathrm{AGL}(n,2)$). A set of explicit tests for such families is supported on 4-tuples from the set $S = \{\langle a,b,c,a+b+c\rangle | a,b,c \in \mathbb{F}_2^n\}$. It is well-known that this set of tests contains a basis for the dual of $\mathrm{RM}(1,n)$. Why does $\mathrm{RM}(1,n)^\perp$ have the single orbit property? Consider the function f supported on $\tau_1 = \langle 0, e_1, e_2, e_1+e_2\rangle$, where e_i is the standard ith basis vector over $\mathbb{F}_2^n$. One can easily check that any 4-tuples $\langle a,b,c,a+b+c\rangle$ in S can be obtained from τ_1 as $\langle a,b,c,a+b+c\rangle = \langle \pi(0), \pi(e_1), \pi(e_2), \pi(e_1+e_2)\rangle$ for some $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n$ with $\pi(x) = Ax + a$, and $A \in \mathbb{F}_2^{n\times n}$. Thus f is a single orbit generator for $\mathrm{RM}(1,n)^\perp$ under the group $\mathrm{AGL}(n,2)$.

In this chapter we will see that $\mathrm{RM}(1,n)^\perp$ (and in fact the more general $\mathrm{RM}(d,n)^\perp$) has a single orbit generator under a much smaller group of invariances, namely $\mathrm{AGL}(1,2^n)$. This observation requires viewing these codes as univariate polynomials over $\mathbb{F}_{2^n}$ (rather than n-variate polynomials over $\mathbb{F}_2$.) In particular, if $w \in \mathbb{F}_{2^n}$ is a primitive element and $\tau_2 = \langle 0, 1, w, 1+w\rangle$ then $\mathrm{RM}(1,n)^\perp$ is generated by the set of all functions supported at $\langle \pi(0), \pi(1), \pi(w), \pi(1+w)\rangle$, with $\pi \in \mathrm{AGL}(1,2^n)$.

In the next chapter we study what families, other than Hadamard, Reed Muller, and $\mathrm{eBCH}(2,n)$ exhibit the single orbit property under the group $\mathrm{AGL}(1,2^n)$. Those results exhibit a large class of families that admit structured tests. There we do not insist on fully explicit tests to the extent of specifying the relations between the elements of the support of a generator. In fact such descriptive level might be hard (if possible) to achieve in that general setting.

Regarding our techniques, we first analyze sufficient conditions for a family to exhibit the single orbit property, in terms of certain 'diagonal' systems of equations. These equations bear a resemblance to the equations arising in the so-called Waring problem. A version of the Waring problem studies ways to express a polynomial as sums of d-th powers of some special polynomials. Our explicit description is inspired by old results

of Paley [85] from the 1930s on the Waring problem. There he describes families of explicit polynomials that satisfy conditions similar to those required by the single orbit property.

5.1 Definitions and main result

Definition 38 (k-Single Orbit) *Let $\mathcal{F} \subseteq \{D \to \mathbb{F}_2\}$ be a linear collection of functions from D to $\mathbb{F}_2$ for some domain D. Let G be a group of functions from D to D. Then $\mathcal{F}$ is said to have the k-single orbit property under the group G if there exists $f \in \mathcal{F}$ with $\mathrm{wt}(f) \leq k$ such that $\mathcal{F} = \mathrm{Span}(\{f \circ \pi | \pi \in G\})$. We say that f is a k-single orbit generator for $\mathcal{F}$.*

We note that in [72] the single-orbit property under the affine group is described as 'formal characterization'.

Definition 39 (Structured Tester) *Let $\mathcal{P}$ be such that its dual has a k-single orbit generator g under a group G. A* structured tester *for $\mathcal{P}$, given a function f: (1) picks $\pi \in G$ uniformly at random, and (2) accepts if $\langle f, g \circ \pi \rangle = 0$, otherwise it rejects.*

When the group of symmetries is an affine group of permutations of the domain, the results of [72] show that the dual family is testable by a structured tester. We state the theorem in the form we need it here and refer to Appendix A for comments regarding the original form from [72].

Theorem 40 ([72]) *If $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ is linear and has the k-single orbit property under the affine group $\mathrm{AGL}(1, 2^n)$, then $\mathcal{F}^\perp$ is $(k, \Omega(1/k^2))$-locally testable by a structured tester.*

BCH codes We next define BCH codes and mention a few of their properties. BCH codes have many alternative and equivalent definitions, for example as subfield

subcodes of Reed Solomon codes with $\mathrm{BCH}(t,n)$ being the subfield subcodes of RS codes of degree $N-2t-1$. They are also commonly defined as cyclic codes whose roots satisfy certain relations. Here we use their representation as evaluations of univariate polynomials, and it is more convenient for us to define them by first defining their duals.

Definition 41 (BCH code) *For every pair of integers n and t, the (binary) dual-BCH code with parameters n and t, denoted $\mathrm{BCH}(t,n)^{\perp} \subseteq \mathbb{F}_2^{2^n-1}$ consists of the evaluations of traces of polynomials of degree $2t$ over $\mathbb{F}_{2^n}^*$. I.e.,*

$$\mathrm{BCH}(t,n)^{\perp} = \{\langle \mathrm{Trace}(f(\alpha))\rangle_{\alpha\in\mathbb{F}_{2^n}^*} | f \in \mathbb{F}_{2^n}[x], \deg(f) \leq 2t\}$$

The BCH code $\mathrm{BCH}(t,n)$ is simply the dual of $\mathrm{BCH}(t,n)^{\perp}$.

$\mathrm{eBCH}(t,n)$ is the extension of $\mathrm{BCH}(t,n)$ by a parity check bit. That is, $\mathrm{eBCH}(t,n)$ is the evaluation of the same polynomials, over the entire domain $\mathbb{F}_{2^n}$. The *design* minimum distance of $\mathrm{BCH}(t,n)$ is $2t+1$ and the minimum distance of $\mathrm{eBCH}(t,n)$ is $2t+2$.

We can now state the main result of this chapter.

Theorem 42 *For any n, there exists $\alpha \in F_{2^n} - \mathbb{F}_2$ such that the function supported at*

$$\langle 0, 1, 1+\alpha^4, \alpha+\alpha^2+\alpha^4, \alpha^2+\alpha^3+\alpha^4, \alpha+\alpha^3+\alpha^4\rangle$$

is [1] *a 6-single orbit generator for $\mathrm{eBCH}(2,n)$ under the affine group $\mathrm{AGL}(1,2^n)$.*

Corollary 43 $\mathrm{eBCH}^{\perp}(2,n)$ *is $(6,\Omega(1))$-locally testable by an explicit structured tester.*

As mentioned, we also describe an explicit structured test (by presenting a single orbit generator under the group $\mathrm{AGL}(1,2^n)$,) for general RM codes.

[1] A variant of these polynomials were suggested in [85] in relation to a problem of Waring.

5.2 Sufficient conditions for single orbit

First recall from Chapter 3 that for $d \in \{1, \ldots, 2^n - 2\}$, $\text{orb}(d) = \{d, 2d \pmod{2^n - 1}, 4d \pmod{2^n - 1}, \ldots, 2^{n-1}d \pmod{2^n - 1}\}$, min-orb$(d)$ denotes the smallest integer in orb(d), and $\mathcal{D} = \{\text{min-orb}(d) \mid d \in \{1, \ldots, 2^n - 2\}\} \cup \{2^n - 1\}$. Also recall that Δs denotes the shadow of degree s. Hereafter, in this chapter affine invariance refers to invariance under the group $\text{AGL}(1, 2^n)$. We will make use of the characterization of affine invariant properties by a set of degrees $D \subseteq \mathcal{D}$ as stated in Proposition 16.

The following lemma is the main tool in proving our theorem.

Lemma 44 *Let $\mathcal{F} \in \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be an affine invariant family, and let D be the set of degrees that describe $\mathcal{F}$. If for some $\langle a_1, a_2, \ldots, a_k \rangle \in \mathbb{F}_{2^n}^k$ the following conditions hold*

1. $\sum_{i=1}^{k} a_i^d = 0$ *for all* $d \in D$
2. $\sum_{i=1}^{k} a_i^{2^l+1} \neq 0$ *for all* $2^l + 1 \in \mathcal{D} - D$
3. $\sum_{i=1}^{k} a_i^s \neq 0$ *for all* $s \in \mathcal{D} - D$ *with* $\text{bwt}(s) \geq 3$ *for which* $\Delta s - \{s\} \subseteq D$,

then $\mathcal{F}^\perp$ has the k-single orbit property with a generator g supported at $\langle a_1, a_2, \ldots, a_k \rangle$.

We first show a simple lemma.

Lemma 45 *Let $\mathcal{F} \in \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be an affine invariant family and let D be the set of degrees that describe $\mathcal{F}$. Then a function g supported at $\langle a_1, a_2, \ldots, a_k \rangle \in \mathbb{F}_{2^n}^k$ belongs to $\mathcal{F}^\perp$ if and only if*

$$\sum_{i=1}^{k} a_i^d = 0 \text{ for all } d \in D.$$

Proof: By definition, $\mathcal{F} = \{\text{Trace}(\sum_{d\in D} \alpha_d x^d), \alpha_d \in \mathbb{F}_{2^n}\}$. Then $g \in \mathcal{F}^\perp$ if and only if for any $f \in \mathcal{F}$, $\langle f, g\rangle = 0$, that is $\sum_{i=1}^{k} f(a_i) = 0$. Let $f_\alpha(x) = \text{Trace}(\alpha x^d) \in \mathcal{F}$, for some $\alpha \in \mathbb{F}_{2^n}$. Then $0 = \sum_{i=1}^{k} f_\alpha(a_i) = \sum_{i=1}^{k} \text{Trace}(\alpha a_i^d) = \text{Trace}(\alpha \sum_{i=1}^{k} a_i^d)$. Since $f_\alpha \in \mathcal{F}$ for all $\alpha \in \mathcal{F}$, it follows that for $\beta = \sum_{i=1}^{k} a_i^d$, the previous identity holds if and only if $\text{Trace}(\beta\alpha) = 0$ for all $\alpha \in \mathbb{F}_{2^n}$. But the function $\text{Trace}(\beta x)$ is linear, and it is identically null only when $\beta = 0$, which concludes the proof.

∎

Proof of Lemma 44: By Lemma 45, condition 1 immediately implies that $g \in \mathcal{F}^\perp$.

We will show that $\mathcal{F}^\perp$ is the smallest affine invariant family that contains g. Notice that if g belongs to some affine invariant family then the set $\{g\circ\pi, \pi \in \text{AGL}(1, 2^n)\}$ belongs to that family, and by linearity, the set of functions $\text{Span}\{g\circ\pi, \pi \in \text{AGL}(1, 2^n)\}$ is included in the family as well. Therefore, g is a single orbit generator for the smallest affine invariant family that contains it, which will conclude the proof.

Assume for the sake of contradiction that $g \in \mathcal{C} \subsetneq \mathcal{F}^\perp$ and $\mathcal{C}$ is affine invariant. Then $\mathcal{F} \subsetneq \mathcal{C}^\perp$, and let $D' \subseteq \mathcal{D}$ be the shadow-closed set of degrees that characterizes $\mathcal{C}^\perp$, i.e. $\mathcal{C}^\perp = \{\text{Trace}(\sum_{d\in D'} \alpha_d x^d), \alpha_d \in \mathbb{F}_{2^n}), \}$. Since $g \in \mathcal{C}$, by Lemma 45 we must have $\sum_{i=1}^{k} a_i^d = 0$ for all $d \in D'$, where by assumption $D \subsetneq D'$. Since D' is shadow closed it follows that all the weight 2 shadows of degrees $d \in D'$ also belong to D'. By condition 2, $\sum_{i=1}^{k} a_i^{2^l+1} \neq 0$ if $2^l + 1 \notin D$, which implies that D' does not contain any extra weight 2 degrees than those in D. Hence the degrees in D' are of the form $d = 2^{l_1} + 2^{l_2} + \ldots + 2^{l_r} + 1$ such that $2^{l_i - l_j} + 1 \in D$ for all i, j. If $D \neq D'$, any affine invariant family with degrees of this form must contain some degree s such that $\Delta s - \{s\} \in D$. Since g satisfies condition 3 it must be the case that $\sum_i a_i^s \neq 0$. Hence, since $s \in D'$ by Lemma 45 $\quad g \notin \mathcal{C}$, a contradiction.

∎

We next show a simple application of the above lemma.

5.2.1 A warm-up: explicit single orbit for the eBCH$(1,n)$

Proposition 46 *Let w be a primitive element of $\mathbb{F}_{2^n}$. Then the function supported on $\langle 0, 1, w, 1+w\rangle$ is a 4-single orbit generator for* eBCH$(1,n)$.

Proof: Notice that the set of degrees that characterizes eBCH$(1,n)^{\perp}$ is $D = \{1\}$. Then condition 1 of Lemma 44 is trivially satisfied. Also condition 3 is vacuously satisfied, since if $\mathrm{bwt}(s) \geq 3$ then $\Delta s - s \not\subset D$.

To verify condition 2, notice that $1 + w^{2^\ell+1} + (1+w)^{2^\ell+1} = w + w^{2^\ell+1} \neq 0$. Indeed, otherwise $w^{2^\ell-1} = 1$, which would imply that w belongs to a subfield of size 2^ℓ, contradicting the assumption that w is a primitive element. ∎

Corollary 47 *If n is prime, then* eBCH$(1,n)$ *is the only affine invariant family that contains functions of weight at most* 4.

Proof: Proposition 46 implies that eBCH$(1,n)$ satisfies the corollary. Suppose that $\mathcal{C}$ is another affine invariant family which contains a function f of weight at most 4, and whose dual is characterized by a set of integer degrees D.

Assume that the support of f is $\{0, 1, \alpha, 1+\alpha\}$, for some $\alpha \notin \mathbb{F}_2$. As shown in Lemma 44, any degree $d \in D$ must satisfy $0^d + 1^d + \alpha^d + (1+\alpha)^d = 0$. If D contains degrees of binary weight at least 2, since D is shadow closed it must contain a degree of binary weight 2. If $2^l + 1 \in D$ then by the same argument as in the proof of Proposition 46 it follows that $\alpha \in \mathbb{F}_{2^l}$, a contradiction to the fact that $\mathbb{F}_{2^n}$ does not contain non-trivial subfields.

It remains to argue the case when $\{0,1\}$ is not included in the support of f. Let a_1, a_2, a_3, a_4 belong to the support of f. Then there exists a permutation $\pi \in \mathrm{AGL}(1, 2^n)$ such that $\pi(a_1) = 0$ and $\pi(a_2) = 1$. Then the function $f \circ \pi$ supported at $\pi(a_1), \pi(a_2), \pi(a_3), \pi(a_4)$ also belongs to $\mathcal{C}$, and hence this case reduces to the above case.

∎

5.3 Explicit single orbit for the eBCH$(2, n)$

In this section we prove our main theorem. As a by-product of our method we show an explicit single orbit under AGL$(1, 2^n)$ for general RM codes.

Proof of Theorem 42: Notice that $D = \{1, 3\}$ is the set of degrees that characterizes eBCH$(2, n)^{\perp}$. We will show that there exists $\alpha \in \mathbb{F}_{2^n}$ such that $a_1 = 0$, $a_2 = 1$, $a_3 = 1 + \alpha^4$, $a_4 = \alpha + \alpha^2 + \alpha^4$, $a_5 = \alpha^2 + \alpha^3 + \alpha^4$, $a_6 = \alpha + \alpha^3 + \alpha^4$, satisfy the conditions of Lemma 44 and therefore they form the support of a 6-single orbit generator for eBCH$(2, n)$.

One can easily check that there is no $s \in \mathcal{D} - D$ with bwt$(s) \geq 3$ such that $\Delta s - s \subseteq D$, and hence condition 3 is vacuously satisfied.

We now proceed to verify condition 2. To that end, for each $2 \leq \ell \leq \lfloor n/2 \rfloor + 1$ define the polynomial

$$P_\ell(x) = 1 + (1 + x^4)^{2^\ell + 1} + (x + x^2 + x^4)^{2^\ell + 1} + (x^2 + x^3 + x^4)^{2^\ell + 1} + (x + x^3 + x^4)^{2^\ell + 1}.$$

Also let,

$$Q(x) = \prod_{\ell=2}^{\lfloor n/2 \rfloor + 1} P_\ell(x).$$

We will argue that $Q(x)$ is not identically 0 over $\mathbb{F}_{2^n}$, which implies the existence of an $\alpha \neq 0, 1$ such that $Q(\alpha) \neq 0$, and concludes the proof.

First notice that the degree in each P_ℓ is at most $4(2^\ell + 1)$, and thus the total degree of Q is at most $4(2^{n/2+1} + 1)n/2 < 2^n - 1$, for large enough n. Hence, no degree is too large to wrap around modulo $x^{2^n} - x$ and cause cancellations with smaller degree terms from the expansion of Q.

Secondly, we argue that each factor is a non-zero polynomial. Hence the product of minimum degree monomials in each P_ℓ results in a non-zero term of $Q(x)$ that cannot be canceled by other terms in the expansion. Indeed, one can easily check that the minimum degree monomial in each P_ℓ is $x^{2^\ell + 2}$. Therefore, the monomial of degree

$\sum_{\ell=2}^{\lfloor n/2 \rfloor+1} 2^\ell + 2 < 2^{n/2+3} + n < 2^n - 1$ is the minimum degree term of Q in the expansion. Since $Q(x)$ is a non-zero polynomial over $\mathbb{F}_{2^n}$, there must exist $\alpha \in \mathbb{F}_{2^n}$ such that $Q(\alpha) \neq 0$ and thus $P_\ell(\alpha) \neq 0$ for all $2 \leq \ell \leq \lfloor n/2 \rfloor + 1$. To finish the proof of condition 2 of Lemma 44, notice that for $\ell > \lfloor n/2 \rfloor + 1$ it is the case that $2^\ell + 1 \notin \mathcal{D}$. Indeed, for $\ell > \lfloor n/2 \rfloor + 1$, $2^\ell + 1 = 2^\ell + 2^n = 2^\ell(1 + 2^{n-\ell}) (\mathrm{mod}\ 2^n - 1)$ and therefore $2^\ell + 1 \in \mathrm{orb}(2^{n-\ell} + 1)$. In fact there are at least $2^n - 1 - 2^{n/2+3} - n$ α's that satisfy the conditions of Lemma 44.

∎

5.4 Explicit single orbit for $\mathrm{RM}(d,n)^\perp$

We will need a generalization of Lucas' identities to multinomial coefficients.

First recall that for $d, e_1, \ldots, e_k \leq 2^n - 2$ with $\sum_{i=1}^k e_i \leq d$ the multinomial coefficient $\binom{d}{e_1, e_2, \ldots, e_k} = \binom{d}{e_1}\binom{d-e_1}{e_2} \cdots \binom{d-e_1-e_2-\ldots-e_{k-1}}{e_k}$.

Fact 48 (Generalized Lucas identity) *For $d, e_1, \ldots, e_k \leq 2^n - 2$ with $d = \sum_{i=0}^{n-1} d_i 2^i$ and $e_i = \sum_{j=0}^{n-1} e_{ij} 2^j$. Then*

$$\binom{d}{e_1, e_2, \ldots, e_k} \mathrm{mod}\ 2 = \prod_{j=1}^{n-1} \binom{d_j}{e_{1j}, e_{2j}, \ldots, e_{kj}}.$$

Hence

$$\binom{d}{e_1, e_2, \ldots, e_k} \mathrm{mod}\ 2 = \begin{cases} 1 & \text{if } \sum_{i=1}^k e_{ij} \leq d_j \ \ \forall \ \ 0 \leq j \leq n-1 \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 49 *For large enough n and integer $d \geq 1$, there exists $\alpha \in \mathbb{F}_{2^n}$ such that the function supported at*

$$\langle \mathrm{Span}\{1, \alpha, \alpha^2, \ldots, \alpha^d\} \rangle,$$

is a 2^{d+1}*-single orbit generator for* $\mathrm{RM}(d,n)^{\perp}$ *under the group* $\mathrm{AGL}(1,2^n)$.

Proof: We proceed with proving that condition 1 of Lemma 44 is satisfied. Recall Definition 27 of $\mathrm{RM}(d,n)$ from Chapter 4, and the fact these codes are characterized by the set of degrees $D(d) = \{j|\mathrm{bwt}(j) \leq d, j \neq 0\}$. Since $\mathrm{RM}(1,n)$ is equivalent to $\mathrm{eBCH}^{\perp}(1,n)$ and we proved this case in Proposition 46, it is enough to consider $d \geq 2$. In this case condition 2 of the lemma becomes vacuously satisfied.

Condition 1 of the lemma is standard (and holds for any α, and in fact for any support of dimension at most $d+1$), but we include its proof here for the sake of completeness (similar arguments appear for example in [81], Chapter 12.) We prove it by induction on d. For $d=1$, $0+1+\alpha+(1+\alpha)=0$. Let $\mathcal{A}(\alpha,l) = \mathrm{Span}\{1,\alpha,\alpha^2,\ldots,\alpha^l\}$. Assume that for any α and any $d' < d$ it is the case that $\sum_{\beta\in\mathcal{A}(\alpha,d')} \beta^r = 0$ for all $\mathrm{bwt}(r) \leq d'$. We show that $\sum_{\beta\in\mathcal{A}(\alpha,d)} \beta^{\ell} = 0$ for all ℓ with $\mathrm{bwt}(\ell) \leq d$. Notice that

$$\begin{aligned}
\sum_{\beta\in\mathcal{A}(\alpha,d)} \beta^{\ell} &= \sum_{\beta\in\mathcal{A}(\alpha,d-1)} \left(\beta^{\ell} + (\beta+\alpha^d)^{\ell}\right) = \sum_{\beta\in\mathcal{A}(\alpha,d-1)} \left(\beta^{\ell} + \sum_{r} \binom{\ell}{r} \beta^r (\alpha^d)^{\ell-r}\right) \\
&= \sum_{\beta\in\mathcal{A}(\alpha,d-1)} \left(\beta^{\ell} + \sum_{r\prec\ell} \beta^r (\alpha^d)^{\ell-r}\right) = \sum_{\beta\in\mathcal{A}(\alpha,d-1)} \sum_{r\prec\ell, r<\ell} \beta^r (\alpha^d)^{\ell-r} \\
&= \sum_{r\prec\ell, r<\ell} (\alpha^d)^{\ell-r} \left(\sum_{\beta\in\mathcal{A}(\alpha,d-1)} \beta^r\right) = \sum_{r\prec\ell, r<\ell} (\alpha^d)^{\ell-r} \cdot 0 \\
&= 0.
\end{aligned}$$

In the above we used Lucas' identity from Chapter 3 and, since $r \prec \ell, r < \ell$ implies $\mathrm{bwt}(s) \leq d-1$ we used the identities given by the induction hypothesis. This concludes the proof for condition 1.

To show condition 3, notice that the set $S = \{s|\Delta s - s \subseteq D\} \cap \mathcal{D} = \{s|\mathrm{bwt}(s) = d+1\} \cap \mathcal{D}$. The existence of α satisfying the theorem follows from an argument similar to the one made in the proof of Theorem 42.

We need to define some notation first. For the purpose of this proof let $\{0,1,\ldots,d\}$

be denoted by $[0,d]$. For $s \in S$, such that $\text{bwt}(s) = d+1$, and $T \subseteq [0,d]$ $(T \neq \emptyset)$ let the polynomial $q_T(x) = \sum_{i \in T} x^i$ and let

$$P_s(x) = \sum_{T \subseteq [0,d]} q_T(x)^s.$$

We first show that P_s is a non-identically zero polynomial, by exhibiting a non-canceling degree. Let $s = 2^{s_0} + 2^{s_1} + \ldots + 2^{s_d}$, with $s_0 < s_1 < \ldots < s_d$.

Then

$$\begin{aligned} P_s(x) &= \sum_{T \subseteq [0,d]} q_T(x)^s \\ &= \sum_{T \subseteq [0,d]} \sum_{j_1 + \ldots + j_{|T|} = s} \binom{s}{j_1, \ldots, j_{|T|}} x^{i_1 j_1 + \ldots + i_{|T|} j_{|T|}}, \end{aligned}$$

where $T = \{i_0, i_1 \ldots i_{|T|}\}$. By the generalized Lucas' identities, a multinomial $\binom{s}{j_1,\ldots,j_t}$ is non-zero only when $j_l \prec s$ for all $1 \leq l \leq t$ and for any l, m it is the case that $j_{li} + j_{mi} \leq s_i$ for all $0 \leq i \leq d$. In other words, this multinomial is non-zero if there exists a function $\pi : [0,d] \to T$ such that $j_l = \sum_{r \in \pi^{-1}(i_l)} 2^r$, for all $1 \leq l \leq |T|$. Also, each function $\pi : [0,d] \to T$ gives rise to a set of degrees $j_1, \ldots, j_t$ s.t. $\binom{s}{j_1,\ldots,j_t} = 1$. Let $\mathcal{F}_T$ be the set of all functions $\pi : [0,d] \to T$.

Hence,

$$\begin{aligned} P_s(x) &= \sum_{T \subseteq [0,d]} \sum_{j_1 + \ldots + j_{|T|} = s} \binom{s}{j_1, \ldots, j_{|T|}} x^{i_1 j_1 + \ldots + i_{|T|} j_{|T|}} \\ &= \sum_{T \subseteq [0,d]} \sum_{\pi \in \mathcal{F}_T} x^{\sum_{t \in T} t \sum_{r \in \pi^{-1}(t)} 2^r}. \end{aligned}$$

Suppose now that $|T| \leq d$ and for some $\pi \in \mathcal{F}_T$ let the formal degree [2] $Q_{T,\pi} = \sum_{t \in T} t \sum_{r \in \pi^{-1}(t)} 2^r$.

Let $T \subseteq T'$ and $\pi \in \mathcal{F}_T$. Then there exists a unique $\pi' \in \mathcal{F}_{T'}$ such that $\pi' = \pi$

[2] by formal we mean that the terms in the summation are exactly the same

(on the entire domain $[0, d]$.) This implies that the formal degree $Q_{T,\pi}$ appears in Equation 5.1 exactly once for each set $T' \supseteq T$, hence it will be counted an even number of times, and thus every such degree vanishes.

The only monomials that remain in the summation are those degrees of the form $Q_{[0,d],\pi}$ where $\pi \in \mathcal{F}_{[0,d]}$ is a bijection. It is now clear that the degree $Q = \sum_{i\in[0,d]} i2^{s_i}$ is the maximum degree of $P_s(x)$ and it is uniquely obtained. This concludes that $P_s(x)$ is a non-trivial polynomial of degree at most $s \cdot d$. Since $s \in \mathcal{D}$ this implies that $s < 2^{\frac{d-1}{d}n+d}$. Indeed, one can notice that $\max_{\mathsf{bwt}(\ell)\leq d+1} \text{min-orb}(\ell)$ is a degree ℓ' whose binary representation contains consecutive 1's (mod $2^n - 1$) at distance roughly d from each other.

Finally, define the polynomial

$$L(x) = \prod_{s\in S} P_s(x).$$

To argue that $L(x)$ is a non-null polynomial, we argue that its total degree is at most $2^n - 1$ and thus no further cancellation can occur from taking modulo $x^{2^n} - x$. To conclude the proof, notice that the sum of maximum degrees in each $P_s(x)$ is a unique degree of L and hence it cannot cancel. Hence, $L's$ total degree is at most $\binom{n}{d} d2^{\frac{d-1}{d}n+d} < 2^n - 1$ for some n large enough.

∎

Chapter 6

Succinct Representation of Codes with Applications to Testing

In the previous chapter we motivated and introduced the single orbit property, and we showed explicit single orbit tests for a few common families of codes. Since this property leads to succinct representations and to nice testing results, namely structured testing, it is important to understand when it is the case that low weight single orbit generators can be encountered in more general settings.

In this chapter we describe our main results in this direction by presenting a set of sufficient conditions that imply single orbit under the affine group $\mathrm{AGL}(1, 2^n)$. In addition we study an even smaller group of permutations, namely permutations under the linear group $\mathrm{GL}(1, 2^n)$. We remark that in this section we are not concerned with providing explicit support for the generators of our families. Our results here focus on families where it appears to be hard to exhibit such fully explicit tests. We also note that while single orbit under affine invariance implies structured testing, there are no known testing implications for families whose duals have single orbit under cyclic groups.

Again, we concentrate on families of binary functions over the domain $\mathbb{F}_{2^n}$. Our first result states that if a family is affine invariant and it contains only a polynomial number of functions (i.e. it is 'sparse') then its dual is generated by a function of

low weight and by its permutations under the affine group. We require that n be prime, a technical consequence of results from number theory that we make use of. These families of functions correspond to codes of very small rate (i.e. *poly* $n/2^n$) and large relative distance (i.e. $1/2 - 2^{-n\epsilon}$.) Therefore their duals are very dense codes, and as it turns out they have very small distance. Our results imply that the duals of sparse, affine invariant codes can be specified succinctly by only $O(n)$ bits - the support elements of a generator. Notice that the dimension of such a dual is $\Omega(2^n)$, and hence a priory in order to specify it one would need to describe $\Omega(2^n)$ basis vectors.

Cyclic invariance is a generalization of affine invariance. Every punctured affine-invariant code (i.e. puncturing means removing one coordinate) is a cyclic code. It is also common to define a cyclic code as an ideal generated by a univariate polynomial in $\mathbb{F}_2[x]$, as commented in Chapter 3. Such an ideal generator exactly corresponds to a generator of the code as a vector space. Hence, every cyclic family has a single-orbit generator function, but it may not have a low weight single orbit generator.

Our results here consider functions defined over the domain $\mathbb{F}_{2^n}^*$. We show that duals of sparse, cyclic invariant families are also generated by a small weight function. In this case we require that not only n be prime, but also $2^n - 1$ have no large non-trivial divisors (in particular Mersenne primes satisfy this condition.)

An immediate application of our results is in the study of BCH codes. Since the duals of BCH codes (and eBCH codes) are sparse and cyclic (or affine, respectively) invariant it follows that BCH (and eBCH) codes are generated by a small weight codeword. These findings improve on the previous state of knowledge regarding their structure in terms of their low weight codewords.

As previously discussed, the single orbit property has applications in testing. In particular, our results imply that general sparse, affine invariant families are testable by strong, structured testers. This gives a large class of testable properties that contains the Hadamard and dual-BCH codes. Our results lead to a couple of conjectures which, if true, would complete the characterization of testable affine invariant properties.

6.1 Main results and implications

First recall the definition of the single orbit property (Definition 38 from Chapter 5.) Also, we will use $N = 2^n$. We now state our results more formally. Our first result considers affine-invariant families.

Theorem 50 (Single orbit property in affine-invariant families) *For every $t > 0$ there exists a $k = k(t)$ such that for every prime n the following holds. Let $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be a linear affine-invariant family containing at most 2^{nt} functions. Then $\mathcal{C}^\perp$ has the k-single orbit property under the affine group* $\mathrm{AGL}(1, 2^n)$.

Next we state our main theorem for cyclic families. We present it here in a general form and then we mention an immediate corollary.

Theorem 51 (Single orbit property in cyclic families) *For every t and $\epsilon > 0$ there exists a $k = k(t, \epsilon)$ such that the following holds. Let n be a prime such that $2^n - 1$ does not have nontrivial divisors larger than $2^{n(1-\epsilon)}$. Let $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be a linear, cyclic invariant family with at most 2^{nt} functions. Then $\mathcal{C}^\perp$ has the k-single orbit property under the cyclic group* $\mathrm{GL}(1, 2^n)$.

We remark that the condition that $2^n - 1$ does not have large divisors implies that, if $2^n - 1$ is not a prime then all its prime divisors are somewhat large (larger than $2^{n\epsilon}$,) and in particular $2^n - 1$ has only a few prime divisors. The following corollary is a simple consequence of our theorem.

Corollary 52 (Single orbit property in cyclic families) *For every t there exists a $k = k(t)$ such that the following holds. Let n be such that $2^n - 1$ is prime. Let $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be a linear, cyclic invariant family with at most 2^{nt} functions. Then $\mathcal{C}^\perp$ has the k-single orbit property under the cyclic group.*

It is not known if there are infinitely many n such that $2^n - 1$ is prime. Nevertheless, as things stand, the question of whether the number of such primes is infinite or not is unresolved (and indeed there are conjectures suggesting there are infinitely many such primes).

6.1.1 Implications to property testing

As mentioned in Chapter 5, it follows from the work of [72] that codes with a single local orbit under the affine symmetry group are locally testable.

Our main theorem, Theorem 50, when combined with Theorem 40 (stated in Chapter 5,) immediately yields the following implication for sparse affine invariant families.

Corollary 53 *For every constant t there exists a constant k such that if $\mathcal{C} \subseteq \{\mathbb{F}_2^{2^n-1} \to \mathbb{F}_2\}$ is a linear, affine-invariant family with at most 2^{nt} functions, then $\mathcal{C}$ is $(k, \Omega(1/k^2))$-locally testable by a structured tester.*

6.1.2 Implications to BCH codes

In addition to the implications for the testability of sparse affine-invariant families/codes, our results also give new structural insight into the classical BCH codes. Even though these codes have been around a long time, and used often in the CS literature, some very basic questions about them are little understood.

Recall from Chapter 5 that $\mathrm{BCH}(t,n)^\perp = \{\langle \mathrm{Trace}(f(\alpha))\rangle_{\alpha \in \mathbb{F}_{2^n}^*} | f \in \mathbb{F}_{2^n}[x], \deg(f) \leq 2t\}$ and $\mathrm{BCH}(t,n) = \{\langle \mathrm{Trace}(f(\alpha))\rangle_{\alpha \in \mathbb{F}_{2^n}^*} | f \in \mathbb{F}_{2^n}[x], \deg(f) \leq 2^n - 2t - 1\}$.

In the previous chapter we showed an explicit single orbit basis for the restricted family of $\mathrm{BCH}(2,n)$ codes. While explicitness would be a nice feature to exhibit in general BCH codes, we do not address this question here. However, in this chapter we make progress in showing the existence of a succinct basis consisting of affine transformations (and in some cases cyclic transformations) of a low-weight function for BCH codes.

Corollary 54 *For every t there exists a k such that for all prime n, $\mathrm{eBCH}(t,n)$ has the k-single orbit property under the affine group.*

The above follows from Theorem 50 using the observation that $\mathrm{eBCH}(t,n)^\perp$ is sparse (has $N^{O(t)}$ codewords) and affine invariant.

Corollary 55 *For every t and ϵ, there exists a k such that for all n prime such that $2^n - 1$ does not contain any nontrivial divisors larger than $2^{n(1-\epsilon)}$, $\mathrm{BCH}(t, n)$ has the k-single orbit property under the cyclic group.*

The above follows from Theorem 51 using the observation that $\mathrm{BCH}(t, n)^{\perp}$ is sparse (has $N^{O(t)}$ codewords) and cyclic invariant.

Finally, we point out that the need for various parameters being prime or having small divisors (n and $2^n - 1$, respectively) is a consequence of the application of some recent results in additive number theory that we use to show that certain codes have very high distance. We do not believe such assumptions ought to be necessary, however we do not see any immediate path to resolving the "stronger" number-theoretic questions that would arise by allowing n to be non-prime.

6.2 Overview of techniques and helpful lemmas

Our main theorems are proved essentially by implementing the following plan:

1. Using the description of affine and cyclic invariant families as polynomials (from Chapter 3), we notice that sparse codes correspond to Traces of *sparse* polynomials.

2. We then apply the recent results in additive number theory to conclude that these families have very high distance. This already suffices to show that sparse affine-invariant families are testable by [71]. However the tests given there are arbitrary and we need to work further to get structured tests for these families, or to show the single-orbit condition.

3. The final, and the novel part of this work, is to show by a counting argument, that there exists one (in fact many) low-weight functions in the dual of the functions we consider such that their orbit spans the dual.

In order to elaborate on these steps, we will need to recall some definitions Chapter 3. For $d \in \{1, \ldots, N-2\}$, $\text{orb}(d) = \{d, 2d(\text{mod } N-1), 4d(\text{mod } N-1), \ldots, 2^{n-1}d(\text{mod } N-1)\}$ and min-orb(d) denotes the smallest integer in orb(d). Also $\mathcal{D} = \{\text{min-orb}(d) \mid d \in \{1, \ldots, N-2\}\} \cup \{N-1\}$, and

$$P_{N,D} = \{\alpha_0 + \sum_{d \in D} \alpha_d x^d \mid \alpha_d \in \mathbb{F}_N, \alpha_0, \alpha_{N-1} \in \{0,1\}\},$$

$$\text{and } P_{N-1,D} = \{\sum_{d \in D} \alpha_d x^d \mid \alpha_d \in \mathbb{F}_N, \alpha_{N-1} \in \{0,1\}\}.$$

Also recall the description of cyclic and affine invariant families from Propositions 11 and 16. Namely, every cyclic family is characterized by a set $D \subseteq \mathcal{D}$ such that $c \in \mathcal{C}$ if and only if there exists a polynomial $p \in P_{N-1,D}$ such that $c(x) = \text{Trace}(p(x))$ for every $x \in \mathbb{F}_N^*$. Similarly, every affine invariant family is described by a set of degrees $D \subseteq \mathcal{D}$ such that $c \in \mathcal{C}$ if and only if there exists a polynomial $p \in P_{N,D}$ such that $c(x) = \text{Trace}(p(x))$ for every $x \in \mathbb{F}_N$.

The first part of the proof is described by the next lemma.

Lemma 56 *For every cyclic-invariant family $\mathcal{C} \subseteq \{\mathbb{F}_N^* \to \mathbb{F}_2\}$ with $|\mathcal{C}| \leq N^t$, the set of degrees $D \subseteq \mathcal{D}$ that characterizes $\mathcal{C}$ satisfies $|D| \leq t$.*

Similarly, for every affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_N \to \mathbb{F}_2\}$ of cardinality N^t, the set of degrees D that characterizes $\mathcal{C}$ satisfies $|D| \leq t$ and $D \subseteq \{1, \ldots, N^{1-1/t}\}$.

Thus in both cases the families/codes are represented by collections of t-sparse polynomials. And in the affine-invariant case, these are also somewhat low-degree polynomials. In what follows we use $\mathcal{C}_N(D)$ to denote the code $\{\text{Trace}(p(x)) | p \in P_{N,D}\}$ and $\mathcal{C}_{N-1}(D)$ to denote the code $\{\text{Trace}(p(x)) | p \in P_{N-1,D}\}$.

Bourgain's bounds We next use a (small variant of a) theorem due to Bourgain [36] to conclude that the families $\mathcal{C}_N(D)$ and $\mathcal{C}_{N-1}(D)$ have very high distance (under the given conditions on D).

Theorem 57 ([36]) *For every $\epsilon > 0$ and $r < \infty$, there is a $\delta = \delta(\epsilon, r) > 0$ such that for every prime n the following holds. Let $N = 2^n$ and $\mathbb{F} = \mathbb{F}_N$ and let $f(x) = \sum_{i=1}^{r} a_i x^{k_i} \in \mathbb{F}[x]$ with $a_i \in \mathbb{F}$, satisfy*

1. $1 \leq k_i \leq N - 1$
2. $\gcd(k_i, N-1) < N^{1-\epsilon}$ *for every* $1 \leq i \leq r$
3. $\gcd(k_i - k_j, N-1) < N^{1-\epsilon}$ *for every* $1 \leq i \neq j \leq r$

Then

$$\left| \sum_{x \in \mathbb{F}} (-1)^{\mathrm{Trace}(f(x))} \right| < N^{1-\delta}.$$

We note that strictly speaking, [36, Theorem 7], only considers the case where N is prime, and considers the sum of any character from $\mathbb{F}$ to the complexes (not just $(-1)^{\mathrm{Trace}(\cdot)}$). We note that the proof extends to cases where $N = 2^n$ where n is prime as well. We comment on the places where the proof in [36] (and related papers) have to be changed to get the result in our case, in Section 6.5. The proof of Bourgain's theorem uses a heavy number theoretic machinery and recent results in additive combinatorics, and it builds on similar results for even sparser polynomials [37, 40, 38].

In our language the above theorem implies that functions represented by traces of sparse polynomials of somewhat low-degree have many non-zeros. Even when the degree is large, but when in addition to n being prime we have that the gcd of the degrees of the polynomial with $N - 1$ is not too large, then we again get functions with a large number of non-zero values in the field.

We thus obtain the following implication.

Lemma 58 *For every t and $\epsilon > 0$ there exists a δ such that the following holds for every $N = 2^n$ for prime n. Let $\mathcal{D} = \mathcal{D}(N)$ and let $D \subseteq \mathcal{D}$ be of size at most t and such that $\max_{d \in D} d < N^{1-\epsilon}$. Then the family $\mathcal{C} = \mathcal{C}_N(D)$ satisfies $\frac{1}{2} - N^{-\delta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\delta}$.*

	degree	sparsity of f	bound
W-C-U [98, 41]	$d_2 = 0$	any	$N/2 \pm d_1 N^{1/2}$
B [36]	$d_2 < N^{1-\epsilon}$	constant	$N/2 \pm N^{1-\delta}$
B [37]	$d_1 < N^{1/2}$ $d_2 < N^{1-\epsilon}$	for g, any for h, constant	$N/2 \pm N^{1-\delta}$

Table 6.1: Weil-Carlitz-Uchiyama and Bourgain bounds on the number of non-zeros of Trace(f); $f(x) = g(x) + h(x)$ with $\deg g = d_1 < N^{1/2}$, $\deg h = d_2$ and $N^{1/2} < \min \deg h$.

Similarly, for every t and $\epsilon' > 0$ there exists a δ such that the following holds for for every $N = 2^n$ such that $N - 1$ does not have any nontrivial divisor larger than $N^{1-\epsilon'}$. Let $\mathcal{D} = \mathcal{D}(N)$ and let $D \subseteq \mathcal{D}$ be of size at most t. Then the $\mathcal{C} = \mathcal{C}_{N-1}(D)$ satisfies $\frac{1}{2} - N^{-\delta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\delta}$.

We remark that Bourgain's theorem above is a generalization of the widely used Weil-Carlitz-Uchiyama bounds [98, 41], which in particular can be employed in estimates of the distance of BCH codes. Those initial bounds however fail to give interesting estimates when the degrees of the polynomials inside the trace are larger than roughly $N^{\frac{1}{2}}$. Since general cyclic/affine invariant codes/families could be characterized by degrees much larger than this, Bourgain's estimates turn to be greatly applicable in obtaining our results. See Table 6.2 for a quick comparison between the bounds.

Main argument We now move to the crucial part of the paper where we attempt to use counting style arguments to claim that the codes we are considering have the single orbit property for small k. Here our plan is as follows.

We first use a result from [71] to show that for any specific family $\mathcal{C}$ we consider and for every sufficiently large k, its dual has roughly $\binom{N}{k}/|\mathcal{C}|$ functions of weight k (this bound is tight to within $1 \pm \Theta(1/N^c)$ factor, for large enough k (where k is independent of N and depends only on t, c and the δ of Lemma 58). Specifically they show:

Theorem 59 ([71] Lemma 3.5) *For every $c, t < \infty$ and $\delta > 0$ there exists a k_0 such that for every $k \geq k_0$ and for every family $\mathcal{C} \subseteq \{\mathbb{F}_N \to \mathbb{F}_2\}$ with at most N^t*

functions, and satisfying $\frac{1}{2} - N^{-\delta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\delta}$ *it is the case that* $\mathcal{C}^{\perp}$ *has* $\binom{N}{k}/|\mathcal{C}| \cdot (1 \pm \theta(N^{-c})$ *functions of weight* k.

Thus for any family $\mathcal{C} = \mathcal{C}(D)$ under consideration, this allows us to conclude that $\mathcal{C}^{\perp}$ has many functions of weight k (for sufficiently large, but constant k). What remains to be shown is that the orbit of one of these, under the appropriate group (affine or cyclic) contains a basis for the whole code $\mathcal{C}^{\perp}$. To do so, we consider any function x of weight k in the dual whose orbit under the group does *not* contain a basis for $\mathcal{C}^{\perp}$ (i.e., $\text{Span}(\{x \circ \pi | \pi\}) \neq \mathcal{C}^{\perp}$). We show that for every such word x there is a set $D' \subseteq \mathcal{D}$ of size $|D'| = |D| + 1$ such that $x \in \mathcal{C}(D')^{\perp}$. The size of $\mathcal{C}(D')$ is roughly a factor of N larger than the size of $\mathcal{C}$ and thus $\mathcal{C}(D')^{\perp}$ is smaller than $\mathcal{C}^{\perp}$ by a factor of roughly N. We argue further that this family $\mathcal{C}(D')$ also satisfies the same invariant structure as $\mathcal{C}$ and so one can apply Lemma 58 and Theorem 59 to it. We can thereby conclude that the number of weight k functions in $\mathcal{C}(D')^{\perp}$ are also smaller than the number weight k functions in $\mathcal{C}^{\perp}$ by a factor of approximately N. Finally, we notice that the number of sets D' to consider is smaller than the factor between the number of functions of weight k in $\mathcal{C}^{\perp}$ and $\mathcal{C}(D')^{\perp}$. That lets us conclude that the set $\cup_{D'}\mathcal{C}(D')^{\perp}$ can not include all possible weight k functions in $\mathcal{C}^{\perp}$, yielding the k-single orbit property for $\mathcal{C}$. This leads to the proofs of Theorem 50 and 51 that appear in Section 6.4.

6.3 Proofs of the helpful lemmas

We now prove Lemma 56 and Lemma 58.

Proof of Lemma 56.: For the cyclic invariant case, the lemma is immediate. Indeed, by Proposition 11 if $\mathcal{C}$ is cyclic invariant and it is characterized by some $D \subseteq \mathcal{D}$, then $\mathcal{C} = \mathcal{C}_{N-1}(D) = \{\text{Trace}(p) | p \in P_{N-1,D}\}$. For every pair of functions it is the case that if $p_1 \neq p_2 \in P_{N-1,D}$ then $\text{Trace}(p_1) \neq \text{Trace}(p_2)$. Hence $|\mathcal{C}| = |P_{N-1,D}| \geq N^{|D|}$ yielding $|D| \leq t$ if $|\mathcal{C}| \leq N^t$.

We now consider the affine invariant case. Consider an affine-invariant family $\mathcal{C}$, which by Proposition 16 is described by the set $D \subseteq \mathcal{D}$ such that $\mathcal{C} = \mathcal{C}_N(D) = \{\text{Trace}(p) | p \in P_{N,D}\}$. As above we also have $|D| \leq t$ if $|\mathcal{C}| \leq N^t$. It remains to be shown that $D \subseteq \{1, \ldots, N^{1-1/t}\}$.

We now use the fact that the set D is shadow-closed, i.e., if $d \in D$ and $e \prec d$ then $e \in D$.

Consider the binary weight of the integers $d \in D$. We claim that for every integer $d \in D$, its binary weight is at most t (or else its shadow and hence D has more than t elements). It follows that the integer $d = \text{min-orb}(d) \leq 2^{n(1-1/t)} = N^{1-1/t}$. Since this holds for every $d \in D$, we conclude that $D \subseteq \{1, \ldots, \lfloor N^{1-1/t} \rfloor\}$. This yields the proof of Lemma 56 for the affine-invariant case. ∎

Proof of Lemma 58: For $p \in P_{N,D}$ such that $\text{Trace}(p) \in \mathcal{C}$ define for the purpose of this proof $\delta(p) = Pr_{x \in \mathbb{F}_N}[\text{Trace}(p(x)) = 1]$. Since the degrees in D are upper bounded by $N^{1-\epsilon}$, by Theorem 57 there exists $\delta' = \delta'(t, \epsilon)$ such that $|\sum_{x \in \mathbb{F}_N} (-1)^{\text{Trace}(p(x))}| < N^{1-\delta'}$. Since $\mathbb{E}_{x \in \mathbb{F}_N}(-1)^{\text{Trace}(p(x))} = 1 - 2\delta(p)$, it follows that there exists δ such that $\frac{1}{2} - N^{-\delta} \leq \delta(p) \leq \frac{1}{2} + N^{-\delta}$. The first part of the lemma is now immediate by noting that $\delta(\mathcal{C}) = \min_{p \in P_{N,D}} \delta(p)$. The second part follows easily by a similar argument. ∎

6.4 Proofs of the main theorems

We now derive the proofs of the main theorems.

6.4.1 Analysis of the cyclic case

Proof of Theorem 51: Let $\delta = \delta(t, \epsilon)$ and $\delta' = \delta'(t+1, \epsilon)$ be as given by Lemma 58 for the cyclic invariant case (so codes of length $N-1$ have distance roughly $1/2 - N^{-\delta}$). Let $c = 2$ and let $k_0 = k_0(c, t, \delta)$ and $k_0' = k_0(c, t+1, \delta')$ be as given by Theorem 59. We prove the theorem for $k = \max\{k_0, k_0'\}$.

Fix $N = 2^n$ such that n is prime and $N-1$ does not have any non-trivial divisor larger than $N^{1-\epsilon}$. Let $\mathcal{C} \subseteq \{\mathbb{F}_N^* \to \mathbb{F}_2\}$ be a cyclic code of cardinality at most N^t. Let $D \subseteq \mathcal{D}$ be the set of degrees that describes $\mathcal{C}$ as given by Proposition 11 so that $\mathcal{C} = \{\text{Trace}(p) | p \in P_{N-1,D}\}$. For $d \in \mathcal{D} - D$, let $\mathcal{C}(d) = \{\text{Trace}(p) | p \in P_{N-1,D\cup\{d\}}\}$. Our analysis below will show that (1) Every function $w \in \mathcal{C}^\perp - \cup_{d\in\mathcal{D}-D}(\mathcal{C}(d)^\perp)$ generates the family $\mathcal{C}^\perp$ by its cyclic shifts, i.e., $\mathcal{C}^\perp = \text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}$, and (2) There is a function of weight k in $\mathcal{C}^\perp - \cup_{d\in\mathcal{D}-D}(\mathcal{C}(d)^\perp)$. Putting the two together we get the proof of the theorem.

We start with the first part. Consider any function $w \in \mathcal{C}^\perp$. We claim that if $\text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}\} \neq \mathcal{C}^\perp$, then there must exist an element $d \in \mathcal{D} - D$ such that $w \in \mathcal{C}(d)^\perp$. To see this, first note that $\text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}\}$ is a family invariant under the cyclic group, and is contained in $\mathcal{C}^\perp$. Thus if $\text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}\} \neq \mathcal{C}^\perp$ then it must be strictly contained in $\mathcal{C}^\perp$. Therefore $(\text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}\})^\perp$ must be a strict superset of $\mathcal{C}$. Using Proposition 11 there must exist a set D' such that $(\text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}\})^\perp = P_{N-1,D'}$. Furthermore D' must be a strict superset of D and so there must exist an element $d \in D' - D$. We claim that $w \in \mathcal{C}(d)^\perp$. This is so since $\mathcal{C}(d) \subseteq (\text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}\})^\perp$ and so $w \in (\text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}\}) \subseteq \mathcal{C}(d)^\perp$. This concludes the proof of the first claim.

It remains to show that there is a function of weight k in $\mathcal{C}^\perp - \cup_{d\in\mathcal{D}-D}(\mathcal{C}(d)^\perp)$. For this we employ simple counting arguments. We first note that, using Lemma 56 we must have that $|D| \leq t$. Further, by Lemma 58 we obtain that $\mathcal{C}$ is a code satisfying $\frac{1}{2} - N^{-\delta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\delta}$. Hence we can apply Theorem 59 to conclude that $\mathcal{C}^\perp$ has at least $\binom{N}{k}/(|\mathcal{C}|) \cdot (1 - O(1/N^2))$ functions of weight k.

On the other hand, for every fixed $d \in \mathcal{D} - D$, we have (by Lemma 56 and Lemma 58 again) $\frac{1}{2} - N^{-\delta'} \leq \delta(\mathcal{C}(d)) \leq \frac{1}{2} + N^{-\delta'}$. Again applying Theorem 59 we have $\mathcal{C}(d)^\perp$ has at most $\binom{N}{k}/(|\mathcal{C}(d)|)(1+O(1/N^2))$ functions of weight k. In case $d = N-1$, then $|\mathcal{C}(d)| = 2 \cdot |\mathcal{C}|$. In case $d \neq N-1$ then $|\mathcal{C}(d)| = N \cdot |\mathcal{C}|$. Thus we can bound the total

number of functions of weight k in $\cup_{d\in\mathcal{D}-D}\mathcal{C}(d)^{\perp}$ from above by

$$\frac{\binom{N}{k}}{(2\cdot|\mathcal{C}|)}(1+O(1/N^2))+|\mathcal{D}|\cdot\frac{\binom{N}{k}}{(N\cdot|\mathcal{C}|)}(1+O(1/N^2))\leq$$

$$\frac{1}{2|\mathcal{C}|}\cdot\binom{N}{k}(1+1/\log_2 N+O(1/N^2)),$$

where above we use the fact that $|\mathcal{D}|\leq N/\log_2 N$.

For sufficiently large N (i.e., when $1/\log_2 N+O(1/N^2)\leq 1/2$) we have that this quantity is strictly smaller than $\frac{\binom{N}{k}}{(|\mathcal{C}|)}\cdot(1-O(1/N^2))$, which was our lower bound on the number of functions of weight k in $\mathcal{C}^{\perp}$. We conclude that there is a function of weight k in $\mathcal{C}^{\perp}-\cup_{d\in\mathcal{D}-D}(\mathcal{C}(d)^{\perp})$ as claimed.

This concludes the proof of the theorem. ∎

6.4.2 Analysis of the affine-invariant case

Proof of Theorem 50: The proof is similar to the proof of Theorem 51 with the main difference being that we need to argue that the polynomials associated with functions in $\mathcal{C}$ and $\mathcal{C}(d)$ are of somewhat low-degree (to be able to conclude that they have high-distance).

Given t, let δ be from Lemma 58, where $\epsilon'=1/t$, and let k be large enough for application of Theorem 59. Fix $N=2^n$ for prime n and and let $\mathcal{C}$ be an affine-invariant family of cardinality N^t. Let $D\subseteq\mathcal{D}$ be a set of cardinality at most t and consisting of integers smaller than $N^{1-1/t}$ such that $\mathcal{C}=\{\text{Trace}(p)|p\in P_{N,D}\}$ (as given by Proposition 16). For $d\in\mathcal{D}-D$, let $\mathcal{C}(d)=\{\text{Trace}(p)|p\in P_{N,D\cup\{d\}}\}$.

Let $\mathcal{D}'=(\mathcal{D}-D)\cap\{1,\ldots,\lfloor N^{1-1/t}\rfloor\}$.

Similar to the proof of Theorem 51 we argue that if there is a weight k function w in $\mathcal{C}^{\perp}$ that is not in some $\mathcal{C}(d)^{\perp}$, but now only for every $d\in\mathcal{D}'$, then $\{\text{Span}(w(\alpha x+\beta)|\alpha\in\mathbb{F}_N^*,\beta\in\mathbb{F}_N\}=\mathcal{C}^{\perp}$. The same counting argument as in the proof of Theorem 51 suffices to show that such a function does exist.

Consider $w \in \mathcal{C}^{\perp}$ and the family $\{\text{Span}(w(\alpha x+\beta)|\alpha \in \mathbb{F}_N^*, \beta \in \mathbb{F}_N\}$. $\{\text{Span}(w(\alpha x+\beta)\}$ is affine invariant and so is given by $P_{N,E}$ for some shadow-closed set E. If $\{\text{Span}(w(\alpha x+\beta)\}^{\perp} \neq \mathcal{C}$ then E strictly contains D and so there must exist some element $d' \in E-D$. Now consider the smallest binary weight element $d \prec d'$ such that $d \in E-D$. We claim that the binary weight of d must be at most $t+1$ (since elements of D have binary weight at most t). We then conclude that $w \in \{\text{Span}(w(\alpha x+\beta)\} \subseteq \mathcal{C}(d)^{\perp}$ yielding the claim.

The counting argument to show there is a function of weight k in $\mathcal{C}^{\perp} - (\cup_{d \in \mathcal{D}'} \mathcal{C}(d)^{\perp}$ is now same as in the proof of Theorem 51 except that we use the affine-invariant part of Lemma 56 and Lemma 58.

This completes the proof of Theorem 50. ∎

6.5 On using results from additive number theory

As pointed out earlier Theorem 7 of [36] only considers the analog of Theorem 57 where the field $\mathbb{F}$ is of prime cardinality N, and shows that for any additive character χ, $|\sum_{x \in \mathbb{F}} \chi(f(x))| \leq N^{1-\delta}$. Here we mention the modifications necessary to extend the proof to the case where $\mathbb{F}_N$ is of cardinality 2^n with n being prime.

In [36] the proof reduces to the two cases $r=1$ and $r=2$. The case $r=1$ in the prime case was obtained in [40]. In our case, where $N=2^n$, the $r=1$ case was shown in [38]. For $r=2$ the proof in the prime case applied the sum-product theorem from [39] and uses Proposition 1 of [37]. We note that Proposition 1 of [37] works also when the field is not of prime cardinality. As argued in [39], the sum-product statement might weaken for more general fields only when the field $\mathbb{F}_N$ contains somewhat large subfields. However, when n is prime $\mathbb{F}_{2^n}$ contains only the constant size base field $\mathbb{F}_2$. We conclude that when $\mathbb{F}=\mathbb{F}_{2^n}$ (n prime) it remains true that if a set $A \subset \mathbb{F}_N$ has size $1 < |A| < N^{1-\epsilon}$ for some given ϵ then $|A+A| + |A \cdot A| > C|A|^{1+\delta}$, for some $\delta = \delta(\epsilon)$. The key ingredient of the proof in [37] is an additional sum-product theorem in the additive/multiplicative group $\mathbb{F}_N \times \mathbb{F}_N$ with N prime, where addition

and multiplication are defined coordinate-wise. The equivalent formulation for our case $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ follows exactly as in [37], and so does the rest of the proof.

6.6 Discussion

Based on the alternate definition of affine/cyclic codes from Chapter 3 as polynomial ideals, our results can be reformulated as describing relations between the sparsity of a polynomial over $\mathbb{F}_2[x]$ (i.e. its number of monomials) and the number of its roots over $\mathbb{F}_{2^n}$.

Theorem 51 essentially says that for any small, arbitrary set $D \subset \mathcal{D}$ (under the specified restriction of n and $2^n - 1$) there exists a polynomial $p \in \mathbb{F}_2[x]/\langle x^{2^n-1} - 1\rangle$ of monomial sparsity at most some $k = k(|D|)$, such that the roots of p in $\mathbb{F}_{2^n}$ are exactly w^d, where $d \in \text{orb}(D)$. Such a polynomial generates a cyclic code characterized by the set of roots $\{w^d | d \in \text{orb}(D)\}$. We note that even when $|D| = 1$ this is not known for general n [49, 96], and understanding the relation between sparsity and number of roots for such polynomials seems to be an important open question, with applications to the theory of cyclic codes.

We next summarize a few immediate open questions suggested by the results presented here.

1. Can Theorems 50 and 51 be extended to non restricted block lengths? Progress in this direction requires either stronger number theoretic versions of the Bourgain theorem we use or a somewhat different approach.

2. Our results give sufficient conditions for binary codes to be generated by a small weight function. Codes that have a constant number of small weight generators under affine transformations are also testable: simply pick random affine transformations and verify that the permuted generators have inner product 0 with the given function. It would be interesting to provide non-trivial sufficient

conditions for codes to have a constant number of low weight single orbits that generate the code.

3. What other groups of permutations could be relevant in the study in boolean functions? Codes that are invariant under larger groups that contain $\mathrm{AGL}(1, 2^n)$ (in particular, codes invariant under 2-transitive groups) could lead to testing implications. These groups have been described explicitly in [28].

Chapter 7

Conclusions and Future Directions

7.1 Towards a full characterization of affine invariant codes

The results of this thesis as well as further supporting evidence [25] lead to the following conjecture which would complete the characterization of testable linear families that are invariant under the affine group $\mathrm{AGL}(1, 2^n)$.

Conjecture 60 *Let $k > 0$ be an integer and let $\mathcal{F} \subseteq \{F_{2^n} \to \mathbb{F}_2\}$ be a linear, and affine invariant family, where n is a prime. Then $\mathcal{F}$ is k-locally testable if and only if the set of degrees that characterize the family is $D = R \cup S$, where the set R characterizes a Reed Muller code of order $r < k$ and S is $g(k)$-sparse (for some function $g : \mathbb{N} \to \mathbb{N}$ independent of n).*

Our Corollary 53 from Chapter 6 confirms this conjecture from the positive perspective.

A few other examples confirm the conjecture from the negative side. A first observation comes from the examples of Reed Muller codes. The dual of RM codes of order d has distance 2^{d+1}, which immediately implies that RM codes of order, say $\omega(1)$, cannot be tested locally.

A second piece of supporting evidence comes from our Lemma 32, which essentially states that certain non-sparse affine invariant families included in $\mathrm{RM}(2,n)$ are not testable. In particular, families described by the set of degrees $D_t = \{2^\ell + 1 \mid 0 \leq \ell \leq t\}$ are such that their duals do not contain functions of weight $\leq t-2$ other than those functions that also belong to $\mathrm{RM}(2,n)^\perp$. Hence, for $t = \omega(1)$ this gives a non-testable family. It also suggests the following conjecture in the same vein.

Conjecture 61 *Let $D_t = \{2^\ell + 1 \mid \ell \in L \text{ and } |L| \leq t\}$ and let $\mathcal{F}$ be the affine invariant family $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ with n prime, characterized by the set of degrees D_t. Then if $f \in \mathcal{F}^\perp - \mathrm{RM}(2,n)^\perp$ it must be the case that $\mathrm{wt}(f) \geq g(t)$, for some strictly increasing function $g : \mathbb{N} \to \mathbb{N}$.*

This conjecture might be key to a similar more general statement that would immediately imply one direction of the Conjecture 60.

Conjecture 62 *Let n be prime and let $D_t = R(d) \cup S_t \subseteq R(d')$ be a shadow closed set such that the degrees in $R(d)$ (and $R(d')$) characterize the Reed Muller code of order d (and d', respectively) and $|S| \leq t$. Then if $f \in \mathcal{F}^\perp - \mathrm{RM}(d',n)^\perp$ it must be the case that $\mathrm{wt}(f) \geq g(t,d')$, for some function $g : \mathbb{N} \to \mathbb{N}$ strictly increasing in t.*

Also pertinent to the role of sparsity in testing affine-invariant families, Ben-Sasson and Sudan in [25] show that if an affine-invariant family is characterized by a set that contains a degree of binary weight t then its dual cannot have functions with weight smaller than $t-2$. This immediately implies that affine families that are $N^{\log N^{\omega(1)}}$-sparse are not testable with a constant number of queries.

The extent to which sparsity determines testability was also studied by Kaufman and Sudan in [71] where they show that sparse codes of large distance are testable. Kopparty and Saraf [76] conjectured that if a linear code is sparse then it is testable, regardless of distance. That conjecture was however disproved by Ben-Sasson and Viderman [26].

7.2 Further related work in testing non-linear, linear-invariant properties

Linear invariance is a common property of many natural collections of boolean functions. Testing dictators [86], juntas [46, 32], halfspaces [82], concise representation [43], Fourier sparsity and dimensionality [58] are just a few other linear invariant families well studied in the literature. These families are not linear and, it turns out that devising and analyzing tests for non-linear families have so far required techniques somewhat broadly different from those used in analyzing linear families that are invariant under linear transformations. Such techniques include Fourier analytic tools and even a machinery specific to learning theory.

As posed by Sudan [93], the question of finding a unifying proof for all testable properties that are invariant under linear transformations is an important open direction of further investigation. A promising perspective comes from analyzing families of functions that are *free* of prescribed patterns. This view is somewhat complementary to that taken in analyzing linear properties, in the following sense. A linear property $\mathcal{P} \subseteq \{f : D \to \mathbb{F}_2\}$ is characterized by a set of vectors in the dual family that generate the dual. Let $F = \{f_1, f_2, \ldots, f_d\}$ be a set of functions generating $\mathcal{P}^\perp$ and let $G = \{\langle \alpha_1, \alpha_2, \ldots, \alpha_k \rangle \in D^k\}$ be the set of supports of all function in F, where k is the maximum size of a support. Then $f \in \mathcal{P}$ if and only if $\sum_i f(\alpha_i) = 0$ for any tuple $\langle \alpha_1, \alpha_2, \ldots, \alpha_k \rangle \in G$. In other words, the tuple $\langle f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_k) \rangle$ can only contain binary patterns with an even number of 1's. Alternately, $\mathcal{P}$ can be defined as the collection of functions that are free of any complementary pattern (i.e. they do not allow patterns that include an odd number of non-zeros at those locations). This particular perspective is amenable to defining non-linear properties, where there is no notion of dual.

Green [59] initiated the study of boolean families defined by forbidding patterns by considering testing *triangle* freeness. More formally, his results show that the family $\mathcal{P} = \{f : \mathbb{F}_2^n \to \mathbb{F}_2 | \langle f(\alpha), f(\beta), f(\alpha + \beta) \rangle \neq \langle 1, 1, 1 \rangle, \quad \forall \alpha, \beta \in \mathbb{F}_2^n\}$ is testable with a

constant number of queries. Remarkably, the technique to prove this result is based on a number theoretic analogy of the Szemeredy regularity lemma for graphs [94], which brought up an entirely new connection between testing properties of boolean functions and graph properties. Follow-up generalizations in this direction were obtained by Shapira [91], Král et al. [77] and Bhattacharya et al.[29].

More recently, Bhattacharyya et al. [31] proposed a conjecture concerning the characterization of all linear invariant properties (under the group $\mathrm{GL}(n,2)$) that are testable with one sided error. Namely, a linear invariant property is testable with one sided error if and only if it is closed under taking restrictions to subspaces. Closure under restrictions to subspaces can be casted in the language of sets being free from solutions to systems of linear equations. This view allows for a great parallelism between algebraic properties and graph properties. Namely, to a large extent, a set free of solutions to a system of linear equations corresponds to a graph that is free of a certain induced subgraph. This analogy allows borrowing techniques from graph property testing [6, 11, 10], and allows the usage of variants of Green's regularity lemma. In [31] we prove that testability with one sided error implies closure under taking subspaces, and moreover we show that freeness from patterns of 'small complexity' lead to testable properties. In upcoming work [30] we attempt to understand the formalism of being free of solutions to systems of equations at a level that allows to argue relations between various classes defined this way.

Appendix A

Missing details from Chapter 5

Theorem 40 is stated in [72] (Theorem 2.9) in a slightly different form and in what follows we show the equivalence between these statements. We start by defining 'formal characterization' from [72] and then describe its relation to the notion of single orbit characterization.

Local affine formal characterization $\mathcal{F} \subseteq \{\mathbb{K} \to \mathbb{F}_2\}$ has a k-local affine formal characterization if there exist integer m, and linear functions $\ell_1, \ell_2, \ldots, \ell_k : \mathbb{K}^m \to \mathbb{K}$ (with $\ell_i(y_1, y_2, \ldots, y_m) = y_1 + \sum_{j=2}^{m} \ell_{ij} y_j$) such that

$$f \in \mathcal{F} \text{ if and only if } \sum_{i=1}^{k} f(\ell_i(y)) = 0 \text{ for all } y \in \mathbb{K}^m.$$

Claim 63 *Let $\mathcal{F} \subseteq \{F_{2^n} \to \mathbb{F}_2\}$ have the k-single orbit property under $\mathrm{AGL}(1, 2^n)$. Then $\mathcal{F}^\perp$ has a k-local affine formal characterization.*

Proof: Let $g \in \mathcal{F}$ be a k-single orbit generator for $\mathcal{F}$, and let $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ be the support of g. Let $\ell_i : \mathbb{F}_{2^n}^2 \to \mathbb{F}_{2^n}$ be defined by $\ell_i(y_1, y_2) = y_1 + \alpha_i y_2$, for all $1 \leq i \leq k$. We show that $f \in \mathcal{F}^\perp$ if and only if $\sum_{i=1}^{k} f(\ell_i(y)) = 0, \quad \forall y \in \mathbb{F}_{2^n}^2$.

Since g is a generator, it follows that $\mathcal{F} = \text{Span}\{g \circ \pi \mid \pi(x) = ax + b, a, b \in \mathbb{F}_{2^n}\}$. Thus if $f \in \mathcal{F}^\perp$ then $\langle f, g \circ \pi \rangle = 0$ for all π. Hence $\sum_{i=1}^{k} f(a\alpha_i + b) = 0$, for all $a, b \in \mathbb{F}_{2^n}$. We only need to show now the opposite direction of the claim. Let $h : \mathbb{F}_{2^n} \to \mathbb{F}_2$ satisfy $\sum_{i=1}^{k} h(\ell_i(y)) = 0, \ \forall y \in \mathbb{F}_{2^n}^2$. We will show that $h \in \mathcal{F}^\perp$. Indeed, for all $a, b \in \mathbb{F}_{2^n}$ $\sum_{i=1}^{k} h(a\alpha_i + b) = 0$ and hence $\langle h, g \circ \pi \rangle = 0, \forall \pi(x) = ax + b$. This immediately implies that $\langle h, f \rangle = 0, \forall f \in \text{Span}\{g \circ \pi \mid \pi(x) = ax + b, a, b \in \mathbb{F}_{2^n}$, and hence $h \in \mathcal{F}^\perp$.

∎

Theorem 2.9 from [72] states that if $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ is linear invariant and has a k-local affine formal characterization then it is k-locally testable. If $f \in \mathcal{F}$ the test accepts, and if f is δ-far from $\mathcal{F}$ the test rejects with probability $\Omega(\min\{\delta/2, 1/k^2\})$. In addition, given the k-local affine formal characterization $\ell_1, \ell_2, \ldots, \ell_k : \mathbb{F}_{2^n}^2 \to \mathbb{F}_{2^n}$ the test picks y_1, y_2 at random and checks whether $\sum_{i=1}^{k} f(\ell_i(y_1, y_2)) = 0$. Hence, if $\mathcal{F}$ has a k-single orbit property under the affine group then it is linear invariant, and by Claim 63 its dual has a k-local affine formal characterization. By Theorem 2.9 from [72] $\mathcal{F}^\perp$ is testable by structured tests in the sense described in Chapter 5. This concludes the proof of the Theorem 40.

Bibliography

[1] Nir Ailon and Bernard Chazelle. Information theory in property testing and monotonicity testing in higher dimension. *Inf. Comput.*, 204(11):1704–1717, 2006.

[2] Noga Alon. Testing subgraphs in large graphs. *Random Struct. Algorithms*, 21(3-4):359–370, 2002.

[3] Noga Alon, Alexandr Andoni, Tali Kaufman, Kevin Matulef, Ronitt Rubinfeld, and Ning Xie. Testing k-wise and almost k-wise independence. In *STOC*, pages 496–505, New York, NY, USA, 2007. ACM.

[4] Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy. Efficient testing of large graphs. *Combinatorica*, 20(4):451–476, 2000.

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

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

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

[8] Noga Alon and Asaf Shapira. A characterization of easily testable induced subgraphs. In *SODA*, pages 942–951, 2004.

[9] Noga Alon and Asaf Shapira. Testing subgraphs in directed graphs. *J. Comput. Syst. Sci.*, 69(3):354–382, 2004.

[10] Noga Alon and Asaf Shapira. A characterization of the (natural) graph properties testable with one-sided error. *SIAM Journal on Computing*, 37(6):1703–1727, 2008.

[11] Noga Alon and Asaf Shapira. Every monotone graph property is testable. *SIAM Journal on Computing*, 38(2):505–522, 2008.

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

[13] Sanjeev Arora and Madhu Sudan. Improved low degree testing and its applications. *Combinatorica*, 23(3):365–426, 2003. Preliminary version in Proceedings of ACM STOC 1997.

[14] László Babai, Lance Fortnow, Leonid A. Levin, and Mario Szegedy. Checking computations in polylogarithmic time. In *Proceedings of the 23rd ACM Symposium on the Theory of Computing*, pages 21–32, New York, 1991. ACM Press.

[15] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1(1):3–40, 1991.

[16] László Babai, Amir Shpilka, and Daniel Stefankovic. Locally testable cyclic codes. *IEEE Transactions on Information Theory*, 51(8):2849–2858, 2005.

[17] Tugkan Batu, Sanjoy Dasgupta, Ravi Kumar, and Ronitt Rubinfeld. The complexity of approximating the entropy. *SIAM J. Comput.*, 35(1):132–150, 2005.

[18] Tugkan Batu, Lance Fortnow, Ronitt Rubinfeld, Warren D. Smith, and Patrick White. Testing that distributions are close. In *FOCS*, pages 259–269, 2000.

[19] Tugkan Batu, Ravi Kumar, and Ronitt Rubinfeld. Sublinear algorithms for testing monotone and unimodal distributions. In *STOC*, pages 381–390, 2004.

[20] Mihir Bellare, Don Coppersmith, Johan Håstad, Marcos Kiwi, and Madhu Sudan. Linearity testing over characteristic two. *IEEE Transactions on Information Theory*, 42(6):1781–1795, November 1996.

[21] Eli Ben-Sasson. Limitations on the rate of families of locally testable codes. *ECCC*, TR10-123, 2010.

[22] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil P. Vadhan. Robust pcps of proximity, shorter pcps, and applications to coding. *SIAM Journal on Computing*, 36(4):889–974, 2006.

[23] Eli Ben-Sasson, Prahladh Harsha, and Sofya Raskhodnikova. Some 3CNF properties are hard to test. *SIAM Journal on Computing*, 35(1):1–21, September 2005. Preliminary version in *Proc. STOC 2003*.

[24] Eli Ben-Sasson and Madhu Sudan. Short PCPs with poly-log rate and query complexity. In *STOC*, pages 266–275, New York, 2005. ACM Press.

[25] Eli Ben-Sasson and Madhu Sudan. Limits on the rate of locally testable affine-invariant codes. *ECCC*, TR10-108, 2010.

[26] Eli Ben-Sasson and Michael Viderman. Low rate is insufficient for local testability. *ECCC*, TR10-004, 2010.

[27] Thierry P. Berger. On the automorphism groups of affine-invariant codes. *Des. Codes Cryptography*, 7(3):215–221, 1996.

[28] Thierry P. Berger and Pascale Charpin. The permutation group of affine-invariant extended cyclic codes. *IEEE Transactions on Information Theory*, 42(6):2194–2209, 1996.

[29] Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie. Testing linear-invariant non-linear properties. In *STACS*, pages 135–146, 2009. Full version at `http://www.eccc.uni-trier.de/report/2008/088/`.

[30] Arnab Bhattacharyya, Elena Grigorescu, Jakob Nordström, Asaf Shapira, and Ning Xie. *Manuscript*, 2010.

[31] Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira. A unified framework for testing linear invariant properties. *FOCS*, To appear, 2010.

[32] Eric Blais. Testing juntas nearly optimally. In *STOC*, pages 151–158, 2009.

[33] 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.

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

[35] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, Balázs Szegedy, and Katalin Vesztergombi. Graph limits and parameter testing. In *STOC*, pages 261–270, 2006.

[36] Jean Bourgain. Mordell's exponential sum estimate revisited. *J. Amer. Math. Soc.*, 18(2):477–499 (electronic), 2005.

[37] Jean Bourgain. Some arithmetical applications of the sum-product theorems in finite fields. In *Geometric aspects of functional analysis*, volume 1910 of *Lecture Notes in Math.*, pages 99–116. Springer, Berlin, 2007.

[38] Jean Bourgain and Mei-Chu Chang. A Gauss sum estimate in arbitrary finite fields. *C. R. Math. Acad. Sci. Paris*, 342(9):643–646, 2006.

[39] Jean Bourgain, Nets Katz, and Terence Tao. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.*, 14(1):27–57, 2004.

[40] Jean Bourgain and Sergei V. Konyagin. Estimates for the number of sums and products and for exponential sums over subgroups in fields of prime order. *C. R. Math. Acad. Sci. Paris*, 337(2):75–80, 2003.

[41] L. Carlitz and S. Uchiyama. Bounds for exponential sums. *Duke math. J.*, 24:37–41, 1957.

[42] Philippe Delsarte. On cyclic codes that are invariant under the general linear group. *IEEE Trans. Inf. Theory*, IT-16(6), 1970.

[43] Ilias Diakonikolas, Homin K. Lee, Kevin Matulef, Krzysztof Onak, Ronitt Rubinfeld, Rocco A. Servedio, and Andrew Wan. Testing for concise representations. In *FOCS*, pages 549–558, 2007.

[44] Irit Dinur. The PCP theorem by gap amplification. *J. ACM*, 54(3):12, 2007.

[45] Eldar Fischer. The art of uninformed decisions: A primer to property testing. In G. Paun, G. Rozenberg, and A. Salomaa, editors, *Current Trends in Theoretical Computer Science: The Challenge of the New Century*, volume 1, pages 229–264. World Scientific Publishing, 2004.

[46] Eldar Fischer, Guy Kindler, Dana Ron, Shmuel Safra, and Alex Samorodnitsky. Testing juntas. *Journal of Computer and System Sciences*, 68(4):753–787, 2004.

[47] Eldar Fischer, Eric Lehman, Ilan Newman, Sofya Raskhodnikova, Ronitt Rubinfeld, and Alex Samorodnitsky. Monotonicity testing over general poset domains. In *STOC*, pages 474–483, 2002.

[48] 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.

[49] Shuhong Gao, Robert J. Lambert, and Ian F. Blake. Construction and distribution problems for irreducible trinomials over finite fields. In *Applications of Finite Fields*. Press, 2001.

[50] Peter Gemmell, Richard Lipton, Ronitt Rubinfeld, Madhu Sudan, and Avi Wigderson. Self-testing/correcting for polynomials and for approximate functions. In *STOC*, pages 32–42, New Orleans, Louisiana, 6-8 May 1991.

[51] Oded Goldreich. Combinatorial property testing (a survey). In *Randomization Methods in Algorithm Design*, pages 45–60, 1998.

[52] Oded Goldreich, Shafi Goldwasser, Eric Lehman, Dana Ron, and Alex Samorodnitsky. Testing monotonicity. *Combinatorica*, 20(3):301–337, 2000.

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

[54] Oded Goldreich and Tali Kaufman. Proximity oblivious testing and the role of invariances. *Manuscript*, 2010.

[55] Oded Goldreich and Dana Ron. Approximating average parameters of graphs. *Random Struct. Algorithms*, 32(4):473–493, 2008.

[56] 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.

[57] Oded Goldreich and Madhu Sudan. Locally testable codes and PCPs of almost-linear length. *J. ACM*, 53(4):558–655, 2006. Preliminary version in *FOCS 2002*.

[58] Parikshit Gopalan, Ryan O'Donnell, Rocco A. Servedio, Amir Shpilka, and Karl Wimmer. Testing Fourier dimensionality and sparsity. In *ICALP*, pages 500–512, 2009.

[59] Ben Green. A Szemerédi-type regularity lemma in abelian groups. *Geometric and Functional Analysis*, 15(2):340–376, 2005.

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

[61] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. Succinct representation of codes with applications to testing. In *APPROX-RANDOM*, pages 534–547, 2009.

[62] Johan Håstad and Avi Wigderson. Simple analysis of graph tests for linearity and PCP. *Random Structures and Algorithms*, 22(2):139–160, 2003.

[63] Charanjit S. Jutla, Anindya C. Patthak, Atri Rudra, and David Zuckerman. Testing low-degree polynomials over prime fields. In *FOCS 2004*, pages 423–432. IEEE Computer Society, 2004.

[64] T. Kasami. An upper bound on k/n for affine-invariant codes with fixed d/n. *IEEE Transactions on Information Theory*, pages 174–176, January, 1969.

[65] 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.

[66] Tali Kaufman, Michael Krivelevich, and Dana Ron. Tight bounds for testing bipartiteness in general graphs. *SIAM J. Comput.*, 33(6):1441–1483, 2004.

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

[68] Tali Kaufman and Simon Litsyn. Long extended bch codes are spanned by minimum weight words. In *AAECC*, pages 285–294, 2006.

[69] Tali Kaufman and Dana Ron. Testing polynomials over general fields. *SIAM J. Comput.*, 36(3):779–802, 2006.

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

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

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

[73] Tali Kaufman and Michael Viderman. Locally testable vs. locally decodable codes. *RANDOM*, To appear, 2010.

[74] Tali Kaufman and Avi Wigderson. Symmetric ldpc codes and local testing. *ICS*, 2010.

[75] Marcos Kiwi. Algebraic testing and weight distribution of dual codes. *Theoretical Computer Science*, 299(1–3):81–106.

[76] Swastik Kopparty and Shubhangi Saraf. Tolerant linearity testing and locally testable codes. In *APPROX-RANDOM*, pages 601–614, 2009.

[77] Daniel Král', Oriol Serra, and Lluís Vena. A combinatorial proof of the removal lemma for groups. *Journal of Combinatorial Theory*, 116(4):971–978, May 2009.

[78] Ravi Kumar and Ronitt Rubinfeld. Algorithms column: Sublinear time algorithms. *SIGACT News*, 34(4):57–67, 2003.

[79] Angsheng Li and Yincheng Pan. Characterizations of locally testable linear and affine invariant families. *Manuscript*, 2010.

[80] Rudolf Lidl and Harald Niederreiter. *Finite fields.* Cambridge University Press, 1983.

[81] Florence J. MacWilliams and Neil J. A. Sloane. *The Theory of Error-Correcting Codes.* Elsevier/North-Holland, Amsterdam, 1981.

[82] Kevin Matulef, Ryan O'Donnell, Ronitt Rubinfeld, and Rocco A. Servedio. Testing halfspaces. In *SODA*, pages 256–264, 2009.

[83] Robert McEliece. On the symmetry of good nonlinear codes. *IEEE Transactions on Information Theory*, IT-16(5):609–611, 1970.

[84] Or Meir. Combinatorial construction of locally testable codes. *SIAM J. Comput.*, 39(2):491–544, 2009.

[85] R. E. A. C. Paley. Theorems on polynomials in a Galois field. *Q J Math*, os-4(1):52–63, 1933.

[86] Michal Parnas, Dana Ron, and Alex Samorodnitsky. Testing basic boolean formulae. *SIAM J. Discrete Math.*, 16(1):20–46, 2002.

[87] Dana Ron. Property testing (a tutorial). *Handbook of Randomization*, 2, 2000.

[88] Dana Ron. Property Testing: A Learning Theory Perspective. In *Foundations and Trends in Machine Learning*, volume 1, pages 307–402. 2008.

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

[90] Alex Samorodnitsky. Low-degree tests at large distances. In *STOC*, pages 506–515, 2007.

[91] Asaf Shapira. Green's conjecture and testing linear-invariant properties. In *Annual ACM Symposium on Theory of Computing*, pages 159–166, 2009.

[92] Amir Shpilka and Avi Wigderson. Derandomizing homomorphism testing in general groups. *SIAM Journal on Computing*, 36(4):1215–1230, 2006.

[93] Madhu Sudan. Invariance in property testing. *Electronic Colloquium on Computational Complexity*, TR 10-051, 2010.

[94] Endre Szemerédi. Regular partitions of graphs. In J.C. Bremond, J.C. Fournier, M. Las Vergnas, and D. Sotteau, editors, *Proc. Colloque Internationaux CNRS 260 - Problèmes Combinatoires et Théorie des Graphes*, pages 399–401, 1978.

[95] Luca Trevisan. Recycling queries in PCPs and in linearity tests (extended abstract). In *STOC*, pages 299–308, New York, NY, USA, 1998. ACM.

[96] John Tromp, Louxin Zhang, and Ying Zhao. Small weight bases for Hamming codes. *Theor. Comput. Sci.*, 181(2):337–345, 1997.

[97] Paul Valiant. Testing symmetric properties of distributions. In *STOC*, pages 383–392, 2008.

[98] André Weil. On some exponential sums. *Proc. Nat. Acad. Sci.*, 34:204–207, 1948.