Succinct Representation of Codes with Applications to Testing[*]

Elena Grigorescu[†] elena-g@purdue.edu

Tali Kaufman[‡] kaufmant@mit.edu

Madhu Sudan[§] madhu@mit.edu

February 4, 2013

Abstract

Motivated by questions in property testing, we search for linear error-correcting codes that have the "single local orbit" property: i.e., they are specified by a single local constraint and its translations under the symmetry group of the code. We show that the dual of every "sparse" binary code whose coordinates are indexed by elements of $\mathbb{F}_{2^n}$ for prime n, and whose symmetry group includes the group of non-singular affine transformations of $\mathbb{F}_{2^n}$, has the single local orbit property. (A code is said to be *sparse* if it contains polynomially many codewords in its block length). In particular this class includes the dual-BCH codes for whose duals (i.e., for BCH codes) simple bases were not known. Our result gives the first short ($O(n)$-bit, as opposed to the natural $\exp(n)$-bit) description of a low-weight basis for BCH codes.

The interest in the "single local orbit" property comes from the recent result of Kaufman and Sudan (STOC 2008) that shows that the duals of codes that have the single local orbit property under the affine symmetry group are locally testable. When combined with our main result, this shows that all sparse *affine-invariant* codes over the coordinates $\mathbb{F}_{2^n}$ for prime n are locally testable. If, in addition to n being prime, $2^n - 1$ does not have large divisors, then we get that every sparse *cyclic-invariant* code also has the single local orbit. In particular this implies that BCH codes of such length are generated by a single low-weight codeword and its cyclic shifts.

1 Introduction

Motivated by questions about the local testability of some well-known error-correcting codes, in this paper we examine their "invariance" properties. Invariances of codes are a well-studied concept (see, for instance, [25, Chapters 7, 8.5, and 13.9]) and yet we reveal some new properties of BCH codes. In the process we also find broad classes of sparse codes that are locally testable. We describe our problems and results in detail below.

A family of codes $\mathcal{C}_N \subseteq \mathbb{F}_2^N$ is said to be locally testable if membership of a word $w \in \mathbb{F}_2^N$ in C_N can be checked probabilistically by a few probes into w. The famed "linearity test" of Blum, Luby and

[*] A preliminary version appeared in the 13th Intl. Workshop on Randomization and Computation, 2009 [16], and in Elena Grigorescu's PhD Thesis [14].

[†] Purdue University, 305 N. University Ave., West Lafayette, IN 47907. Research conducted when this author was at MIT CSAIL. Research supported in part by NSF grant CCR-0829672 and NSF award 1019343 to the Computing Research Association for the Computing Innovation Fellowship Program.

[‡] Bar-Ilan University and the Weizmann Institute of Science, Israel. Research conducted when this author was at MIT CSAIL. Research supported in part by NSF grant CCR-0829672.

[§] Microsoft Research New England, One Memorial Drive, Cambridge, MA 02142. Research conducted when this author was at MIT CSAIL, supported in part by NSF grant CCR-0829672.

Rubinfeld [3] may be considered the first result to show that some code is locally testable. Locally testable codes were formally defined by Rubinfeld and Sudan [27]. The first substantial study of locally testable codes was conducted by Goldreich and Sudan [13], where the principal focus was the construction of locally testable codes of high rate. Local testing of codes is effectively equivalent to property testing [27, 12] with the difference being that the emphasis here is when C is an error-correcting code, i.e., elements of C are pairwise far from each other.

A wide variety of "classical" codes are by now known to be locally testable, including Hadamard codes [3], Reed-Muller codes of various parameters [27, 1, 21, 17], dual-BCH codes [18, 22], turning attention to the question: What broad characteristics of codes are necessary, or sufficient, for codes to be locally testable. One characteristic explored in the recent work of Kaufman and Sudan [23] is the "invariant group" of the code, a well-studied object that we define next.

Let $[N]$ denote the set of integers $\{1, \ldots, N\}$. A code $\mathcal{C} \subseteq \mathbb{F}_2^N$ is said to be *invariant* under a permutation $\tau : [N] \to [N]$ if for every $a = \langle a_1, \ldots, a_N \rangle \in \mathcal{C}$, it is the case that $a \circ \tau = \langle a_{\tau(1)}, \ldots, a_{\tau(N)} \rangle$ is also in $\mathcal{C}$. We will often alternate between viewing $a \in \mathbb{F}_2^N$ as a vector $a = \langle a_1, \ldots, a_N \rangle$ and as a function $a : X \to \mathbb{F}_2$, where X will be some appropriate domain of size N. Two particular domains X of interest to us will be $\mathbb{F}_{2^n}$ and $\mathbb{F}_{2^n}^*$ (where $\mathbb{F}_{2^n}$ is the finite field with 2^n elements and $\mathbb{F}_{2^n}^* = \mathbb{F}_{2^n} \setminus \{0\}$ denotes the non-zero elements in this field). Under this view, a permutation $\tau : [N] \to [N]$ will be associated with a permutation of the domain X, $\pi : X \to X$, and the π-rotation of a is the function $a \circ \pi : X \to \mathbb{F}_2$ given by $a \circ \pi(i) = a(\pi(i))$ for every $i \in X$. As before, a code $\mathcal{C} \subseteq \mathbb{F}_2^N$ is said to be invariant under a permutation $\pi : X \to X$ if for every $a \in \mathcal{C}$, it is the case that $a \circ \pi \in \mathcal{C}$. The set of permutations under which a code $\mathcal{C}$ is invariant forms a group under composition and we refer to it as the *automorphism group* of $\mathcal{C}$.

Kaufman and Sudan [23] suggested that the automorphism group of a code may play an important role in its testability. They supported their suggestion by showing that if the automorphism group is an "affine group", then a linear code whose dual has the "single local orbit" property is locally testable. We explain these terms (in a restricted setting) below.

We say that $\mathcal{C} \subseteq \mathbb{F}_2^N$ is *linear* if it is a $\mathbb{F}_2$-linear subspace of $\mathbb{F}_2^N$. For $a, b \in \mathbb{F}_2^N$, let $a \cdot b = \sum_i a_i b_i$ denote the inner product of a and b. The *dual* of $\mathcal{C}$ is the subspace orthogonal to $\mathcal{C}$ with respect to the inner product, i.e. $\mathcal{C}^\perp = \{b \in \mathbb{F}_2^N \mid b \cdot a = 0,\ \forall a \in \mathcal{C}\}$.

Let $N = 2^n$. In this case we can associate the coordinate set $[N]$ of the code $\mathcal{C} \subseteq \mathbb{F}_2^N$ with the field $\mathbb{F}_{2^n}$. Two groups are of special interest to us in this work. The first is the "affine group" on $\mathbb{F}_{2^n}$ and the second is the "cyclic group" on $\mathbb{F}_{2^n}^*$.

Definition 1 (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 $\pi(x) = \alpha x + \beta$. The* affine group *over $\mathbb{F}_{2^n}$ consists of all the affine permutations over $\mathbb{F}_{2^n}$. 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.*

Definition 2 (Cyclic 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 $\alpha \in \mathbb{F}_{2^n}^*$. The* cyclic group *over $\mathbb{F}_{2^n}^*$ consists of all the cyclic permutations over $\mathbb{F}_{2^n}^*$. 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 cyclic group.*

Many well-known families of codes (with minor variations) are known to be affine-invariant and/or cyclic (under appropriate ordering of the coordinates and with some slight modifications, see [30] or [25]). In particular BCH codes are cyclic (-invariant) and Reed-Muller codes are affine-invariant. Furthermore, under a simple parity check extension operation BCH codes become affine-invariant, and vice versa under a simple puncturing operation, Reed-Muller codes become cyclic. We elaborate on these later.

In this paper our aim is to show that certain families of affine-invariant and cyclic codes have a simple description, that we call a "single-orbit description".

First recall some basic definitions. For a word $a = \langle a_1, \ldots, a_N \rangle \in \mathbb{F}_2^N$ its support is the set $\mathrm{Supp}(a) = \{i | a_i \neq 0\}$ and its *weight* is the quantity $\mathrm{wt}(a) = |\mathrm{Supp}(a)|$. For a set of vectors $S = \{v_1, \ldots, v_k\} \subseteq \mathbb{F}_2^N$, let $\mathrm{Span}(S) = \{\sum_{i=1}^k \alpha_i v_i | \alpha_1, \ldots, \alpha_k \in \mathbb{F}_2\}$ denote the linear span of S.

Definition 3 (k-single orbit code) *Let $k > 0$ and let G be a group of permutations from $[N]$ to $[N]$. A linear code $\mathcal{C} \subseteq \mathbb{F}_2^N$ is said to have the k-*single orbit property *under the group G if there exists $a \in \mathcal{C}$ with $\mathrm{wt}(a) \leq k$ such that $\mathcal{C} = \mathrm{Span}(\{a \circ \pi | \pi \in G\})$.*

In particular the k-single orbit property under the affine group has implications to testing that we discuss in Section 5.1. We emphasize that for these applications N is large ($N \to \infty$) while k needs to be independent of N, and thus is thought of as constant. Therefore, we are interested in asymptotic families of codes $\{\mathcal{C}_N\}_{N \to \infty}$ with $\mathcal{C}_N \subseteq \mathbb{F}_2^N$ or $\mathcal{C}_N \subseteq \mathbb{F}_2^{N-1}$ such that each $\mathcal{C}_N$ in the family has the k-single orbit property. We note that all of our results hold for *families* of codes but we will often omit saying this explicitly, as it will (hopefully) be clear from the context.

Conditions under which codes have the single-orbit property under any given group, seem to be less well-studied. This is somewhat surprising given that the single-orbit property implies very succinct (nearly explicit) descriptions of bases for codes (that have the k-single orbit property under some standard group). More precisely, to specify a code that has the k-single orbit property, it is enough to specify the k non-zero indices of the generating codeword, i.e. $k \log N$ bits (this is the notion of explicitness that we refer to throughout the paper), while for an arbitrary code, one needs to specify the N bits of each basis vector.

Even for such commonly studied codes such as BCH codes such explicit descriptions of low-weight bases[1] were not known prior to this work. In retrospect, the single orbit property was being exploited in previous results in algebraic property testing [3, 27, 1, 21, 17] though this fact was not explicit until the work of [23].

1.1 Results, implications and approaches

In our results we are concerned with "sparse" (families of) codes $\{\mathcal{C}_N\}_{N \to \infty}$, i.e. codes containing at most polynomially many codewords in N. (More formally, $\{\mathcal{C}_N\}_{N \to \infty}$ is sparse if there exists a constant $t > 0$ such that $|\mathcal{C}_N| \leq N^t$ for all $\mathcal{C}_N$'s in the family).

In this work we explore the single orbit property under the affine group for codes on the coordinate set $\mathbb{F}_{2^n}$, as also the single orbit property under the cyclic group for codes over $\mathbb{F}_{2^n}^*$. We show that the dual of every sparse affine-invariant code has the k-single orbit property under the affine group for some constant k, provided $N = 2^n$ for large enough prime n.

Theorem 4 (Single orbit property in affine-invariant codes) *For every integer $t > 0$ there exist k and n_0 such that for every prime $n \geq n_0$ the following holds:*

Let $N = 2^n$ and $\mathcal{C} \subseteq \mathbb{F}_2^N$ be a linear affine-invariant code containing at most N^t codewords. Then $\mathcal{C}^\perp$ has the k-single orbit property under the affine group.

[1] One way to represent a sparse code C (e.g., dual-BCH codes) whose dual $C^\perp$ has a basis among the weight k codewords is to give $\Omega(N)$ codewords that generate $C^\perp$. This requires space $\Omega(kN \log N)$ bits. Alternately, if C is sparse and has N^t codewords, one can give $t \log N$ codewords that generate it; this requires $tN \log N = \Omega(N \log N)$ bits. If $C^\perp$ has the k-single orbit property then it can be specified using $k \log N$ bits. This representation is particularly useful for property testing applications since it provides explicit tests. Namely, a test can be specified by the k non-zero indices of the generating codeword and by a permutation in the automorphism group of the code.

When $N-1$ does not have large divisors, it turns out that the duals of sparse codes have the k-single orbit property under the cyclic group for some constant k yielding an even stronger condition on the basis.

Theorem 5 (Single orbit property in cyclic-invariant codes) *For every integer $t > 0$ and $\epsilon \in (0, 1)$ there exist k and n_0 such that for any prime $n \geq n_0$ for which $2^n - 1$ does not have any non-trivial divisor larger than $2^{n(1-\epsilon)}$ the following holds:*

Let $N = 2^n$ and $\mathcal{C} \subseteq \mathbb{F}_2^{N-1}$ be a linear, cyclic-invariant, code with at most N^t codewords. Then $\mathcal{C}^\perp$ has the k-single orbit property under the cyclic group.

(The proofs of Theorem 4 and Theorem 5 appear in Section 4). Theorem 5 holds in particular when $N-1$ is a Mersenne prime. 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 [28, 31, 33]), and so unconditional result should remain interesting.

Both theorems shed new light on well-studied codes including BCH codes. The actual families considered here are broader, but the BCH codes are typical in these collections. Lemma 6 explicitly characterizes the entire family of codes investigated in this paper.

In particular the first theorem has immediate implications for testing and shows that every sparse affine invariant code is locally testable. This merits comparison with the results of [22] who show that sparse high-minimum-distance codes are locally testable. While syntactically the results seem orthogonal (ours require affine-invariance whereas theirs required high-distance) it turns out (as we show in this paper) that all the codes we consider do have high-distance. Yet for the codes we consider our results are more constructive in that they not only prove the "existence" of a local test, but give a much more "explicit" description of the tester: Our tester is described by a single low-weight word in the dual and tests that a random affine permutation of this word is orthogonal to the word being tested. [2]

Given a code of interest to us, we first study the algebraic structure of the given code by representing codewords as polynomials and studying the degree patterns among the support of these polynomials. We interpret the single orbit property in this language; and this focuses our attention on a collection of closely related codes. We then turn to recent results from additive number theory [4, 5, 6, 7, 8] and apply them to the dual of the given code, as well as the other related codes that arise from our algebraic study, to lower bound their distance. In turn, using the MacWilliams identities (as in prior work [22]) this translates to some information on the weight distribution of the given code and the related ones. Some simple counting then yields that the given code must have the single-orbit property.

We believe that our techniques are of interest, beyond just the theorems they yield. In particular we feel that techniques to assert the single-orbit property are quite limited in the literature. Indeed in all previous results [3, 27, 1, 21, 17] this property was "evident" for the code: The local constraint whose orbit generated a basis for all constraints was explicitly known, and the algebra needed to prove this fact was simple. Our results are the first to consider the setting where the basis is not explicitly known (even after our work) and we manage to bring in non-algebraic tools to handle such settings. We believe that the approach is potentially interesting in broader settings.

1.2 Subsequent developments

Subsequent to our work, Kaufman and Lovett [20] extended our results to affine invariant codes containing a quasi-polynomial number of codewords. Moreover, their results hold for codes over $\mathbb{F}_p$ of length p^n, where p

[2] In contrast the tester of [22] was less "explicit". It merely proved the existence of many low weight codewords in the dual of the code being tested and proved that the test which picked one of these low-weight codewords uniformly at random and tested orthogonality of the given word to this dual codeword was a sound test.

is a (constant) prime and n is arbitrary, thus removing the restriction that n be a prime in our results. The work of [20] follows the footsteps of our analysis. They however manage to replace the tools we use from additive combinatorics (which imply that every sparse affine-invariant code is of high distance) with a new self-contained proof. This allows them to show that affine-invariant codes of quasi-polynomial size have very high distance. In addition, [20] uses a Fourier analysis based approach to replace our use of the MacWilliams identities for estimating the weight distributions of linear codes. This enables them to extend the result to codes over fields with arbitrary small characteristic.

1.3 Organization

In Section 2 we describe our proof strategy and state the basic results that we rely on. In Section 3 we describe the unique representation of affine and cyclic-invariant codes as traces of polynomials. In Section 4 we use this representation to prove our main results. Finally, we conclude with Section 5 on implications of our results to property testing and to the study of BCH codes.

2 Overview of techniques

Theorems 4 and 5 are proved essentially by implementing the following plan:

1. We first show that the codewords in the codes we consider are expressible as the traces of *sparse* polynomials (a sparse polynomial contains only a few monomials). In the affine-invariant case we also show that these polynomials have somewhat low degree, i.e., at most $N^{1-\epsilon}$. This part follows from standard literature in coding theory (and similar steps were employed already in [23]).

2. We then apply the recent results in additive number theory to conclude that these codes have very high distance. This already suffices to show that the affine-invariant codes are testable by [22]. However the tests given there are "non-explicit" and we need to work further to get an "explicit" test for these codes, 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 codewords in the dual of the codes we consider such that their orbit spans the dual.

We elaborate on these steps in detail below, laying out precise statements we will prove.

We start with some notation. Recall $N = 2^n$ and n is prime. For $a = \langle a_i \rangle_i$, and $b = \langle b_i \rangle_i \in \mathbb{F}_2^N$ define the *relative distance* between a, b as $\delta(a,b) = \frac{1}{N} \left| \{ i \mid a_i \neq b_i \} \right|$. Note $\delta(a,b) = \frac{\mathrm{wt}(a-b)}{N}$. The *(relative) distance of the code* $\mathcal{C}$ is $\delta(\mathcal{C}) = \min_{a,b \in \mathcal{C}; a \neq b} \{\delta(a,b)\}$. The relative distance of a vector a to a code $\mathcal{C}$ is $\delta(a, \mathcal{C}) = \min_{b \in \mathcal{C}} \{\delta(a,b)\}$.

Also, we view elements $c \in \mathbb{F}_2^N$ as functions $c : \mathbb{F}_N \to \mathbb{F}_2$. Let $\{\mathbb{F}_N \to \mathbb{F}_2\}$ denote the set of all such functions. Similarly we view elements $c \in \mathbb{F}_2^{N-1}$ as functions $\mathbb{F}_N^* \to \mathbb{F}_2$ and let $\{\mathbb{F}_N^* \to \mathbb{F}_2\}$ denote the set of all such functions.

We will also be using the representation functions as polynomials and in particular of trace polynomials. Recall that the Trace function is defined as

$$\mathrm{Trace}(x) = x + x^2 + x^4 + \cdots + x^{2^{n-1}}.$$

The Trace function is linear over $\mathbb{F}_2$, i.e. $\mathrm{Trace}(\alpha + \beta) = \mathrm{Trace}(\alpha) + \mathrm{Trace}(\beta)\ \forall \alpha, \beta \in \mathbb{F}_N$.

Throughout the paper the notation $k < \infty$ will denote that k is finite.

For $d \in \{1, \dots, N-2\}$, let the cyclotomic coset of d be

$$\mathrm{coset}(d) = \{d, 2d \bmod (N-1), 4d \bmod (N-1), \dots, 2^{n-1}d \bmod (N-1)\}.$$

By the primality of n, we have that $|\mathrm{coset}(d)| = n$ for every d. Let $\mathrm{leader}(d)$ denote the smallest integer in $\mathrm{coset}(d)$, i.e. the coset leader of its cyclotomic class, and let

$$\mathcal{D} = \{\mathrm{leader}(d) \mid d \in \{1, \dots, N-2\}\} \cup \{N-1\}.$$

Note that $|\mathcal{D}| = 1 + (N-2)/n$.

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

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

and

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

The first step in our analysis of codes invariant over the affine group (respectively, cyclic group) is that such codes can be associated uniquely with a set $D \subseteq \mathcal{D}$ so that every codeword in our code is the evaluation of the trace of a polynomial from the associated family $P_{N,D}$ over $\mathbb{F}_N$ (respectively, $P_{N-1,D}$ over $\mathbb{F}_N^*$).

We note that choosing the representatives of the cyclotomic coset be the minimum elements of the cosets is useful not only because this leads to a way of describing unique representations of the codewords in affine/cyclic-invariant codes, but also because in our results about affine-invariant codes the representation of codewords as low-degree polynomials allows us to use the Weil-Calitz-Uchiyama type distance bounds.

Lemma 6 *For every affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_N \to \mathbb{F}_2\}$, there exists a set $D \subseteq \mathcal{D}$ such that $c \in \mathcal{C}$ if and only if there exists a polynomial $p \in P_{N,D}$ such that $c(x) = \mathrm{Trace}(p(x))$ for every $x \in \mathbb{F}_N$. Furthermore, for integer $t \geq 0$, $|\mathcal{C}| \leq 2N^t$ iff $|D| \leq t$ and $D \subseteq \{1, \dots, N^{1-1/t}\}$.*

Similarly, for every cyclic-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_N^* \to \mathbb{F}_2\}$ there exists 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 $c(x) = \mathrm{Trace}(p(x))$ for every $x \in \mathbb{F}_N^*$. Furthermore, for integer $t \geq 0$, $|D| \leq t$ iff $|\mathcal{C}| \leq N^t$.

(The proof of Lemma 6 appears in Section 3.) Thus in both cases codes are represented by collections of t-sparse polynomials (a t-sparse polynomial contains at most t monomials). 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

$$\mathcal{C}_N(D) = \{\mathrm{Trace}(p(x)) | p \in P_{N,D}\}$$

and we use $\mathcal{C}_{N-1}(D)$ to denote

$$\mathcal{C}_{N-1}(D) = \{\mathrm{Trace}(p(x)) | p \in P_{N-1,D}\}.$$

(As mentioned above, the actual codewords of these codes are of the form $\langle \mathrm{Trace}(p(x)) \rangle_{x \in \mathbb{F}_N}$ where $p \in P_{N,D}$, and, respectively, $\langle \mathrm{Trace}(p(x)) \rangle_{x \in \mathbb{F}_N^*}$, where $p \in P_{N-1,D}$).

We next use a variant of a theorem initially due to Bourgain ([4], Theorem 7), whose proof was simplified and extended subsequently in [20], to conclude that the codes $\mathcal{C}_N(D)$ and $\mathcal{C}_{N-1}(D)$ have very high distance (under the given conditions on D).

Theorem 7 (Implied by Theorem 1.3 [20]) *For every $\epsilon > 0$ and $r < \infty$, there exist $n_0, \eta > 0$ such that for every prime $n \geq n_0$ 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 $1 \leq i \leq r$, and (3) $\gcd(k_i - k_j, N-1) < N^{1-\epsilon}$ for every $1 \leq i < j \leq r$. Then*

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

Theorem 7 generalizes the standard "Weil-Carlitz-Uchiyama bound" [9, 32] which states that if $f \in \mathbb{F}[x]$ has degree $\deg(f) \leq N^{1/2-\eta}$ then $\left| \sum_{x \in \mathbb{F}} (-1)^{\mathrm{Trace}(f(x))} \right| < N^{1-\eta}$. Theorem 7 obtains the same qualitative bound as the classical bound but for polynomials of possibly very large degrees. In particular, it holds for polynomials whose degrees are up to $N^{1-\epsilon}$, and even for polynomials of degrees up to $N-1$ but whose monomial degrees are carefully selected as to satisfy the additional constraints of the theorem. In our applications we will also need that η is a constant independent of N, which is ensured by the fact that the polynomials that we focus on are sparse (i.e. have a constant number of terms).

We note that in our preliminary version [16] we attributed Theorem 7 to Bourgain [5], who strictly speaking only considers the case where N is prime, and we commented on how to modify his proof to work when $N = 2^n$ with n prime. A reader interested in the full proof can read [20] for a self-contained and simplified treatment of the more general case $N = p^n$ for arbitrary n, where p is a prime.

In our language the above theorem implies that codes represented by sparse polynomials of somewhat low degree have large distance. Furthermore if the polynomials are sparse, and $N-1$ does not have divisors larger than $N^{1-\epsilon}$, then also the codes have large distance. We thus get the following implication.

Lemma 8 *For every integer $t > 0$ there exist $n_0, \eta > 0$ such that the following holds for every $N = 2^n$ for prime $n \geq n_0$. Let $\mathcal{D} = \mathcal{D}(N)$ and let $D \subseteq \mathcal{D}$ be of size at most t such that the code $\mathcal{C} = \mathcal{C}_N(D)$ is affine-invariant. Then the code $\mathcal{C}$ satisfies $\frac{1}{2} - N^{-\eta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\eta}$.*

Similarly for every integer t and $\epsilon > 0$, there exist $n_0, \eta > 0$ such that the following holds for every $N = 2^n$ such that $n \geq n_0$ is prime, and $N-1$ does not have any non-trivial divisor greater than $N^{1-\epsilon}$. Let $\mathcal{D} = \mathcal{D}(N)$ and let $D \subseteq \mathcal{D}$ be of size at most t. Then the code $\mathcal{C} = \mathcal{C}_{N-1}(D)$(which is always cyclic-invariant) satisfies $\frac{1}{2} - N^{-\eta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\eta}$.

Proof: For $p \in P_{N,D}$ such that $\mathrm{Trace}(p) \in \mathcal{C}$ define for the purpose of this proof

$$\delta(p) = \Pr_{x \leftarrow \mathbb{F}_N} [\mathrm{Trace}(p(x)) = 1],$$

where the probability is taken with respect to the uniform distribution over $\mathbb{F}_N$. Since $|D| \leq t$ and $\mathcal{C}_N(D)$ is affine-invariant, by Lemma 6 it follows that the degrees in D are upper bounded by $N^{1-1/t}$, and by Theorem 7 (with $\epsilon = 1/t$ and $r \leq t+1$) there exists $\eta' = \eta'(t)$ such that

$$| \sum_{x \in \mathbb{F}_N} (-1)^{\mathrm{Trace}(p(x))} | < N^{1-\eta'}.$$

Since $\mathbb{E}_{x \in \mathbb{F}_N} (-1)^{\mathrm{Trace}(p(x))} = 1 - 2\delta(p)$, it follows that there exists η such that

$$\frac{1}{2} - N^{-\eta} \leq \delta(p) \leq \frac{1}{2} + N^{-\eta}.$$

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. ▌

We remark that such use of results from number theory in coding theory is also common. For example, the distance of the sparse dual-BCH codes is inferred by using the Weil-Carlitz-Uchiyama bound on exponential sums in a similar manner.

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 [22] to show that for any specific code $\mathcal{C}$ we consider and for every sufficiently large k, its dual has roughly $\binom{N}{k}/|\mathcal{C}|$ codewords of weight k (this bound is tight to within $1 \pm 1/N^2$ factor, for large enough k, where k is independent of N and depends only on t, and the η of Lemma 8). Specifically they show:

Theorem 9 ([22] Lemma 3.5) *For every $t < \infty$ and $\eta > 0$ there exist k_0, N_0, such that for every $k \geq k_0$ and $N \geq N_0$ the following holds:*

Let $\mathcal{C} \subseteq \mathbb{F}_2^N$ be a code with at most N^t codewords satisfying

$$\frac{1}{2} - N^{-\eta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\eta}.$$

Then the number of codewords of weight k in $\mathcal{C}^\perp$ lies in the interval

$$\left[\frac{\binom{N}{k}}{|\mathcal{C}|} \cdot (1 - N^{-2}), \frac{\binom{N}{k}}{|\mathcal{C}|} \cdot (1 + N^{-2}) \right].$$

Thus for any code $\mathcal{C} = \mathcal{C}(D)$ under consideration, this allows us to conclude that $\mathcal{C}^\perp$ has many codewords 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 codeword x of weight k in the dual whose orbit under the group does *not* contain a basis for $\mathcal{C}^\perp$ (i.e., $\mathrm{Span}(\{x \circ \pi | \pi\}) \neq \mathcal{C}^\perp$). 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 code $\mathcal{C}(D')$ also satisfies the same invariant structure as $\mathcal{C}$ and so one can apply Lemma 8 and Theorem 9 to it and thereby conclude that the number of weight k codewords in $\mathcal{C}(D')^\perp$ are also smaller than the number weight k codewords in $\mathcal{C}^\perp$ by a factor of approximately N. Finally we notice that the number of sets D' is $o(N)$ and so the set $\cup_{D'} \mathcal{C}(D')^\perp$ can not include all possible weight k codewords in $\mathcal{C}^\perp$, yielding the k-single orbit property for $\mathcal{C}$. This leads to the proofs of Theorem 4 and 5 (see Section 4 for the formal proof).

3 Representing sparse invariant codes by sparse polynomials

In this section we study representations of affine-invariant and cyclic-invariant codes by polynomials. That leads to the proof of Lemma 6 which will conclude the section.

We remark that similar results appeared previously in the literature and are well-known [34, 29, 25, 23], but we reprove everything that we need here using our own techniques and in the language of the current exposition, for the sake of completeness.

Recall that every function from $\mathbb{F}_N$ to $\mathbb{F}_N$ and hence every function from $\mathbb{F}_N$ to $\mathbb{F}_2$ is the evaluation of polynomial from $\mathbb{F}_N[x]$. More useful to us is the fact that every function from $\mathbb{F}_N$ to $\mathbb{F}_2$ can also be expressed as the trace of a polynomial from $\mathbb{F}_N[x]$, however this representation is not unique. E.g., $\mathrm{Trace}(x^d) = \mathrm{Trace}(x^{2d}) = \mathrm{Trace}(x^{2^i \cdot d})$. However if we restrict to the setting of polynomials from $P_{N,\mathcal{D}}$ then this representation is unique, as shown below.

Lemma 10 *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 $w(x) = \mathrm{Trace}(p(x))$. In fact, the mapping between codewords in $\{\mathbb{F}_N \to \mathbb{F}_2\}$ and their associated polynomials in $P_{N,\mathcal{D}}$ is 1-1. (And so is the mapping between $\{\mathbb{F}_N^* \to \mathbb{F}_2\}$ and $P_{N-1,\mathcal{D}}$).*

Proof: Since every function w maps $\mathbb{F}_N$ to $\mathbb{F}_N$, we can write $w(x)$ uniquely as $\sum_{i=0}^{N-1} c_i x^i$, where $c_i \in \mathbb{F}_N$. 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 $w(x)^2 = w(x) \bmod (x^N - x)$. But $w(x)^2 = \sum_{i=0}^{N-1} c_i^2 x^{2i}$ and so, equating coefficients we have, $c_0^2 = c_0$, $c_{N-1}^2 = c_{N-1}$, and $c_{2i \bmod (N-1)} = c_i^2$ for every $i \in \{1, \ldots, N-2\}$. Thus writing the set $\{0, \ldots, N-1\}$ (the set of degrees of x) as $\{0, N-1\} \cup (\cup_{d \in \mathcal{D} - \{N-1\}} \mathrm{coset}(d))$, where the sets $\mathrm{coset}(d)$ are disjoint, we have that $w(x) = c_0 x^0 + c_{N-1} x^{N-1} + \sum_{d \in \mathcal{D} \setminus \{N-1\}} \mathrm{Trace}(c_d x^d)$. Furthermore $c_0, c_{N-1} \in \mathbb{F}_2$ (since $c_0^2 = c_0$ and $c_{N-1}^2 = c_{N-1}$). Finally, using the fact that $\mathrm{Trace}(a) = a$ for $a \in \mathbb{F}_2$ (using the fact that n is odd), we have $w(x) = \mathrm{Trace}(p(x))$ where $p(x) = c_0 x^0 + c_{N-1} x^{N-1} + \sum_{d \in \mathcal{D} - \{N-1\}} c_d x^d$, which is by definition a member of $P_{N,\mathcal{D}}$. 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).

The final part of the lemma can be easily shown by a simple counting argument. There are 2^N words $w : \mathbb{F}_N \to \mathbb{F}_2$ and, since $|\mathcal{D}| = 1 + (N-2)/n$, it follows that $|P_{N,\mathcal{D}}| = 4 \cdot N^{\frac{N-2}{n}} = 2^N$. The statement for $w : \mathbb{F}_N^* \to \mathbb{F}_2$ and $P_{N-1,\mathcal{D}}$ follows similarly. ■

Lemma 11 *Suppose $\mathcal{C} \subseteq \{\mathbb{F}_N \to \mathbb{F}_2\}$ is an affine invariant code containing the word $w = \mathrm{Trace}(p(x))$ for some $p \in P_{N,\mathcal{D}}$. Then, for every monomial x^e in the support of p, the function $\mathrm{Trace}(x^e)$ is in $\mathcal{C}$. Furthermore, if $e \notin \{0, N-1\}$ then for every $\beta \in \mathbb{F}_N$, $\mathrm{Trace}(\beta x^e) \in \mathcal{C}$.*

Similarly if $\mathcal{C} \subseteq \{\mathbb{F}_N^* \to \mathbb{F}_2\}$ is 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 $\mathcal{C}$. 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 [23]. 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.

We start with the cyclic-invariant case. 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 $\mathrm{Trace}(\beta x^e)$ 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 $\sum_{\alpha \in \mathbb{F}_N^*} \mathrm{Trace}(\alpha^{-e}) w(\alpha x)$ is also in $\mathcal{C}$ (by linearity). But as we show below this term is simply $\mathrm{Trace}(c_e x^e)$.

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

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

Finally, we need to show that $\mathrm{Trace}(\beta x^e)$ is also in $\mathcal{C}$. To see this, consider the set $S \subseteq \mathbb{F}_N$ defined as $S = \{\gamma|\mathrm{Trace}(c_e\gamma x^e) \in \mathcal{C}\}$. 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 a primitive element 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 $T = \mathbb{F}_N$ (since n is prime and hence the only other subfield of $\mathbb{F}_N$ is $\mathbb{F}_2$, which cannot contain ω^e when $e \in \mathcal{D}$). We thus conclude that both S and T equal $\mathbb{F}_N$ and so for every $\beta \in \mathbb{F}_N$, $\mathrm{Trace}(\beta x_e) \in \mathcal{C}$.

To prove the lemma for the cyclic-invariant case, it 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 $c_{N-1}x^{N-1} = x^{N-1} = \mathrm{Trace}(x^{N-1})$. This concludes the analysis of the cyclic-invariant case.

The affine invariant case is similar (and indeed only needs to use the facts that $w(\alpha x)$ is in $\mathcal{C}$ for every $\alpha \in \mathbb{F}_N$, and the linearity of $\mathcal{C}$). ▌

We now use Lemma 11 to characterize cyclic-invariant families, while also working towards the characterization of affine invariant families.

Lemma 12 *For every affine invariant code $\mathcal{C} \subseteq \{\mathbb{F}_N \to \mathbb{F}_2\}$ there exists a (unique) set $D \subseteq \mathcal{D}$ such that $\mathcal{C} = \{\mathrm{Trace}(p)|p \in P_{N,D}\}$.*

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 $\mathcal{C} = \{\mathrm{Trace}(p)|p \in P_{N-1,D}\}$.

Proof: We start with the affine-invariant case (the cyclic case is almost identical). We 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 $\mathrm{Trace}(p) \in \mathcal{C}$. By Lemma 11 we have that every function $\mathrm{Trace}(\beta x^d) \in \mathcal{C}$ for every $\beta \in \mathbb{F}_N$, if $d \notin \{0, N-1\}$. Furthermore since $\mathrm{Trace}((x+1)^d)$ 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 $\mathcal{C}$. Conversely, it can also be verified that every function in $\mathcal{C}$ is a trace of a polynomial in $P_{N,D}$.

The cyclic-invariant case is similar. ▌

Lemma 12 essentially suffices to yield Lemma 6 for the cyclic case (though we still need to verify that $|D|$ is small as claimed). For the affine case we need to work a little harder to bound the size of the integers in D. To do so we note that affine-invariant properties have further constraints on the set D.

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. We note that affine-invariant codes are characterized by codes with a "shadow-closure" property described below.

Lemma 13 *If $\mathcal{C}$ is an affine-invariant code,* $\mathrm{Trace}(x^d) \in \mathcal{C}$ *and* $e \prec d$ *then* $\mathrm{Trace}(x^e) \in \mathcal{C}$.

Proof: Since $\mathrm{Trace}(x^d) \in \mathcal{C}$ and $\mathcal{C}$ is affine invariant, then $\mathrm{Trace}((x+1)^d) \in \mathcal{C}$. But $(x+1)^d = \prod_i (1+x)^{d_i 2^i} = \prod_i (1+x^{d_i 2^i}) = \sum_{e \prec d} x^e$. Therefore, $\mathrm{Trace}(\sum\limits_{e \prec d} x^e) \in \mathcal{C}$ and by Lemma 11 $\mathrm{Trace}(x^e) \in \mathcal{C}$. ▮

We can now complete the proof of Lemma 6.

Proof of Lemma 6.: We show the second part of the lemma first as it is easier to argue. Namely, for the cyclic-invariant case, the lemma is immediate from Lemma 12 which claims that every cyclic-invariant code $\mathcal{C} = \mathcal{C}_{N-1}(D) = \{\mathrm{Trace}(p) | p \in P_{N-1,D}\}$ for some $D \subseteq \mathcal{D}$. Conversely, it can be verified that for every $D \subseteq \mathcal{D}$, the code $\mathcal{C}(D)$ is cyclic-invariant and maps $\mathbb{F}_N^*$ to $\mathbb{F}_2$. Finally, since for every pair of functions $p_1 \neq p_2 \in P_{N-1,D}$ $\mathrm{Trace}(p_1) \neq \mathrm{Trace}(p_2)$, we have that $|\mathcal{C}| = |P_{N-1,D}| = 2N^{|D|-1}$, yielding $|D| \leq t$ if $|\mathcal{C}| \leq N^t$.

Consider an affine-invariant code $\mathcal{C}$. By Lemma 12 there is a set $D \subseteq \mathcal{D}$ such that $\mathcal{C} = \mathcal{C}_N(D) = \{\mathrm{Trace}(p) | p \in P_{N,D}\}$. As above we also have $|D| \leq t$ if $|\mathcal{C}| \leq N^t$. It remains to be shown that $D \subseteq \{1, \ldots, N^{1-1/t}\}$.

For this part we use Lemma 13 to note first that the set D should be "shadow-closed", i.e., if $d \in D$ and $e \prec d$ then $e \in D$. Now consider the "binary weight" of the integers $d \in D$, i.e., the number of non-zero bits in the binary representation of d. We claim that for every integer $d \in D$, its binary weight is (very crudely) at most t (or else its shadow and hence D has more than t elements). It follows that the integer $d = \mathrm{leader}(d) \leq 2^{n(1-1/t)} = N^{1-1/t}$. 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 6 for the affine-invariant case. ▮

4 Proofs of Main theorems

In this section we prove Theorem 4 and Theorem 5. While in our exposition so far we have described the results about affine-invariant codes before the similar results for cyclic-invariant codes, mainly because our results on affine-invariant codes are cleaner to state, in this section we will however prove Theorem 5 before Theorem 4 since the former is simpler to prove. The essential difficulty in the analysis of affine-invariant codes can be better emphasized by comparison to the analysis of cyclic-invariant codes.

4.1 Analysis of the cyclic case

Proof of Theorem 5: Let $\eta = \eta(t, \epsilon)$ and $\eta' = \eta'(t+1, \epsilon)$ be as given by applications of Lemma 8 (for integers t and $t+1$), for the cyclic-invariant case (so codes of length $N-1$ have distance roughly $1/2 - N^{-\eta}$). Let $k_0 = k_0(t, \eta)$ and $k_0' = k_0(t+1, \eta')$ be as given by Theorem 9. We prove the theorem for $k = \max\{k_0, k_0'\}$.

Fix $N = 2^n$ such that n is a large enough 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-invariant code of cardinality at most N^t. Let $D \subseteq \mathcal{D}$ be as given by

Lemma 6, so that $\mathcal{C} = \{\mathrm{Trace}(p) | p \in P_{N-1,D}\}$. For $d \in \mathcal{D} - D$, let $\mathcal{C}(d) = \{\mathrm{Trace}(p) | p \in P_{N-1,D\cup\{d\}}\}$. Our analysis below will show that:

1. Every codeword $w \in \mathcal{C}^\perp \setminus \cup_{d\in\mathcal{D}\setminus D}(\mathcal{C}(d)^\perp)$ generates the code $\mathcal{C}^\perp$ by its cyclic shifts, i.e., $\mathcal{C}^\perp = \mathrm{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}$, and
2. There is a codeword of weight k in $\mathcal{C}^\perp \setminus \cup_{d\in\mathcal{D}\setminus D}(\mathcal{C}(d)^\perp)$.

Putting the two together we get the proof of the theorem.

We start with the first part. Consider any codeword $w \in \mathcal{C}^\perp$. We claim that if $\mathrm{Span}\{w(\alpha x)\} \neq \mathcal{C}^\perp$, then there must exist an element $d \in \mathcal{D} \setminus D$ such that $w \in \mathcal{C}(d)^\perp$. To see this, first note that $\mathrm{Span}\{w(\alpha x)\}$ is a code invariant over the cyclic group, and is contained in $\mathcal{C}^\perp$. Thus if $\mathrm{Span}\{w(\alpha x)\} \neq \mathcal{C}^\perp$ then it must be strictly contained in $\mathcal{C}^\perp$ and so $(\mathrm{Span}\{w(\alpha x)\})^\perp$ must be a strict superset of $\mathcal{C}$. Using Lemma 6 there must exist a set D' such that $(\mathrm{Span}\{w(\alpha x)\})^\perp = P_{N-1,D'}$. Furthermore D' must be a strict superset of D and so there must exist an element $d \in D' \setminus D$. We claim that $w \in \mathcal{C}(d)^\perp$. This is so since $\mathcal{C}(d) \subseteq (\mathrm{Span}\{w(\alpha x)\})^\perp$ and so $w \in (\mathrm{Span}\{w(\alpha x)\}) \subseteq \mathcal{C}(d)^\perp$. This concludes the proof of the first claim.

It remains to show that there is a codeword of weight k in $\mathcal{C}^\perp \setminus \cup_{d\in\mathcal{D}\setminus D}(\mathcal{C}(d)^\perp)$. For this we employ simple counting arguments. We first note that, using Lemma 8, $\mathcal{C}$ is a code satisfying $\frac{1}{2} - N^{-\eta} \le \delta(\mathcal{C}) \le \frac{1}{2} + N^{-\eta}$. Hence we can apply Theorem 9 to conclude that $\mathcal{C}^\perp$ has at least $\binom{N}{k}/(|\mathcal{C}|) \cdot (1 - 1/N^2)$ codewords of weight k. On the other hand, for every fixed $d \in \mathcal{D} \setminus D$, we have (by Lemma 8 again) that $\frac{1}{2} - N^{-\eta'} \le \delta(\mathcal{C}(d)) \le \frac{1}{2} + N^{-\eta'}$. Again applying Theorem 9 we have that $\mathcal{C}(d)^\perp$ has at most $\binom{N}{k}/(|\mathcal{C}(d)|)(1 + 1/N^2)$ codewords 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 codewords of weight k in $\cup_{d\in\mathcal{D}\setminus D}\mathcal{C}(d)^\perp$ from above by

$$\frac{\binom{N}{k}}{2 \cdot |\mathcal{C}|}(1 + \frac{1}{N^2}) + |\mathcal{D}| \cdot \frac{\binom{N}{k}}{N \cdot |\mathcal{C}|}(1 + \frac{1}{N^2}) \le \frac{1}{2|\mathcal{C}|} \cdot \binom{N}{k}(1 + \frac{2}{\log_2 N} + \frac{3}{N^2}),$$

where above we use the fact that $|\mathcal{D}| \le N/\log_2 N$. For sufficiently large N (i.e., when $2/\log_2 N + 3/N^2 \le 1/2$) we have that this quantity is strictly smaller than $\binom{N}{k}/(|\mathcal{C}|) \cdot (1 - 1/N^2)$, which was our lower bound on the number of codewords of weight k in $\mathcal{C}^\perp$. We conclude that there is a codeword of weight k in $\mathcal{C}^\perp \setminus \cup_{d\in\mathcal{D}\setminus D}(\mathcal{C}(d)^\perp)$ as claimed. This concludes the proof of the theorem. ▌

4.2 Analysis of the affine-invariant case

Proof of Theorem 4: The proof is similar to the proof of Theorem 5 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).

Let $\eta = \eta(t)$ and $\eta' = \eta'(t + 1)$ be as given by Lemma 8 for the affine invariant case (so codes of length N have distance roughly $1/2 - N^{-\eta}$). Let $k_0 = k_0(t, \eta)$ and $k_0' = k_0(t + 1, \eta')$ be as given by Theorem 9. We prove the theorem for $k = \max\{k_0, k_0'\}$.

Fix $N = 2^n$ for prime large enough n and let $\mathcal{C}$ be an affine-invariant code of cardinality N^t. Let $D \subseteq \mathcal{D}$ be a set of cardinality at most t and consisting of integers smaller that $N^{1-1/t}$ such that $\mathcal{C} = \{\mathrm{Trace}(p) | p \in P_{N,D}\}$ (as given by Lemma 6). For $d \in \mathcal{D} \setminus D$, let $\mathcal{C}(d) = \{\mathrm{Trace}(p) | p \in P_{N,D\cup\{d\}}\}$. Let $\mathcal{D}' = (\mathcal{D} \setminus D) \cap \{1, \ldots, \lfloor N^{1-\frac{1}{t+1}} \rfloor\}$. We will proceed as in the proof of Theorem 5 with the difference being that now we focus on integers $d \in \mathcal{D}'$. Namely, we first claim that if there is a weight k codeword w in $\mathcal{C}^\perp$ that is not in some $\mathcal{C}(d)^\perp$, for every $d \in \mathcal{D}'$, then $\{\mathrm{Span}(w(\alpha x + \beta) | \alpha \in \mathbb{F}_N^*, \beta \in \mathbb{F}_N\} = \mathcal{C}^\perp$. Then the same counting argument as in the proof of Theorem 5 suffices to show that such a word does exist.

To show the first claim, consider $w \in \mathcal{C}^{\perp}$ and the code $\{\mathrm{Span}(w(\alpha x + \beta) | \alpha \in \mathbb{F}_N^*, \beta \in \mathbb{F}_N\}$, which is affine invariant and so is given by $P_{N,E}$ for some shadow-closed set E. If $\{\mathrm{Span}(w(\alpha x + \beta)\}^{\perp} \neq \mathcal{C}$ then E strictly contains D and so there must exist some element $d' \in E \setminus D$. Now consider the smallest positive integer $d \prec d'$ such that $d \in E \setminus D$. We claim that the binary weight of d must be at most $t+1$. Indeed, this is true since the entire shadow of d' is in E and since elements of D have binary weight at most t. It follows that $d \leq \lfloor N^{1-\frac{1}{t+1}} \rfloor$ and $\mathcal{C}(d)$ is affine invariant. We conclude that $w \in \{\mathrm{Span}(w(\alpha x + \beta)\} \subseteq \mathcal{C}(d)^{\perp}$, yielding the claim.

We now repeat the counting argument from the proof of Theorem 5 to show that there is a codeword of weight k in $\mathcal{C}^{\perp} \setminus (\cup_{d \in \mathcal{D}'} \mathcal{C}(d)^{\perp})$. Using the affine-invariant part of Lemma 8 it follows that $\mathcal{C}$ is a code satisfying $\frac{1}{2} - N^{-\eta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\eta}$. Hence we can apply Theorem 9 to conclude that $\mathcal{C}^{\perp}$ has at least $\binom{N}{k}/(|\mathcal{C}|) \cdot (1 - 1/N^2)$ codewords of weight k. On the other hand, for every fixed $d \in \mathcal{D}'$, we have (by Lemma 8 again) that $\frac{1}{2} - N^{-\eta'} \leq \delta(\mathcal{C}(d)) \leq \frac{1}{2} + N^{-\eta'}$. Again applying Theorem 9 we have $\mathcal{C}(d)^{\perp}$ has at most $\binom{N}{k}/(|\mathcal{C}(d)|)(1 + 1/N^2)$ codewords of weight k. Since $d \leq N^{1-\frac{1}{t+1}}$ it follows that $|\mathcal{C}(d)| = N \cdot |\mathcal{C}|$. Thus we can bound the total number of codewords of weight k in $\cup_{d \in \mathcal{D}'} \mathcal{C}(d)^{\perp}$ from above by

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

where above we use the fact that $|\mathcal{D}'| \leq N / \log_2 N$. For sufficiently large N (i.e., when $\frac{1}{\log_2 N}(1 + \frac{1}{N^2}) \leq 3/4$), we have that this quantity is strictly smaller than $\binom{N}{k}/(|\mathcal{C}|) \cdot (1 - 1/N^2)$, which was our lower bound on the number of codewords of weight k in $\mathcal{C}^{\perp}$. We conclude that there is a codeword of weight k in $\mathcal{C}^{\perp} \setminus \cup_{d \in \mathcal{D}'} (\mathcal{C}(d)^{\perp})$ as claimed, and so this codeword generates $\mathcal{C}^{\perp}$.

∎

5 Implications of Our Results to Property Testing and BCH Codes

5.1 Implications to property testing

It follows from the work of [23] that codes with a single local orbit under the affine symmetry group are locally testable. We recall some basic definitions below and summarize the implication of our main theorem to testability.

Definition 14 (Locally testable code [13]) *A family of codes $\mathcal{C}_N \subseteq \mathbb{F}_2^N$ is $(k(N), \tau(N))$-*locally testable *if there exists a probabilistic algorithm T called the* tester *that, given oracle access to a vector $v \in \mathbb{F}_2^N$ makes at most k queries to the oracle for v and accepts $v \in \mathcal{C}_N$ with probability 1, while rejecting $v \notin \mathcal{C}_N$ with probability at least $\tau(N) \cdot \delta(v, \mathcal{C}_N)$. A family $\mathcal{C}_N$ is said to be* locally testable *if there exist constants $k, \tau > 0$ such that $\mathcal{C}_N$ is (k, τ)-locally testable.*

We note that the above definition corresponds to the strong definition of local testability ([13, Definition 2.2]). We now state the result of [23] on the testability of affine-invariant codes with the single local orbit property.

Theorem 15 (Restatement of Theorem 2.9 from [23]) *If for some fixed integer $k > 0$, $\mathcal{C}_N \subseteq \mathbb{F}_2^N$ is linear and has the k-single orbit property under the affine group, then $\mathcal{C}_N^{\perp}$ is $(k, \frac{1}{2k^2})$-locally testable.*

We note that in [23] the single-orbit property under the affine group is denoted "strong formal characterization." In what follows we make some side comments on the equivalence between Theorem 15 and Theorem 2.9 from [23].

We start by defining "formal characterization" from [23] and then we describe its relation to the notion of single orbit characterization.

Definition 16 (Local affine formal characterization [23]) *$\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ has a k-local affine formal characterization if there exist integer m, and linear functions $\ell_1, \ell_2, \ldots, \ell_k : \mathbb{F}_{2^n}^m \to \mathbb{F}_{2^n}$ (with $\ell_i(y_1, y_2, \ldots, y_m) = y_1 + \sum_{j=2}^{m} \ell_{ij} y_j$) such that*

$$f \in \mathcal{F} \textit{ if and only if } \sum_{i=1}^{k} f(\ell_i(y)) = 0 \textit{ for all } y \in \mathbb{F}_{2^n}^m.$$

Claim 17 *Let $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ have the k-single orbit property under the affine group. 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 $1 \le i \le k$. We show that $f \in \mathcal{F}^\perp$ if and only if $\sum_{i=1}^{k} f(\ell_i(y)) = 0, \ \forall y \in \mathbb{F}_{2^n}^2$.

Since g is a generator, it follows that $\mathcal{F} = \mathrm{Span}\{g \circ \pi \mid \pi(x) = ax + b, a, b \in \mathbb{F}_{2^n}\}$. 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 $\langle h, f \rangle = 0, \forall f \in \mathrm{Span}\{g \circ \pi \mid \pi(x) = ax + b, a, b \in \mathbb{F}_{2^n}$, and hence $h \in \mathcal{F}^\perp$.

▮

Theorem 18 (Theorem 2.9 from [23]) *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 $(k, \frac{1}{2k^2})$-locally testable.*

Proof of Theorem 15: Suppose $\mathcal{C}_N$ has the single-orbit characterization under the affine group. Then by Claim 17 $\mathcal{C}_N^\perp$ has a k-local formal characterization, and by Theorem 18 $\mathcal{C}_N^\perp$ is $(k, \frac{1}{2k^2})$-locally testable.

▮

Returning to the applications of our results to testing, our main theorem, Theorem 4, when combined with Theorem 15 immediately yields the following implication for sparse affine invariant codes.

Corollary 19 *For every constant t there exist constants k, n_0 such that if $n \ge n_0$ is prime, $N = 2^n$ and $\mathcal{C}_N \subseteq \mathbb{F}_2^N$ is a linear, affine-invariant code with at most N^t codewords, then $\mathcal{C}_N$ is $(k, \frac{1}{2k^2})$-locally testable.*

5.2 Implications to BCH codes

In addition to the implications for the testability of sparse affine-invariant codes, our results also give new structural insight into the classical BCH codes. Even though these codes have been around a long time, some very basic questions about them are little understood. We describe the codes, the unanswered questions about them, and the implications of our work in this context below.

We start by defining the BCH codes and the extended-BCH codes. The former are classical cyclic (-invariant) codes, and the latter are affine-invariant. They are also commonly defined as subfield-subcodes of Reed-Solomon codes. We refer to [25, 26] for further details.

Definition 20 *For every pair of integers n and t, the (binary) BCH code with parameters n and t, denoted* $\mathrm{BCH}(n,t) \subseteq \mathbb{F}_2^{N-1}$ *is*

$$\mathrm{BCH}(n,t) = \{\langle f(\alpha)\rangle_{\alpha \in \mathbb{F}_{2^n}^*} | f \in \mathbb{F}_{2^n}[x], \deg(f) \leq N - 2t + 1\} \cap \mathbb{F}_2^{N-1}.$$

The extended BCH code is the evaluation of the same polynomials over the entire field $\mathbb{F}_{2^n}$.

$$\mathrm{eBCH}(n,t) = \{\langle f(\alpha)\rangle_{\alpha \in \mathbb{F}_{2^n}} | f \in \mathbb{F}_{2^n}[x], \deg(f) \leq N - 2t + 1\} \cap \mathbb{F}_2^{N}.$$

The duals of $\mathrm{BCH}(n,t)$ and of $\mathrm{eBCH}(n,t)$ can also be described in terms of traces of sparse polynomials, a standard result in coding theory by now, initially proved by Delsarte [11].

Definition 21 *For every pair of integers n and t, the (binary) dual-BCH code with parameters n and t, denoted* $\mathrm{BCH}(n,t)^{\perp} \subseteq \mathbb{F}_{2^n}^{N-1}$ *consists of the evaluations of traces of polynomials consisting of monomials of odd degrees $\leq 2t-1$ and constant term 0 over $\mathbb{F}_{2^n}^*$, i.e.,*

$$\mathrm{BCH}(n,t)^{\perp} = \{\langle \mathrm{Trace}(f(\alpha))\rangle_{\alpha \in \mathbb{F}_{2^n}^*} | f = \sum_{i=0}^{t-1} f_{2i+1} x^{2i+1} \in \mathbb{F}_{2^n}[x]\}.$$

The extended dual-BCH code $\mathrm{eBCH}(n,t)^{\perp} \subseteq \mathbb{F}_2^N$ *is*

$$\mathrm{eBCH}(n,t)^{\perp} = \{\langle \mathrm{Trace}(f(\alpha))\rangle_{\alpha \in \mathbb{F}_{2^n}^*} | f = f_0 + \sum_{i=0}^{t-1} f_{2i+1} x^{2i+1} \in \mathbb{F}_{2^n}[x], f_0 \in \mathbb{F}_2\}.$$

Even though the BCH codes are very classical codes, much is unknown about them. Some open problems about BCH codes are formulated in Chapter 11 of [10], where it is noted that even the exact minimum distance of BCH codes for relatively small n's (say $n = 511$) is unknown. In the asymptotic regime some similar barriers have also been identified. For instance, while it is easy to see (by a counting argument) that the BCH code $\mathrm{BCH}(n,t)$ must have codewords of weight $2t+1$ [24] (for constant t and large enough n), such words are not known "explicitly" in the general settings (the notion of "explicitness" was discussed in more details in Section 1). Examples of explicit BCH codewords are rare in the literature and only for particular settings of t and n (see for eg. Chapter 9.2 in [25]).

These observations lead to the first question: "What is an explicit low-weight codeword of $\mathrm{BCH}(n,t)$?" Till recently it was not known that the set of codewords of low weight even generate the BCH code, and this was answered affirmatively only recently by Kaufman and Litsyn [19] who showed that words of weight $2t+1$ and $2t+2$ certainly include a basis for the BCH code. This proof remains "non-explicit" and the most "succinct" description of this basis is via $O(tN)$ field elements of $\mathbb{F}_{2^n}$ (i.e. by specifying the $2t+1$ or $2t+2$

non-zero indices (as field elements) of each of the $O(N)$ basis vectors). This leads to the second, harder question: "What is an explicit basis of $\mathrm{BCH}(n, t)$?"

Our result manages to make progress on the second question without making progress on the first, by showing that the affine orbit (or in some cases the cyclic orbit) of a single low-weight codeword gives a basis for the BCH code. While this single codeword is still not explicit, the rest of the basis is explicit given the codeword. We note here that in the recent work of Grigorescu and Kaufman [15] they describe explicit codewords of $\mathrm{BCH}(n, 2)$ that generate the code under the orbit of the affine group. We also remark that understanding what explicit bases for $\mathrm{BCH}(n, t)$ codes look like is a hard problem even when t is not a constant. In particular, in [2] it is shown that when $t = 2^{m-2} - 1$ for any m, the minimum weight codewords of $\mathrm{BCH}(n, t)$ are the same as the minimum weight codewords of punctured Reed-Muller codes of order 2 (which can be described explicitly), but these don't form a basis for the code.

We formalize next the implications discussed above.

Corollary 22 *For every t there exist constants k, n_0 such that for all prime $n \geq n_0$, $\mathrm{eBCH}(n, t)$ has the k-single orbit property under the affine group.*

The above follows from Theorem 4 using the observation that $\mathrm{eBCH}(n, t)^{\perp}$ is sparse (has $2N^t$ codewords) and affine invariant.

Corollary 23 *For every t and $0 < \epsilon < 1$ there exist constants k, n_0 such that for all prime $n \geq n_0$ such that $2^n - 1$ does not have any non-trivial divisors greater than $2^{n(1-\epsilon)}$, $\mathrm{BCH}(n, t)$ has the k-single orbit property under the cyclic group.*

The above follows from Theorem 5 using the observation that $\mathrm{BCH}(n, t)^{\perp}$ is sparse (has N^t codewords) and cyclic-invariant.

We remark that questions of this nature are relevant not only to coding theory, but also to computing. For instance a recurring question in CS is to find explicit balls of small radius in tightly packed codes that contain many codewords. In such problems, the goal is to find an explicit vector (not in the code) along with explicit description of a large set of nearby codewords. Our study, in contrast, attempts to find an explicit description of a large set of codewords near the zero vector (a codeword).

Finally, we point out that the need for various parameters (n and $2^n - 1$) being prime, or not having large divisors, 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. Indeed, as discussed above, in the recent work of Kaufman and Lovett [20] the restriction on n being prime was removed using different techniques. However, the work of [20] could not eliminate the restriction on $2^n - 1$ being prime, or not having large divisors, as required in our result for cyclic-invariant codes.

Acknowledgments

We would like to thank Oded Goldreich for valuable suggestions, and Ronny Roth and the anonymous reviewers for detecting several inconsistencies, omissions, and typos in prior versions of this paper and for pointing out several relevant references from the literature. We thank Swastik Kopparty for helpful discussions.

References

[1] 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.

[2] Daniel Augot, Pascale Charpin, and Nicolas Sendrier. Studying the locator polynomials of minimum weight codewords of BCH codes. *IEEE Transactions on Information Theory*, 38(3):960–973, 1992.

[3] 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.

[4] Jean Bourgain. Mordell's exponential sum estimate revisited. *Journal of the American Mathematical Society*, 18(2):477–499 (electronic), 2005.

[5] 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.

[6] Jean Bourgain and Mei-Chu Chang. A Gauss sum estimate in arbitrary finite fields. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 342(9):643–646, 2006.

[7] Jean Bourgain, Nets Katz, and Terence Tao. A sum-product estimate in finite fields, and applications. *Geometric and Functional Analysis*, 14(1):27–57, 2004.

[8] Jean Bourgain and S. V. Konyagin. Estimates for the number of sums and products and for exponential sums over subgroups in fields of prime order. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 337(2):75–80, 2003.

[9] L. Carlitz and S. Uchiyama. Bounds for exponential sums. *Duke Mathematical Journal*, 24(1):37–41, 1957.

[10] Pascale Charpin. Open problems on cyclic codes. *Handbook of Coding Theory*, 1, 1998.

[11] Phillipe Delsarte. On subfield subcodes of Reed-Solomon codes. *IEEE Transactions on Information Theory*, 21:575–576, 1975.

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

[13] Oded Goldreich and Madhu Sudan. Locally testable codes and PCPs of almost-linear length. *Journal of ACM*, 53(4):558–655, 2006.

[14] Elena Grigorescu. *Symmetries in Algebraic Property Testing*. PhD thesis, MIT, 2010.

[15] Elena Grigorescu and Tali Kaufman. Explicit low-weight bases for bch codes. *IEEE Transactions on Information Theory*, 58(1):78–81, 2012.

[16] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. Succinct representation of codes with applications to testing. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. 13th International Workshop, RANDOM 2009*, pages 534–547, 2009.

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

[18] Tali Kaufman and Simon Litsyn. Almost orthogonal linear codes are locally testable. In *Annual IEEE Symposium on Foundations of Computer Science*, pages 317–326, 2005.

[19] Tali Kaufman and Simon Litsyn. Long extended BCH codes are spanned by minimum weight words. In *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 16th International Symposium,*, pages 285–294, 2006.

[20] Tali Kaufman and Shachar Lovett. New extension of the Weil bound for character sums with applications to coding. In *Annual IEEE Symposium on Foundations of Computer Science*, pages 788–796, 2011.

[21] Tali Kaufman and Dana Ron. Testing polynomials over general fields. *SIAM Journal of Computing*, 36(3):779–802, 2006.

[22] Tali Kaufman and Madhu Sudan. Sparse random linear codes are locally decodable and testable. In *Annual IEEE Symposium on Foundations of Computer Science*, pages 590–600, 2007.

[23] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In Richard E. Ladner and Cynthia Dwork, editors, *Annual ACM Symposium on Theory of Computing*, pages 403–412. ACM, 2008.

[24] Ilia Krasikov and Simon Litsyn. On the distance distributions of BCH codes and their duals. *Designs, Codes and Cryptography*, 23(2):223–232, 2001.

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

[26] Ron M. Roth. *Introduction to coding theory*. Cambridge University Press, 2006.

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

[28] Daniel Shanks. *Solved and unsolved problems in Number Theory*, volume 1. Spartan Books, 1962.

[29] Henning Stichtenoth. *Algebraic function fields and codes*, volume 254. Springer-Verlag, 2nd edition, Berlin, 2009.

[30] Jacobus H. van Lint. *Introduction to Coding Theory*. Graduate Texts in Mathematics **86**, (3rd Edition) Springer-Verlag, Berlin, 1999.

[31] Samuel S. Wagstaff. Divisors of Mersenne numbers. *Mathematics of Computation*, 40(161):385–397, 1983.

[32] A. Weil. Sur les courbes algébriques et les variétés qui s'en déduisent. *Actualités Scientifiques et Industrielles*, (1041), 1948.

[33] Wikipedia. Lenstra-Pomerance-Wagstaff conjecture. http://en.wikipedia.org/wiki/Mersenne_conjectures. Online; accessed 30-August-2012.

[34] J. Wolfmann. Polynomial description of binary linear codes and related properties. *Applicable Algebra in Engineering, Communication and Computing*, 2:119–138, 1991.