Succinct Representation of Codes with Applications to Testing

Elena Grigorescu[1⋆], Tali Kaufman[1⋆⋆], and Madhu Sudan[2⋆⋆⋆]

[1] MIT, Cambridge, MA, USA, {elena_g,kaufmant}@mit.edu
[2] Microsoft Research, Cambridge, MA, USA, madhu@microsoft.com

Abstract. Motivated by questions in property testing, we search for linear error-correcting codes that have the "single local orbit" property: 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 *sparse* if it contains polynomially many codewords in its block length.) In particular this class includes the dual-BCH codes for whose duals (BCH codes) simple bases were not known. Our result gives the first short ($O(n)$-bit, as opposed to $\exp(n)$-bit) description of a low-weight basis for BCH codes. If $2^n - 1$ is a Mersenne prime, then we get that every sparse *cyclic* code also has the single local orbit.

Key words: Locally testable codes, affine/cyclic invariance, single orbit

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, [16, 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 code $C \subseteq \mathbb{F}_2^N$ is said to be locally testable if membership of a word $w \in \mathbb{F}_2^N$ in the code C can be checked probabilitistically by a few probes into w. The famed "linearity test" of Blum, Luby and Rubinfeld [2] may be considered the first result to show that some code is locally testable. Locally testable codes were formally defined by Rubinfeld and Sudan [17]. The first substantial study of locally testable codes was conducted by Goldreich and Sudan [9], where the principal focus was the construction of locally testable codes of high rate. Local

⋆ Funded in part by NSF grant CCR-0829672.
⋆⋆ Funded in part by NSF grant CCR-0829672 and NSF-0729011.
⋆⋆⋆ Research conducted when this author was at MIT CSAIL. Research was funded in part by NSF award CCR-0829672.

testing of codes is effectively equivalent to property testing [17, 8] 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 [2], Reed-Muller codes of various parameters [17, 1, 13, 10], dual-BCH codes [11, 14], 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 [15] is the "invariance group" of the code, which we describe next.

Let $[N]$ denote the set of integers $\{1, \ldots, N\}$. A code $C \subseteq \mathbb{F}_2^N$ is said to be invariant under a permutation $\pi : [N] \to [N]$ if for every $a = \langle a_1, \ldots, a_N \rangle \in C$, it is the case that $a \circ \pi = \langle a_{\pi(1)}, \ldots, a_{\pi(N)} \rangle$ is also in C. The set of permutations under which any code C is invariant forms a group under composition and we refer to it as the invariant group. [15] suggested that the invariant group of a code may play an important role in its testability. They supported their suggestion by showing that if the invariant 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.

Let $N = 2^n$ and let $C \subseteq \mathbb{F}_2^N$ be a code. In this case we can associate the coordinate set $[N]$ of the code C with the field $\mathbb{F}_{2^n}$. Now consider the permutations $\pi : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ of the form $\pi(x) = \alpha x + \beta$ where $\alpha \in \mathbb{F}_{2^n} - \{0\}$ and $\beta \in \mathbb{F}_{2^n}$. This set is closed under composition and we refer to this as the *affine group*. If C is invariant under every π in the affine group, then we say that C is *affine-invariant*. We say that C is *linear* if it is a vector subspace of $\mathbb{F}_2^N$. The *dual* of C, denoted $C^\perp$, is the null space of C as a vector space.

We now define the final term above, namely, the "single local orbit property". Let G be a group of permutations mapping $[N]$ to $[N]$. For $b \in \mathbb{F}_2^N$, let its *weight*, denoted $\mathrm{wt}(b)$, be the number of non-zero elements of b. A code C is said to have the *k-single orbit property* under G if there exists an element $b \in \mathbb{F}_2^N$ of weight at most k such that $C = \mathrm{Span}(\{b \circ \pi | \pi \in G\})$, where $\mathrm{Span}(S) = \{\sum_i c_i b_i | c_i \in \mathbb{F}_2, b_i \in S\}$. Two groups are of special interest to us in this work. The first is the affine group on $\mathbb{F}_{2^n}$. A second group of interest to us is the "cyclic group" on $\mathbb{F}_{2^n}^* = \mathbb{F}_{2^n} - \{0\}$ given by the permutations $\pi_a(x) = ax$ for $a \in \mathbb{F}_{2^n}^*$. (Note that if ω is a multiplicative generator of $\mathbb{F}_{2^n}^*$ and the coordinates of C are ordered $\langle \omega, \omega^2, \ldots, \omega^{2^n - 1} = 1 \rangle$ then each π_a is simply a cyclic permutation.)

The invariance groups of codes are well-studied objects. In particular codes that are invariant under cyclic permutations, known as cyclic codes, are widely studied and include many common algebraic codes (under appropriate ordering of the coordinates and with some slight modifications, see [18] or [16].) The fact that many codes are also affine-invariant is also explicitly noted and used in the literature [16].

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 size

$k \log N$ as opposed to $\omega(N))$[3] of bases for codes (that have the k-single orbit property under some standard group.) Even for such commonly studied codes such as the BCH codes such explicit descriptions of bases were not known prior to this work. In retrospect, the single orbit property was being exploited in previous results in algebraic property testing [2, 17, 1, 13, 10] though this fact was not explicit until the work of [15].

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 (i.e., codes with at most polynomially many codewords in N) has the k-single orbit property under the affine group for some constant k, provided $N = 2^n$ for prime n (see Theorem 1.) When $N - 1$ is also prime, 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 (see Theorem 2.) 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 1 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 [14] who show that sparse high-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. [4]

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 focusses our attention on a collection of closely related codes. We then turn to recent results from additive number theory [4, 3, 6, 5, 7] 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 [14]) this translates to some

[3] One way to represent a sparse code C 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.

[4] In contrast the tester of [14] 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.

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 [2, 17, 1, 13, 10] 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 manages to bring in non-algebraic tools to handle such cases. We believe that the approach is potentially interesting in broader settings.

2 Definitions and main results

We recall some basic notation. $[N]$ denotes the set $\{1, \ldots, N\}$. $\mathbb{F}_q$ denotes the finite field with q elements and $\mathbb{F}_q^*$ will denote the non-zero elements of this field. We will consider codes contained in the vector space $\mathbb{F}_2^N$. 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 $a = \langle a_i \rangle_i$, and $b = \langle b_i \rangle_i \in \mathbb{F}_2^N$ define the *relative distance* between a, b as $\delta(a, b) = \frac{1}{N} |\{i \mid a_i \neq b_i\}|$. Note $\delta(a, b) = \frac{\mathrm{wt}(a-b)}{N}$. A binary code $\mathcal{C}$ is a subset of $\mathbb{F}_2^N$. The (relative) distance of $\mathcal{C}$ is $\delta(\mathcal{C}) = \min_{a,b \in \mathcal{C}; a \neq b} \{\delta(a, b)\}$. 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. $\mathcal{C}$ is a *linear* code if its codewords form a vector space in $\{0, 1\}^N$ over $\mathbb{F}_2$, i.e., if $\mathrm{Span}(\mathcal{C}) = \mathcal{C}$. 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 code $\mathcal{C}^\perp = \{b \in \mathbb{F}_2^N \mid b \cdot a = 0, \ \forall a \in \mathcal{C}\}$. We will alternate between viewing $a \in \mathbb{F}_2^N$ as a vector $a = \langle a_1, \ldots, a_N \rangle$ and as a function $a : D \to \mathbb{F}_2$ where D will be some appropriate domain of size N. Two particular domains of interest to us will be $\mathbb{F}_{2^n}$ and $\mathbb{F}_{2^n}^*$.

2.1 Invariance and the single local orbit property

Let $a \in \mathbb{F}_2^N$ be viewed as a function $a : D \to \mathbb{F}_2$ for some domain D of size N. Let $\pi : D \to D$ be a permutation of D. The π-rotation of a is the function $a \circ \pi : D \to \mathbb{F}_2$ given by $a \circ \pi(i) = a(\pi(i))$ for every $i \in D$.

Let D be a set of size N and let $\mathbb{F}_2^N$ denote the set of functions from $D \to \mathbb{F}_2$. A code $\mathcal{C} \subseteq \mathbb{F}_2^N$ is said to be *invariant* under a permutation $\pi : D \to D$ 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 invariant group of a code.

We will be interested in studying codes that are invariant under some well-studied groups (i.e., whose invariant groups contain some well-studied groups.) Two groups of interest to us are the affine group over $\mathbb{F}_{2^n}$ and the cyclic group

over $\mathbb{F}_{2^n}^*$. In what follows we let $N = 2^n$ and view $\mathbb{F}_2^N$ as the set of functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ and $\mathbb{F}_2^{N-1}$ as the set of functions from $\mathbb{F}_{2^n}^*$ to $\mathbb{F}_2$.

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 invariant 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 $\alpha \in \mathbb{F}_{2^n}^*$.* [5] *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 invariant 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. In particular BCH codes are cyclic and Reed-Muller codes are affine-invariant. Furthermore under a simple "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". We define this term next.

Definition 3 (k-single orbit code). *Let $\mathbb{F}_2^N$ be the collection of functions from D to $\mathbb{F}_2$ for some domain D. Let G be a group of permutations from D to D. 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 2.3.

2.2 Main results

Our main results show that, under certain conditions, duals of "sparse" codes have the single orbit property for small k. By "sparse" we mean that the code has only polynomially many codewords in the length of the codewords.

Our first result considers affine-invariant codes.

Theorem 1 (Single orbit property in affine-invariant codes). *For every $t > 0$ there exists a $k = k(t)$ such that for every prime n 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.*

Next we present our main theorem for cyclic codes.

[5] 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 multiplicative generator of $\mathbb{F}_{2^n}^*$.

Theorem 2 (Single orbit property in cyclic codes). *For every t there exists a k such that the following holds: Let n be such that $2^n - 1$ is prime. Let $\mathcal{C} \subseteq \mathbb{F}_2^{N-1}$ 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.*

We remark that it is not known if there are infinitely many n such that $2^n - 1$ is prime. Of course if there are only finitely many such primes then our theorem becomes "trivial". 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), and so unconditional result should remain interesting.

2.3 Implications to property testing

It follows from the work of [15] 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 4 (Locally testable code [9]). *A code $\mathcal{C} \subseteq \mathbb{F}_2^N$ is* (k, α)-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 $v \in \mathcal{C}$ with probability 1, while rejecting $v \notin \mathcal{C}$ with probability at least $\alpha \cdot \delta(v, \mathcal{C})$. $\mathcal{C}$ is said to be locally testable if there exist $k < \infty$ and $\alpha > 0$ such that $\mathcal{C}$ is (k, α)-locally testable.*

We note that the above definition corresponds to the strong definition of local testability ([9, Definition 2.2].) We now state the result of [15] on the testability of affine-invariant codes with the single local orbit property.

Theorem 3 ([15]). *If $\mathcal{C} \subseteq \mathbb{F}_2^N$ is linear and has the k-single orbit property under the affine group, then $\mathcal{C}$ is $(k, \Omega(1/k^2))$-locally testable.*

We note that in [15] the single-orbit property under the affine group is described as the "strong formal characterization."

Our main theorem, Theorem 1, when combined with the above theorem, immediately yields the following implication for sparse affine invariant codes.

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

2.4 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, and used often in the CS literature, 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 codes, and the latter are affine-invariant. Let Trace : $\mathbb{F}_{2^n} \to \mathbb{F}_2$ be the function $\mathrm{Trace}(x) = x + x^2 + \ldots + x^{2^{n-1}}$. We define the BCH codes by defining their dual.

Definition 5. *For every pair of integers n and t, the (binary) dual-BCH code with parameters n and t, denoted* $\mathrm{BCH}(n,t)^\perp \subseteq \mathbb{F}_2^{N-1}$ *consists of the evaluations of traces of polynomials of degree $2t$ over $\mathbb{F}_{2^n}^*$. I.e.,*

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

The BCH code $\mathrm{BCH}(n,t)$ *is simply the dual of* $\mathrm{BCH}(n,t)^\perp$*. The extended dual-BCH code* $\mathrm{eBCH}(n,t)^\perp \subseteq \mathbb{F}_2^N$ *is simply the evaluation of the same functions over all of $\mathbb{F}_{2^n}$, and* $\mathrm{eBCH}(n,t)$ *is its dual.*

(We note that the more common definition of BCH codes is as the subfield subcodes of Reed Solomon codes, with $\mathrm{BCH}(n,t)$ being the subfield subcodes of RS codes of degree $N - 2t - 1$. But it is a folklore fact that the two definitions are equivalent.)

Even though the BCH codes are very classical codes, much is unknown about them. 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 $2t+1$, such words are not known "explicitly," leading to the first question: "What is an explicit low-weight codeword of $BCH(n,t)$?" Till recently it was not known that the set of codes of low weight even generate the BCH code, and this was answered affirmatively only recently by Kaufman and Litsyn [12] 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 $O(Nt)$ field elements of $\mathbb{F}_{2^n}$. This leads to the second, harder question: "What is an explicit basis of $BCH(n,t)$?"

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 state these implications formally below.

Corollary 2. *For every t there exists a k such that for all prime n,* $\mathrm{eBCH}(n,t)$ *has the k-single orbit property under the affine group.*

The above follows from Theorem 1 using the observation that $\mathrm{eBCH}(n,t)^\perp$ is sparse (has $N^{O(t)}$ codewords) and affine invariant.

Corollary 3. *For every t there exists a k such that for all n such that $2^n - 1$ is prime,* $\mathrm{BCH}(n,t)$ *has the k-single orbit property under the cyclic group.*

The above follows from Theorem 2 using the observation that $\mathrm{BCH}(n,t)^{\perp}$ is sparse (has $N^{O(t)}$ 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 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.

3 Overview of techniques

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

1. We first show that every codeword in the codes we consider are expressible as the Traces of *sparse* polynomials. 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 standard literature in coding theory (and similar steps were employed already in [15].)
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 [14]. 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. 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.

For $d \in \{1, \ldots, N-2\}$, let $\mathrm{orb}(d) = \{d, 2d (\mathrm{mod}\ N-1), 4d (\mathrm{mod}\ N-1), \ldots, 2^{n-1}d (\mathrm{mod}\ N-1)\}$. By the primality of n, we have that $|\mathrm{orb}(d)| = n$ for every d. Let $\text{min-orb}(d)$ denote the smallest integer in $\mathrm{orb}(d)$, and let $\mathcal{D} = \{\text{min-orb}(d) \mid d \in \{1, \ldots, N-2\}\} \cup \{N-1\}$. Note that $|\mathcal{D}| = 1 + (N-2)/n$.

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

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

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

The first step in our analysis of codes invariant over the affine group (resp. 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 $\mathbb{F}_N$ (resp. $P_{N-1,D}$ over $\mathbb{F}_N^*$.)

Lemma 1. *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 $|D| \leq t$ iff $|\mathcal{C}| \leq N^t$.*

Similarly, for every affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_N \to \mathbb{F}_2\}$ of cardinality N^t, 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, $|\mathcal{C}| \leq 2N^t$ iff $|D| \leq t$ and $D \subseteq \{1, \ldots, N^{1-1/t}\}$.

Thus in both cases 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 $\{\mathrm{Trace}(p(x))|p \in P_{N,D}\}$ and $\mathcal{C}_{N-1}(D)$ to denote the code $\{\mathrm{Trace}(p(x))|p \in P_{N-1,D}\}$.

We next use a (small variant of a) theorem due to Bourgain [3] 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 4 ([3]). *For every $\epsilon > 0$ and $r < \infty$, there is a $\delta > 0$ 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 $1 \leq i \leq r$, and (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, [3, 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 [3] (and related papers) have to be changed to get the result in our case, in Appendix A.

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 $N-1$ is prime, then also the codes have large distance. We thus get the following implication.

Lemma 2. *For every t 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. Then the code $\mathcal{C} = \mathcal{C}_N(D)$ satisfies $\frac{1}{2} - N^{-\delta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\delta}$.*

Similarly for every t there exists a δ such that the following holds for for every $N = 2^n$ such that $N-1$ is prime. 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 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 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 [14] 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 $1 \pm \Theta(1/N^c)$ factor, for large enough k, where k is independent of N and depends only on t, c and the δ of Lemma 2.) Specifically they show:

Theorem 5 ([14] 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 code $\mathcal{C} \subseteq \mathbb{F}_2^N$ with at most N^t codewords satisfying $\frac{1}{2} - N^{-\delta} \leq \delta(\mathcal{C}) \leq \frac{1}{2} + N^{-\delta}$ it is the case the $\mathcal{C}^\perp$ has $\binom{N}{k}/|\mathcal{C}| \cdot (1 \pm \Theta(N^{-c})$ codewords of weight k.*

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., $\text{Span}(\{x \circ \pi | \pi\}) \neq \mathcal{C}^\perp$.) 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 code $\mathcal{C}(D')$ also satisfies the same invariant structure as $\mathcal{C}$ and so one can apply Lemma 2 and Theorem 5 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 1 and 2 - see Section 5.

4 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 1, which we defer to the full version, along with the other missing proofs of this section. (We will be using the definitions of the sets $\mathcal{D}$, $P_{N,D}$, and $P_{N-1,D}$ as defined in Section 3 heavily throughout this section.)

We start by recalling some standard properties of the Trace function. Recall that $\text{Trace}(x) = x + x^2 + x^4 + \cdots + x^{2^{n-1}}$. The Trace function is linear, i.e. $\text{Trace}(\alpha+\beta) = \text{Trace}(\alpha) + \text{Trace}(\beta)\ \forall \alpha, \beta \in \mathbb{F}_N$. 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 3. *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 $w(x) = \mathrm{Trace}(p(x))$.*

Lemma 4. *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 $\beta \in \mathbb{F}_N$, $\mathrm{Trace}(\beta x^e) \in \mathcal{C}$.

We now use Lemma 4 to characterize cyclic invariant families, while also working towards the characterization of affine invariant families.

Lemma 5. *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 $\mathcal{C} = \{\mathrm{Trace}(p) | p \in P_{N-1,D}\}$.

Lemma 5 essentially suffices to yield Lemma 1 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 6. *If $\mathcal{C}$ is an affine-invariant code, $\mathrm{Trace}(x^d) \in \mathcal{C}$ and $e \prec d$ then $\mathrm{Trace}(x^e) \in \mathcal{C}$.*

5 Proofs of Main theorems

5.1 Analysis of the cyclic case

Proof (of Theorem 2). Let $\delta = \delta(t)$ and $\delta' = \delta'(t+1)$ be as given by Lemma 2 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 5. We prove the theorem for $k = \max\{k_0, k_0'\}$.

Fix N so that $N-1$ is prime and 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 as given by Lemma 1, 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 codeword in $w \in \mathcal{C}^\perp - \cup_{d\in\mathcal{D}-D}(\mathcal{C}(d)^\perp)$ generates the code $\mathcal{C}^\perp$ by its cyclic shifts, i.e., $\mathcal{C}^\perp = \text{Span}\{w(\alpha x) | \alpha \in \mathbb{F}_N^*\}$, and (2) There is a codeword 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 codeword $w \in \mathcal{C}^\perp$. We claim that if $\text{Span}\{w(\alpha x)\} \neq \mathcal{C}^\perp$, 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 $\text{Span}\{w(\alpha x)\}$ is a code invariant under the cyclic group, and is contained in $\mathcal{C}^\perp$. Thus if $\text{Span}\{w(\alpha x)\} \neq \mathcal{C}^\perp$ then it must be strictly contained in $\mathcal{C}^\perp$ and so $(\text{Span}\{w(\alpha x)\})^\perp$ must be a strict superset of $\mathcal{C}$. Using Lemma 1 there must exist a set D' such that $(\text{Span}\{w(\alpha x)\})^\perp = P_{N-1,D'}$. Furthermore D' must be a strict superset of D and so there must exist an element $d \in D' - D$. We claim that $w \in \mathcal{C}(d)^\perp$. This is so since $\mathcal{C}(d) \subseteq (\text{Span}\{w(\alpha x)\})^\perp$ and so $w \in (\text{Span}\{w(\alpha x)\}) \subseteq \mathcal{C}(d)^\perp$. This concludes the proof of the first claim.

It remains to show that there is a codeword of weight k in $\mathcal{C}^\perp - \cup_{d\in\mathcal{D}-D}(\mathcal{C}(d)^\perp)$. For this we employ simple counting arguments. We first note that, using Lemma 2, 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 5 to conclude that $\mathcal{C}^\perp$ has at least $\binom{N}{k}/(|\mathcal{C}|) \cdot (1 - O(1/N^2))$ codewords of weight k. On the other hand, for every fixed $d \in \mathcal{D} - D$, we have (by Lemma 2 again) $\frac{1}{2} - N^{-\delta'} \leq \delta(\mathcal{C}(d)) \leq \frac{1}{2} + N^{-\delta'}$. Again applying Theorem 5 we have $\mathcal{C}(d)^\perp$ has at most $\binom{N}{k}/(|\mathcal{C}(d)|)(1 + O(1/N^2))$ 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 $\cup_{d\in\mathcal{D}-D}\mathcal{C}(d)^\perp$ from above by

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

where above we use the fact that $|\mathcal{D}| \leq N/\log_2 N$. For sufficiently large N (i.e., when $1/\log_2 N + O(1/N^2) \leq 1/2$) we have that this quantity is strictly smaller than $\binom{N}{k}/(|\mathcal{C}|) \cdot (1 - O(1/N^2))$, 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 $\mathcal{C}^\perp - \cup_{d\in\mathcal{D}-D}(\mathcal{C}(d)^\perp)$ as claimed. This concludes the proof of the theorem.

5.2 Analysis of the affine-invariant case

Proof (of Theorem 1). The proof is similar to the proof of Theorem 2 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.) Details below.

Given t, let δ be from Lemma 2 and let k be large enough for application of Theorem 5. Fix $N = 2^n$ for prime n and 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 $\mathcal{C} = \{\text{Trace}(p)|p \in P_{N,D}\}$ (as given by Lemma 1.) 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 2 we argue that if there is a weight k codeword 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 2 suffices to show that such a word does exist.

Consider $w \in \mathcal{C}^\perp$ and the code $\{\text{Span}(w(\alpha x + \beta)|\alpha \in \mathbb{F}_N^*, \beta \in \mathbb{F}_N\}$, which is affine invariant and so is given by $P_{N,E}$ for some shadow-closed set E. If $\{\text{Span}(w(\alpha x + \beta)\}^\perp \neq \mathcal{C}$ then E strictly contains D and so there must exist some element $d' \in E - D$. Now consider 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 codeword 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 2 except that we use the affine-invariant part of Lemma 2. This completes the proof of Theorem 1.

Acknowledgments

We would like to thank Oded Goldreich for valuable suggestions and anonymous reviewers for detecting several omissions and errors in prior versions of this paper. 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. 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.
3. J. Bourgain. Mordell's exponential sum estimate revisited. *J. Amer. Math. Soc.*, 18(2):477–499 (electronic), 2005.
4. J. 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.
5. J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.*, 14(1):27–57, 2004.
6. 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.
7. 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. *C. R. Math. Acad. Sci. Paris*, 337(2):75–80, 2003.

8. Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *JACM*, 45(4):653–750, 1998.
9. 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*.
10. 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.
11. Tali Kaufman and Simon Litsyn. Almost orthogonal linear codes are locally testable. In *FOCS*, pages 317–326. IEEE Computer Society, 2005.
12. Tali Kaufman and Simon Litsyn. Long extended BCH codes are spanned by minimum weight words. In Marc P. C. Fossorier, Hideki Imai, Shu Lin, and Alain Poli, editors, *AAECC*, volume 3857 of *Lecture Notes in Computer Science*, pages 285–294. Springer, 2006.
13. Tali Kaufman and Dana Ron. Testing polynomials over general fields. *SIAM J. Comput.*, 36(3):779–802, 2006.
14. Tali Kaufman and Madhu Sudan. Sparse random linear codes are locally decodable and testable. In *FOCS*, pages 590–600. IEEE Computer Society, 2007.
15. Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In Richard E. Ladner and Cynthia Dwork, editors, *STOC*, pages 403–412. ACM, 2008.
16. F. J. MacWilliams and Neil J. A. Sloane. *The Theory of Error-Correcting Codes*. Elsevier/North-Holland, Amsterdam, 1981.
17. Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, April 1996.
18. Jacobus H. van Lint. *Introduction to Coding Theory*. Graduate Texts in Mathematics **86**, (Third Edition) Springer-Verlag, Berlin, 1999.

A On using results from additive number theory

As pointed out earlier Theorem 7 of [3] only considers the analog of Theorem 4 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 [3] the proof reduces to the two cases $r = 1$ and $r = 2$. The case $r = 1$ in the prime case was obtained in [7]. In our case, where $N = 2^n$, the $r = 1$ case was shown in [6]. For $r = 2$ the proof in the prime case applied the sum-product theorem from [5] and uses Proposition 1 of [4]. We note that Proposition 1 of [4] works also when the field is not of prime cardinality. As argued in [5], 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 [4] 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 [4], and so does the rest of the proof.