Limitations on Testable Affine-Invariant Codes in the High-Rate Regime

Venkatesan Guruswami[*] Madhu Sudan[†] Ameya Velingker[‡] Carol Wang[§]

Abstract

Locally testable codes (LTCs) of constant minimum (absolute) distance that allow the tester to make a nearly linear number of queries have become the focus of attention recently, due to their connections to central questions in approximability theory. In particular, the binary Reed-Muller code of block length N and absolute distance d is known to be testable with $O(N/d)$ queries, and has a dimension of $\approx N - (\log N)^{\log d}$. The polylogarithmically small co-dimension is the basis of constructions of small set expanders with many "bad" eigenvalues, and size-efficient PCPs based on a shorter version of the long code. The smallest possible co-dimension for a distance d code (without any testability requirement) is $\approx \frac{d}{2} \log N$, achieved by BCH codes. This raises the natural question of understanding where in the spectrum between the two classical families, Reed-Muller and BCH, the optimal co-dimension of a distance d LTC lies — in other words the "price" one has to pay for local testability.

One promising approach for constructing LTCs is to focus on affine-invariant codes, whose structure makes testing guarantees easier to deduce than for general codes. Along these lines, the authors of [HRZS13] and [GKS13] recently constructed an affine-invariant family of high-rate LTCs with slightly smaller co-dimension than Reed-Muller codes. In this work, we show that their construction is essentially optimal among linear affine-invariant LTCs that contain the Reed-Muller code of the appropriate degree.

1 Introduction

Locally testable codes (LTCs) have received much attention in recent years. They are error-correcting codes equipped with a *tester*, a randomized algorithm that queries the received word at a few judiciously chosen positions and decides whether the word is a valid codeword. The tester must accept valid codewords with probability 1 and reject words that are far from the code in Hamming distance with nontrivial probability. LTCs have garnered much interest due to their connections to probabilistically checkable proofs (PCPs) and property testing (see the surveys [Gol11, Tre04]). Many PCP constructions are based on or related to LTCs [BSGH$^+$06, GS06, Din07, BSS08]. The primary focus thus far has been on LTCs in which the number of queries is *constant*, and much progress has been made on constructions in this regime (see for example the line of work culminating in [Vid13]). There has also been work on LTCs with a sub-linear number of queries (i.e., N^ϵ queries where N is the block length and $\epsilon > 0$ is arbitrary) [BSS06, GKS13].

The study of locally testable codes naturally fits in the intersection of discrete mathematics and computer science. Recently, high-rate LTCs, in which the tester is allowed to make a nearly *linear* number of queries (i.e., ϵN queries), have been shown to have surprising connections to central questions in the theory of approximation algorithms. This renewed interest leads to new questions about the limits of LTCs that motivate the main question of this work. Specifically, in [BGH$^+$12] a beautiful connection between such LTCs and the construction of small set expander graphs is presented. Instantiating this connection with the binary Reed-Muller (RM) code, the authors of [BGH$^+$12] construct small set expanders whose Laplacian has many small eigenvalues. They also derandomize the "long code" (hypercube) which underlies all optimal PCP constructions to give a shorter low-degree version (which they called the "short code"). The low-degree long code has since been used to construct more size-efficient PCPs, leading to improved hardness results for hypergraph coloring [DG13, GHH$^+$14, KS14].

The binary Reed-Muller code RM(r, n) of degree

[*] Carnegie Mellon University, guruswami@cmu.edu. Part of this work was done while visiting Microsoft Research New England. Research supported in part by NSF grant CCF-0963975.

[†] Microsoft Research New England, madhu@mit.edu.

[‡] Carnegie Mellon University, avelingk@cs.cmu.edu. Part of this work was done while visiting Microsoft Research New England. Research supported by NSF grant CCF-0963975 and MSR-CMU Center for Computational Thinking.

[§] Carnegie Mellon University, wangc@cs.cmu.edu. Research supported by NSF grant CCF-0963975 and NSF Graduate Fellowship.

r in n variables encodes a (multilinear) polynomial $f \in \mathbb{F}_2[X_1, \ldots, X_n]$ of total degree at most r by the vector of its evaluations $\big(f(\alpha)\big)_{\alpha \in \mathbb{F}_2^n}$. The (minimum) distance of $\mathrm{RM}(r, n)$ equals 2^{n-r}. A central ingredient in the above exciting recent developments is a local testability result for binary RM codes due to [BKS$^+$10]. In the high-rate regime of relevance to the above connections, the result of [BKS$^+$10] shows the following (one should think of s as constant, and n as growing in the statement below):

Theorem 1.1. ([BKS$^+$10]) *There exists an absolute constant $\xi > 0$ such that the Reed-Muller code $\mathrm{RM}(n-s, n)$ (of distance 2^s) can be tested with 2^{n-s+1} queries, rejecting a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ that is $2^s/3$-far from $\mathrm{RM}(n-s, n)$ with probability at least ξ.*

The n-variate binary RM code of constant distance d, namely $\mathrm{RM}(n - \log d, n)$, has dimension $\approx N - (\log N)^{\log d - 1}$, where $N = 2^n$, and is testable with $2N/d$ queries.[1] For the connection to small set expansion in [BGH$^+$12], a binary linear code $\mathcal{C}$ of block length N that is testable with ϵN queries results in a graph with vertex set $\mathcal{C}^\perp$ (the dual code to $\mathcal{C}$) whose Laplacian has $\Omega(N)$ eigenvalues smaller than $O(\epsilon)$. To get many "bad" eigenvalues as a function of the graph size, we would like $\mathcal{C}^\perp$ to be small compared to N, i.e., we would like the dimension of $\mathcal{C}$ to be as large as possible. This leads to the following question, which motivates our work: *What is the largest dimension of a distance d binary linear code $\mathcal{C} \subset \mathbb{F}_2^N$ that is testable with $O(N/d)$ queries?*

Reed-Muller codes give a construction with dimension $\approx N - (\log N)^{\log d - 1}$. Achieving higher dimension would imply small set expanders (SSEs) whose Laplacians have even larger number of small eigenvalues, and in particular, a dimension of $N - O_d(\log N)$ would imply polynomially many small eigenvalues (the existence of such SSEs is necessary if the SSE intractability hypothesis of [RS10] holds). The only known upper bound on dimension is the Hamming bound $\approx N - \frac{d}{2}\log N$, based just on the distance (*without* using the testability condition). BCH codes achieve (up to lower order terms) the Hamming bound; however, as all codewords in the dual of the BCH code have Hamming weight close to $N/2$, the BCH code is not testable with $O(N/d)$ queries.[2]

In other words, there is a gap between the dimension of the testable distance d Reed-Muller code, which is $\approx N - (\log N)^{\log d - 1}$, and the dimension of the BCH code of distance d, which is $\approx N - \frac{d}{2}\log N$ (optimal for distance d). The natural question motivating this work is to understand how significant a limitation the testability requirement poses on the dimension, and whether the highest possible dimension of a *testable* code with distance d is closer to that of BCH or RM.

Unfortunately, this seems to be a difficult problem in general. Our understanding of upper bounds on the dimension of LTCs in general is fairly limited. Works such as [BGK$^+$10] and [BSV12] have made some progress in showing limitations of constant query LTCs, but the existence of asymptotically good linear LTCs has still not been ruled out. In fact, it is consistent with current knowledge that there are constant query LTCs whose rate vs relative distance trade-off is close to the Gilbert-Varshamov bound.

Given this state of affairs regarding limits of LTCs, as a first step toward our challenging goal, in this work we focus on proving limitations in the special case of *affine-invariant codes.* Affine-invariance generalizes many popular families of algebraic codes and is a well-studied concept in coding theory. The investigation of the role of affine-invariance, and invariance in general, in the context of testability were initiated by Kaufman and Sudan [KS07] and there have been many further works in the area (see, for instance, the survey by Sudan [Sud11, Section 5] and references therein). Affine-invariant codes are subsets of functions from $\mathbb{F}_Q^n$ to $\mathbb{F}_q$ that are invariant under affine transformations of the domain, where $\mathbb{F}_Q$ and $\mathbb{F}_q$ are finite fields with $\mathbb{F}_Q$ extending $\mathbb{F}_q$ (see Section 2.1 for a more formal definition in the case of $Q = q$).

As it turns out, both Reed-Muller and BCH are affine-invariant codes. Furthermore, Guo et al. [GKS13] as well as Haramaty et al. [HRZS13] show constructions of additional classes of codes that are testable with $O(N/d)$ queries and provide slight improvements to the dimension of the Reed-Muller code. Interestingly, these improved codes, too, are affine-invariant. It seems worthwhile, therefore, to initially restrict our attention to affine-invariant codes and gain further insights into the problem. The rich structure of affine-invariance gives us some handle for understanding the constraints imposed by local testability. For example, although we know virtually no lower bounds for LTCs in the constant-query regime, it was shown in [BSS11] that affine-invariant LTCs for a constant number of queries cannot have constant rate.

Affine-invariance also offers many advantages for *constructing* locally testable codes. It turns out that their structure means that only fairly weak conditions have to be satisfied in order for a code to be testable. For example, it has been shown that *any*

[1] The dimension of $\mathrm{RM}(n - \log d, n)$ can be calculated as $\sum_{i=0}^{n-\log d} \binom{n}{i} \approx N - \left(\frac{en}{\log d - 1}\right)^{\log d - 1}$. Logs are base 2.

[2] It is known that for linear codes, one can assume without loss of generality that the tester checks orthogonality to a set of dual codewords (see [BSHR05]).

affine-invariant linear code which is characterized by constant-weight constraints is testable with constantly many queries [KS07].[3]

In the constant distance (linear query) regime, affine-invariant codes have yielded LTCs with the highest dimension known thus far, and improving slightly upon binary Reed-Muller codes. By using a technique known as *lifting* of affine-invariant codes, the works [GKS13, HRZS13] give constructions of a class of affine-invariant linear-query LTCs that improve upon the dimension of the RM code. For some of these codes, with lower dimension, [HRZS13] shows the soundness guarantee that is necessary to allow them to replace the RM code in the application of [BGH+12]. Without this stronger guarantee, [GKS13] gives a code $\mathcal{C} \subseteq \{0,1\}^N$ of distance d and dimension

$$\dim(\mathcal{C}) \geq N - \left(1 + \frac{\log N}{\log d - 1}\right)^{\log d - 1} \tag{1.1}$$

that is testable with $2N/d$ queries (where $N = 2^n$). This code contains the binary code $\mathrm{RM}(n - \log d, n)$, as do the corresponding codes of [HRZS13]. Hence, it is natural to ask for the optimal dimension of a code containing the RM code that still has the desired testability properties. Note that the (extended) BCH code of distance d (which does not satisfy the testability requirements) also contains $\mathrm{RM}(n - \log d, n)$.

In this work, we show that the code of [GKS13] is essentially optimal. That is, we show for constant d that any linear affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ of distance d which is testable with $2N/d$ queries (the number of queries needed for testing the RM code of the same distance) and contains $\mathrm{RM}(n - \log d, n)$ has dimension at most

$$\dim(\mathcal{C}) \leq N - \left(\frac{\log N}{\log^2 d}\right)^{\log d - 1},$$

where $N = 2^n$ (see Theorem 4.4 for the formal statement of the result). We also show that any linear affine-invariant code $\mathcal{C}$ satisfying (1.1) must contain the RM code of degree $\log N - (\log d - 1)\log(n + \log d - 1) + \Omega_d(1)$, implying that our assumptions are not far from the truth.

Our results suggest that any LTC which improves noticeably on the Reed-Muller code in the linear query regime would need techniques beyond the known ones based on affine-invariance.

Paper Organization. In Section 2, we give definitions and preliminaries on affine-invariant LTCs. Section 3 then describes previous work that complements our results. In Section 4, we prove our lower bound on affine-invariant codes that contain high-order Reed-Muller codes. Finally, Appendix B provides some justification for why containment of a high-order Reed-Muller code is a reasonable assumption. Additional omitted proofs appear in the appendices.

2 Preliminaries

2.1 Our Setup We begin by defining some basic terms about codes.

Definition 2.1. (Dual of a code) *If $\mathcal{C} \subseteq \mathbb{F}^N$ is a linear code, then the* dual code *of $\mathcal{C}$ is $\mathcal{C}^\perp = \{u \in \mathbb{F}^N : \langle u, w \rangle = 0$ for all $w \in \mathcal{C}\}$.*

Definition 2.2. (δ-far) *A word $w \in \mathbb{F}^N$ is said to be* δ-far *from a linear code $\mathcal{C} \subseteq \mathbb{F}^N$ if $\min_{c \in \mathcal{C}} \Delta(w, c) \geq \delta N$, where $\Delta(x, y)$ denotes the Hamming distance between two vectors.*

We now define the notion of a (weak) LTC and a canonical tester.

Definition 2.3. (Canonical testers) *Suppose $\mathcal{C} \subseteq \mathbb{F}^N$ is a linear code. A k*-query canonical tester *for $\mathcal{C}$ is a distribution $\mathcal{D}$ over subsets $I \subseteq \{1, 2, \ldots, N\}$ satisfying $|I| \leq k$; invoking the tester on a word $w \in \mathbb{F}^N$ consists of sampling $I \sim \mathcal{D}$ and accepting w if and only if $w|_I \in \mathcal{C}|_I$.*

Definition 2.4. (LTCs) *A linear code $\mathcal{C} \subseteq \mathbb{F}^N$ is said to be a* (k, ϵ, ρ)-LTC *if there exists a probabilistic algorithm that queries at most k positions of an input word $w \in \mathbb{F}^N$ and (1) accepts with probability 1 if $w \in \mathcal{C}$, and (2) rejects with probability at least ϵ if w is ρ-far from $\mathcal{C}$.*

It is known that a necessary condition for a linear code to be testable is the existence of a dual codeword of low Hamming weight. The proof of the following fact appears in Appendix D.

Fact 2.1. (Low weight dual word) *Let $\mathcal{C} \subseteq \mathbb{F}^N$ be a linear LTC that is testable with k queries. Then, there must exist a nonzero $w \in \mathcal{C}^\perp$ such that $|\{i \in \{1, \ldots, N\} : w_i \neq 0\}| \leq k$, i.e., w has Hamming weight at most k.*

In this work, we will view binary linear codes of block length 2^n as functions $\mathbb{F}_{2^n} \to \mathbb{F}_2$, and we will often write $N = 2^n$. All logarithms will be base 2 unless otherwise specified.

We next define affine-invariant codes, which are the focus of this work.

[3] In fact, the testability also extends to non-linear codes [BFH+13], but with an enormous price in the error analysis.

Definition 2.5. *Let $\mathbb{F}_Q$ be a field of size Q. We call a function $A : \mathbb{F}_Q^t \to \mathbb{F}_Q^t$ an* affine transformation *if $A(x) = Mx + b$ for some matrix $M \in \mathbb{F}_Q^{t\times t}$ and vector $b \in \mathbb{F}_Q^t$.*

Definition 2.6. *Let $\mathbb{F}_q$ be a field of size q, and let $\mathbb{F}_Q$ be its extension field of size $Q = q^m$. Then, we call a code $\mathcal{F} \subseteq \{\mathbb{F}_Q^t \to \mathbb{F}_q\}$* affine-invariant *if for every $f \in \mathcal{F}$ and affine transformation $A : \mathbb{F}_Q^t \to \mathbb{F}_Q^t$, the function $f \circ A$ is in $\mathcal{F}$.*

The task is to consider binary affine-invariant codes $\mathcal{C} \subseteq \{f : \mathbb{F}_{2^n} \to \mathbb{F}_2\}$ with fixed distance d such that $\mathcal{C}$ is an LTC with locality $O\left(\frac{N}{d}\right)$. We wish to find the optimal rate of such a code $\mathcal{C}$.

2.2 Affine-Invariant Codes

From now on, we will only consider *univariate* affine-invariant codes, that is, subsets of $\{f\colon \mathbb{F}_{2^n} \to \mathbb{F}_2\}$. This is without loss of generality, as $(\mathbb{F}_Q)^t$ is isomorphic to $\mathbb{F}_{Q^t}$ for all t and prime powers Q, and this preserves affine-invariance ([BSS11]). testability is preserved.

We refer the reader to Appendix A for basic facts about affine-invariant codes and their degree sets. Here we repeat the most important facts, specialized to our case in which $q = 2$.

Definition 2.7. *For a function $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$, write it as the unique polynomial $f(x) = \sum_{e=0}^{2^n-1} c_e x^e$ of degree at most $2^n - 1$ which agrees with f on $\mathbb{F}_{2^n}$. Then, the* support *of f, denoted* $\mathrm{Supp}(f)$*, is the set of degrees with non-zero coefficients in f, that is,* $\mathrm{Supp}(f) = \{e : c_e \neq 0\}$.

Definition 2.8. *Let $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be a code. We define* $\mathrm{Deg}(\mathcal{F})$*, the* degree set *of $\mathcal{F}$, to be the set* $\mathrm{Deg}(\mathcal{F}) = \bigcup_{f\in\mathcal{F}} \mathrm{Supp}(f)$.

Theorem 2.1. *If $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ is a linear affine-invariant code, then* $\dim(\mathcal{F}) = |\mathrm{Deg}(\mathcal{F})|$.

3 Background and Previous Work

We now state some results on binary affine-invariant codes that motivate our work.

Definition 3.1. *The* 2-weight *of a degree $e \in \{0, 1, \ldots, 2^n-1\}$, denoted* $\mathrm{wt}_2(e)$*, is the number of ones in the binary representation of e.*

Recall the definition of a trace code (Definition A.3, with $q = 2$). It is a folklore fact that the Reed-Muller code is equivalent to the univariate code

$$\mathrm{RM}(r, n) = \mathcal{T}(\{e \in \{0, 1, \ldots, 2^n - 1\} : \mathrm{wt}_2(e) \le r\}).$$

Furthermore, the dual of the extended BCH code of distance $d = 2t + 2$ can be expressed as

$$\text{dual-eBCH}(n, t) = \mathcal{T}(\{0, 1, \ldots, t\}).$$

Similarly, the extended BCH code itself is expressible as

$$\mathrm{eBCH}(n, t) = \mathcal{T}(D),$$

where $D \subseteq \{0, 1, \ldots, 2^n - 1\}$ is the set of all degrees e such that the zeros in the n-bit binary representation of e do not all lie within a cyclic block of length $\log d - 1$. Note that we have

$$\mathrm{RM}(n - \log d, n) \subseteq \mathrm{eBCH}(n, t),$$

and both are linear affine-invariant codes of distance d. Moreover, $\mathrm{RM}(n - \log d, n)$ has dimension

$$\sum_{i=0}^{n-\log d} \binom{n}{i} \approx N - \left(\frac{en}{\log d - 1}\right)^{\log d - 1},$$

while $\mathrm{eBCH}(n, t)$ has rate roughly $N - \frac{dn}{2}$. However, $\mathrm{RM}(n - \log d, n)$ can be tested with $\frac{2N}{d}$ queries [BKS+10, AKK+05]; on the other hand, we cannot hope to test $\mathrm{eBCH}(n, t)$ with the same number of queries (for $d > 4$), due to Fact 2.1 and the fact that dual-eBCH(n, t) has relative distance close to $1/2$ (see [MS81]).

3.1 Testable Codes Surpassing Reed-Muller

Guo, et al. [GKS13] show the existence of linear affine-invariant codes of linear locality that contain the generalized Reed-Muller code of appropriate order. In particular, for $n = \ell m$, where $\ell = \log d - 1$ and m is any positive integer, they present a multivariate affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_{2^\ell}^m \to \mathbb{F}_2\}$ of block length $N = 2^n$ which satisfies $\dim(\mathcal{C}) = N - (m+1)^\ell = N - \left(1 + \frac{n}{\log d - 1}\right)^{\log d - 1}$.

There is also a univariate analogue of the above codes with identical distance and dimension. See E.1 of Appendix E for details.

3.2 Consequence of the Extended Weil Bound

In [KL11], the authors prove an extension of the Weil bound, which implies that sparse linear affine-invariant codes have relative distance close to $1/2$. This yields a lower bound on the dimension of any sparse linear affine-invariant code that has relative distance much less than $1/2$.

Theorem 3.1. (consequence of [KL11]) *Let $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be a linear affine-invariant code of relative distance $\le \frac{1}{2} - \delta$, for some $\delta > 0$. Then for any $\epsilon > 0$, $|\mathcal{F}| \ge 2^{\Omega(n^{\frac{3}{2}-\epsilon})}$, i.e., $\dim(\mathcal{F}) = \Omega(n^{\frac{3}{2}-\epsilon})$.*

This theorem does not appear explicitly in [KL11], but it can be deduced from their techniques. For completeness, we include the details in Appendix C.

Because we are interested in very large codes $\mathcal{C}$ whose duals are sparse, we can apply Theorem 3.1 to $\mathcal{F} = \mathcal{C}^{\perp}$ to obtain an upper bound on the dimension of $\mathcal{C}$.

Corollary 3.1. *If $\mathcal{C}$ is a linear affine-invariant code of distance $d \geq 5$ testable with $\frac{2N}{d}$ queries, then* $\dim(\mathcal{C}^{\perp}) \geq n^{3/2-o(1)}$.

Remark 3.1. *Although we are able to prove much stronger lower bounds in the following section, our results only hold when $\mathcal{C}^{\perp}$ contains (the indicator of) a low-dimensional affine subspace. The work of [KL11] does not require this assumption.*

4 Upper Bounds on the Dimension of $\mathcal{C}$

By Theorem A.2, to show that $\dim(\mathcal{C})$ is small, it suffices to show that $\operatorname{Deg}(\mathcal{C})$ is small. Thus, we will show that under our assumptions, we can find *many* degrees which cannot be in $\operatorname{Deg}(\mathcal{C})$.

We will assume throughout that $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ is affine-invariant, contains the Reed-Muller code $\mathrm{RM}(n - \log d, n)$, and is testable with $2N/d$ queries. Note that the containment assumption implies that $\operatorname{Deg}(\mathcal{C})$ contains all degrees of 2-weight at most $n-\log d$. Furthermore, Fact 2.1 guarantees the existence of some $f \in \mathcal{C}^{\perp} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ of Hamming weight $2N/d$. Since $\mathcal{C}^{\perp}$ is affine-invariant, we have that $g = f \circ A \in \mathcal{C}^{\perp}$ for any affine transformation $A : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$. In particular, choose A to be an invertible transformation that maps 0 to some $x \in \mathbb{F}_{2^n}$ with $f(x) \neq 0$. Then, g has Hamming weight $2N/d$ and is supported on 0.

Since $\mathcal{C}^{\perp}$ is contained in the dual of $\mathrm{RM}(n-\log d, n)$, all dual codewords of Hamming weight $2N/d$ correspond to (indicators of) affine subspaces of dimension $n - \log d + 1$ (see [PHE98]). Therefore, g must be (the indicator of) an affine subspace S of dimension $n - \log d + 1$. Moreover, since g is supported on 0, S must in fact be a *linear* subspace.

4.1 Matrix Determinant Formulation

Recall that an affine-invariant code is specified by its degree set. Thus, if $e \in \operatorname{Deg}(\mathcal{C})$ and the indicator vector of a subspace S is in the dual code $\mathcal{C}^{\perp}$, then we must have

$$\sum_{\alpha \in S} \alpha^e = 0. \tag{4.2}$$

We will often abuse notation and say that if (4.2) holds, then S is *orthogonal* to e, or e *passes* S.

We have assumed that any degree e of 2-weight at most $n - \log d$ is in $\operatorname{Deg}(\mathcal{C})$. Thus, let us consider which degrees e of 2-weight exactly $n - \log d + 1$ can be contained in $\operatorname{Deg}(\mathcal{C})$. The following lemma, whose proof appears in Appendix D, gives an equivalent condition for when a subspace of dimension $n - \log d + 1$ is orthogonal to e.

Lemma 4.1. *Suppose $e = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_k}$ is a degree of 2-weight $k = n - \log d + 1$, for i_j distinct. Suppose S is a subspace of dimension k, and let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be an $\mathbb{F}_2$-basis for S. Then S is orthogonal to e if and only if the following determinant is zero:*

$$M_e(\alpha_1, \alpha_2, \ldots, \alpha_k) := \begin{pmatrix} \alpha_1^{2^{i_1}} & \alpha_2^{2^{i_1}} & \cdots & \alpha_k^{2^{i_1}} \\ \alpha_1^{2^{i_2}} & \alpha_2^{2^{i_2}} & \cdots & \alpha_k^{2^{i_2}} \\ \vdots & \vdots & & \vdots \\ \alpha_1^{2^{i_k}} & \alpha_2^{2^{i_k}} & \cdots & \alpha_k^{2^{i_k}} \end{pmatrix}$$

4.2 Containment in Extended BCH

The determinant formulation of Lemma 4.1 allows us to show that a code satisfying our desired conditions must lie inside the extended BCH code of the same distance.

Theorem 4.1. *Suppose $\mathcal{C}$ is a linear affine-invariant code of distance $d = 2t+2$ that contains $\mathrm{RM}(n-\log d, n)$ and is locally testable with $\frac{2N}{d}$ queries. Then, $\mathcal{C} \subseteq$ $\mathrm{eBCH}(n, t)$.*

Proof. First, we consider the degree $e^* = 2^0 + 2^1 + 2^2 + \cdots + 2^{n-\log d}$ of 2-weight $n - \log d + 1$. We will show that $e^* \notin \operatorname{Deg}(\mathcal{C})$.

Let S be an arbitrary subspace of dimension $k = n - \log d + 1$. We will show that S cannot be orthogonal to e^*. Let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be an $\mathbb{F}_2$-basis for S. Then,

$$M_{e^*}(\alpha_1, \ldots, \alpha_k) = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_k^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{2^{k-2}} & \alpha_2^{2^{k-2}} & \cdots & \alpha_k^{2^{k-2}} \\ \alpha_1^{2^{k-1}} & \alpha_2^{2^{k-1}} & \cdots & \alpha_k^{2^{k-1}} \end{pmatrix},$$

which has been studied as the (transpose) Moore matrix, whose (i,j) entry is $\alpha_j^{2^{i-1}}$ (see [Moo96]). The determinant of the matrix is known to be

$$\prod_{\lambda_1, \ldots, \lambda_k \in \{0,1\} \text{ not all zero}} (\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_k \alpha_k),$$

i.e., the product of all non-trivial $\mathbb{F}_2$-linear combinations of $\alpha_1, \alpha_2, \ldots, \alpha_k$. Since the α_i are $\mathbb{F}_2$-linearly independent by choice, it follows that the above determinant is nonzero. Thus, Lemma 4.1 implies that S cannot be orthogonal to e^*. Since S was arbitrary, any $\mathcal{C}$ whose degree set contains e^* cannot have dual distance $\frac{2N}{d}$ and would therefore not be locally testable with the desired locality.

Now recall that for $d = 2t + 2$, dual-eBCH(n, t) has degree set

$$\mathrm{Deg}(\text{dual-eBCH}(n,t)) = \overline{\mathrm{shift}}(\{0, 1, 2, \dots, t\}),$$

which implies that $\mathrm{Deg}(\mathrm{eBCH}(n, t)) = \{0, 1, \dots, 2^n - 1\} \setminus T$, where T is the set of all degrees e for which the zeros in the n-bit base-2 representation of e are contained in a consecutive (cyclic) block of size $\log d - 1$. Note that for any $e \in T$, there is some cyclic shift of e^* in its shadow. Since $\mathrm{Deg}(\mathcal{C})$ does not contain e^*, and affine-invariant codes are closed under shifts and shadows, it follows that $\mathrm{Deg}(\mathcal{C}) \cap T = \emptyset$. Hence, $\mathrm{Deg}(\mathcal{C}) \subseteq \mathrm{Deg}(\mathrm{eBCH}(n, t))$, and so, $\mathcal{C} \subseteq \mathrm{eBCH}(n, t)$.

4.3 Dimension Bound via Local Transformations of Degree Now, we show that for any degree e of 2-weight $n - \log d + 1$ that does not pass a fixed subspace S of dimension $n - \log d + 1$, we can perform a slight perturbation to e to obtain another degree e' of 2-weight $n - \log d + 1$ that does not pass S. In other words, for any subspace S, the existence of one degree that does not pass S implies many others.

First, let us state some facts which will be useful for the proof of the main result.

Fact 4.1. *Let $\lambda \in \mathbb{F}_{2^n}$ be nonzero. Then, a subspace S is orthogonal to a degree e if and only if the subspace $\lambda S = \{\lambda s : s \in S\}$ is orthogonal to e.*

Lemma 4.2. *Let $m < n$ and $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{F}_{2^n}$. There exists a nonzero $\lambda \in \mathbb{F}_{2^n}$ such that*

$$\mathrm{Tr}(\lambda\alpha_1) = \cdots = \mathrm{Tr}(\lambda\alpha_m) = 0. \tag{4.3}$$

The proof of Lemma 4.2 appears in Appendix D. Now, we prove one of the main technical theorems.

Theorem 4.2. *Suppose S is a subspace of dimension $k = n - \log d + 1$. Let $e = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_k}$ be a degree of 2-weight k that does not pass S. Then, for any integer $1 \le r \le k$, there exists $u \in \{0, 1, \dots, n-1\} \setminus \{i_1, i_2, \dots, i_k\}$ such that $e' = e - 2^{i_r} + 2^u$ does not pass S.*

Proof. Let $\{j_1, j_2, \dots, j_\ell\} = \{0, 1, \dots, n-1\} \setminus \{i_1, i_2, \dots, i_k\}$. Let $\alpha_1, \alpha_2, \dots, \alpha_k$ be a basis for S. Then, by Lemma 4.2, there exists some nonzero $\lambda \in \mathbb{F}_{2^n}$ such that $\mathrm{Tr}(\lambda\alpha_i) = 0$ for each i. Scaling S by λ, we may assume that $\mathrm{Tr}(\alpha_1) = \mathrm{Tr}(\alpha_2) = \cdots = \mathrm{Tr}(\alpha_k) = 0$.

For ease of notation, we will write $\alpha^{[i]}$ for α^{2^i}. Consider the matrix

$$M = \begin{pmatrix} \alpha_1^{[i_1]} & \alpha_2^{[i_1]} & \cdots & \alpha_k^{[i_1]} \\ \vdots & \vdots & & \vdots \\ \alpha_1^{[i_{r-1}]} & \alpha_2^{[i_{r-1}]} & \cdots & \alpha_k^{[i_{r-1}]} \\ \sum_{t=1}^{\ell} \alpha_1^{[j_t]} & \sum_{t=1}^{\ell} \alpha_2^{[j_t]} & \cdots & \sum_{t=1}^{\ell} \alpha_k^{[j_t]} \\ \alpha_1^{[i_{r+1}]} & \alpha_2^{[i_{r+1}]} & \cdots & \alpha_k^{[i_{r+1}]} \\ \vdots & \vdots & & \vdots \\ \alpha_1^{[i_k]} & \alpha_2^{[i_k]} & \cdots & \alpha_k^{[i_k]} \end{pmatrix}.$$

We observe that $\det M$ is equal to the determinant of the following matrix M' which is obtained by replacing the r^{th} row of M with the sum of all rows of M:

$$M' = \begin{pmatrix} \alpha_1^{[i_1]} & \cdots & \alpha_k^{[i_1]} \\ \vdots & & \vdots \\ \alpha_1^{[i_{r-1}]} & \cdots & \alpha_k^{[i_{r-1}]} \\ \sum_{0 \le t < n, t \ne i_r} \alpha_1^{[t]} & \cdots & \sum_{0 \le t < n, t \ne i_r} \alpha_k^{[t]} \\ \alpha_1^{[i_{r+1}]} & \cdots & \alpha_k^{[i_{r+1}]} \\ \vdots & & \vdots \\ \alpha_1^{[i_k]} & \cdots & \alpha_k^{[i_k]} \end{pmatrix}$$

However, note that $\sum_{0 \le t < n, t \ne i_r} \alpha_s^{[t]} = \alpha_s^{[i_r]} + \mathrm{Tr}(\alpha_s) = \alpha_s^{[i_r]}$ for $s = 1, 2, \dots, k$. Hence, $M' = M_e(\alpha_1, \dots, \alpha_k)$. By Lemma 4.1, since S is not orthogonal to e, we must have $\det(M_e(\alpha_1, \dots, \alpha_k)) \ne 0$. It follows that $\det M \ne 0$. Noting that

$$\det M = \sum_{s=1}^{\ell} \det M_{e_s}(\alpha_1, \dots, \alpha_k),$$

where $e_s = e - 2^{i_r} + 2^{j_s}$. Thus, there exists some s for which $\det M_{e_s}(\alpha_1, \dots, \alpha_k) \ne 0$. Hence, we conclude that the desired statement holds for $u = j_s$.

A corollary of the above theorem is that a code $\mathcal{C}$ with our desired parameters must have dual dimension $\Omega(n^2)$, which already improves on the bound of Corollary 3.1 obtained from [KL11].

Corollary 4.1. *Suppose $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ is a linear affine-invariant code of distance d that contains* $\mathrm{RM}(n - \log d, n)$ *and is locally testable with $\frac{2N}{d}$ queries. Then, $\dim(\mathcal{C}^\perp) = \Omega(n^2)$.*

Proof. Again, write $k = n - \log d + 1$. Recall that $e^* = 2^0 + 2^1 + \cdots + 2^{k-1}$. We know that $e^* \notin \mathrm{Deg}(\mathcal{C})$ from the proof of Theorem 4.1. Therefore, by Theorem 4.2, we see that there exist $j_0, j_1, \dots, j_{k-1} \in \{k, k+1, \dots, n-1\}$, such that $e_i = (e^* - 2^i + 2^{j_i}) \notin \mathrm{Deg}(\mathcal{C})$ for $i = 0, 1, \dots, k-1$. Also, at least a constant fraction of $e_0, \dots, e_{k-1}$ are shift independent (i.e., not shifts of each other); this is because the number of ones in a maximal consecutive cyclic block of ones in the binary representation of e_i can be either 1, i, $i+1$, $k-i-1$, or $k-i$. Thus, at least $n \cdot \Omega(k) = \Omega(n^2)$ degrees of 2-weight k do not lie in $\mathrm{Deg}(\mathcal{C})$, which proves the result.

Theorem 4.2 shows that for a given degree e that does not pass a fixed subspace S, one can shift any 1 in the binary representation of e to some position that is currently occupied by a 0 and obtain another degree that does not pass S. Next, we try to prove an analogue (Theorem 4.3) which allows us to shift any desired 0 to a position occupied by a 1. First, we prove a lemma.

Lemma 4.3. *Suppose $\alpha_1, \alpha_2, \dots, \alpha_k$ are $\mathbb{F}_2$-linearly independent, and let $v_0, \dots, v_{n-1} \in \mathbb{F}_{2^n}^k$ be defined as*

$$v_i = (\alpha_1^{2^i}, \alpha_2^{2^i}, \dots, \alpha_k^{2^i}),$$

where $k = n - \log d + 1$. Then, there is no set of at most $\frac{n}{\log d}$ of the v_i that are linearly dependent.

Proof. Suppose, for the sake of contradiction, that there exists $t \leq \frac{n}{\log d}$ such that $\lambda_1 v_{i_1} + \lambda_2 v_{i_2} + \cdots + \lambda_t v_{i_t} = 0$, where $i_1, i_2, \dots, i_t$ are distinct, and not all of the $\lambda_1, \lambda_2, \dots, \lambda_t \in \mathbb{F}_{2^n}$ are zero. Without loss of generality, suppose $0 \leq i_1 < i_2 < \cdots < i_t \leq n-1$. Let $j_r = (i_{r+1} - i_r) \pmod{n}$, where $i_{t+1} = i_1$. Since $j_1 + j_2 + \cdots + j_t = n$, there exists some r such that $j_r \geq \frac{n}{t} \geq \log d$. Then, note that $v_{i_{r+1}}, v_{i_{r+1}+1}, \cdots, v_{i_{r+1}+(k-1)}$ are linearly independent (where subscripts on v are modulo n): Letting $e^* = 2^0 + 2^1 + \cdots + 2^{k-1}$, we have

$$\begin{aligned} \det \begin{pmatrix} v_{i_{r+1}} \\ v_{i_{r+1}+1} \\ \vdots \\ v_{i_{r+1}+(k-1)} \end{pmatrix} &= \left(\det M_{e^*}(\alpha_1, \dots, \alpha_k)\right)^{2^{i_{r+1}}} \\ &\neq 0, \end{aligned}$$

where the last statement is shown in the proof of Theorem 4.1. However, $v_{i_1}, v_{i_2}, \dots, v_{i_t}$ appear among $v_{i_{r+1}}, v_{i_{r+1}+1}, \dots, v_{i_{r+1}+(k-1)}$. Thus, we obtain a contradiction.

Theorem 4.3. *Suppose S is a subspace of dimension $k = n - \log d + 1$. Let $e = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_k}$ be a degree of 2-weight k that does not pass S. Then, for any integer $0 \leq u \leq n-1$ with $u \notin \{i_1, i_2, \dots, i_k\}$, there exist at least $\frac{n}{\log d} - 1$ values of $r \in [k]$ for which $e + 2^u - 2^{i_r}$ is a degree that does not pass S.*

Proof. Let $u \notin \{i_1, i_2, \dots, i_k\}$, and let $\alpha_1, \alpha_2, \dots, \alpha_k$ be a basis for S. Because e does not pass S, we know that the matrix $M = M_e(\alpha_1, \alpha_2, \dots, \alpha_k)$ has a nonzero determinant. Write $w_t = (\alpha_1^{2^{i_t}}, \alpha_2^{2^{i_t}}, \dots, \alpha_k^{2^{i_t}})$ for $t = 1, 2, \dots, k$, i.e., w_t is the t^{th} row of M. Also, let $v = (\alpha_1^{2^u}, \alpha_2^{2^u}, \dots, \alpha_k^{2^u})$. Since M has nonzero determinant, its row span is all of $\mathbb{F}_{2^n}^k$, and we can find $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}_{2^n}$ such that $v = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k$.

Suppose $\lambda_j \neq 0$. Then, the linear dependence

$$\lambda_j(w_j + \lambda_j^{-1} v) + \sum_{\substack{1 \leq t \leq k \\ t \neq j}} \lambda_t w_t = 0$$

implies that

$$\begin{aligned} 0 &= \det \begin{pmatrix} w_1 \\ \vdots \\ w_{j-1} \\ w_j + \lambda_j^{-1} v \\ w_{j+1} \\ \vdots \\ w_k \end{pmatrix} \\ &= \det M + \lambda_j^{-1} \det M_{e'}(\alpha_1, \alpha_2, \dots, \alpha_k), \end{aligned}$$

where $e' = e + 2^u - 2^{i_j}$. Since $\det M \neq 0$, we have $\det M_{e'}(\alpha_1, \alpha_2, \dots, \alpha_k) \neq 0$, implying that e' does not pass S. To conclude, simply note that Lemma 4.3 implies that there are at least $\frac{n}{\log d} - 1$ values of j for which $\lambda_j \neq 0$. Thus, the desired conclusion follows.

Remark 4.1. *The bounds in Theorems 4.2 and 4.3 are tight, as they are achieved by the (univariate analogue) of the codes of [GKS13]. See E.2 in Appendix E for details.*

Now, we are ready to prove the main theorem, which proves a lower bound on $\dim(\mathcal{C}^\perp)$.

Theorem 4.4. (Main) *Let $\mathcal{C} \supseteq \mathrm{RM}(n - \log d, n)$ be a linear affine-invariant code of block length $N = 2^n$ that has distance d and is testable with $\frac{2N}{d}$ queries. Then, $\dim(\mathcal{C}^\perp) \geq \left(\frac{n}{\log^2 d}\right)^{\log d - 1}$.*

Proof. Fix a subspace S of dimension $n - \log d + 1$ whose indicator is in $\mathcal{C}^\perp$. Let $k = n - \log d + 1$. Recall that $e^* = 2^0 + 2^1 + \cdots + 2^{k-1}$ does not pass S.

Consider the following procedure. Let $e_k = e^*$. Then, for $j = k, k+1, \dots, n-1$ (in succession), we perform either one of the following steps:

- Set $e_{j+1} = e_j$.

- Choose an $i_j \in \{0, 1, \ldots, n-1\}$ such that 2^{i_j} appears in the binary representation of e_j and so that $e_j + 2^j - 2^{i_j}$ does not pass S. Set $e_{j+1} = e_j + 2^j - 2^{i_j}$.

It is clear that at the end of the procedure, e_n will be a degree of 2-weight k that does not pass S. Moreover, for each j in the procedure, there will be at least $\frac{n}{\log d}$ choices for setting e_{j+1} (by Theorem 4.3). On the other hand, any final e_n could have been obtained in at most $(\log d)^{\log d - 1}$ ways. Thus, it follows that there are at least $\left(\frac{n}{\log d}\right)^{\log d - 1} \Big/ (\log d)^{\log d - 1} = \left(\frac{n}{\log^2 d}\right)^{\log d - 1}$ degrees that do not pass S.

5 Conclusion

In this work, we have proven limitations on certain classes of locally testable affine-invariant codes, showing that currently known constructions are essentially optimal. Our work shows that any improved constructions in the regime of linear-query LTCs must use additional ideas. We hope that our techniques will also provide a starting point for future study towards the goal of eventually understanding general LTCs.

References

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

[BFH$^+$13] Arnab Bhattacharyya, Eldar Fischer, Hamed Hatami, Pooya Hatami, and Shachar Lovett. Every locally characterized affine-invariant property is testable. In *Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing*, STOC '13, pages 429–436, 2013.

[BGH$^+$12] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, FOCS, pages 370–379, 2012.

[BGK$^+$10] Eli Ben-Sasson, Venkatesan Guruswami, Tali Kaufman, Madhu Sudan, and Michael Viderman. Locally testable codes require redundant testers. *SIAM Journal on Computing*, 39(7):3230–3247, 2010. Preliminary version appeared in Proc. IEEE CCC 2009.

[BKS$^+$10] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. Optimal testing of Reed-Muller codes. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science*, FOCS, pages 488–497, 2010.

[BSGH$^+$06] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil P. Vadhan. Robust PCPs of proximity, shorter PCPs, and applications to coding. *SIAM J. Comput.*, 36(4):889–974, 2006.

[BSHR05] Eli Ben-Sasson, Prahladh Harsha, and Sofya Raskhodnikova. Some 3CNF properties are hard to test. *SIAM Journal on Computing*, 35(1):1–21, September 2005.

[BSS06] Eli Ben-Sasson and Madhu Sudan. Robust locally testable codes and products of codes. *Random Structures and Algorithms*, 28(4):387–402, 2006.

[BSS08] Eli Ben-Sasson and Madhu Sudan. Short PCPs with polylog query complexity. *SIAM J. Comput.*, 38(2):551–607, 2008.

[BSS11] Eli Ben-Sasson and Madhu Sudan. Limits on the rate of locally testable affine-invariant codes. In Leslie Ann Goldberg, Klaus Jansen, R. Ravi, and José D. P. Rolim, editors, *APPROX-RANDOM*, volume 6845 of *Lecture Notes in Computer Science*, pages 412–423. Springer, 2011.

[BSV12] Eli Ben-Sasson and Michael Viderman. Towards lower bounds on locally testable codes via density arguments. *Computational Complexity*, 21(2):267–309, 2012.

[DG13] Irit Dinur and Venkatesan Guruswami. PCPs via low-degree long code and hardness for constrained hypergraph coloring. In *Proceedings of the 54th Annual Symposium on Foundations of Computer Science*, FOCS, pages 340–349, 2013.

[Din07] Irit Dinur. The PCP theorem by gap amplification. *J. ACM*, 54(3):12, 2007.

[GHH$^+$14] Venkatesan Guruswami, Johan Håstad, Prahladh Harsha, Srikanth Srinivasan, and Girish Varma. Super-polylogarithmic hypergraph coloring hardness via low-degree long codes. In *Proceedings of the 46th annual ACM Symposium on Theory of Computing*, STOC, 2014.

[GKS13] Alan Guo, Swastik Kopparty, and Madhu Sudan. New affine-invariant codes from lifting. In *Proceedings of ITCS 2013*, pages 529–540, 2013.

[Gol11] Oded Goldreich. Short locally testable codes and proofs: a survey in two parts. In *Property testing*, pages 65–104. Springer, 2011.

[GS06] Oded Goldreich and Madhu Sudan. Locally testable codes and PCPs of almost-linear length. *Journal of the ACM*, 53(4):558–655, 2006.

[HRZS13] Elad Haramaty, Noga Ron-Zewi, and Madhu Sudan. Absolutely sound testing of lifted codes. In *Proceedings of APPROX-RANDOM 2013*, pages 671–682, 2013.

[KL11] Tali Kaufman and Shachar Lovett. New extension of the Weil bound for character sums with applications to coding. In *52nd Annual IEEE Symposium on Foundations of Computer Science*, FOCS, pages 788–796, 2011.

[KS07] Tali Kaufman and Madhu Sudan. Algebraic property testing: The role of invariance. *Electronic Colloquium on Computational Complexity (ECCC)*, 14(111), 2007.

[KS14] Subhash Khot and Rishi Saket. Hardness of coloring 2-colorable 12-uniform hypergraphs with $2^{(\log n)^{\Omega(1)}}$ colors. *Electronic Colloquium on Computational Com-*

plexity (ECCC), 21:51, 2014.

[Moo96] Eliakim Hastings Moore. A two-fold generalization of Fermat's theorem. *Bull. Am. Math. Soc.*, 2(7):189–199, 1896. MR:1557441. JFM:27.0139.05.

[MS81] F. J. MacWilliams and Neil J. A. Sloane. *The Theory of Error-Correcting Codes.* Elsevier/North-Holland, Amsterdam, 1981.

[PHE98] Vera S. Pless and W. Cary Huffman (Eds.). *Handbook of Coding Theory (2 Volumes).* Elsevier, 1998.

[RS10] Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *Proceedings of the 42nd ACM Symposium on Theory of Computing*, STOC, pages 755–764, 2010.

[Sud11] Madhu Sudan. Guest column: Testing linear properties: Some general themes. *SIGACT News*, 42(1):59–80, March 2011.

[Tre04] Luca Trevisan. Some applications of coding theory in computational complexity. *Quaderni di Matematica*, pages 347–424, 2004.

[Vid13] Michael Viderman. Strong LTCs with inverse polylog rate and constant soundness. In *54th Annual Symposium on Foundations of Computer Science*, FOCS, pages 330–339, 2013.

A Properties of Affine-Invariant Codes and their Degree Sets

Recall that we are focusing on univariate affine-invariant codes. Here we collect the definitions and results which allow us to study affine-invariant codes by analyzing their degree sets (see, for example, [KS07]).

Definition A.1. *For a function $f : \mathbb{F}_{q^n} \to \mathbb{F}_q$, write it as the unique polynomial $f(x) = \sum_{e=0}^{q^n-1} c_e x^e$ of degree at most $q^n - 1$ which agrees with f on $\mathbb{F}_{q^n}$. Then, the* support *of f, denoted* $\mathrm{Supp}(f)$*, is the set of degrees with non-zero coefficients in f; that is,* $\mathrm{Supp}(f) = \{e : c_e \neq 0\}$.

Definition A.2. *Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a code. We define* $\mathrm{Deg}(\mathcal{F})$*, the* degree set *of $\mathcal{F}$, to be the set* $\mathrm{Deg}(\mathcal{F}) = \bigcup_{f \in \mathcal{F}} \mathrm{Supp}(f)$.

Definition A.3. *Suppose $D \subseteq \{0, 1, \ldots, q^n - 1\}$. We define the $\mathcal{T}(D) \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ to be the* trace code *on D defined by*

$$\mathcal{T}(D) = \left\{ \left(\sum_{e \in D} \mathrm{Tr}(c_e x^e) \right) \in (\mathbb{F}_{q^n} \to \mathbb{F}_q) : c_e \in \mathbb{F}_{q^n} \right\},$$

where $\mathrm{Tr} : \mathbb{F}_{q^n} \to \mathbb{F}_q$ *is the field trace function given by* $\mathrm{Tr}(x) = x + x^q + x^{q^2} + \cdots + x^{q^{n-1}}$.

Let $(\mathrm{mod}^* \ Q)$ refer to the operation that maps non-negative integers into $\{0, 1, \ldots, Q - 1\}$ such that $a \ (\mathrm{mod}^* \ Q) = 0$ if $a = 0$, while if $a \neq 0$, then $a \ (\mathrm{mod}^* \ Q) = b$, where $b \in \{1, 2, \ldots, Q - 1\}$ is the unique integer such that $a \equiv b \pmod{Q - 1}$.

Definition A.4. *For any $e \in \{0, 1, \ldots, q^n - 1\}$, we say that $e' \in \{0, 1, \ldots, q^n - 1\}$ is a q-shift of e if there exists some nonnegative integer i such that $e' = q^i \cdot e \ (mod^* \ q^n)$. Furthermore, we define the* shift closure *of e to be the set of all shifts of e:*

$$\overline{\mathrm{shift}}(e) = \{eq^i \ (mod^* \ q^n) : i \in \{0, 1, \ldots, n - 1\}.$$

The shift closure *of a set $D \subseteq \{0, 1, \ldots, q^n - 1\}$ is then defined to be the union of the shift closures of its elements:*

$$\overline{\mathrm{shift}}(D) = \bigcup_{e \in D} \overline{\mathrm{shift}}(e).$$

Finally, D is said to be shift-closed *if $D = \overline{\mathrm{shift}}(D)$.*

An alternate view of shift-closed sets arises by viewing an element $e \in D$ as a vector in $\{0, 1, \ldots, q - 1\}^n$ given by the base q representation of e. The q-shifts of e are precisely the integers whose corresponding vectors (obtained by taking the base q representation) are cyclic shifts of the vector associated with e. A set D is, therefore, shift-closed if the set is closed under taking "cyclic" shifts of the associated base q representations.

Definition A.5. *Let $e, e' \in \{0, 1, \ldots, q^n - 1\}$. Let $e = \sum_{i=0}^{n-1} e_i q^i$ and $e' = \sum_{i=0}^{n-1} e'_i q^i$ be the base q representations of e and e', respectively. We say that e' lies in the q-*shadow *of e if $e'_i \leq e_i$ for all $0 \leq i \leq n - 1$. We will denote this as $e' \leq_q e$.*

*A set $D \subseteq \{0, 1, \ldots, q^n - 1\}$ is said to be q-*shadow-closed *if*

$$\{e' : e' \leq_q e \textit{ for some } e \in D\} = D.$$

When q is understood, we will say D is shadow-closed.

It is known that known that linear affine-invariant codes can be characterized by their corresponding degree sets.

Theorem A.1. *Let $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ be a linear affine-invariant code. Then, $D = \mathrm{Deg}(\mathcal{F})$ is the unique set $D \subseteq \{0, 1, \ldots, q^n - 1\}$ that is shift-closed and shadow-closed such that $\mathcal{F}$ equals the trace code $\mathcal{T}(D)$. Conversely, if $D \subseteq \{0, 1, \ldots, q^n - 1\}$ is shift-closed and shadow-closed, then $\mathcal{T}(D)$ is a linear affine-invariant code with degree set D.*

Moreover, the dimension of a linear affine-invariant code is given by the size of its degree set.

Theorem A.2. *If $\mathcal{F} \subseteq \{\mathbb{F}_{q^n} \to \mathbb{F}_q\}$ is a linear affine-invariant code, then* $\dim(\mathcal{F}) = |\mathrm{Deg}(\mathcal{F})|$.

B Reed-Muller Containment Assumption

In this work, we have analyzed affine-invariant codes $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ that contain $\mathrm{RM}(n - \log d, n)$. Let us provide some justification for this assumption by showing that any linear affine-invariant code with large dimension must contain a Reed-Muller code of large order.

Theorem B.1. *Suppose $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ is a linear affine-invariant code such that $\mathrm{RM}(s,n) \not\subseteq \mathcal{C}$, for some $s = n - (\log d - 1)\log(n + \log d - 1) + \Omega_d(1)$. Then, $\dim(\mathcal{C}) \leq 2^n - \left(1 + \frac{n}{\log d - 1}\right)^{\log d - 1}$.*

Proof. Suppose $\mathcal{C}$ satisfies the conditions of the hypothesis. Recall that $\mathrm{RM}(s,n)$ is the trace code with degree set consisting of precisely those $0 \leq e < 2^n$ of 2-weight at most s. Thus, there exists some degree of 2-weight at most s that does not appear in $\mathrm{Deg}(\mathcal{C})$. Since the degree set of $\mathcal{C}$ is shadow-closed, it then follows that there exists e of 2-weight *exactly* s that does not appear in $\mathrm{Deg}(\mathcal{C})$. Note that there are n shifts of e (possibly repeated). For any shift e' of e, there are 2^{n-s} degrees that contain e' in their shadow, for a total of $n \cdot 2^{n-s}$. Moreover, any of these degrees may appear up to n times (since each degree contains at most n shifts of e in its shadow). Thus, there are at least $n \cdot 2^{n-s}/n = 2^{n-s}$ distinct degrees that cannot be in $\mathrm{Deg}(\mathcal{C})$. This shows that

$$\dim(\mathcal{C}) \leq 2^n - 2^{n-s} \leq 2^n - \left(1 + \frac{n}{\log d - 1}\right)^{\log d - 1}$$

for $s = n - (\log d - 1)\log(n + \log d - 1) + \Omega_d(1)$.

Therefore, any affine-invariant code that is expected to improve on the testable codes of [GKS13] and [HRZS13] must contain a Reed-Muller code of order $n - O_d(1)\log n$. However, note that the above theorem holds for *any* linear affine-invariant code and does not use testability. It seems that using the testability assumption should yield a tighter bound, which is a promising direction for future work.

C Application of the Extended Weil Bound

Let us show how Theorem 3.1 follows from the extended Weil bound. The main theorem of [KL11], specialized to our setting (where we set $p = 2$, $\chi(x) \equiv (-1)^{\mathrm{Tr}(x)}$ and $g(x) \equiv 0$), can be stated as follows.

Theorem C.1. ([KL11]) *Let $f(x)$ be the sum of $k \geq 1$ monomials, each of 2-weight at most d. Then, either $\mathrm{Tr}(f(x))$ is constant over all $x \in \mathbb{F}_{2^n}$, or*

$$\left|\mathbf{E}_{x \in \mathbb{F}_{2^n}}[(-1)^{\mathrm{Tr}(f(x))}]\right| \leq 2^{-\frac{n}{4d^2 2^d k}}.$$

For the remainder of this section, assume $\mathcal{F}$ is a linear affine-invariant code, and let $D = \mathrm{Deg}(\mathcal{F})$. Let $R = \{1, 3, 5, \ldots, 2^n - 1\}$ be the set of odd degrees, and set $R' = D \cap R$.

Let us bound the maximum possible 2-weight of a degree in D in terms of $|D|$ and $|R'|$.

Lemma C.1. *Let r be the maximum 2-weight of a degree in D. Then, $r \leq \log|D|$.*

Proof. Pick a degree $e \in D$ of 2-weight r. There are exactly 2^r degrees in the shadow of e. Since D is shadow-closed, $2^r \leq |D|$, as desired.

Lemma C.2. *Suppose $|D| > 1$. Let r be the maximum 2-weight of a degree in D. Then, $r \leq \log|R'| + 1$.*

Proof. Observe that we can pick a degree $e \in R'$ of weight $r \geq 1$ (since $|D| > 1$ and D is shift-closed). Note that there are 2^{r-1} odd degrees in the shadow of e. Thus, $2^{r-1} \leq |R'|$, which implies the claim.

Next, we prove an upper bound on $|R'|$ in terms of $|D|$.

Lemma C.3. *Suppose $|D| > 1$. Then, $|R'| \leq \frac{|D|\log^2|D|}{n}$.*

Proof. Note that for any nonzero degree $e \in D$, there are at least $\frac{n}{\mathrm{wt}_2(e)} \geq \frac{n}{\log|D|}$ distinct shifts of e, by Lemma C.1. Moreover, for any nonzero degree $e \in D$, there are at most $\mathrm{wt}_2(e) \leq \log|D|$ shifts of e that lie in R'. Since D contains $|D| - 1$ nonzero degrees, we see that

$$|R'| \leq \frac{|D| - 1}{n/\log|D|} \cdot \log|D| \leq \frac{|D|\log^2|D|}{n},$$

as desired.

Now, we are ready to prove Theorem 3.1. Recall the statement:

Theorem 3.1 (*restated*)**.** Let $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ be a linear affine-invariant code of relative distance $\leq \frac{1}{2} - \delta$, for some $\delta > 0$. Then for any $\epsilon > 0$, $|\mathcal{F}| \geq 2^{\Omega(n^{\frac{3}{2}-\epsilon})}$, i.e., $\dim(\mathcal{F}) > \Omega(n^{\frac{3}{2}-\epsilon})$.

Proof. Suppose the code $\mathcal{F} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ satisfies the hypothesis of Theorem 3.1. Let $D = \mathrm{Deg}(\mathcal{F})$. Since the code has relative distance $\frac{1}{2} - \delta$, $|D| > 1$.

For the sake of contradiction, assume that $\dim(\mathcal{F}) \leq O(n^{\frac{3}{2}-\epsilon})$ for some $\epsilon > 0$. Then,

$$|D| \leq O(n^{\frac{3}{2}-\epsilon}). \tag{C.1}$$

We have that $\mathcal{F} = \mathcal{T}(R' \cup \{0\})$, since each nonzero degree in D has some shift contained in R'. Therefore, any $h(x) \in \mathcal{F}$ can be written as $\mathrm{Tr}(f(x))$ for some $f(x)$ that is a sum of at most $k = |R'| + 1$ monomials. Moreover, by Lemma C.2, we can guarantee that each of these monomials has 2-weight at most $d = \log |R'| + 1$. Then, Theorem C.1 implies that either h is constant or

$$|\mathbf{E}_{x \in \mathbb{F}_{2^n}}[(-1)^{h(x)}]| \leq 2^{-\frac{n}{8(\log|R'|+1)^2 \cdot |R'|(|R'|+1)}}.$$

Assume h is not constant. By Lemma C.3, we have $|R'| \leq \frac{|D| \log^2 |D|}{n}$, and so,

$$\begin{aligned}&|\mathbf{E}_{x \in \mathbb{F}_{2^n}}[(-1)^{h(x)}]| \\ &\leq \exp\left(-\Omega\left(\frac{n^3/|D|^2 \log^4 |D|}{(\log(|D| \log^2 |D|) - \log n + 1)^2}\right)\right).\end{aligned}$$

It is now straightforward to observe that (C.1) implies that

$$|\mathbf{E}_{x \in \mathbb{F}_{2^n}}[(-1)^{h(x)}]| \to 0$$

as $n \to \infty$. However, this implies that the relative Hamming weight of any nonconstant $h(x)$ approaches $\frac{1}{2}$ in the limit $n \to \infty$. Furthermore, any (nonzero) constant $h(x)$ has relative Hamming weight 1. Hence, the relative distance of $\mathcal{F}$ approaches $\frac{1}{2}$ in the limit $n \to \infty$, which contradicts the assumption that the relative distance is $\leq \frac{1}{2} - \delta$. This concludes the proof of Theorem 3.1.

D Omitted Proofs

Proof. [Proof of Fact 2.1] By Theorem 3.3 in [BSHR05], we know that if $\mathcal{C}$ is a (k, ϵ, ρ)-LTC, then $\mathcal{C}$ has a k-query *canonical* tester $\mathcal{T}$ that accepts all $v \in \mathcal{C}$ with probability 1 and rejects any v that is ρ-far from $\mathcal{C}$ with probability at least ϵ. Consider an arbitrary v that is ρ-far from $\mathcal{C}$. There exists some $I \subseteq \{1, 2, \dots, N\}$ in the support of the underlying distribution of $\mathcal{T}$ such that $v|_I \notin \mathcal{C}|_I$. Thus, $\mathcal{C}|_I$ is a linear subspace of $\mathbb{F}^N|_I$ that is strictly contained in $\mathbb{F}^N|_I$. It follows that there exists a nonzero $w' \in \mathbb{F}^N|_I$ that is orthogonal to all of $\mathcal{C}|_I$. Hence, the word $w \in \mathbb{F}^N$ that is supported on I and satisfies $w|_I = w'$ is an element of $\mathcal{C}^\perp$ with Hamming weight at most k.

Proof. [Proof of Lemma 4.1] S is orthogonal to e if and only if $\sum_{\alpha \in S} \alpha^e = 0$. Note that

$$\begin{aligned}\sum_{\alpha \in S} \alpha^e &= \sum_{\lambda_1, \dots, \lambda_k \in \{0,1\}} \prod_{j=1}^{k} (\lambda_1 \alpha_1 + \cdots + \lambda_k \alpha_k)^{2^{i_j}} \\ &= \sum_{\lambda_1, \dots, \lambda_k \in \{0,1\}} \prod_{j=1}^{k} (\lambda_1 \alpha_1^{2^{i_j}} + \cdots + \lambda_k \alpha_k^{2^{i_j}}) \\ &= \sum_{\pi \in S_n} \prod_{j=1}^{k} \alpha_j^{2^{i_{\pi(j)}}},\end{aligned}$$

where the last sum ranges over all permutations of $\{1, 2, \dots, n\}$. The final line follows because any term $\alpha_1^{t_1} \alpha_2^{t_2} \cdots \alpha_k^{t_k}$ that has some t_j of 2-weight at least 2 must also have some $t_j = 0$, hence implying that such a term must occur an even number of times in the sum. Since we are working over fields of characteristic 2, it follows that such a term cannot have a nonzero coefficient. Moreover, the above quantity is equal to the permanent of $M_e(\alpha_1, \alpha_2, \dots, \alpha_k)$, which, over fields of characteristic 2, is equal to $\det M_e(\alpha_1, \alpha_2, \dots, \alpha_k)$. This proves the claim.

Proof. [Proof of Lemma 4.2] Note that $(\mathrm{Tr}(\lambda\alpha_1), \mathrm{Tr}(\lambda\alpha_2), \dots, \mathrm{Tr}(\lambda\alpha_m)) \in \{0,1\}^m$ for all $\lambda \in \mathbb{F}_{2^n} \setminus \{0\}$. Thus, by the pigeonhole principle, there exist two distinct $\lambda_1, \lambda_2 \in \mathbb{F}_{2^n} \setminus \{0\}$ for which

$$\begin{aligned}&(\mathrm{Tr}(\lambda_1\alpha_1), \dots, \mathrm{Tr}(\lambda_1\alpha_m)) \\ &\qquad = (\mathrm{Tr}(\lambda_2\alpha_1), \dots, \mathrm{Tr}(\lambda_2\alpha_m)).\end{aligned}$$

Thus, by linearity of trace, we see that (4.3) holds for $\lambda = \lambda_1 - \lambda_2$.

E Univariate Constructions of Codes

Recall that [GKS13] gives a linear affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_{2^\ell}^m \to \mathbb{F}_2\}$ with block length $N = 2^n$, where $n = \ell m$. For $\ell = \log d - 1$, the code has distance d and is testable with $2N/d$ queries. Moreover, $\mathcal{C}$ contains the multivariate Reed-Muller code $\mathrm{RM}(n - \log d, n)$.

The above code is obtained by "lifting" a parity check code of smaller block length and happens to be *multivariate*. In our work, we are concerned with dimension bounds on *univariate* codes. As it turns out, the code of [GKS13] has a univariate analogue, i.e., a subset of $\{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$. We provide a construction of this univariate code which does not involve lifting.

E.1 Subspaces from Subfields Let us try to construct a univariate linear affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$ that has distance $d > 1$ and is testable with $2N/d$ queries, where $N = 2^n$. Again, we consider

$\mathcal{C} \supseteq \mathrm{RM}(n - \log d, n)$. From Fact 2.1 and affine invariance of the $\mathcal{C}^\perp$, we know that in order for $\mathcal{C}$ to be testable with the desired locality, there must be a codeword in $w \in \mathcal{C}^\perp$ of Hamming weight at most $2N/d$ such that $w_0 \neq 0$ (i.e., the coordinate of w corresponding to the zero element of $\mathbb{F}_{2^n}$ is nonzero). It is known that $\mathrm{RM}(n - \log d, n)$ has dual distance $2N/d$, and the dual codewords of minimum weight are precisely the affine subspaces of dimension $n - \log d + 1$. Hence, w must be (the indicator of) a *linear* subspace S of the aforementioned dimension.

Hence, we will consider a fixed subspace S of dimension $n - \log d + 1$ and consider which degrees we can take in $\mathrm{Deg}(\mathcal{C})$. We will say that a degree e *passes* the subspace S if

$$\sum_{a \in S} a^e = 0.$$

The above condition is necessary for us to be able to take e in $\mathrm{Deg}(\mathcal{C})$.

Let us again write $\ell = \log d - 1$. Assume $\ell \mid n$, so that $\mathbb{F}_{2^\ell}$ is a subfield of $\mathbb{F}_{2^n}$. Write $n = \ell m$. We can then consider subspaces S of the form

$$S = \lambda_1 \mathbb{F}_{2^\ell} + \lambda_2 \mathbb{F}_{2^\ell} + \cdots + \lambda_{m-1} \mathbb{F}_{2^\ell}, \tag{E.2}$$

where λA is used to mean $\{\lambda a : a \in A\}$, and $\lambda_1, \lambda_2, \ldots, \lambda_{m-1}$ are $\mathbb{F}_{2^\ell}$-linearly independent.

Now, a degree $e = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_u}$ passes S if and only if

$$\begin{aligned}
0 &= \sum_{a \in S} a^e \\
&= \sum_{a \in S} \prod_{j=1}^{u} a^{2^{i_j}} \\
&= \sum_{a_1, \ldots, a_{m-1} \in \mathbb{F}_{2^\ell}} \prod_{j=1}^{u} \left(\sum_{k=1}^{m-1} \lambda_k a_k \right)^{2^{i_j}} \\
&= \sum_{a_1, \ldots, a_{m-1} \in \mathbb{F}_{2^\ell}} \prod_{j=1}^{u} \sum_{k=1}^{m-1} (\lambda_k a_k)^{2^{i_j}} \\
&= \sum_{a_1, \ldots, a_{m-1} \in \mathbb{F}_{2^\ell}} \sum_{e_1, \ldots, e_{m-1}} \prod_{k=1}^{m-1} (\lambda_k a_k)^{e_k} \\
&= \sum_{e_1, \ldots, e_{m-1}} \prod_{k=1}^{m-1} \lambda_k^{e_k} \left(\sum_{a \in \mathbb{F}_{2^\ell}} a^{e_k} \right), \qquad \text{(E.3)}
\end{aligned}$$

where in the last two equations, $e_1, \ldots, e_{m-1}$ range over all $e_1, \ldots, e_{m-1}$ with distinct supports in their binary expansion, such that $e_1 + \cdots + e_{m-1} = e$. Observe that $\sum_{a \in \mathbb{F}_{2^\ell}} a^{e_j} \neq 0$ if and only if e_j is a positive integral multiple of $2^\ell - 1$. Hence, the above condition would be guaranteed for e if there happens to be no way to write e as a sum $e = e_1 + e_2 + \cdots + e_{m-1}$ such that **(1.)** $e_1, \ldots, e_{m-1}$ have distinct supports in their binary expansion, and **(2.)** $e_1, e_2, \ldots, e_{m-1}$ are all positive multiples of $2^\ell - 1$.

Now, it will be convenient to reason about degrees in terms of a matrix form.

Definition E.1. *Let $0 \le e < 2^n$. Moreover, let $e = b_0 2^0 + b_1 2^1 + \cdots + b_{n-1} 2^{n-1}$ be the binary representation of e (where $b_0, b_1, \ldots, b_{n-1} \in \{0, 1\}$). Then, define the* block matrix representation *of e to be the following $m \times \ell$ matrix:*

$$\begin{pmatrix}
b_{n-\ell} & b_{n-\ell+1} & \cdots & b_{n-1} \\
\vdots & \vdots & & \vdots \\
b_\ell & b_{\ell+1} & \cdots & b_{2\ell-1} \\
b_0 & b_1 & \cdots & b_{\ell-1}
\end{pmatrix}.$$

*Furthermore, for $j = 0, 1, \ldots, \ell - 1$, we define the j-*shifted row projection *of e, denoted $\mathrm{proj}_j(e)$, as*

$$\mathrm{proj}_j(e) = \sum_{i=0}^{n-1} b_i 2^{((i+j)\ mod\, \ell)}.$$

In other words, $\mathrm{proj}_j(e)$ is obtained by taking the block matrix representation of e, cyclically shifting its columns by j to the right, and then taking the inner product of $(2^0, 2^1, \ldots, 2^{\ell-1})$ with the row sum of the resulting matrix.

Note the following easy property about row projections.

Lemma E.1. *For any $j = 0, 1, \ldots, \ell - 1$, we have that $\mathrm{proj}_j(e) \equiv 2^j e \pmod{2^\ell - 1}$. In particular, $\mathrm{proj}_j(e) \equiv 0 \pmod{2^\ell - 1}$ if and only if $e \equiv 0 \pmod{2^\ell - 1}$.*

Proof. As usual, let $e = b_0 2^0 + \cdots + b_{n-1} 2^{n-1}$ be the binary representation of e. Note that

$$\begin{aligned}
\mathrm{proj}_j(e) &= \sum_{i=0}^{n-1} b_i 2^{((i+j) \bmod \ell)} \\
&\equiv \sum_{i=0}^{n-1} b_i 2^{i+j} && \pmod{2^\ell - 1} \\
&\equiv 2^j \sum_{i=0}^{n-1} b_i 2^i && \pmod{2^\ell - 1} \\
&\equiv 2^j e && \pmod{2^\ell - 1},
\end{aligned}$$

which proves the first part of the claim. The second part of the claim is a simple consequence of the first part.

THEOREM E.1. *Suppose e is a degree whose block matrix representation has at least two zeros in some column. Then, e passes any $(n - \log d + 1)$-dimensional subspace S of the form (E.2).*

Proof. Recall (E.3). Suppose e satisfies the hypothesis of the claim. As noted before, it suffices to show that there is no way to write e as a sum $e = e_1 + e_2 + \cdots + e_{m-1}$ such that **(1.)** $e_1, \ldots, e_{m-1}$ have distinct supports in their binary expansion, and **(2.)** $e_1, e_2, \ldots, e_{m-1}$ are all positive multiples of $2^\ell - 1$.

For the sake of contradiction, assume that there does exist a decomposition $e = e_1 + e_2 + \cdots + e_{m-1}$ satisfying **(1.)** and **(2.)**. Also, suppose the j^{th} column of the block matrix representation of e contains at least two zeros. Then, by Lemma E.1, we have that for $i = 1, 2, \ldots, m-1$,

$$\text{proj}_{\ell-j}(e_i) \equiv 2^{\ell-j} e_i \equiv 0 \pmod{2^\ell - 1}.$$

Moreover, since e_i is positive, we must have that $\text{proj}_{\ell-j}(e_i) > 0$. Thus, $\text{proj}_{\ell-j}(e_i) \geq 2^\ell - 1$. It follows that

$$\begin{aligned} \text{proj}_{\ell-j}(e) &= \sum_{i=1}^{m-1} \text{proj}_{\ell-j}(e_i) \\ &\geq (m-1)(2^\ell - 1). \end{aligned} \tag{E.4}$$

On the other hand, since there are at least two zeros in the j^{th} column of the block matrix representation of e, we have

$$\begin{aligned} \text{proj}_{\ell-j}(e) &\leq m(2^0 + 2^1 + \cdots + 2^{\ell-1}) - 2 \cdot 2^{\ell-1} \\ &= (m-1)(2^\ell - 1) - 1, \end{aligned}$$

which contradicts (E.4). Hence, **(1.)** and **(2.)** cannot be satisfied, and desired result follows.

Thus, let us define $D \subseteq \{0, 1, \ldots, 2^n - 1\}$ as the set of all $0 \leq e \leq 2^n - 1$ such that the block matrix representation of e contains at least two zeros in some column. Then, it is easy to see that D is shift-closed and shadow-closed. Thus, $\mathcal{T}(D) \subseteq \{\mathbb{F}_{2^n} \to \mathbb{F}_2\}$. Moreover, for none of the degrees in D can the zeros in the n-bit binary representation lie in a cyclic block of length $\log d - 1$ (this is guaranteed by the condition that there are two zeros in some column of the block matrix representation). Thus, $\mathcal{T}(D) \subseteq \text{eBCH}(n, (d-2)/2)$. Combining this with $\text{RM}(n - \log d, n) \subseteq \mathcal{T}(D)$ shows that $\mathcal{T}(D)$ has distance d. Moreover, by Theorem E.1, all $e \in D$ simultaneously pass a common subspace S of dimension $n - \log d + 1$, which means that the distance of the dual code is $2N/d$.

Finally, recall from Theorem A.2 that $\dim(\mathcal{T}(D)) = |\text{Deg}(D)|$. The degrees that are *not* in D are precisely those that have at most one zero in each column of their block matrix representation. Hence, a simple counting argument shows that

$$\dim(\mathcal{T}(D)) = N - \left(1 + \frac{\log N}{\log d - 1}\right)^{(\log d - 1)}.$$

REMARK E.1. *The above code $\mathcal{T}(D)$ turns out to be the univariate analogue of the multivariate linear locality LTC presented in [GKS13]. The criterion for the degree set in the multivariate code is virtually the same "two zeros in some column" criterion here, except that the degrees for the multivariate code are m-tuples, and each component of the m-tuple corresponds to a row (viewed as a binary representation) of our block matrix representation. Testability of our univariate analogue follows from [GKS13], with the use of an isomorphism between $\mathbb{F}_{2^n}$ and $\mathbb{F}_{2^\ell}^m$.*

REMARK E.2. *The linear locality code of [HRZS13] is a code $\mathcal{C} \subseteq \{\mathbb{F}_{2^t}^{n/t} \to \mathbb{F}_2\}$ for general $t \mid n$. It is a generalization of the code in [GKS13] (the latter follows by setting $t = \ell$ for n that are multiples of ℓ). The procedure of this section can be applied in a similar fashion to obtain univariate analogues of the codes of [HRZS13], except that one uses subspaces constructed using the subfield $\mathbb{F}_{2^t}$ instead of $\mathbb{F}_{2^\ell}$, and the block matrix representation will have to be defined as an $(n/t) \times t$ matrix. We omit the details, since the technique is similar enough, and the specific construction of [GKS13] is the one that matches the lower bound on co-dimension given by Theorem 4.4.*

E.2 Optimality Results Now, we show that the technical results of Theorems 4.2 and 4.3 are tight by showing that the univariate construction of the previous section matches those bounds.

Again, take $\ell = \log d - 1$ and $n = \ell m$, and let D be as defined in Section E.1. Moreover, choose S to be a subspace whose indicator lies in the dual of $\mathcal{T}(D)$. Let $e^* = 2^0 + 2^1 + \cdots + 2^{n-\ell-1}$.

LEMMA E.2. *For any $0 \leq r \leq n - \ell - 1$, there exists at most one value of $u \in \{n - \ell, n - \ell + 1, \ldots, n-1\}$ such that $e' = e^* - 2^r + 2^u$ does not pass S.*

Proof. Let $s = r \bmod \ell$. Note that for any $u \in \{n - \ell, \ldots, n-1\}$ such that $u \neq n - \ell + s$, the block matrix representation of $e' = e^* - 2^r + 2^u$ contains two zeros in some column, and thus, e' would pass S. This implies that the only admissible value of u for which $e' = e^* - 2^r + 2^u$ does not pass S is $u = n - \ell + s$, as desired.

From the proof of Theorem 4.1, we know that e^* does not pass S. Therefore, the result of Theorem 4.2 shows that there must exist *at least* one value of $e' = e - 2^r + 2^u$ that does not pass S. Thus, Lemma E.2 matches this lower bound.

Next, we note the following lemma.

LEMMA E.3. *For any $n - \ell \leq u \leq n - 1$, there exist at most $m - 1 = \frac{n}{\log d - 1} - 1$ values of $r < n - \ell$ such that $e' = e^* + 2^u - 2^r$ does not pass S.*

Proof. Let $s = u \bmod \ell$. Then, note that for any $r < n - \ell$ such that $r \bmod \ell \neq s$, the block matrix representation of $e' = e^* + 2^u - 2^r$ contains two zeros in some column, and hence, e' would pass S. Thus, the only possible values of $r < n - \ell$ for which $e' = e^* + 2^u - 2^r$ may not pass S are those for which $r \bmod \ell = s$. There are precisely $m - 1$ such values of r, which proves the desired claim.

Since the result of Theorem 4.3 shows that there must exist *at least* $\frac{n}{\log d - 1} - 1$ values of $e' = e + 2^u - 2^r$ that do not pass S, we see that Lemma E.3 matches this lower bound.

REMARK E.3. *Straightforward generalizations of the above lemmas (generalized to any e^* whose block matrix representation has exactly one zero in each column) actually show that $\mathcal{T}(D)$ is* maximal *among affine-invariant codes with the desired properties. In particular, any degree set D' strictly containing D that is both shift-closed and shadow-closed would have to contain some degree whose block matrix representation contains exactly one zero in each column, which is impossible.*