New Error Correcting Codes from Lifting

by

Alan Xinyu Guo

B.S., Duke University (2011)
S.M., Massachusetts Institute of Technology (2013)

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

Doctor of Philosophy in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

Author .....
Department of Electrical Engineering and Computer Science
May 1, 2015

Certified by .....
Madhu Sudan
Adjunct Professor
Thesis Supervisor

Accepted by .....
Leslie A. Kolodziejski
Chair, Department Committee on Graduate Students

New Error Correcting Codes from Lifting

by

Alan Xinyu Guo

Submitted to the Department of Electrical Engineering and Computer Science
on May 1, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

Error correcting codes have been widely used for protecting information from noise. The theory of error correcting codes studies the range of parameters achievable by such codes, as well as the efficiency with which one can encode and decode them. In recent years, attention has focused on the study of sublinear-time algorithms for various classical problems, such as decoding and membership verification. This attention was driven in part by theoretical developments in probabilistically checkable proofs (PCPs) and hardness of approximation. Locally testable codes (codes for which membership can be verified using a sublinear number of queries) form the combinatorial core of PCP constructions and thus play a central role in computational complexity theory. Historically, low-degree polynomials (the Reed-Muller code) have been the locally testable code of choice. Recently, "affine-invariant" codes have come under focus as providing potential for new and improved codes.

In this thesis, we exploit a natural algebraic operation known as "lifting" to construct new affine-invariant codes from shorter base codes. These lifted codes generically possess desirable combinatorial and algorithmic properties. The lifting operation preserves the distance of the base code. Moreover, lifted codes are naturally locally decodable and testable. We tap deeper into the potential of lifted codes by constructing the "lifted Reed-Solomon code", a supercode of the Reed-Muller code with the same error-correcting capabilities yet vastly greater rate.

The lifted Reed-Solomon code is the first high-rate code known to be locally decodable up to half the minimum distance, locally list-decodable up to the Johnson bound, and robustly testable, with robustness that depends only on the distance of the code. In particular, it is the first high-rate code known to be both locally decodable and locally testable. We also apply the lifted Reed-Solomon code to obtain new bounds on the size of Nikodym sets, and also to show that the Reed-Muller code is robustly testable for all field sizes and degrees up to the field size, with robustness that depends only on the distance of the code.

Thesis Supervisor: Madhu Sudan
Title: Adjunct Professor

Acknowledgments

First and foremost, I thank my advisor, Madhu Sudan, for his patient guidance, support, and encouragement throughout my four years at MIT. Our shared taste in finding clean, general solutions to algebraic problems has led to many collaborations. I thank Scott Aaronson and Dana Moshkovitz for serving on my thesis committee.

I am also indebted to Ezra Miller, who guided my undergraduate research at Duke, and to Vic Reiner and Dennis Stanton who mentored me at their REU in Minnesota.

I thank my co-authors during graduate school: Greg Aloupis, Andrea Campagna, Erik Demaine, Elad Haramaty, Swastik Kopparty, Ronitt Rubinfeld, Madhu Sudan, and Giovanni Viglietta. I especially thank Ronitt and Piotr Indyk for their guidance early on, Erik for his excitement and willingness to entertain my more fun research ideas (i.e. hardness of video games), and Swastik for hosting my visit to Rutgers and teaching me about codes and PCPs.

I also thank friends and faculty with whom I have shared conversations at MIT: Pablo Azar, Mohammad Bavarian, Eric Blais, Adam Bouland, Mahdi Cheraghchi, Henry Cohn, Matt Coudron, Michael Forbes, Badih Ghazi, Pritish Kamath, Ameya Velingker, Henry Yuen, and so many others.

I thank my friends outside of the field for their friendship and the good times during the past four years. I especially thank Vivek Bhattacharya, Sarah Freitas, Henry Hwang, Steven Lin, Ann Liu, Lakshya Madhok, and Roger Que.

I am grateful to my family. Without my parents, I would not exist, and neither would this thesis. Moreover, they always supported my education and encouraged me to pursue my dreams. I thank my younger sister Julia for encouraging me to set a good example.

Finally, my biggest thanks goes to Lisa — my fiancée, best friend, and companion for the rest of my life. Graduate school was not always easy, but you were always there to support me and pick me up when I was down, and to share in my triumphs. Thank you for your unwavering love and loyalty, without which I do not believe I would have survived through graduate school. Only you know every twist and turn my journey has taken. Without hesitation, I dedicate this thesis to you.

Contents

Chapter 1

Introduction

1.1 Background

1.1.1 Error Correcting Codes

Error correcting codes arise as a solution to the problem of communicating over a noisy channel. The sender first encodes the message using an error correcting code, which adds redundancy to the message, into a codeword. The codeword is then sent over the noisy channel. The receiver receives a word which is a corruption of the codeword. The receiver then decodes the received word and hopefully retrieves the original message. The actual *code* is the set of possible codewords.

Two classical parameters of interest are the *rate* and *distance* of the code. The rate is the ratio of the message length to the codeword length, and measures the efficiency of the encoding. The distance measures the error-correcting capability of the code. The Hamming distance between two strings is the number of symbols in which they differ. The distance of a code is the minimum Hamming distance between two distinct codewords. If the distance of a code is d, then in principle one can detect up to $d-1$ errors and correct up to $\lfloor d/2 \rfloor - 1$ errors: if the codeword has been corrupted in at most $d-1$ locations, then it cannot have been corrupted into a different codeword, so to detect errors one "merely" checks if the received word is a codeword; if the codeword has been corrupted in at most $\lfloor d/2 \rfloor - 1$ locations, then

there is at most one codeword within Hamming distance $\lfloor d/2 \rfloor - 1$ of the received word, so to correct errors one "merely" finds the nearest codeword to the received word. We often prefer to work with the *relative distance* of a code, which is simply its distance divided by the length of the codewords. There is a fundamental tradeoff between rate and distance, which is still not fully understood. In addition, there is the problem of designing codes which support efficient algorithms for encoding and decoding.

Another notion of decoding is *list-decoding*. We just showed that if a code $\mathcal{C}$ has distance d, then for any word, there is at most one codeword within Hamming distance $\lfloor d/2 \rfloor - 1$. If we wish to correct more than $\lfloor d/2 \rfloor - 1$ errors, we cannot guarantee that there is a unique codeword. However, if the radius is not too large, then we can hope that there are not too many codewords within the radius from the received word. Instead of outputting the correct codeword or message, a list-decoding would output a list of potential codewords or messages.

Another important feature of an error correcting code is the *alphabet*. Designing a code is easier using a larger alphabet. If the alphabet size is allowed to grow with the code length n, then the Singleton bound asserts that the rate R and relative distance δ satisfies $R + \delta \leq 1 + 1/n$. This bound is tight, as demonstrated by the Reed-Solomon code.

The *Reed-Solomon code* is perhaps the most ubiquitous code in the literature. The idea is simple. Let $k \geq 1$ and let $q \geq k$ be a prime power. Let $\mathbb{F}_q$ be the finite field of size q. Each message $m = (m_0, \ldots, m_{k-1})$ of length k over the alphabet $\mathbb{F}_q$ is interpreted as a degree $k-1$ polynomial $m(X) = m_0 + m_1 X + \cdots + m_{k-1} X^{k-1}$. Let $\alpha_1, \ldots, \alpha_q$ be the elements of $\mathbb{F}_q$. The encoding of m is simply the evaluation of m at every point: $(m(\alpha_1), \ldots, m(\alpha_q))$. The rate of this code is clearly $R = \frac{k}{q}$, and it follows from the Fundamental Theorem of Algebra that $\delta = 1 - \frac{k-1}{q}$, so that $R + \delta = 1 + 1/q$, meeting the Singleton bound. Furthermore, the Reed-Solomon code can be efficiently decoded up to half its distance using, for instance, the Welch-Berlekamp algorithm [WB86] (see [GS92] for an exposition). Guruswami and Sudan [GS99] showed that the Reed-Solomon code with distance $\delta > 0$ can be list-decoded up to the "Johnson bound" ($1 - \sqrt{1-\delta}$ fraction errors). More precisely, they gave an efficient algorithm which, on input a received word, outputs a list of $O(1/\epsilon^2)$ codewords that

are within (relative) distance $1 - (1+\epsilon)\sqrt{1-\delta}$ of the received word.

A related code is the *Reed-Muller code*. This code, parameterized by a degree d and a number m of variables, consists of polynomials $f : \mathbb{F}_q^m \to \mathbb{F}_q$ of degree at most d (or rather, their evaluations). Its message length is $k = \binom{d+m}{m}$ and its code length is $n = q^m$. When $m = 1$, this is simply the Reed-Solomon code. If $d < q$, it follows from the Schwartz-Zippel lemma that the distance of the Reed-Muller code is $1 - \frac{d}{q}$. Although the Reed-Muller code's rate-distance tradeoff is worse than that of the Reed-Solomon code, the Reed-Muller code offers *locality*, which we discuss in Section 1.1.2. Like the Reed-Solomon code, the Reed-Muller code can be list-decoded up to the Johnson bound [PW04].

1.1.2 PCPs and Local Algorithms

In the late 1980s and early 1990s, there was an explosion of interest in sublinear-time algorithms. Blum, Luby, and Rubinfeld [BLR93] showed that one can probabilistically test whether a given function $f : \{0,1\}^n \to \{0,1\}$ is linear — that is, $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ for every $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$ — by querying f at only 3 points. If f is linear, then their test accepts with probability one, while if f is "ϵ-far" from linear (it disagrees with every linear function in at least ϵ-fraction of the domain), then their test rejects with probability $\Omega(\epsilon)$. The space of linear functions $f : \{0,1\}^n \to \{0,1\}$ forms an error correcting code with rate $n/2^n$ and distance $1/2$ — this code is known as the *Hadamard code*. This was in fact the first code shown to be *locally testable* — that is, using a sublinear number of queries to the received word, one can verify membership in the code. It is also easy to show that the Hadamard code is *locally correctable* — one can correct any given symbol of the received word with high probability using a sublinear number of queries. To correct f at $\mathbf{x} \in \{0,1\}^n$, select random $\mathbf{y} \in \{0,1\}^n$ and output the value $f(\mathbf{x}+\mathbf{y}) - f(\mathbf{y})$. A related notion if *local decodability* — one can correct any given symbol of the *message* with high probability by making a sublinear number of queries to the received word.

We will often consider codes $\mathcal{C} \subseteq \mathbb{F}_q^n$ that are *linear* (i.e. $\mathcal{C}$ forms a vector space over $\mathbb{F}_q$). The Reed-Solomon code, Reed-Muller code, and Hadamard code are all linear codes.

Rather than thinking of words in $\mathbb{F}_q^n$ as sequences of length n, we view them as functions from some fixed set S of cardinality $|S| = n$ to the range $\mathbb{F}_q$. The structure of the set S and symmetries will play a role later. We use $\{S \to \mathbb{F}_q\}$ to denote the set of all such functions. We say a function f is τ-far from $\mathcal{C}$ if $\delta(f, \mathcal{C}) \triangleq \min_{g \in \mathcal{C}} \delta(f, g) \geq \tau$.

Given a code $\mathcal{C} \subseteq \{S \to \mathbb{F}_q\}$ and integer ℓ, an *ℓ-local tester* $\mathcal{T}$ is a distribution $\mathcal{D}$ on $(S^\ell, 2^{\mathbb{F}_q^\ell})$ with the semantics as follows: given oracle access to $f : S \to \mathbb{F}_q$, the tester $\mathcal{T}$ samples $(\pi, V) \leftarrow \mathcal{D}$, where $\pi = (\pi_1, \ldots, \pi_\ell) \in S^\ell$ and $V \subseteq \mathbb{F}_q^\ell$, and accepts f if and only if $f|_\pi \triangleq (f(\pi_1), \ldots, f(\pi_\ell)) \in V$. The tester is *$\epsilon$-sound* if $\mathcal{T}$ accepts $f \in \mathcal{C}$ with probability one, while rejecting f that is δ-far from $\mathcal{C}$ with probability at least $\epsilon \cdot \delta$.

We will also be interested in a stronger property of testers known as their robustness, formally defined by Ben-Sasson and Sudan [BSS06], based on analogous notions in complexity theory due to Ben-Sasson et al. [BSGH+04] and Dinur and Reingold [DR04]. The hope with a robust tester is that, while it may make a few more queries than the minimum possible, the rejection is "more emphatic" in that functions that are far from $\mathcal{C}$ typically yield views that are far from acceptable, i.e. if $\delta(f, \mathcal{C})$ is large, then so is $\delta(f|_\pi, V)$ for typicaly choices of $(\pi, V) \leftarrow \mathcal{D}$. Formally, a tester $\mathcal{D}$ is *α-robust* if $\mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}}\left[\delta(f|_\pi, V)\right] \geq \alpha \cdot \delta(f, \mathcal{C})$. Robustness can be a much stronger property than mere soundnesss since it allows for composition with other local testers. In particular, if there is an α-robust tester for f with distribution $\mathcal{D}$ and if for every (π, V) in the support of $\mathcal{D}$, the property of being in V has an ℓ'-local tester that is ϵ-sound, then $\mathcal{C}$ has an ℓ'-local tester that is $\alpha \cdot \epsilon$-sound. The hope that membership in V has a nice local tester for *every* V in the support of $\mathcal{D}$ may seem overly optimistic, but for many symmetric codes (such as affine-invariant codes, to be discussed later), all the V's are isomorphic — so this is really just one hope.

The interest in sublinear-time algorithms is obvious from a practical perspective. As the amount of data stored and transmitted by society continues its explosive growth, even linear-time algorithms may be too slow, and also unnecessary. Indeed, statisticians understood this decades ago when estimating the population averages. If we want an approximate answer, often $O(1)$ queries suffice to give a good approximation with high confidence. Moreover, in

the context of error correction, if we have a very large encoded file and only wish to decode a small portion of it, then we need our code to be locally decodable.

Surprisingly, sublinear-time algorithms for algebraic codes have played a prominent role in computational complexity theory. In particular, locally testable codes form the "combinatorial core" of probabilistically checkable proofs (PCPs). A PCP for a language L is a protocol involving two parties, a Prover and a Verifier, and an input x, whereby the Prover supplies a proof, depending on x, in an attempt to convince the Verifier that $x \in L$. The Verifier makes a small number of random queries to the proof and then either accepts or rejects the proof based on what it sees. A valid PCP for L satisfies the following: if $x \in L$, then there is some proof that the Prover can provide such that V accepts with probability 1; if $x \notin L$, then regardless of the proof provided by the Prover, the Verifier will reject with high probability (over its random queries). The celebrated PCP Theorem, proved in [AS98, ALM+98], characterizes the complexity class **NP** as the class of languages with PCPs where the Verifier makes only $O(1)$ queries to the proof and uses only $O(\log n)$ bits of randomness in determining its random queries, where n is the length of the input x. Not only did this theorem elucidate the class of **NP** a bit more, but it also paved the way for proving that, for many combinatorial optimization problems, even approximating the solution is **NP**-hard.

At the heart of the proof of the PCP theorem lies the problem of low-degree testing — the problem of testing whether a given function $f : \mathbb{F}_q^m \to \mathbb{F}_q$ has total degree $\deg(f) \leq d$ or is far from any such function. Low-degree testing has been the most extensively studied algebraic property testing problem. First studied in the work of Rubinfeld and Sudan [RS96], low-degree testing and many variations have been analyzed in many subsequent works — a partial list includes [ALM+98, FS95, AS03, RS97, MR06, AKK+05, KR06, JPRZ09, BKS+10, HSS11]. When $d \ll q$, low-degree tests making as few as $d+2$ queries are known, that have $1/\operatorname{poly}(d)$-soundness (see, for instance, Friedl-Sudan [FS95]). However, tests that make $O(d)$ queries achieve constant soundness (a universal constant independent of m, d, q provided q is sufficiently larger than d), and even constant robustness. This constant robustness is central

to the PCP construction of Arora et al. [ALM+98]. In all cases with $d \ll q$, low-degree tests operate by considering the restriction of a function to a random line, or sometimes plane, in the domain, and accepting a function if its restriction to the chosen subspace is a polynomial of degree at most d. Thus, the different restrictions π are different affine subspaces of low dimension (one or two) and the acceptable pattern V is the same for all π. In particular, the robust analysis of the low-degree test allows for low-query tests, or even proofs, of membership in V in constant dimensional spaces to be composed with the low-degree test in high dimensions to yield low-query PCPs. Robustness turns out to be much more significant as a parameter to analyze in these results than the query complexity of the outer test. Indeed, subsequent strengthenings of the PCP theorem in various senses (e.g. in [AS03, RS97, MR06]) rely on improving the robustness to a quantity close to 1, and this leads to PCPs of arbitrarily small constant, and then even $o(1)$, error.

1.1.3 Affine-Invariance

In an attempt to understand what exactly makes codes like the Reed-Muller code testable, Kaufman and Sudan set out to systematically study *affine-invariant codes* [KS08]. By then, it was known that symmetry in properties contributed to their testability, as in graph property testing. The hope was to find the analogous notion of symmetry for algebraic properties such as linearity or low-degree, and affine-invariance seemed to be a promising abstraction. A code $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ is *affine-invariant* if, for every codeword $f \in \mathcal{C}$, the codeword $f \circ A$ obtained by first applying an affine permutation $A : \mathbb{F}_q^m \to \mathbb{F}_q^m$ to the domain before evaluating f, is also a codeword. Kaufman and Sudan showed in [KS08] that any linear affine-invariant code that is "k-single-orbit characterized" is testable with k queries and soundness $\Omega(k^{-2})$. For example, the Hadamard code is 3-single-orbit characterized by $f(\mathbf{x})+f(\mathbf{y})-f(\mathbf{x}+\mathbf{y}) = 0$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^m$ (the characterization is by constraints on 3 points, and the constraints are all in a single orbit by the group action of affine permutations). Their proof simplifies, unifies, and generalizes the proofs found in [BLR93, RS96, AKK+05, KR06, JPRZ09] and unearths some of the fundamental underlying reasons for why testing works. Additionally,

Kaufman and Sudan initiated the systematic study of affine-invariant properties, and laid the groundwork for our structural understanding of affine-invariant properties. The hope was that eventually the study of affine-invariance would lead to constructions of new affine-invariant codes with desirable testability properties.

While the locality properties (testability and correctability) of Reed-Muller codes are well-studied, they are essentially the only rich class of symmetric codes that are well-studied. The only other basic class of symmetric codes that are studied seem to be sparse ones (codes with few codewords).

1.2 This Thesis

In this thesis, we use an algebraic operation, known as "lifting", to construct new affine-invariant codes. This thesis includes results from [GKS13, GK14, GHS15], though we omit some results and generalize other results from these papers. [GKS13] is joint work with Swastik Kopparty and Madhu Sudan. [GK14] is joint work with Swastik Kopparty. [GHS15] is joint work with Elad Haramaty and Madhu Sudan.

1.2.1 Main Result

The main result of our thesis is the construction of high-rate LCCs and LTCs.

Theorem 1.2.1 (Main theorem, informal)**.** $\forall \epsilon, \beta > 0\ \exists \delta, \alpha > 0$ *such that for infinitely many* n *there exists* $q = q(n) = O(n^{\epsilon})$ *and a linear code* $\mathcal{C} \subseteq \mathbb{F}_q^n$ *of distance* δ *and rate* $1 - \beta$ *such that*

- $\mathcal{C}$ *is locally correctable and decodable from* $\Omega(\delta)$ *fraction errors with* $O(n^{\epsilon})$ *queries;*
- $\mathcal{C}$ *is list-decodable from* $1 - \sqrt{1-\delta}$ *fraction errors in polynomial time, and is locally list-decodable from* $1 - \sqrt{1-\delta}$ *fraction errors with* $O(n^{3\epsilon})$ *queries;*
- $\mathcal{C}$ *has an* α*-robust tester using* $O(n^{2\epsilon})$ *queries.*

Theorem 1.2.1 is proved in Section 7.1.5, and is the culmination of the work in this thesis.

1.2.2 Lifting

The lifting operation was first defined and used in [BSMSS11] to prove negative results — in particular, to construct "symmetric LDPC codes" that are not testable. The work of [GKS13] initiated the systematic study of lifting, and also was the first work to use lifting to prove positive results — in particular, to construct new high-rate locally correctable and locally testable codes. Our definition of lifting is more general and somewhat cleaner than that of [BSMSS11]. Starting from a base linear affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_{q^t} \to \mathbb{F}_q\}$, we define the *$m$-dimensional lift* $\mathcal{C}^{t \nearrow m} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ to be the code consisting of all $f : \mathbb{F}_q^m \to \mathbb{F}_q$ satisfying the following: f is in the lift if and only if, for every t-dimensional affine subspace $A \subseteq \mathbb{F}_q^m$, the restriction $f|_A \in \mathcal{C}$.

Lifting is a natural algebraic operation on affine-invariant codes, which is evident in the generic properties that lifted codes naturally possess. We show that if the base code $\mathcal{C}$ has (relative) distance δ, then the lifted code $\mathcal{C}^{t \nearrow m}$ has distance $\delta - q^{-t}$, and in fact if $q \in \{2, 3\}$, then the lifted distance is also δ. So, lifting approximately preserves distance. Moreover, $\mathcal{C}^{t \nearrow m}$ is naturally q^t-single-orbit characterized by construction, and so, by [KS08], the natural t-dimensional test — choose a random t-dimensional affine subspace $A \subseteq \mathbb{F}_q^m$ and accept if and only if $f|_A \in \mathcal{C}$ — is $\Omega(q^{-2t})$-sound. Finally, lifted codes are naturally locally correctable — to correct f at a point $\mathbf{x} \in \mathbb{F}_q^m$, choose a random t-dimensional subspace $A \subseteq \mathbb{F}_q^m$ passing through $\mathbf{x}$, use the correction algorithm for $\mathcal{C}$ to correct $f|_A$ to a codeword $c \in \mathcal{C}$, and then output $c(\mathbf{x})$.

1.2.3 Robust Testing of Lifted Codes

In [GHS15], we consider robust testing of lifted codes. We propose and analyze the following test for $\mathcal{C}^{t \nearrow m}$: Pick a random $2t$-dimensional subspace A in $\mathbb{F}_q^m$ and accept if $f|_A \in \mathcal{C}^{t \nearrow 2t}$. Our main theorem relates the robustness of this test to the distance of the code $\mathcal{C}$.

Theorem 1.2.2. *$\forall \delta > 0$ $\exists \alpha > 0$ such that the following holds: For every finite field $\mathbb{F}_q$, for every pair of positive integers t and m, and for every affine-invariant code $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ satisfying $\delta(\mathcal{C}) \geq \delta$, the code $\mathcal{C}^{t \nearrow m}$ has a q^{2t}-local test that is α-robust.*

Theorem 1.2.2 is proved in Section 6.1. As we elaborate below, Theorem 1.2.2 immediately implies a robust analysis for low-degree tests. Whereas almost all previous robust analyses of low-degree tests had more complex conditions on the relationship between the robustness, the degree, and the field size, our relationship is extremely clean. The dependence of α on δ that we prove is polynomial but of fairly high degree $\alpha = \Omega(\delta^{74})$. We do not attempt to improve this relationship in this thesis and choose instead to keep the intermediate statements simple and general. We note that a significant portion of this complexity arises due to our desire to lift t-dimensional codes for general t, and here the fact that the robustness lower-bound is independent of t is itself significant.

Comparing with other testing results for lifted codes, there are only two prior works to compare with: Kaufman and Sudan [KS08] analyze a tester for a broader family of codes that they call "single-orbit" codes. Their result would yield a robustness of $\Theta(q^{-3t})$. (See Corollary 6.1.2.)

Haramaty et al. [HRS13] also give a tester for lifted codes. They do not state their results in terms of robustness but their techniques would turn into a robustness of $\epsilon_q \cdot \delta$, where the ϵ_q is a positive constant for constant q but goes to zero extremely quickly as $q \to \infty$. Thus for growing q (and even slowly shrinking δ) our results are much stronger.

Proof approach and some technical contributions

In order to describe our test and analysis techniques, we briefly review the two main tests proposed in the literature for "low-degree testing", when the field size is much larger than the degree. The most natural test for this task is the one that picks a random line in $\mathbb{F}_q^m$ and computes the proximity of the function restricted to this line to the space of univariate degree d polynomials. This is the test proposed by Rubinfeld and Sudan [RS96] and analyzed in [RS96, ALM+98, AS03]. A second low-degree test is somewhat less efficient in its query complexity (quadratically so) but turns out to have a much simpler analysis — this test would pick a random two-dimensional (affine) subspace in $\mathbb{F}_q^m$ and verify that the function is a bivariate polynomial of degree at most d on this subspace. This is the test proposed

by Raz and Safra [RS97] and analyzed in [RS97, MR06]. Both tests can be analyzed by first reducing the testing problem to that of testing constant variate functions (at most four variate functions) and then analyzing the constant dimensional problem as a second step.

The first step is completely generic or at least it was sensed to be so. However there was no prior formalization of the fact that it is generic. The only class of functions to which it has been applied is the class of low-degree polynomials and a priori it is not clear how to even justify the claim of genericity. Here we show that the first step applies to all lifted codes, thus giving the first justification of the presumed genericity of this step, which we consider to be a conceptual contribution.

For the second step, the robust analyses in [ALM+98, AS03] are quite algebraic and there seems to be no hope to use them on general lifted codes. The test and analysis of Raz and Safra [RS97] on the other hand feels much more generic. In this work we use their test, and extend it to general lifted codes and show that it is robust. Even the extension of the test is not completely obvious. In particular, to test low-degree polynomials they look at restrictions of the given function to 2-dimensional "planes". When lifting t-dimensional properties, it is not immediate what would be the dimension of the restrictions the test should look at: Should it be $t+1$? Or $2t$ or maybe $3t-1$ (each of which does make logical sense)? We show that the $2t$ dimensional tests are robust, with robustness being independent of t.

Next we turn to our analysis. In showing robustness of their test, applied to generic lifted codes there is a major barrier: Almost all analyses of low-degree tests, for polynomials of degree at most d, attempt to show first that a function passing the test with high probability is close to a polynomial of degree *twice* the degree, i.e., at most $2d$, with some additional features. They then use the distance of the space of polynomials of degree $2d$ and the additional features to establish that the function being tested is really close to a degree d polynomial. In extending such analyses to our setting we face two obstacles: In the completely generic setting, there is no nice notion corresponding to the set of degree $2d$ polynomials. One approach might be to consider the linear space spanned by products of functions in our basic space and work with them, but the algebra gets hairy to understand and

analyze. Even if we abandon the complete genericity and stick to the space of polynomials of degree d, but now allow $d > q/2$ we hit a second obstacle: The space of polynomials of degree $2d$ have negligible relative distance compared to the space of polynomials of degree d.

Thus we need to search for a new proof technique and we find one by unearthing a new connection between "lifted codes" and "tensor product" codes. The tensor product is a natural operation in linear algebra and when applied to two linear codes, it produces a new linear code in a natural way. Tensor products of codes are well-studied in the literature on coding theory. The testing of tensor product codes was initiated by Ben-Sasson and Sudan [BSS06] and subsequently has been well-studied [DSW06, Val05, BSV09b, BSV09a, GGR09]. Specifically, a recent result of Viderman [Vid12] gives a powerful analysis which we are able to reproduce in a slightly different setting to get our results. In particular this is the ingredient that allows us to work with base codes whose distance is less than $1/2$. Also, for the sake of the exposition we pretend that this test can test two-dimensional tensor products of one dimensional codes, with one-dimensional tests. (Actually, the test works with three dimensional tensors and tests them by looking at two-dimensional planes, but by suppressing this difference, our exposition becomes a little simpler.)

To explain the connection between lifted codes and tensor product codes, and the idea that we introduce to test the former, we turn to the simple case of testing a bivariate lift of a univariate Reed-Solomon code. Specifically, let $\mathcal{C}$ be the family of univariate polynomials of degree at most d mapping $\mathbb{F}_q$ to $\mathbb{F}_q$. Let $\mathcal{C}_2$ be the family of bivariate polynomials that become a univariate polynomial of degree at most d on every restriction to a line. The tensor product of $\mathcal{C}$ with itself, which we denote $\mathcal{C}^{\otimes 2}$ corresponds to the set of bivariate polynomials of degree at most d in each variable. Clearly $\mathcal{C}_2 \subseteq \mathcal{C}^{\otimes 2}$ but such subset relationships are not of immediate use in testing a code. (Indeed locally testable codes contain many non-LTCs.) To get a tighter relationship, now fix two "directions" d_1 and d_2 and let $\mathcal{C}_{d_1,d_2}$ be the code containing all bivariate polynomials over $\mathbb{F}_q$ that on every restriction to lines in directions d_1 and d_2 form univariate degree d polynomials. On the one hand the code $\mathcal{C}_{d_1,d_2}$ is just isomorphic to the tensor product code $\mathcal{C}^{\otimes 2}$ which is testable by the natural test,

by our assumption. On the other hand, we now have $\mathcal{C}_2 = \cap_{d_1,d_2}\mathcal{C}_{d_1,d_2}$ so we now have a characterization of the lifted codes in terms of the tensor product. One might hope that one could use this characterization to get a (robust) analysis of the lifted test since it tests membership in $\mathcal{C}_{d_1,d_2}$ for random choices of d_1 and d_2, but unfortunately we do not see a simple way to implement this hope.

Our key idea is look instead at a more complex family of codes $\mathcal{C}_{d_1,d_2,d_3}$ that consists of functions of degree d in directions d_1, d_2 and d_3. (Of course now d_1, d_2, d_3 are linearly dependent and so $\mathcal{C}_{d_1,d_2,d_3}$ is not a tensor product code. We will return to this issue later.) We still have $\mathcal{C}_2 = \cap_{d_1,d_2,d_3}\mathcal{C}_{d_1,d_2,d_3}$. Indeed we can even fix d_1, d_2 arbitrarily (only requiring them to be linearly independent) and we have $\mathcal{C}_2 = \cap_{d_3}\mathcal{C}_{d_1,d_2,d_3}$. This view turns out to be more advantageous since we now have that for any d_3 and d_3' we have $\mathcal{C}_{d_1,d_2,d_3} \cup \mathcal{C}_{d_1,d_2,d_3'} \subseteq \mathcal{C}_{d_1,d_2}$ which is a code of decent distance. This allows us to show that if the function being tested is close to $\mathcal{C}_{d_1,d_2,d_3}$ for many choices of d_3 then the nearest codewords for all these choices of d_3 are *the same*. An algebraic analysis of lifted codes tells us that a codeword of $\mathcal{C}_{d_1,d_2}$ can not be in $\mathcal{C}_{d_1,d_2,d_3}$ for many choices of d_3 without being a codeword of the lifted code and this lends promise to our idea. But we are not done, since we still need to test the given function for proximity to $\mathcal{C}_{d_1,d_2,d_3}$ and this is no longer a tensor product code so Viderman's result does not apply directly. Fortunately, we are able to develop the ideas from Viderman's analysis for tensor product codes [Vid12] and apply them also to our case and this yields our test and analysis. We note that this extension is not immediate — indeed one of the central properties of tensor product codes is that they are decodable from some clean erasure patterns and this feature is missing in our codes. Nevertheless the analysis can be modified to apply to our codes and this suffices to complete the analysis.

In the actual implementation, as noted earlier, we can't work with univariate tests even for the simple case above, and work instead by using a bivariate test for trivariate and 4-variate functions. (This is similar to the reasons why Raz and Safra used a bivariate test.) This complicates the notations a bit, but the idea remains similar to the description above. Our task gets more complicated when the base code being lifted is t-dimensional for $t > 1$.

The most natural adaptation of our analysis leads to dependencies involving δ (the distance of the base code) and t. We work somewhat harder in this case to eliminate any dependence on t while working within the framework described above.

1.2.4 Applications

Lifted Reed-Solomon Code

The most interesting construction arising from lifting is the "lifted Reed-Solomon code". As its name suggests, the lifted Reed-Solomon code is obtained by simply lifting the Reed-Solomon code, a univariate code, to m dimensions. The lifted Reed-Solomon, by definition, contains the Reed-Muller code. However, if the degree d of the Reed-Solomon code is sufficiently large relative to the field size — in particular, if $d \geq q - q/p$, where $\mathbb{F}_q$ has characteristic p — then the lifted Reed-Solomon code contains polynomials of high degree as well. This fact follows from the characterization of low-degree polynomials as proven in [KR06]. In fact, using structural results about affine-invariance and lifts, we easily re-prove a special case of the characterization of low-degree polynomials in Theorem 7.1.6. The lifted Reed-Solomon code is arguably more natural than the Reed-Muller code, the latter of which is an unnecessarily sparse subcode of the former. We therefore expect the lifted Reed-Solomon code to exhibit the same versatility as that of the Reed-Muller code, and indeed since the lifted Reed-Solomon code is a lifted code, it generically has good distance and is locally decodable and testable. We show that it is also in fact (locally) list-decodable up to the Johnson radius, just like the Reed-Muller code. What sets the lifted Reed-Solomon code apart from the Reed-Muller code is its vastly greater rate. If one insists on having positive distance $\delta > 0$, then the m-variate Reed-Muller code has rate bounded by $1/m!$, whereas the rate of the lifted Reed-Solomon code approaches 1 as $\delta \to 0$.

The first family of high-rate locally correctable codes known were the multiplicity codes of Kopparty, Saraf, and Yekhanin [KSY14]. Kopparty [Kop12] showed that multiplicity codes are also locally list-decodable up to the Johnson radius. The only prior construction of high-rate codes that are robustly testable are the tensor product codes of Viderman [Vid12].

Thus, the lifted Reed-Solomon code is the first high-rate code known to be locally correctable, locally list-decodable up to the Johnson radius, and robustly testable. This is the code in Theorem 1.2.1.

Robust Low-Degree Testing

An almost direct corollary of our robustness result for lifted codes is a q^4-local robust low-degree test for the setting $d \leq (1-\delta)q$. To see why we get q^4 queries, note that when $d \geq q - q/p$, the m-variate Reed-Muller code of degree d is not equal to the m-dimensional lift of the degree d Reed-Solomon code. But the latter turns out to be the m-dimensional lift of the bivariate Reed-Muller code of degree d. Applying our robust testing result to this lifted family yields a robust test making q^4 queries. But with some slight extra work we can get a better tester that makes only q^2 queries and this yields the following theorem.

Theorem 1.2.3. *$\forall \delta > 0$ $\exists \alpha > 0$ such that the following holds: For every finite field $\mathbb{F}_q$, for every integer $d \leq (1-\delta)q$ and every positive integer m, there is a q^2-query α-robust low-degree test for the class of m-variate polynomials of degree at most d over $\mathbb{F}_q$.*

We note that previous works on low-degree testing worked only when $d < q/2$. This ratio seems to be achieved by Friedl and Sudan (see [FS95, Theorem 13]). Other works [RS96, ALM+98, RS97, AS03, MR06] seem to achieve weaker ratios for a variety of reasons discussed above.

Nikodym Set Size Bounds

One of the applications of lifted codes is to bounding, from below, the size of "Nikodym sets" over finite fields (of small characteristic). A set $S \subseteq \mathbb{F}_q^m$ is a *Nikodym set* if every point $\mathbf{x}$ has a line passing through it such that all points of the line, except possibly the point $\mathbf{x}$ itself, are elements of S. Nikodym sets are closely related to "Kakeya sets" — the latter contain a line in every direction, while the former contain almost all of a line through every point. A lower bound for Kakeya sets over finite fields was proved by Dvir [Dvi08] using the polynomial method and further improved by using the "method of multiplicites" by Saraf and

Sudan [SS08] and Dvir et al. [DKSS09]. Kakeya sets have seen applications connecting its study to the study of randomness extractors, especially [DS07, DW11]. Arguably, Nikodym sets are about as natural in this connection as Kakeya sets.

Previous lower bounds on Kakeya sets were typically also applicable to Nikodym sets and led to bounds of the form $|S| \geq (1 - o(1))q^m/2^m$ where the $o(1)$ term goes to zero as $q \to \infty$. In particular, previous lower bounds failed to separate the growth of Nikodym sets from those of Kakeya sets. In our work, we present a simple connection that shows that the existence of the lifted Reed-Solomon code yields a large lower bound on the size of Nikodym sets, thereby significantly improving the known lower bound on the size of Nikodym sets over fields of constant characteristic.

Theorem 1.2.4. *For every prime p, and every integer m, there exists $\epsilon = \epsilon(p, m) > 0$ such that for every finite field $\mathbb{F}_q$ of characteristic p, if $S \subseteq \mathbb{F}_q^m$ is a Nikodym set, then $|S| \geq q^m - q^{(1-\epsilon)m}$. In particular, if $q \to \infty$, then $|S| \geq (1 - o(1)) \cdot q^m$.*

Thus, whereas previous lower bounds on the size of Nikodym sets allowed for the possibility that the density of the Nikodym sets vanishes as m grows, ours show that Nikodym sets occupy almost all of the space. One way to view our results is that they abstract the polynomial method in a more general way, and thus lead to stronger lower bounds in some cases.

1.2.5 Organization

In Chapter 2, we establish notation and preliminary definitions. In Chapter 3, we formally define error correcting codes and relevant models: local decoding, correcting, list-decoding, and testing. In Chapter 4, we review some structural properties of affine-invariant codes. In Chapter 5, we formally define the lifting operation and prove generic properties of lifted codes. In Chapter 6, we prove that lifted codes are robustly testable with robustness parameter that depends only on the distance of the base code. In Chapter 7, we discuss applications of lifted codes: construction of the lifted Reed-Solomon code, robust low-degree testing, and new lower bounds on the size of Nikodym sets.

Chapter 2

Preliminaries

2.1 Notation

Letters. We will typically use lower-case italic letters (e.g. a, b, c) to denote scalar values, lower-case bold letters (e.g. $\mathbf{a}, \mathbf{b}, \mathbf{c}$) to denote vectors, and upper-case bold letters (e.g. $\mathbf{A}, \mathbf{B}, \mathbf{C}$) to denote matrices. If $\mathbf{A}$ is a matrix, then we denote by $\mathbf{A}_{i*}$ the i-th row of $\mathbf{A}$ and by $\mathbf{A}_{*j}$ the j-th row of $\mathbf{A}$.

Sets and functions. For a set S and $n \in \mathbb{N}$, let $\binom{S}{n}$ be the collection of subsets $T \subseteq S$ with $|T| = n$. Let 2^S be the collection of all subsets of S. For a positive integer n, define $[n] \triangleq \{1, \ldots, n\}$ and $[\![n]\!] \triangleq \{0, 1, \ldots, n-1\}$. For sets A and B, let $\{A \to B\}$ denote the set of functions from A to B.

Vectors and Hamming distance. If $\mathbf{a} \in \mathbb{N}^m$, let $\|\mathbf{a}\| \triangleq \sum_{i=1}^{m} a_i$. If Σ is a finite set, $n \in \mathbb{N}$, and $\mathbf{s} \in \Sigma^n$ is a string, then let s_i denote the i-th component of $\mathbf{s}$, for $i \in [n]$, so that $\mathbf{s} = (s_1, \ldots, s_n)$. For two vectors $\mathbf{a}, \mathbf{b} \in \Sigma^n$, denote their Hamming distance by $\Delta(\mathbf{a}, \mathbf{b}) \triangleq \#\{i \in [n] \mid a_i \neq b_i\}$ and their normalized Hamming distance by $\delta(\mathbf{a}, \mathbf{b}) \triangleq \frac{|\Delta(\mathbf{a},\mathbf{b})|}{n}$. If $S \subseteq \Sigma^n$ is a set, then $\delta(\mathbf{a}, S) \triangleq \min_{\mathbf{b} \in S} \delta(\mathbf{a}, \mathbf{b})$ is the distance from $\mathbf{a}$ to S. If Σ is a field, then the (resp. normalized) Hamming weight of $\mathbf{a} \in \Sigma^n$ is (resp. $\delta(\mathbf{a}, \mathbf{0})$) $\Delta(\mathbf{a}, \mathbf{0})$. We will frequently think of functions $f : A \to B$ as vectors in B^A and so we extend the vector notations to

functions as well. In particular, if $f, g : A \to B$, then $\Delta(f,g) \triangleq \#\{x \in A \mid f(x) \neq g(x)\}$ and $\delta(f,g) \triangleq \frac{\Delta(f,g)}{|A|}$. Note that $\delta(f,g) = \Pr_{x \in A}[f(x) \neq g(x)]$.

Minkowski sums and affine subspaces. For sets $A, B \subseteq \mathbb{F}^m$, their Minkowski sum is denoted $A + B \triangleq \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in A, \mathbf{b} \in B\}$. For $A \subseteq \mathbb{F}^m$, the span of A is denoted $\mathrm{span}(A) \triangleq \{\sum_{\mathbf{a} \in A} c_{\mathbf{a}} \cdot \mathbf{a} \mid c_{\mathbf{a}} \in \mathbb{F}\}$. For $\mathbf{x} \in \mathbb{F}^m$ and $A \subseteq \mathbb{F}^m$, let $(\mathbf{x}, A) \triangleq \{\mathbf{x}\} + \mathrm{span}(A)$ be the *affine subspace through* $\mathbf{x}$ *in directions* A.

Shadows. Let $p, a, b \in \mathbb{N}$ and $a = \sum_{i \geq 0} a^{(i)} p^i$, $b = \sum_{i \geq 0} b^{(i)} p^i$ with $a^{(i)}, b^{(i)} \in [\![p]\!]$ for $i \geq 0$. Then a *is in the* p*-shadow of* b, denoted by $a \leq_p b$, if $a^{(i)} \leq b^{(i)}$ for $i \geq 0$. If $\mathbf{a}, \mathbf{b} \in \mathbb{N}^m$, then $\mathbf{a} \leq_p \mathbf{b}$ means that $a_i \leq_p b_i$ for every $i \in [m]$. If $\mathbf{a} \in \mathbb{N}^n$ and $b \in \mathbb{N}$, then $\mathbf{a} \leq_p b$ means that $\sum_{i \in S} a_i \leq_p b$ for any subset $S \subseteq [n]$. If $\mathbf{A} \in \mathbb{N}^{m \times n}$ and $\mathbf{b} \in \mathbb{N}^m$, then $\mathbf{A} \leq_p \mathbf{b}$ means that $\mathbf{A}_{i*} \leq_p \mathbf{b}$ for every $i \in [m]$.

Finite fields and polynomials. If p is a prime and q is a power of p, then $\mathbb{F}_q$ denotes the finite field of size q. If $\mathbf{X} = (X_1, \ldots, X_m)$ are variables and $\mathbf{d} \in \mathbb{N}^m$, then $\mathbf{X}^{\mathbf{d}}$ denotes the monomial $\prod_{i=1}^m X_i^{d_i}$. Any polynomial $h(\mathbf{X}) \in \mathbb{F}_q[\mathbf{X}]$ with such that $h(\mathbf{X}) = f(\mathbf{X}) + g(\mathbf{X}) \prod_{i=1}^m (X_i^q - X_i)$ for some $g(\mathbf{X}) \in \mathbb{F}_q[\mathbf{X}]$ defines the same function $\mathbb{F}_q^m \to \mathbb{F}_q$, since the polynomial $X^q - X$ is identically zero on $\mathbb{F}_q$. Observe that every function $f : \mathbb{F}_q^m \to \mathbb{F}_q$ can be expressed uniquely as a linear combination $f(\mathbf{X}) = \sum_{\mathbf{d} \in [\![q]\!]^m} f_{\mathbf{d}} \cdot \mathbf{X}^{\mathbf{d}}$ of monomials $\mathbf{X}^{\mathbf{d}}$ with $\mathbf{d} \in [\![q]\!]^m$. Throughout this thesis, when we refer to a "polynomial" $f : \mathbb{F}_q^m \to \mathbb{F}_q$, we mean the unique $f(\mathbf{X})$ defined above. The *support of* f is $\mathrm{supp}(f) \triangleq \{\mathbf{d} \in [\![q]\!]^m \mid f_{\mathbf{d}} \neq 0\}$. The *degree of a polynomial* $f : \mathbb{F}_q^m \to \mathbb{F}_q$ is $\deg(f) \triangleq \max\{\|\mathbf{d}\| \mid \mathbf{d} \in \mathrm{supp}(f)\}$. When we refer to a polynomial $f : \mathbb{F}_q^m \to \mathbb{F}_q$ in the context of codewords, we refer to the q^m-dimensional vector $(f(\mathbf{x}))_{\mathbf{x} \in \mathbb{F}_q^m}$.

Mod-star. For $a \in \mathbb{N}$ and $b > 1$, define the operation $\bmod^*$ by

$$a \bmod^* b = \begin{cases} a & a < b \\ a \bmod (b-1) & a \geq b \end{cases}$$

so that $X^a \equiv X^{a \bmod^* b} \pmod{X^b - X}$.

2.2 Probability and Concentration bounds

If X is a random variable, we use $\mathbb{E}[X]$ and $\mathrm{Var}[X]$ to denote the expectation and variance of X, respectively. If the probability space of X is not clear from context, we use subscripts, e.g. if $X = X(a, b)$, then $\mathbb{E}_a[X]$ is the average over a with b fixed.

Proposition 2.2.1 (Markov inequality)**.** *Let $X \geq 0$ be a random variable with $\mathbb{E}[X] < \infty$ and let $a > 0$. Then $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$.*

Proposition 2.2.2 (Chebyshev inequality)**.** *Let X be a random variable with $\mathbb{E}[X] < \infty$ and $\mathrm{Var}[X] < \infty$, and let $a > 0$. Then $\Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}$.*

Proposition 2.2.3 (Hoeffding inequality)**.** *Let $X_1, \ldots, X_n \in [0, 1]$ be independent random variables and let $\overline{X} \triangleq \frac{1}{n}\sum_{i=1}^n X_i$. Let $\epsilon > 0$. Then $\Pr\left[\left|\overline{X} - \mathbb{E}[\overline{X}]\right| > \epsilon\right] \leq 2\exp(-2n\epsilon^2)$.*

Proposition 2.2.4. *Let $f, g : \mathbb{F}_q^m \to \mathbb{F}_q$. Let $\widetilde{\delta}(f, g)$ be the estimate of $\delta(f, g)$ by random sampling, i.e. independent uniformly random $\mathbf{x} \in \mathbb{F}_q^m$ are chosen and the average of $\mathbb{1}_{f(\mathbf{x}) \neq g(\mathbf{x})}$ is output. If $\Theta(\ln(1/\eta)/\epsilon^2)$ queries are used in the sample, then with probability at least $1 - \eta$, the estimate $\widetilde{\delta}(f, g)$ has additive error at most ϵ.*

Proof. Let $n \geq \ln(2/\eta)/(2\epsilon^2)$. For each $i \in [n]$, let $X_i \triangleq \mathbb{1}_{f(\mathbf{x}_i) \neq g(\mathbf{x}_i)}$, and define $\overline{X} \triangleq \frac{1}{n}\sum_{i=1}^n X_i$ so that $\widetilde{\delta}(f, g) = \overline{X}$, and $\delta(f, g) = \mathbb{E}[\overline{X}]$. By Proposition 2.2.3,

$$\Pr\left[\left|\widetilde{\delta}(f, g) - \delta(f, g)\right| > \epsilon\right] \leq 2\exp(-2n\epsilon^2) \leq \eta.$$

□

Chapter 3

Error Correcting Codes

3.1 Classical Parameters

Error correcting codes are schemes for encoding messages as codewords to protect them from noise. The code itself is the subset of valid codewords. The rate of a code is the ratio of the message length to the encoding length, and measures the efficiency of the encoding. The distance of a code measures the minimum distance between valid codewords, and indicates the error-correcting capability of the code. In this section, we formally define these notions.

Definition 3.1.1 (Code)**.** Let Σ be a finite set and let n be a natural number. A *code over* Σ *of block length* n is a subset $\mathcal{C} \subseteq \Sigma^n$. If there exists a set Σ_0, integer $k \leq n$, and injective function $\mathsf{Enc} : \Sigma_0^k \to \Sigma^n$ such that $\mathsf{Enc}(\Sigma_0^k) = \mathcal{C}$, then Enc is an *encoding function for* $\mathcal{C}$.

Definition 3.1.2 (Rate of a code)**.** The *rate* of a code $\mathcal{C} \subseteq \Sigma^n$ is $\frac{\log |\mathcal{C}|}{n \log |\Sigma|}$.

Definition 3.1.3 (Distance of a code)**.** The *(normalized) distance* of a code $\mathcal{C} \subseteq \Sigma^n$ is $\delta(\mathcal{C}) \triangleq \min_{\mathbf{x} \neq \mathbf{y} \in \mathcal{C}} \delta(\mathbf{x}, \mathbf{y})$.

The following proposition shows that, algorithmic efficiency aside, any code supports unique decoding up to half its minimum distance.

Proposition 3.1.4. *For every* $\mathbf{r} \in \Sigma^n$*, there exists at most one* $\mathbf{c} \in \mathcal{C}$ *with* $\delta(\mathbf{r}, \mathbf{c}) < \frac{\delta(\mathcal{C})}{2}$.

Proof. Suppose $\mathbf{c}, \mathbf{c}' \in \mathcal{C}$ and $\delta(\mathbf{r}, \mathbf{c}), \delta(\mathbf{r}, \mathbf{c}') < \frac{\delta(\mathcal{C})}{2}$. Then, by the triangle inequality, $\delta(\mathbf{c}, \mathbf{c}') \leq \delta(\mathbf{c}, \mathbf{r}) + \delta(\mathbf{r}, \mathbf{c}') < \delta(\mathcal{C})$, so $\mathbf{c} = \mathbf{c}'$. □

Linear codes are codes whose alphabet is a (finite) field and whose codewords form a vector space over the field. Every code of interest to us in this thesis is a linear code.

Definition 3.1.5 (Linear code)**.** A code $\mathcal{C} \subseteq \Sigma^n$ is *linear* if $\Sigma = \mathbb{F}$ is a field and $\mathcal{C}$ is a linear subspace of $\mathbb{F}^n$.

Proposition 3.1.6. *If $\mathcal{C} \subseteq \mathbb{F}^n$ is a linear code, then $\delta(\mathcal{C})$ is equal to the minimal normalized Hamming weight of nonzero $\mathbf{c} \in \mathcal{C}$.*

Proof. Let $\mathbf{c} \in \mathcal{C}$ be nonzero of minimal Hamming weight. By definition, $\delta(\mathcal{C}) \leq \delta(\mathbf{c}, \mathbf{0})$. On the other hand, for any two distinct $\mathbf{x}, \mathbf{y} \in \mathcal{C}$, $\delta(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}, \mathbf{0}) \geq \delta(\mathbf{c}, \mathbf{0})$, and so minimizing over $\mathbf{x} \neq \mathbf{y}$, we have $\delta(\mathcal{C}) \geq \delta(\mathbf{c}, \mathbf{0})$. □

3.2 Local decoding, correcting, and list-decoding

In this section, we formally define the models of local decoding, local correcting, and local list-decoding. Intuitively, local decoding entails recovering a symbol of the original message using few queries, while local correcting entails recovering a symbol of the original codeword using few queries.

Definition 3.2.1 (Local decoding)**.** A code $\mathcal{C} \subseteq \Sigma^n$ with encoding function $\mathsf{Enc} : \Sigma_0^k \to \Sigma^n$ is (ℓ, τ, ϵ)-*locally decodable* if there exists a randomized oracle $\mathcal{A} : [k] \to \Sigma_0$ with oracle access to a received word $\mathbf{r} \in \Sigma^n$ such that

1. $\mathcal{A}^r$ queries at most ℓ symbols of $\mathbf{r}$;
2. if there is $\mathbf{m} \in \Sigma_0^k$ with $\delta(\mathsf{Enc}(\mathbf{m}), \mathbf{r}) \leq \tau$, then $\Pr[\mathcal{A}^r(i) = m_i] \geq 1 - \epsilon$ for every $i \in [k]$.

Definition 3.2.2 (Local correcting)**.** A code $\mathcal{C} \subseteq \Sigma^n$ is (ℓ, τ, ϵ)-*locally correctable* if there exists a randomized oracle $\mathcal{A} : [n] \to \Sigma$ with oracle access to a received word $\mathbf{r} \in \Sigma^n$ such that

1. $\mathcal{A}^r$ queries at most ℓ symbols of $\mathbf{r}$;

2. if there is $\mathbf{c} \in \mathcal{C}$ with $\delta(\mathbf{c}, \mathbf{r}) \leq \tau$, then $\Pr[\mathcal{A}^r(i) = c_i] \geq 1 - \epsilon$ for every $i \in [n]$.

If $\mathcal{C}$ is a linear code, then it is possible to encode $\mathcal{C}$ in *systematically*, i.e. such that the original message is part of the codeword. Of course, this is not algorithmically satisfying unless the systematic encoding function is *explicit*, i.e. computable in polynomial time. When $\mathcal{C}$ is linear, we can think of it as a space of functions $\{S \to \mathbb{F}_q\}$. A systematic encoding is then equivalent to finding an interpolating set for $\mathcal{C}$ in S.

Definition 3.2.3. If $\mathcal{C} \subseteq \{S \to \mathbb{F}_q\}$ is linear, then $I \subseteq S$ is an *interpolating set for* $\mathcal{C}$ if, for every $f : I \to \mathbb{F}_q$, there exists a unique extension $g \in \mathcal{C}$ such that $g|_I = f$.

Remark 3.2.4. If $\mathcal{C} \subseteq \{S \to \mathbb{F}_q\}$ has an interpolating set $I \subseteq S$, then $|\mathcal{C}| = q^{|I|}$. In particular, if $\mathcal{C}$ is linear, then $|I| = \dim_{\mathbb{F}_q}(\mathcal{C})$.

Proposition 3.2.5. *If $\mathcal{C} \subseteq \{S \to \mathbb{F}_q\}$ is linear and has an explicit interpolating set $I \subseteq S$, and $\mathcal{C}$ is a (ℓ, τ, ϵ)-locally correctable code, then $\mathcal{C}$ is a (ℓ, τ, ϵ)-locally decodable code.*

Proof. Let Enc be the map which takes $f : I \to \mathbb{F}_q$ to its unique extension $g \in \mathcal{C}$ such that $g|_I = f$, guaranteed by the fact that I is an interpolating set. Then the local correcting algorithm for $\mathcal{C}$ also serves as the local decoding algorithm, when restricted to I. □

A local list-decoding algorithm outputs a list of oracles, such that each valid codeword within the given radius is computed by some oracle in the output list.

Definition 3.2.6 (Local list-decoding)**.** A code $\mathcal{C} \subseteq \Sigma^n$ is $(\ell_1, \ell_2, \tau, L, \epsilon, \eta)$-*locally list-decodable* if there exists a randomized algorithm $\mathcal{A}$ with oracle access to a received word $\mathbf{r} \in \Sigma^n$ that outputs a list $M_1, \ldots, M_L : [n] \to \Sigma$ of randomized oracles with oracle access to $\mathbf{r}$, such that

1. $\mathcal{A}^r$ queries at most ℓ_1 symbols of $\mathbf{r}$;

2. for each $j \in [L]$, M_j queries at most ℓ_2 symbols of $\mathbf{r}$;

3. with probability at least $1 - \eta$, the following holds: if there is $\mathbf{c} \in \mathcal{C}$ with $\delta(\mathbf{c}, \mathbf{r}) \leq \tau$, then there is some $j \in [L]$ such that $\Pr[M^r(i) = c_i] \geq 1 - \epsilon$ for every $i \in [n]$.

3.3 Local testing and robust testing

In this section, we formally define the model of testing. Since we will be solely interested in the testing of linear codes, we only present the definition of local testing in the context of linear codes. We also define the notions of soundness and robustness, and prove some simple relationships between the two.

Definition 3.3.1 (Local testing)**.** A *ℓ-local tester* for a code $\mathcal{C} \subseteq \{S \to \mathbb{F}_q\}$ is a randomized algorithm $\mathcal{T}$ with oracle access to a received word $f : S \to \mathbb{F}_q$, which randomly samples (π, V) according to some distribution $\mathcal{D}$ on $(S^\ell, 2^{\mathbb{F}_q^\ell})$, with $\pi = (\pi_1, \ldots, \pi_\ell) \in S^\ell$ and $V \subseteq \mathbb{F}_q^\ell$ and accepts if and only if $f|_\pi \triangleq (f(\pi_1), \ldots, f(\pi_\ell)) \in V$.

The tester is *ϵ-sound* if $\mathcal{T}$ accepts $f \in \mathcal{C}$ with probability one, and rejects $f \notin \mathcal{C}$ with probability at least $\epsilon \cdot \delta(f, \mathcal{C})$.

The tester is *α-robust* if $\mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}}\left[\delta\left(f|_\pi, V\right)\right] \geq \alpha \cdot \delta(f, \mathcal{C})$.

Proposition 3.3.2. *Let $\mathcal{T}$ be an ℓ-local tester for $\mathcal{C}$. If $\mathcal{T}$ is ϵ-sound, then $\mathcal{T}$ is (ϵ/ℓ)-robust. If $\mathcal{T}$ is α-robust, then $\mathcal{T}$ is α-sound.*

Proof. Suppose $\mathcal{T}$ is ϵ-sound. Observe that if $f|_\pi \notin V$, then $\delta(f|_\pi, V) \geq 1/\ell$. Therefore,

$$\begin{aligned}
\mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}}\left[\delta\left(f|_\pi, V\right)\right] &= \mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}}\left[\delta\left(f|_\pi, V\right) \mid f|_\pi \notin V\right] \cdot \Pr_{(\pi,V)\leftarrow\mathcal{D}}\left[f|_\pi \notin V\right] && (3.1)\\
&\geq (1/\ell) \cdot \Pr_{(\pi,V)\leftarrow\mathcal{D}}\left[f|_\pi \notin V\right] && (3.2)\\
&\geq (1/\ell) \cdot \epsilon \cdot \delta(f, \mathcal{C}). && (3.3)
\end{aligned}$$

Now suppose $\mathcal{T}$ is α-robust. Then

$$\begin{aligned}
\Pr_{(\pi,V)\leftarrow\mathcal{D}}\left[f|_\pi \notin V\right] &\geq \mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}}\left[\delta\left(f|_\pi, V\right) \mid f|_\pi \notin V\right] \cdot \Pr_{(\pi,V)\leftarrow\mathcal{D}}\left[f|_\pi \notin V\right] && (3.4)\\
&= \mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}}\left[\delta\left(f|_\pi, V\right)\right] && (3.5)\\
&\geq \alpha \cdot \delta(f, \mathcal{C}). && (3.6)
\end{aligned}$$

□

Proposition 3.3.3. *Let $\mathcal{T}_{\mathcal{C}}$ be an ℓ_1-local tester for $\mathcal{C}$ with distribution $\mathcal{D}_{\mathcal{C}}$, and suppose for every (π, V) in the support of $\mathcal{D}$, V has an ℓ_2-local tester $\mathcal{T}_V$ with distribution $\mathcal{D}_V$, and $\ell_2 \leq \ell_1$.*

1. *If $\mathcal{T}_{\mathcal{C}}$ is α_1-robust and $\mathcal{T}_V$ is α_2-robust for every (π, V) in the support of $\mathcal{D}_{\mathcal{C}}$, then $\mathcal{C}$ has an ℓ_2-local tester that is $(\alpha_1 \cdot \alpha_2)$-robust.*

2. *If $\mathcal{T}_{\mathcal{C}}$ is α-robust and $\mathcal{T}_V$ is ϵ-sound for every (π, V) in the support of $\mathcal{D}_{\mathcal{C}}$, then $\mathcal{C}$ has an ℓ_2-local tester that is $(\alpha \cdot \epsilon)$-sound.*

Proof. Let $\mathcal{D}$ be the following distribution: choose $(\pi, V) \leftarrow \mathcal{D}_{\mathcal{C}}$ and then choose and output $(\pi', V') \leftarrow \mathcal{D}_V$. Let $\mathcal{T}$ be the ℓ_2-tester for $\mathcal{C}$ with distribution $\mathcal{D}$.

1. If $\mathcal{T}_{\mathcal{C}}$ is α_1-robust and $\mathcal{T}_V$ is α_2-robust, then

$$\begin{aligned}
\mathbb{E}_{(\pi',V')\leftarrow\mathcal{D}}\left[\delta\left(f|_{\pi'}, V'\right)\right] &= \mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}_{\mathcal{C}}}\left[\mathbb{E}_{(\pi',V')\leftarrow\mathcal{D}_V}\left[\delta\left(f|_{\pi'}, V'\right)\right]\right] && (3.7)\\
&\geq \alpha_2 \cdot \mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}_{\mathcal{C}}}\left[\delta\left(f|_{\pi}, V\right)\right] && (3.8)\\
&= \alpha_1 \cdot \alpha_2 \cdot \delta(f, \mathcal{C}). && (3.9)
\end{aligned}$$

2. If $\mathcal{T}_{\mathcal{C}}$ is α-robust and $\mathcal{T}_V$ is ϵ-sound, then

$$\begin{aligned}
\Pr_{(\pi',V')\leftarrow\mathcal{D}}\left[f|_{\pi'} \notin V'\right] &= \mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}_{\mathcal{C}}}\left[\Pr_{(\pi',V')\leftarrow\mathcal{D}_V}\left[f|_{\pi'} \notin V'\right]\right] && (3.10)\\
&\geq \epsilon \cdot \mathbb{E}_{(\pi,V)\leftarrow\mathcal{D}_{\mathcal{C}}}\left[\delta\left(f|_{\pi}, V\right)\right] && (3.11)\\
&\geq \alpha \cdot \epsilon \cdot \delta(f, \mathcal{C}). && (3.12)
\end{aligned}$$

□

3.4 Codes

In this section, we present two of the most ubiquitous linear codes: the Reed-Solomon code and the Reed-Muller code. We will directly use the Reed-Solomon code in our constructions,

whereas the Reed-Muller code serves as a benchmark for comparison.

3.4.1 Reed-Solomon code

The Reed-Solomon code consists of evaluations of low-degree univariate polynomials over a finite field $\mathbb{F}_q$.

Definition 3.4.1. Let q be a prime power, and let $d \in \mathbb{N}$. The *Reed-Solomon code* $\mathrm{RS}(q,d)$ *of degree* d *over* $\mathbb{F}_q$ is the code $\mathrm{RS}(q,d) \triangleq \{f : \mathbb{F}_q \to \mathbb{F}_q \mid \deg(f) \leq d\}$.

Proposition 3.4.2. *If* $d < q$*, then the Reed-Solomon code* $\mathrm{RS}(q,d)$ *has distance* $1 - \frac{d}{q}$ *and rate* $\frac{d+1}{q}$.

In [GS99], Guruswami and Sudan showed that the Reed-Solomon code of distance $\delta > 0$ can be efficiently list-decoded up to the Johnson radius $1-\sqrt{1-\delta}$. We will use this algorithm as a subroutine in our list-decoding and local list-decoding algorithms for the lifted Reed-Solomon code in Sections 7.1.3 and 7.1.4, respectively.

Theorem 3.4.3 (Guruswami-Sudan list-decoding [GS99])**.** *For every* $\delta, \epsilon > 0$*, there is a polynomial time algorithm taking as input a function* $r : \mathbb{F}_q \to \mathbb{F}_q$ *and outputs a list* $\mathcal{L}$ *of size* $|\mathcal{L}| = O(1/\epsilon^2)$ *satisfying the following: if* $c \in \mathrm{RS}(q,(1-\delta)q)$ *and* $\delta(r,c) < 1 - \sqrt{1-\delta} - \epsilon$*, then* $c \in \mathcal{L}$.

3.4.2 Reed-Muller code

The Reed-Muller code consists of evaluations of low-degree multivariate polynomials over a finite field $\mathbb{F}_q$.

Definition 3.4.4. Let q be a prime power, and let $d, m \in \mathbb{N}$. The *m-variate Reed-Muller code* $\mathrm{RM}(q,d,m)$ *of degree* d *over* $\mathbb{F}_q$ is the code $\mathrm{RM}(q,d,m) \triangleq \{f : \mathbb{F}_q^m \to \mathbb{F}_q \mid \deg(f) \leq d\}$.

Remark 3.4.5. It follows immediately from definitions that $\mathrm{RM}(q,d,1) = \mathrm{RS}(q,d)$.

Proposition 3.4.6. *If* $d < q$*, then the Reed-Muller code* $\mathrm{RM}(q,d,m)$ *has distance* $1 - \frac{d}{q}$ *and rate* $\frac{\binom{d+m}{m}}{q^m}$.

Reed-Muller codes play a prominent role in complexity theory due to their locality features. Reed-Muller codes are locally decodable/correctable, and are also list-decodable and locally list-decodable up to the Johnson radius [PW04, STV99, BK09]. Moreover, Reed-Muller codes are testable, even robustly [RS96, ALM+98, FS95, AS03, RS97, MR06, AKK+05, KR06, JPRZ09, BKS+10, HSS11].

Note that if we want a family of Reed-Muller codes with positive distance $\delta > 0$, we need the degree $d = (1-\delta)q$. The rate is therefore roughly $\frac{(1-\delta)^m}{m!} < \frac{1}{m!}$. In particular, the rate never exceeds $\frac{1}{2}$. The multiplicity codes of [KSY14] were the first locally correctable codes with rate close to 1. The highlight of our work is the construction of codes with the same distance as that of the Reed-Muller code and same locality features (decodability, correctability, list-decodability, and testability), but with rate close to 1.

Chapter 4

Affine-invariance

4.1 Affine-invariant Codes

In [KS08], Kaufman and Sudan examine the role of symmetry in algebraic property testing (testing of linear codes). The type of symmetry they focus on is *affine-invariance*. Viewing codewords as functions $f : \mathbb{F}_Q^m \to \mathbb{F}_q$, where $\mathbb{F}_Q$ is an extension field of $\mathbb{F}_q$, one can permute the symbols of f by applying a permutation $\pi : \mathbb{F}_Q^m \to \mathbb{F}_Q^m$ to the domain, resulting in a new word $f \circ \pi$. Affine-invariance is simply the property of being closed under applying affine permutations to the domain, i.e. if f is a codeword, then $f \circ A$ is a codeword for any affine permutation A.

Definition 4.1.1 (Affine-invariance)**.** A code $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ is *affine-invariant* if $f \circ A \in \mathcal{C}$ whenever $f \in \mathcal{C}$ and $A : \mathbb{F}_q^m \to \mathbb{F}_q^m$ is an affine permutation.

Affine-invariance appears to be the right abstraction of the low-degree property. In fact, when $Q = q$ is prime, then the only affine-invariant codes are the Reed-Muller codes [KS08]. However, when q is a prime power or if Q is a power of q, then there is a richer collection of affine-invariant codes. Affine-invariant codes are particularly appealing because they possess rich structure. The main structural feature of affine-invariant codes we will use is that they are spanned by monomials.

Definition 4.1.2 (Degree set)**.** A code $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ has a *degree set* $\mathrm{Deg}(\mathcal{C}) \subseteq [\![q]\!]^m$ if $\mathcal{C} = \{f : \mathbb{F}_q^m \to \mathbb{F}_q \mid \mathrm{supp}(f) \subseteq \mathrm{Deg}(\mathcal{C})\}$. The degree set $\mathrm{Deg}(\mathcal{C})$ is *p-shadow-closed* if, whenever $\mathbf{d} \in \mathrm{Deg}(\mathcal{C})$ and $\mathbf{e} \leq_p \mathbf{d}$, we have $\mathbf{e} \in \mathrm{Deg}(\mathcal{C})$.

Proposition 4.1.3 ([KS08])**.** *If $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ is linear affine-invariant, where $\mathbb{F}_q$ has characteristic p, then $\mathcal{C}$ has a p-shadow-closed degree set.*

Proposition 4.1.4 ([BGM+11a])**.** *If $\mathcal{C} \subseteq \{\mathbb{F}_{q^m} \to \mathbb{F}_q\}$ is a $\mathbb{F}_q$-linear affine-invariant code, then $\dim_{\mathbb{F}_q}(\mathcal{C}) = |\mathrm{Deg}(\mathcal{C})|$.*

4.2 Equivalence of Invariance under Affine Transformations and Permutations

In their work initiating the study of the testability of affine-invariant properties (codes), Kaufman and Sudan [KS07] studied properties closed under general affine transformations and not just permutations. While affine transformations are nicer to work with when available, they are not mathematical elegant (they do not form a group under composition). Furthermore in the case of codes they also do not preserve the code — they only show that every codeword stays in the code after the transformation. Among other negative features affine transformations do not even preserve the weight of non-zero codewords, which can lead to some rude surprises. Here we patch the gap by showing that families closed under affine permutations are also closed under affine transformations. So one can assume the latter, without restricting the class of properties under consideration. We note that such a statement was proved in [BGM+11b] for the case of univariate functions. Unfortunately their proof does not extend to the multivariate setting and forces us to rework many steps from [KS08].

Theorem 4.2.1. *If $\mathcal{C} \subseteq \{\mathbb{F}_Q^m \to \mathbb{F}_q\}$ is an $\mathbb{F}_q$-linear code invariant under affine permutations, then $\mathcal{C}$ is invariant under all affine transformations.*

The central lemma (Lemma 4.2.2) that we prove is that every non-trivial function can be split into more basic ones. This leads to a proof of Theorem 4.2.1 fairly easily.

We first start with the notion of a basic function. For $Q = q^n$, let $\mathrm{Tr} : \mathbb{F}_Q \to \mathbb{F}_q$ denote the *trace* function $\mathrm{Tr}(x) = x + x^q + \cdots + x^{q^{n-1}}$. We say that $f : \mathbb{F}_Q^m \to \mathbb{F}_q$ is a *basic* function if $f(\mathbf{X}) = \mathrm{Tr}(\lambda \mathbf{X}^{\mathbf{d}})$ for some $\mathbf{d} \in [\![Q]\!]^m$. For $\mathcal{C} \subseteq \{\mathbb{F}_Q^m \to \mathbb{F}_q\}$ and $f \in \mathcal{C}$ we say f can be *split* (in $\mathcal{C}$) if there exist functions $g, h \in \mathcal{C}$ such that $f = g + h$ and $\mathrm{supp}(g), \mathrm{supp}(h) \subsetneq \mathrm{supp}(f)$.

Lemma 4.2.2. *If $\mathcal{C} \subseteq \{\mathbb{F}_Q^m \to \mathbb{F}_q\}$ is an $\mathbb{F}_q$-linear code invariant under affine permutations, then for every function $f \in \mathcal{C}$, f is either basic or f can be split.*

We first prove Theorem 4.2.1 from Lemma 4.2.2.

Proof of Theorem 4.2.1. First we assert that it suffices to prove that for every function $f \in \mathcal{C}$ the function $\tilde{f} = f(X_1, \ldots, X_{m-1}, 0)$ is also in $\mathcal{C}$. To see this, consider $f \in \mathcal{C}$ and $A : \mathbb{F}_Q^m \to \mathbb{F}_Q^m$ which is not a permutation. Then there exists affine permutations $B, C : \mathbb{F}_Q^m \to \mathbb{F}_Q^m$ such that $A(\mathbf{X}) = B(C(\mathbf{X})_1, \ldots, C(\mathbf{X})_r, 0, \ldots, 0)$ where $r < m$ is the dimension of the image of A. By closure under affine permutations, it follows $f \circ B \in \mathcal{C}$. Applying the assertion above $m-r$ times we have that $f'(\mathbf{X}) \triangleq (f \circ B)(X_1, \ldots, X_r, 0, \ldots, 0)$ is also in $\mathcal{C}$. Finally $f \circ A = f' \circ C$ is also in $\mathcal{C}$. So we turn to proving that for every $f \in \mathcal{C}$ the function $\tilde{f} = f(X_1, \ldots, X_{m-1}, 0)$ is also in $\mathcal{C}$.

Let $f(\mathbf{X}) = \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{X}^{\mathbf{d}}$. Notice $\tilde{f}(\mathbf{X}) = \sum_{\mathbf{d} | d_m = 0} c_{\mathbf{d}} \mathbf{X}^{\mathbf{d}}$. Writing $f = \tilde{f} + f_1$, we use Lemma 4.2.2 to split f till we express it as a sum of basic functions $f = \sum_{i=1}^{N} b_i$, where each b_i is a basic function in $\mathcal{C}$. Note that for every b_i, we have $\mathrm{supp}(b_i) \subseteq \mathrm{supp}(\tilde{f})$ or $\mathrm{supp}(b_i) \subseteq \mathrm{supp}(f_1)$ (since the trace preserves $d_m = 0$). By reordering the b_i's assume the first M b_i's have their support in the support of $\tilde{f}$. Then we have $\tilde{f} = \sum_{i=1}^{M} b_i \in \mathcal{C}$. □

We thus turn to the proof of Lemma 4.2.2. We prove the lemma in a sequence of cases, based on the kind of monomials that f has in its support.

We say that $\mathbf{d}$ and $\mathbf{e}$ are equivalent (modulo q), denoted $\mathbf{d} \equiv_q \mathbf{e}$ if there exists a j such that for every i, $d_i = q^j e_i \bmod^* Q$. The following proposition is immediate from previous works (see, for example, [BGM+11b]). We include a proof for completeness.

Proposition 4.2.3. *If every pair* $\mathbf{d}, \mathbf{e}$ *in the support of* $f : \mathbb{F}_Q^m \to \mathbb{F}_q$ *are equivalent, then* f *is a basic function.*

Proof. We first note that since the $\mathrm{Tr} : \mathbb{F}_Q \to \mathbb{F}_q$ is a (Q/q)-to-one function, we have in particular that for every $\beta \in \mathbb{F}_q$ there is an $\alpha \in \mathbb{F}_Q$ such that $\mathrm{Tr}(\alpha) = \beta$. As an immediate consequence we have that every function $f : \mathbb{F}_Q^m \to \mathbb{F}_q$ can be expressed $\mathrm{Tr} \circ g$ where $g : \mathbb{F}_Q^m \to \mathbb{F}_Q$. Finally we note that we can view g as an element of $\mathbb{F}_Q[\mathbf{X}]$, to conclude that $f = \mathrm{Tr} \circ g$ for some polynomial g.

Now fix $f : \mathbb{F}_Q^m \to \mathbb{F}_q$ all of whose monomials are equivalent. By the above we can express $f = \mathrm{Tr} \circ g$ for some polynomial g. By inspection we can conclude that all monomials in the support of g are equivalent to the monomials in the support of f. Finally, using the fact that $\mathrm{Tr}(\alpha \mathbf{X}^{\mathbf{d}}) = \mathrm{Tr}(\alpha^q \mathbf{X}^{q\mathbf{d} \bmod^* Q})$ we can assume that g is supported on a single monomial and so $f = \mathrm{Tr}(\lambda \mathbf{X}^{\mathbf{d}})$ for some $\lambda \in \mathbb{F}_Q$. □

So it suffices to show that every function that contains non-equivalent degrees in its support can be split. We first prove that functions with "non-weakly-equivalent" monomials can be split.

We say that $\mathbf{d}$ and $\mathbf{e}$ are weakly equivalent if there exists a j such that for every i, $d_i = q^j e_i \pmod{Q-1}$.

Lemma 4.2.4. *If* $\mathcal{F} \subseteq \{\mathbb{F}_Q^m \to \mathbb{F}_q\}$ *is an* $\mathbb{F}_q$*-linear code invariant under affine permutations and* $f \in \mathcal{C}$ *contains a pair of non-weakly equivalent monomials in its support, then* f *can be split.*

Proof. Let $\mathbf{d}$ and $\mathbf{e}$ be two non weakly-equivalent monomials in the support of f. Fix j and consider the function $f_j(\mathbf{X}) = \sum_{\mathbf{a} \in (\mathbb{F}_Q^*)^m} \prod a_i^{-q^j d_i} f(a_1 X_1, \ldots, a_m X_m)$. We claim that (1) the support of f_j is a subset of the support of f, (2) $q^j \mathbf{d}$ is in the support of f_j, (3) $\mathbf{d}$ is in the support of f_j only if for every i $f_i = q^j d_i \pmod{Q-1}$ and in particular (4) $\mathbf{e}$ is not in the support of f_j.

Now let $b = b(\mathbf{d})$ be the smallest positive integer such that $q^b d_i = d_i \bmod^* Q$ for every i. Now consider the function $g = \sum_{j=0}^{b-1} f_j$. We have that $g \in \mathcal{C}$ since it is an $\mathbb{F}_q$-linear

combination of linear transforms of functions in $\mathcal{C}$. By the claims about the f_j's we also have that $\mathbf{d}$ is in the support of g, the support of g is contained in the support of f and $\mathbf{e}$ is not in the support of f. Expressing $f = g + (f - g)$ we now have that f can be split. $\square$

The remaining cases are those where some of coordinates of $\mathbf{d}$ are zero or $Q-1$ for every $\mathbf{d}$ in the support of f. We deal with a special case of such functions next.

Lemma 4.2.5. *Let $\mathcal{C}$ be a linear affine-invariant code. Let $f \in \mathcal{C}$ be given by $f(\mathbf{X}, \mathbf{Y}) = \mathrm{Tr}(\mathbf{Y}^{\mathbf{d}} p(\mathbf{X}))$ where every variable in $p(\mathbf{X})$ has degree in $\{0, Q-1\}$ in every monomial, and $\mathbf{d}$ is arbitrary. Further, let degree of $p(\mathbf{X})$ be $a(Q-1)$. Then for every $0 \leq b \leq a$ and for every $\lambda \in \mathbb{F}_Q$, the function $(X_1 \cdots X_b)^{Q-1} \mathrm{Tr}(\lambda \mathbf{Y}^{\mathbf{d}}) \in \mathcal{C}$.*

Note that in particular the lemma above implies that such f's can be split into basic functions.

Proof. We prove the lemma by a triple induction, first on a, then on b, and then on the number of monomials in p. The base case is $a = 0$ and that is trivial. So we consider general $a > 0$.

First we consider the case $b < a$. Assume without loss of generality that the monomial $(X_1 \cdots X_a)^{Q-1}$ is in the support of p and write $p = p_0 + X_1^{Q-1} p_1$ where p_0, p_1 do not depend on X_1. Note that $p_1 \neq 0$ and $\deg(p_1) = (a-1)(Q-1)$. We will prove that $-\mathrm{Tr}(\mathbf{Y}^{\mathbf{d}} p_1(\mathbf{X})) \in \mathcal{C}$ and this will enable us to apply the inductive hypothesis to p_1. Let $g(\mathbf{X}, \mathbf{Y}) = \sum_{\beta \in \mathbb{F}_Q} f(X_1 + \beta, X_2, \ldots, X_m, \mathbf{Y})$. By construction $g \in \mathcal{C}$. By linearity of the Trace we have

$$g = \mathrm{Tr}\left(\mathbf{Y}^{\mathbf{d}} \left(\sum_{\beta \in \mathbb{F}_Q} p_0 + (X_1 + \beta)^{Q-1} p_1\right)\right) = \mathrm{Tr}(\mathbf{Y}^{\mathbf{d}}(-p_1(\mathbf{X}))),$$

where the second equality follows from the fact that $\sum_{\beta \in \mathbb{F}_Q} (z + \beta)^{Q-1} = -1$. Thus we can now use induction to claim $(X_1 \ldots X_b)^{Q-1} \mathrm{Tr}(\lambda \mathbf{Y}^{\mathbf{d}}) \in \mathcal{C}$.

Finally we consider the case $b = a$. Now note that since the case $b < a$ is known, we can assume without loss of generality that p is homogeneous (else we can subtract off the lower degree terms). Now if $a = m$ there is nothing to be proved since p is just a single monomial. So

assume $a < m$. Also if p has only one monomial then there is nothing to be proved, so assume p has at least two monomials. In particular assume p is supported on some monomial that depends on X_1 and some monomial that does not depend on X_1. Furthermore, assume without loss of generality that a monomial depending on X_1 does not depend on X_2. Write $p = X_1^{Q-1}p_1 + X_2^{Q-1}p_2 + (X_1X_2)^{Q-1}p_3 + p_4$ where the p_i's don't depend on X_1 or X_2. By assumption on the monomials of p we have that $p_1 \neq 0$ and at least one of $p_2, p_3, p_4 \neq 0$. Now consider the affine transform A that sends X_1 to $X_1 + X_2$ and preserves all other X_i's. We have $g = f \circ A = \mathrm{Tr}\left(\mathbf{Y}^{\mathbf{d}}(X_1^{Q-1}p_1 + X_2^{Q-1}(p_1 + p_2) + (X_1X_2)^{Q-1}p_3 + p_4 + r)\right)$ where the X_1-degree of every monomial in r is in $[Q-2]$. Now consider $g'(\mathbf{x}, \mathbf{y}) = \sum_{\alpha \in \mathbb{F}_Q^*} g(\alpha x_1, x_2, \ldots, x_m, \mathbf{y})$. The terms of r vanish in g' leaving

$$g' = -(f \circ A - r) = \mathrm{Tr}\left(\mathbf{Y}^{\mathbf{d}}\left(-X_1^{Q-1}p_1 - X_2^{Q-1}(p_1 + p_2) - (X_1X_2)^{Q-1}p_3 - p_4\right)\right).$$

Finally we consider the function $\tilde{g} = f + g' = \mathrm{Tr}(\mathbf{Y}^{\mathbf{d}}(-X_2^{Q-1}p_1))$ which is a function in $\mathcal{C}$ of degree $a(Q-1)$ supported on a smaller number of monomials than f, so by applying the inductive hypothesis to $\tilde{g}$ we have that $\mathcal{C}$ contains the monomial $(X_1 \cdots X_a)^{Q-1}$. □

The following lemma converts the above into the final piece needed to prove Lemma 4.2.2.

Lemma 4.2.6. *If $\mathcal{F} \subseteq \{\mathbb{F}_Q^m \to \mathbb{F}_q\}$ is an $\mathbb{F}_q$-linear code invariant under affine permutations and all monomials in $f \in \mathcal{C}$ are weakly equivalent, then f can be split.*

Proof. First we describe the structure of a function $f : \mathbb{F}_Q^m \to \mathbb{F}_q$ that consists only of weakly equivalent monomials. First we note that the m variables can be separated into those in which every monomial has degree in $[Q-2]$ and those in which every monomial has degree in $\{0, Q-1\}$ (since every monomial is weakly equivalent). Let us denote by $\mathbf{X}$ the variables in which the monomials of f have degree in $\{0, Q-1\}$ and $\mathbf{Y}$ be the remaining variables. Now consider some monomial of the form $M = c\mathbf{X}^{\mathbf{e}}\mathbf{Y}^{\mathbf{d}}$ in f. Since f maps to $\mathbb{F}_q$ we must have that the coefficient of $(\mathbf{X}^{\mathbf{e}}\mathbf{Y}^{\mathbf{d}})^{q^j}$ is c^{q^j}. Furthermore, we have every other monomial M' in the support of f is of the form $c'\mathbf{Y}^{q^j\mathbf{d}}\mathbf{X}x^{\mathbf{e}'}$. Thus f can be written as $\mathrm{Tr}(\mathbf{Y}^{\mathbf{d}}p(\mathbf{X}))$ where $p(X_1, \ldots, X_m) = \tilde{p}(X_1^{Q-1}, \ldots, X_m^{Q-1})$. But, by Lemma 4.2.5, such an f can be split. □

Proof of Lemma 4.2.2. If f contains a pair of non-weakly equivalent monomials then f can be split by Lemma 4.2.4. If not, then f is either basic or, by Lemma 4.2.6 is can be split. □

We also prove an easy consequence of Lemma 4.2.2.

Lemma 4.2.7. *Let $\mathcal{C} \subseteq \{\mathbb{F}_Q^m \to \mathbb{F}_q\}$ be affine invariant. If $\mathbf{d} \in \mathrm{Deg}(\mathcal{C})$, then $\mathrm{Tr}(\lambda \mathbf{X}^{\mathbf{d}}) \in \mathcal{C}$ for all $\lambda \in \mathbb{F}_Q$.*

Proof. We first claim that Lemma 4.2.2 implies that there exists $\beta \in \mathbb{F}_Q$ such that $\mathrm{Tr}(\beta \mathbf{X}^{\mathbf{d}})$ is a non-zero function in $\mathcal{C}$. To verify this, consider a "minimal" function (supported on fewest monomials) $f \in \mathcal{C}$ with $\mathbf{d} \in \mathrm{supp}(f)$. Since f can't be split in $\mathcal{C}$ (by minimality), by Lemma 4.2.2 f must be basic and so equals (by definition of being basic) $\mathrm{Tr}(\beta \mathbf{X}^{\mathbf{d}})$.

Now let $b = b(\mathbf{d})$ be the smallest positive integer such that $q^b \mathbf{d} \bmod^* Q = \mathbf{d}$. If $Q = q^n$, note that b divides n and so one can write $\mathrm{Tr} : \mathbb{F}_Q \to \mathbb{F}_q$ as $\mathrm{Tr}_1 \circ \mathrm{Tr}_2$ where $\mathrm{Tr}_1 : \mathbb{F}_{q^b} \to \mathbb{F}_q$ is the function $\mathrm{Tr}_1(z) = z + z^q + \cdots + z^{q^{b-1}}$ and $\mathrm{Tr}_2 : \mathbb{F}_Q \to \mathbb{F}_{q^b}$ is the function $\mathrm{Tr}_2(z) = z + z^{q^b} + \cdots + z^{Q/q^b}$. (Both Tr_1 and Tr_2 are trace functions mapping the domain to the range.) It follows that $\mathrm{Tr}(\beta \mathbf{X}^{\mathbf{d}}) = \mathrm{Tr}_1(\mathrm{Tr}_2(\beta)\mathbf{X}^{\mathbf{d}})$.

We first claim that $\mathrm{Tr}_1(\tau \mathbf{X}^{\mathbf{d}}) \in \mathcal{C}$ for every $\tau \in \mathbb{F}_{q^b}$. Let $S = \{\sum_{\alpha \in (\mathbb{F}_Q^*)^m} a_\alpha \cdot \alpha^{\mathbf{d}} \mid a_\alpha \in \mathbb{F}_q\}$. We note that by linearity and affine-invariance of $\mathcal{C}$, we have that $\mathrm{Tr}_1(\mathrm{Tr}_2(\beta) \cdot \eta \mathbf{X}^{\mathbf{d}}) \in \mathcal{C}$ for every $\eta \in S$. By definition S is closed under addition and multiplication and so is a subfield of $\mathbb{F}_Q$. In fact, since every $\eta \in S$ satisfies $\eta^{q^b} = \eta$ (which follows from the fact that $\alpha^{\mathbf{d}} = \alpha^{q^b \mathbf{d}}$), we have that $S \subseteq \mathbb{F}_{q^b}$. It remains to show $S = \mathbb{F}_{q^b}$. Suppose it is a strict subfield of size q^c for $c < b$. Consider γ^{d_i} for $\gamma \in \mathbb{F}_Q$ and $i \in [m]$. Since $\gamma^{d_i} \in S$, we have that $\gamma^{d_i q^c} = \gamma^{d_i}$ for every $\gamma \in \mathbb{F}_Q$ and so we get $X_i^{q^c d_i} = X_i \pmod{X_i^Q - X_i}$. We conclude that $\mathbf{X}^{q^c \mathbf{d}} = \mathbf{X}^{\mathbf{d}} \pmod{\mathbf{X}^Q - \mathbf{X}}$ which contradicts the minimality of $b = b(\mathbf{d})$. We conclude that $S = \mathbb{F}_{q^b}$. Since $\mathrm{Tr}_2(\beta) \in \mathbb{F}_{q^b}^*$, we conclude that the set of coefficients τ such that $\mathrm{Tr}_1(\tau \mathbf{X}^{\mathbf{d}}) \in \mathcal{C}$ is all of $\mathbb{F}_{q^b}$ as desired.

Finally consider any $\lambda \in \mathbb{F}_Q$. since $\mathrm{Tr}_2(\lambda) \in \mathbb{F}_{q^b}$, we have that $\mathrm{Tr}_1(\mathrm{Tr}_2(\lambda)\mathbf{X}^{\mathbf{d}}) \in \mathcal{C}$ (from the previous paragraph), and so $\mathrm{Tr}(\lambda \mathbf{X}^{\mathbf{d}}) = \mathrm{Tr}_1(\mathrm{Tr}_2(\lambda)\mathbf{X}^{\mathbf{d}}) \in \mathcal{C}$ □

Chapter 5

Lifting Codes

5.1 The Lift Operator

First defined and used in [BSMSS11] to prove the existence of certain codes that are *not* locally testable, the lift operator takes short codes and creates longer codes from them. The original lift operator took codes defined over the domain $\mathbb{F}_q$ and "lifted" them to codes defined over the domain $\mathbb{F}_{q^m}$. We generalize the definition, allowing one to lift codes over $\mathbb{F}_q^t$ to codes over $\mathbb{F}_q^m$, for any $m \geq t$ (in particular, t does not need to divide m).

Definition 5.1.1 (Lift)**.** Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be a linear affine-invariant code. Let $m \geq t$ be an integer. The *m-dimensional lift of $\mathcal{C}$*, denoted by $\mathcal{C}^{t \nearrow m}$, is the code

$$\mathcal{C}^{t \nearrow m} \triangleq \left\{ f : \mathbb{F}_q^m \to \mathbb{F}_q \;\middle|\; f|_A \in \mathcal{C} \text{ for every } t\text{-dimensional affine subspace } A \subseteq \mathbb{F}_q^m \right\}.$$

Proposition 5.1.2. *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant, and let $m \geq t$. If $f \in \mathcal{C}^{t \nearrow m}$ and $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ is an affine map, then $f \circ A \in \mathcal{C}$.*

Proof. By Proposition B.1.1, there exists an invertible affine map $A'' : \mathbb{F}_q^t \to \mathbb{F}_q^m$ and a linear map $A' : \mathbb{F}_q^t \to \mathbb{F}_q^t$ such that $A = A'' \circ A'$. By the definition of lift, $g \triangleq f \circ A'' \in \mathcal{C}$. Since $\mathcal{A}$

is affine-invariant, it follows from Theorem 4.2.1 that $g \circ A' \in \mathcal{C}$. Therefore,

$$f \circ A = f \circ (A'' \circ A') = (f \circ A'') \circ A' = g \circ A' \in \mathcal{C}.$$

□

5.2 Algebraic and Combinatorial Properties

We begin by exploring the algebraic and combinatorial (as opposed to algorithmic) properties of codes lifted from linear affine-invariant codes. First, we show that the lifting operator is a natural algebraic operation, preserving linearity and affine-invariance and composing well. We then present the degree set of a lifted code in terms of the degree set of the base code. Finally, we conclude by showing that lifting preserves distance.

5.2.1 Algebraic Properties

Proposition 5.2.1. *Let $t \leq m$ be integers. If $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ is linear affine-invariant, then $\mathcal{C}^{t \nearrow m}$ is linear affine-invariant.*

Proof. First, we show that $\mathcal{C}^{t \nearrow m}$ is linear. Let $f, g \in \mathcal{C}^{t \nearrow m}$ and let $\alpha \in \mathbb{F}_q$. If $A \subseteq \mathbb{F}_q^m$ is a t-dimensional affine subspace, then, since $\mathcal{C}$ is linear,

$$(\alpha f + g)|_A = \alpha \cdot (f|_A) + (g|_A) \in \mathcal{C}$$

Next, we show that $\mathcal{C}^{t \nearrow m}$ is affine-invariant. Let $f(\mathbf{X}) \in \mathcal{C}^{t \nearrow m}$, and let $\mathbf{M} \in \mathbb{F}_q^{m \times m}$ and $\mathbf{c} \in \mathbb{F}_q^m$. Let $g(\mathbf{X}) = f(\mathbf{MX} + \mathbf{c})$. Let $\mathbf{Y} = (Y_1, \ldots, Y_t)$, let $\mathbf{A} \in \mathbb{F}_q^{m \times t}$ and $\mathbf{b} \in \mathbb{F}_q^m$. Then

$$g(\mathbf{AX} + \mathbf{b}) = f(\mathbf{M}(\mathbf{AX} + \mathbf{b}) + \mathbf{c}) = f((\mathbf{MA})\mathbf{X} + (\mathbf{Mb} + \mathbf{c})) \in \mathcal{C}.$$

Since $\mathbf{A}, \mathbf{b}$ were arbitrary, this implies that $g \in \mathcal{C}^{t \nearrow m}$. Since $\mathbf{M}, \mathbf{c}$ were arbitrary, this implies that $\mathcal{C}^{t \nearrow m}$ is affine-invariant. □

Proposition 5.2.2 (Composition of lift)**.** *Let $t \leq m \leq n$ be integers. Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant. Then*

$$\mathcal{C}^{t \nearrow n} = \left(\mathcal{C}^{t \nearrow m}\right)^{m \nearrow n}.$$

Proof. This follows immediately from the fact that choosing a t-dimensional affine subspace $A \subseteq \mathbb{F}_q^n$ is equivalent to first choosing an m-dimensional affine subspace $B \subseteq \mathbb{F}_q^n$ and then choosing a t-dimensional affine subspace $A \subseteq B$. □

Proposition 5.2.3 (Degree set of lift)**.** *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant with degree set $\mathrm{Deg}(\mathcal{C})$. If $m \geq t$ is an integer, then the lift $\mathcal{C}^{t \nearrow m}$ has degree set*

$$\mathrm{Deg}(\mathcal{C}^{t \nearrow m}) = \{\mathbf{d} \in [\![q]\!]^m \mid \mathbf{E} \leq_p \mathbf{d} \implies (\|\mathbf{E}_{*1}\|, \ldots, \|\mathbf{E}_{*t}\|) \bmod^* q \in \mathrm{Deg}(\mathcal{C})\}$$

where $\mathbf{E}$ has rows $[m]$ and columns $[\![t+1]\!]$.

Proof. Let $f \in \mathcal{C}^{t \nearrow m}$. Let $D = \mathrm{Deg}(\mathcal{C}^{t \nearrow m})$. Let $\mathbf{X} = (X_1, \ldots, X_m)$ and let $\mathbf{Y} = (Y_1, \ldots, Y_t)$. Write $f(\mathbf{X}) = \sum_{\mathbf{d} \in D} f_{\mathbf{d}} \cdot \mathbf{X}^{\mathbf{d}}$. For any matrix $\mathbf{A} \in \mathbb{F}_q^{m \times t}$ and $\mathbf{b} \in \mathbb{F}_q^m$, it follows from Proposition A.1.4 that

$$f(\mathbf{AY} + \mathbf{b}) = \sum_{\mathbf{d} \in D} f_{\mathbf{d}} \cdot \sum_{\mathbf{E} \leq_p \mathbf{d}} \binom{\mathbf{d}}{\mathbf{E}} \prod_{i=1}^{m} \left(b_i^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}} \right) \cdot \prod_{j=1}^{t} Y_j^{\|\mathbf{E}_{*j}\|}$$

Since $f(\mathbf{AY} + \mathbf{b}) \in \mathcal{C}$ for any $\mathbf{A}, \mathbf{b}$, and $\mathcal{C}$ is affine-invariant, it follows by Proposition 4.1.3 that for every $\mathbf{d} \in D$ and $\mathbf{E} \leq_p \mathbf{d}$, with columns $[\![t+1]\!]$, it holds that the monomial $\prod_{j=1}^{t} Y_j^{\|\mathbf{E}_{*j}\|} \in \mathcal{C}$, hence $(\|\mathbf{E}_{*1}\|, \ldots, \|\mathbf{E}_{*t}\|) \bmod^* q \in \mathrm{Deg}(\mathcal{C})$. Conversely, if $\mathbf{d}$ satisfies that for every $\mathbf{E} \leq_p \mathbf{d}$ with columns $[\![t+1]\!]$ that $(\|\mathbf{E}_{*1}\|, \ldots, \|\mathbf{E}_{*t}\|) \bmod^* q \in \mathrm{Deg}(\mathcal{C})$, then it is easy to see that $\mathbf{d} \in D$ by considering $f(\mathbf{X}) = \mathbf{X}^{\mathbf{d}}$. □

5.2.2 Distance of Lifted Codes

Proposition 5.2.4 (Distance of lift)**.** *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant, and let $m \geq t$. The following hold:*

1. $\delta(\mathcal{C}^{t\nearrow m}) \leq \delta(\mathcal{C})$;
2. $\delta(\mathcal{C}^{t\nearrow m}) \geq \delta(\mathcal{C}) - q^{-t}$;
3. *if $q \in \{2,3\}$ and $\delta(\mathcal{C}) > q^{-t}$, then $\delta(\mathcal{C}^{t\nearrow m}) = \delta(\mathcal{C})$.*

We prove several lemmas, which in turn will help us prove Proposition 5.2.4. The first lemma proves Proposition 5.2.4 (1).

Lemma 5.2.5. *If $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ is linear affine-invariant, and $m \geq t$, then $\delta(\mathcal{C}^{t\nearrow m}) \leq \delta(\mathcal{C})$.*

Proof. The case $m = t$ is trivial, so assume $m > t$. By Proposition 5.2.2 and induction, it suffices to consider the case $m = t+1$. Let $f \in \mathcal{C}$ and let $\delta \triangleq \delta(f, 0)$. Let $\mathbf{X} = (X_1, \ldots, X_t)$. Consider the function $g : \mathbb{F}_q^{t+1} \to \mathbb{F}_q$ defined by $g(\mathbf{X}, X_{t+1}) = f(\mathbf{X})$. Clearly, we have $\delta(g, 0) = \delta$. We claim that $g \in \mathcal{C}^{t\nearrow(t+1)}$. Let $A : \mathbb{F}_q^t \to \mathbb{F}_q^{t+1}$ be an affine map. Let $A' : \mathbb{F}_q^t \to \mathbb{F}_q^t$ be the affine map given by the projection of A onto its first t coordinates, i.e. $A'(\mathbf{X}) = (A(\mathbf{X})_1, \ldots, A(\mathbf{X})_t)$. Since $f \in \mathcal{C}$ and $\mathcal{C}$ is affine-invariant, by Theorem 4.2.1 we have $f \circ A' \in \mathcal{C}$. Therefore, $g \circ A = f \circ A' \in \mathcal{C}$. Since A was arbitrary, this shows that $g \in \mathcal{C}^{t\nearrow(t+1)}$. □

The next lemma proves Proposition 5.2.4 (2), and uses a simple probabilistic argument.

Lemma 5.2.6. *If $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ is linear affine-invariant, and $m \geq t$, then $\delta(\mathcal{C}^{t\nearrow m}) \geq \delta(\mathcal{C}) - q^{-t}$.*

Proof. Let $f, g \in \mathcal{C}^{t\nearrow m}$ be distinct. Let $\mathbf{x} \in \mathbb{F}_q^m$ such that $f(\mathbf{x}) \neq g(\mathbf{x})$. Let $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ be a random affine map such that $A(\mathbf{0}) = \mathbf{x}$, so that $f \circ A, g \circ A \in \mathcal{C}$ (by Proposition 5.1.2) and are distinct. Then

$$\begin{aligned} \delta(\mathcal{C}) &\leq \mathbb{E}_A\left[\delta\left(f \circ A, g \circ A\right)\right] && (5.1) \\ \text{(Proposition B.1.3)} &\leq \delta(f, g) + q^{-t}. && (5.2) \end{aligned}$$

□

The next two lemmas will help in proving Proposition 5.2.4, for the cases $q = 2$ and $q = 3$ respectively.

Lemma 5.2.7. *For all $m \geq 2$, if $\delta > \frac{1}{2^{m-1}}$ and $f : \mathbb{F}_2^m \to \mathbb{F}_2$ such that $0 < \delta(f, 0) < \delta$, then there exists a hyperplane $H \subset \mathbb{F}_2^m$ such that $0 < \delta(f|_H, 0) < \delta$.*

Proof. We proceed by induction on m. The base case $m = 2$ is straightforward to verify.

Now suppose $m > 2$ and our assertion holds for $m-1$. Let H_0, H_1 be the affine subspaces given by $X_m = 0$ and $X_m = 1$, respectively. For $i \in \{0, 1\}$, let $\delta_i \triangleq \delta(f|_{H_i}, 0)$. Note that $\delta > \delta(f, 0) = (\delta_0 + \delta_1)/2$. If both $\delta_0, \delta_1 > 0$, then by averaging we have $0 < \delta_i < \delta$, for some $i \in \{0, 1\}$, and so $H = H_i$ does the job. Otherwise, suppose without loss of generality that $\delta_1 = 0$. Note that $0 < \delta_0 < 2\delta$ and $2\delta > \frac{1}{2^{m-2}}$. Thus, by the induction hypothesis, there exists an $(m-2)$-dimensional affine subspace $H_0' \subseteq H_0$ such that $0 < \delta(f|_{H_0'}, 0) < 2\delta$. Let $H_1' \triangleq H_0' + \mathbf{e}_m$, where $\mathbf{e}_m$ is the m-th standard basis vector. Note that $\delta(f|_{H_1'}, 0) = 0$. Let $H = H_0' \cup H_1'$. Then H is an $(m-1)$-dimensional subspace of $\mathbb{F}_2^m$ (spanned by H_0' and $\mathbf{e}_m$) such that $0 < \delta(f|_H, 0) = \left(\delta(f|_{H_0'}, 0) + \delta(f|_{H_1'}, 0)\right)/2 < \delta$. □

Lemma 5.2.8. *For all $m \geq 2$, if $f : \mathbb{F}_3^m \to \mathbb{F}_3$ and $\delta(f, 0) \geq \frac{1}{3^{m-1}}$, then there exists a hyperplane $H \subseteq \mathbb{F}_3^m$ such that $0 < \delta(f|_H, 0) \leq \delta(f, 0)$.*

Proof. Let $\delta \triangleq \delta(f, 0)$. We proceed by induction on m.For the base case, $m = 2$, $\delta \geq \frac{1}{3}$. Suppose $f : \mathbb{F}_3^2 \to \mathbb{F}_3$ and for each $i \in \mathbb{F}_3$, let f_i be the restriction of f to the hyperplane defined by $X_2 = i$. If $f_i \neq 0$ for every $i \in \mathbb{F}_3$, then by averaging there is some $i \in \mathbb{F}_3$ for which $0 < \delta(f_i, 0) \leq \delta$. Otherwise, without loss of generality, suppose $f_2 = 0$. Further, without loss of generality, suppose $f_0 \neq 0$ and $f(0, 0) \neq 0$. Now, if $\delta \geq \frac{2}{3}$, then the line $H = \{(x, y) \in \mathbb{F}_2^3 \mid x = 0\}$ does the job, since $0 < \delta(f|_H, 0) \leq \frac{2}{3} \leq \delta$. If $\delta < \frac{2}{3}$, then there must exist some $a, b \in \mathbb{F}_3$ and $c \in \{0, 1\}$ such that $f(a, c) \neq 0$ and $f(b, 1-c) = 0$. Then the line $H = \{(a, c), (b, 1-c), (2b-a, 2)\}$ does the job, since $0 < \delta(f|_H, 0) = \frac{1}{3} \leq \delta$.

Now suppose $m > 2$ and the assertion holds for $m-1$. For $i \in \mathbb{F}_3$, let H_i be the hyperplane cut out by $X_m = i$ and let $\delta_i \triangleq \delta(f|_{H_i}, 0)$ for each $i \in \mathbb{F}_3$. Then $\delta_0 + \delta_1 + \delta_3 = 3\delta$.

If $\delta_i > 0$ for all $i \in \mathbb{F}_3$, then by simple averaging, for some $i \in \mathbb{F}_3$, we have $0 < \delta_i \leq \delta$, so assume without loss of generality that $\delta_2 = 0$ and $\delta_0 \geq \delta_1$.

First, suppose $\delta_0 \geq \frac{1}{3^{m-2}}$. Then, by the inductive hypothesis, there exists an $(m-2)$-dimensional affine subspace $H^{(0)} \subset H_0$ such that $0 < \delta(f|_H, 0) \leq \delta_0$. Suppose $H^{(0)}$ is defined by the linear equations $\sum_{i=1}^m a_i X_i - a_0 = 0$ and $X_m = 0$ for some $(a_0, \ldots, a_m) \in \mathbb{F}_q^{m+1}$. For each $i, j \in \mathbb{F}_3$, let $H_j^{(i)} \subset H_i$ be the affine subspace defined by $\sum_{i=1}^m a_i X_i - a_0 = j$ and $X_m = i$. By averaging, for some $j \in \mathbb{F}_3$, we have $\delta\left(f|_{H_j^{(1)}}, 0\right) \leq \delta_1$. Take $H \triangleq H^{(0)} \cup H_j^{(1)} \cup H_{2i}^{(2)}$. Then

$$0 < \delta(f|_H, 0) = \frac{\delta\big(f|_{H^{(0)}}, 0\big) + \delta\big(f|_{H_j^{(1)}}, 0\big) + \delta\big(f|_{H_{2j}^{(2)}}, 0\big)}{3} \leq \frac{\delta_0 + \delta_1}{3} = \delta.$$

Now, suppose $\delta_0 < \frac{1}{3^{m-2}}$, so $\delta_0, \delta_1 \leq \frac{2}{3^{m-1}}$. There exists an $(m-2)$-dimensional affine subspace $H^{(0)} \subset H_0$ such that $\delta(f|_{H^{(0)}}, 0) = \frac{1}{3^{m-1}}$. To see this, let $\mathbf{a}, \mathbf{b} \in H_0$ such that $f(\mathbf{a}), f(\mathbf{b}) \neq 0$, and suppose $a_k \neq b_k$, for some $k \in [m]$. Then take $H^{(0)}$ to be the subspace defined by $X_k = a_k$ and $X_m = 0$. For $i, j \in \mathbb{F}_3$, let $H_j^{(i)}$ be the $(m-2)$-dimensional subspace defined by $X_k = a_k + j$ and $X_m = i$. Since $\delta_1 \leq \frac{2}{3^{m-2}}$, there is $j \in \mathbb{F}_3$ such that $f|_{H_j^{(1)}} = 0$. Then, taking $H \triangleq H^{(0)} \cup H_j^{(1)} \cup H_{2j}^{(2)}$, we have

$$0 < \delta(f|_H, 0) = \frac{\delta\big(f|_{H^{(0)}}, 0\big) + \delta\big(f|_{H_j^{(1)}}, 0\big) + \delta\big(f|_{H_{2j}^{(2)}}, 0\big)}{3} = \frac{1}{3^m} < \delta.$$

□

Now, we are ready to prove Proposition 5.2.4.

Proof of Proposition 5.2.4. Parts 1 and 2 follow immediately from Lemmas 5.2.5 and 5.2.5, respectively, so we proceed with proving part 3. In light of part 1, it suffices to show that $\delta(\mathcal{C}^{t\nearrow m}) \geq \delta(\mathcal{C})$.

We start with the $q = 2$ case. We proceed by induction on $m - t$. Indeed, the inductive step is straightforward since, by Proposition 5.2.2, we have $\mathcal{C}^{t\nearrow m} = \left(\mathcal{C}^{t\nearrow m-1}\right)^{m-1\nearrow m}$, so by induction $\delta(\mathcal{C}^{t\nearrow m}) \geq \delta(\mathcal{C}^{t\nearrow m-1}) \geq \delta(\mathcal{C})$. The main case is therefore the base case $m = t+1$. Suppose $f \in \mathcal{C}^{t\nearrow t+1}$ such that $0 < \delta(f, 0) < \delta(\mathcal{C})$. By assumption, $\delta(\mathcal{C}) > \frac{1}{2^t}$, so we

may apply Lemma 5.2.7 to conclude that there exists a hyperplane $H \subset \mathbb{F}_2^m$ such that $9 < \delta(f|_H, 0) < \delta(\mathcal{C})$, contradicting the fact that $f|_H \in \mathcal{C}$ (since $f \in \mathcal{C}^{t\nearrow t+1}$).

Finally, we consider the $q = 3$ case. Again, we proceed by induction on $m - t$, and again the main case is the base case $m = t + 1$. Suppose $f \in \mathcal{C}^{t\nearrow t+1}$ such that $0 < \delta(f, 0) < \delta(\mathcal{C})$. If $\delta(f, 0) \geq \frac{1}{3^{m-1}}$, then, by Lemma 5.2.8, there exists a hyperplane $H \subset \mathbb{F}_q^m$ such that $0 < \delta(f|_H, 0) \leq \delta(f, 0) < \delta(\mathcal{C})$, contradicting the fact that $f|_H \in \mathcal{C}$. If $\delta(f, 0) < \frac{1}{3^{m-1}}$, then there are at most two points $\mathbf{a}, \mathbf{b} \in \mathbb{F}_3^m$ such that $f(\mathbf{a}), f(\mathbf{b}) \neq 0$. Let $i \in [m]$ such that $a_i \neq b_i$ and let H be the hyperplane defined by $X_i = a_i$. Then $f|_H$ is nonzero only on $\mathbf{a}$, so $0 < \delta(f|_H) = \frac{1}{3^{m-1}} < \delta(\mathcal{C})$, again contradicting the fact that $f|_H \in \mathcal{C}$. □

5.3 Local Decoding and Correcting

In this section, we explore some of the algorithmic decoding properties of lifted codes. In particular, for codes of distance $\delta > 0$, we give simple algorithms for locally correcting up to $\delta/4$ and then $\delta/2$ fraction errors. We then show that all linear affine-invariant codes — in particular, lifted codes — have explicit interpolating sets, thereby showing that lifted codes are locally decodable as well.

5.3.1 Local Correcting up to 1/4 Distance

The following algorithm is an abstraction of a simple algorithm for locally correcting Reed-Muller codes of distance $\delta > 0$ from $\delta/4$ fraction errors.

Theorem 5.3.1. *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant with distance $\delta \triangleq \delta(\mathcal{C})$, and let $m \geq t$. Then, for every $\epsilon > 0$, the lifted code $\mathcal{C}^{t\nearrow m}$ is $\left(q^t, (\frac{1}{4} - \epsilon)\delta - q^{-t}, \frac{1}{2} - 2\epsilon\right)$-locally correctable.*

Proof. Let $\mathsf{Corr}_{\mathcal{C}}$ be a correction algorithm for $\mathcal{C}$, so that for every $f : \mathbb{F}_q \to \mathbb{F}_q$ that is $\delta/2$-close to some $g \in \mathcal{C}$, $\mathsf{Corr}_{\mathcal{C}}(f) = g$. The following algorithm is a local correction algorithm that achieves the desired parameters.

Local correction algorithm: Oracle access to received word $r : \mathbb{F}_q^m \to \mathbb{F}_q$.

On input $\mathbf{x} \in \mathbb{F}_q^m$:

1. Choose uniform random affine map $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ such that $A(\mathbf{0}) = \mathbf{x}$.
2. Set $f \triangleq r \circ A$.
3. Compute $\widehat{f} \triangleq \mathsf{Corr}_{\mathcal{C}}(f)$.
4. Output $\widehat{f}(\mathbf{0})$.

Analysis: Let $\tau \triangleq (\frac{1}{4} - \epsilon)\delta$. Fix a received word $r : \mathbb{F}_q^m \to \mathbb{F}_q$. Let $c \in \mathcal{C}$ be a codeword such that $\delta(r, c) < \tau - q^{-t}$. Let $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ be a random affine map such that $A(\mathbf{0}) = \mathbf{x}$. By Proposition B.1.3, the expected distance is $\mathbb{E}_A\left[\delta\left(r \circ A, c \circ A\right)\right] \leq \delta(r, c) + q^{-t} < \tau$, so, by Markov's inequality, with probability at least $\frac{1}{2} + 2\epsilon$, we have $\delta(r \circ A, c \circ A) < \frac{\delta}{2}$. For such A, setting $f \triangleq r \circ A$, we have $\widehat{f} = \mathsf{Corr}_{\mathcal{C}}(f) = c \circ A$, hence $\widehat{f}(\mathbf{0}) = c(A(\mathbf{0})) = c(\mathbf{x})$. □

5.3.2 Local Correcting up to 1/2 Distance

The following algorithm can locally correct lifted codes up to half the minimum distance. The basic idea is to decode along multiple lines and weight their opinions based on distance to the base code. This is a direct translation of the elegant line-weight local decoding algorithm for matching-vector codes [BET10] to the Reed-Muller code and lifted codes setting.

Theorem 5.3.2. *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant with distance $\delta \triangleq \delta(\mathcal{C})$, and let $m \geq t$. Then, for every $\epsilon, \eta > 0$, the lifted code $\mathcal{C}^{t \nearrow m}$ is $\left(Q, (\frac{1}{2} - \epsilon)\delta - q^{-t}, \eta\right)$-locally correctable for $Q = O\left(\ln(1/\eta)/\epsilon^2 \cdot q^t\right)$.*

Proof. Let $\mathsf{Corr}_{\mathcal{C}}$ be a correction algorithm for $\mathcal{C}$, so that for every $f : \mathbb{F}_q \to \mathbb{F}_q$ that is $\delta/2$-close to some $g \in \mathcal{C}$, $\mathsf{Corr}_{\mathcal{C}}(f) = g$. The following algorithm is a local correction algorithm that achieves the desired parameters.

Local correction algorithm: Oracle access to received word $r : \mathbb{F}_q^m \to \mathbb{F}_q$.

On input $\mathbf{x} \in \mathbb{F}_q^m$:

1. Let $c = \left\lceil \frac{2}{\epsilon^2} \ln \frac{4}{\eta} \right\rceil$ and choose affine maps $A_1, \ldots, A_c \in \mathbb{F}_q^m$ such that $A_i(\mathbf{0}) = \mathbf{x}$ for each $i \in [c]$ independently and uniformly at random.

2. For each $i \in [c]$:

   (a) Set $r_i \triangleq r \circ A_i$.

   (b) Compute $s_i \triangleq \mathsf{Corr}_{\mathcal{C}}(r_i)$ and $\delta_i \triangleq \delta(r_i, s_i)$.

   (c) Assign the value $s_i(\mathbf{0})$ a weight $W_i \triangleq \max\left(1 - \frac{\delta_i}{\delta/2}, 0\right)$.

3. Set $W \triangleq \sum_{i=1}^{c} W_i$. For every $\alpha \in \mathbb{F}_q$, let $w(\alpha) := \frac{1}{W} \sum_{i : s_i(\mathbf{0}) = \alpha} W_i$. If there is an $\alpha \in \mathbb{F}_q$ with $w(\alpha) > \frac{1}{2}$, output α, otherwise fail.

Analysis: Let $\tau \triangleq (\frac{1}{2} - \epsilon)\delta$. Fix a received word $r : \mathbb{F}_q^m \to \mathbb{F}_q$. Let $c \in \mathcal{C}$ be a codeword such that $\delta(r, c) < \tau - q^{-t}$. The query complexity follows from the fact that the algorithm queries $O(\ln(1/\eta)/\epsilon^2)$ subspaces, each consisting of at most q^t points. Fix an input $\mathbf{x} \in \mathbb{F}_q^m$. We wish to show that, with probability at least $1 - \eta$, the algorithm outputs $c(\mathbf{x})$, i.e. $w(c(\mathbf{x})) > \frac{1}{2}$.

For each affine map $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$, define the following:

$$
\begin{aligned}
\tau_A &\triangleq \delta(r \circ A, c \circ A) \\
s_A &\triangleq \mathsf{Corr}_{\mathcal{C}}(r \circ A) \\
\delta_A &\triangleq \delta(r \circ A, s_A) \\
W_A &\triangleq \max\left(1 - \frac{\delta_A}{\delta/2}, 0\right) \\
X_A &\triangleq \begin{cases} W_A & s_A = c \circ A \\ 0 & s_A \neq c \circ A. \end{cases}
\end{aligned}
$$

Let $p \triangleq \Pr_A[s_A = c \circ A]$. Note that if $s_A = c \circ A$, then $\delta_A = \tau_A$, otherwise $\delta_A \geq \delta - \tau_A$. Hence, if $s_A = c \circ A$, then $W_A \geq 1 - \frac{\tau_A}{\delta/2}$, otherwise $W_A \leq \frac{\tau_A}{\delta/2} - 1$.

Define

$$\begin{aligned}
\tau_{\text{good}} &\triangleq \mathbb{E}_A[\tau_A \mid s_A = c \circ A] \\
\tau_{\text{bad}} &\triangleq \mathbb{E}_A[\tau_A \mid s_A \neq c \circ A] \\
W_{\text{good}} &\triangleq \mathbb{E}_A[W_A \mid s_A = c \circ A] \geq 1 - \frac{\tau_{\text{good}}}{\delta/2} \\
W_{\text{bad}} &\triangleq \mathbb{E}_A[W_A \mid s_A \neq c \circ A] \leq \frac{\tau_{\text{bad}}}{\delta/2} - 1.
\end{aligned}$$

Observe that

$$\begin{aligned}
\mathbb{E}_A[\tau_A] &\leq \frac{1 + (\tau - \frac{1}{q})(q-1)}{q} \leq \tau \\
\mathbb{E}_A[X_A] &= p \cdot W_{\text{good}} \\
\mathbb{E}_A[W_A] &= p \cdot W_{\text{good}} + (1-p) \cdot W_{\text{bad}}.
\end{aligned}$$

We claim that

$$p \cdot W_{\text{good}} \geq (1-p) \cdot W_{\text{bad}} + 2\epsilon. \tag{5.3}$$

To see this, we start from

$$\left(\frac{1}{2} - \epsilon\right)\delta = \tau \geq \mathbb{E}_A[\tau_A] = p \cdot \tau_{\text{good}} + (1-p) \cdot \tau_{\text{bad}}.$$

Dividing by $\delta/2$ yields

$$1 - 2\epsilon \geq p \cdot \frac{\tau_{\text{good}}}{\delta/2} + (1-p) \cdot \frac{\tau_{\text{bad}}}{\delta/2}.$$

Re-writing $1 - 2\epsilon$ on the left-hand side as $p + (1-p) - 2\epsilon$ and re-arranging, we get

$$p \cdot \left(1 - \frac{\tau_{\text{good}}}{\delta/2}\right) \geq (1-p) \cdot \left(\frac{\tau_{\text{bad}}}{\delta/2} - 1\right) + 2\epsilon.$$

The left-hand side is bounded from above by $p \cdot W_{\text{good}}$ while the right-hand side is bounded from below by $(1-p) \cdot W_{\text{bad}} + 2\epsilon$, hence (5.3) follows.

Consider the random affine maps $A_1, \ldots, A_c$ chosen by the algorithm. Note that the X_A

are defined such that A_i contributes weight $\frac{X_{A_i}}{W}$ to $w(c(\mathbf{x}))$, so it suffices to show that, with probability at least $1-\eta$,

$$\frac{\sum_{i=1}^{c} X_{A_i}}{\sum_{i=1}^{c} W_{A_i}} > \frac{1}{2}.$$

Each $X_A, W_A \in [0,1]$, so by Proposition 2.2.3,

$$\begin{aligned}
\Pr\left[\left|\frac{1}{c}\sum_{i=1}^{c} X_{A_i} - \mathbb{E}_A[X_A]\right| > \epsilon/2\right] &\leq 2\exp(-\epsilon^2 c/2) \leq \eta/2 \\
\Pr\left[\left|\frac{1}{c}\sum_{i=1}^{c} W_{A_i} - \mathbb{E}_A[W_A]\right| > \epsilon/2\right] &\leq 2\exp(-\epsilon^2 c/2) \leq \eta/2.
\end{aligned}$$

Therefore, by a union bound, with probability at least $1-\eta$ we have

$$\begin{aligned}
\frac{\sum_{i=1}^{c} X_{A_i}}{\sum_{i=1}^{c} W_{A_i}} &\geq \frac{\mathbb{E}_A[X_A] - \epsilon/2}{\mathbb{E}[W_A] + \epsilon/2} \\
&= \frac{p \cdot W_{\text{good}} - \epsilon/2}{p \cdot W_{\text{good}} + (1-p) \cdot W_{\text{bad}} + \epsilon/2} \\
\text{by (5.3)} \quad &\geq \frac{(1-p) \cdot W_{\text{bad}} + 3\epsilon/2}{2(1-p) \cdot W_{\text{bad}} + 5\epsilon/2} \\
&> \frac{1}{2}.
\end{aligned}$$

□

5.3.3 Local Decoding

We show that all linear affine-invariant codes, in particular lifted codes, have explicit interpolating sets, which allows us to immediately translate the local correctability of lifted codes into local decodability.

Finite field isomorphisms. Let $\mathrm{Tr} : \mathbb{F}_{q^m} \to \mathbb{F}_q$ be the $\mathbb{F}_q$-linear trace map $z \mapsto \sum_{i=0}^{q-1} z^{q^i}$. Let $\alpha_1, \ldots, \alpha_m \in \mathbb{F}_{q^m}$ be linearly independent over $\mathbb{F}_q$ and let $\phi : \mathbb{F}_{q^m} \to \mathbb{F}_q$ be the map $z \mapsto (\mathrm{Tr}(\alpha_1 z), \ldots, \mathrm{Tr}(\alpha_m z))$. Since Tr is $\mathbb{F}_q$-linear, ϕ is $\mathbb{F}_q$-linear, and in fact ϕ is an isomorphism. Observe that ϕ induces a $\mathbb{F}_q$-linear isomorphism $\phi^* : \{\mathbb{F}_q^m \to \mathbb{F}_q\} \to \{\mathbb{F}_{q^m} \to \mathbb{F}_q\}$ defined by

$\phi^*(f) = f \circ \phi$.

It is straightforward to verify that if $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$, and $S \subseteq \mathbb{F}_{q^m}$ is an interpolating set for $\phi^*(\mathcal{C})$, then $\phi(S)$ is an interpolating set for $\mathcal{C}$.

Theorem 5.3.3. *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ be a nontrivial affine-invariant code with $\dim_{\mathbb{F}_q}(\mathcal{C}) = D$. Let $\omega \in \mathbb{F}_{q^m}$ be a generator, i.e. ω has order $q^m - 1$. Let $S = \{\omega, \omega^2, \ldots, \omega^D\} \subseteq \mathbb{F}_{q^m}$. Then $\phi(S) \subseteq \mathbb{F}_q^m$ is an interpolating set for $\mathcal{C}$.*

Proof. The map ϕ induces a map $\phi^* : \{\mathbb{F}_q^m \to \mathbb{F}_q\} \to \{\mathbb{F}_{q^m} \to \mathbb{F}_q\}$ defined by $\phi^*(f) = f \circ \phi$. It suffices to show that S is an interpolating set for $\mathcal{C}' \triangleq \phi^*(\mathcal{C})$. Observe that $\mathcal{C}'$ is affine-invariant over $\mathbb{F}_{q^m}$, and hence has a degree set $\mathrm{Deg}(\mathcal{C}')$, by Proposition 4.1.3. By Proposition 4.1.4, $\dim_{\mathbb{F}_q}(\mathcal{C}') = |\mathrm{Deg}(\mathcal{C}')|$, so suppose $\mathrm{Deg}(\mathcal{C}') = \{i_1, \ldots, i_D\}$. Every $g \in \mathcal{C}'$ is of the form $g(Z) = \sum_{j=1}^{D} a_j Z^{i_j}$, where $a_j \in \mathbb{F}_{q^m}$. By linearity, it suffices to show that if $g \in \mathcal{C}'$ is nonzero, then $g(z) \neq 0$ for some $z \in S$. We have

$$\begin{bmatrix} \omega^{i_1} & \omega^{i_2} & \cdots & \omega^{i_D} \\ \omega^{2i_1} & \omega^{2i_2} & \cdots & \omega^{2i_D} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{Di_1} & \omega^{Di_2} & \cdots & \omega^{Di_D} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_D \end{bmatrix} = \begin{bmatrix} g(\omega) \\ g(\omega^2) \\ \vdots \\ g(\omega^D) \end{bmatrix}$$

and the leftmost matrix is invertible since it is a generalized Vandermonde matrix. Therefore, if $g \neq 0$, then the right-hand side, which is simply the vector of evaluations of g on S, is nonzero. □

Theorem 5.3.4. *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant with distance $\delta \triangleq \delta(\mathcal{C})$, and let $m \geq t$. Then, for every $\epsilon, \eta > 0$, the lifted code $\mathcal{C}^{t \nearrow m}$ is $\left(Q, (\frac{1}{2} - \epsilon)\delta - q^{-t}, \eta\right)$-locally decodable where $Q = O(\ln(1/\eta)/\epsilon^2)$.*

Proof. Follows immediately from Theorems 5.3.2 and 5.3.3 and Proposition 3.2.5. □

5.4 Local Testing and Robust Testing

5.4.1 Local Testing

Many testing results may be viewed as robustness statements about properties. A *characterization* statement would say "an object X has a property P globally if and only if it has property P locally". For example, degree d polynomials over $\mathbb{F}_q$ of characteristic p are characterized by the fact that, when restricted to t-dimensional subspaces, for $t = \left\lceil \frac{d+1}{q-q/p} \right\rceil$, they are degree d polynomials. This was proved in [KR06]. A *testing* statement would go further to say "if X does not have property P globally, then locally it often fails to have property P". For example, building on our previous example, a low-degree testing statement would say that if a polynomial over $\mathbb{F}_q$ of characteristic p has degree greater than d, then on *many* t-dimensional subspaces, it has degree greater than d. This was also proved in [KR06]. A *robust testing* statement would go even further to say "if X is far from having property P globally, then locally its average distance from P is also far".

The lift operator is natural for many reasons, but primarily because it suggests such a natural test for lifted codes. The lifted code $\mathcal{C}^{t\nearrow m}$, by construction, is locally characterized by the fact that any codeword, when restricted to t-dimensional subspaces, is a codeword of $\mathcal{C}$. Along with the symmetry provided by affine-invariance, the work of Kaufman and Sudan [KS08] immediately imply that the natural t-dimensional test is $\Omega(q^{-2t})$-sound, i.e. if $f \notin \mathcal{C}^{t\nearrow m}$, then on at least $\Omega(q^{-2t})$-fraction of t-dimensional subspaces A, $f|_A \notin \mathcal{C}$.

Definition 5.4.1. Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant, and let $t \leq k \leq m$. The *natural k-dimensional test for $\mathcal{C}^{t\nearrow m}$* is the q^k-local tester with the following distribution: it selects a uniformly random k-flat $A \subseteq \mathbb{F}_q^m$ and accepts a function $f : \mathbb{F}_q^m \to \mathbb{F}_q$ if and only if $f|_A \in \mathcal{C}^{t\nearrow k}$.

Theorem 5.4.2 ([KS08])**.** *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant, and let $m \geq t$. Then the natural t-dimensional test for the lifted code $\mathcal{C}^{t\nearrow m}$ is $(q^{-2t}/2)$-sound.*

5.4.2 Robust Testing

However, for our high-rate code constructions, we often take m to be constant and let $q \to \infty$ to make our code longer. As such, a soundness of $\Omega(q^{-O(t)})$ is insufficient as this quickly approaches zero. In fact, in Chapter 6, we prove that the natural $2t$-dimensional (as opposed to t-dimensional) test is α-robust, where α is a polynomial in the distance of the code, but does not depend on q, t, or m. In particular, by Proposition 3.3.2, this implies that the $2t$-dimensional test is α-sound.

Theorem 5.4.3. *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant, and let $m \geq t$. Then the natural $(2t)$-dimensional test for the lifted code $\mathcal{C}^{t \nearrow m}$ is $\frac{\delta(\mathcal{C})^{72}}{2 \cdot 10^{52}}$-robust.*

The proof of Theorem 5.4.3 is somewhat long and technical, so we defer it to Chapter 6.

Chapter 6

Robust Testing of Lifted Codes

6.1 Robustness of Lifted Codes

In this section, we prove Theorem 6.1.18, which is simply a more precise restatement of Theorem 1.2.2. That is, we prove that the natural $2t$-dimensional test for the m-dimensional lift of a t-dimensional code over $\mathbb{F}_q$ is α-robust, where α depends only on the distance of the code and not on t or m or the field size. Our approach is a standard one — we first analyze the test for low-dimensional settings (Theorem 6.1.4), and then use a general projection argument ("bootstrapping") to get an analysis for all dimensions (Theorem 6.1.18).

6.1.1 Preliminaries

We begin by presenting some basic definitions and results on robust testing. We define the robustness of a lifted code, specializing the definition to robustness with respect to subspace testers. We include the dimension of the testing subspace as a parameter in the robustness since this will be convenient later.

Definition 6.1.1. Let $t \leq k \leq m$. The code $\mathcal{C}^{t \nearrow m}$ is (α, k)-*robust* if, for every $r : \mathbb{F}_q^m \to \mathbb{F}_q$,

$$\mathbb{E}_A \left[\delta \left(r|_A, \mathcal{C}^{t \nearrow k} \right) \right] \geq \alpha \cdot \delta \left(r, \mathcal{C}^{t \nearrow m} \right)$$

where the expectation is over uniformly random k-dimensional subspaces $A \subseteq \mathbb{F}_q^m$. When k is clear from context, we say the code is α-robust.

Observe that if A is a random k_1-dimensional affine subspace and B is a random k_2-dimensional affine subspace, where $k_2 \geq k_1$, then

$$\mathbb{E}_A\left[\delta\left(r|_A, \mathcal{C}^{t\nearrow k_1}\right)\right] = \mathbb{E}_B\left[\mathbb{E}_{A\subseteq B}\left[\delta\left(r|_A, \mathcal{C}^{t\nearrow k_1}\right)\right]\right] \leq \mathbb{E}_B\left[\delta\left(r|_B, \mathcal{C}^{t\nearrow k_2}\right)\right]$$

so if $\mathcal{C}^{t\nearrow m}$ is (α, k_1)-robust, then it is also (α, k_2)-robust.

As a corollary to Theorem 5.4.2, the k-dimensional test (for $k \geq t$) for $\mathcal{C}^{t\nearrow m}$ is $O(q^{-3t})$-robust.

Corollary 6.1.2. *If $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ is linear affine-invariant, then $\mathcal{C}^{t\nearrow m}$ is $(\frac{q^{-3t}}{2}, k)$-robust for $k \geq t$.*

Proof. It suffices to show that $\mathcal{C}^{t\nearrow m}$ is $(\frac{q^{-3t}}{2}, t)$-robust, which follows immediately from Theorem 5.4.2 and Proposition 3.3.2. □

Proposition 6.1.3 (**Robustness composes multiplicatively**). *Let $t \leq k_1 \leq k_2 \leq m$ and let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant. If $\mathcal{C}^{t\nearrow m}$ is (α_2, k_2)-robust and $\mathcal{C}^{t\nearrow k_2}$ is (α_1, k_1)-robust, then $\mathcal{C}^{t\nearrow m}$ is $(\alpha_1 \cdot \alpha_2, k_1)$-robust.*

6.1.2 Robustness for Small Dimension

Throughout Sections 6.1.2 and 6.1.3, fix a linear affine invariant code $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ with relative distance $\delta \triangleq \delta(\mathcal{C})$. Let $n \geq 1$ be an integer (we will use $n = 3$ or 4) and let $m = nt$.

Two codes that play a prominent role are the lift $\mathcal{C}^{t\nearrow m}$ of $\mathcal{C}$ to m dimensions, and the n-fold tensor product $\mathcal{C}^{\otimes n}$ of $\mathcal{C}$, which is also an m-dimensional code.

We begin by giving a tester with robust analysis for $\mathcal{C}^{t\nearrow m}$ for this restricted choice of m. We will show that the $(m-t)$-dimensional test is $\left(\frac{\delta^n}{n}\right)^{O(1)}$-robust. (Note the robustness degrades poorly with $n = m/t$ and so can only be applied for small n). It is important, for Section 6.1.4, that there is no dependence on t.

Theorem 6.1.4. *Let $n \geq 3$ and $t \geq 1$ and set $m = nt$. Then $\mathcal{C}^{t\nearrow^m}$ is $(\alpha_0, m-t)$-robust for $\alpha_0 = \frac{\delta^{3n}}{16(n^2+3n+2)^3}$.*

Overview. For simplicity, we describe the proof idea for $t = 1$. Suppose the average local distance to $\mathcal{C}^{1\nearrow^{(m-1)}}$ on random hyperplanes is small. For $\mathbf{a} \in \mathbb{F}_q^m$, let $\mathcal{C}_{\mathbf{a}}$ be the code consisting of tensor codewords in $\mathcal{C}^{\otimes n}$ whose restrictions to lines in direction $\mathbf{a}$ are also codewords of $\mathcal{C}$. Note that $\bigcap_{\mathbf{a}} \mathcal{C}_{\mathbf{a}} = \mathcal{C}^{1\nearrow^m}$. Our main technical result (Theorem 6.1.11) of this section shows that $\mathcal{C}_{\mathbf{a}}$ is $\left(\frac{\delta^n}{n}\right)^{O(1)}$-robust. Now, observe that choosing a random hyperplane can be done by choosing m random linearly independent directions, choosing an additional random direction $\mathbf{a}$ that is not spanned by any $m-1$ of the former, and choosing a random hyperplane spanned by $m-1$ of these $m+1$ random directions (call such a hyperplane "special"). Viewing the first m chosen directions as the standard basis directions, we see that the average local distance to $\mathcal{C}^{1\nearrow^{(m-1)}}$, and hence to $\mathcal{C}^{\otimes(m-1)}$, when restricting to special hyperplanes, is still small. Therefore, for most $\mathbf{a}$, the average local distance to $\mathcal{C}^{\otimes(m-1)}$ on special hyperplanes is small. By the robustness of $\mathcal{C}_{\mathbf{a}}$, this implies that our received word is close to some codeword $c_{\mathbf{a}} \in \mathcal{C}_{\mathbf{a}}$ for most $\mathbf{a}$. But these codewords $c_{\mathbf{a}}$ are all codewords of $\mathcal{C}^{\otimes m}$ and close to each other, so they must be the same codeword $c \in \mathcal{C}^{\otimes m}$. So we have shown that we are close to $c \in \mathcal{C}^{\otimes m}$. We proceed by showing that in fact $c \in \mathcal{C}^{1\nearrow^m}$. Note that $c \in \mathcal{C}_{\mathbf{a}}$ for most $\mathbf{a}$. Another technical result, Corollary 6.2.7, implies that in fact $c \in \mathcal{C}_{\mathbf{a}}$ for all $\mathbf{a}$ and we are done.

To generalize to $t > 1$, we replace dimension k subspaces with dimension kt subspaces throughout. Some work needs to be done to ensure that Theorem 6.1.11 still works, and also Corollary 6.2.7 must be generalized appropriately to remove the dependence on t. These issues will be discussed in the corresponding sections.

Definition 6.1.5. Let $\ell \leq m$ be an integer. A collection $D \subseteq \binom{\mathbb{F}_q^m}{t}$ is *ℓ-proper* if for every k and every distinct $A_1, \ldots, A_k \in D$, the union $\bigcup_{i=1}^k A_i$ contains at least $\min\{kt, m-\ell\}$ linearly independent vectors.

Definition 6.1.6. For a set $D \subseteq \binom{\mathbb{F}_q^m}{t}$, for every $\mathbf{x} \in \mathbb{F}_q^m$ define $\mathcal{V}_D^k(\mathbf{x})$ to be the collection

of subspaces through $\mathbf{x}$ in directions from k different sets from D. More precisely,

$$\mathcal{V}_D^k(\mathbf{x}) \triangleq \left\{ (\mathbf{x}, A) \;\middle|\; A = \bigcup_{i=1}^{k} D_i, \quad \{D_1, \ldots, D_k\} \in \binom{D}{k} \right\}$$

Define

$$\mathcal{V}_D^k \triangleq \bigcup_{\mathbf{x} \in \mathbb{F}_q^n} \mathcal{V}_D^k(\mathbf{x}).$$

The *testing subspaces through* $\mathbf{x}$ are $\mathcal{T}_D(\mathbf{x}) \triangleq \mathcal{V}_D^{m-1}(\mathbf{x})$ and the *decoding subspaces through* $\mathbf{x}$ are $\mathcal{D}_D(\mathbf{x}) \triangleq \mathcal{V}_D^1(\mathbf{x})$. Similarly, the *testing subspaces* are $\mathcal{T}_D \triangleq \mathcal{V}_D^{m-1}$ and the *decoding subspaces* are $\mathcal{D}_D \triangleq \mathcal{V}_D^1$. If $S = (\mathbf{x}, \cup_{i=1}^k D_i) \in \mathcal{V}_D^k$, then the *blocks of* S are the sets $D_1, \ldots, D_k$. Two testing subspaces are *adjacent* if they differ in at most one block.

Remark 6.1.7. If D is ℓ-proper for $\ell \leq t$ then for any $k < n$ we have that $\mathcal{V}_D^k$ consists of kt-dimensional subspaces.

Definition 6.1.8. Define $\mathcal{C}_D^n$ to be the code of all words $f : \mathbb{F}_q^m \to \mathbb{F}_q$ such that $f|_u \in \mathcal{C}$ for every decoding subspace $u \in \mathcal{D}_D$.

Remark 6.1.9. Observe that $\mathcal{C}^{t \nearrow m}$ is a subcode of $\mathcal{C}_D^n$ for any D. If $\bigcup D$ contains the standard basis vectors, then $\mathcal{C}_D^n$ is a subcode of $\mathcal{C}^{\otimes n}$.

Proof of Theorem 6.1.4. Define $\ell \triangleq \left\lfloor \frac{n \log\left(\frac{1}{\delta}\right) + \log(n^2+3n+2)+1}{\log(q)} \right\rfloor$. We note that for the most interesting cases, where $\delta > 0$ and n are fixed and $q \to \infty$, $\ell = 0$. Our first step handles the less interesting cases (by appealing to a known result). Specifically, if $\ell \geq t$ then by Corollary 6.1.2 we are done since

$$\frac{q^{-3t}}{2} \geq \frac{q^{-3\ell}}{2} \geq \frac{q^{-3\left(\frac{n \log\left(\frac{1}{\delta}\right) + \log(n^2+3n+2)+1}{\log(q)}\right)}}{2} = \frac{\delta^{3n}}{16(n^2+3n+2)^3} = \alpha_0 \ .$$

Now assume $\ell < t$ and let $\rho \triangleq \mathbb{E}_v \left[\delta\left(r|_v, \mathcal{C}^{t \nearrow (m-t)}\right)\right]$, where $v \subseteq \mathbb{F}_q^m$ is a uniformly random $(m-t)$-dimensional affine subspace. We will assume without lost of generality that $\rho \leq \alpha_0$ and in particular $\rho \leq \frac{\delta^{3n}}{16\binom{n+1}{n-1}^2(n^2+3n+2)} \leq \frac{\delta^{2n} q^{-\ell}}{8\binom{n+1}{n-1}^2}$.

Observe that

$$\rho = \mathbb{E}_{A_1,\ldots,A_n \in \binom{\mathbb{F}_q^m}{t}} \mathbb{E}_{A \in \binom{\mathbb{F}_q^n}{t}} \mathbb{E}_{v \in \mathcal{T}_{\{A_1,\ldots,A_n,A\}}} \left[\delta\left(r|_v, \mathcal{C}^{t\nearrow(m-t)}\right)\right]$$

where $A_1, \ldots, A_n$ are random sets such that their union is linearly independent, and A is a random set such that $\{A_1, \ldots, A_n, A\}$ is ℓ-proper. Fix $A_1, \ldots, A_n$ such that

$$\mathbb{E}_{A \in \binom{\mathbb{F}_q^m}{t}} \mathbb{E}_{v \in \mathcal{T}_{\{A_1,\ldots,A_n,A\}}} \left[\delta\left(r|_v, \mathcal{C}^{t\nearrow(m-t)}\right)\right] \leq \rho.$$

Since $\mathcal{C}^{t\nearrow(m-t)} \subseteq \mathcal{C}^{\otimes(n-1)}$,

$$\mathbb{E}_{A \in \binom{\mathbb{F}_q^m}{t}} \mathbb{E}_{v \in \mathcal{T}_{\{A_1,\ldots,A_n,A\}}} \left[\delta\left(r|_v, \mathcal{C}^{\otimes(n-1)}\right)\right] \leq \rho.$$

By affine-invariance, we may assume without loss of generality that $A_1, \ldots, A_n$ form the standard basis vectors for $\mathbb{F}_q^m$. For any $A \in \binom{\mathbb{F}_q^m}{t}$, let $D_A \triangleq \{A_1, \ldots, A_n, A\}$. By Markov's inequality,

$$\Pr_A\left[\mathbb{E}_{v \in \mathcal{T}_{D_A}}\left[\delta\left(r|_v, \mathcal{C}^{\otimes(n-1)}\right) \geq 2\delta^{-n}\rho\right]\right] < \frac{1}{2}\delta^n.$$

So, for more than $1 - \frac{1}{2}\delta^n$ fraction of blocks A such that D_A is ℓ-proper, we have a codeword $c_A \in \mathcal{C}^n_{D_A} \subseteq \mathcal{C}^{\otimes n}$ such that (by Theorem 6.1.11) $\delta(r, c_A) < 2\delta^{-n}\rho\binom{n+1}{n-1} < \frac{1}{2}\delta^n$. For every two such blocks A, A', we have $\delta(c_A, c_{A'}) \leq \delta(c_A, r) + \delta(r, c_{A'}) < \delta^n = \delta(\mathcal{C}^{\otimes n})$, so there is some codeword $c \in \mathcal{C}^{\otimes n}$ such that $c_A = c$ for every such A. For such A, it follows that for every $\mathbf{b} \in \mathbb{F}_q^m$, the restriction of c to the subspace $(\mathbf{b}, A)$ is a codeword of $\mathcal{C}$, i.e. $c|_{(\mathbf{b},A)} \in \mathcal{C}$. By Claim 6.1.17, for more than $1 - \frac{1}{2}\delta^n - \binom{n}{2}\frac{q^{-t}}{q-1} - n\frac{q^{-l}}{q-1}$ fraction of blocks A (without the requirement that D_A be proper), $c|_{(\mathbf{b},A)} \in \mathcal{C}$ for every $\mathbf{b} \in \mathbb{F}_q^m$. In particular, $c \in \mathcal{C}^{\otimes n}$ and for every $\mathbf{b} \in \mathbb{F}_q^m$, $c|_{(\mathbf{b},A)} \notin \mathcal{C}$ for less than $\frac{1}{2}\delta^n + \binom{n}{2}\frac{q^{-t}}{q-1} + n\frac{q^{-\ell}}{q-1}$ fraction of A. It sufficient to show that $\frac{1}{2}\delta^n + \binom{n}{2}\frac{q^{-t}}{q-1} + n\frac{q^{-l}}{q-1} \leq \delta^n - (n+1)q^{-t}$. Then it will follow from Corollary 6.2.7

that $c \in \mathcal{C}^{t\nearrow m}$ and since $\delta(r,c) \le 2\delta^{-n}\rho\binom{n+1}{n-1}$ we are done. Calculating:

$$\begin{aligned}
\binom{n}{2}\frac{q^{-t}}{q-1} + n\frac{q^{-l}}{q-1} + (n+1)q^{-t} &\le \binom{n}{2}\frac{q^{-l}}{q-1} + n\frac{q^{-l}}{q-1} + (n+1)\frac{q^{-l}}{q-1} \\
&= \frac{q^{-l}}{q-1}\left(\frac{n^2+3n+2}{2}\right) \\
&\le \frac{q\delta^n}{4(q-1)} \\
&\le \frac{1}{2}\delta^n.
\end{aligned}$$

□

The composability of robust tests immediately yields the following corollary where the test is now $2t$ dimensional.

Corollary 6.1.10. *$\mathcal{C}^{t\nearrow 4t}$ is $(\alpha_1, 2t)$-robust, where $\alpha_1 \ge \frac{\delta^{21}}{6\cdot 10^{10}}$.*

Proof. By Theorem 6.1.4, $\mathcal{C}^{t\nearrow 4t}$ is $\left(\frac{\delta^{12}}{432{,}000}, 3t\right)$-robust and $\mathcal{C}^{t\nearrow 3t}$ is $\left(\frac{\delta^{9}}{128{,}000}, 2t\right)$-robust. Therefore, by composing, the $2t$-dimensional robustness of $\mathcal{C}^{t\nearrow 4t}$ is at least $\frac{\delta^{12}}{432{,}000}\cdot\frac{\delta^{9}}{128{,}000} = \frac{\delta^{21}}{55{,}296{,}000{,}000}$ □

6.1.3 Robustness of Special Tensor Codes

In this section, we prove the main technical result (Theorem 6.1.11) used in Section 6.1.2.

Theorem 6.1.11. *Let $n \ge 3$ and $\ell \le t$. Set $m = nt$. Let $D \subseteq \binom{\mathbb{F}_q^n}{t}$ be ℓ-proper with $|D| \ge n$ blocks. Let $r : \mathbb{F}_q^m \to \mathbb{F}_q$ be a word with $\rho \triangleq \mathbb{E}_{v\in\mathcal{T}_D}\left[\delta\left(r|_v, \mathcal{C}^{\otimes(n-1)}\right)\right]$. If $\rho < \frac{\delta^n q^{-\ell}}{4\binom{|D|}{n-1}^2}$, then $\delta(r, \mathcal{C}_D^n) \le \rho\binom{|D|}{n-1}$.*

Overview. Our analysis is an adaptation of Viderman's [Vid12]. For simplicity, assume $t = 1$. We define a function $c : \mathbb{F}_q^m \to \mathbb{F}_q$, which we show is both close to r and a codeword of $\mathcal{C}_D^n$. Following Viderman's analysis, we partition $\mathbb{F}_q^m$ into "good", "fixable", and "bad" points. Each hyperplane $v \in \mathcal{T}_D$ has an associated codeword $c_v \in \mathcal{C}^{\otimes(m-1)}$, the nearest codeword to $r|_v$, and an opinion $c_v(\mathbf{x})$ about $\mathbf{x}$. "Good" points are points for which any

hyperplane agrees with r. "Fixable" points are points for which hyperplanes agree with each other, but not with r. "Bad" points are points for which at least two hyperplanes disagree with each other. For good or fixable $\mathbf{x}$, we naturally define $c(\mathbf{x})$ to be the common opinion $c_v(\mathbf{x})$ of any hyperplane v through $\mathbf{x}$. Claim 6.1.13 implies that there are not many bad points, which immediately shows that c is close to r.

So far, our proof has been a straightforward adaptation of Viderman's. However, at this point, we are forced to depart from Viderman's proof. A hyperplane is "bad" if it has more than $\frac{1}{2}\delta^{m-1}$ fraction bad points. Claim 6.1.12 shows that every bad point is in a bad hyperplane, and Claim 6.1.14 shows that there are less than $\frac{1}{2}\delta q$ bad hyperplanes. In [Vid12], which analyses $\mathcal{C}^{\otimes m}$ and axis-parallel hyperplanes instead of $\mathcal{C}^n_D$ and $\mathcal{T}_D$, this is already enough, since this implies that in each axis-parallel direction, there are less than δq bad hyperplanes, so the remaining points are all good or fixable and with a little bit more work, one can show that c can be extended uniquely to a tensor codeword using the erasure-decoding properties of tensor codes. Unfortunately, we do not have this structure.

We say a line is "good" if it is contained in some good hyperplane, otherwise it is bad. We must further partition the bad points into merely bad and "super-bad" points, which are points such that either every hyperplane is bad, or there are two disagreeing good hyperplanes. For merely bad $\mathbf{x}$, we define $c(\mathbf{x})$ to be the common opinion $c_v(\mathbf{x})$ of any good hyperplane v through $\mathbf{x}$. For super-bad $\mathbf{x}$, we pick any line u through $\mathbf{x}$, take the restriction of c to the non-super-bad points on u, and extend it to a codeword $c_u \in \mathcal{C}$, and define $c(\mathbf{x}) \triangleq c_u(\mathbf{x})$. Two non-trivial steps remain: showing that $c(\mathbf{x})$ is well-defined for super-bad $\mathbf{x}$, and showing that $c \in \mathcal{C}^n_D$.

Claim 6.1.15 shows that, for any special plane, there are less than $\frac{1}{2}\delta q$ lines in each direction that are bad (not contained in any good hyperplane) or contain a super-bad point. This is proved by exhibiting, for each such undesirable line, a bad hyperplane in a fixed direction containing the line. If there were too many undesirable lines, this would result in too many parallel bad hyperplanes, contradicting Claim 6.1.14. Finally, Claim 6.1.16 shows that if u is a line with no super-bad points, then $c|_u \in \mathcal{C}$ is a codeword.

Now, we show that c is well-defined on super-bad $\mathbf{x}$. Let u_1, u_2 be two lines through $\mathbf{x}$. Let P be the plane through $\mathbf{x}$ containing u_1, u_2. On this plane, by Claim 6.1.15, in each direction we have enough lines u with no super-bad points, for which $c|_u \in \mathcal{C}$ (by Claim 6.1.16), so that we can uniquely extend c onto the entire plane (by Proposition A.2.5). This gives a well-defined value for $c(\mathbf{x})$.

Finally, we show that $c \in \mathcal{C}_D^n$. Let u be any line. If u has no super-bad points, then $c|_u \in \mathcal{C}$ follows from Claim 6.1.16. If c does have a super-bad point $\mathbf{x}$, then $c|_u \in \mathcal{C}$ by the way we defined $c(\mathbf{x})$.

This completes our analysis for the case $t = 1$. To generalize to $t > 1$, we replace lines with "decoding subspaces" (subspaces of dimension t), planes with subspaces of dimension $2t$, and hyperplanes with "testing subspaces" (subspaces of codimension t). Some care must be taken when proving Claim 6.1.15, because the intersection of two decoding subspaces may have non-trivial dimension. We therefore require the notion of "ℓ-properness" of D, and must modify Claim 6.1.14 and also prove Claim 6.1.17 to accommodate this notion. Details follow.

Proof of Theorem 6.1.11. For each testing subspace $v \in \mathcal{T}_D$, define $c_v \in \mathcal{C}^{\otimes(n-1)}$ to be the closest codeword to $r|_v$ (break ties arbitrarily). We will partition $\mathbb{F}_q^m$ into three disjoint sets G, F, B (*good*, *fixable*, and *bad points*, respectively) as follows:

$$\begin{aligned} G &\triangleq \left\{\mathbf{x} \in \mathbb{F}_q^m \mid c_v(\mathbf{x}) = r(\mathbf{x}) \text{ for every } v \in \mathcal{T}_D(\mathbf{x})\right\} \\ F &\triangleq \left\{\mathbf{x} \in \mathbb{F}_q^m \mid c_v(\mathbf{x}) = c_{v'}(\mathbf{x}) \neq r(\mathbf{x}) \text{ for every } v, v' \in \mathcal{T}_D(\mathbf{x})\right\} \\ B &\triangleq \left\{\mathbf{x} \in \mathbb{F}_q^m \mid c_v(\mathbf{x}) \neq c_{v'}(\mathbf{x}) \text{ for some } v, v' \in \mathcal{T}_D(\mathbf{x})\right\}. \end{aligned}$$

Call a testing subspace *bad* if at least $\frac{1}{2}\delta^{n-1}$ fraction of its points are in B, and *good* otherwise. A decoding subspace is *good* if it is contained in some good testing subspace, and *bad* otherwise. Further, define the set B' of *super-bad* points

$$B' \triangleq \{\mathbf{x} \in B \mid \text{every } v \in \mathcal{T}_D(\mathbf{x}) \text{ is bad or } \exists \text{ good } v, v' \in \mathcal{T}_D(\mathbf{x}) \text{ such that } c_v(\mathbf{x}) \neq c_{v'}(\mathbf{x})\}.$$

Claim 6.1.12. *If $v, v' \in \mathcal{T}_D$ are adjacent good testing subspaces, then $c_v|_{v\cap v'} = c_{v'}|_{v\cap v'}$. In particular, every bad point is in a bad testing subspace.*

Proof. Suppose $\mathbf{b} \in v \cap v'$ and $c_v(\mathbf{b}) \neq c_{v'}(\mathbf{b})$. Since v, v' are adjacent, they have $n-2$ blocks $A_1, \ldots, A_{n-2}$ in common. Let v have blocks $A_1, \ldots, A_{n-2}, A$ and let v' have blocks $A_1, \ldots, A_{n-2}, A'$.

Let $u \in \mathcal{D}_D$ be the decoding subspace $(\mathbf{b}, A_1)$. Since $c_v|_u, c_{v'}|_u \in \mathcal{C}$ disagree on $\mathbf{b}$, they are distinct codewords and hence disagree on at least δq^t points of u, say $\mathbf{x}_1, \ldots, \mathbf{x}_{\delta q^t}$. For each $i \in [\delta q^t]$, let $v_i \in \mathcal{T}_D$ be the testing subspace $(\mathbf{x}_i, A_2 \cup \cdots \cup A_{n-2} \cup A \cup A')$.

Since $c_v(\mathbf{x}_i) \neq c_{v'}(\mathbf{x}_i)$, that means c_{v_i} disagrees with one of $c_v, c_{v'}$ at $\mathbf{x}_i$. Without loss of generality, suppose c_v disagrees with $c_{v_1}, \ldots, c_{v_{\delta q^t/2}}$. We will show that v is bad, which proves the first part of the claim.

For each $i \in [\delta q^t]$, let $w_i = (\mathbf{x}_i, A_2 \cup \cdots \cup A_{n-2} \cup A)$. Note that $u \cap w_i = \{\mathbf{x}_i\}$ (since D is ℓ-proper, for $\ell \leq t$), so all w_i are different parallel subspaces and hence disjoint. Since $w_i \in \mathcal{V}_D^{n-2}$, the restrictions $c_v|_{w_i}, c_{v_i}|_{w_i} \in \mathcal{C}^{\otimes n-2}$ are codewords and are distinct because they disagree on $\mathbf{x}_i$, therefore, by Proposition A.2.4, they disagree on at least $\delta^{n-2}q^{m-2t}$ points in w_i, which are therefore bad. Thus, each v_i contributes $\delta^{n-2}q^{m-2t}$ bad points to v, for a total of $\frac{1}{2}\delta^{n-1}q^{m-t}$ bad points since the w_i are disjoint.

For the second part, suppose $\mathbf{b} \in B$ is a bad point. We will show that $\mathbf{b}$ lies in a bad testing subspace. By definition, there are two testing subspaces $v, v' \in \mathcal{T}_D(\mathbf{b})$ such that $c_v(\mathbf{b}) \neq c_{v'}(\mathbf{b})$. Suppose v has blocks $A_1, \ldots, A_{n-1}$ and v' has directions $A'_1, \ldots, A'_{n-1}$. Assume, without lost of generality, that if $A_i = A'_j$ then $i = j$. Define $v_0 \triangleq v$, and for $i \in [n-1]$, define $v_i \in \mathcal{T}_D$ to be the testing subspace through $\mathbf{b}$ in directions $A'_1, \ldots, A'_i, A_{i+1}, \ldots, A_{n-1}$. Consider the sequence $v_0, v_1, \ldots, v_{n-1}$ of testing subspaces. For each i, the testing subspaces v_i, v_{i+1} are adjacent. Since $c_{v_0}(\mathbf{b}) \neq c_{v_{n-1}}(\mathbf{b})$, there exists some i such that $c_{v_i}(\mathbf{b}) \neq c_{v_{i+1}}(\mathbf{b})$, and by the first part of the claim it follows that one of v_i, v_{i+1} is bad. □

Claim 6.1.13. $\rho \geq \frac{|F|}{q^m} + \frac{|B|}{q^m\binom{|D|}{n-1}}$

Proof. Observe that $|\mathcal{T}_D| = q^t\binom{|D|}{n-1}$. Therefore,

$$\begin{aligned}
\rho &= \mathbb{E}_{v\in\mathcal{T}_D}\left[\delta\left(r|_v, \mathcal{C}^{\otimes(n-1)}\right)\right] \\
&= \mathbb{E}_{v\in\mathcal{T}_D}\left[\delta\left(r|_v, c_v\right)\right] \\
&= \frac{1}{q^t\binom{|D|}{n-1}}\sum_{v\in\mathcal{T}_D}\frac{1}{q^{m-t}}\sum_{\mathbf{x}\in v}\mathbb{1}_{c_v(\mathbf{x})\neq r(\mathbf{x})} \\
&= \frac{1}{q^m\binom{|D|}{n-1}}\sum_{\mathbf{x}\in\mathbb{F}_q^m}\#\{v\in\mathcal{T}_D(\mathbf{x}) \mid c_v(\mathbf{x})\neq r(\mathbf{x})\} \\
&\geq \frac{1}{q^m\binom{|D|}{n-1}}\left(\sum_{\mathbf{x}\in G}0 + \sum_{\mathbf{x}\in F}\binom{|D|}{m-1} + \sum_{\mathbf{x}\in B}1\right) \\
&= \frac{|F|}{q^m} + \frac{|B|}{q^m\binom{|D|}{n-1}}.
\end{aligned}$$

□

Claim 6.1.14. *There are less than $\frac{1}{2}\delta q^{t-\ell}$ bad testing subspaces.*

Proof. By Claim 6.1.13, there are at most $|B| \leq \rho\binom{|D|}{n-1}q^m$ bad points. Each bad testing subspace has at least $\delta^{n-1}q^{m-t}/2$ bad points by definition. Each bad point has at most $\binom{|D|}{n-1}$ bad testing subspaces through it. Therefore, the number of testing subspaces is bounded by

$$\frac{|B|}{\frac{1}{2}\delta^{n-1}q^{m-t}}\cdot\binom{|D|}{n-1} \leq \frac{2\rho}{\delta^{n-1}}\binom{|D|}{n-1}^2 q^t < \frac{1}{2}\delta q^{t-\ell}.$$

□

Now we proceed to prove the lemma. We construct a codeword $c \in \mathcal{C}_D^n$ with $\delta(r,c) \leq \rho\binom{|D|}{n-1}$ in stages, as follows. First, for $\mathbf{x} \in G \cup F$, we define $c(\mathbf{x}) \triangleq c_v(\mathbf{x})$ for any testing subspace $v \in \mathcal{T}_D(\mathbf{x})$. This is well-defined by the definition of G and F. Furthermore, since $c(\mathbf{x}) = c_v(\mathbf{x}) = r(\mathbf{x})$ for $\mathbf{x} \in G$, we already guarantee that $\delta(r,c) \leq \frac{|F|+|B|}{q^m} \leq \rho\binom{|D|}{n-1}$.

For $\mathbf{x} \in B \setminus B'$, define $c(\mathbf{x}) \triangleq c_v(\mathbf{x})$ for any good testing subspace $v \in \mathcal{T}_D(\mathbf{x})$, whose existence is guaranteed by the fact that $\mathbf{x} \notin B'$. This is well-defined because if $v, v' \in \mathcal{T}_D(\mathbf{x})$ are both good, then it follows from the fact that $\mathbf{x} \notin B'$ that $c_v(\mathbf{x}) = c_{v'}(\mathbf{x})$.

Claim 6.1.15. *Let $w \in \mathcal{V}_D^2$ be a subspace in directions $A_1, A_2 \in D$. For each $i \in \{1,2\}$, w contains less than $\frac{1}{2}\delta q^t$ decoding subspaces in direction A_i which intersect B' or are bad (not contained in any good testing subspace).*

Proof. By symmetry, it suffices to consider $i = 2$. Let $A_3, \ldots, A_n \in D$ be blocks in some other direction. Let $u_1, \ldots, u_k \subseteq w$ be decoding subspaces in direction A_2 such that, for each $j \in [k]$, u_j intersects B' or is bad. It suffices to exhibit, for each $j \in [k]$, a bad testing subspace $v \in \mathcal{T}_D$ containing u_j which has block A_2 but not A_1. In this case we will show that $|v \cap w| \leq q^{t+\ell}$ so each such bad testing subspace contain at most q^ℓ of the u_i-s. Since, by Claim 6.1.14, there are at most $\frac{1}{2}\delta q^{t-\ell}$ such subspaces, we get that $k \leq \frac{1}{2}\delta q^t$. Indeed, since D is ℓ-proper, the subspace $u + v \in \mathcal{V}_D^n$ has dimension at least $m - \ell$. Therefore,

$$\dim(v \cap w) = \dim(v) + \dim(w) - \dim(v + w) \leq m - t + 2t - (m - l) = t + l$$

and $|v \cap w| \leq q^{t+\ell}$.

Fix $j \in [l]$ and $u \triangleq u_j$ and we will show that u is contained in a bad testing subspace. If u is bad, then we are done, since any testing subspace containing u, in particular the testing subspace in directions $A_2, \ldots, A_n$, is bad. Now suppose u has a point $\mathbf{x} \in u \cap B'$. Let $v \in \mathcal{T}_D$ be the testing subspace $(\mathbf{x}, A_2 \cup \cdots \cup A_n)$. If v is bad, we are done. Otherwise, since $\mathbf{x} \in B'$, there exists another good hyperplane $v' \in D$, in directions $A'_1, \ldots, A'_{n-1}$, such that $c_v(\mathbf{x}) \neq c_{v'}(\mathbf{x})$. Without loss of generality, assume that if $A_i = A'_j$ then $i = j$ (in particular $A_1 \notin \{A'_2, \ldots, A'_{n-1}\}$). For each $i \in [n-1]$, if $A_2 = A'_2$ define $v_i \in \mathcal{T}_D(\mathbf{x})$ to be the testing subspace $(\mathbf{x}, A'_2, \ldots, A'_i, A_{i+1}, \ldots, A_n)$, and if $A_2 \neq A'_2$ define v_1 to be v and v_i to be $(\mathbf{x}, A_2, A'_2, \ldots, A'_i, A_{i+1}, \ldots, A_{n-1})$. In any case define $v_n \triangleq v'$. For every $i \in [n-1]$, v_i and v_{i+1} are adjacent. Note that for every $i \in [n-1]$, v_i contains the direction A_2 and does not contain the direction A_1. We will show that v_i is bad for some $i \in [n-1]$. Since $c_{v_1}(\mathbf{x}) \neq c_{v_n}(\mathbf{x})$, there exists some $i \in [n-1]$ such that $c_{v_i}(\mathbf{x}) \neq c_{v_{i+1}}(\mathbf{x})$, and therefore, by Claim 6.1.12, one of v_i, v_{i+1} is bad. If $i < n-1$, then $i, i+1 \leq n-1$, and so we are done. If $i = n-1$, then by assumption $v_n = v'$ is good, so it must be that v_{n-1} is bad. □

Claim 6.1.16. *If $u \in \mathcal{D}_D$ is a decoding subspace and $u \cap B' = \emptyset$, then for every $\mathbf{x} \in u$ there is a codeword $c_{\mathbf{x}} \in \mathcal{C}$ defined on u such that $c_{\mathbf{x}}(\mathbf{x}) = c(\mathbf{x})$ and $\delta(c_{\mathbf{x}}, c|_u) < \frac{\delta}{2}$. In particular, $c|_u \in \mathcal{C}$.*

Proof. Fix $\mathbf{x} \in u$. Let $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_t\} \in D$ be the directions of u. Since $\mathbf{x} \notin B'$, there is a good testing subspace $v \in \mathcal{T}_D(\mathbf{x})$. Let $A' = \{\mathbf{a}'_1, \ldots, \mathbf{a}'_t\}$ be some block in v not equal to A and consider the subspace $w = (\mathbf{x}, A \cup A') \in \mathcal{V}_D^2$. For $\mathbf{s}, \mathbf{s}' \in \mathbb{F}_q^t$, let $w(\mathbf{s}, \mathbf{s}') \triangleq \mathbf{x} + \sum_{i=1}^t s_i \mathbf{a}_i + \sum_{i=1}^t s'_i \mathbf{a}'_i$. Let

$$w(\mathbf{s}, *) \triangleq \{w(\mathbf{s}, \mathbf{s}') \mid \mathbf{s}' \in \mathbb{F}_q^t\} \in \mathcal{D}_D,$$

$$w(*, \mathbf{s}') \triangleq \{w(\mathbf{s}, \mathbf{s}') \mid \mathbf{s} \in \mathbb{F}_q^t\} \in \mathcal{D}_D.$$

Let $I \subseteq \mathbb{F}_q^t \setminus \{\mathbf{0}\}$ be the set of points $\mathbf{s} \neq \mathbf{0}$ such that $w(\mathbf{s}, *)$ intersects B' or is bad. Similarly, let $I' \subseteq \mathbb{F}_q^t \setminus \{\mathbf{0}\}$ be the set of points $\mathbf{s}' \neq \mathbf{0}$ such that $w(*, \mathbf{s}')$ intersects B' or is bad. By Claim 6.1.15, $|I|, |I'| < \frac{1}{2}\delta q^t$. Note that for each $(\mathbf{s}, \mathbf{s}') \in (\mathbb{F}_q^t \setminus I) \times (\mathbb{F}_q^t \setminus I')$, we have $w(\mathbf{s}, \mathbf{s}') \notin B'$: if $\mathbf{s} \neq \mathbf{0}$ or $\mathbf{s}' \neq \mathbf{0}$, this follows from the definition of I and I'; if $\mathbf{s} = \mathbf{s}' = \mathbf{0}$, then $w(\mathbf{s}, \mathbf{s}') = \mathbf{x} \notin B'$. Thus c is defined on $w((\mathbb{F}_q^t \setminus I) \times (\mathbb{F}_q^t \setminus I'))$. Note that for each $\mathbf{s} \in \mathbb{F}_q^t \setminus I$, the decoding subspace $w(\mathbf{s}, *)$ is good and hence contained in a good testing subspace $v_{\mathbf{s}} \in \mathcal{T}_D$, therefore $c_{v_{\mathbf{s}}}|_{w(\mathbf{s},*)} \in \mathcal{C}$. Similarly, for each $\mathbf{s}' \in \mathbb{F}_q^t \setminus I'$, the decoding subspace $w(*, \mathbf{s}')$ is contained in a good testing subspace $v_{\mathbf{s}'} \in \mathcal{T}_D$, hence $c_{v_{\mathbf{s}'}}|_{w(*,\mathbf{s}')} \in \mathcal{C}$. Since $|I|, |I'| < \frac{1}{2}\delta q^t$, by Proposition A.2.5, c can be extended uniquely into $c_w \in \mathcal{C}^{\otimes 2}$ defined on w. Define $c_{\mathbf{x}} \triangleq c_w|_u$. Note that $c_{\mathbf{x}} \in \mathcal{C}$ since it is the restriction of $c_w \in \mathcal{C}^{\otimes 2}$ to $u = w(*, \mathbf{0})$. Also, if $\mathbf{s} \in \mathbb{F}_q^t \setminus I$, then $c|_{w(\mathbf{s},*)} = c_{v_{\mathbf{s}}}|_{w(\mathbf{s},*)} = c_w|_{w(\mathbf{s},*)}$ and in particular, $c(w(\mathbf{s}, \mathbf{0})) = c_w(w(\mathbf{s}, \mathbf{0})) = c_{\mathbf{x}}(w(\mathbf{s}, \mathbf{0}))$. So $\delta(c, c_{\mathbf{x}}) \leq \frac{|I|}{q^t} < \frac{\delta}{2}$. Finally, since $\mathbf{0} \notin I, I'$, we have $c(\mathbf{x}) = c(w(\mathbf{0}, \mathbf{0})) = c_{\mathbf{x}}(w(\mathbf{0}, \mathbf{0})) = c_{\mathbf{x}}(\mathbf{x})$. This proves the first part of the claim.

For the second part (showing $c|_u \in \mathcal{C}$), fix some $\mathbf{x}_0 \in u$. For each $\mathbf{x} \in u$, let $c_{\mathbf{x}}$ be the codeword guaranteed by the previous part. Then, for every $\mathbf{x} \in u$, $\delta(c_{\mathbf{x}_0}, c_{\mathbf{x}}) \leq \delta(c_{\mathbf{x}_0}, c|_u) + \delta(c|_u, c_{\mathbf{x}}) < \delta$, therefore $c_{\mathbf{x}_0} = c_{\mathbf{x}}$. Moreover, for all $\mathbf{x} \in u$, $c_{\mathbf{x}_0}(\mathbf{x}) = c_{\mathbf{x}}(\mathbf{x}) = c(\mathbf{x})$, so $c|_u = c_{\mathbf{x}_0} \in \mathcal{C}$. □

We proceed to define $c(\mathbf{x})$ for $x \in B'$. For such an $\mathbf{x}$, pick any decoding subspace $u \in \mathcal{D}_D(\mathbf{x})$, extend $c|_{u \setminus B'}$ to a codeword $c_u \in \mathcal{C}$, and define $c(\mathbf{x}) \triangleq c_u(\mathbf{x})$. We now argue that this is well-defined. Suppose $u_1, u_2 \in \mathcal{D}_D(\mathbf{x})$ in directions $A_1, A_2 \in D$, respectively. We need to show that c_{u_1}, c_{u_2} are well-defined and that $c_{u_1}(\mathbf{x}) = c_{u_2}(\mathbf{x})$. Let $w \in \mathcal{V}_D^2$ be the unique subspace containing u_1, u_2, so $w = (\mathbf{x}, A_1 \cup A_2)$. By Claim 6.1.15, in each direction A_1, A_2, there are less than $\frac{1}{2}\delta q^t$ decoding subspaces in that direction in w which intersect B'. In particular, this implies that u_1, u_2 each contain less than δq^t points from B'. By what we just showed, there are sets $J_1, J_2 \subseteq \mathbb{F}_q^t$ of size $|J_1|, |J_2| > (1-\delta)q^t$ such that the "sub-rectangle" $R \triangleq w(J_1 \times J_2)$ contains no points from B', and therefore c has already been defined on R. By Claim 6.1.16, on each decoding subspace u in R in either direction A_1 or A_2, $c|_u \in \mathcal{C}$. Applying Proposition A.2.5, we see that $c|_R$ can be uniquely extended to a tensor codeword $c_w \in \mathcal{C}^{\otimes 2}$ on w, and this gives a way to extend $c|_{u_i \setminus B'}$ to the codeword $c_{u_i} \triangleq c_w|_{u_i} \in \mathcal{C}$ for $i \in \{1, 2\}$. Therefore, the extensions c_{u_1}, c_{u_2} agree on $\mathbf{x}$ since $c_{u_1}(\mathbf{x}) = c_w(\mathbf{x}) = c_{u_2}(\mathbf{x})$, and moreover for each decoding subspace u_i this extension is unique since each decoding subspace has less than δq^t points from B'.

Now that we have defined $c : \mathbb{F}_q^m \to \mathbb{F}_q$ and have shown that $\delta(r, c) \leq \rho\binom{|D|}{n-1}$, it only remains to show that $c \in \mathcal{C}_D^n$. Let $u \in \mathcal{D}_D$ be a decoding subspace. If $u \cap B' = \emptyset$, then $c|_u \in \mathcal{C}$ follows from Claim 6.1.16. If u intersects B', by the way we defined $c(\mathbf{x})$ for $\mathbf{x} \in B'$, we showed that for any decoding subspace u through $\mathbf{x} \in B'$, $c|_u \in \mathcal{C}$ by extending $c|_{u \setminus B'}$ to a codeword. □

Claim 6.1.17. *Let $\ell \leq t$ be a natural number and $A_1, \ldots, A_n \in \binom{\mathbb{F}_q^m}{t}$ be such that their union is linearly independent. Then at least $1 - \binom{n}{2}\frac{q^{-t}}{q-1} - n\frac{q^{-\ell}}{q-1}$ fraction of $A \in \binom{\mathbb{F}_q^m}{t}$ satisfy that $A, A_1, \ldots, A_n$ is ℓ-proper.*

Proof. Let $\mathbf{a}_1, \ldots, \mathbf{a}_t$ be the random elements comprising A. By a union bound, it suffices to show for any $S \in \binom{[n]}{n-2}$ that $\left(\bigcup_{i \in S} A_i\right) \cup A$ is linearly independent with probability at least $1 - \frac{q^{-t}}{q-1}$ and for any $T \in \binom{[n]}{n-1}$ the probability that $\left(\bigcup_{i \in T} A_i\right) \cup A$ contains at least $m - \ell$ linearly independent elements is at least $1 - \frac{q^{-\ell}}{q-1}$.

Fix $S \in \binom{[n]}{n-2}$ .For any $j \in [t]$, the probability that $\mathbf{a}_j \in \mathbb{F}_q^m$ is in the span of $\left(\bigcup_{i \in S} A_i\right) \cup$

$\{\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}\}$, conditioned on the event that the latter is linearly independent, is $\frac{q^{m-2t+j-1}}{q^m} = q^{-2t+j-1}$. So the probability that $\left(\bigcup_{i\in S} A_i\right) \cup A$ is linearly independent is

$$\begin{aligned}\prod_{j=1}^{t}\left(1-q^{-2t+j-1}\right) &\geq 1-\sum_{j=1}^{t} q^{-2t+j-1} \\ &\geq 1-q^{-t}\sum_{j=1}^{\infty} q^{-j} \\ &= 1-\frac{q^{-t}}{q-1}.\end{aligned}$$

Now fix $T \in \binom{[n]}{n-1}$. Similarly, The probability that $a_j \in \left(\bigcup_{i\in T} A_i\right) \cup \{\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}\}$, condition on the event that the later linearly independent is q^{-t+j-1}. So we get that $\left(\bigcup_{i\in S} A_i\right) \cup \{\mathbf{a}_1, \ldots, \mathbf{a}_{t-\ell}\}$ are linearly independent is

$$\begin{aligned}\prod_{j=1}^{t-\ell}\left(1-q^{-t+j-1}\right) &\geq 1-\sum_{j=1}^{t-\ell} q^{-t+j-1} \\ &\geq 1-q^{-\ell}\sum_{j=1}^{\infty} q^{-j} \\ &= 1-\frac{q^{-\ell}}{q-1}.\end{aligned}$$

□

6.1.4 Robustness for Large Dimension

In this section, we prove our main result of the chapter:

Theorem 6.1.18. *Let $\rho \triangleq \mathbb{E}_v[\delta(r|_v, \mathcal{C}^{t\nearrow 2t})]$, where v is a random affine subspace of dimension $2t$. Let α_1 be the $2t$-dimensional robustness of $\mathcal{C}^{t\nearrow 4t}$ given by Corollary 6.1.10. If $\rho < \frac{\alpha_1\delta^3}{400} - 3q^{-t}$, then $\rho \geq \left(1-\frac{\delta}{4}\right)\cdot\delta(r, \mathcal{C}^{t\nearrow m})$. In particular, $\mathcal{C}^{t\nearrow m}$ is $(\alpha_2, 2t)$-robust, where $\alpha_2 \geq \frac{\delta^{72}}{2\cdot 10^{52}}$.*

Notation. Throughout Section 6.1.4, fix the received word $r : \mathbb{F}_q^m \to \mathbb{F}_q$ and define $\rho \triangleq \mathbb{E}_v[\delta(r|_v, \mathcal{C}^{t\nearrow 2t})]$, and we will assume that $0 < \rho < \frac{\alpha_1\delta^3}{400} - 3q^{-t}$. The case where $\frac{\alpha_1\delta^3}{400} > 3q^{-t}$ is easily dealt with at the end of the proof by using Corollary 6.1.2. Note that, since $\alpha_1, \delta \leq 1$, this implies $q^{-t} \leq \frac{\delta}{1200}$. Throughout this section we will assume $m \geq 4t$. If $m < 4t$ we can pad the function f to get a function $\hat{f} : \mathbb{F}_q^{4t} \to \mathbb{F}_q$ (by setting $\hat{f}(\mathbf{x}, \mathbf{y}) = f(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{F}_q^m$ and $\mathbf{y} \in \mathbb{F}_q^{4t-m}$) and applying our tester to $\hat{f}$. We will typically use u, v, w to denote affine subspaces of dimension t, $2t$, and $4t$ respectively. For any affine subspace $A \subseteq \mathbb{F}_q^m$, let $c_A \in \mathcal{C}^{t\nearrow \dim(A)}$ be the codeword nearest to $r|_A$, breaking ties arbitrarily. Let $\rho_A \triangleq \mathbb{E}_{v\subseteq A}[\delta(r|_v, \mathcal{C}^{t\nearrow 2t})]$, where the expectation is taken over uniformly random $2t$-dimensional subspaces $v \subseteq A$. Fix the following constants:

$$\begin{aligned} \gamma &\triangleq \frac{\alpha_1\delta^2}{40} - \alpha_1 q^{-t} \\ \epsilon &\triangleq \frac{\rho + 2q^{-t}}{\gamma}. \end{aligned}$$

In particular, these constants are chosen so that the following bounds hold:

$$\begin{aligned} 20\delta^{-1}(\alpha_1^{-1}\gamma + q^{-t}) &\leq \frac{\delta}{2} \\ \epsilon &\leq \frac{\delta}{10}. \end{aligned}$$

Overview. This proof is a straightforward generalization of "bootstrapping" proofs originating in the work of Rubinfeld and Sudan [RS96] and which also appears in [ALM+98, AS03, Aro94]. Our writeup in particular follows [Aro94]. For simplicity, assume $t = 1$. Our approach is to define a function $c : \mathbb{F}_q^m \to \mathbb{F}_q$ and then show that it is both close to r and a codeword of $\mathcal{C}^{1\nearrow m}$. The definition of c is simple: for every $\mathbf{x} \in \mathbb{F}_q^m$, consider the opinion $c_u(\mathbf{x})$ for every line u through $\mathbf{x}$, and define $c(\mathbf{x})$ as the majority opinion. We need to show that c is well-defined (the majority is actually a majority). Our main technical lemma (Lemma 6.1.21) of this section shows that most lines agree with each other, so c is well-defined. Lemma 6.1.21 uses Claim 6.1.19, which shows that for a 4-dimensional affine

subspace w, if ρ_w is small, then for every $\mathbf{x} \in w$, most lines $u \subseteq w$ satisfy $c_u(\mathbf{x}) = c_w(\mathbf{x})$. To prove Claim 6.1.19 we use the results of Section 6.1.2, in particular the robustness of the plane test in $m = 4$ dimensions (Corollary 6.1.10). Since the average $\delta(r|_u, c_w|_u)$ over u through $\mathbf{x}$ is about $\delta(r|_w, c_w)$, by robustness this is less than $\alpha_1^{-1}\rho_w$, which is small since ρ_w is small. Therefore, for most u, $\delta(r|_u, c_w|_u)$ is small and so it must be that $c_u = c_w|_u$.

Once we have shown that c is well-defined, showing that c is close to r requires just a bit of calculation. Showing that $c \in \mathcal{C}^{1\nearrow m}$ involves more work. For each line u, define $c'_u \in \mathcal{C}$ to be the nearest codeword to $c|_u$. Fix a line u and a point $\mathbf{x} \in u$. We want to show that $c|_u(\mathbf{x}) = c'_u(\mathbf{x})$. The idea is to show the existence of a "good" 4-dimensional $w \supseteq u$ such that ρ_w is small and for more than $1 - \frac{\delta}{2}$ fraction of points $\mathbf{y} \in u$ (including $\mathbf{x}$) are "good" in the sense that $c(\mathbf{y}) = c_{u'}(\mathbf{y})$ for a non-negligible fraction of lines u' through $\mathbf{y}$. Once we have such a w, we show that for every good $\mathbf{y} \in u$, $c(\mathbf{y}) = c_w(\mathbf{y})$. Since u has more than $1 - \frac{\delta}{2}$ fraction good points, this implies that $\delta(c|_u, c_w|_u) < \frac{\delta}{2}$, hence $c'_u = c|_w$, so $c'_u(\mathbf{x}) = c|_w(\mathbf{x}) = c(\mathbf{x})$, as desired.

Claim 6.1.19. *If $w \subseteq \mathbb{F}_q^m$ be a $4t$-dimensional affine subspace with $\rho_w \leq \gamma$, then for every $\mathbf{x} \in w$, at least $1 - \frac{\delta}{20}$ fraction of t-dimensional subspaces $u \subseteq w$ satisfy $c_u(\mathbf{x}) = c_w(\mathbf{x})$.*

Proof. Fix $\mathbf{x} \in w$. Let U be the set of t-dimensional subspaces u containing $\mathbf{x}$ such that $\delta(r|_u, c_w|_u) < 20\delta^{-1}(\alpha_1^{-1}\rho_w + q^{-t})$. By Corollary 6.1.10, $\mathbb{E}_{\substack{u \subseteq w \\ u \ni \mathbf{x}}}[\delta(r|_u, c_w|_u)] \leq \delta(r|_w, c_w) + q^{-t} \leq \alpha_1^{-1}\rho_w + q^{-t}$, so by Markov's inequality, the probability that $\delta(r|_u, c_w|_u) \geq 20\delta^{-1}(\alpha_1^{-1}\rho_w + q^{-t})$ is at most $\frac{\alpha_1^{-1}\rho_w + q^{-t}}{20\delta^{-1}(\alpha_1^{-1}\rho_w + q^{-t})} = \frac{\delta}{20}$. For $u \in U$, since $\delta(r|_u, c_w|_u) < 20\delta^{-1}(\alpha_1^{-1}\rho_w + q^{-t}) \leq \frac{\delta}{2}$ and $c_w|_u \in \mathcal{C}$, we have $c_u = c_w|_u$ and therefore $c_u(\mathbf{x}) = c_w(\mathbf{x})$. □

The following claim says that $\mathbb{E}_w[\rho_w] \approx \rho$, even if we insist that w contains a fixed t-dimensional subspace.

Claim 6.1.20. *For any t-dimensional affine subspace $u \subseteq \mathbb{F}_q^m$, $\mathbb{E}_{w \supseteq u}[\rho_w] \leq \rho + 2q^{-t}$, where w is a random $4t$-dimensional affine subspace containing u. In particular, for any point $\mathbf{x} \in \mathbb{F}_q^m$, $\mathbb{E}_{w \ni \mathbf{x}}[\rho_w] \leq \rho + 2q^{-t}$.*

Proof. Observe that

$$\begin{aligned}
\rho &= \mathbb{E}_v\left[\delta\left(r|_v, \mathcal{C}^{t\nearrow 2t}\right)\right] \\
&\geq \mathbb{E}_{v:u\cap v=\emptyset}\left[\delta\left(r|_v, \mathcal{C}^{t\nearrow 2t}\right)\right] \cdot \Pr_v[u\cap v=\emptyset] \\
\text{(Lemma B.2.1)} &\geq \mathbb{E}_{v:u\cap v=\emptyset}\left[\delta\left(r|_v, \mathcal{C}^{t\nearrow 2t}\right)\right] \cdot \left(1-q^{-(m-3t)}\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}_{w\supseteq u}[\rho_w] &= \mathbb{E}_{w\supseteq u}\left[\mathbb{E}_{v\subseteq w}\left[\delta(r|_v, \mathcal{C}^{t\nearrow 2t})\right]\right] \\
&\leq \mathbb{E}_{w\supseteq u}\left[\mathbb{E}_{v\subseteq w}\left[\delta(r|_v, \mathcal{C}^{t\nearrow 2t}) \;\middle|\; u\cap v=\emptyset\right] + \Pr_{v\subseteq w}[u\cap v\neq\emptyset]\right] \\
\text{(Lemma B.2.1)} &\leq \mathbb{E}_{w\supseteq u}\left[\mathbb{E}_{v\subseteq w}\left[\delta(r|_v, \mathcal{C}^{t\nearrow 2t}) \;\middle|\; u\cap v=\emptyset\right]\right] + q^{-t} \\
&= \mathbb{E}_{v:u\cap v=\emptyset}[\delta(r|_v, \mathcal{C}^{t\nearrow 2t})] + q^{-t} \\
&\leq \frac{\rho}{1-q^{-(m-3t)}} + q^{-t} \\
&\leq \rho + 2q^{-t}
\end{aligned}$$

□

Lemma 6.1.21 (Main)**.** *For every $\mathbf{x}\in\mathbb{F}_q^m$, there is a collection U_1 of at least $1-\frac{\delta}{5}-\frac{\delta}{600}$ fraction of the t-dimensional affine subspaces through $\mathbf{x}$, such that $c_u(\mathbf{x}) = c_{u'}(\mathbf{x})$ for every $u, u'\in U_1$.*

Proof. Let U be the set of all t-dimensional affine subspaces u through $\mathbf{x}$. Partition U into disjoint collections $U_1,\ldots,U_k$ with $|U_1|\geq\cdots\geq|U_k|$ according to the value of $c_u(\mathbf{x})$. We will show that $\Pr_{u\ni x}[u\in U_1]\geq 1-\frac{\delta}{5}-\frac{\delta}{600}$. For every $4t$-dimensional subspace w, let U_w be the collection of t-dimensional subspaces u through $\mathbf{x}$, guaranteed by Claim 6.1.19, satisfying

$c_u(\mathbf{x}) = c_w(\mathbf{x})$. Then

$$\begin{aligned}
\Pr_{u\ni\mathbf{x}}[u \in U_1] &\geq \Pr_{u,u'\ni\mathbf{x}}[\exists i \;\; u, u' \in U_i] \\
&= \Pr_{u,u'\ni\mathbf{x}}[c_u(\mathbf{x}) = c_{u'}(\mathbf{x})] \\
\text{(Lemma B.2.2)} \quad &\geq \Pr_{u\cap u'=\{\mathbf{x}\}}[c_u(\mathbf{x}) = c_{u'}(\mathbf{x})] - q^{-(m-2t)} \\
&= \mathbb{E}_{w\ni\mathbf{x}}\left[\Pr_{\substack{u,u'\subseteq w \\ u\cap u'=\{\mathbf{x}\}}}[c_u(\mathbf{x}) = c_{u'}(\mathbf{x})]\right] - q^{-(m-2t)} \\
\text{(Lemma B.2.2)} \quad &\geq \mathbb{E}_{w\ni\mathbf{x}}\left[\Pr_{\substack{u,u'\subseteq w \\ u,u'\ni\mathbf{x}}}[c_u(\mathbf{x}) = c_{u'}(\mathbf{x})\right] - q^{-2t} - q^{-(m-2t)} \\
&\geq \mathbb{E}_{w\ni\mathbf{x}}\left[\Pr_{\substack{u,u'\subseteq w \\ u,u'\ni\mathbf{x}}}[c_u(\mathbf{x}) = c_{u'}(\mathbf{x})\right] - \frac{\delta}{600} \\
&\geq \mathbb{E}_{w\ni\mathbf{x}}\left[\Pr_{\substack{u,u'\subseteq w \\ u,u'\ni\mathbf{x}}}[c_u(\mathbf{x}) = c_{u'}(\mathbf{x})] \;\middle|\; \rho_w \leq \gamma\right] \cdot \Pr_{w\ni\mathbf{x}}[\rho_w \leq \gamma] - \frac{\delta}{600} \\
&\geq \mathbb{E}_{w\ni\mathbf{x}}\left[\Pr_{\substack{u,u'\subseteq w \\ u,u'\ni\mathbf{x}}}[u, u' \in U_w] \;\middle|\; \rho_w \leq \gamma\right] \cdot \Pr_{w\ni\mathbf{x}}[\rho_w \leq \gamma] - \frac{\delta}{600} \\
\text{(Claim 6.1.19)} \quad &\geq \left(1 - \frac{\delta}{20}\right)^2 \cdot \Pr_{w\ni\mathbf{x}}[\rho_w \leq \gamma] - \frac{\delta}{600} \\
\text{(Markov)} \quad &\geq \left(1 - \frac{\delta}{20}\right)^2 \cdot \left(1 - \frac{\mathbb{E}_{w\ni\mathbf{x}}[\rho_w]}{\gamma}\right) - \frac{\delta}{600} \\
\text{(Claim 6.1.20)} \quad &\geq \left(1 - \frac{\delta}{20}\right)^2 \cdot \left(1 - \frac{\rho + 2q^{-t}}{\gamma}\right) - \frac{\delta}{600} \\
&\geq 1 - \frac{\delta}{10} - \frac{\rho + 2q^{-t}}{\gamma} - \frac{\delta}{600} \\
&= 1 - \frac{\delta}{10} - \epsilon - \frac{\delta}{600} \\
&\geq 1 - \frac{\delta}{5} - \frac{\delta}{600}
\end{aligned}$$

□

We are now ready to prove the main theorem.

Proof of Theorem 6.1.18. We will define a function $c : \mathbb{F}_q^m \to \mathbb{F}_q$ and then show that it is close to r and is a codeword of $\mathcal{C}^{t\nearrow m}$. For $\mathbf{x} \in \mathbb{F}_q^m$, define $c(\mathbf{x}) \triangleq \text{Majority}_{u \ni \mathbf{x}}\{c_u(\mathbf{x})\}$, where the majority is over t-dimensional affine subspaces u through $\mathbf{x}$. Since $\frac{\delta}{5} + \frac{\delta}{600} < \frac{1}{2}$, it follows from Lemma 6.1.21 that c is well-defined.

Next, we show that c is close to r. Indeed,

$$\begin{aligned}
\rho &= \mathbb{E}_v[\delta(r|_v, c_v)] \\
&\geq \mathbb{E}_u[\delta(r|_u, c_u)] \\
&= \mathbb{E}_u\left[\mathbb{E}_{\mathbf{x}\in u}\left[\mathbb{1}_{c_u(\mathbf{x})\neq r(\mathbf{x})}\right]\right] \\
&= \mathbb{E}_{\mathbf{x}}\left[\mathbb{E}_{u\ni \mathbf{x}}\left[\mathbb{1}_{c_u(\mathbf{x})\neq r(\mathbf{x})}\right]\right] \\
&\geq \mathbb{E}_{\mathbf{x}}\left[\mathbb{E}_{u\ni \mathbf{x}}\left[\mathbb{1}_{c_u(\mathbf{x})\neq r(\mathbf{x})}\right] \;\middle|\; c(\mathbf{x}) \neq r(\mathbf{x})\right] \cdot \Pr_{\mathbf{x}}[c(\mathbf{x}) \neq r(\mathbf{x})] \\
&\geq \mathbb{E}_{\mathbf{x}}\left[\Pr_{u\ni \mathbf{x}}[c_u(\mathbf{x}) = c(\mathbf{x})] \;\middle|\; c(\mathbf{x}) \neq r(\mathbf{x})\right] \cdot \delta(r, c) \\
\text{(Lemma 6.1.21)} \quad &\geq \left(1 - \frac{\delta}{4}\right) \cdot \delta(r, c).
\end{aligned}$$

Finally, we show that $c \in \mathcal{C}^{t\nearrow m}$. Let $u \subseteq \mathbb{F}_q^m$ a t-dimensional affine subspace. We wish to show that $c|_u \in \mathcal{C}$. Let $c'_u \in \mathcal{C}$ be the codeword of $\mathcal{C}$ nearest to $c|_u$ (not to be confused with c_u, the nearest codeword to $r|_u$). Let $\mathbf{x} \in u$. We will show that $c'_u(\mathbf{x}) = c|_u(\mathbf{x})$. For a $4t$-dimensional affine subspace $w \subseteq \mathbb{F}_q^m$, we say a point $\mathbf{y} \in w$ is *good for* w if $\Pr_{\substack{u' \subseteq w \\ u' \ni \mathbf{y}}}[c_{u'}(\mathbf{y}) = c(\mathbf{y})] > \frac{\delta}{20}$. We will show, by a union bound, that there exists a $4t$-dimensional affine subspace $w \supseteq u$ such that

1. $\rho_w \leq \gamma$;
2. $\mathbf{x}$ is good for w;
3. more than $1 - \frac{\delta}{2}$ fraction of points $\mathbf{y} \in u$ are good for w.

Observe that for any $\mathbf{y} \in u$, picking a random $4t$-dimensional w containing u and then picking a random t-dimensional $u' \subseteq w$ through $\mathbf{y}$ that intersect u only on y is equivalent to picking a random t-dimensional u' through $\mathbf{y}$ that intersect u only on y and then picking a

random $4t$-dimensional w containing both u, u'. Therefore, for any fixed $\mathbf{y} \in u$

$$\begin{aligned}
\mathbb{E}_{w \supseteq u}\left[\Pr_{\substack{u' \subseteq w \\ u' \ni \mathbf{y}}}[c_{u'}(\mathbf{y}) \neq c(\mathbf{y})]\right] &= \mathbb{E}_{\substack{w \supseteq u \\ u' \subseteq w, u' \ni \mathbf{y}}}\left[\mathbb{1}_{c_{u'}(\mathbf{y}) \neq c(\mathbf{y})}\right] \\
&\leq \mathbb{E}_{\substack{w \supseteq u \\ u' \subseteq w, u' \ni \mathbf{y}}}\left[\mathbb{1}_{c_{u'}(\mathbf{y}) \neq c(\mathbf{y})} \mid u \cap u' = \{\mathbf{y}\}\right] \\
&\quad + \Pr_{\substack{w \supseteq u \\ u' \subseteq w, u' \ni \mathbf{y}}}[u \cap u' \neq \{\mathbf{y}\}] \\
\text{(Lemma B.2.2)} \quad &\leq \mathbb{E}_{u' \ni \mathbf{y}}\left[\mathbb{1}_{c_{u'}(\mathbf{y}) \neq c(\mathbf{y})} \mid u \cap u' = \{y\}\right] + q^{-2t} \\
\text{(Lemma B.2.2)} \quad &\leq \mathbb{E}_{u' \ni \mathbf{y}}\left[\mathbb{1}_{c_{u'}(\mathbf{y}) \neq c(\mathbf{y})}\right] + q^{-(m-2t)} + q^{-2t} \\
\text{(Lemma 6.1.21 and definition of } c) \quad &\leq \frac{\delta}{5} + \frac{\delta}{600} + 2q^{-2t} \leq \frac{\delta}{5} + \frac{\delta}{300} \leq \frac{\delta}{4}.
\end{aligned}$$

Therefore, by Markov's inequality, for any fixed $\mathbf{y} \in u$,

$$\begin{aligned}
\Pr_{w \supseteq u}[\mathbf{y} \text{ is not good for } w] &= \Pr_{w \supseteq u}\left[\Pr_{u' \supseteq w, u' \ni \mathbf{y}}[c_{u'}(\mathbf{y}) \neq c(\mathbf{y})] \geq 1 - \frac{\delta}{20}\right] \\
&\leq \frac{\frac{\delta}{4}}{1 - \frac{\delta}{20}} \\
&\leq \frac{5}{19} \cdot \delta.
\end{aligned}$$

In particular, this applies for $\mathbf{y} = \mathbf{x}$. Further applying Markov's inequality, we find that

$$\Pr_{w \supseteq u}\left[\text{fraction of not good } \mathbf{y} \text{ in } u \geq \frac{\delta}{2}\right] \leq \frac{5\delta/19}{\delta/2} = \frac{10}{19}.$$

Finally, since $\mathbb{E}_{w \supseteq u}[\rho_w] \leq \rho + 2q^{-t}$ (by Claim 6.1.20), we have

$$\Pr_{w \supseteq u}[\rho_w > \gamma] \leq \frac{\rho + 2q^{-t}}{\gamma} = \epsilon \leq \frac{\delta}{10}.$$

Since $\delta \leq 1$ and $\frac{5}{19} + \frac{10}{19} + \frac{1}{10} < 1$, by the union bound such a desired w exists.

Now that we have such a subspace w, consider c_w. We claim that it suffices to prove that if $\mathbf{y} \in u$ is good, then $c_w(\mathbf{y}) = c(\mathbf{y})$. Indeed, since more than $1 - \frac{\delta}{2}$ fraction of points in u are good, we have $\delta(c_w|_u, c|_u) < \frac{\delta}{2}$. Therefore $c_w|_u = c'_u$, and since $\mathbf{x}$ is good, we

have $c(\mathbf{x}) = c_w(\mathbf{x}) = c'_u(\mathbf{x})$ as desired. It remains to prove that $c_w(\mathbf{y}) = c(\mathbf{y})$ for good $\mathbf{y} \in u$. By Claim 6.1.19, at least $1 - \frac{\delta}{20}$ fraction of t-dimensional $u' \subseteq w$ through $\mathbf{y}$ satisfy $c_{u'}(\mathbf{y}) = c_w(\mathbf{y})$. Since $\mathbf{y}$ is good, more than $\frac{\delta}{20}$ fraction of t-dimensional $u' \subseteq w$ through $\mathbf{y}$ satisfy $c_{u'}(\mathbf{y}) = c(\mathbf{y})$. Therefore, there must be some t-dimensional $u' \subseteq w$ through $\mathbf{y}$ which satisfies $c_w(\mathbf{y}) = c_{u'}(\mathbf{y}) = c(\mathbf{y})$.

Finally, for the robustness statement: if $q^{-t} \geq \frac{\delta^{24}}{10^{14}}$, then by Corollary 6.1.2, the robustness is at least $\frac{q^{-3t}}{2} \geq \frac{\delta^{72}}{2 \cdot 10^{52}}$. Otherwise, the robustness is at least $\frac{\alpha_1 \delta^3}{57{,}600 \cdot 400} - 3q^{-t} \geq \frac{\delta^{24}}{2 \cdot 10^{14}}$. □

6.2 Technical Algebraic Results

The purpose of this section is to prove Theorem 6.2.6 and its Corollaries 6.2.7 and 6.2.8. If we allow our robustness in Theorem 6.1.18 to depend on t, the dimension of the base code, then proving what we need for Theorem 6.2.6 is easy. However, removing the dependence on t requires some new ideas, including the definition of a new operation ("degree lifting") on codes, and the analysis of the distance of degree lifted codes. In Section 6.2.1, we define degree lifting and analyze the degree lifted codes (Proposition 6.2.4). In Section 6.2.2, we prove Theorem 6.2.6 and its corollaries.

6.2.1 Degree Lift

In this section, we define the degree lift operation on codes with degree sets. The operation can be thought of as "Reed-Mullerization", in the sense that the degree lift of the Reed-Solomon code of degree d is the Reed-Muller code of degree d. This resembles the degree lift operation of Ben-Sasson et al. [BGK+13] who defined a "Reed-Mullerization" for algebraic-geometry codes (in contrast, we want to define it for codes over $\mathbb{F}_q^m$ spanned by monomials).

Definition 6.2.1 (Degree lift)**.** Let $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ have degree set $\mathrm{Deg}(\mathcal{C})$. For positive integer $s \geq 1$, define the *s-wise degree lift* $\mathcal{C}(s) \subseteq \{\mathbb{F}_q^{ms} \to \mathbb{F}_q\}$ of $\mathcal{C}$ to be the code with

degree set

$$\mathrm{Deg}(\mathcal{C}(s)) \triangleq \left\{ (\mathbf{d}_1, \ldots, \mathbf{d}_s) \in \{0, 1, \ldots, q-1\}^{m \times s} \;\middle|\; \sum_{j=1}^{s} \mathbf{d}_j \in \mathrm{Deg}(\mathcal{C}) \right\}.$$

Our goal with this definition is to prove Proposition 6.2.4, which says that the distance of $\mathcal{C}(s)$ is nearly the same as the distance of $\mathcal{C}$. One can show that $\delta(\mathcal{C}(s)) \geq \delta(\mathcal{C}) - mq^{-1}$ To do so, we will use the following fact.

Proposition 6.2.2. *Let $t, n \geq 1$ and let $m = nt$. Let $\mathcal{C}_0 \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant and let $\mathcal{C} \triangleq (\mathcal{C}_0)^{\otimes n}$. For each $i \in [n]$, let $\mathbf{X}_i = (X_{i1}, \ldots, X_{it})$. If $f(\mathbf{X}_1, \ldots, \mathbf{X}_n) \in \mathcal{C}$, and $A_1, \ldots, A_n : \mathbb{F}_q^t \to \mathbb{F}_q^t$ are affine transformations, then $f(A_1(\mathbf{X}_1), \ldots, A_n(\mathbf{X}_n)) \in \mathcal{C}$.*

Proof. By linearity, it suffices to consider the case where $f(\mathbf{X}_1, \ldots, \mathbf{X}_n) = \prod_{i=1}^n \mathbf{X}_i{}^{\mathbf{d}_i}$ is a monomial, where $\mathbf{d}_i = (d_{i1}, \ldots, d_{it}) \in \{0, 1, \ldots, q-1\}^t$. Each $\mathbf{X}_i{}^{\mathbf{d}_i} \in \mathcal{C}_0$, so by affine-invariance $A_i(\mathbf{X}_i)^{\mathbf{d}_i} \in \mathcal{C}_0$. Therefore, by Proposition A.2.2, $f(A_1(\mathbf{X}_1), \ldots, A)_n(\mathbf{X}_n)) = \prod_{i=1}^n A_i(\mathbf{X}_i)^{\mathbf{d}_i} \in (\mathcal{C}_0)^{\otimes n} = \mathcal{C}$. □

Overview. To prove Proposition 6.2.4, we show, through Lemma 6.2.3, that there is a special subset of m-dimensional subspaces A, such that for any $f \in \mathcal{C}(s)$, $f|_A \in \mathcal{C}$. Then, we analyze the distance of from f to the zero function by looking at the distance on a random special m-dimensional A. This will yield a distance of $\delta(\mathcal{C}(s)) \geq \delta(\mathcal{C}) - o(1)$ as long as the special subspaces sample $\mathbb{F}_q^m$ well. However, we require the $o(1)$ term to be $(nq^{-t})^{O(1)}$, otherwise we would not be able to remove the dependence on t in the robustness of Theorem 6.1.18. In order to do so, we need to further assume that $\mathcal{C}$ is the tensor product $(\mathcal{C}_0)^n$ of some t-dimensional code $\mathcal{C}_0$ (which is satisfied by our use case).

We now describe the special subspaces we consider in Lemma 6.2.3. Label the variables of $\mathbb{F}_q^{ms} = \mathbb{F}_q^{nts}$ by X_{cij}, where $c \in [n]$, $i \in [t]$, $j \in [s]$. Let Y_{ci}, for $c \in [n]$, $i \in [t]$, be the variables parameterizing A. Note that an arbitrary subspace restriction corresponds to substituting, for each X_{cij}, an affine function of all of the variables $Y_{11}, \ldots, Y_{nt}$. This is too much to hope for. However, if we substitute for X_{cij} an affine function of just $Y_{c1}, \ldots, Y_{ct}$, this works.

Lemma 6.2.3. *Let $t, n \geq 1$ and $m = nt$. Let $\mathcal{C}_0 \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant and let $\mathcal{C} \triangleq (\mathcal{C}_0)^{\otimes n}$. Let $s \geq 1$, and let $f(\mathbf{X}) \in \mathcal{C}(s)$, with variables $\mathbf{X} = (X_{cij})_{c\in[n],i\in[t],j\in[s]}$. Let $g(Y_{11}, \ldots, Y_{nt})$ be the m-variate polynomial obtained from $f(\mathbf{X})$ by setting, for each $c \in [n], i \in [t]$, and $j \in [s]$, $X_{cij} = \sum_{k=1}^{t} a_{cijk} Y_{ck} + b_{cij}$, for some $a_{cijk}, b_{cij} \in \mathbb{F}_q$. That is, for all $(c, i, j) \in [n] \times [t] \times [s]$ X_{cij} is an affine function of $Y_{c1}, \ldots, Y_{ct}$. Then $g \in \mathcal{C}$.*

Proof. By linearity, it suffices to consider the case where $f(\mathbf{X}) = \prod_{c=1}^{n} \prod_{i=1}^{t} \prod_{j=1}^{s} X_{cij}^{d_{cij}}$ is a monomial, for some $0 \leq d_{cij} \leq q-1$. For each $j \in [s]$, define $\mathbf{d}_j \triangleq (d_{11j}, \ldots, d_{ntj})$, so that $(\mathbf{d}_1, \ldots, \mathbf{d}_s) \in \mathrm{Deg}(\mathcal{C}(s))$, i.e. $\sum_{i=1}^{s} \mathbf{d}_i \in \mathrm{Deg}(\mathcal{C})$. Then

$$\begin{aligned} g(Y_{11}, \ldots, Y_{nt}) &= \prod_{c=1}^{n} \prod_{i=1}^{t} \prod_{j=1}^{s} \left(\sum_{k=1}^{t} a_{cijk} Y_{ck} + b_{cij} \right)^{d_{cij}} \\ &= \prod_{c=1}^{n} \prod_{i=1}^{t} \prod_{j=1}^{s} \sum_{\mathbf{e}_{cij} \leq d_{cij}} \binom{d_{cij}}{\mathbf{e}_{cij}} b_{cij}^{e_{cij0}} \prod_{k=1}^{t} a_{cijk}^{e_{cijk}} Y_{ck}^{e_{cijk}} \\ &= \sum_{\substack{\mathbf{e}_{cij} \leq_p d_{cij} \\ \forall\ i,j}} (\cdots) \prod_{c=1}^{n} \prod_{k=1}^{t} Y_{ck}^{\sum_{i=1}^{t} \sum_{j=1}^{s} e_{cijk}} \end{aligned}$$

where the $(\cdots)$ denotes constants in $\mathbb{F}_q$. So, it suffices to show that each monomial of the form $\prod_{c=1}^{n} \prod_{k=1}^{t} Y_{ck}^{\sum_{i=1}^{t} \sum_{j=1}^{n} e_{cijk}} \in \mathcal{C}$, which we show in the remainder of the proof.

Let $h(Y_{11}, \ldots, Y_{nt})$ be the m-variate polynomial obtained from f by substituting $X_{cij} = Y_{ci}$ for each $c \in [n]$, $i \in [t]$, $j \in [s]$. Then $h(Y_{11}, \ldots, Y_{nt}) = \prod_{c=1}^{n} \prod_{i=1}^{t} \prod_{j=1}^{s} Y_{ci}^{d_{cij}} = \prod_{c=1}^{n} \prod_{i=1}^{t} Y_{ci}^{\sum_{j=1}^{s} d_{cij}}$ is a monomial with degree $\left(\sum_{j=1}^{s} d_{11j}, \ldots, \sum_{j=1}^{s} d_{ntj} \right) = \sum_{j=1}^{s} \mathbf{d}_j \in \mathrm{Deg}(\mathcal{C})$, hence $h \in \mathcal{C}$. Now, consider applying an affine transformation as follows: for each $1 \leq c \leq n$ and each $1 \leq i \leq t$, substitute $Y_{ci} \leftarrow \sum_{k=1}^{t} \alpha_{cik} Y_{ck} + \beta_{ci}$, and call the new

polynomial h'. By Proposition 6.2.2, $h' \in \mathcal{C}$. On the other hand,

$$\begin{aligned}
h'(Y_1, \ldots, Y_m) &= \prod_{c=1}^{n} \prod_{i=1}^{t} \left(\sum_{k=1}^{t} \alpha_{cik} Y_{ck} + \beta_{ci} \right)^{\sum_{j=1}^{s} d_{cij}} \\
&= \prod_{c=1}^{n} \prod_{i=1}^{t} \prod_{j=1}^{s} \left(\sum_{k=1}^{t} \alpha_{cik} Y_{ck} + \beta_{ci} \right)^{d_{cij}} \\
&= \prod_{c=1}^{n} \prod_{i=1}^{t} \prod_{j=1}^{s} \sum_{\mathbf{e}_{cij} \leq_p d_{cij}} \binom{d_{cij}}{\mathbf{e}_{cijk}} \beta_{ci}^{e_{cij0}} \prod_{k=1}^{t} \alpha_{cik}^{e_{cijk}} Y_{ck}^{e_{cijk}} \\
&= \sum_{\substack{\mathbf{e}_{cij} \leq_p d_{cij} \\ \forall\ c,i,j}} \left(\prod_{c,i,j} \binom{d_{cij}}{\mathbf{e}_{cijk}} \beta_{ci}^{e_{cij0}} \right) \prod_{c=1}^{n} \prod_{k=1}^{t} \left(\prod_{i=1}^{t} \alpha_{cik}^{\sum_{j=1}^{s} e_{cijk}} \right) Y_{ck}^{\sum_{i=1}^{t} \sum_{j=1}^{s} e_{cijk}}
\end{aligned}$$

and since the α_{cik} and the β_{ci} are arbitrary and $\mathcal{C}$ has a degree set $\mathrm{Deg}(\mathcal{C}) = \mathrm{Deg}(\mathcal{C}_0)^n$, each monomial $\prod_{c=1}^{n-} \prod_{k=1}^{t} Y_{ck}^{\sum_{i=1}^{t} \sum_{j=1}^{n} e_{cijk}} \in \mathcal{C}$, as desired. □

Proposition 6.2.4. *Let $t, n \geq 1$ and $m = nt$. Let $\mathcal{C}_0 \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant and let $\mathcal{C} \triangleq (\mathcal{C}_0)^{\otimes n}$. For any positive integer $s \geq 1$, $\delta(\mathcal{C}(s)) \geq \delta(\mathcal{C}) - nq^{-t} = \delta(\mathcal{C}_0)^n - nq^{-t}$.*

Proof. Let $f(\mathbf{X}) \in \mathcal{C}(s)$ be a nonzero codeword with variables $\mathbf{X} = (X_{cij})_{c \in [n], i \in t], j \in [s]}$. For $c \in [n]$, $i \in [m]$, $j \in [s]$, $k \in [t]$, let $a_{cijk}, b_{cij} \in \mathbb{F}_q$, and let $\mathbf{a} \triangleq (a_{cijk})_{c \in [n], i \in [t], j \in [s], k \in [t]}$ and $\mathbf{b} \triangleq (b_{cij})_{c \in [n], i \in [t], j \in [s]}$. Let $g_{\mathbf{a},\mathbf{b}}(Y_{11}, \ldots, Y_{nt})$ be the m-variate polynomial obtained by setting $X_{cij} = \sum_{k=1}^{t} a_{cijk} Y_{ck} + b_{cij}$ for each $1 \leq c \leq n$ and $1 \leq i \leq t$.

By linearity of $\mathcal{C}$ and thus of $\mathcal{C}(s)$, it suffices to show that $\delta(f, 0) \geq \delta(\mathcal{C}) - nq^{-t}$. Let $\mathbf{b} \in \mathbb{F}_q^{nts}$ be a point such that $f(\mathbf{b}) \neq 0$. Consider choosing $\mathbf{a}$ uniformly at random. Then $g_{\mathbf{a},\mathbf{b}} \neq 0$ since $g_{\mathbf{a},\mathbf{b}}(\mathbf{0}) = f(\mathbf{b}) \neq 0$. For fixed $y_{11}, \ldots, y_{nt}$, as long as for each $c \in [n]$ there is some $k \in [t]$ such that $y_{ck} \neq 0$, then the points $\sum_{k=1}^{t} a_{cijk} y_{ck} + b_{cij}$ are independent and

uniform over $\mathbb{F}_q$. This occurs with probability at least $1 - nq^{-t}$. Therefore,

$$\begin{aligned}
\delta(\mathcal{C}) &\leq \mathbb{E}_{\mathbf{a}}\left[\delta\left(g_{\mathbf{a},\mathbf{b}}, 0\right)\right] \\
&= \mathbb{E}_{\mathbf{a}}\left[\mathbb{E}_{\mathbf{y}}\left[\mathbb{1}_{g_{\mathbf{a},\mathbf{b}}(\mathbf{y})\neq 0}\right]\right] \\
&= \mathbb{E}_{\mathbf{y}}\left[\mathbb{E}_{\mathbf{a}}\left[\mathbb{1}_{g_{\mathbf{a},\mathbf{b}}(\mathbf{y})\neq 0}\right]\right] \\
&\leq nq^{-t} + \mathbb{E}_{\mathbf{y}\neq\mathbf{0}}\left[\mathbb{1}_{g_{\mathbf{a},\mathbf{b}}(\mathbf{y})\neq 0}\right] \\
&= nq^{-t} + \mathbb{E}_{\mathbf{y}\neq\mathbf{0}}\left[\mathbb{1}_{f((\sum_{k=1}^{t} a_{cijk} y_{ck} + b_{cij}))\neq 0}\right] \\
&= nq^{-t} + \delta(f, 0).
\end{aligned}$$

□

6.2.2 Analysis of Subspace Restrictions

In this section we prove Theorem 6.2.6 and its corollaries.

Overview. Corollary 6.2.7 says that if a codeword f of the tensor product $\mathcal{C}^{\otimes n}$ of a t-dimensional code $\mathcal{C}$ is not a codeword of $\mathcal{C}^{t\nearrow nt}$, then on there is a point $\mathbf{b}$ such that on many t-dimensional subspaces u through $\mathbf{b}$, the restriction $f|_u \notin \mathcal{C}$. We use this in the proof of Theorem 6.1.4 when arguing that if a tensor codeword $c \in \mathcal{C}^{\otimes m}$ satisfies $c \in \mathcal{C}_{\mathbf{a}}$ (see overview) for many $\mathbf{a}$, then $c \in \bigcap_{\mathbf{a}} \mathcal{C}_{\mathbf{a}} = \mathcal{C}^{1\nearrow m}$. A special case of Corollary 6.2.8 says that if f is a lifted Reed-Solomon codeword but not a Reed-Muller codeword, then on many planes f is not a bivariate Reed-Muller code. The actual corollary merely generalizes this to arbitrary t and codes $\mathcal{C}_0, \mathcal{C}_1$.

Both Corollaries 6.2.7 and 6.2.8 are proved in a similar manner. Note that both are statements of the form "if f is in some big code but not in a lifted code, then on many subspaces it is not a codeword of the base code". A natural approach is to write f out as a linear combination of monomials, restrict to an arbitrary subspace of the appropriate dimension, re-write the restriction as a linear combination of monomials in the parameterizing variables, and note that the coefficients of the monomials are functions in the parameterization coeffi-

cients. Since f is not in the lift, there is a monomial outside the base code whose coefficient (the "offending coefficient") is a nonzero function. Then, one shows that these functions belong to a code with good distance, so for many choices of parameterizing coefficients, the offending coefficient is nonzero.

Theorem 6.2.6 abstracts the above approach and shows that, in the case of Corollary 6.2.7, the offending coefficient is a codeword of the degree lift $(\mathcal{C}^{\otimes n})(t)$ of $\mathcal{C}^{\otimes n}$, and in the case of Corollary 6.2.8, the offending coefficient is a codeword of a lifted code. This necessitates the analysis of the distance of degree lifted codes, hence the need for Section 6.2.1.

Lemma 6.2.5. *Let $\mathcal{C} \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ be a linear code with a p-shadow-closed degree set. If $f \in \mathcal{C}$, and*

$$f\left(a_{10} + \sum_{j=1}^{t} a_{1j} Y_j, \ldots, a_{m0} + \sum_{j=1}^{t} a_{mj} Y_j\right) = \sum_{\mathbf{e} \in \{0,1,\ldots,q-1\}^t} f_{\mathbf{e}}(\mathbf{a}) \cdot \mathbf{Y}^{\mathbf{e}}$$

where $\mathbf{a} = (a_{ij})_{1 \le i \le m; 0 \le j \le t} \in \mathbb{F}_q^{m(t+1)}$, then, for every $\mathbf{e} \in \{0, 1, \ldots, q-1\}^t$,

$$f_{\mathbf{e}}(\mathbf{a}) = \sum_{\substack{\mathbf{d} \in D \\ \mathbf{E} \le_p \mathbf{d} \\ \|\mathbf{E}_{*j}\| = e_j \ \forall j}} f_{\mathbf{d}} \cdot \prod_{i=1}^{m} \binom{d_i}{\mathbf{E}_{i*}} a_{i0}^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}}.$$

In particular,

1. *$f_{\mathbf{e}} \in \mathcal{C}(t+1)$, the $(t+1)$-wise degree lift of $\mathcal{C}$ (see Definition 6.2.1);*

2. *if $\mathcal{C} = (\mathcal{C}_0)^{1 \nearrow m}$ for some linear affine-invariant code $\mathcal{C}_0 \subseteq \{\mathbb{F}_q \to \mathbb{F}_q\}$, then $f_{\mathbf{e}} \in (\mathcal{C}_0)^{1 \nearrow m(t+1)}$*

Proof. Let D be the degree set of $\mathcal{C}$. Write $f(\mathbf{X}) = \sum_{\mathbf{d} \in D} f_{\mathbf{d}} \cdot \mathbf{X}^{\mathbf{d}}$. Let $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ be the

affine map $\mathbf{Y} \mapsto \left(a_{10} + \sum_{j=1}^{t} a_{1j}Y_j, \ldots, a_{m0} + \sum_{j=1}^{t} a_{mj}Y_j\right)$. Expanding, we get

$$\begin{aligned}
(f \circ A)(\mathbf{Y}) &= \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot \prod_{i=1}^{m} \left(a_{i0} + \sum_{j=1}^{t} a_{ij}Y_j\right)^{d_i} \\
&= \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot \prod_{i=1}^{m} \left(\sum_{\mathbf{e}_i \le_p d_i} \binom{d_i}{\mathbf{e}_i} a_{i0}^{e_{i0}} \cdot \prod_{j=1}^{t} a_{ij}^{e_{ij}} Y_j^{e_{ij}}\right) \\
&= \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot \sum_{\mathbf{E} \le_p \mathbf{d}} \left(\prod_{i=1}^{m} \binom{d_i}{\mathbf{E}_{i*}} a_{i0}^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}}\right) \cdot \prod_{j=1}^{t} Y_j^{\|\mathbf{E}_{*j}\|} \\
&= \sum_{\mathbf{e}\in\{0,1,\ldots,q-1\}^t} \mathbf{Y}^{\mathbf{e}} \cdot \sum_{\substack{\mathbf{d}\in D \\ \mathbf{E}\le_p \mathbf{d} \\ \|\mathbf{E}_{*j}\| \bmod^* q = e_j \ \forall j}} f_{\mathbf{d}} \cdot \prod_{i=1}^{m} \binom{d_i}{\mathbf{E}_{i*}} a_{i0}^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}}
\end{aligned}$$

and therefore, for each $\mathbf{e} \in \{0, 1, \ldots, q-1\}^t$,

$$f_{\mathbf{e}}(\mathbf{a}) = \sum_{\substack{\mathbf{d}\in D \\ \mathbf{E}\le_p \mathbf{d} \\ \|\mathbf{E}_{*j}\| \bmod^* q = e_j \ \forall j}} f_{\mathbf{d}} \cdot \prod_{i=1}^{m} \binom{d_i}{\mathbf{E}_{i*}} a_{i0}^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}}.$$

View the variables $\mathbf{a} = (a_{ij})$ in the order $(a_{10}, \ldots, a_{m0}, \ \ldots \ , a_{1t}, \ldots, a_{mt})$, and interpret $\mathbf{E}$ as $(\mathbf{E}_{*0}, \ldots, \mathbf{E}_{*t})$. If $\mathbf{E} \le_p \mathbf{d}$ and $\mathbf{d} \in D = \mathrm{Deg}(\mathcal{C})$, then $\mathbf{E} \in \mathrm{Deg}(\mathcal{C}(t+1))$. Therefore, $f_{\mathbf{e}} \in \mathcal{C}(t+1)$.

Now suppose $\mathcal{C} = (\mathcal{C}_0)^{1\nearrow m}$ for some linear affine-invariant $\mathcal{C}_0 \subseteq \{\mathbb{F}_q \to \mathbb{F}_q\}$. It suffices to show that if $\mathbf{d} = (d_1, \ldots, d_m) \in \mathrm{Deg}(\mathcal{C})$ and $\mathbf{E} \le_p \mathbf{d}$ with entries e_{ij}, $i \in [m]$, $0 \le j \le t$, then the length $m(t+1)$ vector $(\mathbf{E}_{*0}, \mathbf{E}_{*1}, \ldots, \mathbf{E}_{*t}) \in \mathrm{Deg}((\mathcal{C}_0)^{1\nearrow m(t+1)})$. By Proposition 5.2.3, it suffices to show that, if $u_{ij} \le_p e_{ij}$ for every $i \in [m]$ and $0 \le j \le t$, then $\sum_{ij} u_{ij} \bmod^* q \in \mathrm{Deg}(\mathcal{C}_0)$. Since $\mathbf{d} \in \mathrm{Deg}(\mathcal{C}) = \mathrm{Deg}((\mathcal{C}_0)^{1\nearrow m})$, this implies that if $e'_i \le_p d_i$ for $i \in [m]$, then $\sum_i e'_i \bmod^* q \in \mathrm{Deg}(\mathcal{C}_0)$. Set $e'_i \triangleq \sum_{j=0}^{t} u_{ij}$. Observe that, since $(e_{i0}, e_{i1}, \ldots, e_{it}) \le_p d_i$, this implies that $e'_i \le_p d_i$. Therefore, $\sum_{ij} u_{ij} \bmod^* q = \sum_i e'_i \bmod^* q \in \mathrm{Deg}(\mathcal{C}_0)$, as desired. □

Theorem 6.2.6. *Let $1 \le t < m$. Let $\mathcal{C}_1 \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be a linear affine-invariant code, and let $\mathcal{C}_2 \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ have a p-shadow-closed degree set. Suppose $f \in \mathcal{C}_2 \setminus {\mathcal{C}_1}^{t\nearrow m}$. Then the*

following hold:

1. *if* $\mathcal{C}_2 = (\mathcal{C}_0)^{\otimes n}$ *for some linear affine-invariant code* $\mathcal{C}_0 \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$*, where* $m = nt$*, then there exists a point* $\mathbf{b} \in \mathbb{F}_q^m$ *such that for at least* $\delta(\mathcal{C}_0)^n - (n+1)q^{-t}$ *fraction of* t*-dimensional affine subspaces* $A \subseteq \mathbb{F}_q^m$ *passing through* $\mathbf{b}$*, the restriction* $f|_A \notin \mathcal{C}_1$*;*

2. *if* $\mathcal{C}_2 = (\mathcal{C}_0)^{1\nearrow m}$ *for some linear affine-invariant code* $\mathcal{C}_0 \subseteq \{\mathbb{F}_q \to \mathbb{F}_q\}$*, then for at least* $\delta(\mathcal{C}_0) - q^{-1}$ *fraction of* t*-dimensional affine subspaces* $A \subseteq \mathbb{F}_q^m$*, the restriction* $f|_A \notin \mathcal{C}_1$*.*

Proof. Let p be the characteristic of $\mathbb{F}_q$. Let A be parameterized by $X_i = a_{i0} + \sum_{j=1}^t a_{ij} Y_j$, where the matrix $\{a_{ij}\}_{i=1,j=1}^{m,t} \in \mathbb{F}_q^{m \times t}$ has full rank. Write

$$f|_A(\mathbf{Y}) = \sum_{\mathbf{e} \in \{0,1,\ldots,q-1\}^t} f_{\mathbf{e}}(\mathbf{a}) \cdot \mathbf{Y}^{\mathbf{e}}.$$

Since $f \notin \mathcal{C}_1^m$, there exists $\mathbf{e} \notin \mathrm{Deg}(\mathcal{C}_1)$ such that $f_{\mathbf{e}} \neq 0$.

1. By Corollary A.2.3, $\mathcal{C}_2$ has a p-shadow-closed degree set. By Lemma 6.2.5 (1), $f_{\mathbf{e}} \in \mathcal{C}_2(t+1)$. For each $\mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{F}_q^m$, let $f_{\mathbf{e},\mathbf{b}}$ denote the polynomial $f_{\mathbf{e}}$ with the variable a_{i0} fixed to value b_i for each $i \in [m]$ (i.e. insisting that A passes through $\mathbf{b}$). Observe that each $f_{\mathbf{e},\mathbf{b}} \in \mathcal{C}_2(t)$. Since $f_{\mathbf{e}} \neq 0$, there exists $\mathbf{b} \in \mathbb{F}_q^m$ such that $f_{\mathbf{e},\mathbf{b}} \neq 0$. By Proposition 6.2.4, for at least $\delta(\mathcal{C}_2(t)) \geq \delta(\mathcal{C}_0)^n - nq^{-t}$ fraction of matrices $\{a_{ij}\}_{i\in[m];j\in[t]}$, we have $f_{\mathbf{e},\mathbf{b}}(\{a_{ij}\}_{i\in[m];j\in[t]}) \neq 0$. Since, by Lemma B.2.2, at least $1 - q^{t-m}$ fraction of such matrices have full rank, we get that for at least $\delta(\mathcal{C}_0)^n - nq^{-t} - q^{t-m} \geq \delta(\mathcal{C}_0)^n - (n+1)q^{-t}$ of the full rank matrices satisfy $f_{\mathbf{e},\mathbf{b}}\left(\{a_{ij}\}_{i=1,j=1}^{m,t}\right) \neq 0$, and therefore $f|_A(\mathbf{Y}) \notin \mathcal{C}_1$.

2. By Proposition 4.1.3 it has a p-shadow-closed degree set. By Lemma 6.2.5 (2), $f_{\mathbf{e}} \in (\mathcal{C}_0)^{1\nearrow m(t+1)}$, so $f_{\mathbf{e}}(\mathbf{a}) \neq 0$ for at least $\delta((\mathcal{C}_0)^{1\nearrow m(t+1)}) \geq \delta(\mathcal{C}_0) - q^{-1}$ fraction of choices $\mathbf{a}$ (including such that the corresponding matrix does not have full rank), and therefore, by Lemma B.2.2, $f|_A(\mathbf{Y}) \notin \mathcal{C}_1$ for at least $\delta(\mathcal{C}_0) - q^{-1} - q^{t-m} \geq \delta(\mathcal{C}_0) - 2q^{-1}$.

□

Corollary 6.2.7. *Let $t, n \geq 1$ and let $m = nt$. Let $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be a linear affine-invariant code. If $f \in \mathcal{C}^{\otimes n} \setminus \mathcal{C}^{t\nearrow m}$, then there is a point $\mathbf{b} \in \mathbb{F}_q^m$ such that for $\delta(\mathcal{C})^n - (n+1)q^{-t}$ fraction of t-dimensional subspaces u through $\mathbf{b}$, the restriction $f|_u \notin \mathcal{C}$.*

Proof. Follows immediately from Theorem 6.2.6 (1) with $\mathcal{C}_0 = \mathcal{C}_1 = \mathcal{C}$, and $\mathcal{C}_2 = \mathcal{C}^{\otimes n}$. □

Corollary 6.2.8. *Let $1 \leq t \leq m$. Let $\mathcal{C}_0 \subseteq \{\mathbb{F}_q \to \mathbb{F}_q\}$ be a linear affine-invariant code. Let $\mathcal{C}_1 \subsetneq (\mathcal{C}_0)^{1\nearrow t}$ be a linear affine-invariant code that is a strict subcode of $(\mathcal{C}_0)^{1\nearrow t}$. If $f \in (\mathcal{C}_0)^{1\nearrow m} \setminus (\mathcal{C}_1)^{t\nearrow m}$, then for at least $\delta(\mathcal{C}_0) - 2q^{-1}$ fraction of t-dimensional subspaces $A \subseteq \mathbb{F}_q^m$, the restriction $f|_A \notin \mathcal{C}_1$.*

Proof. Follows immediately from Theorem 6.2.6 (2) with $\mathcal{C}_2 = (\mathcal{C}_0)^{1\nearrow m}$. □

Chapter 7

Applications

7.1 Lifted Reed-Solomon Code

We begin by describing the central construction of the thesis — the lifted Reed-Solomon. As its name suggests, the lifted Reed-Solomon is simply obtained by lifting the Reed-Solomon code. With judicious choice of parameters, the lifted Reed-Solomon code is the first code to achieve a combination of parameters never achieved by one code before. For now, we describe our construction for an arbitrary choice of parameters.

Let q be a prime power, let $m \geq 2$ be an integer, and let $d < q$. Let $\mathrm{RS} \triangleq \mathrm{RS}(q, d)$ and let $\mathcal{C} \triangleq \mathrm{RS}^{1 \nearrow m}$. The code $\mathcal{C}$ is the *lifted Reed-Solomon code*. It automatically inherits distance, a well-structured degree set, local correctability, decodability, and robust testability by virtue of being a lifted code.

Proposition 7.1.1. *The distance of $\mathcal{C}$ is $\delta(\mathcal{C}) \geq 1 - \frac{d+1}{q}$.*

Proof. By Proposition 5.2.4, $\delta(\mathcal{C}) \geq \delta(\mathrm{RS}) - q^{-1}$, and by Proposition 3.4.2, $\delta(\mathrm{RS}) = 1 - \frac{d}{q}$, so $\delta(\mathcal{C}) \geq \delta(\mathrm{RS}) - q^{-1} \geq 1 - \frac{d+1}{q}$. □

Proposition 7.1.2. *The code $\mathcal{C}$ is linear affine-invariant and $\mathbf{d} \in \mathrm{Deg}(\mathcal{C})$ if and only if, for every $\mathbf{e} \leq_p \mathbf{d}$, the sum $\sum_{i=1}^{m} e_i \bmod^* q \leq d$.*

Proof. Follows immediately from Proposition 5.2.3. □

Proposition 7.1.3. *For every $\epsilon, \eta > 0$, the code $\mathcal{C}$ is $\left(Q, (\frac{1}{2} - \epsilon)\delta - q^{-1}, \eta\right)$-locally correctable and decodable, where $\delta \triangleq 1 - \frac{d}{q}$ and $Q = O(\ln(1/\eta)/\epsilon^2)$.*

Proof. The local correctability follows immediately from Theorem 5.3.2 while local decodability follows from Theorem 5.3.4. □

Proposition 7.1.4. *Let $\delta \triangleq 1 - \frac{d}{q}$. The code $\mathcal{C}$ has a q^2-local tester that is $\frac{\delta^{72}}{2 \cdot 10^{52}}$-robust.*

Proof. Follows immediately from Theorem 5.4.3. □

7.1.1 Relationship to Reed-Muller

We proceed by examining the relationship between the lifted Reed-Solomon code and the Reed-Muller code. It follows immediately from definitions that the Reed-Muller code is contained in the lifted Reed-Solomon code, i.e. $\mathrm{RM}(q, d, m) \subseteq \mathrm{RS}(q, d)^{1 \nearrow m}$. Kaufman and Ron [KR06] proved the following characterization: a polynomial over $\mathbb{F}_q^m$ has degree d if and only if on every t-dimensional affine subspace its restriction has degree d, where $t = \left\lceil \frac{d+1}{q - q/p} \right\rceil$ and p is the characteristic of $\mathbb{F}_q$. This generalizes a result of Friedl and Sudan [FS95], which says that if $d \leq q - q/p$ and a polynomial has degree d on every line, then its global degree is d, and that this is not necessarily true if $d \geq q - q/p$. Observe that, in the language of lifting, this statement simply says that $\mathrm{RS}(q, d)^{1 \nearrow m} = \mathrm{RM}(q, d, m)$ if and only if $d < q - q/p$. In Theorem 7.1.6, we give a more efficient proof of the above result of [FS95] using the technology of affine-invariance and lifting. In Theorem 7.1.8, we prove a specialization of the characterization of [KR06]: if $q - q/p \leq d < q$, then polynomials of degree d are characterized by having degree d on planes, i.e. $\mathrm{RM}(q, d, 2)^{2 \nearrow m} = \mathrm{RM}(q, d, m)$, again using affine-invariance and lifting.

First, we prove a handy lemma.

Lemma 7.1.5. *Let $m \geq t$, let q be a prime power, and let $d < q$. If $\mathrm{RM}(q, d, m) = \mathrm{RM}(q, d, t)^{t \nearrow m}$, then $\mathrm{RM}(q, d-1, m) = \mathrm{RM}(q, d-1, t)^{t \nearrow m}$.*

Proof. Let $f \in \mathrm{RM}(q, d-1, t)^{t \nearrow m}$. Define $g : \mathbb{F}_q^m \to \mathbb{F}_q$ by $g(\mathbf{X}) \triangleq X_1 \cdot f(\mathbf{X})$. Since $f \in \mathrm{RM}(q, d-1, t)^{t \nearrow m}$, by Proposition 5.1.2 it follows that $\deg(f \circ A) \leq d - 1 < q - 1$

for every affine $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ Note that $(g \circ A)(\mathbf{Y}) = g(A(\mathbf{Y})) = (A(\mathbf{Y}))_1 \cdot (f \circ A)(\mathbf{Y})$, so $\deg(g \circ A) \leq \deg(f \circ A) + 1 \leq d$ for every affine $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$, hence $g \in \mathrm{RM}(q, d, t)^{t \nearrow m} = \mathrm{RM}(q, d, m)$, so $\deg(g) \leq d$. Therefore, $\deg(f) = \deg(g) - 1 \leq d - 1$. □

Theorem 7.1.6. *Let p be a prime and let q be a power of p. Let $d < q$. Let $m \geq 2$. Then $\mathrm{RM}(q, d, m) = \mathrm{RS}(q, d)^{1 \nearrow m}$ if and only if $d < q - q/p$.*

Proof. By Lemma 7.1.5, it suffices to show that equality holds for $d = q - q/p - 1$ and does not hold for $d = q - q/p$. Let $d = q - q/p - 1 = (1 - p^{-1})q - 1$. This corresponds to the construction of $\mathcal{C} \triangleq \mathrm{RS}(q, d)^{1 \nearrow m}$ in Section 7.1 with $c = 1$. Let $s \geq 1$ be such that $q = p^s$. Let $\mathbf{d} = (d_1, \ldots, d_m) \in \mathrm{Deg}(\mathcal{C})$. We will show that $\sum_{i=1}^m d_i \leq d = q - q/p - 1$. Suppose, for the sake of contradiction, that $\sum_{i=1}^m d_i \geq q - q/p$. We will exhibit $\mathbf{e} \leq_p \mathbf{d}$ such that $q - q/p \leq \sum_{i=1}^m e_i < q$, which contradicts the fact that $\mathbf{d} \in \mathrm{Deg}(\mathcal{C})$ by Proposition 5.2.3. If $\sum_{i=1}^m d_i < q$, then we are done by taking $\mathbf{e} = \mathbf{d}$, so assume $\sum_{i=1}^m d_i \geq q$. Let $a \triangleq \sum_{i=1}^m d_i^{(s)}$. By Claim 7.1.9, $a \leq p - 2$. For each $i \in [m]$, let $\overline{d_i} = d_i - d_i^{(s)} p^{s-1}$. Then

$$\sum_{i=1}^m \overline{d_i} = \sum_{i=1}^m d_i - p^{s-1} \sum_{i=1}^m d_i^{(s)} \geq (p - a) p^{s-1}.$$

Let $k \in [m]$ be the minimal integer such that $\sum_{i=1}^k \overline{d_i} \geq (p - a - 1) p^{s-1}$. Since each $\overline{d_i} < p^{s-1}$, this implies that $\sum_{i=1}^k \overline{d_i} < (p - a) p^{s-1}$. Now, for $i \in [k]$, define $e_i \triangleq d_i$, and for $i > k$, define $e_i \triangleq d_i^{(s)} p^{s-1}$. By construction, $\mathbf{e} \leq_p \mathbf{d}$. On the other hand,

$$\sum_{i=1}^m e_i = \sum_{i=1}^k \overline{d_i} + p^{s-1} \sum_{i=1}^m d_i^{(s)} = \sum_{i=1}^k \overline{d_i} + a p^{s-1}$$

which is at least $(p - a - 1)p^{s-1} + ap^{s-1} = (p - 1)p^{s-1} = q - q/p$ and strictly less than $(p - a)p^{s-1} + ap^{s-1} = p^s = q$.

Now, let $d = q - q/p$. Let $f(\mathbf{X}) = X_1^{q-q/p} X_2^{q-q/p}$. Clearly $\deg(f) = 2(q - q/p) > q - q/p$.

We claim that $f \in \mathrm{RS}(q,d)^{1\nearrow m}$. For any $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^m$, we have

$$\begin{aligned} f(\mathbf{a}T + \mathbf{b}) &= (a_1 T + b_1)^{q - q/p} (a_2 T + b_2)^{q - q/p} && (7.1) \\ &= (a_1^{q/p} T^{q/p} + b_1^{q/p})^{p-1} (a_2^{q/p} T^{q/p} + b_2^{q/p})^{p-1} && (7.2) \\ & && (7.3) \end{aligned}$$

so, since $T^q = T$, every monomial in $f(\mathbf{a}T + \mathbf{b})$ is of the form $T^{c \cdot q/p}$ for some $0 \leq c \leq p-1$, thus $\deg(f(\mathbf{a}T + \mathbf{b})) \leq q - q/p = d$. □

As a corollary, we see that, over prime fields, lifting the Reed-Solomon code simply yields the Reed-Muller code.

Corollary 7.1.7. *If p is a prime and $d < q$, then $\mathrm{RS}(p,d)^{1\nearrow m} = \mathrm{RM}(p,d,m)$ for every $m \geq 1$.*

Proof. If $m = 1$, then both codes are equal to $\mathrm{RS}(p,d)$, so there is nothing to prove. Assume $m \geq 2$. If $d = p-1$, then both codes are equal to everything, so again there is nothing to prove. If $d < p-1$, then the result follows from Theorem 7.1.6. □

Theorem 7.1.8. *Let q be a prime power, let $d < q$, and let $m \geq 2$. Then $\mathrm{RM}(q,d,2)^{2\nearrow m} = \mathrm{RM}(q,d,m)$.*

Proof. By Lemma 7.1.5, it suffices to consider the case where $d = q-1$. It follows immediately from definitions that $\mathrm{RM}(q,d,m) \subseteq \mathrm{RM}(q,d,2)^{2\nearrow m}$, so it only remains to show the reverse inclusion. We do so by showing that if $\mathbf{d} \in \mathrm{Deg}(\mathrm{RM}(q,d,2)^{2\nearrow m})$, then $\|\mathbf{d}\| \leq d$. Actually, we show the converse. By Proposition 5.2.3, it suffices to show that if $\|\mathbf{d}\| \geq q$, then there exists $\mathbf{E} \leq_p \mathbf{d}$ such that $(\|\mathbf{E}_{*1}\| \bmod^* q) + (\|\mathbf{E}_{*2}\| \bmod^* q) \geq q$ where $\mathbf{E}$ has rows $[m]$ and columns $\{0,1,2\}$.

Without loss of generality, assume $d_1 \geq d_2 \geq \cdots \geq d_m$. Let $k \geq 2$ be the smallest integer such that $d_2 + \cdots + d_k \geq q - d_1$. Then $d_2 + \cdots + d_k \leq q-1$, for otherwise $d_k \geq d_1 + 1$, which is impossible. Construct $\mathbf{E}$ as follows. Define $\mathbf{E}_{1*} \triangleq (0, d_1, 0)$ and for $i \in \{2, \ldots, k\}$, define $\mathbf{E}_{i*} \triangleq (0, 0, d_i)$. Finally, for $i > k$, define $\mathbf{E}_{i*} = (0,0,0)$. By construction, $\mathbf{E} \leq_p \mathbf{d}$,

and moreover $\|\mathbf{E}_{*1}\| = d_1$ and $\|\mathbf{E}_{*2}\| = d_2 + \cdots + d_k$. Since $\|\mathbf{E}_{*1}\|, \|\mathbf{E}_{*2}\| \leq q-1$, applying $\bmod^* q$ does not reduce them, hence $(\|\mathbf{E}_{*1}\| \bmod^* q) + (\|\mathbf{E}_{*2}\| \bmod^* q) = \|\mathbf{E}_{*1}\| + \|\mathbf{E}_{*2}\| \geq d_1 + (q - d_1) = q$, as desired. □

7.1.2 Rate

Rate is the one key parameter that is not guaranteed by the lifting operator. However, the potential in lifted codes comes from the fact that it is possible to get extremely dense codes by lifting. From Section 7.1.1, we see that if $q - q/p \leq d < q$, then there are lifted Reed-Solomon codewords that are not Reed-Muller codewords. We show that, in fact, there are *many* such codewords. Our strategy is to lower bound the rate of the code. Observe that the lifted Reed-Solomon code, being a linear affine-invariant code, is spanned by monomials, i.e. has a degree set. Therefore, its dimension is equal to the size of its degree set. We use our knowledge of the structure of its degree set to lower bound the number of such degrees.

In this section, we analyze the rate of $\mathcal{C}$ for special values of d. Let p be the characteristic of $\mathbb{F}_q$, let $s \geq 1$ be such that $q = p^s$, let $c \leq s$ and let $d = (1 - p^{-c})q - 1$.

We begin by re-interpreting what it means for a number e to be less than d, in terms of the p-ary expansion of e. Through this section, for any $a \in \mathbb{N}$, let $a = \sum_{i \geq 0} a^{(i)} p^i$ be the p-ary expansion of a.

Claim 7.1.9. *If $e \in [\![q]\!]$, then $e \leq d$ if and only if $e^{(i)} < p - 1$ for some $s - c \leq i < s$.*

Proof. Since $e < q$, $e^{(i)} = 0$ for $i \geq s$. Note that if $e^{(i)} = p - 1$ for $s - c \leq i < s$, then $e \geq (p-1)\sum_{i=s-c}^{s-1} p^i = p^{s-c}(p-1)\sum_{i=0}^{c-1} p^i = (1 - p^{-c})q$, while if $e^{(i)} < p - 1$ for some $s - c \leq i < s$, this only decreases the value of e. □

Next, we use our knowledge of the degree set of the lifted Reed-Solomon code to provide a sufficient condition for a degree to be in the degree set.

Claim 7.1.10. *Let $b = \lceil \log_p m \rceil + 1$ and let $\mathbf{d} \in [\![q]\!]^m$. If there is some $s - c \leq j \leq s - b$ such that $d_i^{(k)} = 0$ for every $i \in [m]$ and every $j \leq k < j + b$, then $\mathbf{d} \in \mathrm{Deg}(\mathcal{C})$.*

Proof. Let $\mathbf{e} \leq_p \mathbf{d}$, and let $e = \sum_{i=1}^{m} e_i \bmod^* q$. We claim that $e^{(j+b-1)} = 0$, which implies $e \leq d$ Claim 7.1.9 since $s - c \leq j + b - 1 < s$. Note that $a \mapsto p \cdot a \bmod^* q$ results in a cyclic permutation of the digits $a^{(i)}$. So we may multiply $\mathbf{d}$ and $\mathbf{e}$ by an appropriate power of p, namely p^{s-b-j}, so that $j = s - b$. Therefore, we may assume without loss of generality that $j = s-b$, and we wish to show that $e^{(s-1)} = 0$, i.e. $e < p^{s-1}$. Note that since $\mathbf{e} \leq_p \mathbf{d}$, for each $i \in [m]$ and $s - b \leq k \leq s - 1$ we have $e_i^{(k)} = 0$, i.e. $e_i < p^{s-b}$ for every $i \in [m]$. Therefore $\sum_{i=1}^{m} e_i < mp^{s-b} < p^{b-1}p^{s-b} = p^{s-1}$. □

Finally, we lower bound the rate.

Theorem 7.1.11. *Let $b = \lceil \log_p m \rceil + 1$ The rate of the code $\mathcal{C}$ is at least $1 - e^{-(c+b)/(bp^{mb})}$.*

Proof. Consider choosing $\mathbf{d} \in [\![q]\!]^m$ uniformly at random. Let $a = \lfloor c/b \rfloor$. For $j \in [a]$, let E_j be the event that $d_i^{(k)} = 0$ for every $i \in [m]$ and $s - jb \leq k < s - (j-1)b$. By Claim 7.1.10, $\bigvee_{j=1}^{a} E_j$ is sufficient for $\mathbf{d} \in \mathrm{Deg}(\mathcal{C})$, so we wish to lower bound $\Pr\left[\bigvee_{j=1}^{a} E_j\right]$. We have $\Pr[E_j] = p^{-mb}$ for every $j \in [a]$, therefore

$$\begin{aligned}
\Pr\left[\bigvee_{j=1}^{a} E_j\right] &= 1 - \Pr\left[\bigwedge_{j=1}^{a} \overline{E_j}\right] && (7.4)\\
&= 1 - \prod_{j=1}^{a} \Pr\left[\overline{E_j}\right] && (7.5)\\
&= 1 - \prod_{j=1}^{a} (1 - \Pr[E_j]) && (7.6)\\
&= 1 - (1 - p^{-mb})^a && (7.7)\\
&\geq 1 - (1 - p^{-mb})^{(c+b)/b} && (7.8)\\
&\geq 1 - e^{-(c+b)/(bp^{mb})}. && (7.9)
\end{aligned}$$

□

7.1.3 Global List-Decoding

In this section, we present an efficient global list decoding algorithm for $\mathrm{RS}(q,d)^{1\nearrow m}$. Define $\alpha_1,\ldots,\alpha_m \in \mathbb{F}_{q^m}$, ϕ, and ϕ^* as in Section 5.3.3. Our main result states that $\mathrm{RS}(q,d)^{1\nearrow m}$ is isomorphic to a subcode of $\mathrm{RS}(q^m,(d+m)q^{m-1}) \subseteq \{\mathbb{F}_{q^m} \to \mathbb{F}_q\}$. In particular, one can simply list decode $\mathrm{RS}(q,d)^{1\nearrow m}$ by list-decoding $\mathrm{RS}(q^m,(d+m)q^{m-1})$ up to the Johnson radius. We will use this algorithm for $m=2$ as a subroutine in our local list decoding algorithm in Section 7.1.4.

Theorem 7.1.12. *If $f \in \mathrm{RS}(q,d)^{1\nearrow m}$, then $\deg(\phi^*(f)) \le (d+m)q^{m-1}$.*

Proof. By linearity, it suffices to prove this for a monomial $f(X_1,\ldots,X_m)=\prod_{i=1}^m X_i^{d_i}$. We have

$$\phi^*(f)(Z) = \sum_{\substack{(e_{11},\ldots,e_{1m})\le_p d_1 \\ \vdots \\ (e_{m1},\ldots,e_{mm})\le_p d_m}} (\cdots) Z^{\sum_{ij} e_{ij} q^{m-i}},$$

so it suffices to show that $\sum_{j=1}^m \sum_{i=1}^m e_{ij} q^{m-j} \bmod^* q^m \le (d+m)q^{m-1}$. By Proposition 7.1.2, for every $e_i \le_p d_i$, $i \in [m]$, we have $\sum_{i=1}^m e_i \bmod^* q \le d$. Therefore, there is some integer $0 \le k < m$ such that $\sum_{i=1}^m e_{i1} \in [kq, k(q-1)+d]$. Thus,

$$\begin{aligned}
kq^m &\le q^{m-1}\sum_{i=1}^m e_{i1} + \sum_{j=2}^m \sum_{i=1}^m e_{ij} q^{m-j} && (7.10)\\
&\le q^{m-1}\sum_{i=1}^m e_{i1} + q^{m-2}\sum_{j=2}^m \sum_{i=1}^m e_{ij} && (7.11)\\
&\le (k(q-1)+d)q^{m-1} + mq^{m-1} && (7.12)\\
&= k(q^m-1) + (d+m-k)q^{m-1} + k && (7.13)\\
&\le k(q^m-1) + (d+m)q^{m-1} && (7.14)
\end{aligned}$$

and hence $\sum_{j=1}^m \sum_{i=1}^m e_{ij} q^{m-j} \bmod^* q^m \le (d+m)q^{m-1}$. □

Corollary 7.1.13. *For every $m \ge 2$ and ever $\epsilon > 0$, there is a polynomial time algorithm that takes as input a function $r : \mathbb{F}_q^m \to \mathbb{F}_q$ and outputs a list $\mathcal{L}$ of size $|\mathcal{L}| = O(1/\epsilon^2)$ which*

contains all $c \in \mathrm{RS}(q,d)^{1\nearrow m}$ *such that* $\delta(r,c) < 1 - \sqrt{\frac{d+m}{q}} - \epsilon$.

Proof. Given $r : \mathbb{F}_q^m \to \mathbb{F}_q$, convert it to $r' = \phi^*(r)$, and then run the Guruswami-Sudan list decoder (Theorem 3.4.3) for $\mathrm{RS} \triangleq \mathrm{RS}(q^m, (d+m)q^{m-1})$ on r' to obtain a list $\mathcal{L}$ of size $|\mathcal{L}| = O(1/\epsilon^2)$ with the guarantee that any $c \in \mathrm{RS}$ with $\delta(r',c) < 1 - \sqrt{\frac{d+m}{q}} - \epsilon$ lies in $\mathcal{L}$. We require that any $c \in \mathrm{RS}(q,d)^{1\nearrow m}$ satisfying $\delta(r',c) < 1 - \sqrt{\frac{d+m}{q}} - \epsilon$ lies in $\mathcal{L}$, and this follows immediately from Theorem 7.1.12. □

7.1.4 Local List-Decoding

In this section, we present a local list decoding algorithm for $\mathrm{LiftedRS}(q,d,m)$, where $d = (1-\delta)q$ which decodes up to radius $1 - \sqrt{1-\delta} - \epsilon$ for any constant $\epsilon > 0$, with list size $O(1/\epsilon^2)$ and query complexity q^3.

Theorem 7.1.14. *For every* $m \geq 2$ *and every* $\delta, \epsilon > 0$, *setting* $d = (1-\delta)q$, $\mathrm{RS}(q,d)^{1\nearrow m}$ *is* $\left(q, q^3, 1 - \sqrt{1-\delta} - \epsilon, O(1/\epsilon^2), 0.2 + \frac{2}{\delta q}, O\left(\frac{1}{\epsilon^6 q \delta}\right)\right)$*-locally list-decodable.*

Proof. The following algorithm is the outer local list-decoding algorithm.

Local list decoder: Oracle access to received word $r : \mathbb{F}_q^m \to \mathbb{F}_q$.

1. Pick an affine transformation $\ell : \mathbb{F}_q \to \mathbb{F}_q^m$ uniformly at random.
2. Run Reed-Solomon list decoder (e.g. Guruswami-Sudan) on $r \circ \ell$ from $1 - \sqrt{1-\delta} - \frac{\epsilon}{2}$ fraction errors to get list $g_1, \ldots, g_L : \mathbb{F}_q \to \mathbb{F}_q$ of Reed-Solomon codewords.
3. For each $i \in [L]$, output $\mathsf{Correct}(A_{\ell,g_i})$

where $\mathsf{Correct}$ is a local correction algorithm for the lifted codes for 0.1δ fraction errors given by Theorem 5.3.1, which has a failure probability of $0.2 + \frac{2}{\delta q}$, and A is an oracle which takes as advice a univariate affine map and a univariate polynomial and simulates oracle access to a function which is supposed to be $\ll 0.1\delta$ close to a lifted RS codeword.

Oracle $A_{\ell,g}(\mathbf{x})$:

1. Let $P : \mathbb{F}_q^2 \to \mathbb{F}_q^m$ be the unique affine map such that $P(t,0) = \ell(t)$ for $t \in \mathbb{F}_q$ and $P(0,1) = \mathbf{x}$.

2. Use the global list decoder for $\mathrm{RS}(q,d)^{1\nearrow 2}$ given by Corollary 7.1.13 to decode $r \circ P$ from $1 - \sqrt{1-\delta} - \frac{\epsilon}{2}$ fraction errors and obtain a list $\mathcal{L}$.

3. If there exists a unique $h \in \mathcal{L}$ such that $h \circ \ell = g$, output $h(0,1)$, otherwise fail.

Analysis: For the query complexity, note that the local list-decoder makes q queries to r to output its oracles. For each oracle, $\mathsf{Correct}$ makes q queries to A_{ℓ,g_i}, which itself makes q^2 queries to r, so the entire oracle makes q^3 queries to r.

To show correctness, we just have to show that, with high probability over the choice of ℓ, for every lifted RS codeword f such that $\delta(r,f) < 1 - \sqrt{1-\delta} - \epsilon$, there is $i \in [L]$ such that $\mathsf{Correct}(A_{\ell,g_i}) = f$, i.e. $\delta(A_{\ell,g_i}, f) \le 0.1\delta$.

Fix such a function f. We will proceed in two steps:

1. First, we show that with high probability over ℓ, there is some $i \in [L]$ such that $f|_\ell = g_i$.

2. Next, we show that $\delta(A_{\ell,f\circ\ell}, f) \le 0.1\delta$ with high probability.

For the first step, note that $f|_\ell \in \{g_1, \ldots, g_L\}$ if $\delta(f \circ \ell, r \circ \ell) \le 1 - \sqrt{1-\delta} - \frac{\epsilon}{2}$. By Proposition B.1.4, $\delta(f \circ \ell, r \circ \ell)$ has mean less than $1 - \sqrt{1-\delta} - \epsilon$ and variance less than $1/(4q)$. By Chebyshev's inequality, the probability that $\delta(f|_\ell, r|_\ell) \le 1 - \sqrt{1-\delta} - \frac{\epsilon}{2}$ is at least $1 - \frac{1}{\epsilon^2 q}$.

For the second step, we want to show that $\Pr_{\mathbf{x}\in\mathbb{F}_q^m}[A_{\ell,f\circ\ell}(\mathbf{x}) \neq f(\mathbf{x})] \le 0.1\delta$. First consider the probability when we randomize ℓ as well. Then $P : \mathbb{F}_q^2 \to \mathbb{F}_q^m$ is a uniformly random affine map. We get $A_{\ell,f\circ\ell}(\mathbf{x}) = f(\mathbf{x})$ as long as $f \circ P \in \mathcal{L}$ and no other element $h \in \mathcal{L}$ has $h \circ \ell = f \circ \ell$. By Proposition B.1.4, $\delta(f \circ P, r \circ P)$ has mean less than $1 - \sqrt{1-\delta} - \epsilon$ and variance less than $1/(4q^2)$. By Chebychev's inequality, with probability at least $1 - \frac{1}{\epsilon^2 q^2}$ over ℓ and $\mathbf{x}$, we have $\delta(f|_P, r|_P) \le 1 - \sqrt{1-\delta} - \frac{\epsilon}{2}$ and hence $f|_P \in \mathcal{L}$. For the probability

that no two codewords in $\mathcal{L}$ agree on ℓ, view this as first choosing a random affine map $P : \mathbb{F}_q^2 \to \mathbb{F}_q^m$ and then choosing random affine $\ell' : \mathbb{F}_q \to \mathbb{F}_q^2$ and defining $\ell \triangleq P' \circ \ell'$. For any distinct $\phi, \psi : \mathbb{F}_q^2 \to \mathbb{F}_q$, the probability over ℓ' that $\phi \circ \ell' = \psi \circ \ell'$ is at most $2/q$, by Proposition B.1.5. Since $|\mathcal{L}| = O(1/\epsilon^2)$, there are at most $O(1/\epsilon^4)$ pairs of functions from $\mathcal{L}$, so by the union bound, with probability at least $1 - O(1/(\epsilon^4 q))$, all the functions in $\mathcal{L}$ are distinct on ℓ and hence $f \circ P$ is the unique codeword in $\mathcal{L}$ which is consistent with $f \circ \ell$.

Therefore, the probability, over ℓ and $\mathbf{x}$, that $A_{\ell.f\circ\ell}(\mathbf{x}) \neq f(\mathbf{x})$ is $O\left(\frac{1}{\epsilon^4 q}\right)$, and thus

$$\begin{aligned}
\Pr_{\ell}\left[\delta(A_{\ell,f\circ\ell}, f\circ\ell) > 0.1\delta\right] &= \Pr_{\ell}\left[\Pr_{\mathbf{x}}[A_{\ell,f\circ\ell}(\mathbf{x}) \neq f(\mathbf{x})] > 0.1\delta\right] \\
&\leq \frac{\Pr_{\ell,\mathbf{x}}[A_{\ell,f\circ\ell}(\mathbf{x}) \neq f(\mathbf{x})]}{0.1\delta} \\
&= O\left(\frac{1}{\epsilon^4 q\delta}\right).
\end{aligned}$$

So, for a fixed codeword $f \in \mathrm{RS}(q,d)^{1\nearrow m}$ such that $\delta(r,f) < 1-\sqrt{1-\delta}-\epsilon$, the probability of success is $O\left(\frac{1}{\epsilon^4 q\delta}\right)$. Since there are $O(1/\epsilon^2)$ such codewords, the overall probability of success is $O\left(\frac{1}{\epsilon^6 q\delta}\right)$.

□

As a corollary, we get the following testing result.

Theorem 7.1.15. *For every $\delta > 0$ and for any $\alpha < \beta < 1 - \sqrt{1-\delta}$, there is an algorithm which, given oracle access to a function $r : \mathbb{F}_q^m \to \mathbb{F}_q$, distinguishes between the cases where r is α-close to $\mathrm{RS}(q,d)^{1\nearrow m}$ (where $d \triangleq (1-\delta)q$), and where r is β-far, while making $O\left(\frac{\ln(1/(\beta-\alpha))q^4}{(\beta-\alpha)^4\delta^{72}}\right)$ queries to r.*

Proof. Let $\rho \triangleq (\alpha+\beta)/2$, and let $\epsilon \triangleq (\beta-\alpha)/8$, so that $\alpha = \rho - 4\epsilon$ and $\beta = \rho + 4\epsilon$. Set $\eta \triangleq O(\epsilon^2)$, where the constant is sufficiently large. Let T' be the q^2-query $\Omega(\delta^{72})$-robust tester algorithm for $RS(q,d)^{1\nearrow m}$, which rejects words that are ϵ-far with probability $\Omega(\epsilon\delta^{72})$, by Proposition 3.3.2. Let T be the $O(\ln(1/\eta)q^2/(\epsilon\delta^{72}))$-query tester which runs T' repeatedly $O(\ln(1/\eta)q^2/(\epsilon\delta^{72}))$ times and accepts if and only if every iteration accepts, to increase the rejection probability for ϵ-far words to $1-\eta$. Our algorithm is to run the local list-decoding

algorithm on r with error radius ρ, to obtain a list of oracles $M_1, \ldots, M_L$. For each M_i, we use random sampling to compute an estimate $\widetilde{\delta}(r, f_i)$ of the distance between r and the function f_i computed by M_i to within ϵ additive error with failure probability η, and keep only the ones with estimated distance less than $\rho + \epsilon$. Then, for each remaining M_i, we run T on M_i. We accept if T accepts some M_i, otherwise we reject.

The number of queries required to run the local list-decoding algorithm is q. For each $i \in [L]$, the number of queries to M_i we need to make for computing $\widetilde{\delta}(r, f_i)$ is $O(\ln(1/\eta)/\epsilon^2)$, by Proposition 2.2.4. Each query to M_i makes q^2 queries to r, for a total of $O(\ln(1/\eta)q^2/\epsilon^2)$ queries for each M_i and therefore $O(\ln 1/\eta)q^2/\epsilon^4)$ queries for the distance estimation step. The tester T makes $O(\ln(1/\eta)q^2/(\epsilon\delta^{72}))$ queries to each M_i, for a total of $O(\ln(1/\eta)q^4/(\epsilon\delta^{72}))$ queries to r for each $i \in [L]$ and therefore a grand total of $O(\ln(1/\eta)q^4/(\epsilon^3\delta^{72}))$ queries to r made by T. Therefore, the total number of queries to r made by our testing algorithm is $O\left(\frac{\ln(1/\eta)q^4}{\epsilon^4\delta^{72}}\right) = O\left(\frac{\ln(1/\epsilon)q^4}{\epsilon^4\delta^{72}}\right)$.

If r is α-close to $\mathrm{RS}(q,d)^{1\nearrow m}$, then it is α-close to some codeword f, and by the guarantee of the local list-decoding algorithm, there is some $j \in [L]$ such that M_j computes f. Since, $\delta(r, f) \leq \alpha$, with probability at least $1 - \eta$ we have $\widetilde{\delta}(r, f) \leq \alpha + \epsilon < \rho$, in which case M_j will not be pruned by our distance estimation. Since f is a codeword, this M_j will pass the testing algorithm T and so our algorithm will accept. The total failure probability is therefore $\eta + O\left(\frac{1}{\epsilon^6 q\delta}\right)$.

Now suppose r is β-far from $\mathrm{RS}(q,d)^{1\nearrow m}$. With probability $1 - O\left(\frac{1}{\epsilon^6 q\delta}\right)$, the local list-decoding algorithm succeeds, so let us condition on its success. With probability $1 - O(\eta/\epsilon^2)$, all the distance estimations for all the M_i simultaneously succeed, and let us condition on this as well. Consider any oracle M_i output by the local list-decoding algorithm and not pruned by our distance estimation. Let f_i be the function computed by M_i. Then the estimated distance is $\widetilde{\delta}(r, f_i) < \rho + \epsilon$, so the true distance is $\delta(r, f_i) < \rho + 2\epsilon$. Since r is β-far from any codeword, that means the distance from f_i to any codeword is at least $\beta - (\rho + 2\epsilon) > \epsilon$, and hence T will reject M_i with probability $1 - \eta$. By the union bound, all the M_i not pruned will be rejected with probability $1 - O(\eta/\epsilon^2)$. So the total failure probability in this case is

at most $O(\eta/\epsilon^2) + O\left(\frac{1}{\epsilon^6 q\delta}\right)$.

In either case, the failure probability is at most some constant. □

7.1.5 Main Result: The Code That Does It All

The following is our main theorem, Theorem 7.1.16, a more precise restatement of Theorem 1.2.1, which is the culmination of the lifting technology results of Chapters 5 and 6. It states that there is a high-rate code that is simultaneously locally correctable and decodable, locally list-decodable, and robustly testable.

Theorem 7.1.16. *For every $\alpha, \beta > 0$ there exists $\delta > 0$ such that the following holds: For infinitely many $n \in \mathbb{N}$, there is $q = q(n) = O(n^\beta)$ and a linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ with distance $\delta(\mathcal{C}) \geq \delta$, rate at least $1 - \alpha$, that is locally correctable and decodable up to $\frac{\delta}{2}$ fraction errors with $O(q)$ queries, locally list-decodable up to the Johnson bound with $O(q^3)$ queries, and $\Omega(\delta^{72})$-robustly testable with q^2 queries.*

Proof. Let $m = \lceil 1/\beta \rceil$ and let $b = \lceil \log_2 m \rceil + 1$. Let $c = b2^{mb} \cdot \lceil \ln(1/\alpha) \rceil + b$, and set $\delta \triangleq 2^{-c}$. For $s \geq c$, set $q = 2^s$ and $n \triangleq q^m$. Set $d \triangleq (1-\delta)q - 1$. Let $\mathcal{C} \triangleq \mathrm{RS}(q,d)^{1 \nearrow m}$.

By Proposition 7.1.1, $\delta(\mathcal{C}) \geq \delta$. By Theorem 7.1.11, the rate is at least $1 - \alpha$. By Proposition 7.1.3, $\mathcal{C}$ is locally correctable and decodable up to $\frac{\delta}{2}$ fraction errors with q queries. By Theorem 7.1.14, $\mathcal{C}$ is locally list-decodable up to the Johnson bound with q^3 queries. By Proposition 7.1.4, $\mathcal{C}$ is $\Omega(\delta^{72})$-robustly testable with q^2 queries. □

7.2 Robust Low-Degree Testing

In this section, we prove Theorem 7.2.2, which is simply a more precise restatement of Theorem 1.2.3. We do so by proving Theorem 7.2.1, a generalization from which Theorem 7.2.2 follows immediately. Theorem 7.2.1 replaces the Reed-Solomon code with an arbitrary univariate linear affine-invariant code $\mathcal{C}_0$ and replaces the bivariate Reed-Muller code with an arbitrary t-variate linear affine-invariant code $\mathcal{C}_1$ which is a strict subcode of $(\mathcal{C}_0)^{1 \nearrow t}$.

Theorem 7.2.1. *Let $t > 1$ and let $m \geq 3$. Let $\mathcal{C}_0 \subseteq \{\mathbb{F}_q \to \mathbb{F}_q\}$ and $\mathcal{C}_1 \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be linear affine-invariant codes such that $\mathcal{C}_1 \subsetneq (\mathcal{C}_0)^{1 \nearrow t}$. Let $\delta \triangleq \delta(\mathcal{C}_0)$. Fix $r : \mathbb{F}_q^m \to \mathbb{F}_q$. Let $\rho \triangleq \mathbb{E}_u \left[\delta\left(r|_u, \mathcal{C}_1\right)\right]$, where the expectation is taken over random t-dimensional $u \subseteq \mathbb{F}_q^m$. Let α_2 be the t-dimensional robustness of $(\mathcal{C}_0)^{1 \nearrow m}$. Then $\rho \geq \min\left\{\frac{\alpha_2 \delta^2}{4}, \left(\frac{\delta}{2} - 2q^{-1}\right) \cdot \frac{\delta}{2}\right\} \cdot \delta(r, (\mathcal{C}_1)^{t \nearrow m})$.*

Theorem 7.2.2 (Robust plane testing for Reed-Muller). *Let $m \geq 3$. Fix a positive constant $\delta > 0$ and a degree $d = (1 - \delta) \cdot q$. Let $\mathrm{RM}(m)$ be the m-variate Reed-Muller codes of degree d over $\mathbb{F}_q$. Then $\mathrm{RM}(m)$ is $\left(\frac{\delta^{74}}{8 \cdot 10^{52}}, 2\right)$-robust.*

Proof. Let RS be the Reed-Solomon code over $\mathbb{F}_q$ of degree d. Let p be the characteristic of q. Let α_1 be the 2-dimensional robustness of $\mathrm{RS}^{1 \nearrow m}$. Then $\alpha_2 \geq \frac{\delta^{72}}{2 \cdot 10^{52}}$ by Theorem 6.1.4 if $m = 3$, and by Theorem 6.1.18 if $m \geq 4$.

If $d < q - q/p$, then $\mathrm{RM}(m) = \mathrm{RS}^{1 \nearrow m}$ by Theorem 7.1.6, and so in this case the theorem follows immediately from Theorem 6.1.18. If $d \geq q - q/p$ and $q \geq \frac{8}{\delta}$, then $\mathrm{RM}(2) \subsetneq \mathrm{RS}^{1 \nearrow 2}$ (by Theorem 7.1.6) but $\mathrm{RM}(m) = \mathrm{RM}(2)^{2 \nearrow m}$ (by Theorem 7.1.8), and so in this case the theorem follows immediately from Theorem 7.2.1, with $\mathcal{C}_0 = \mathrm{RS}$, $t = 2$, and $\mathcal{C}_1 = \mathrm{RM}(2)$. If $q < \frac{8}{\delta}$ then the theorem follows from Corollary 6.1.2. □

Overview of Proof of Theorem 7.2.1. We illustrate the idea for the case where $t = 2$, $\mathcal{C}_0$ is the Reed-Solomon code, and $\mathcal{C}_1$ is the bivariate Reed-Muller code of the same degree. The generalization to arbitrary t and codes $\mathcal{C}_0, \mathcal{C}_1$ is straightforward. If r is far from the lifted code, then on random planes r will be far from the bivariate lifted code and hence also from the bivariate Reed-Muller code. So the remaining case is when r is close to the lifted code. If the nearest function is a Reed-Muller codeword, then the theorem follows from the robustness of the lifted code. Otherwise, if the nearest function c is not Reed-Muller, then we show (through Corollary 6.2.8) that on many planes c is not a bivariate Reed-Muller codeword, and so r (being close to c) is not close to a bivariate Reed-Muller codeword (by the distance of the code).

Proof of Theorem 7.2.1. Observe that, since $(\mathcal{C}_1)^{t \nearrow m} \subset (\mathcal{C}_0)^{1 \nearrow m}$, we have $\delta(r, (\mathcal{C}_0)^{1 \nearrow m}) \leq$

$\delta(r, (\mathcal{C}_1)^{t\nearrow m})$. If $\delta(r, (\mathcal{C}_0)^{1\nearrow m}) \geq \min\left\{\frac{\delta^2}{4}, \delta(r, (\mathcal{C}_1)^{t\nearrow m})\right\}$, then we are done since

$$\begin{aligned}
\rho &= \mathbb{E}_u\left[\delta\left(r|_u, \mathcal{C}_1\right)\right] \\
&\geq \mathbb{E}_u\left[\delta\left(r|_u, (\mathcal{C}_0)^{1\nearrow t}\right)\right] \\
&\geq \alpha_2 \cdot \delta(r, (\mathcal{C}_0)^{1\nearrow m}) \\
&\geq \alpha_2 \cdot \min\left\{\frac{\delta^2}{4}, \delta(r, (\mathcal{C}_1)^{t\nearrow m})\right\} \\
&\geq \frac{\alpha_2\delta^2}{4} \cdot \delta(r, (\mathcal{C}_1)^{t\nearrow m}).
\end{aligned}$$

Therefore, suppose $\delta(r, (\mathcal{C}_0)^{1\nearrow m}) < \min\left\{\frac{\delta^2}{4}, \delta(r, (\mathcal{C}_1)^{t\nearrow m})\right\}$. Let $f \in (\mathcal{C}_0)^{1\nearrow m}$ be the nearest codeword to r, so that $f \notin (\mathcal{C}_1)^{t\nearrow m}$ and $\delta(r, f) < \frac{\delta^2}{4}$. If u is a t-dimensional subspace for which $f|_u \notin \mathcal{C}_1$, then, since $\mathcal{C}_1$ is a subcode of $(\mathcal{C}_0)^{1\nearrow t}$, $\delta(r|_u, \mathcal{C}_1) \geq \delta - \delta(r|_u, f|_u)$. Since

$$\mathbb{E}_u\left[\delta\left(r|_u, f|_u\right)\right] = \delta(r, f) < \frac{\delta^2}{4},$$

by Markov,

$$\Pr_u\left[\delta(r|_u, f|_u) \geq \frac{\delta}{2}\right] \leq \frac{\delta}{2}$$

By Corollary 6.2.8,

$$\Pr_u\left[f|_u \in \mathcal{C}_1\right] \leq 1 - \delta + 2q^{-1}.$$

By the union bound, it follows that for at least $\frac{\delta}{2} - 2q^{-1}$ fraction of the t-dimensional $u \subseteq \mathbb{F}_q^m$, it holds that

$$\delta(r|_u, \mathcal{C}_1) \geq \delta - \delta(r|_u, f|_u) \geq \frac{\delta}{2}.$$

Therefore,

$$\rho = \mathbb{E}_u\left[\delta\left(r|_u, \mathcal{C}_1\right)\right] \geq \left(\frac{\delta}{2} - 2q^{-1}\right) \cdot \frac{\delta}{2}.$$

□

7.3 Nikodym Sets

In this section, we use the lifted Reed-Solomon code to improve upon the polynomial method to show that Nikodym sets are large.

Definition 7.3.1. A subset $S \subseteq \mathbb{F}_q^m$ is a *Nikodym set* if for every $\mathbf{x} \in \mathbb{F}_q^m$ there exists a nonzero $\mathbf{a} \in \mathbb{F}_q^m$ such that $\mathbf{x} + t\mathbf{a} \in S$ for every nonzero $t \in \mathbb{F}_q$.

Before we prove our lower bound, let us recall how the polynomial method yields a good lower bound in the first place. Let S be a Nikodym set. Let R be the rate of the code $\mathrm{RM}(q, q-2, m)$, which is approximately $R \approx \frac{1}{m!}$. Now, assume for the sake of contradiction that $|S| < Rq^m$. Then there exists a nonzero f with $\deg(f) \leq q-2$ that vanishes on every point of S. This is because the coefficients of f provide Rq^m degrees of freedom, but we only have $|S|$ linear constraints. Now, we claim that f actually vanishes everywhere, contradicting the fact that it is nonzero. To see this, let $\mathbf{x} \in \mathbb{F}_q^m$ be an arbitrary point. By the Nikodym property of S, there is a line L through S such that $(L \setminus \{\mathbf{x}\}) \subseteq S$. Consider the univariate polynomial $f|_L$. It vanishes on $q-1$ points $L \setminus \{\mathbf{x}\}$ but has degree $\deg(f|_L) < q-1$, so $f|_L$ is identically zero. But then $f(\mathbf{x}) = f|_L(\mathbf{x}) = 0$. Therefore $|S| \geq Rq^m \approx \frac{q^m}{m!}$.

Observe that, in the above argument, the only property we actually used was the fact that, on every line L, the restriction $f|_L$ has degree $\deg(f|_L) < q-1$. So pulling f from the Reed-Muller code was needlessly restrictive. We could use a larger code if possible, and in fact the largest code satisfying this property is none other than the lifted Reed-Solomon code!

Theorem 7.3.2. *Let p be a prime and let $q = p^s$. If $S \subseteq \mathbb{F}_q^m$ is a Nikodym set, then $|S| \geq \left(1 - e^{-(s+b)/(bp^{mb})}\right) \cdot q^m$ where $b = \lceil \log_p m \rceil + 1$. In particular, if p is fixed and $q \to \infty$, then $|S| \geq (1 - o(1)) \cdot q^m$.*

Proof. Let $R = 1 - e^{-(s+b)/(bp^{mb})}$. Suppose, for the sake of contradiction, that $|S| < Rq^m$. Let $\mathcal{C} = \mathrm{RS}(q, q-2)^{\nearrow m}$. This is the construction of Section 7.1 with $c = s$. By Theorem 7.1.11, $\dim_{\mathbb{F}_q}(\mathcal{C}) \geq Rq^m > |S|$, so there exists a nonzero $c \in \mathcal{C}$ such that $c(\mathbf{x}) = 0$ for every $\mathbf{x} \in S$. This follows from the fact that each equation of the form $c(\mathbf{x}) = 0$ is a linear constraint on

c, and there are $|S|$ linear constraints on c which has $\dim_{\mathbb{F}_q}(\mathcal{C}) > |S|$ degrees of freedom. We proceed to show that $c = 0$, a contradiction. Let $\mathbf{x} \in \mathbb{F}_q^m$. Since S is a Nikodym set, there is a nonzero $\mathbf{a} \in \mathbb{F}_q^m$ such that $\mathbf{x} + t\mathbf{a} \in S$ for every nonzero $t \in \mathbb{F}_q$. Define $c_{\mathbf{a}}(t) \triangleq c(\mathbf{x} + t\mathbf{a})$. Since $c \in \mathcal{C}$, we have $\deg(c_{\mathbf{a}}) \leq q - 2$. However, $c_{\mathbf{a}}(t) = 0$ for every $t \neq 0$, so $c_{\mathbf{a}}$ vanishes on $q - 1 > \deg(c_{\mathbf{a}})$ points, so $c_{\mathbf{a}} = 0$. In particular, $c(\mathbf{x}) = c_{\mathbf{a}}(0) = 0$. □

Appendix A

Algebra Background

This appendix contains well-known facts about arithmetic over finite fields and tensor codes that are used in the thesis.

A.1 Arithmetic over finite fields

Much of our work involves manipulating polynomials over finite fields. Expanding multinomials over a finite field, which has positive characteristic, is different from expanding over a field of zero characteristic, because binomial coefficients may vanish due to the characteristic of the field. For example, over a field of characteristic 2, the binomial $(X+Y)^2$ expands to $X^2 + 2XY + Y^2 = X^2 + Y^2$.

The key fact we will use is (a generalization of) Lucas' Theorem, which tells us what a multinomial coefficient looks like modulo a prime p.

Theorem A.1.1 (Generalized Lucas' Theorem)**.** *Let $a_0, a_1, \ldots, a_m \in \mathbb{N}$ and let p be a prime. Write $a_j = \sum_{i=1}^{n} a^{(i)} \cdot p^i$ where $a_j^{(i)}, b_j^{(i)} \in [\![p]\!]$ for every $0 \le i \le n$ and $0 \le j \le m$. Then*

$$\binom{a_0}{a_1, \ldots, a_m} \equiv \prod_{i=0}^{n} \binom{a_0^{(i)}}{a_1^{(i)}, \ldots, a_m^{(i)}} \pmod{p}.$$

Proof. Let $S = \{1, \ldots, a_0\}$. Partition S into $S_{i,j} \subset S$, for each $i \in [\![n+1]\!]$ and $j \in \left[a_0^{(i)}\right]$,

of size $|S_{i,j}| = p^i$. For each $i \in [\![n+1]\!]$ and $j \in \left[a_0^{(i)}\right]$, let $G_{i,j} = \mathbb{Z}_{p^i}^{a_0^{(i)}}$ act on $S_{i,j}$ be cyclic permutation. Let $G = \bigoplus_{i=0}^{n} \bigoplus_{j=1}^{a_0^{(i)}} G_{i,j}$, which acts on S. Let C be the collection of all m-colorings $c : S \to [m]$ such that $|c^{-1}(k)| = a_k$. Observe that $|C| = \binom{a_0}{a_1,\ldots,a_m}$, and the action of G on S naturally induces an action of G on C, as follows: if $c \in C$, and $g \in G$, then for $x \in S$, $(gc)(x) = c(g^{-1}x)$. Since $|G|$ is a power of p, so is the size of any orbit of C under G. So, to compute $\binom{a_0}{a_1,\ldots,a_m} \pmod p$, it suffices to count the number of fixed points of C under G. Since G fixes each $S_{i,j}$, it is easy to see that a coloring is fixed under G if and only if each $S_{i,j}$ is monochromatic. It therefore suffices to show that for each $i \in [\![n+1]\!]$ and each $k \in [m]$, there are exactly $a_k^{(i)}$ values for j such that $S_{i,j}$ has color k. Fix $k \in [m]$. Observe that there are $a_k = \sum_{i=1}^{n} a_k^{(i)}$ elements of color k in total, and each set $S_{i,j}$ of color k contributes p^i elements. The claim then follows easily by induction on $n - i$. □

In particular, it characterizes which multinomial coefficients are nonzero modulo p.

Corollary A.1.2. *If $d, e_1, \ldots, e_n \in \mathbb{N}$, then $\binom{d}{e_1,\ldots,e_n} \not\equiv 0 \pmod p$ only if $(e_1, \ldots, e_n) \leq_p d$.*

This allows us to expand multinomials over a finite field and know which terms of the usual expansion disappear.

Corollary A.1.3. *Let p be a prime and let $\mathbb{F}$ be a field of characteristic p. Let $d \in \mathbb{N}$ and let $x_1, \ldots, x_n \in \mathbb{F}$. Then*

$$\left(\sum_{i=1}^{n} x_i\right)^d = \sum_{(e_1,\ldots,e_n) \leq_p d} \binom{d}{e_1, \ldots, e_n} \prod_{i=1}^{n} x_i^{e_i}$$

Expanding multinomials over finite fields is particularly important for us since we frequently look at the restricting polynomials to affine subspaces, which entails composing with an affine function.

Proposition A.1.4. *Let p be a prime and let $\mathbb{F}$ be a field of characteristic p. Let $\mathbf{X} = (X_1, \ldots, X_m)$ and let $\mathbf{Y} \in (Y_1, \ldots, Y_t)$. Let $f(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]$, and let $\mathbf{A} \in \mathbb{F}^{m \times t}$ and $\mathbf{b} \in \mathbb{F}_q^m$. If*

$f(\mathbf{X}) = \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot \mathbf{X}^{\mathbf{d}}$, *then*

$$f(\mathbf{AY}+\mathbf{b}) = \sum_{\mathbf{d}\in \mathrm{supp}(f)} f_{\mathbf{d}} \cdot \sum_{\mathbf{E}\leq_p \mathbf{d}} \binom{\mathbf{d}}{\mathbf{E}} \prod_{i=1}^{m} \left(b_i^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}} \right) \cdot \prod_{j=1}^{t} Y_j^{\|\mathbf{E}_{*j}\|}$$

where $\mathbf{E}$ *in the summation is a* $m \times (t+1)$ *matrix with rows indexed by* $[m]$ *and columns indexed by* $[\![t+1]\!]$.

Proof.

$$\begin{aligned}
f(\mathbf{AY}+\mathbf{b}) &= \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot (\mathbf{AY}+\mathbf{b})^{\mathbf{d}} \\
&= \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot \prod_{i=1}^{m} \left(\sum_{j=1}^{t} a_{ij} Y_j + b_i \right)^{d_i} \\
\text{(Corollary A.1.3)} &= \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot \prod_{i=1}^{m} \sum_{(e_{i0},e_{i1},\ldots,e_{it})\leq_p d_i} \binom{d_i}{e_{i0}, e_{i1}, \ldots, e_{it}} b_i^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}} Y_j^{e_{ij}} \\
&= \sum_{\mathbf{d}\in D} f_{\mathbf{d}} \cdot \sum_{\mathbf{E}\leq_p \mathbf{d}} \binom{\mathbf{d}}{\mathbf{E}} \prod_{i=1}^{m} \left(b_i^{e_{i0}} \prod_{j=1}^{t} a_{ij}^{e_{ij}} \right) \cdot \prod_{j=1}^{t} Y_j^{\|\mathbf{E}_{*j}\|}
\end{aligned}$$

□

A.2 Tensor codes

The tensor product is a natural operation in linear algebra that, when applied to two linear codes, produces a new linear code in a natural way. There are many equivalent ways to define the tensor product of two codes. Since in this thesis we think of codes as linear subspaces of functions in $\{\mathbb{F}_q^m \to \mathbb{F}_q\}$, we define the tensor product in this context.

Definition A.2.1. Let $n \geq 2$, let $t_1, \ldots, t_n \geq 1$ and $m = \sum_{i=1}^{n} t_i$, and for each $i \in [n]$, let the code $\mathcal{C}_i \subseteq \{\mathbb{F}_q^{t_i} \to \mathbb{F}_q\}$ be linear and let $V_{i,\mathbf{a}} \subseteq \mathbb{F}_q^m$ be the t_i dimensional subspace consisting of all points where the i-th block (of t_i coordinates) is free and all the $[n] \setminus \{i\}$ blocks are fixed to $\mathbf{a} \in \prod_{j\neq i} \mathbb{F}_q^{t_j}$. The *tensor product code* $\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_n \subseteq \{\mathbb{F}_q^m \to \mathbb{F}_q\}$ is the

code

$$\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_n \triangleq \left\{ f : \mathbb{F}_q^m \to \mathbb{F}_q \;\middle|\; f|_{V_{i,\mathbf{a}}} \in \mathcal{C}_i \text{ for every } i \in [n] \text{ and } \mathbf{a} \in \prod_{j \neq i} \mathbb{F}_q^{t_j} \right\}$$

Define $\mathcal{C}^{\otimes n} \triangleq \overbrace{\mathcal{C} \otimes \cdots \otimes \mathcal{C}}^{n}$.

The following characterization of tensor product codes will be helpful.

Proposition A.2.2. *Let $n \geq 2$, let $t_1, \ldots, t_n \geq 1$ and $m = \sum_{i=1}^n t_i$, and for each $i \in [n]$, let the code $\mathcal{C}_i \subseteq \{\mathbb{F}_q^{t_i} \to \mathbb{F}_q\}$ be linear, and let $\mathbf{X}_i = (X_{i1}, \ldots, X_{it_i})$ be variables. Then*

$$\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_n = \mathrm{span}_{\mathbb{F}_q} \left\{ \prod_{i=1}^n f_i(\mathbf{X}_i) \;\middle|\; f_i \in \mathcal{C}_i \right\}$$

Corollary A.2.3. *If $\mathcal{C} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ has a degree set $\mathrm{Deg}(\mathcal{C})$, and $n \geq 1$, then $\mathcal{C}^{\otimes n}$ has degree set $\mathrm{Deg}(\mathcal{C}^{\otimes n}) = \mathrm{Deg}(\mathcal{C})^n$. In particular, if $\mathcal{C}$ is linear affine-invariant, and $\mathbb{F}_q$ has characteristic p, then $\mathcal{C}^{\otimes n}$ has a p-shadow-closed degree set.*

Proposition A.2.4. *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be codes with distance δ_1 and δ_2 repectively. Then $\delta(\mathcal{C}_1 \otimes \mathcal{C}_2)$ is at least $\delta_1\delta_2$. In particular, $\delta\left(\mathcal{C}^{\otimes n}\right) \geq \delta(\mathcal{C})^n$.*

The following is a statement about the erasure decoding properties of tensor product codes.

Proposition A.2.5. *Let $\mathcal{C} = \mathcal{C}_1 \otimes \ldots \mathcal{C}_n \in \left\{\mathbb{F}_q^m \to \mathbb{F}_q\right\}$ and $S \subseteq \mathbb{F}_q^m$ be a subset such that for every $i \in [n]$ and $\mathbf{a} \in \prod_{j \neq i} \mathbb{F}_q^{t_j}$ satisfy $|S \cap V_{i,\mathbf{a}}| \geq (1 - \delta(\mathcal{C}_i))q^{t_i}$. Let $r : S \to \mathbb{F}_q$ be such that for every $i \in [n]$ and $\mathbf{a} \in \prod_{j \neq i} \mathbb{F}_q^{t_j}$ satisfy that $r|_{S \cap V_{i,\mathbf{a}}}$ can be extended into a codeword of $\mathcal{C}_i$ on $V_{i,\mathbf{a}}$. Then there exists a unique $r' \in \mathcal{C}$ such that $r'|_S = r$.*

Appendix B

Finite field geometry

This appendix describes some basic structural facts about affine maps, as well as basic geometry of affine subspaces over finite fields.

B.1 Affine maps

The following proposition allows us to decompose an arbitrary affine map (not necessarily injective) into a composition of a linear map and an injective affine map.

Proposition B.1.1. *Let $t \leq m$. For every affine map $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$, there exists a linear map $A' : \mathbb{F}_q^t \to \mathbb{F}_q^t$ and injective affine map $A'' : \mathbb{F}_q^t \to \mathbb{F}_q^m$ such that $A = A'' \circ A'$.*

Proof. If $A : \mathbf{x} \mapsto A_L(\mathbf{x}) + \mathbf{b}$ for some linear $A_L : \mathbb{F}_q^t \to \mathbb{F}_q^m$ and $\mathbf{b} \in \mathbb{F}_q^m$, and if $A_L = A''_L \circ A'$ for some injective affine $A''_L : \mathbb{F}_q^t \to \mathbb{F}_q^m$ and linear $A' : \mathbb{F}_q^t \to \mathbb{F}_q^t$, then $A = A'' \circ A'$ where $A'' : \mathbf{x} \mapsto A''_L(\mathbf{x}) + \mathbf{b}$. Therefore, we reduce to the case where A is linear, i.e. $A(\mathbf{0}) = \mathbf{0}$.

Fix a basis $\mathbf{e}_1, \ldots, \mathbf{e}_t \in \mathbb{F}_q^t$. For $j \in [t]$, let $\mathbf{v}_j \triangleq A(\mathbf{e}_i)$. After re-labeling, we may assume without loss of generality that, for some $0 \leq r \leq t$, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent and $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_t \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. There exist (unique) $a_{ij} \in \mathbb{F}_q$, for $i \in [r]$ and $r+1 \leq j \leq t$, such that $\mathbf{v}_j = \sum_{i=1}^r a_{ij}\mathbf{v}_j$. Define $A' : \mathbb{F}_q^t \to \mathbb{F}_q^t$ as the unique linear map such that $A'(\mathbf{e}_i) = \mathbf{e}_i$ for $i \in [r]$ and $A'(\mathbf{e}_j) = \sum_{i=1}^r a_{ij}\mathbf{e}_i$ for $r+1 \leq j \leq t$. For $i \in [r]$, set

$\mathbf{w}_i \triangleq \mathbf{v}_i$, and extend $\mathbf{w}_1, \ldots, \mathbf{w}_r$ to a set of linearly independent vectors $\mathbf{w}_1, \ldots, \mathbf{w}_t \in \mathbb{F}_q^m$. Define $A'' : \mathbb{F}_q^t \to \mathbb{F}_q^m$ to be the unique linear map such that $A''(\mathbf{e}_j) = \mathbf{w}_j$.

By construction, A'' is injective since its image is $\text{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_t\}$. Moreover, for every $i \in [r]$,

$$(A'' \circ A')(\mathbf{e}_i) = A''(A'(\mathbf{e}_i)) = A''(\mathbf{e}_i) = \mathbf{w}_i = \mathbf{v}_i = A(\mathbf{e}_i)$$

and for every $r + 1 \leq j \leq t$,

$$(A'' \circ A')(\mathbf{e}_j) = A''(A'(\mathbf{e}_j)) = A''\left(\sum_{i=1}^{r} a_{ij}\mathbf{e}_i\right) = \sum_{i=1}^{r} a_{ij}A''(\mathbf{e}_i) = \sum_{i=1}^{r} a_{ij}\mathbf{v}_i = \mathbf{v}_j = A(\mathbf{e}_j),$$

so, by linearity, $A'' \circ A' = A$. □

The next three propositions describe the behavior of random affine maps. We are particularly interested in how $\delta(f \circ A, g \circ A)$ behaves, where f, g are distinct functions and A is a random subspace. Proposition B.1.3 covers the case where $A(\mathbf{0})$ is fixed, while Proposition B.1.4 allows A to be truly free.

Proposition B.1.2. *Let $t \leq m$, let $\mathbf{x} \in \mathbb{F}_q^m$, and let $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ be a random affine map such that $A(\mathbf{0}) = \mathbf{x}$. For any fixed nonzero $\mathbf{y} \in \mathbb{F}_q^t$, $A(\mathbf{y})$ is a uniformly random point in $\mathbb{F}_q^m$.*

Proof. We can choose A by choosing $a_{ij} \in \mathbb{F}_q$ for $i \in [m], j \in [t]$ independently and uniformly at random, and setting

$$A(\mathbf{y}) = \mathbf{x} + \begin{bmatrix} a_{11} & \cdots & a_{1t} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mt} \end{bmatrix} \mathbf{y}$$

The i-th coordinate of $A(\mathbf{y})$ is therefore $x_i + \sum_{j=1}^{t} a_{ij}y_j$ which is uniformly random in $\mathbb{F}_q$, and the coordinates of $A(\mathbf{y})$ are independent since the a_{ij} are independent. □

Proposition B.1.3. *Let $t \leq m$, let $\mathbf{x} \in \mathbb{F}_q^m$, and let $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ be a random affine map such that $A(\mathbf{0}) = \mathbf{x}$. If $f, g : \mathbb{F}_q^m \to \mathbb{F}_q$, then $\mathbb{E}_A\left[\delta(f \circ A, g \circ A)\right] \leq \delta(f, g) + q^{-t}$.*

Proof. We have

$$\begin{aligned}
\mathbb{E}_A\left[\delta(f\circ A, g\circ A)\right] &= \mathbb{E}_A\left[\mathbb{E}_{\mathbf{y}\in\mathbb{F}_q^t}\left[\mathbb{1}_{f(A(\mathbf{y}))\neq g(A(\mathbf{y}))}\right]\right] && \text{(B.1)}\\
&= \mathbb{E}_{\mathbf{y}\in\mathbb{F}_q^t}\left[\mathbb{E}_A\left[\mathbb{1}_{f(A(\mathbf{y}))\neq g(A(\mathbf{y}))}\right]\right] && \text{(B.2)}\\
&\leq q^{-t} + \mathbb{E}_{\mathbf{y}\in\mathbb{F}_q^t\setminus\{\mathbf{0}\}}\left[\mathbb{E}_A\left[\mathbb{1}_{f(A(\mathbf{y}))\neq g(A(\mathbf{y}))}\right]\right] && \text{(B.3)}\\
\text{(Proposition B.1.2)} &= q^{-t} + \delta(f,g). && \text{(B.4)}
\end{aligned}$$

□

Proposition B.1.4. *Let $t \leq m$, and let $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ be a uniformly random affine map. If $f, g : \mathbb{F}_q^m \to \mathbb{F}_q$ and $\delta \triangleq \delta(f,g)$, then $\delta(f\circ A, g\circ A)$ has mean δ and variance $\delta(1-\delta)q^{-t} \leq q^{-t}/4$.*

Proof. Observe that for any fixed $\mathbf{y} \in \mathbb{F}_q^m$, if $A : \mathbb{F}_q^t \to \mathbb{F}_q^m$ is uniformly random, then $A(\mathbf{y})$ is a uniformly random point in $\mathbb{F}_q^m$, and for distinct $\mathbf{y}$, the $A(\mathbf{y})$ are pairwise independent. Therefore, $\mathbb{1}_{f(A(\mathbf{y}))\neq g(A(\mathbf{y}))}$ has mean δ and variance $\delta(1-\delta)$. Since $\delta(f\circ A, g\circ A) = q^{-t}\sum_{\mathbf{y}\in\mathbb{F}_q^t}\mathbb{1}_{f(A(\mathbf{y}))\neq g(A(\mathbf{y}))}$, it follows that the mean is δ and variance is $\delta(1-\delta)q^{-t} \leq q^{-t}/4$. □

Finally, we prove that if f, g are distinct polynomials on a plane, then they cannot agree on too many lines.

Proposition B.1.5. *Let $f, g : \mathbb{F}_q^2 \to \mathbb{F}_q$ be distinct. For any $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^2$, define $f_{\mathbf{a},\mathbf{b}}, g_{\mathbf{a},\mathbf{b}} : \mathbb{F}_q \to \mathbb{F}_q$ by $f_{\mathbf{a},\mathbf{b}}(T) \triangleq f(\mathbf{a}T + \mathbf{b})$ and similarly $g_{\mathbf{a},\mathbf{b}}(T) \triangleq g(\mathbf{a}T + \mathbf{b})$. Then $f_{\mathbf{a},\mathbf{b}} = g_{\mathbf{a},\mathbf{b}}$ for at most $2q^3$ pairs $(\mathbf{a},\mathbf{b}) \in \mathbb{F}_q^2 \times \mathbb{F}_q^2$.*

Proof. Let $h = f - g$ and $h_{\mathbf{a},\mathbf{b}} = f_{\mathbf{a},\mathbf{b}} - g_{\mathbf{a},\mathbf{b}}$. We have four cases to consider.

1. $a_1, a_2 \neq 0$: if $h_{\mathbf{a},\mathbf{b}} = 0$, then the polynomial $h(X,Y)$ is divisible by $Y - \frac{a_2}{a_1}X - \frac{a_2 b_1}{a_1} - b_2$. There are $q(q-1)$ pairs $(\mathbf{a},\mathbf{b})$ which correspond to this factor ($q-1$ choices for $\mathbf{a}$ and then q choices for $\mathbf{b}$ given $\mathbf{a}$).

2. $a_1 \neq 0, a_2 = 0$: if $h_{\mathbf{a},\mathbf{b}} = 0$, then the polynomial $h(X,Y)$ is divisible by $Y - b_2$. There are q^2 pairs corresponding to this factor (q choices for a_1 and q choices for b_1).

3. $a_1 = 0, a_2 \neq 0$: same as previous case, by symmetry.

4. $a_1 = a_2 = 0$: there are at most q^2 such pairs (q^2 choices for $\mathbf{b}$).

Say a pair $(\mathbf{a}, \mathbf{b})$ is *bad* if $h_{\mathbf{a},\mathbf{b}} = 0$. Trivially, $\deg(h) \leq 2(q-1)$, so h has at most $2(q-1)$ linear factors. Each factor from cases 1, 2, and 3 corresponds to at most q^2 pairs $(\mathbf{a}, \mathbf{b})$, resulting in at most $2q^2(q-1)$ bad pairs $(\mathbf{a}, \mathbf{b})$ from those cases, and there are at most q^2 bad pairs from case 4. So, the total number of bad pairs is at most $2q^2(q-1)+q^2 \leq 2q^3$. □

B.2 Affine subspaces

The next two lemmas analyze the behavior of random affine subspaces in $\mathbb{F}_q^m$. Lemma B.2.1 shows that a low-dimensional subspace is disjoint from almost all low-dimensional affine subspaces.

Lemma B.2.1. *Let $t \leq k < m$. Let $u \subseteq \mathbb{F}_q^m$ be a fixed affine subspace of dimension t, and let $v \subseteq \mathbb{F}_q^m$ be a uniformly random affine subspace of dimension k. Then $\Pr_v[u \cap v \neq \emptyset] < q^{-(m-k-t)}$.*

Proof. By affine symmetry, we may assume that v is fixed and u is random. Furthermore, we can assume that $v = (\mathbf{0}, B)$, where B is a basis and hence $|B| = k$. We choose u by choosing random $\mathbf{x} \in \mathbb{F}_q^m$, random basis $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_t\}$, and setting $u \triangleq (\mathbf{x}, A)$. Let E be the event that $u \cap v \neq \emptyset$.

Re-arrange $\mathbf{a}_1, \ldots, \mathbf{a}_t$ so that for some $0 \leq s \leq t-1$, $\mathbf{a}_i \in \mathrm{span}(B)$ if and only if $i \leq s$. Note that E holds if and only if there exist $c_1, \ldots, c_t \in \mathbb{F}_q^m$ such that $\mathbf{x} + c_1\mathbf{a}_1 + \cdots + c_t\mathbf{a}_t \in \mathrm{span}(B)$. Let $P : \mathbb{F}_q^m \to \mathbb{F}_q^{m-k}$ be the linear map that projects onto the last $m-k$ coordinates. Note that $\ker(P) = B$. For each $i \in [t]$, let $\mathbf{a}'_i \triangleq P\mathbf{a}_i \in \mathbb{F}_q^{m-k}$. Then E holds if and only if $P\mathbf{x} \in \mathrm{span}(\mathbf{a}'_{s+1}, \ldots, \mathbf{a}'_t)$. Therefore, there are at most q^t choices for $P\mathbf{x}$, hence at most q^{k+t} choices for $\mathbf{x}$, out of q^m total choices for $\mathbf{x}$, so $\Pr[E] \leq \frac{q^{k+t}}{q^m} = q^{-(m-k-t)}$. □

Lemma B.2.2 shows that if $k \ll m$, then k random vectors are likely to be linearly independent, and in particular two random low-dimensional subspaces through a fixed point

$\mathbf{x}$ are likely to intersect only at $\mathbf{x}$.

Lemma B.2.2. *Let $k < m$ and $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{F}_q^m$ be uniformly chosen vectors. Then the probability that $\{\mathbf{a}_i\}_{i=1}^k$ are linearly independent is at least $1 - q^{-(m-k)}$. In particular, the probability that two t-dimensional subspaces through a point $\mathbf{x} \in \mathbb{F}_q^m$ will intersect only on $\mathbf{x}$ is at least $1 - q^{-(m-2t)}$.*

Proof. The probability that $\mathbf{a}_{i+1} \notin \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_i\}$ given that the latter are linearly independent is $1 - q^{-(m-i)}$. Therefore the probability that all of them are independent is

$$\prod_{i=0}^{k-1}(1 - q^{-(m-i)}) \geq 1 - \sum_{i=0}^{k-1} q^{-(m-i)} = 1 - q^{-(m-k)} \sum_{i=1}^{k} q^{-i} \geq 1 - q^{-(m-k)} \ .$$

For the last part, observe that choosing two t-dimensional subspaces through $\mathbf{x}$ is equivalent to choose $2t$ basis vectors, given that each t are linearly independent. So the probability that they intersect only on $\mathbf{x}$, is the same as that those vectors are linearly independent. Hence, by the first part, this probability is at least $1 - q^{-(m-2t)}$. □

Bibliography

[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.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, May 1998.

[Aro94] Sanjeev Arora. *Probabilistic checking of proofs and the hardness of approximation problems*. PhD thesis, University of California at Berkeley, 1994.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, January 1998.

[AS03] Sanjeev Arora and Madhu Sudan. Improved low degree testing and its applications. *Combinatorica*, 23(3):365–426, 2003. Preliminary version in Proceedings of ACM STOC 1997.

[BET10] Avraham Ben-Aroya, Klim Efremenko, and Amnon Ta-Shma. Local list decoding with a constant number of queries. In *51st IEEE Symposium on Foundations of Computer Science (FOCS)*, 2010.

[BGK+13] Eli Ben-Sasson, Ariel Gabizon, Yohay Kaplan, Swastik Kopparty, and Shubhangi Saraf. A new family of locally correctable codes based on degree-lifted algebraic geometry codes. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 833–842. ACM, 2013.

[BGM+11a] Eli Ben-Sasson, Elena Grigorescu, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. On sums of locally testable affine invariant properties. In *APPROX-RANDOM*, pages 400–411, 2011.

[BGM+11b] Eli Ben-Sasson, Elena Grigorescu, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. On sums of locally testable affine invariant properties. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:79, 2011.

[BK09] K. Brander and S. Kopparty. List-decoding Reed-Muller over large fields upto the Johnson radius. *Manuscript*, 2009.

[BKS+10] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. Optimal testing of Reed-Muller codes. In *FOCS*, pages 488–497. IEEE Computer Society, 2010.

[BLR93] 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.

[BSGH+04] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. Robust PCPs of proximity, shorter PCPs and applications to coding. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 1–10, New York, 2004. ACM Press.

[BSMSS11] Eli Ben-Sasson, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. Symmetric LDPC codes are not necessarily locally testable. In *IEEE Conference on Computational Complexity*, pages 55–65. IEEE Computer Society, 2011.

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

[BSV09a] Eli Ben-Sasson and Michael Viderman. Composition of semi-LTCs by two-wise tensor products. In Irit Dinur, Klaus Jansen, Joseph Naor, and José D. P. Rolim, editors, *APPROX-RANDOM*, volume 5687 of *Lecture Notes in Computer Science*, pages 378–391. Springer, 2009.

[BSV09b] Eli Ben-Sasson and Michael Viderman. Tensor products of weakly smooth codes are robust. *Theory of Computing*, 5(1):239–255, 2009. Preliminary version in Proc. APPROX-RANDOM 2008.

[DKSS09] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. Extensions to the method of multiplicities, with applications to kakeya sets and mergers. In *FOCS*, pages 181–190. IEEE Computer Society, 2009.

[DR04] Irit Dinur and Omer Reingold. Assignment testers: Towards a combinatorial proof of the PCP-theorem. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 155–164, Loc Alamitos, CA, USA, 2004. IEEE Press.

[DS07] Zeev Dvir and Amir Shpilka. An improved analysis of linear mergers. *Computational Complexity*, 16(1):34–59, 2007.

[DSW06] Irit Dinur, Madhu Sudan, and Avi Wigderson. Robust local testability of tensor products of LDPC codes. In Josep Díaz, Klaus Jansen, José D. P. Rolim,

and Uri Zwick, editors, *APPROX-RANDOM*, volume 4110 of *Lecture Notes in Computer Science*, pages 304–315. Springer, 2006.

[Dvi08] Zeev Dvir. On the size of Kakeya sets in finite fields. *Journal of the American Mathematical Society*, (to appear), 2008. Article electronically published on June 23, 2008.

[DW11] Zeev Dvir and Avi Wigderson. Kakeya sets, new mergers, and old extractors. *SIAM J. Comput.*, 40(3):778–792, 2011.

[FS95] Katalin Friedl and Madhu Sudan. Some improvements to total degree tests. In *Proceedings of the 3rd Annual Israel Symposium on Theory of Computing and Systems*, pages 190–198, Washington, DC, USA, 4-6 January 1995. IEEE Computer Society. Corrected version available online at `http://people.csail.mit.edu/madhu/papers/friedl.ps`.

[GGR09] Parikshit Gopalan, Venkatesan Guruswami, and Prasad Raghavendra. List decoding tensor products and interleaved codes. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 13–22, 2009.

[GHS15] Alan Guo, Elad Haramaty, and Madhu Sudan. Robust testing of lifted codes with applications to low-degree testing. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:34, 2015.

[GK14] Alan Guo and Swastik Kopparty. List-decoding algorithms for lifted codes. *CoRR*, abs/1412.0305, 2014.

[GKS13] Alan Guo, Swastik Kopparty, and Madhu Sudan. New affine-invariant codes from lifting. In Robert D. Kleinberg, editor, *Innovations in Theoretical Computer Science, ITCS '13, Berkeley, CA, USA, January 9-12, 2013*, pages 529–540. ACM, 2013.

[GS92] Peter Gemmell and Madhu Sudan. Highly resilient correctors for multivariate polynomials. *Information Processing Letters*, 43(4):169–174, September 1992.

[GS99] Venkatesan Guruswami and Madhu Sudan. Improved decoding of Reed-Solomon and algebraic-geometric codes. *IEEE Transactions on Information Theory*, 45:1757–1767, 1999. Preliminary version appeared in Proc. of FOCS 1998.

[HRS13] Elad Haramaty, Noga Ron-Zewi, and Madhu Sudan. Absolutely sound testing of lifted codes. In Prasad Raghavendra, Sofya Raskhodnikova, Klaus Jansen, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 16th International Workshop, APPROX 2013, and 17th International Workshop, RANDOM 2013, Berkeley,*

CA, USA, August 21-23, 2013. Proceedings, volume 8096 of *Lecture Notes in Computer Science*, pages 671–682. Springer, 2013.

[HSS11] Elad Haramaty, Amir Shpilka, and Madhu Sudan. Optimal testing of multivariate polynomials over small prime fields. In Rafail Ostrovsky, editor, *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 629–637. IEEE, 2011.

[JPRZ09] Charanjit S. Jutla, Anindya C. Patthak, Atri Rudra, and David Zuckerman. Testing low-degree polynomials over prime fields. *Random Struct. Algorithms*, 35(2):163–193, 2009.

[Kop12] Swastik Kopparty. List-decoding multiplicity codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:44, 2012.

[KR06] Tali Kaufman and Dana Ron. Testing polynomials over general fields. *SIAM Journal of Computing*, 36(3):779–802, 2006.

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

[KS08] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In *STOC*, pages 403–412, 2008.

[KSY14] Swastik Kopparty, Shubhangi Saraf, and Sergey Yekhanin. High-rate codes with sublinear-time decoding. *J. ACM*, 61(5):28, 2014.

[MR06] Dana Moshkovitz and Ran Raz. Sub-constant error low degree test of almost-linear size. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 21–30, 2006.

[PW04] Ruud Pellikaan and Xin-Wen Wu. List decoding of q-ary reed-muller codes. *IEEE Transactions on Information Theory*, 50(4):679–682, 2004.

[RS96] Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, April 1996.

[RS97] Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 475–484, New York, NY, 1997. ACM Press.

[SS08] Shubhangi Saraf and Madhu Sudan. Improved lower bound on the size of Kakeya sets over finite fields. *ArXiv e-prints*, August 2008.

[STV99] Madhu Sudan, Luca Trevisan, and Salil Vadhan. Pseudorandom generators without the XOR lemma. *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 537–546, 1999.

[Val05] Paul Valiant. The tensor product of two codes is not necessarily robustly testable. In Chandra Chekuri, Klaus Jansen, José D. P. Rolim, and Luca Trevisan, editors, *APPROX-RANDOM*, volume 3624 of *Lecture Notes in Computer Science*, pages 472–481. Springer, 2005.

[Vid12] Michael Viderman. A combination of testability and decodability by tensor products. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 15th International Workshop, APPROX 2012, and 16th International Workshop, RANDOM 2012, Cambridge, MA, USA, August 15-17, 2012. Proceedings*, pages 651–662, 2012.

[WB86] Lloyd R. Welch and Elwyn R. Berlekamp. Error correction of algebraic block codes. *US Patent Number 4,633,470*, December 1986.