Ideal-theoretic Explanation of Capacity-achieving Decoding [*]

Siddharth Bhandari[†]   Prahladh Harsha [‡]   Mrinal Kumar[§]   Madhu Sudan[¶]

Abstract

In this work, we present an abstract framework for some algebraic error-correcting codes with the aim of capturing codes that are list-decodable to capacity, along with their decoding algorithms. In the polynomial ideal framework, a code is specified by some ideals in a polynomial ring, messages are polynomials and the encoding of a message polynomial is the collection of residues of that polynomial modulo the ideals. We present an alternate way of viewing this class of codes in terms of linear operators, and show that this alternate view makes their algorithmic list-decodability amenable to analysis.

Our framework leads to a new class of codes that we call *affine Folded Reed-Solomon codes* (which are themselves a special case of the broader class we explore). These codes are common generalizations of the well-studied Folded Reed-Solomon codes and Univariate Multiplicity codes as well as the less-studied Additive Folded Reed-Solomon codes, and lead to a large family of codes that were not previously known/studied.

More significantly our framework also captures the algorithmic list-decodability of the constituent codes. Specifically, we present a unified view of the decoding algorithm for ideal-theoretic codes and show that the decodability reduces to the analysis of the distance of some related codes. We show that a good bound on this distance leads to a capacity-achieving performance of the underlying code, providing a unifying explanation of known capacity-achieving results. In the specific case of affine Folded Reed-Solomon codes, our framework shows that they are efficiently list-decodable up to capacity (for appropriate setting of the parameters), thereby unifying the previous results for Folded Reed-Solomon, Multiplicity and Additive Folded Reed-Solomon codes.

1 Introduction

Reed-Solomon codes are obtained by evaluations of polynomials of degree less than k at n distinct points in a finite field $\mathbb{F}$. Folded-Reed-Solomon (FRS) codes are obtained by evaluating a polynomial at sn (carefully chosen) points that are grouped into n bundles of size s each, and then viewing the resulting sn evaluations as n elements of $\mathbb{F}^s$. Multiplicity codes are obtained by evaluating the polynomial and $s-1$ of its derivatives, and again viewing the resulting sn evaluations as n elements of $\mathbb{F}^s$.

This "bundling" (or folding, as it is called for FRS codes) in FRS codes and multiplicity codes may be viewed at best as a harmless operation — it does not hurt the rate and (relative) distance of the codes, which are already optimal in these parameters. But far from merely being harmless, in the context of algorithmic list-decoding, bundling has led to remarkable improvements and to two of the very few explicit capacity achieving codes in the literature. Indeed the only other codes that achieve list-decoding capacity algorithmically and do not use one of the above codes as an ingredient are the Folded Algebraic-Geometric codes, which also use bundling. Despite this central role, the bundling operation is not well-understood algebraically. Indeed it seems like an "adhoc" operation rather than a principled one. Unearthing what bundling is and understanding when and why it turns out to be so powerful is the primary goal of this paper, and we make some progress towards this.

Turning to the algorithms for list-decoding the above codes close to capacity, there are two significantly different ones in the literature. A (later) algorithm due to Guruswami and Wang [GW13][1] which seems more generalizable, and the original algorithm of Guruswami and Rudra [GR08] which is significantly more challenging to apply to multiplicity codes (see [Kop15]). In both cases, while the algorithm for FRS works in all (reasonable) settings, the algorithms for multiplicity codes only work when the characteristic of the field is larger than the degrees of the polynomials in question. Looking more closely at FRS codes, part of the careful choice of bundling in FRS codes is to pick each bundle to be a geometric progression. If one were to switch this to an arithmetic progression, then one would get a less-studied family of codes called the Additive-FRS codes. It turns out the Additive-FRS codes are also known to be list-decodable to capacity but only via the original algorithm. We note that the skew polynomial machinery developed by Gopi and Guruswami [GG22] in the context of local reconstruction codes provides yet another proof of list-decodibility of these codes (See Section 8 for more details). Thus, the short summary of algorithmic list-decoding is that there is no short summary! Algorithms tend to work but we need to choose carefully and read the fine print.

The goal of this write-up is to provide a unifying algebraic framework that (a) captures bundling

---

[1]We note that the Guruswami-Wang algorithm is inspired by an idea due to Vadhan [Vad12, Theorem 5.24] that shows that it suffices to interpolate a polynomial Q which is linear in the y-variables. However, the algorithm from [Vad12] is not applicable to our setting since it uses polynomial factorization as well as analysis tools that are specific to Reed-Solomon codes. The further simplifications developed in [GW13] are key to the applicability in our setting.

algebraically, (b) captures most of the algorithmic success also algebraically, leaving well-defined parts for combinatorial analysis and (c) leads to new codes that also achieve capacity. In this work we use basic notions from linear algebra and polynomial rings to present a unifying definition (see Definitions 3.1 and 4.5) that captures the codes very generally, and also their efficient decoding properties (see Theorem 1.1). We also describe some new variants of these codes (see Section 3.1), that can be analyzed using this unifying framework and shown to achieve capacity We elaborate on these below.

Polynomial ideal codes. Our starting point is what we term "polynomial ideal codes". A *polynomial ideal code* over a finite field $\mathbb{F}$ and parameters k, s is specified by n pairwise relatively prime monic polynomials $E_0(X), \ldots, E_{n-1}(X) \in \mathbb{F}[X]$ of degree equal to s.[2] The encoding maps a message $p \in \mathbb{F}^k$ (interpreted as a polynomial of degree less than k) to n symbols as follows:

$$\mathbb{F}_{<k}[X] \longrightarrow (\mathbb{F}_{<s}[X])^n$$
$$p(X) \longmapsto (p(X) \pmod{E_i(X)})_{i=0}^{n-1}$$

The codes described above, Reed-Solomon, FRS, Multiplicity and Additive-FRS, are all examples of polynomial ideal codes. (A rigorous proof can be found in Section 3.1). For Reed-Solomon codes, this is folklore knowledge: the evaluation point a_i corresponds to going mod $E_i(X) = (X - a_i)$. By this we mean that value a polynomial at the evaluation point a_i is the same as the remainder obtained when the polynomial is divided by $E_i(X) = (X - a_i)$. For any bundling of the Reed-Solomon codes this follows by taking the product of the corresponding polynomials. For multiplicity codes of order s, the evaluation of a polynomial and its derivatives at a_i corresponds to going modulo $E_i(X) = (X - a_i)^s$.

The abstraction of polynomial ideal codes is not new to this work. Indeed Guruswami, Sahai and Sudan [GSS00, Appendix A] already proposed these codes as a good abstraction of algebraic codes. Their framework is even more general, in particular they even consider non-polynomial ideals such as in $\mathbb{Z}$. They suggest algorithmic possibilities but do not flesh out the details. In this work we show (see Section 6) that polynomial ideal codes, as we define them, are indeed list-decodable up to the Johnson radius. We note that the proof involves some steps not indicated in the previous work but for the most part this confirms the previous thinking.

The abstraction above also captures "bundling" (or folding) nicely - we get this by choosing $E_i(X)$ to be a product of some $E_{ij}(X)$. But the above abstraction thus far fails to capture the capacity-achieving aspects of the codes (i.e., the benefits of this bundling) and the decoding algorithms. This leads us to the two main *novel* steps of this write-up:

- We present an alternate viewpoint of polynomial ideal *codes* in terms of *linear operators*.

---

[2]Here $\mathbb{F}[X]$ refers to the ring of univariate polynomials in the variable X over the field $\mathbb{F}$ while $\mathbb{F}_{<k}[X]$ refers to the vector-space of polynomials in $\mathbb{F}[X]$ of degree strictly less than k.

- We abstract the Guruswami-Wang linear-algebraic list-decoding *algorithm* in terms of *linear operators*.

The two sets of "linear operators", in the codes and in the decoding algorithm, are not the same. But the linearity of both allows them to interact nicely with each other. We elaborate further below after introducing them.

Linear operator codes. In this write-up, a linear operator is an $\mathbb{F}$-linear function $L : \mathbb{F}[X] \to \mathbb{F}[X]$. A *linear operator code* is characterized by a family of linear operators $\mathcal{L} = (L_0, \ldots, L_{s-1})$, a set $A = \{a_0, \ldots, a_{n-1}\} \subseteq \mathbb{F}$ of evaluation points and k a degree parameter such that $k \leq s \cdot n$. The corresponding linear operator code, denoted by $LO_k^A(\mathcal{L})$, is given as follows:

$$\mathbb{F}_{<k}[X] \longrightarrow (\mathbb{F}^s)^n$$
$$p(X) \longmapsto (\mathcal{L}(p)(a_i))_{i=0}^{n-1}$$

Linear operator codes easily capture polynomial ideal codes. For instance, the multiplicity codes are linear operator codes wherein the linear operators are the successive derivative operators. But they are also too general — even if we restrict the operators to map $\mathbb{F}_{<k}[X]$ to itself, an operator allows k^2 degrees of freedom.

We narrow this broad family by looking at subfamilies of linear operators and codes. The specific subfamily we turn to are what we call "ideal linear operators". We say that linear operators $L_0, \ldots, L_{s-1}$ are *ideal linear operators* with respect to a set A of evaluation points if for every $a \in A$, the vector space

$$I^a(\mathcal{L}) = \{p \in \mathbb{F}[X] \mid \mathcal{L}(p)(a) = \bar{0}\}$$

is an ideal. (When the set of evaluation points is clear from context, we drop the phrase "with respect to A".) Linear operator codes corresponding to ideal linear operators are called *ideal linear operator codes* (see Definitions 4.1 and 4.5 for precise definitions).

It is not hard to see that a family of linear operators $\mathcal{L} = (L_0, \ldots, L_{s-1})$ has the ideal property if it satisfies the following *linearly-extendibility* property: There exists a matrix $M(X) \in \mathbb{F}[X]^{s \times s}$ such that for all $p \in \mathbb{F}[X]$ we have

$$\mathcal{L}(X \cdot p(X)) = M(X) \cdot \mathcal{L}(p(X)).$$

This motivates yet another class of linear operators and codes: We say that an operator family $\mathcal{L}$ is a *linearly-extendible linear operator* if such a matrix $M(X)$ exists and the resulting code is said to be a *linearly-extendible linear operator code* (see Definitions 4.2 and 4.5 for precise definitions).

It turns out that these three definitions of codes — polynomial ideal codes, ideal linear operator codes and linearly-extendible linear operator codes — are equivalent (see Propositions 4.7 and 4.9 and Corollary 4.10). And while the notion of polynomial ideal codes captures the codes mentioned thus far naturally, the equivalent notion of linearly-extendible codes provides a path to understanding the applicability of the linear-algebraic list-decoding algorithm of Guruswami and Wang.

While it is not the case that every linearly-extendible linear operator code (and thus every polynomial ideal code) is amenable to this list-decoding algorithm, it turns out that one can extract a nice sufficient condition on the linear-extendibility for the algorithm to be well-defined. This allows us to turn the question of list-decodability into a quantitative one — how many errors can be corrected. And the linear operator framework now converts this question into analyzing the rank of an associated matrix.

The sufficient condition we extract is the following: we say that an operator $L : \mathbb{F}[X] \to \mathbb{F}[X]$ is degree-preserving if $\deg_X(Lf) \leq \deg_X(f)$ for all $f \in \mathbb{F}[X]$. Observe that any degree-preserving linear operator when restricted to $\mathbb{F}_{<k}[X]$ can be represented by an upper-triangular matrix in $\mathbb{F}^{k \times k}$. A family of linear operators obtained by repeated iteration, $\mathcal{L} = (I = L^0, L = L^1, L^2, \ldots, L^{s-1})$ is called an *iterative* family. We associate with any degree-preserving family $\mathcal{L} = (L_0, \ldots, L_{s-1})$ of linear operators a simple matrix in $\mathbb{F}^{s \times k}$ called $\mathrm{Diag}(\mathcal{L})$, whose ith row is the diagonal of L_i and consider the code in $\mathbb{F}^k$ generated by $\mathrm{Diag}(\mathcal{L})$.

The following theorem now shows that for any degree-preserving iterative linearly-extendible operator codes, the lower bound on the distance of $\mathrm{Diag}(\mathcal{L})$ yields an upper bound on the list size obtained by the Guruswami-Wang algorithm, even when the number of errors approaches $(1 - R)$ where R is the rate of the code.

Theorem 1.1. *Suppose $\mathbb{F}$ is a field of size q and $L : \mathbb{F}[X] \to \mathbb{F}[X]$ a degree-preserving linear operator and A a set of evaluation points such that for $\mathcal{L} = (L^0, L^1, \ldots, L^{s-1})$ the corresponding code $\mathcal{C}$ is a linearly-extendible linear operator code. Furthermore, if the matrix $\mathrm{Diag}(\mathcal{L}) \in \mathbb{F}^{s \times k}$ formed by stacking the diagonals of the s linear operators as the rows is the generator matrix of a code with distance $1 - \frac{\ell}{k}$, then, $\mathcal{C}$ is code with rate $\frac{k}{sn}$ and relative distance $1 - \frac{k-1}{sn}$ over an alphabet of size q^s, and it is list-decodable up to the distance $1 - \frac{k}{(s-w+1)n} - \frac{1}{w}$ with list size q^ℓ for any $1 \leq w \leq s$.*

We remark that our actual theorem is more general (see Theorem 5.2) where we further separate the role of linear operators used to build the code from those that seed the decoding algorithm. But it immediately implies Theorem 1.1 above, which in turn already suffices to capture the capacity achieving decodability of FRS, multiplicity and additive-FRS codes. The list-decodability of multipliciy and additive-FRS codes can be proved using Theorem 1.1 by working with the linear operators $L(f(X)) = Xf'(X)$ and $L(f(X)) = X \cdot (f(X + \beta) - f(X))$ instead of the more natural operators $L(f(X)) = f'(X)$ and $L(f(X)) = f(X + \beta)$. It is to be noted that using these alternate operators does not change the underlying codes. However, this is not the approach we follow as

we use the more general Theorem 5.2 to establish the capacity achieving decodability of the above mentioned codes. Indeed the generality of the arguments allows us to capture broader families of codes uniformly, as described next.

A Common Generalization. Our framework leads very naturally to a *new* class of codes that we call the *Affine Folded Reed-Solomon* (Affine-FRS) codes: these are codes defined by ideals of the form $\prod_{i=0}^{s-1}(X - \ell^{(i)}(a))$ where $\ell(z) = \alpha z + \beta$ is any linear form and $\ell^{(i)}(z) = \underbrace{\ell(\ell \ldots \ell(z) \ldots)}_{i \text{ times}}$ denotes the i-fold composition of the linear form $\ell(z)$. These codes generalize all the previously considered codes: The case $\ell(z) = \gamma z$ are the FRS codes, the case $\ell(z) = z$ are the Multiplicity codes, and the case $\ell(z) = z + \beta$ are the Additive FRS codes!

Theorem 1.2 (Informal statement – see Theorem 8.8)**.** *Let ℓ be any linear form such that either* $\text{ord}(\ell) \geq k$ *or* $(\text{char}(\mathbb{F}) \geq k$ *and* $\beta \neq 0)$ [3]*. Then the Affine-FRS codes corresponding to the linear form ℓ are list-decodable up to capacity.*

Previously, even for the special case of the Additive FRS codes, list-decodability close to capacity was only achieved by the more involved algorithm of [GR08] and Kopparty [Kop15] (see paragraph on Additive Folding and Footnote 4 in [Gur11, Section III]). (A similar approach can be extended to cover the case of $\text{ord}(\ell) \geq k$ in Theorem 1.2: however, it seems difficult to do so for the case when $\text{ord}(\ell)$ is small.)

Thus, our Affine-FRS codes lead to the first common abstraction of the three codes as well as the first common algorithm for solving the list-decoding problem for these codes. (Furthermore, this algorithm is linear-algebraic.)' Arguably thus, even if the Affine-FRS codes had been studied previously, it is not clear that the ability to decode them for every choice of $\ell(z)$ would be obvious.

Organization

The rest of the write-up is organized as follows. We begin with some preliminaries in Section 2. We then formally define polynomial ideal codes and linear operator codes in Sections 3 and 4 respectively. In Section 5, we discuss list-decoding algorithms for polynomial ideal codes. We first present the list-decoding algorithm for *all* polynomial ideal codes up to the Johnson radius in Section 5.1 and then the list-decoding algorithm beyond the Johnson radius for special families of linear operator codes in Section 5.2. The proofs of these algorithms can be found in Sections 6 and 7 respectively. Finally, we conclude by demonstrating how these results can be used to show that several well-known families of codes (Folded Reed-Solomon, multiplicity, additive Folded Reed-Solomon codes) as well as their common generalization affine folded Reed-Solomon achieve list-decoding capacity in Section 8.

---

[3] $\text{ord}(\ell)$ refers to the smallest positive integer u such that $\ell^{(u)}(z) = z$.

2 Notation & Preliminaries

We start with some notations that we follow in the rest of this write-up.

- For a natural number n, $[n]$ denotes the set $\{0, 1, \ldots, n-1\}$.

- $\mathbb{F}$ denotes a field.

- For $a, b, i, j \in \mathbb{Z}$, where $a, b, i, j \geq 0$ the bivariate monomial $X^i Y^j$ is said to have (a, b)-weighted degree at most d if $ai + bj \leq d$. $N(a, b)$ denotes the number of bivariate monomials of $(1, a)$-weighted degree at most b.

- For $a, b \in \mathbb{Z}$, a bivariate polynomial $Q(X, Y)$ is said to have (a, b)-weighted degree at most d, if it is supported on monomials of (a, b)-weighted degree at most d.

- We say that a function $f(n) : \mathbb{N} \to \mathbb{N}$ is poly(n), if there are constants $c, n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $f(n) \leq n^c$.

- $\mathbb{F}[X]$ is the ring of univariate polynomials with coefficients in $\mathbb{F}$, and for every $k \in \mathbb{N}$, $\mathbb{F}_{<k}[X]$ denotes the set of polynomials in $\mathbb{F}[X]$ of degree strictly less than k.

- For a multivariate polynomial $f(X_0, X_1, \ldots, X_{n-1}) \in \mathbb{F}[X_0, X_1, \ldots, X_{n-1}]$, $\deg_{X_i}(f)$ denotes the degree of f, when viewing it as a univariate in X_i, with coefficients in the polynomial ring on the remaining variables over the field $\mathbb{F}$.

Estimates on number of bivariate monomials: We rely on the following simple lemma to estimate the number of bivariate monomials with $(1, a)$-weighted degree at most b.

Lemma 2.1. *For every $a, b \in \mathbb{N}$, let $N(a, b)$ denote the number of bivariate monomials with $(1, a)$-weighted degree at most b. Then, the following are true.*

1. *$N(a - 1, b) \geq b^2 / (2a)$.*

2. *For every $\eta \in \mathbb{N}$, if a divides b, then*

$$N(a, b) - N(a, b - a\eta) - \eta(b - a\eta + 1) = a\eta(\eta + 1)/2.$$

Proof. Note that from definition of $N(a, b)$, it follows that $N(a - 1, b) \geq N(a, b)$. So to prove the first item of the claim, it suffices to prove a lower bound on $N(a, b)$. Let $\tau = \lfloor \frac{b}{a} \rfloor$. Then,

$$N(a, b) = \sum_{j=0}^{\tau} \sum_{i=0}^{b-aj} 1$$

7

$$= \sum_{j=0}^{\tau} (b - aj + 1)$$

$$= (b+1)(\tau + 1) - a\tau(\tau + 1)/2$$

$$= (\tau + 1)/2 \cdot (2b + 2 - a\tau).$$

Now, plugging in the value of τ, we get the lower bound on $N(a - 1, b)$.

For the second item, we know that $\tau = b/a$ is an integer. Thus,

$$N(a, b) = (b + a)(b + 2)/(2a).$$

Now, we plug in the exact simplified expression obtained for $N(a, b)$ above in the expression $N(a, b) - N(a, b - a\eta) - \eta(b - a\eta + 1)$ that we aim to estimate, to get the following.

$$N(a, b) - N(a, b - a\eta) - \eta(b - a\eta + 1)$$

$$= (b + a)(b + 2)/(2a) - (b - a\eta + a)(b - a\eta + 2)/(2a) - \eta(b - a\eta + 1)$$

$$= \frac{1}{2a} \left((b^2 + (2 + a)b + 2a) - ((b - a\eta)^2 + (2 + a)(b - a\eta) + 2a) \right) - \eta(b - a\eta + 1)$$

$$= \frac{1}{2a} \left((2 + a)a\eta + 2ab\eta - a^2\eta^2 \right) - \eta(b - a\eta + 1)$$

$$= ((1 + a/2)\eta + b\eta - a\eta^2/2) - (b\eta - a\eta^2 + \eta)$$

$$= a\eta/2 + a\eta^2/2$$

$$= a\eta(\eta + 1)/2. \qquad \square$$

Coding theory basics:

Definition 2.2 (codes, rate, distance). *Let Σ be a finite alphabet and n be a positive integer. Given a subset $C \subseteq \Sigma^n$, define the following quantities R_c and δ_C:*

$$R_C := \frac{\log_{|\Sigma|}(|C|)}{n}, \qquad \delta_C := \min_{\substack{x,y \in C \\ x \neq y}} \left\{ \frac{\Delta(x, y)}{n} \right\}$$

where $\Delta(x, y) = |\{i \in \{1, 2, \ldots, n\} : x_i \neq y_i\}|$ denotes the Hamming distance between x and y. Then, C is said to a code *of relative distance δ_C and rate R_C with blocklength n over the alphabet Σ.* ⌟

Definition 2.3 (linear codes). *Let $\mathbb{F}_q$ be a field and let $\Sigma = \mathbb{F}_q^s$ for some positive integer s. We say that $C \subseteq (\Sigma)^n$ is a* linear code *if C is an $\mathbb{F}_q$-linear space when viewed as a subset of $\mathbb{F}_q^{sn}$.* ⌟

We note that the *standard* definition of linear codes corresponds to the case when $s = 1$. All the codes we consider in this paper will be linear under this slightly more general definition of linear codes. The above more general definition allows the alphabet to be a linear space $\mathbb{F}_q^s$ instead of

just a field $\mathbb{F}_q$. It is easy to check that the rate and distance of linear codes satisfy $R_C = \frac{\dim_{\mathbb{F}_q}(C)}{sn}$ and $\delta_C = \min_{\substack{x \in C \\ x \neq 0}} \left\{ \frac{|x|}{n} \right\}$ where $|x| = |\{i \in \{1, 2, \ldots, n\}: x_i \neq 0\}|$ denotes the Hamming weight of $x \in (\mathbb{F}_q^s)^n$.

As a consequence of the triangle inequality for Hamming distance, we have that for any code $C \subseteq \Sigma^n$ with relative distance δ and for all $x \in \Sigma^n$ the number of codewords in the ball of radius $\delta/2$ (in terms of relative distance) centered at x, i.e., $\{y \in \Sigma^n: \Delta(x, y)/n < \delta/2\}$, is at most 1. Hence, $\delta/2$ is the so-called *unique-decoding radius* for a code with relative distance δ. A natural question is to ask what happens to the number of codewords within a ball of radius larger than $\delta/2$ centered at $x \in \Sigma^n$. The following well-known fact shows that the number remains polynomial in n even when the radius of the ball grows to $1 - \sqrt{(1-\delta)}$.

Theorem 2.4 (list-decoding up to Johnson radius [GRS, Theorem 7.3.1]). *Let $q \in \mathbb{N}$ be a natural number. Any code with block length n and relative distance δ over an alphabet of size q is (combinatorially) list decodable from $(1 - \sqrt{(1-\delta)})$ fraction of errors with list size at most $n^2 q \delta$.*

We have the following bound for codes, referred to popularly as the *Singleton bound* [Sin64], though the bound appears earlier in the works of Joshi [Jos58] and Komamiya [Kom53].

Theorem 2.5 (Komamiya-Joshi-Singleton bound [GRS, Theorem 4.3.1]). *The rate R and the relative distance δ of a code satisfy $R + \delta \leq 1 + o(1)$.*

In particular, for codes which lie on the Komamiya-Joshi-Singleton bound, we have that they are combinatorially list decodable from $1 - \sqrt{R} - o(1)$ fraction of errors with polynomial list size.

List-decoding upto capacity:

Definition 2.6 (list-decoding capacity). *Consider a family of codes $\mathcal{C} = \{C_1, \ldots, C_n, \ldots\}$ where C_n has rate ρ_n and block length n with alphabet Σ_n. Then, $\mathcal{C}$ is said to achieve list-decoding capacity if $\forall \varepsilon > 0$ there exists an n_0 such that $\forall n \geq n_0$ and all received words $w \in \Sigma^n$, there exists at most a polynomial number of codewords $c \in C_n$ such that $\delta(c, w) \leq (1 - \rho_n(1 + \varepsilon))$ where $\delta(c, w)$ is the fractional distance between c and w, i.e., $\Delta(c, w)/n$.*

Further, if there exists an efficient algorithm for finding all these codewords, then, $\mathcal{C}$ is said to achieve list-decoding capacity efficiently. Ideally, we want to keep Σ_n as small as possible. ⌋

Chinese remainder theorem: We also rely on the following version of the Chinese Remainder Theorem for the polynomial ring.

Theorem 2.7 (Chinese remainder theorem [GG13, Corollary 5.3]). *Let $E_0(X), E_1(X), \ldots, E_{n-1}(X)$ be univariate polynomials of degree equal to s over a field $\mathbb{F}$ such that for every distinct $i, j \in [n]$, E_i and E_j are relatively prime. Then, for every n-tuple of polynomials $(r_0(X), \ldots, r_{n-1}(X)) \in \mathbb{F}[X]^n$ such that*

each r_i is of degree strictly less than s (or zero), there is a unique polynomial $p(X) \in \mathbb{F}[X]$ of degree at most $ns - 1$ such that for all $i \in [n]$,

$$p(X) = r_i(X) \mod E_i(X).$$

Polynomial ideals:

Definition 2.8. *A subset $I \subseteq \mathbb{F}[X]$ of polynomials is said to be an ideal if the following are true.*

- *$0 \in I$.*

- *For all $p(X), q(X) \in I$, $p + q \in I$.*

- *For every $p(X) \in I$ and $q(X) \in \mathbb{F}[X]$, $p(X) \cdot q(X) \in I$.*

For the univariate polynomial ring $\mathbb{F}[X]$, we also know that every ideal I is principal, i.e., there exists a polynomial $p(X) \in I$ such that

$$I = \{p(X)q(X) : q(X) \in \mathbb{F}[X]\}.$$

3 Polynomial Ideal Codes

In this section, we discuss polynomial ideal codes in more detail, and see how this framework captures some of the well studied families of algebraic error correcting codes.

We start with the formal definition of polynomial ideal codes.

Definition 3.1 (polynomial ideal codes)**.** *Given a field $\mathbb{F}$, parameters s, k and n satisfying $k < s \cdot n$, the polynomial ideal code is specified by a family of n polynomials $E_0, \ldots, E_{n-1}$ in the ring $\mathbb{F}[X]$ of univariate polynomials over the field $\mathbb{F}$ satisfying the following properties.*

1. *For all $i \in [n]$, polynomial E_i has degree exactly s.*

2. *The E_i's are* monic *polynomials.*

3. *The polynomials E_i's are pairwise relatively prime.*

The encoding of the polynomial ideal code maps is as follows:

$$\mathbb{F}_{<k}[X] \longrightarrow (\mathbb{F}_{<s}[X])^n$$
$$p(X) \longmapsto (p(X) \pmod{E_i(X)})_{i=0}^{n-1}$$

As is clear from the definition, polynomial ideal codes are linear over $\mathbb{F}$ and have rate code $k/(sn)$ and relative distance $(1 - (k-1)/(sn))$ (there can't be too many zeros in the encoding of a message polynomial $p(X)$, as the product of all E_i's where the encoding of $p(X)$ is zero, divides $p(X)$). Since the sum of rate and relative distance satisfy the Singleton bound, these codes are maximal-distance separable (MDS) codes.

We note that in general, E_i's need not have the same degree, but for notational convenience, we work in the setting when each of them is of degree equal to s. We also note that these codes continue to be well-defined even if the E_i's are not relatively prime. In this case, the condition, $k < s \cdot n$ is replaced by k being less than the degree of the lowest common multiple of $E_0, E_1, \ldots, E_{n-1}$. However, the distance of the code suffers in this case, and such codes need not approach the Singleton bound. We now observe that some of the standard and well studied family of algebraic error correcting codes are in fact instances of polynomial ideal codes for appropriate choice of $E_0, E_1, \ldots, E_{n-1}$.

3.1 Some Well Known Codes via Polynomial Ideals

The message space for all these codes is identified with univariate polynomials of degree at most $k - 1$ in $\mathbb{F}[X]$. We assume that the underlying field $\mathbb{F}$ is of size at least n for this discussion, else, we work over a large enough extension of $\mathbb{F}$.

Reed-Solomon Codes. Let $a_0, a_1, \ldots, a_{n-1}$ be n distinct elements of $\mathbb{F}$. In a Reed-Solomon code, we encode a message polynomial $p(X) \in \mathbb{F}[X]_{<k}$ by its evaluation on $a_0, a_1, \ldots, a_{n-1}$. To view these as a polynomial ideal code, observe that $p(a_i) = p(X) \mod (X - a_i)$. Thus, we can set the polynomials $E_i(X)$ in Definition 3.1 to be equal to $(X - a_i)$ for each $i \in [n]$. Thus, $s = 1$. Clearly, the E_i's are relatively prime since $a_0, a_1, \ldots, a_{n-1}$ are distinct.

Folded Reed-Solomon Codes [Kra03, GR08]. Let $\gamma \in \mathbb{F}_q^*$ be an element of multiplicative order at least s, i.e., $\gamma^0, \gamma, \ldots, \gamma^{s-1}$ are all distinct field elements. Further, let the set of evaluation points be $A = \{a_0, \ldots, a_{n-1}\}$ such that for any two distinct i and j the sets $\{a_i, a_i\gamma, \ldots, a_i\gamma^{s-1}\}$ and $\{a_j, a_j\gamma, \ldots, a_j\gamma^{s-1}\}$ are disjoint. In a Folded Reed-Solomon code, with block length n and folding parameter s is defined by the following encoding function.

$$p(X) \longmapsto \left(p(a_i), p(a_i\gamma^1), \ldots, p(a_i\gamma^{s-1}) \right)_{i=0}^{n-1}$$

Thus, these are codes over the alphabet $\mathbb{F}^s$.

To view these as polynomial ideal codes, we set $E_i(X) = \prod_{j=0}^{s-1}(X - a_i\gamma^j)$. Clearly, each such E_i is a polynomial of degree equal to s, and since for any two distinct i and j the sets $\{a_i, a_i\gamma, \ldots, a_i\gamma^{s-1}\}$ and $\{a_j, a_j\gamma, \ldots, a_j\gamma^{s-1}\}$ are disjoint, the polynomials $E_0, E_1, \ldots, E_{n-1}$ are all

relatively prime.

To see the equivalence between these two viewpoints observe that $p(a_i\gamma^j) = p(X) \mod (X - a_i\gamma^j)$. Moreover, $(X - a_i\gamma^j)$ are all relatively prime as j varies in $[s]$ for every $i \in [n]$. Thus, by the Chinese Remainder Theorem over $\mathbb{F}[X]$, there is a bijection between remainders of a polynomial modulo $\{(X - a_i\gamma^j) : j \in [s]\}$ and the remainder modulo the product $E_i = \prod_{j\in[s]}(X - a_i\gamma^j)$ of these polynomials.

Additive Folded Reed-Solomon Codes [GR08].   Additive Folded Reed-Solomon codes are a variant of the Folded Reed-Solomon codes defined above. Let $\mathbb{F}_q$ have characteristic at least s and let $\beta \in \mathbb{F}_q^*$. Further, let the set of evaluation points be $A = \{a_0, \ldots, a_{n-1}\}$ where $a_i - a_j \notin \{0, \beta, 2\beta, \ldots, (s-1)\beta\}$ for distinct i and j. Here, s denotes the folding parameter. The encoding is defined as follows.

$$p(X) \longmapsto (p(a_i), p(a_i + \beta), \ldots, p(a_i + \beta(s-1)))_{i=0}^{n-1}$$

Thus, these are also codes over the alphabet $\mathbb{F}^s$.

To view these as polynomial ideal codes, we set $E_i(X) = \prod_{j=0}^{s-1}(X - a_i - \beta j)$. Clearly, each such E_i is a polynomial of degree equal to s, and since $a_i - a_j \notin \{0, \beta, 2\beta, \ldots, (s-1)\beta\}$ for distinct i and j, the polynomials $E_0, E_1, \ldots, E_{n-1}$ are all relatively prime.

To see the equivalence between the two definitions, the argument is again identical to that for Folded Reed-Solomon codes discussed earlier in this section. We just observe $(X - a_i - \beta j)$ are all relatively prime j varies in $[s]$ for every $i \in [n]$, and thus by the Chinese Remainder Theorem over $\mathbb{F}[X]$, there is a bijection between remainders of a polynomial modulo $\{(X - a_i - \beta j) : j \in [s]\}$ and the remainder modulo the product $E_i = \prod_{j\in[s]}(X - a_i - \beta j)$ of these polynomials.

Univariate Multiplicity Codes [RT97, Nie01, KSY14].   Univariate multiplicity codes, or simply multiplicity codes are a variant of Reed-Solomon, where in addition to the evaluation of the message polynomial at every a_i, we also give the evaluation of its derivatives of up to order $s - 1$. While they can be defined over all fields, for the exposition in this write-up, we consider these codes over fields $\mathbb{F}$ of characteristic at least sn. Moreover, we also work with the standard derivatives (from analysis), as opposed to Hasse derivatives which is typically the convention in coding theoretic context. Let $a_0, a_1, \ldots, a_{n-1} \in \mathbb{F}$ be distinct field elements.

The encoding is defined as follows.

$$p(X) \longmapsto \left(p(a_i), \frac{\partial p}{\partial X}(a_i), \ldots, \frac{\partial^{s-1} p}{\partial X^{s-1}}(a_i) \right)_{i=0}^{n-1}$$

Here, $\frac{\partial^j p}{\partial X^{j-1}}$ denotes the (standard) jth order derivative of p with respect to X.

To view these as polynomial ideal codes, we set $E_i(X) = (X - a_i)^s$. Clearly, each such E_i is a polynomial of degree equal to s, and since a_i's are all distinct, these polynomials $E_0, E_1, \ldots, E_{n-1}$ are all relatively prime.

The equivalence of these two definitions follows from an application of Taylor's theorem to univariate polynomials, which says the following.

$$p(X) = p(a_i + X - a_i) = p(a_i) + (X - a_i)\frac{\partial p}{\partial X}(a_i) + \cdots + \frac{1}{(s-1)!}(X - a_i)^{s-1}\frac{\partial^{s-1}p}{\partial X^{s-1}}(a_i) + (X - a_i)^s \cdot q(X),$$

for some polynomial $q(X) \in \mathbb{F}[X]$. Thus,

$$p(X) \mod (X - a_i)^s = p(a_i) + (X - a_i)\frac{\partial p}{\partial X}(a_i) + \cdots + \frac{1}{(s-1)!}(X - a_i)^{s-1}\frac{\partial^{s-1}p}{\partial X^{s-1}}(a_i).$$

Therefore, we can *read* off the evaluations of the derivatives of p of order up to $s - 1$ at a_i by explicitly writing $p(X) \mod (X - a_i)^s$ as a polynomial in $(X - a_i)$ (via interpolation for instance), and reading off the various coefficients. Similarly, using the above expression, given the evaluation of all the derivatives of order up to $s - 1$ of p at a_i, we can also reconstruct $p(X) \mod (X - a_i)^s$.

Affine Folded Reed-Solomon Codes   We now describe a common generalization of the codes defined above, which we call Affine Folded Reed-Solomon Codes. Fix integers k, n, q with $n \le q$. Let $\alpha \in \mathbb{F}_q^*$ and $\beta \in \mathbb{F}_q$ such that the multiplicative order of α is u. Further, define $\ell(X) = \alpha X + \beta$ and

$$\ell^{(i)}(X) = \underbrace{\ell(\ell \ldots \ell(X))}_{i \text{ times}} = \alpha^i X + \beta \cdot \sum_{j=0}^{i-1} \alpha^j = \alpha_i X + \beta_i.$$

In fact, if $\alpha \neq 1$, i.e, $u > 1$ then, $\beta_u = \beta \cdot \sum_{i=0}^{u-1} \alpha^i = 0$, and hence, $\ell^{(u)}(X) = \ell^{(0)}(X)$. Let $\text{ord}(\ell)$ denote the smallest positive integer t such that $\ell^{(t)}(X) = X$. Note that if $\alpha \neq 1$ then $\text{ord}(\ell) = u$. The message space of the Affine Folded Reed-Solomon code of degree k with block length n and folding parameter s is polynomials of degree at most $k - 1$ over $\mathbb{F}[X]$, i.e., $\mathbb{F}_{<k}[X]$ where $\mathbb{F} = \mathbb{F}_q$. Let the set of evaluation points be $A = \{a_0, \ldots, a_{n-1}\}$ such that for distinct i, j the sets $\{\ell^{(0)}(a_i), \ldots, \ell^{(s-1)}(a_i)\}$ and $\{\ell^{(0)}(a_j), \ldots, \ell^{(s-1)}(a_j)\}$ are disjoint.

The encoding function of Affine Folded Reed-Solomon Codes is given as: (Recall that $t = \text{ord}(\ell)$; let $s = v \cdot t + r$ where $r < t$.)

13

$$p(X) \longmapsto \begin{pmatrix} p(\ell^{(0)}(a_i)) & \frac{\partial p}{\partial X}(\ell^{(0)}(a_i)) & \cdots & \frac{\partial^{v-1} p}{\partial X^{v-1}}(\ell^{(0)}(a_i)) & \frac{\partial^v p}{\partial X^v}(\ell^{(0)}(a_i)) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \frac{\partial^v p}{\partial X^v}(\ell^{(r-1)}(a_i)) \\ p(\ell^{(t-1)}(a_i)) & \frac{\partial p}{\partial X}(\ell^{(t-1)}(a_i)) & \cdots & \frac{\partial^{v-1} p}{\partial X^{v-1}}(\ell^{(t-1)}(a_i)) & \end{pmatrix}_{i=0}^{n-1}.$$

Thus, these are also codes over the alphabet $\mathbb{F}^s$.

To view these as polynomial ideal codes we set

$$E_i(X) = \prod_{j=0}^{s-1}(X - \alpha_j a_i - \beta_j) = \prod_{j=0}^{r-1}(X - \ell^{(j)}(a_i))^{v+1} \cdot \prod_{j=r}^{t-1}(X - \ell^{(j)}(a_i))^{v}.$$

For the choice of A as above, the polynomials $E_i = E_i(X)$ are pairwise co-prime. Similar to the previous cases of Folded/Additive Reed-Solomon and Multiplicy codes we have a bijection between the remainders of a polynomial modulo E_i and the encoding of the polynomial at a_i.

3.2 An Alternate Definition

We now discuss an alternate definition of polynomial ideal codes; the advantage being that this definition ties together the polynomials $E_0, E_1, \ldots, E_{n-1}$ into a single bivariate polynomial. This would be useful later on when we discuss the connection between polynomial ideal codes and linear operator codes.

Definition 3.2 (polynomial ideal codes (in terms of bivariate polynomials))**.** *Given a field $\mathbb{F}$, parameters s, k and n satisfying $k < s \cdot n$, the polynomial ideal code is specified by a bivariate polynomial $E(X, Y)$ over the field $\mathbb{F}$ and a set of n field elements $a_0, a_1 \ldots, a_{n-1}$ in $\mathbb{F}$ satisfying the following properties.*

1. *$\deg_X E(X, Y) = s$.*

2. *$E(X, Y)$ is a monic polynomial in the variable X.*

3. *The polynomials $E(X, a_i)$'s are pairwise relatively prime.*

Since E is monic and has (exact) degree s in the variable X, any polynomial $p \in \mathbb{F}[X]$ has the following unique representation.

$$p(X) = Q^{(p)}(X, Y) \cdot E(X, Y) + R^{(p)}(X, Y) \qquad \text{where } \deg_X(R^{(p)}(X, Y)) < s.$$

The encoding of the polynomial ideal code maps is as follows:

$$\mathbb{F}_{<k}[X] \longrightarrow (\mathbb{F}_{<s}[X])^n$$

$$p(X) \longmapsto \left(R^{(p)}(X, a_i) \right)_{i=0}^{n-1} .$$

⌟

The equivalence of Definitions 3.1 and 3.2 is not hard to see. We summarize this in the simple observation below.

Observation 3.3. *Definitions 3.1 and 3.2 are equivalent.*

Proof. Given a code as per Definition 3.1, we can view this as a code according to Definition 3.2 by picking n distinct $a_0, a_1, \ldots, a_{n-1} \in \mathbb{F}$ (or in a large enough extension of $\mathbb{F}$ of size at least n) and use standard Lagrange interpolation to find a bivariate polynomial $E(X, Y)$ such that for every $i \in [n]$,

$$E(X, a_i) = E_i .$$

More precisely, we define $E(X, Y)$ as follows.

$$E(X, Y) := \sum_{i \in [n]} \left(\prod_{j \in [n] \setminus \{i\}} \frac{(Y - a_j)}{(a_i - a_j)} \right) \cdot E_i(X) .$$

Clearly, $E(X, a_i)$'s are relatively prime, and their degree in X equals s and $E(X, Y)$ is monic in X. This is because the coefficient of X^s is a polynomial of degree at most $n - 1$ which takes the value 1 at $a_1, \ldots, a_n$, and so has to be the constant 1. The equivalence of the encoding function also follows immediately from the definitions.

The other direction is even simpler. Given a code as per Definition 3.2, we can view this as a code as per Definition 3.1 by just setting $E_i(X)$ to be equal to $E(X, a_i)$ for every $i \in [n]$. The condition on the degree of E_i and their relative primality follows immediately from the fact that $E(X, Y)$ is monic in X of degree s, and $E(X, a_i)$'s are relatively prime. Once again, the encoding map can be seen to be equivalent in both the cases. □

From Observation 3.3 and the discussion in Section 3.1, the Reed-Solomon codes, Folded Reed-Solomon codes, Additive Folded Reed-Solomon codes and Multiplicity codes can also be viewed as polynomial ideal codes as per Definition 3.2.

- **Reed-Solomon codes**: We take $E(X, Y)$ to be equal to $(X - Y)$, the set of points $a_0, \ldots, a_{n-1}$ remain the same.

- **Folded Reed-Solomon codes**: We take $E(X, Y) = \prod_{j \in [s]} (X - \gamma^j Y)$ and the set of evaluation points $a_0, \ldots, a_{n-1}$ are set as before, and $\gamma \in \mathbb{F}^*$ is an element of high enough order.

- **Additive Folded Reed-Solomon codes**: We take $E(X, Y) = \prod_{j \in [s]} (X - Y + \beta j)$ and the set of evaluation points $a_0, \ldots, a_{n-1}$ are set as before. Recall that $\mathbb{F}$ is taken to be a field of characteristic at least s for these codes.

- **Multiplicity codes**: We take $E(X, Y)$ to be equal to $(X - Y)^s$, the set of points $a_0, \ldots, a_{n-1}$ are distinct.

- **Affine Folded Reed-Solomon codes:** We take $E(X, Y) = \prod_{i=0}^{s-1}(X - \ell^{(i)}(Y))$ where $\ell(Y) = \alpha Y + \beta$ with $\alpha \in \mathbb{F}_q^*$ and $\beta \in \mathbb{F}_q$. Recall that the set of evaluation points $A = \{a_0, \ldots, a_{n-1}\}$ is such that for distinct i, j the sets $\left\{\ell^{(0)}(a_i), \ldots, \ell^{(s-1)}(a_i)\right\}$ and $\left\{\ell^{(0)}(a_j), \ldots, \ell^{(s-1)}(a_j)\right\}$ are disjoint.

It follows immediately from these definitions that all the desired properties in Definition 3.2 are indeed satisfied. We skip the remaining details.

4 Linear Operator Codes

In this section, we give an alternate viewpoint of polynomial ideal codes in terms of codes defined based on linear operators on the ring of polynomials.

Definition 4.1 (linear operators). *Let $\mathcal{L} = (L_0, \ldots, L_{s-1})$ be a tuple of s linear operators where each $L_i : \mathbb{F}[X] \to \mathbb{F}[X]$ is a $\mathbb{F}$-linear operator over the ring $\mathbb{F}$. For any $f \in \mathbb{F}[X]$, it will be convenient to denote by $\mathcal{L}(f)$ the (row) vector $(L_0(f), \ldots, L_{s-1}(f)) \in \mathbb{F}[X]^s$.*

Given any such family $\mathcal{L}$ and element $a \in \mathbb{F}$, define

$$I^a(\mathcal{L}) = \{p(X) \in \mathbb{F}[X] \mid \mathcal{L}(p)(a) = \bar{0}\}.$$

If the family $\mathcal{L}$ of linear operators family and the set of field elements $A \subseteq \mathbb{F}$ further satisfy the property that $I^a(\mathcal{L})$ is an ideal for each $a \in A$, we refer to the family $\mathcal{L}$ as an ideal family of linear operators *with respect to A.*

In this case, since $\mathbb{F}[X]$ is a principal ideal domain, for each $a \in A$, $I^a(\mathcal{L}) = \langle E^a(\mathcal{L})(X) \rangle$ for some monic polynomial $E^a(\mathcal{L}) \in F[X]$. ⌟

We now define a special condition on the family of linear operators $\mathcal{L}$ which will help us capture when $I^a(\mathcal{L})$ forms an ideal.

Definition 4.2 (linearly-extendible linear operators). *The family $\mathcal{L}$ of linear operators is said to be* linearly-extendible *if there exists a matrix $M(X) \in \mathbb{F}[X]^{s \times s}$ such that for all $p \in F[X]$ we have*

$$\mathcal{L}(X \cdot p(X)) = M(X) \cdot \mathcal{L}(p(X)). \tag{4.3}$$

⌟

We give two examples to illustrate the definition:

- Let $L_0(f(X)) = f(X)$ and $L_1(f(X)) = f'(X)$ where f' is the formal derivative of f. Then, by the product rule $L_1(Xf(X)) = X \cdot f'(X) + f(X)$. Hence, in this case $M(X) = \left(\begin{smallmatrix} X & 0 \\ 1 & X \end{smallmatrix}\right)$.

16

- Let $L_0(f(X)) = f(X)$ and $L_1(f(X)) = f(\gamma X)$ where $\gamma \in \mathbb{F}_q$ is non-zero. Then, we have $L_1(Xf(X)) = \gamma Xf(\gamma X)$. Hence, in this case $M(X) = \left(\begin{smallmatrix} X & 0 \\ 0 & \gamma X \end{smallmatrix}\right)$.

Observation 4.4. *Suppose $\mathcal{L}$ is linearly-extendible and $M(X)$ is the corresponding matrix from Eq. (4.3).*

- *For any $j \geq 0$ we have $\mathcal{L}(X^j \cdot p(X)) = (M(X))^j \cdot \mathcal{L}(p(X))$. Thus, by linearity we have that for any $q \in \mathbb{F}[X]$:*

$$\mathcal{L}(q(X) \cdot p(X)) = q(M(X)) \cdot \mathcal{L}(p(X)).$$

For instance if $q(X) = X^j$ then $\mathcal{L}(X^j \cdot p(X)) = (M(X))^j \cdot \mathcal{L}(p(X))$.

- *The family $\mathcal{L}$ is completely specified by $\mathcal{L}(1)$ and $M(X)$. In other words, $\mathcal{L}(p(X)) = p(M(X)) \cdot \mathcal{L}(1)$.*

- *For every set A of evaluation points, $\mathcal{L}$ is an ideal family of linear operators with respect to A. This is because if at a point a we have $\mathcal{L}(p)(a) = 0$ then $\mathcal{L}(Xp)(a) = (M(X) \cdot \mathcal{L}(p(X)))(a) = M(X = a) \cdot \mathcal{L}(p)(a) = 0$. This means that if $p(X) \in I^a(\mathcal{L})$ then $Xp(X) \in I^a(\mathcal{L})$, and hence by linearity for any $q(X) \in \mathbb{F}[X]$ we have $q(X) \cdot p(X) \in I^a(\mathcal{L})$.*

Definition 4.5 (linear operator codes). *Let $\mathcal{L} = (L_0, \ldots, L_{s-1})$ be a family of linear operators, $A = \{a_1, \ldots, a_n\} \subseteq \mathbb{F}$ be a set of evaluation points and k a degree parameter such that $k \leq s \cdot n$. Then the linear operator code generated by $\mathcal{L}$ and A, denoted by $LO_k^A(\mathcal{L})$ is given as follows:*

$$\mathbb{F}_{<k}[X] \longrightarrow (\mathbb{F}^s)^n$$
$$p(X) \longmapsto (\mathcal{L}(p)(a_i))_{i=1}^n$$

- *If $\mathcal{L}$ is an ideal family of linear operators with respect to A where the polynomials $E_i := E^{a_i}(\mathcal{L})$, which are the monic generator polynomials for the ideals $I^{a_i}(\mathcal{L})$, further satisfy the following:*

  1. *For all $i \in [n]$, polynomial E_i has degree exactly s.*
  2. *The polynomials E_i's are pairwise relatively prime.*

  Then the linear operator code is said to be an ideal linear operator code and denoted by $ILO_k^A(\mathcal{L})$.

- *If the ideal linear operator code $ILO_k^A(\mathcal{L})$ further satisfies that $\mathcal{L}$ is linearly-extendible, then the ideal linear operator code is said to be a linearly-extendible linear operator code, denoted by $LELO_k^A(\mathcal{L})$.*

⌟

Remark 4.6. *The rate of the $LO_k^A(\mathcal{L})$ code is $k/(sn)$. Further, if the the code is an ideal linear operator code, i.e., $ILO_k^A(\mathcal{L})$, then its distance is $1 - \frac{k-1}{sn}$. This is because for any message polynomial $p(X)$, the product of all E_i's where the encoding of $p(X)$ is zero, divides $p(X)$, and hence there can't be too many zeros in the encoding of $p(X)$. Hence, $ILO_k^A(\mathcal{L})$ is an MDS code.* ⌐

Proposition 4.7. *Any polynomial ideal code is a linearly-extendible linear operator code.*

Proof. Consider a polynomial ideal code given by a bivariate polynomial $E(X,Y)$ and a set of evaluation points $\{a_1, \ldots, a_n\}$ as in Definition 3.2. Recall that $E(X,Y)$ is a monic polynomial in the variable X, $\deg_X E(X,Y) = s$ and the $E(X,a_i)$'s are relatively prime. Further, any polynomial $p(X) \in \mathbb{F}[X]$ has the following unique representation.

$$p(X) = Q^{(p)}(X,Y) \cdot E(X,Y) + R^{(p)}(X,Y) \qquad \text{where } \deg_X(R^{(p)}(X,Y)) < s.$$

The encoding map of the polynomial ideal code is as follows:

$$\mathbb{F}_{<k}[X] \longrightarrow (\mathbb{F}_{<s}[X])^n$$
$$p(X) \longmapsto \left(R^{(p)}(X,a_i)\right)_{i=0}^{n-1}.$$

Let $E(X,Y) = X^s - \sum_{i=0}^{s-1} H_i(Y)X^i$ and $R^{(p)}(X,Y) = \sum_{i=0}^{s-1} R_i^p(Y)X^i$. Define $\mathcal{L} = (L_0, \ldots, L_{s-1})$ as $L_i(p(X)) = R_i^p(X)$[4]. Therefore, at any point $a \in \{a_1, \ldots, a_n\}$ we have $R^{(p)}(X,a) = \sum_{i=0}^{s-1} L_i(p(X))(a) \cdot X^i$.

Notice, that for $p(X), q(X) \in \mathbb{F}[X]$ we have $R^{(p+q)}(X,Y) = R^{(p)}(X,Y) + R^{(q)}(X,Y)$ and thus $R_i^{p+q}(Y) = R_i^p(Y) + R_i^q(Y)$ for $i < s$. This shows that L_i's are indeed linear operators. Also, $R^{(Xp)}(X,Y) = \sum_{i=1}^{s-1} R_{i-1}^p(Y)X^i + R_{s-1}^p(Y) \cdot \sum_{i=0}^{s-1} H_i(Y)X^i$. Therefore, we have $\mathcal{L}(Xp(X)) = M(X)\mathcal{L}(p(X))$ where $M(X)_{ij} = \mathbb{I}[i-1=j] + \mathbb{I}[j=s-1] \cdot H_i(X)$ for $i,j \in \{0,1,\ldots,s-1\}$. More descriptively,

$$M_{s\times s} = \begin{bmatrix} 0 & 0 & 0 & \ldots & H_0(X) \\ 1 & 0 & 0 & \ldots & H_1(X) \\ 0 & 1 & 0 & \ldots & H_2(X) \\ \vdots & \vdots & 1 & \ldots & H_3(X) \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 1 & H_{s-1}(X) \end{bmatrix}$$

Hence $\mathcal{L}$ forms a linearly-extendible set of linear operators. □

---

[4]Note that we have changed the formal variable from Y to X in the definition of $R_i^p(X)$ here.

Remark 4.8. *(degree preserving) If the bivariate polynomial $E(X, Y)$ has total degree s, then, the linear operator in the LELO code obtained above has the property that $\deg_X L_i(X^j) \le j$: in fact, $\deg_X L_i(X^j) \le j - i$.* ⌟

Proposition 4.9. *Any ideal linear operator code is a polynomial ideal code.*

Proof. Consider an ideal linear operator code $ILO_k^A(\mathcal{L})$. For any polynomial $p(X) \in \mathbb{F}[X]$ and a point $a_i \in A$, giving $\mathcal{L}(p(X))(a_i)$ is equivalent to giving $p(X) \mod \langle E_i \rangle$ where $\langle E_i \rangle = I^{a_i}(\mathcal{L})$. However, the E_is readily satisfy Definition 3.1. $\square$

Now, we state a corollary which further corroborates the notion of linear-extendibility.

Corollary 4.10 (Equivalence of *ILO* and *LELO*). *From Propositions 4.7 and 4.9 it follows that every ideal linear operator code is also a linearly-extendible linear operator code.*

Below we state some well known codes in their linear operator descriptions (a more formal treatment is given in Section 8):

- **Reed-Solomon Codes**: Let $A = \{a_0, \ldots, a_{n-1}\}$ be distinct elements in $\mathbb{F}_q$ These are $LELO_{\mathcal{L},A}$ where $\mathcal{L} = (I)$. That is the encoding of the message polynomial $p(X) \in \mathbb{F}_{<k}[X]$ at a point a is $L(f(X))(a) = f(a)$.

- **Folded Reed-Solomon Codes**: Let $\gamma \in F_q^*$ with multiplicative order at least s. $FRS[k, n]$ with folding parameter s are linearly-extendible linear operator codes $LELO_{\mathcal{L},A}$ where:

  - $\mathcal{L} = (L_0, \ldots, L_{s-1})$ with $L_1(f(X)) = f(\gamma X)$ for $f(X) \in \mathbb{F}_q[X]$ and $L_i = L_1^i$ for $i \in \{0, 1, \ldots, s-1\}$.

  - For the above family of operators $M(X)$ is given by $M(X)_{ij} = \gamma^i X \cdot \mathbb{I}[i = j]$ for $i, j \in [s]$.

  - The set of evaluation points is $A = \{a_0, \ldots, a_{n-1}\}$ where for any two distinct i and j the sets $\{a_i, a_i\gamma, \ldots, a_i\gamma^{s-1}\}$ and $\{a_j, a_j\gamma, \ldots, a_j\gamma^{s-1}\}$ are disjoint.

- **Multiplicity Codes**: Then, $MULT[k, n]$ codes of order s are linearly-extendible linear operator codes $LELO_{\mathcal{L},A}$ where:

  - $\mathcal{L} = (L_0, \ldots, L_{s-1})$ with $L_1(f(X)) = \frac{\partial f(X)}{\partial X}$ for $f(X) \in \mathbb{F}_q[X]$ and $L_i = L_1^i$ for $i \in \{0, 1, \ldots, s-1\}$.

  - For the above family of operators $M(X)$ is given by $M(X)_{ij} = X \cdot \mathbb{I}[i = j] + i \cdot \mathbb{I}[i - 1 = j]$ for $i, j \in [s]$.

  - The set of evaluation points is $A = \{a_0, \ldots, a_{n-1}\}$ where a_is are all distinct.

- **Additive Folded Reed-Solomon Codes**: Let $\beta \in \mathbb{F}_q$ be a non-zero element and the characteristic of $\mathbb{F}_q$ be at least s. Then, Additive-FRS$[k, n]$ codes with folding parameter s are linearly-extendible linear operator codes $LELO_{\mathcal{L},A}$ where:

- $\mathcal{L} = (L_0, \ldots, L_{s-1})$ with $L_1(f(X)) = f(X + \beta)$ for $f(X) \in \mathbb{F}_q[X]$ and $L_i = L_1^i$ for $i \in \{0, 1, \ldots, s-1\}$.

- For the above family of operators $M(X)$ is given by $M(X)_{ij} = (X + i\beta) \cdot \mathbb{I}[i = j]$ for $i, j \in [s]$.

- The set of evaluation points is $A = \{a_0, \ldots, a_{n-1}\}$ where $a_i - a_j \notin \{0, \beta, 2\beta, \ldots, (s-1)\beta\}$ for distinct i and j.

- **Affine Folded Reed-Solomon Codes:** Let $\alpha \in \mathbb{F}_q^*$ and $\beta \in \mathbb{F}_q$. Further, let $\ell(X) = \alpha X + \beta$ with $\mathrm{ord}(\ell) = u$. Then Affine-FRS$[k, n]$ codes with folding parameter s are linearly-extendible codes $LELO_{\mathcal{L}, A}$ described below. (See Observation 8.7 for more details.)

  Define $D_1 : \mathbb{F}[X] \to \mathbb{F}[X]$ as $D_1(f(X)) = \frac{\partial f(X)}{\partial X}$ and $S_1 : \mathbb{F}[X] \to \mathbb{F}[X]$ as $S_1(f(X)) = f(\ell(X))$. Further, for $i \geq 0$ let $D_i = D_1^i$ and $S_i = S_1^i$. Recall, that the order of α is u. For any integer $r \in [s]$ let $r = r_1 u + r_0$, with $r_0 < u$, be the unique representation of r.

  - Define $L_r : \mathbb{F}[X] \to \mathbb{F}[X]$ as $L_r(f(X)) = S_{r_0}(D_{r_1} f(X))$. Set $\mathcal{L} = (L_0, \ldots, L_{s-1})$. Clearly, $\mathcal{L}$ is a family of linear operators.

  - $L_r(Xf) = S_{r_0}(D_{r_1} Xf) = S_{r_0}(r_1 \cdot D_{r_1-1}f + X \cdot D_{r_1}f) = r_1 \cdot L_{r-u}f + S_{r_0}(X) \cdot L_r f$: hence, $\mathcal{L}$ is a set of linearly-extendible linear operators.

  - The set of evaluation points $A = \{a_0, \ldots, a_{n-1}\}$ is such that for distinct i, j the sets $\left\{ \ell^{(0)}(a_i), \ldots, \ell^{(s-1)}(a_i) \right\}$ and $\left\{ \ell^{(0)}(a_j), \ldots, \ell^{(s-1)}(a_j) \right\}$ are disjoint.

5 List-Decoding of Polynomial Ideal Codes

In this section, we discuss the list-decoding of polynomial ideal codes.

5.1 List-decoding Up to the Johnson Radius

We first observe that polynomial ideal codes are list decodable in polynomial time, up to the Johnson radius.

Theorem 5.1. *Let $k, s, n \in \mathbb{N}$ be such that $k < sn$ and $s < k - 1$. Let $E_0(X), E_1(X), \ldots, E_{n-1}(X) \in \mathbb{F}[X]$ be relatively prime monic polynomials of degree equal to s each. Let* $\mathrm{Enc} : \mathbb{F}_{<k}[X] \longrightarrow (\mathbb{F}_{<s}[X])^n$ *be the encoding function defined as*

$$p(X) \longmapsto (p(X) \pmod{E_i(X)})_{i=0}^{n-1}.$$

Then, there is an algorithm, which takes as input a received word $\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_n) \in \mathbb{F}_{<s}[X]^n$ and for every $\varepsilon > 0$ outputs all polynomials $f \in \mathbb{F}_{<k}[X]$ such that $\mathrm{Enc}(f)$ and $\mathbf{c}$ agree on at least $(k/(sn))^{1/2} + \varepsilon$ fraction of coordinates in time $\mathrm{poly}(n, 1/\varepsilon)$.

Observe that the rate of this code is $k/(sn)$ and distance is $1 - (k-1)/(sn)$, and thus Theorem 5.1 gives us an algorithmic analog of Theorem 2.4 for these codes.

The list-decoding algorithm for polynomial ideal codes is an (almost immediate) extension of an algorithm of Guruswami, Sahai and Sudan [GSS00] for list-decoding codes based on Chinese Remainder Theorem to this setting. This algorithm, in turn, relies on ideas in an earlier algorithm of Guruswami and Sudan [GS99] for list-decoding Reed-Solomon codes up to the Johnson radius.

As noted in the introduction, most of the ideas for the proof of Theorem 5.1 were already there in the work of Guruswami, Sahai and Sudan [GSS00] and all we do in this section is to flush out some of the details. The proof of this theorem is deferred to Section 6.

5.2 List-decoding beyond the Johnson Radius

In this section, we use the linear operator viewpoint of polynomial ideal codes to study their list-decodability beyond the Johnson radius. We show that if the family of linear operators $\mathcal{L}$ and the evaluation points satisfy some further properties, then the linear operator code is list-decodable all the way up to the distance of the code.

Let $\mathcal{G} = (G_0, \ldots, G_{w-1})$ and $\mathcal{T} = (T_0, T_1, \ldots, T_{r-1})$ be two families of linear operators such that $G_i : \mathbb{F}[X] \to \mathbb{F}[X]$ and $\mathcal{T}$ is a linearly-extendible family of linear operators. We say that the pair $(\mathcal{T}, \mathcal{G})$ *list-composes* in terms of $\mathcal{L}$ at the set of evaluation points A if we have the following. For every linear operator $G \in \mathcal{G}$ and field element $a \in A$, there exists a linear function $h_{G,a} : \mathbb{F}^s \to \mathbb{F}^r$ such that for every polynomial $f \in \mathbb{F}[X]$ we have

$$\mathcal{T}(G(f))(a) = h_{G,a}(\mathcal{L}(f)(a)).$$

For instance, consider the *FRS* code over $\mathbb{F}_q$ with folding parameter s and with the set of evaluation points being $A = \{a_0, \ldots, a_{n-1}\}$. The message space is polynomials of degree at most $k-1$ over $\mathbb{F}_q[X]$. This code is a linearly-extendible linear operator code where $\mathcal{L} = (L_0, \ldots, L_{s-1})$ with $L_1(f(X)) = f(\gamma X)$ for $f(X) \in \mathbb{F}_q[X]$ and $L_i = L_1^i$ for $i \in \{0, 1, \ldots, s-1\}$. Set $\mathcal{G} = (L_0, \ldots, L_{w-1})$ for some integer $w < s$ and $\mathcal{T} = (T_0, \ldots, T_{r-1})$ with $r = s - w + 1$ and $T_i = L_i$. Then, for all $G_i \in \mathcal{G}$, $T_j \in \mathcal{T}$ and $a \in A$, we have that for every polynomial $f \in \mathbb{F}[X]$: $T_j(G_i(f))(a) = L_{i+j}(f)(a)$. Notice that $L_{i+j} \in \mathcal{L}$ as $i + j \leq s - 1$. Hence, the pair $(\mathcal{T}, \mathcal{G})$ *list-composes* in terms of $\mathcal{L}$ at the set of evaluation points A.

Theorem 5.2. *If $LO_k^A(\mathcal{L})$ is a linear operator code and there exists two families of linear operators $\mathcal{G} = (G_0, \ldots, G_{w-1})$ and $\mathcal{T} = (T_0, \ldots, T_{r-1})$ such that*

1. *$(\mathcal{T}, A)$ forms a linearly-extendible linear operator code $LELO_{k+nr/w}^A(\mathcal{T})$*

2. *The pair $(\mathcal{T}, \mathcal{G})$ list-composes in terms of $\mathcal{L}$ at the set of evaluation points*

3. *$\mathcal{G}$ is degree-preserving*

4. $\mathrm{Diag}(\mathcal{G}) \in \mathbb{F}^{|\mathcal{G}| \times k}$ is the generator matrix of a code with distance $k - \ell$.

Then, $LO_k^A(\mathcal{L})$ is list-decodable up to the distance $1 - \frac{k}{rn} - \frac{1}{w}$ with list size q^ℓ.

This theorem clearly implies Theorem 1.1. Recall the hypothesis of Theorem 1.1. We instantiate $\mathcal{G}, \mathcal{T}$ in Theorem 5.2 as $(L_0, \ldots, L_{w-1})$ and $(L_0, \ldots, L_{r-1})$ respectively, with $r = s - w + 1$: properties 1, 3 and 4 above follow directly from the hypothesis of Theorem 1.1. For property 2 notice that for all $G_i \in \mathcal{G}$, $T_j \in \mathcal{T}$ and $a \in A$, we have that for every polynomial $f \in \mathbb{F}[X]$: $T_j(G_i(f))(a) = L_{i+j}(f)(a)$, and $L_{i+j} \in \mathcal{L}$ as $i + j \leq s - 1$. Hence, the pair $(\mathcal{T}, \mathcal{G})$ *list-composes* in terms of $\mathcal{L}$. Theorem 5.2 is proved in Section 7. We then use this theorem to demonstrate that several families of linear operator codes are list-decodable up to capacity in Section 8.

6 List-decoding Polynomial Ideal Codes (Proof of Theorem 5.1)

In this section we prove Theorem 5.1. The proof proceeds in three steps. In the first step, we find a bivariate polynomial $Q(X, Y)$ such that for every polynomial $f \in \mathbb{F}_{<k}[X]$, if $f \mod E_i = \mathbf{c}_i$, then $Q(X, f(X)) = 0 \mod E_i^r$. In the second step of the argument, we show that if f is such that $\mathrm{Enc}(f)$ is close enough to $\mathbf{c}$, then $Q(X, f(X))$ must be the identically zero polynomial, and therefore, $(Y - f(X))$ is a factor of $Q(X, Y)$ in the ring $\mathbb{F}[X, Y]$. In the final step of the algorithm, we factor $Q(X, Y)$ to output all factors of the form $(Y - f(X))$, where f has degree less than k and $\mathrm{Enc}(f)$ and $\mathbf{c}$ have a large agreement.

We now describe some of the details.

6.1 Interpolating a Polynomial of an Appropriate Form

Lemma 6.1. *Let $r \in \mathbb{N}$ be a parameter. Let D be be an integer such that $(nsr(r+1)k)^{1/2} < D \leq (nsr(r+1)k)^{1/2} + 1$ and let $D' \geq D$ be an integer divisible by s. Then, there exist bivariate polynomials $Q(X, Y), \{B_i(X, Y) : i \in [n]\}$ and univariate polynomials $\{A_{i,j}(X) : i \in [n], j \in \{1, 2, \ldots, r\}\}$ such that the following conditions hold.*

- *Q is not identically zero.*

- *For every $i \in [n]$,*

$$Q(X, Y) - (Y - \mathbf{c}_i)^r \cdot B_i(X, Y) + \sum_{j=1}^{r} E_i(X)^j (Y - \mathbf{c}_i)^{r-j} \cdot A_{i,j}(X) = 0.$$

- *The $(1, k-1)$-weighted degree of Q is at most D.*

- *For each $i \in [n]$, the $(1, s)$-weighted degree of B_i is at most $D' - rs$.*

- *For each $i \in [n], j \in [r] \setminus \{0\}$, the degree of $A_{i,j}(X)$ is at most $D' - rs$.*

Moreover, these polynomials can be found deterministically in time $\mathrm{poly}(n, r, s)$.

Before proceeding further, we remark that this slightly mysterious form of Q in the Lemma 6.1 is to ensure that for any $f \in \mathbb{F}[X]$, and $i \in [n]$, if $f(X) \mod E_i(X) = \mathbf{c}_i$, then $Q(X, f(X)) = 0$ mod $E_i(X)^r$. We note this in the following claim.

Claim 6.2. *Let $Q(X, Y)$, $\{B_i : i \in [n]\}$, $\{A_{i,j} : i \in [n], j \in \{1, 2, \dots, r\}\}$ be polynomials satisfying the conditions in Lemma 6.1. For any $f \in \mathbb{F}[X]$, and $i \in [n]$, if $f(X) \mod E_i(X) = \mathbf{c}_i$, then*

$$Q(X, f(X)) = 0 \mod E_i(X)^r.$$

Proof. The condition $f(X) \mod E_i(X) = \mathbf{c}_i$ implies that $(f(X) - \mathbf{c}_i) = 0 \mod E_i(X)$. Thus, for every $j \in \{0, \dots, r\}$, $E_i^j(X) \cdot (f(X) - \mathbf{c}_i)^{r-j} = 0 \mod E_i(X)^r$. Therefore, in the expression,

$$Q(X, f(X)) = (f(X) - \mathbf{c}_i)^r \cdot B_i(X, Y) + \sum_{j=1}^{r} E_i(X)^j (f(X) - \mathbf{c}_i)^{r-j} \cdot A_{i,j}(X),$$

each of the summands is divisible by $E_i(X)^r$, and hence $Q(X, f(X)) = 0 \mod E_i(X)^r$. $\square$

We now move on to the proof of Lemma 6.1.

Proof of Lemma 6.1. The proof of the lemma is by a fairly standard argument of viewing the conditions in the second item of Lemma 6.1 as homogeneous linear constraints on the coefficients of the polynomials involved and observing that there are more variables than homogeneous linear constraints, and hence there is a non-zero solution which can be found algorithmically by standard linear algebra. One subtlety here is to note that any non-zero solution of this linear system leads to a non-zero polynomial Q (required by the first item in Lemma 6.1). This observation is crucial to ensure that Q is non-zero, as apriori an arbitrary non-zero solution to this linear system could just mean that some of the coefficients of the other polynomials ($A_{i,j}$'s and B_i's) are non-zero, but they somehow cancel each other out to ensure that Q remains zero. We now argue that this cannot be the case.

Let $i \in [n]$ be such that there is a solution to this linear system where $B_i, A_{i,1}, \dots, A_{i,r}$ are not all identically zero. If Q is non-zero, then we are done. So, let us now assume that Q is zero, and argue that this cannot be the case. If B_i is non-zero, then, observe that $(Y - \mathbf{c}_i)^r B_i(X, Y)$ contains a monomial with Y-degree at least r, and this cannot be cancelled by any monomial in $\sum_{j=1}^{r} E_i(X)^j (f(X) - \mathbf{c}_i)^{r-j} \cdot A_{i,j}(X)$ since the $\deg_Y$ for this polynomial is strictly less than r. Thus, Q cannot be identically zero. Else, if B_i is identically zero, then let $j \in \{1, 2, \dots, r\}$ be the smallest index such that $A_{i,j}(X)$ is non-zero. Then, the summand $E_i(X)^j (Y - \mathbf{c}_i)^{r-j} A_{i,j}(X)$ contains a non-zero monomial with Y-degree equal to $r - j$, which cannot be cancelled out by the rest of the summands. Therefore, Q is non-zero.

23

We now count the number of homogeneous linear constraints in the system. Since $s < k - 1$, it follows that Q must also have $(1, s)$-weighted degree at most D. Moreover, since $D \leq D'$, we have that for each i, the equation

$$Q(X, Y) - (Y - \mathbf{c}_i)^r \cdot B_i(X, Y) + \sum_{j=1}^{r} E_i(X)^j (Y - \mathbf{c}_i)^{r-j} \cdot A_{i,j}(X) = 0$$

only involves monomials of $(1, s)$-weighted degree at most D'. Thus, from Lemma 2.1, each such linear constraint leads to at most $N(s, D')$ homogeneous constraints on the coefficients, where for natural numbers a, b, $N(a, b)$ denotes the number of bivariate monomials of $(1, a)$-weighted degree at most b. Since there are n such equations, the total number of homogeneous linear constraints is at most $nN(s, D')$.

To get an upper bound on the number of variables in this homogeneous linear system, observe from the weighted degree conditions and Lemma 2.1 that the number of variables to this system contributed by Q is at least $N(k - 1, D)$, by each B_i is at least $N(s, D' - rs)$ coefficients and each $A_{i,j}$ is at least $(D' - rs + 1)$ coefficients. Thus, the total number of variables is at least

$$N(k - 1, D) + n \left(N(s, D' - rs) + (D' - rs + 1) \right).$$

Thus, there exists a non-zero solution to this system of homogeneous linear equations if

$$N(k - 1, D) + n \left(N(s, D' - rs) + r(D' - rs + 1) \right) > nN(s, D'),$$

or, equivalently,

$$N(k - 1, D) > n \left(N(s, D') - \left(N(s, D' - rs) + r(D' - rs + 1) \right) \right).$$

Now, from Lemma 2.1, we know that

$$N(k - 1, D) \geq D^2/(2k) > nsr(r + 1)/2,$$

and

$$n \left(N(s, D') - \left(N(s, D' - rs) + r(D' - rs + 1) \right) \right) = nsr(r + 1)/2.$$

This last inequality follows by invoking the second item of Lemma 2.1 with $a = s, b = D', \eta = r$ (recall that D' is divisible by s). So, we get that

$$n \left(N(s, D') - \left(N(s, D' - rs) + r(D' - rs + 1) \right) \right) = nsr(r + 1)/2.$$

Thus, for our choice of parameters, we have

$$N(k-1, D) + n \left(N(s, D' - rs) + r(D' - rs + 1) \right) > nN(s, D),$$

and the system of equations must have a non-zero solution.

We can find such a non-zero solution by solving the linear system, for instance, by Gaussian elimination over $\mathbb{F}$, which runs in time polynomial in the size of the system. This completes the proof of the lemma. $\qquad \square$

6.2 Close Enough Codewords Satisfy the Equation

We now prove the following lemma which is the second step for the proof of Theorem 5.1.

Lemma 6.3. *Let D be as in Lemma 6.1 and let $Q(X, Y), \{B_i : i \in [n]\}, \{A_{i,j} : i \in [n], j \in \{1, 2, \ldots, r\}\}$ be polynomials satisfying the conditions in Lemma 6.1. And, let $f \in \mathbb{F}_{<k}[X]$ be such that $\mathrm{Enc}(f)$ and $\mathbf{c}$ agree on greater than $D/(rs)$ coordinates. Then, $Q(X, f(X))$ is identically zero.*

Proof. From Claim 6.2, we know for any $i \in [n]$, $f(X) \mod E_i = \mathbf{c}_i$ implies that $Q(X, f(X)) = 0 \mod E_i^r$. We also know from the statement of Theorem 5.1 that $E_0, E_1, \ldots, E_{n-1}$ are relatively prime. Therefore, if $S \subset [n]$ such that for all $i \in S$, $i \in [n]$, $f(X) \mod E_i = \mathbf{c}_i$, then by the Chinese Remainder Theorem (see Theorem 2.7), we have

$$Q(X, f(X)) = 0 \mod \prod_{i \in S} E_i^r.$$

We know that the degree of $Q(X, f(X))$ is at most the $(1, k-1)$-weighted degree of Q which is at most D. Moreover, the degree of $\prod_{i \in S} E_i^r$ equals $|S|sr$, which is strictly larger than D if $|S| > D/(rs)$. Thus, in this case, $Q(X, f(X)) = 0 \mod \prod_{i \in S} E_i^r$ implies that $Q(X, f(X))$ must be identically zero as a polynomial in $\mathbb{F}[X]$. $\qquad \square$

6.3 Reconstruction of All Close Enough Codewords

Finally, from Lemma 6.3, we know that for any $f \in \mathbb{F}_{<k}[X]$ such that $\mathrm{Enc}(f)$ and $\mathbf{c}$ agree on at least $D/(rs)$ coordinates, $Q(X, f(X))$ must be identically zero. Thus, to recover all such f, we use any standard polynomial factorization algorithm (e.g. the algorithm due to Kaltofen [Kal85]) to factor $Q(X, Y)$, and for every factor of the form $Y - f(X)$ such that $f(X)$ has degree less than k and $\mathrm{Enc}(f)$ and $\mathbf{c}$ agree on greater than $D/(rs)$ coordinates, include f in the output list. The list size is clearly bounded by the degree of Q, which is $\mathrm{poly}(n, r, s)$.

Thus, we have an efficient algorithm which outputs all codewords which agree with the received word on greater than

$$D/(nrs) \leq 1/(nrs) \cdot ((nsr(r+1)k)^{1/2} + 1) \leq (1/(nrs) + (k/(sn))^{1/2} \cdot (1 + 1/r)^{1/2})$$

fraction of coordinates. Choosing r to be large enough, based on ε, e.g. $r = \Theta(1/\varepsilon)$, we get Theorem 5.1.

7 List-decoding of Linear Operator Codes (Proof of Theorem 5.2)

In this section, we prove Theorem 5.2. To this end, we follow the framework of Guruswami and Wang; the key observation being that the framework is general enough to be applicable to all families of codes with properties as stated in Theorem 5.2, and not just Folded Reed-Solomon codes and Multiplicity codes, as shown by Guruswami and Wang. Before we proceed, we need some notation.

For a natural number n, $[n]$ denotes the set $\{0, 1, \ldots, n-1\}$. Recall that the alphabet of the code is $\mathbb{F}^s$, and the block length is $|A| = n$. We denote the received word by $\mathbf{c} \in \mathbb{F}^{sn}$. For notational convenience, we identify the set $[n]$ with the set A via an arbitrary ordering of the elements of A. Thus, for every $a \in A$, we use $\mathbf{c}_a \in \mathbb{F}^s$ to denote the ath coordinate of $\mathbf{c}$.

Recall that since $\mathcal{T}$ is linearly-extendible, it follows that there exists a matrix $M_{\mathcal{T}}$ such that for every polynomial $q(X) \in \mathbb{F}[X]$

$$\mathcal{T}(q(X) \cdot p(X)) = q(M_{\mathcal{T}}(X)) \cdot \mathcal{T}(p(X)).$$

The proof of Theorem 5.2, which follows the high level outline of the proof of Guruswami and Wang, follows from Lemmas 7.1 to 7.3. Lemma 7.1 shows that we can interpolate a low degree polynomial Q, with appropriately nice structure and low enough degree, which *explains* the received word $\mathbf{c}$ in some sense. We then move on to observe in Lemma 7.2 that any polynomial f such that $\mathrm{Enc}(f)$ is close enough to the received word $\mathbf{c}$ in Hamming distance *satisfies* an equation depending upon Q. Finally, in Lemma 7.3, we *solve* this equation, which is a system of homogeneous linear equations on the coefficients of f to recover all low degree polynomials f such that $\mathrm{Enc}(f)$ and $\mathbf{c}$ are close enough. Because of the linear nature of constraints, all such solutions are contained in a low dimensional linear space. As we shall observe, each of these steps in the decoding procedure just involves doing some basic linear algebra over the underlying field, and hence the decoding can be done in polynomial time by a deterministic algorithm.

We now proceed with the details of each of these steps.

7.1 Interpolating a Polynomial.

Lemma 7.1. *There exists a non-zero polynomial $Q(X, U_0, U_1, \ldots, U_{w-1}) \in \mathbb{F}[X, \mathbf{u}]$ of the form*

$$Q(X, \mathbf{u}) = \sum_{i=0}^{w-1} Q_i(X) \cdot U_i,$$

such that

- *For every $i \in [w]$, $\deg(Q_i)$ is at most $D = nr/w$.*

- *For every $a \in A$,*

$$\left(\sum_{i \in [w]} Q_i(M_{\mathcal{T}})(a) \cdot h_{G_i,a} \cdot \mathbf{c}_a \right) = 0,$$

where, (by a slight abuse of notation), we also use $h_{G_i,a}$ to denote the matrix associated with the linear transformation $h_{G_i,a}$.

Moreover, such a polynomial Q can be constructed deterministically with at most $\mathrm{poly}(n)$ operations over the underlying field $\mathbb{F}$.

Proof. The properties desired from Q in the lemma can be viewed as a system of linear constraints on the coefficients of Q. More precisely, for every $a \in A$, the condition

$$\left(\sum_{i \in [w]} Q_i(M_{\mathcal{T}})(a) \cdot h_{G_i,a} \cdot \mathbf{c}_a \right) = 0$$

imposes r homogeneous linear constraints on the coefficients of Q. The existence of a non-zero polynomial $Q = \sum_{i \in [w]} Q_i U_i$ satisfying these constraints now just follows from the fact that the number of homogeneous linear constraints is at most nr, whereas the number of variables is $(D + 1)w = (nr/w + 1)w > nr$. Thus, there is always a non-zero solution.

Since the size of the linear system is polynomially bounded in n, a non-zero Q satisfying the conditions can be found by solving the linear system, which can be done with $\mathrm{poly}(n)$ field operations using standard linear algebra algorithms. $\qquad \square$

7.2 Close Enough Codewords Satisfy the Equation.

We now argue that any polynomial $f \in \mathbb{F}[x]$ of degree at most $k - 1$, whose encoding is close enough to the received word $\mathbf{c}$ must satisfy an appropriate equation (depending upon Q).

Lemma 7.2. *If $f \in \mathbb{F}[X]$ is a polynomial of degree less than k such that for at least $n \cdot (1/w + k/(nr)) + 1$ points $a \in A$, $\mathrm{Enc}(f)(a) = \mathbf{c}_a$, then the polynomial $Q(X, G_0(f), G_1(f), \ldots, G_{w-1}(f)) \in \mathbb{F}[X]$ is identically zero.*

Proof. Let $R(X)$ be defined as

$$R(X) := Q(X, G_0(f), G_1(f), \ldots, G_{w-1}(f)) = \sum_{i \in [w]} Q_i(X) \cdot G_i(f).$$

Since f is of degree at most $k-1$ and the operators G_i do not increase the degree, R is a polynomial of degree at most $D + k - 1 \leq nr/w + k - 1$.

Let $a \in A$ be such that $\text{Enc}(f)(a) = \mathbf{c}_a$, then, we will show that $\mathcal{T}(R)$ is zero at a. Also, from the linearity of $\mathcal{T}$, it follows that

$$\mathcal{T}(R) = \mathcal{T}\left(\sum_{i \in [w]} Q_i(x) \cdot G_i(f)\right) = \left(\sum_{i \in [w]} \mathcal{T}(Q_i(X) \cdot G_i(f))\right).$$

Using linear extendibility of $\mathcal{T}$, we get

$$\mathcal{T}(R) = \left(\sum_{i \in [w]} Q_i(M_{\mathcal{T}}) \cdot \mathcal{T}(G_i(f))\right).$$

Now, since $(\mathcal{T}, \mathcal{G})$ list composes in terms of $\mathcal{L}$ at the set of evaluation points A, we know that for every $i \in [w]$, and $a \in A$, $\mathcal{T}(G_i(f))(a) = h_{G_i,a}(\mathcal{L}(f)(a))$. Therefore,

$$\mathcal{T}(R)(a) = \left(\sum_{i \in [w]} Q_i(M_{\mathcal{T}})(a) \cdot h_{G_i,a}(\mathcal{L}(f)(a))\right).$$

Since $\text{Enc}(f)(a) = \mathcal{L}(f)(a) = \mathbf{c}_a$, we get that

$$\mathcal{T}(R)(a) = \left(\sum_{i \in [w]} Q_i(M_{\mathcal{T}})(a) \cdot h_{G_i,a} \cdot \mathbf{c}_a\right).$$

Here, we abuse notation and also use $h_{G_i,a}$ to denote the matrix associated to the linear transformation given by $h_{G_i,a}$. Now, from the constraints on the polynomial Q in Lemma 7.1, we know that the right-hand side of the above equation is zero for all $a \in A$. Thus, $\mathcal{T}(R)(a)$ is zero, whenever $\text{Enc}(f)(a) = \mathbf{c}_a$ for an $a \in A$.

We now recall that since the operators $\mathcal{T}$ give us a code with rate $(D+k)/(rn) = 1/w + k/(rn)$ and distance $(1 - 1/w - k/(rn))$. Thus, if $\text{Enc}(f)$ and $\mathbf{c}$ have agreed on greater than $(1/w + k/(rn))$ fraction of points in A, R must be identically zero. □

7.3 Solving the Equation to Recover the Codewords

We now show that we can solve equations of the form

$$Q(X, G_0(f), G_1(f), \ldots, G_{w-1}(f)) = 0,$$

to recover a (small) list of all polynomials f of degree at most $k-1$ which satisfy the above equation.

Lemma 7.3. *The set of polynomials $f(X) \in \mathbb{F}[X]$ of degree at most $k - 1$ such that the polynomial $Q_0(X)G_0(f) + Q_1(X)G_1(f) + \cdots + Q_{w-1}(X) \cdot G_{w-1}(f)$ is identically zero form a linear space of dimension at most ℓ over the underlying field $\mathbb{F}$.*

Moreover, there is a deterministic algorithm which runs in polynomial time and given $Q, \mathcal{G}$ as input outputs a basis for this linear space.

Proof. From the linearity of $\mathcal{G}$, and the fact that $Q(X, \mathbf{u})$ is linear in the u variables, it immediately follows that the set of polynomials f of degree at most $k - 1$ such that

$$Q_0(X)G_0(f) + Q_1(X)G_1(f) + \cdots + Q_{w-1}(X) \cdot G_{w-1}(f) \equiv 0$$

form a linear space. Moreover, given the polynomial Q, and a description of $\mathcal{G}$, we can set up this linear system in terms of the coefficients of f and solve the system in time $\text{poly}(n)$. So, all that remains for the proof of the lemma is to argue that the dimension of this solution space is not too large. For this, we will crucially rely on the property of $\mathcal{G}$ that $\text{Diag}(\mathcal{G})$ is the generator matrix of a code of distance $k - \ell$. We start with setting up some notation.

Let $d = \max_{j \in [w]} \deg(Q_j)$, and $Q_j(X) = \sum_{i=0}^{d} q_{j,i} X^i$. From the definition of d, it follows that the vector $\tilde{q} = (q_{0,d}, q_{1,d}, \ldots, q_{w-1,d})$ is *not* the all zeros vector. For every $i, i' \in [k]$, let $g_i^j \in \mathbb{F}^k$ denote the ith row of the matrix G_j (here we are interpreting $G_j : F^{<k}[X] \to F^{<k}[x]$ as a $k \times k$ matrix) and let $g_{i,i'}^j$ denotes the (i, i') element of G_j. We also note that since $\mathcal{G}$ is a degree preserving set of linear operators, each of these matrices G_j are upper triangular (here we interpret G_j acting on vectors whose i^{th} coordinate corresponds to the coefficient of X^i and so on). Our goal is to find the set of all vectors $\text{coeff}(f) = (f_{k-1}, f_{k-2}, \ldots, f_0)$ where $f = \sum_{j=0}^{k-1} f_j x^j$ satisfies the equation

$$Q_0(X)G_0(f) + Q_1(X)G_1(f) + \cdots + Q_{w-1}(X) \cdot G_{w-1}(f) \equiv 0 \, .$$

We note that this is equivalent to saying that the coefficient of every monomial in X on the left hand side is zero. Moreover, the $Q_0(X)G_0(f) + Q_1(X)G_1(f) + \cdots + Q_{w-1}(X) \cdot G_{w-1}(f)$ is a polynomial of degree $d + k - 1$. We now chase down some of these coefficients in decreasing order of their degree, as summarised in the following simple claim.

Claim 7.4. *For each $i < k$, the coefficient of $X^{d+k-1-i}$ in $Q_0(X)G_0(f) + Q_1(X)G_1(f) + \cdots + Q_{w-1}(X) \cdot G_{w-1}(f)$ equals*

$$\sum_{j \in [w]} \left(\sum_{i'=0}^{i} q_{j,d-i'} \cdot \langle g_{k-1-(i-i')}^j, \text{coeff}(f) \rangle \right) \, .$$

From the degree preserving property of $\mathcal{G}$, we also know that the coefficient of $X^{d+k-1-i}$ in the above claim only depends on $f_{k-1-i}, f_{k-i}, \ldots f_{k-1}$. In particular, if we set up a linear system where the ith constraint equates the coefficient of $X^{d+k-1-i}$ obtained in Claim 7.4 to zero,

then resulting linear system in $(f_{k-1}, f_{k-2}, \ldots, f_1, f_0)$ is lower triangular, and the diagonal elements of the matrix of linear constraints which equals the coefficient of f_{k-1-i} in the expression $\sum_{j \in [w]} \left(\sum_{i'=0}^{i} q_{j,d-i'} \cdot \langle g^j_{k-1-(i-i')}, \mathrm{coeff}(f) \rangle \right)$ is precisely

$$\sum_{j \in [w]} q_{j,d} g^j_{k-1-i,k-1-i}.$$

We can view $\sum_{j \in [w]} q_{j,d} g^j_{k-1-i,k-1-i}$ as an inner product of the (non-zero) vector $\tilde{q} = (q_{0,d}, \ldots, q_{w-1,d})$ with the vector $v_i = (g^0_{k-1-i,k-1-i}, \ldots, g^{w-1}_{k-1-i,k-1-i})$. Now, we know that q is a non-zero vector. So, if we can ensure that at most ℓ of the coordinates of the vector $(\langle \tilde{q}, v_i \rangle : i \in [k])$ are zero, we would have the desired bound of ℓ on the dimension of the solution space.

To this end, consider the $k \times w$ matrix W, whose ith row is the vector v_i. From the definition of v_i, we can observe that the jth column of this matrix are precisely the main diagonal of the matrix G_j. From the last item in the hypothesis of Theorem 5.2, we know that this matrix W is code of distance $k - \ell$. Thus, the vector $W \cdot \tilde{q}$ can be zero on at most ℓ of its coordinates. This gives us an ℓ dimensional linear space containing all the solutions f of this equation, and therefore a bound of $|\mathbb{F}|^\ell$ on the size of this solution space. □

Remark 7.5. *Notice that in Lemma 7.3 we recovered the coefficients of the polynomial f in decreasing order of their degree. The advantage to recovering the coefficients in this order is as follows. The degree preserving nature of the G's ensures that the coefficient f_i doesn't play a role in the coefficient of $X^{d'}$ in $Q_0(X)G_0(f) + Q_1(X)G_1(f) + \cdots + Q_{w-1}(X) \cdot G_{w-1}(f)$ when d' is larger that $d - i$. This leads to a triangular system of equations whose rank can be easily inferred from the diagonal elements.* ⌟

8 Example of Codes Achieving List-Decoding Capacity

In this section we will use Theorem 5.2 to (re)prove the list-decoding capacity of the Folded Reed-Solomon codes, multiplicity codes and additive Folded Reed-Solomon codes. We then introduce a common generalization of all these codes, which we refer to as affine Folded Reed-Solomon codes and prove the list-decoding up to capacity of these codes.

We recall that Guruswami and Rudra [GR08] proved the list-decoding capacity of FRS codes, first introduced by Krachkovsky [Kra03] while Kopparty [Kop15] proved the list-decoding capacity of multiplicity codes. Guruswami and Wang [GW13] then gave an alternative and simpler linear-algebraic framework to prove the list-decoding capacity of both FRS and multiplicity codes. The list-decoding capacity of additive FRS codes is proved using the more involved algorithm of Guruswami & Rudra [GR08] and an observation of Kopparty [Kop15] (see paragraph on Additive Folding and Footnote 4 in [Gur11, Section III]). More recently (subsequent to the conference version of this paper), Gopi and Guruswami [GG22] used skew polynomials to construct improved

maximally recoverable local reconstruction codes (MR LRCs). It can be shown that the machinery of skew polynomials can be used to give yet another proof of list-decodability of FRS, multiplicity and additive FRS codes (see [GG22, Appendix C]). We remark that it is apriori unclear how to prove list-decodability of affine FRS codes using skew polynomials (or via the previous frameworks of [GR08, Kop15, GW13]).

8.1 Folded Reed-Solomon (FRS) Codes

Fix integers k, n, q with $n \leq q$. Fix $\gamma \in \mathbb{F}_q^*$ of multiplicative order at least s. The message space of the $FRS_s^{\gamma}[k, n]$ code with folding parameter s is polynomials of degree at most $k - 1$ over $\mathbb{F}[X]$, i.e., $\mathbb{F}_{<k}[X]$ where $\mathbb{F} = \mathbb{F}_q$. Then, FRS codes are linearly-extendible linear operator codes $LELO_{\mathcal{L}, A}$ where:

- $\mathcal{L} = (L_0, \ldots, L_{s-1})$ with $L_1(f(X)) = f(\gamma X)$ for $f(X) \in \mathbb{F}_q[X]$ and $L_i = L_1^i$ for $i \in \{0, 1, \ldots, s-1\}$.

- For the above family of operators $M(X)$ is given by $M(X)_{ij} = \gamma^i X \cdot \mathbb{I}[i = j]$ for $i, j \in [s]$.

- The set of evaluation points is $A = \{a_0, \ldots, a_{n-1}\}$ where for any two distinct i and j the sets $\{a_i, a_i \gamma, \ldots, a_i \gamma^{s-1}\}$ and $\{a_j, a_j \gamma, \ldots, a_j \gamma^{s-1}\}$ are disjoint.

Remark 8.1.

1. *Recall that the bivariate polynomial $E(X, Y)$ corresponding to the polynomial ideal code representation is $E(X, Y) = \prod_{i=0}^{s-1} (X - \gamma^i Y)$.*

2. *For the choice of A as above, the rate of the code is $\frac{k}{sn}$ and its distance is $1 - \frac{k-1}{sn}$ as the polynomials $E_i = E(X, a_i)$ are pairwise co-prime.*

⌟

Theorem 8.2 ([GW13]). *Let $\gamma \in \mathbb{F}_q^*$ be an element of order at least k. Further, let $A = \{a_0, \ldots, a_{n-1}\}$ be a set of evaluation points where for any two distinct i and j the sets $\{a_i, a_i \gamma, \ldots, a_i \gamma^{s-1}\}$ and $\{a_j, a_j \gamma, \ldots, a_j \gamma^{s-1}\}$ are disjoint. For every $\varepsilon > 0$ there exists s large enough ($s \geq \Omega(1/\varepsilon^2)$) such that $FRS_s^{\gamma}[k, n]$ at the set of evaluation points A can be efficiently list-decoded up to distance $1 - \frac{k}{sn} - \varepsilon$.*

Proof. We will prove this by applying Theorem 5.2. Set $\mathcal{G} = (L_0, \ldots, L_{w-1})$ for some integer $w < s$ to be set later and $\mathcal{T} = (T_0, \ldots, T_{r-1})$ with $r = s - w + 1$ and $T_i = L_i$.

Theorem 5.2-Item 1: Clearly, $(\mathcal{T}, A)$ forms a linearly-extendible linear operator code $LELO_{k+nr/w}^A(\mathcal{T})$ which is $FRS_r^{\gamma}[k + nr/w, n]$ at the set of evaluation points A.

Theorem 5.2-Item 2: For all $G_i \in \mathcal{G}$, $T_j \in \mathcal{T}$ and $a \in A$, we have that for every polynomial $f \in \mathbb{F}[X]$: $T_j(G_i(f))(a) = L_{i+j}(f)(a)$. Notice that $L_{i+j} \in \mathcal{L}$ as $i + j \leq s - 1$.

Theorem 5.2-Item 3: $G_i(x^j) = \gamma^{ij} X^j$, and hence $\mathcal{G}$ is degree preserving.

Theorem 5.2-Item 4: The matrix $\text{Diag}(\mathcal{G})$ is given by $\text{Diag}(\mathcal{G})_{ij} = \gamma^{ij}$ for $i \in [w]$ and $j \in [k]$. Hence, as long as γ has order at least k this is the generator matrix of $RS[w-1,k]$ and hence its distance is $k - w + 1$.

Thus $FRS_s^\gamma[k,n]$ can be efficiently list-decoded up to distance $1 - \frac{k-1}{rn} - \frac{1}{w}$ with list size q^{w-1}. By choosing a large enough w and s we can ensure that $1 - \frac{k-1}{rn} - \frac{1}{w} > 1 - \frac{k}{sn} - \varepsilon$. $\qquad\square$

8.2 Multiplicity (*MULT*) Codes

Fix integers k, n, q with $n \leq q$. The message space of the $MULT_s[k,n]$ code of order s is polynomials of degree at most $k - 1$ over $\mathbb{F}[X]$, i.e., $\mathbb{F}_{<k}[X]$ where $\mathbb{F} = \mathbb{F}_q$. Then, $MULT_s[k,n]$ codes are linearly-extendible linear operator codes $LELO_{\mathcal{L},A}$ where:

- $\mathcal{L} = (L_0, \ldots, L_{s-1})$ with $L_1(f(X)) = \frac{\partial f(X)}{\partial X}$ for $f(X) \in \mathbb{F}_q[X]$ and $L_i = L_1^i$ for $i \in \{0, 1, \ldots, s-1\}$.

- For the above family of operators $M(X)$ is given by $M(X)_{ij} = X \cdot \mathbb{I}[i = j] + i \cdot \mathbb{I}[i - 1 = j]$ for $i, j \in [s]$.

- The set of evaluation points is $A = \{a_0, \ldots, a_{n-1}\}$ where a_is are all distinct.

Remark 8.3.

1. *Recall that the bivariate polynomial $E(X, Y)$ corresponding to the polynomial ideal code representation is $E(X, Y) = (X - Y)^s$.*

2. *For the choice of A as above, $MULT_s[k,n]$ is a code with rate $\frac{k}{sn}$ and distance $1 - \frac{k-1}{sn}$ as the polynomials $E_i = E(X, a_i)$ are pairwise co-prime.*

$\lrcorner$

Theorem 8.4 ([GW13]). *Let the characteristic of $\mathbb{F}_q$ be at least $\max(s, k)$. Further, let the set of evaluation points be $A = \{a_0, \ldots, a_{n-1}\}$ where a_is are all distinct. Then, for every $\varepsilon > 0$ there exists s large enough $(s \geq \Omega(1/\varepsilon^2))$ such that $MULT_s[k,n]$ can be efficiently list-decoded up to distance $1 - \frac{k}{sn} - \varepsilon$.*

Proof. We will again appeal to Theorem 5.2. Set $\mathcal{G} = (G_0, \ldots, G_{w-1})$ where $G_i = \frac{X^i}{i!} \cdot L_i$ for $i \in \{0, 1, \ldots, w-1\}$ for some integer $w < s$ to be set later and $\mathcal{T} = (T_0, \ldots, T_{r-1})$ with $r = s - w + 1$ and $T_i = L_i$.

Theorem 5.2-Item 1: Clearly, $(\mathcal{T}, A)$ forms a linearly-extendible linear operator code $LELO_{k+nr/w}^A(\mathcal{T})$ which is $MULT_r[k + nr/w, n]$ of order r at the set of evaluation points A.

Theorem 5.2-Item 2: For all $G_i \in \mathcal{G}$, $T_j \in \mathcal{T}$ and $a \in A$, we have that for every polynomial $f \in \mathbb{F}[X]$:

$$T_j(G_i(f))(a) = (\sum_{b=0}^{j} \binom{j}{b}\binom{i}{b} \cdot (b!/i!) \cdot X^{i-b} L_{i+b}(f))(a).$$

Notice that the above expression only involves L_is where $i < s$.

Theorem 5.2-Item 3: $G_i(X^j) = \binom{j}{i} \cdot X^j$, and hence $\mathcal{G}$ is degree preserving.

Theorem 5.2-Item 4: The matrix $\text{Diag}(\mathcal{G})$ is given by $\text{Diag}(\mathcal{G})_{ij} = \binom{j}{i}$ for $i \in [w]$ and $j \in [k]$. This matrix can be transformed via elementary row operations to a $RS[w,k]$ generator matrix with points of evaluations as $0, 1, \ldots, k-1$; thus, as long as the characteristic of $\mathbb{F}_q$ is at least k we have that the distance of $\text{Diag}(\mathcal{G})$ is $k - w + 1$.

Thus $MULT_s[k,n]$ can be efficiently list-decoded up to distance $1 - \frac{k-1}{rn} - \frac{1}{w}$ with list size q^{w-1}. By choosing a large enough w and s we can ensure that $1 - \frac{k-1}{rn} - \frac{1}{w} > 1 - \frac{k}{sn} - \varepsilon$. $\qquad\square$

8.3 Additive Folded Reed-Solomon (Additive-FRS) Codes

Fix integers k, n, q with $n \leq q$. Let $\beta \in \mathbb{F}_q$ be a non-zero element and characteristic of $\mathbb{F}_q$ is at least s. The message space of the Additive-FRS$_s^\beta[k,n]$ code with folding parameter s is polynomials of degree at most $k - 1$ over $\mathbb{F}[X]$, i.e., $\mathbb{F}_{<k}[X]$ where $\mathbb{F} = \mathbb{F}_q$. Then, Additive-FRS$_s^\beta[k,n]$ codes are linearly-extendible linear operator codes $LELO_{\mathcal{L},A}$ where:

- $\mathcal{L} = (L_0, \ldots, L_{s-1})$ with $L_1(f(X)) = f(X + \beta)$ for $f(X) \in \mathbb{F}_q[X]$ and $L_i = L_1^i$ for $i \in \{0, 1, \ldots, s-1\}$.

- For the above family of operators $M(X)$ is given by $M(X)_{ij} = (X + i\beta) \cdot \mathbb{I}[i = j]$ for $i, j \in [s]$.

- The set of evaluation points is $A = \{a_0, \ldots, a_{n-1}\}$ where $a_i - a_j \notin \{0, \beta, 2\beta, \ldots, (s-1)\beta\}$ for distinct i and j.

Remark 8.5.

1. Recall that the bivariate polynomial $E(X, Y)$ corresponding to the polynomial ideal code representation is $E(X, Y) = \prod_{i=0}^{s-1} (X - Y - i\beta)$.

2. For the choice of A as above, Additive-FRS$_s^\beta[k,n]$ is a code with rate $\frac{k}{sn}$ and distance $1 - \frac{k-1}{sn}$ as the polynomials $E_i = E(X, a_i)$ are pairwise co-prime.

$\lrcorner$

Theorem 8.6. *Let the characteristic of $\mathbb{F}_q$ be at least $\max(s, k)$ and $\beta \in \mathbb{F}_q$ be a non-zero element. Further, let the set of evaluation points $A = \{a_0, \ldots, a_{n-1}\}$ be such that $a_i - a_j \notin \{0, \beta, 2\beta, \ldots, (s-1)\beta\}$ for distinct i and j. Then, for every $\varepsilon > 0$ there exists s large enough ($s \geq \Omega(1/\varepsilon^2)$) such that Additive-FRS$_s^\beta[k,n]$ over the set of evaluation points A can be efficiently list-decoded up to distance $1 - \frac{k}{sn} - \varepsilon$.*

Proof. We will again appeal to Theorem 5.2. To define $\mathcal{G} = (G_0, \ldots, G_{w-1})$ for some integer $w < s$, we need the following definitions. Let $B \in \mathbb{F}_q^{w \times w}$ be a matrix where $B_{ij} = (j)^i$ for $i, j \in [w]$, i.e, the transpose of the Vandermonde matrix at the points $\{0, 1, \ldots, w-1\}$: these points are distinct since the characteristic of the field is at least k. Further, let $\mathbf{b}_i \in \mathbb{F}_q^w$ be a vector such that $B\mathbf{b}_i = e_i$

33

for $i \in [w]$ where e_is are the standard basis vectors: $\mathbf{b}_i$s exist because B is full rank. Now, define $G_i = X^i \cdot \sum_{c=0}^{w-1} \mathbf{b}_i(c) L_c$ for $i \in [w]$. Set $\mathcal{T} = (T_0, \ldots, T_{r-1})$ with $r = s - w + 1$ and $T_i = L_i$.

Theorem 5.2-Item 1: Clearly, $(\mathcal{T}, A)$ forms a linearly-extendible linear operator code $LELO^A_{k+nr/w}(\mathcal{T})$ which is Additive-FRS$^\beta_r[k + nr/w, n]$ with folding parameter r at the set of evaluation points A.

Theorem 5.2-Item 2: For all $G_i \in \mathcal{G}$, $T_j \in \mathcal{T}$ and $a \in A$, we have that for every polynomial $f \in \mathbb{F}[X]$:

$$T_j(G_i(f))(a) = T_j\left(X^i \cdot \sum_{c=0}^{w-1} \mathbf{b}_i(c) L_c\right)(a)$$

$$= \left((X + j\beta)^i \cdot \sum_{c=0}^{w-1} \mathbf{b}_i(c) L_{c+j}\right)(a).$$

Notice that the above expression only involves L_is where $i < s$. Theorem 5.2-Item 3:

$$G_i(X^j) = X^i \cdot \sum_{c=0}^{w-1} \mathbf{b}_i(c) L_c(X^j)$$

$$= X^i \cdot \sum_{c=0}^{w-1} \mathbf{b}_i(c)(X + c\beta)^j$$

$$= X^i \cdot \sum_{c=0}^{w-1} \mathbf{b}_i(c) \sum_{h \le j} \binom{j}{h} X^h \cdot (c\beta)^{j-h}$$

$$= X^i \cdot \left(\binom{j}{i}\beta^i X^{j-i} + \sum_{h \le j-w} \alpha_h X^h\right)$$

(this is because $B\mathbf{b}_i = e_i$ which weans that for $h > j - w$ we have $\sum_{c=0}^{w-1} \mathbf{b}_i(c) \cdot (c)^{j-h} = \mathbb{I}[j - h = i]$; α_h are field constants)

$$= \binom{j}{i}\beta^{i-1} X^j + \ldots,$$

and hence $\mathcal{G}$ is degree preserving.

Theorem 5.2-Item 4: By the above, the matrix $\text{Diag}(\mathcal{G})$ is given by $\text{Diag}(\mathcal{G})_{ij} = \binom{j}{i}\beta^i$ for $i \in [w]$ and $j \in [k]$. Up to scaling this is the same code as $\text{Diag}(\mathcal{G})$ in Theorem 8.4: and hence, if the characteristic of the field is at least k then its distance is $k - w + 1$.

Thus Additive-FRS$^\beta_s[k, n]$ can be efficiently list-decoded up to distance $1 - \frac{k-1}{rn} - \frac{1}{w}$ with list size q^{w-1}. By choosing a large enough w and s we can ensure that $1 - \frac{k-1}{rn} - \frac{1}{w} > 1 - \frac{k}{sn} - \varepsilon$. $\qquad\square$

8.4 Affine Folded Reed-Solomon (Affine-FRS) Codes

We first recall the defintion of Affine-FRS codes. Fix integers k, n, q with $n \leq q$. Let $\alpha \in \mathbb{F}_q^*$ and $\beta \in \mathbb{F}_q$ such that the multiplicative order of α is u. Further, define $\ell(X) = \alpha X + \beta$ and

$$\ell^{(i)}(X) = \underbrace{\ell(\ell \ldots \ell(X))}_{i \text{ times}} = \alpha^i X + \beta \cdot \sum_{j=0}^{i-1} \alpha^j = \alpha_i X + \beta_i.$$

In fact, if $\alpha \neq 1$, i.e, $u > 1$ then, $\ell^{(u)}(X) = \ell^{(0)}(X)$. Let $\text{ord}(\ell)$ denote the smallest positive integer t such that $\ell^{(t)}(z) = z$. The message space of the Affine-FRS$_s^{\alpha, \beta}[k, n]$ code with folding parameter s is polynomials of degree at most $k - 1$ over $\mathbb{F}[X]$, i.e., $\mathbb{F}_{<k}[X]$ where $\mathbb{F} = \mathbb{F}_q$. Let the set of evaluation points be $A = \{a_0, \ldots, a_{n-1}\}$ such that for distinct i, j the sets $\left\{\ell^{(0)}(a_i), \ldots, \ell^{(s-1)}(a_i)\right\}$ and $\left\{\ell^{(0)}(a_j), \ldots, \ell^{(s-1)}(a_j)\right\}$ are disjoint. Then, Affine-FRS$_s^{\alpha, \beta}[k, n]$ codes are polynomial ideal codes where:

- The bivariate polynomial $E(X, Y)$ corresponding to the polynomial ideal code representation is $E(X, Y) = \prod_{i=0}^{s-1}(X - \alpha_i Y - \beta_i)$.

- For the choice of A as above, Affine-FRS$_s^{\alpha, \beta}[k, n]$ is a code with rate $\frac{k}{sn}$ and distance $1 - \frac{k-1}{sn}$ as the polynomials $E_i = E(X, a_i)$ are pairwise co-prime.

We will now recall the description of Affine-FRS codes in terms of linear operators which will be helpful while list-decoding. Define $D_1 : \mathbb{F}[X] \to \mathbb{F}[X]$ as $D_1(f(X)) = \frac{\partial f(X)}{\partial X}$ and $S_1 : \mathbb{F}[X] \to \mathbb{F}[X]$ as $S_1(f(X)) = f(\ell(X))$. Further, for $i \geq 0$ let $D_i = D_1^i$ and $S_i = S_1^i$. Recall, that the order of α is u. For any integer $r \in [s]$ let $r = r_1 u + r_0$, with $r_0 < u$, be the unique representation of r. Then, define $L_r : \mathbb{F}[X] \to \mathbb{F}[X]$ as $L_r(f(X)) = S_{r_0}(D_{r_1} f(X))$. Set $\mathcal{L} = (L_0, \ldots, L_{s-1})$. Clearly, $\mathcal{L}$ is a family of linear operators. Further, $L_r(Xf) = S_{r_0}(D_{r_1} Xf) = S_{r_0}(r_1 \cdot D_{r_1-1}f + X \cdot D_{r_1}f) = r_1 \cdot L_{r-u}f + S_{r_0}(X) \cdot L_r f$: hence, $\mathcal{L}$ is a set of linearly-extendible linear operators.

Observation 8.7. *If $u > 1$ then at an evaluation point $a \in F_q$ the following pieces of information are the same:*

- $f(X) \mod \prod_{i=0}^{s-1}(X - \alpha_i a - \beta_i)$

- $\mathcal{L}(f)(a)$.

Hence, if $u > 1$, then, Affine-FRS$_s^{\alpha, \beta}[k, n]$ at the points of evaluation A is $LELO_{\mathcal{L}, A}$.

Theorem 8.8. *For every $\varepsilon > 0$, there exists a large enough s such that the follow holds. Let $\mathbb{F}_q$ be a field, k a parameter and $\ell(X) = \alpha \cdot X + \beta$ such that $\alpha \in \mathbb{F}_q^*$ and $\beta \in \mathbb{F}_q$. Furthermore, let the evaluation points $A = \{a_0, \ldots, a_{n-1}\}$ be such that for distinct i, j the sets $\left\{\ell^{(0)}(a_i), \ldots, \ell^{(s-1)}(a_i)\right\}$ and $\left\{\ell^{(0)}(a_j), \ldots, \ell^{(s-1)}(a_j)\right\}$ are disjoint. Then, if either:*

- $\mathrm{ord}(\ell) \geq k$ *or*

- $\mathrm{char}(\mathbb{F}_q) > k$ *and* $\beta \neq 0$

holds, Affine-FRS$_s^{\alpha,\beta}[k, n]$ *over the set of evaluation points A can be efficiently list-decoded up to distance* $1 - \frac{k}{sn} - \varepsilon$.

Proof. We will again appeal to Theorem 5.2. Let u be the multiplicative order of α. Let $v = \lfloor s/u \rfloor$.

Case $\mathrm{ord}(\ell) \geq k$: This means that $u \geq k$. This is similar to decoding *FRS* codes. We skip the details.

Henceforth, we assume that $\mathrm{char}(\mathbb{F}_q) \geq k$ *and* $\beta \neq 0$.

Case $u = 1$: This is the same case as for Additive-FRS codes. Thus, by Theorem 8.6 we are done.

Case $u > 1$ **and** $v \geq \sqrt{s}$: (This case is similar to $MULT_v[k, n]$.)
Define $\mathcal{G} = (G_0, \ldots, G_{w-1})$ for some integer $w < s$, as $G_i(f) = (X^i/i!) \cdot D_i f$. Let $r = (v - w)u$ and set $\mathcal{T} = \{L_0, L_1, \ldots, L_{r-1}\}$.
Theorem 5.2-Item 1: Clearly, $(\mathcal{T}, A)$ forms a linearly-extendible linear operator code $LELO_{k+nr/w}^A(\mathcal{T})$ which is Affine-FRS$_r^{\alpha,\beta}[k + nr/w, n]$ at the set of evaluation points A.
Theorem 5.2-Item 2: For all $G_i \in \mathcal{G}$, $T_j \in \mathcal{T}$ and $a \in A$ we have that for every polynomial $f \in \mathbb{F}[X]$:

$$
\begin{aligned}
T_j(G_i(f))(a) &= \left(S_{j_0} D_{j_1} \left(\frac{X^i}{i!} \cdot D_i(f) \right) \right)(a) \\
&= \left(S_{j_0} \sum_{b=0}^{j_1} \binom{j_1}{b} \binom{i}{b} \cdot (b!/i!) \cdot X^{i-b} D_{i+b}(f) \right)(a) \\
&= \left(\sum_{b=0}^{j_1} \binom{j_1}{b} \binom{i}{b} \cdot (b!/i!) \cdot (S_{j_0} X^{i-b}) \cdot L_{j_0+(i+b)u}(f) \right)(a).
\end{aligned}
$$

Notice that the above expression only involves L_is where $i < s$.
Theorem 5.2-Items 3 and 4: are identical to the corresponding items in Theorem 8.4.
Thus Affine-FRS$_s^{\beta}[k, n]$ can be efficiently list-decoded up to distance $1 - \frac{k-1}{rn} - \frac{1}{w}$ with list size q^{w-1}. By choosing a large enough w and s we can ensure that $1 - \frac{k-1}{rn} - \frac{1}{w} > 1 - \frac{k}{sn} - \varepsilon$.

Case $u > \sqrt{s}$: (This case is similar to Additive-FRS$_u^{\beta}[k, n]$.) As in Theorem 8.6, to define $\mathcal{G} = (G_0, \ldots, G_{w-1})$ for some integer $w < u$, we need the following definitions. Let $B \in \mathbb{F}_q^{w \times w}$ be a matrix where $B_{ij} = (\beta(\alpha^j - 1)/(\alpha^j))^i$ for $i, j \in [w]$, i.e, the transpose of the Vandermonde matrix at the points $\{\beta(\alpha^j - 1)/(\alpha^j) \mid j \in [w]\}$: these points are distinct since the order of u is at least w.

Further, let $\mathbf{b}_i \in \mathbb{F}_q^w$ be a vector such that $B\mathbf{b}_i = e_i$ for $i \in [w]$ where e_is are the standard basis vectors: $\mathbf{b}_i$s exist because B is full rank.

Define $\mathcal{G} = (G_0, \ldots, G_{w-1})$ for some integer $w < s$, as $G_i = X^i \cdot \sum_{c=0}^{w-1} b_i(c) S_c$. Let $r = s - w + 1$ and set $\mathcal{T} = \{L_0, \ldots, L_{r-1}\}$.

Theorem 5.2-Item 1: Clearly, $(\mathcal{T}, A)$ forms a linearly-extendible linear operator code $LELO_{k+nr/w}^A(\mathcal{T})$ which is Affine-FRS$_r^{\alpha,\beta}[k + nr/w, n]$ at the set of evaluation points A.

Theorem 5.2-Item 2: For all $G_i \in \mathcal{G}$, $T_j \in \mathcal{T}$ and $a \in A$ we have that for every polynomial $f \in \mathbb{F}[X]$:

$$
\begin{aligned}
T_j(G_i(f))(a) &= \left(S_{j_0} D_{j_1} \left(X^i \cdot \sum_{c=0}^{w-1} b_i(c) S_c f \right) \right)(a) \\
&= \left(S_{j_0} \sum_{b=0}^{j_1} \binom{j_1}{b} \binom{i}{b} \cdot (b!) \cdot X^{i-b} D_b \left(\sum_{c=0}^{w-1} b_i(c) S_c f \right) \right)(a) \\
&= \left(S_{j_0} \sum_{b=0}^{j_1} \binom{j_1}{b} \binom{i}{b} \cdot (b!) \cdot X^{i-b} \left(\sum_{c=0}^{w-1} (b_i(c) \alpha_c^b) S_c D_b f \right) \right)(a) \\
&= \left(S_{j_0} \sum_{b=0}^{j_1} \binom{j_1}{b} \binom{i}{b} \cdot (b!) \cdot X^{i-b} \left(\sum_{c=0}^{w-1} (b_i(c) \alpha_c^b) L_{bu+c} f \right) \right)(a).
\end{aligned}
$$

Notice that the above expression only involves L_is where $i < s$.

Theorem 5.2-Items 3 and 4: follow almost identically to the corresponding items in Theorem 8.6.

Thus Affine-FRS$_s^\beta[k, n]$ can be efficiently list-decoded up to distance $1 - \frac{k-1}{rn} - \frac{1}{w}$ with list size q^{w-1}. By choosing a large enough w and s we can ensure that $1 - \frac{k-1}{rn} - \frac{1}{w} > 1 - \frac{k}{sn} - \varepsilon$. $\qquad\square$

References

[BHKS21] SIDDHARTH BHANDARI, PRAHLADH HARSHA, MRINAL KUMAR, and MADHU SUDAN. *Ideal-theoretic explanation of capacity-achieving decoding*. In MARY WOOTTERS and LAURA SANITÀ, eds., *Proc. 25th International Conf. on Randomization and Computation (RANDOM)*, volume 207 of *LIPIcs*, pages 56:1–56:21. Schloss Dagstuhl, 2021. `arXiv:2103.07930`, `eccc:2021/TR21-036`. 1

[GG13] JOACHIM VON ZUR GATHEN and JÜRGEN GERHARD. *Modern Computer Algebra*. Cambridge University Press, 3 edition, 2013. 9

[GG22] SIVAKANTH GOPI and VENKATESAN GURUSWAMI. *Improved maximally recoverable LRCs using skew polynomials*. IEEE Trans. Inform. Theory, 68(11):7198–7214, 2022. `arXiv:2012.07804`, `eccc:2021/TR21-025`. 2, 30, 31

[GR08] VENKATESAN GURUSWAMI and ATRI RUDRA. *Explicit codes achieving list decoding capacity: Error-correction with optimal redundancy*. IEEE Trans. Inform. Theory, 54(1):135–150, 2008. (Preliminary version in *38th STOC*, 2006). `arXiv:cs/0511072`, `eccc:2005/TR05-133`. 2, 6, 11, 12, 30, 31

[GRS]   VENKATESAN GURUSWAMI, ATRI RUDRA, and MADHU SUDAN. *Essential coding theory*. (draft of book). 9

[GS99]   VENKATESAN GURUSWAMI and MADHU SUDAN. *Improved decoding of Reed-Solomon and algebraic-geometry codes*. IEEE Trans. Inform. Theory, 45(6):1757–1767, 1999. (Preliminary version in *39th FOCS*, 1998). `eccc:1998/TR98-043`. 21

[GSS00]   VENKATESAN GURUSWAMI, AMIT SAHAI, and MADHU SUDAN. *"Soft-decision" decoding of Chinese Remainder Codes*. In *Proc. 41st IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 159–168. 2000. 3, 21

[Gur11]   VENKATESAN GURUSWAMI. *Linear-algebraic list decoding of folded Reed-Solomon codes*. In *Proc. 26th IEEE Conf. on Comput. Complexity*, pages 77–85. 2011. `arXiv:1106.0436`. 6, 30

[GW13]   VENKATESAN GURUSWAMI and CAROL WANG. *Linear-algebraic list decoding for variants of Reed-Solomon codes*. IEEE Trans. Inform. Theory, 59(6):3257–3268, 2013. (Preliminary version in *26th IEEE Conference on Computational Complexity*, 2011 and *15th RANDOM*, 2011). `eccc:2012/TR12-073`. 2, 30, 31, 32

[Jos58]   DURGA DATT JOSHI. *A note on upper bounds for minimum distance codes*. Information and Control, 1(3):289–295, 1958. 9

[Kal85]   ERICH L. KALTOFEN. *Polynomial-time reductions from multivariate to bi- and univariate integral polynomial factorization*. SIAM J. Comput., 14(2):469–489, 1985. 25

[Kom53]   YASUO KOMAMIYA. *Application of logical mathematics to information theory*. Proc. 3rd Japan. Nat. Cong. Appl. Math, 437, 1953. 9

[Kop15]   SWASTIK KOPPARTY. *List-decoding multiplicity codes*. Theory of Computing, 11:149–182, 2015. `eccc:2012/TR12-044`. 2, 6, 30, 31

[Kra03]   VICTOR YU. KRACHKOVSKY. *Reed-Solomon codes for correcting phased error bursts*. IEEE Trans. Inform. Theory, 49(11):2975–2984, 2003. 11, 30

[KSY14]   SWASTIK KOPPARTY, SHUBHANGI SARAF, and SERGEY YEKHANIN. *High-rate codes with sublinear-time decoding*. J. ACM, 61(5):28:1–28:20, 2014. (Preliminary version in *43rd STOC*, 2011). `eccc:2010/TR10-148`. 12

[Nie01]   RASMUS REFSLUND NIELSEN. *List decoding of linear block codes*. Ph.D. thesis, Technical University of Denmark, 2001. 12

[RT97]   M. YU ROSENBLOOM and MICHAEL ANATOLÉVICH TSFASMAN. Коды для m-метрики *(Russian) [Codes for the m-metric]*. Probl. Peredachi Inf., 33(1):55–63, 1997. (English translation in *Problems Inform. Transmission*, 33(1):45–52, 1997). 12

[Sin64]   RICHARD COLLOM SINGLETON. *Maximum distance q-nary codes*. IEEE Trans. Inform. Theory, 10(2):116–118, 1964. 9

[Vad12]   SALIL P. VADHAN. *Pseudorandomness*. Found. Trends Theor. Comput. Sci., 7(1-3):1–336, 2012. `doi:10.1561/0400000010`. 2