Decoding Analog Subspace Codes: Algorithms for Character-Polynomial Codes

Samin Riasat and Hessam Mahdavifar
Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA
Email: {riasat.s, h.mahdavifar}@northeastern.edu

***Abstract*—We propose efficient minimum-distance decoding and list-decoding algorithms for a certain class of analog subspace codes, referred to as character-polynomial (CP) codes, recently introduced by Soleymani and the second author. In particular, a CP code without its character can be viewed as a subcode of a Reed–Solomon (RS) code, where a certain subset of the coefficients of the message polynomial is set to zeros. We then demonstrate how classical decoding methods, including list decoders, for RS codes can be leveraged for decoding CP codes. For instance, it is shown that, in almost all cases, the list decoder behaves as a unique decoder. We also present a probabilistic analysis of the improvements in list decoding of CP codes when leveraging their certain structure as subcodes of RS codes.**

I. Introduction

Let W be an ambient vector space and let $\mathcal{P}(W)$ denote the set of all of subspaces of W. Then a subspace code associated with this ambient vector space is defined as a subset of $\mathcal{P}(W)$. In general, subspace codes become relevant in non-coherent communication settings where the communication medium only preserves the subspace spanned by the input vectors into the medium rather than the individual entries of the vectors. The notion of subspace codes and their application in non-coherent communication was first developed in the seminal work by Koetter and Kschischang in the context of randomized network coding [1]. Specifically, [1] develops a coding theory framework for subspace coding over finite fields. Recently, this framework is extended to the field of real/complex numbers in [2], i.e., when $W = \mathbb{R}^M$ or $W = \mathbb{C}^M$. In particular, it is shown in [2] that *analog* subspace codes enable reliable communication over wireless networks in a non-coherent fashion, where the communication medium is modeled as an *analog operator channel*. This was done by demonstrating that the subspace error- and erasure-correction capability of an analog subspace code is characterized by the minimum subspace distance, a variation of the chordal distance introduced in [2]. A new algebraic construction for one-dimensional complex subspace codes, referred to as character-polynomial codes, is introduced in [2] and later extended to higher dimensions in [3].

Generally speaking, subspace codes are used for reliable communication over analog operator channels in the same way block codes are conventionally used for reliable communication over point-to-point channels. Such codes can also be viewed as codes in "Grassmann space" (i.e. *Grassmann codes*), provided that the dimensions of all the subspace codewords are equal. Grassmann codes have also found applications in the design of communication system in the context of space-time code design for multiple-input multiple-output (MIMO) channels [4]. However, the problem of Grassmann coding in MIMO settings is often viewed as a modulation constellation design. Furthermore, even though asymptotic limits (as n grows large, and the dimension of subspace codewords is fixed) on packing in Grassmann manifolds are known [5], there are only a few explicit constructions known. The problem of designing efficient decoders for analog subspace codes/Grassmann codes is even less studied/understood in the literature.

In this paper, we study CP subspace codes from a decoding perspective. In particular, we estimate the minimum distance of one-dimensional CP codes over complex numbers and provide a minimum-distance decoding algorithm. We also present a list-decoding algorithm and show that it is efficient by analyzing its complexity. Finally, we demonstrate that the list decoder almost always returns a list of size one. This is done by a probabilistic analysis of the average list size against the well-known Guruswami–Sudan list decoder for Reed–Solomon codes [6], demonstrating the advantages of leveraging CP code structures when utilizing off-the-shelf RS decoders.

The rest of this paper is organized as follows. In § II we provide some preliminaries and background on subspace codes. In § III we present decoding algorithms for CP codes. The capabilities of the list-decoding algorithm are analyzed in § IV. Finally, the paper is concluded in § V.

II. Preliminaries: Operator Channel and Subspace Codes

A. Notation Convention

- We view the elements of W as row vectors and denote by $\mathcal{P}_k(W)$ the set of all k-dimensional subspaces of W.
- Polynomial and power series coefficients are expressed using subscripts on the symbol of the polynomial or power series, e.g. $f(X) = \sum_k f_k X^k$. This is done similarly for vector and codeword coordinates, e.g. $v = (v_1, \ldots, v_n)$.

This work was supported by NSF under Grant CCF-2415440 and the Center for Ubiquitous Connectivity (CUbiC) under the JUMP 2.0 program.

B. Analog Operator Channel

Definition 1 ([2, Definition 1])**.** An *(analog) operator channel* C associated with W is a channel where the input subspace $U \in \mathcal{P}(W)$ and the output subspace $V \in \mathcal{P}(W)$ are related by

$$V = \mathcal{H}_k(U) \oplus E,$$

where

- $\mathcal{H}_k$ is a stochastic operator that returns a random k-dimensional subspace of U. In particular, $\mathcal{H}_k(U) \in \mathcal{P}_k(U)$ if $\dim(U) > k$, and $\mathcal{H}_k(U) = U$ otherwise.
- $E \in \mathcal{P}(W)$ is an *interference/error subspace*, where, without loss of generality, $E \cap U = \{0\}$.

In transforming U to V, we say C *introduces* $t = \dim(E)$ *insertions/errors* and $\rho = \dim(V) - k$ *deletions/erasures*.

C. Grassmann Space and Subspace Codes

Definition 2 ([3, § I], [2, § II])**.** $\mathcal{P}_m(\mathbb{C}^n)$ is denoted $G_{m,n}(\mathbb{C})$ and is called the *Grassmann space*. The elements of $G_{m,n}(\mathbb{C})$ are called m*-planes*. Any m-plane U is equipped with the natural inner product $\langle u, v\rangle := uv^\dagger$ for $u, v \in U$.

Definition 3. Let $U, V \in G_{m,n}(\mathbb{C})$ be m-planes with respective orthonormal bases $\{u_i\}_{1\le i\le m}$ and $\{v_i\}_{1\le i\le m}$ (i.e. u_i, v_i are unit row vectors). The *principal angle* θ_i for $1 \le i \le m$ is $\theta_i := \arccos|u_i v_i^\dagger|$. Then the *chordal distance* [7], [2] between U and V is

$$d_{\mathrm{c}}(U, V) := \sqrt{\sum_{i=1}^{m} \sin^2\theta_i}. \tag{1}$$

We remark that this is not the only possible way to define distances between subspaces. However, as was shown in [2], a variation of this notion can perfectly capture the error- and erasure-correction capabilities of subspace codes for the analog operator channel, as defined below.

Definition 4 ([2, Definition 2])**.** The *subspace distance* of two subspaces $U, V \in \mathcal{P}(W)$ is defined as

$$d^{(\mathrm{s})}(U, V) := 2d_c(U, V)^2. \tag{2}$$

Definition 5 ([2, Definition 3])**.** A *subspace code* $\mathcal{C}$ is a collection of subspaces of W, i.e. $\mathcal{C} \subseteq \mathcal{P}(W)$. The *minimum distance* of $\mathcal{C}$ is

$$d^{(\mathrm{s})}_{\min}(\mathcal{C}) := \min_{\substack{U,V\in\mathcal{C}\\ U\neq V}} d^{(\mathrm{s})}(U, V).$$

D. Minimum-Distance Decoding

As with conventional block codes, one can associate a minimum-distance decoder to a subspace code $\mathcal{C}$ for communication over an analog operator channel in order to recover from subspace errors and erasures. Such a decoder returns the nearest codeword $V \in \mathcal{C}$ given $U \in \mathcal{P}(W)$ as its input, i.e. for any $V' \in \mathcal{C}$, $d^{(\mathrm{s})}(U, V) \le d^{(\mathrm{s})}(U, V')$.

Theorem 1 ([2, Theorem 1])**.** *Consider a subspace code $\mathcal{C}$ used for communication over an analog operator channel as in Definition 1. Then the minimum-distance decoder successfully recovers the transmitted codeword $U \in \mathcal{C}$ from the received subspace V if*

$$2(\rho + t) < d^{(\mathrm{s})}_{\min}(\mathcal{C}).$$

Theorem 1 implies that erasures and errors have equal costs in the subspace domain as far as the minimum-distance decoder is concerned. In other words, the minimum-distance decoder for a subspace code $\mathcal{C}$ can correct up to $\left\lfloor \frac{d^{(\mathrm{s})}_{\min}(\mathcal{C})-1}{2} \right\rfloor$ errors and erasures.

E. Construction of Analog Subspace Codes

We recall below the novel approach [2] based on character sums resulting in explicit constructions with better rate-minimum distance trade-off compared to the other known constructions for a wide range of parameters. We also introduce some new notation that simplifies our presentation and arguments compared to earlier work. Henceforth, $\mathbb{F}_q$ denotes a finite field of characteristic p with q elements.

Definition 6. For $k < q$, the *message space*

$$\mathcal{F}(k, q) := \{f \in \mathbb{F}_q[X] : \deg(f) \le k\}$$

is the set of all polynomials of degree at most k over $\mathbb{F}_q$. The elements of $\mathcal{F}(k, q)$ are called *message polynomials*, whose coefficients represent message symbols. For $f \in \mathcal{F}(k, q)$, we denote $f^{(p)}(X) := \sum_{0\le j\le k/p} f_{jp}X^j$ and define the additional message spaces

$$\mathcal{F}_p(k, q) := \{f(X) - f^{(p)}(X^p) : f \in \mathcal{F}(k, q)\}$$
$$\mathcal{F}_p(k, q)' := \{f(X)/X : f \in \mathcal{F}_p(k, q)\}.$$

In other words, $\mathcal{F}_p(k, q)$ is the set of all message polynomials $f \in \mathcal{F}(k, q)$ with $f_{jp} = 0$ for all integers $j \ge 0$.

Definition 7 (CP Code [3, Definition 6])**.** Fix $k \le n = q - 1$, a non-trivial character χ of $\mathbb{F}_q$, and any ordering $\alpha_1, \ldots, \alpha_n$ of $\mathbb{F}_q^\times := \mathbb{F}_q \setminus \{0\}$. Then the encoding of $f \in \mathcal{F}_p(k, q)$ in $\mathrm{CP} := \mathrm{CP}(\mathcal{F}_p(k, q), \chi) \subseteq G_{1,n}(\mathbb{C})$ is given by

$$\mathrm{CP}(f) := (\chi(f(\alpha_1)), \ldots, \chi(f(\alpha_n))).$$

We remark that all choices for the non-trivial character χ result in the same codebook [2]. Hence, we will sometimes omit one or more of the parameters k, p, q, χ when they are clear from the context. Note also that

$$|\,\mathrm{CP}\,| = |\mathcal{F}_p(k, q)| = |\mathcal{F}_p(k, q)'| = q^{k-\lfloor k/p\rfloor}, \tag{3}$$

which can be obtained from Definition 6. (See also [2, Theorem 9].)

III. Decoding CP Codes

One can observe that $\mathrm{CP}(\mathcal{F}_p(k, q), \chi)$ without χ applied to the codeword coordinates can be regarded as a subcode of a Reed–Solomon (RS) code, where a certain subset of the coefficients of the message polynomial is set to zeros. Hence, RS decoders may be used as one component in decoding CP

codes. We recall the following definitions of the well-known RS codes that will be referred to throughout this work.

Definition 8 (RS Code)**.** Fix $k \leq n \leq q$ and distinct $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$. Then the encoding of $f \in \mathcal{F}(k-1, q)$ in $\mathrm{RS} := \mathrm{RS}(\mathcal{F}(k-1,q))$ is given by

$$\mathrm{RS}(f) := (f(\alpha_1), \ldots, f(\alpha_n)). \tag{4}$$

Definition 9 (Generalized RS Code)**.** Fix $k \leq n \leq q$, distinct $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ and not necessarily distinct $v_1, \ldots, v_n \in \mathbb{F}_q^\times$. Then the encoding of $f \in \mathcal{F}(k-1,q)$ in $\mathrm{GRS} := \mathrm{GRS}(\mathcal{F}(k-1,q))$ is given by

$$\mathrm{GRS}(f) := (v_1 f(\alpha_1), \ldots, v_n f(\alpha_n)). \tag{5}$$

For the rest of this section, let $n = q-1$, $d = n-k+1 \geq 1$ and $v_i = \alpha_i$ for $i \in \{1, \ldots, n\}$. Note that $n = q-1$ is a common special case for RS and GRS, and it is well known that RS in Definition 8 and GRS in Definition 9 are $[n,k,d]_q$ codes.

Note that as in Definition 7, there was a slight abuse of notation by identifying the RS and GRS codes above with their encoding maps (4) and (5). By doing so, $\mathrm{CP}(\mathcal{F}_p(k,q),\chi)$ without χ applied to the codeword coordinates is simply expressed as $\mathrm{RS}(\mathcal{F}_p(k,q)) = \mathrm{GRS}(\mathcal{F}_p(k,q)')$ (i.e. see (9) below), which is a linear subcode of $\mathrm{GRS} \subseteq \mathrm{RS}(\mathcal{F}(k,q))$.

Throughout the following, let $\mathcal{C} = \mathrm{GRS}(\mathcal{F}_p(k,q)')$.

A. Unique Decoding

1) Minimum Distance of CP Codes: Since the minimum distance of RS and GRS as defined above is d, a RS or GRS decoder can correct $< d/2$ block errors. Consequently, the decoder applied to $\mathcal{C}$ can also correct $< d/2$ block errors. It is therefore natural to ask whether polynomials in $\mathcal{F}_p(k,q)'$ having $1 + \lfloor k/p \rfloor$ guaranteed zero coefficients can allow for correcting more errors.

Computational evidence suggests that $\mathcal{C}$

- is MDS for $p = 2$ or $q = p$,
- has minimum distance $d + \lfloor k/p \rfloor$ or $d - 1 + \lfloor k/p \rfloor$.

We try to prove some of these observations below.

Theorem 2. *The minimum distance of $\mathcal{C}$ satisfies*

$$d \leq d_{\min}(\mathcal{C}) \leq d + \left\lfloor \frac{k}{p} \right\rfloor.$$

with equality on the left if $k < p$ or $k-1$ divides $q-1$, and on the right iff $\mathcal{C}$ is MDS.

Proof. The right half of the inequality follows directly from the Singleton bound, so we prove the other half.

Since $\mathcal{C} \subseteq \mathrm{GRS}$,

$$d_{\min}(\mathcal{C}) \geq d_{\min}(\mathrm{GRS}) = d. \tag{6}$$

In the other direction, let $f(X) = X^{k-1} - 1 \in \mathcal{F}_p(k,q)'$. When $k-1$ divides $q-1$, the $k-1$ distinct roots of $f(X)$ all lie in $\mathbb{F}_q^\times$. Hence, $k-1$ of the coordinate values in $\mathcal{C}(f)$ are 0, which implies $d_{\min}(\mathcal{C}) \leq \mathrm{wt}(\mathcal{C}(f)) = d$.

Finally, if $k < p$, then $\mathcal{C} = \mathrm{GRS}$, so (6) is an equality. □

Note that one-dimensional CP subspace codes can be regarded as classical block codes by ensuring that all codewords have equal norms. Hence, their minimum Hamming distance $d_{\min}(\mathrm{CP})$ can be also studied when regarded as block codes, as in the next theorem.

Theorem 3. *The minimum distance of* CP *satisfies*

$$d_{\min}(\mathrm{CP}) \leq d + \left\lfloor \frac{k}{p} \right\rfloor \tag{7}$$

with equality iff CP *is MDS. Furthermore, if $q = p$, then*

$$d_{\min}(\mathrm{CP}) \geq d. \tag{8}$$

In particular, CP *is MDS when $k < p$ or $q = p$.*

Proof. Observe that

$$\mathrm{CP}(\mathcal{F}_p(k,q),\chi) = \{(\chi(c_1), \ldots, \chi(c_n)) : c \in \mathcal{C}\}. \tag{9}$$

Since $\chi(\mathbb{F}_q)$ is the multiplicative group of the p-th roots of unity, it follows that $d_{\min}(\mathrm{CP}) \leq d_{\min}(\mathcal{C})$, with equality if $k < p$ or $q = p$. Now (7) follows from the Singleton bound (via (3)), and (8) follows from Theorem 2. □

Note that Theorem 2 gives only sufficient conditions for equality in (6). In other words, for other values of $k \geq p$ where $k-1$ does not divide $q-1$, one may still find examples of $f \in \mathcal{F}_p(k,q)'$ with all roots in $\mathbb{F}_q^\times$, which would also imply equality in (6).

Theorem 3 implies that when q is prime, RS decoders may be used as a major component in decoding CP codes without losing any error-correction capability of the code in the worst-case scenario. The idea, based on (9), is as follows. Given any received word, we first map each coordinate to its closest point in $\Psi := \chi(\mathbb{F}_q)$ and apply χ^{-1}. Finally, we invoke any RS decoder to recover the original message. This is explained below in more detail.

Henceforth, we take $q = p$.

2) CP Decoding Algorithm: Observe that

$$\begin{aligned}
&\mathrm{CP}(\mathcal{F}_p(k,q),\chi) \\
&\quad = \{(\chi(\alpha_1 c_1), \ldots, \chi(\alpha_n c_n)) : c \in \mathrm{RS}(\mathcal{F}_p(k,q)')\} \\
&\quad \subseteq \{(\chi(\alpha_1 c_1), \ldots, \chi(\alpha_n c_n)) : c \in \mathrm{RS}(\mathcal{F}(k-1,q))\}.
\end{aligned}$$

Since Ψ is a finite cyclic subgroup of $\{z \in \mathbb{C} : |z| = 1\}$ of order q, we can map each coordinate $y_i \in \mathbb{C}$ of the received message y to its closest point $e^{2\pi i r/q}$ in Ψ, where r is the closest integer to $q \arg(y_i)/(2\pi)$. More precisely, define $\phi : \mathbb{C} \to \mathbb{F}_q$ as

$$z \mapsto \chi^{-1}\left(\exp\left(\frac{2\pi i}{q}\left\lfloor \frac{q \arg(z)}{2\pi} + \frac{1}{2} \right\rfloor\right)\right). \tag{10}$$

This leads to Algorithm 1 below.

Algorithm 1 CP-decode(m')

Input: Received word $m' \in \mathbb{C}^n$ with $< d/2$ errors
Output: Unique f with $d_{\mathrm{H}}(m', \mathrm{CP}(f)) < d/2$
Available Functions: GRS-decode: any GRS decoder

1: **for** $i = 1, \ldots, n$ **do**
2: $\quad y_i \leftarrow \phi(m'_i)$
3: **end for**
4: $g \leftarrow \mathrm{GRS\text{-}decode}(y)$
5: **return** $Xg(X)$

B. List Decoding

In list decoding, the decoder is relaxed by allowing to return a small list of codewords containing the original codeword, rather than a single codeword. This often provides better error-correction capabilities at the expense of sacrificing uniqueness. List decoding of RS codes has been extensively studied in the literature (see e.g. [6]). In particular, the well-known Guruswami–Sudan (GS) algorithm [6] for $\mathrm{GRS}(\mathcal{F}(k-1, q))$ is used as a key step in our CP list-decoding algorithm below. Here, for any received vector $y \in \mathbb{C}^n$ and an integer parameter $s > 0$ (called *interpolation multiplicity*), $\mathrm{GS}(y, s)$ returns all $f \in \mathcal{F}(k-1, q)$ such that $d_{\mathrm{H}}(y, \mathrm{GRS}(f)) \leq n - \tau_s$, where $\tau_s := \lfloor c/s \rfloor + 1$ and $c := \lfloor \sqrt{(k-1)ns(s+1)} \rfloor$. There are two major steps in the GS algorithm: the *interpolation* step and the *factorization* step.

1) CP List-Decoding Algorithm: As with unique decoding, one can use GS to list decode CP codes by following the template of Algorithm 1. However, the advantage for CP codes, as will be shown in § IV, is that one can expect a much smaller list size and error probability compared to a plain RS code of the same length and degree. The resulting algorithm is formally stated below as Algorithm 2.

Algorithm 2 CP-list-decode(m', s)

Input: Received word $m' \in \mathbb{C}^n$ with possibly $\geq d/2$ errors; interpolation multiplicity $s > 0$
Output: All $f \in \mathcal{F}_p(k, q)$ with $d_{\mathrm{H}}(m', \mathrm{CP}(f)) \leq n - \tau_s$

1: **for** $i = 1, \ldots, n$ **do**
2: $\quad y_i \leftarrow \phi(m'_i)$
3: **end for**
4: $L \leftarrow \mathrm{GS}(y, s)$
5: **return** $\{Xg(X) \in \mathcal{F}_p(k, q) : g \in L\}$

Theorem 4. CP-list-decode(m', s) *produces all codewords within distance $n - \tau_s$ of m'.*

Proof. This is a direct consequence of [6, Theorem 4.8]. □

2) Complexity Analysis: (Lines 1-3) By pre-computing $\chi(\mathbb{F}_q)$ in a hash table (using $O(n)$ space), we can compute $\phi(m'_j)$ using (10) for each j in $O(1)$ time. Therefore, we can compute y in $O(n)$ time, or $O(1)$ latency with parallelization.

(Line 4) Solving the interpolation problem in GS takes $O(n^3)$ time using the Feng–Tzeng algorithm [8, § VIII] or naive Gaussian elimination. This was improved by Kötter to $O(s^4 n^2)$ [8, § VII], which is the most efficient known solution [9, § 4].

The most efficient known solution to the factorization problem of GS is due to Gao and Shokrollahi [10] with a time complexity of $O(\ell^3 k^2)$, although the Roth–Ruckenstein algorithm [11, § V] is quite competetive [9, § 4] (see also [8, § IX]). Here ℓ is a design parameter, typically a small constant [11], which is an upper bound on the size of the list of decoded codewords, which is bounded above by the degree of the interpolation polynomial in the second variable. [6].

In general, $\ell \leq \lfloor c/(k-1) \rfloor = O(s\sqrt{n/k})$ (see also [6, Theorem 4.8]), which gives $O(\ell^3 k^2) = O(s^3 n \sqrt{kn})$, yielding a $O(s^4 n^2)$ time complexity for Line 4 of the algorithm.

IV. Probabilistic Analysis of List Decoding

In this section, we analyze the decoder error probability and the average list size (averaged over all error patterns of a given weight) of Algorithm 2.

Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a linear code. Following McEliece [8, Appendix D], for an arbitrary vector $u \in \mathbb{F}_q^n$, define

$$B_{\mathcal{C}}(u, t) := \{c \in \mathcal{C} \setminus \{0\} : d(c, u) \leq t\},$$
$$D_{\mathcal{C}}(w, t) := \sum_{\mathrm{wt}(u)=w} |B_{\mathcal{C}}(u, t)|,$$

i.e., if the all-zero vector is the transmitted codeword and u is received, then $B_{\mathcal{C}}(u, t)$ is the set of non-zero codewords at distance $\leq t$ (the *decoding radius*) from u. By linearity, if an arbitrary codeword c is the transmitted and u is the error pattern, then $|B_{\mathcal{C}}(u, t)|$ is the number of codewords $\neq c$ (the *non-casual codewords*) at distance $\leq t$ from the received word $c + u$. If $|B_{\mathcal{C}}(u, t)| = s$, we say u is *s-tuply falsely decodable*. Then, $D_{\mathcal{C}}(w, t)$ is the total number of falsely decodable words u of weight w, where an s-tuply falsely decodable word is counted s times.

Recall that a bounded distance decoder with decoding radius t returns all codewords within distance t of the received word.

Theorem 5. *Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a linear code with minimum distance d. Consider a bounded distance decoder for $\mathcal{C}$ with decoding radius t. If the error vector has weight w, then the average number of non-casual codewords (over all error patterns of weight w) in the decoding sphere is given by*

$$\overline{L}_{\mathcal{C}}(w, t) := \frac{D_{\mathcal{C}}(w, t)}{\binom{n}{w}(q-1)^w}.$$

Moreover, the probability that there exists at least one non-casual codeword within distance t of the received word is

$$P_{\mathcal{C}}(w, t) \leq \overline{L}_{\mathcal{C}}(w, t)$$

for all w, t, with equality if $t < d/2$.

Proof. This is essentially [8, Theorem D-1] (see also [9, Theorem 5.1], [12] and [13]), whose proof, although given for $[n, k]$ RS codes only, applies more generally to $\mathcal{C}$ as above. □

A. Bounds on Average List Size and Error Probability

Using the RS subcode view of CP mentioned at the beginning of § III, any bounds obtained for RS using Theorem 5 automatically apply to CP (e.g. see (11) below). However, any $f \in \mathcal{F}_p(k,q)$ has $k+1$ coefficients, $1+\lfloor k/p \rfloor$ of which are guaranteed to be zero. Therefore, we can expect a factor of $q^{1+\lfloor k/p \rfloor}$ improvement on the list size of CP over RS on average, as discussed below.

Let $\mathcal{C} = \mathrm{GRS}(\mathcal{F}_p(k,q)')$ and $\mathcal{C}' = \mathrm{GRS}(\mathcal{F}(k-1,q))$. Then CP-list-decode$(m',s)$ and $\mathrm{GS}(m',s)$ are bounded distance decoders for $\mathcal{C}$ and $\mathcal{C}'$, respectively, with decoding radius

$$\begin{aligned} t_s &:= n - \tau_s \\ &= n - 1 - \left\lfloor \frac{\lfloor \sqrt{(k-1)n(s+1)} \rfloor}{s} \right\rfloor \\ &= n - 1 - \left\lfloor \sqrt{(k-1)n\left(1+\frac{1}{s}\right)} \right\rfloor, \end{aligned}$$

where (and henceforth) $n = q-1$. Note that $t_1, t_2, \ldots$ is a bounded non-decreasing sequence of positive integers. Hence, there exists a positive integer s_0 such that

$$t_{s_0} = t_{s_0+1} = \cdots = t_\infty := n - 1 - \lfloor \sqrt{(k-1)n} \rfloor.$$

Since $\mathcal{C} \subseteq \mathcal{C}'$, we have $B_{\mathcal{C}}(u,t) \subseteq B_{\mathcal{C}'}(u,t)$. Consequently, $D_{\mathcal{C}}(w,t) \le D_{\mathcal{C}'}(w,t)$ and $\overline{L}_{\mathcal{C}}(w,t) \le \overline{L}_{\mathcal{C}'}(w,t)$. Moreover, since $|B_{\mathcal{C}}(w,t)|$ is non-decreasing in t, so are $D_{\mathcal{C}}(w,t)$ and $\overline{L}_{\mathcal{C}}(w,t)$. Hence, by Theorem 5,

$$P_{\mathcal{C}}(w,t_s) \le \overline{L}_{\mathcal{C}}(w,t_s) \le \overline{L}_{\mathcal{C}'}(w,t_s), \tag{11}$$

for all $s \ge 1$.

B. Average List Size Improvement

Assuming that the coefficients of the polynomials corresponding to the codewords in the decoding sphere of $\mathrm{GS}(m',s)$ are random, only a $(1/q)^{1+\lfloor k/p \rfloor}$ fraction of them belong to $\mathcal{F}_p(k,q)$ on average. Hence,

$$\overline{L}_{\mathcal{C}}(w,t_s) = q^{-1-\lfloor k/p \rfloor} \overline{L}_{\mathcal{C}'}(w,t_s).$$

Then, by Theorem 5,

$$P_{\mathcal{C}}(w,t) \le q^{-1-\lfloor k/p \rfloor} \overline{L}_{\mathcal{C}'}(w,t). \tag{12}$$

McEliece and Swanson [12] proved that for all w and t,

$$\overline{L}_{\mathcal{C}'}(w,t) \le \frac{1}{(q-1)^{n-k}} \sum_{i=d-w}^{t} \binom{n}{i} (q-1)^i.$$

Writing $q-1 = \theta/(1-\theta)$ for $\theta = 1-1/q \in (0,1)$ gives

$$\begin{aligned} \overline{L}_{\mathcal{C}'}(w,t) &\le \frac{(1-\theta)^{n-k}}{\theta^{n-k}} \sum_{i=d-w}^{t} \binom{n}{i} \frac{\theta^i}{(1-\theta)^i} \\ &= \frac{(1-\theta)^{-k}}{\theta^{n-k}} \sum_{i=d-w}^{t} \binom{n}{i} \theta^i (1-\theta)^{n-i} \\ &= \frac{(1-\theta)^{-k}}{\theta^{n-k}} [F(t;n,\theta) - F(d-w-1;n,\theta)] \\ &\le \frac{(1-\theta)^{-k}}{\theta^{n-k}} F(t;n,\theta), \end{aligned}$$

where F is the cumulative distribution function of the binomial distribution. Hence, by (12),

$$\begin{aligned} q^{1+\lfloor k/p \rfloor} \overline{L}_{\mathcal{C}}(w,t_s) &\le \frac{(1-\theta)^{-k}}{\theta^{n-k}} F(t_s;n,\theta) && (13) \\ &\le \frac{(1-\theta)^{-k}}{\theta^{n-k}} \exp\left(-nD\left(\frac{t_s}{n}\middle\|\theta\right)\right) && (14) \\ &= \frac{n^{k+t_s}}{t_s^{t_s}(n-t_s)^{n-t_s}} \le \frac{2^n}{n^{n-k-t_s}} && (15) \end{aligned}$$

for all $s \ge 1$. Here, (14) is by Chernoff's bound, and (15) is by Jensen's inequality. Therefore,

$$\begin{aligned} \overline{L}_{\mathcal{C}}(w,t_\infty) &\le \frac{2^n}{n^{\lfloor \sqrt{(k-1)n} \rfloor - (k-1)} (n+1)^{1+\lfloor k/p \rfloor}} \\ &= \exp\Big(n \log 2 - \Big(\lfloor \sqrt{(k-1)n} \rfloor - (k-1)\Big) \log n \\ &\qquad - (1+\lfloor k/p \rfloor) \log(n+1)\Big). \end{aligned}$$

Writing $k = Rn$ and $q = p = n+1$, the argument of the exponential is asymptotically

$$n\left(\log 2 - (\sqrt{R} - R)\log n - \frac{(1+R)}{n}\log(n+1)\right),$$

which tends to $-\infty$ as $n \to \infty$ for any fixed rate $R \in (0,1)$. Therefore, by (12), the output of CP-list-decode is $q^{1+\lfloor k/p \rfloor}$ times less likely to contain more than one codeword than that of GS for a GRS code of the same length and degree.

Simulation Results: For each prime $q \in (5,50)$, 5 randomly chosen $f \in \mathcal{F}_p(k,q)'$ for $k = \lfloor q/2 \rfloor - 1$ were encoded in $\mathcal{C}$ and $\mathcal{C}'$, and transmitted 200 times with t_∞ random errors. The average list sizes of the corresponding list decoders are shown in Table I.

TABLE I
MONTE CARLO SIMULATED AVERAGE LIST SIZES

q	$\mathcal{C}$	$\mathcal{C}'$	q	$\mathcal{C}$	$\mathcal{C}'$
7	1	1	29	1.001	1.072
11	1	1	31	1	1.002
13	1	1.585	37	1	1.004
17	1	1.121	41	1	1
19	1.005	1.711	43	1	1
23	1.004	1.115	47	1	1

V. CONCLUSION AND FUTURE DIRECTIONS

In this work, we studied and analyzed efficient minimum-distance and list decoding for one-dimensional CP subspace codes over prime fields. Possible directions for future work include adapting our decoding algorithms in § III to CP codes over arbitrary finite fields and to higher-dimensional CP codes [3]. Note that the problem becomes inherently more challenging in the latter case as higher-dimensional CP codes cannot be directly mapped to block codes. Hence, studying this problem may require developing entirely new techniques.

Note also that we only considered hard-decision decoding as the first step in decoding CP codes. Thus, another direction is to look at soft-decision decoding (see, e.g. [14]) of CP codes, when soft information is available from the mapping of the message coordinates to the finite field elements (i.e. via (10)).

REFERENCES

[1] R. Koetter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3579–3591, 2008.

[2] M. Soleymani and H. Mahdavifar, "Analog subspace coding: A new approach to coding for non-coherent wireless networks," *IEEE Transactions on Information Theory*, vol. 68, no. 4, pp. 2349–2364, 2022. [Online]. Available: https://doi.org/10.1109/TIT.2021.3133071

[3] ——, "New packings in Grassmannian space," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 807–812.

[4] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 359–383, 2002.

[5] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2450–2454, 2002.

[6] V. Guruswami, "Algorithmic results in list decoding," *Foundations and Trends in Theoretical Computer Sciences*, vol. 2, no. 2, pp. 107–195, 2006. [Online]. Available: https://www.cs.cmu.edu/~venkatg/pubs/papers/listdecoding-NOW.pdf

[7] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing Lines, Planes, etc.: Packings in Grassmannian Spaces," *Experimental Mathematics*, vol. 5, no. 2, pp. 139–159, 1996. [Online]. Available: https://doi.org/10.1080/10586458.1996.10504585

[8] R. J. McEliece, "The Guruswami–Sudan Decoding Algorithm for Reed–Solomon Codes," *IPN Progress Report 42-153*, 2003. [Online]. Available: https://tmo.jpl.nasa.gov/progress_report/42-153/153F.pdf

[9] ——, "On the average list size for the Guruswami-Sudan decoder," in *7th International Symposium on Communications Theory and Applications (ISCTA)*, vol. 194, 2003.

[10] S. Gao and M. A. Shokrollahi, "Computing Roots of Polynomials over Function Fields of Curves," in *Coding Theory and Cryptography*, D. Joyner, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 214–228.

[11] R. Roth and G. Ruckenstein, "Efficient decoding of Reed–Solomon codes beyond half the minimum distance," *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 246–257, 2000.

[12] R. J. McEliece and L. Swanson, "On the decoder error probability for Reed–Solomon codes," *IEEE Transactions on Information Theory*, vol. 32, no. 5, pp. 701–703, 1986.

[13] K.-M. Cheung, "More on the decoder error probability for Reed–Solomon codes," *IEEE Transactions on Information Theory*, vol. 35, no. 4, pp. 895–900, 1989.

[14] R. Koetter and A. Vardy, "Algebraic soft-decision decoding of Reed-Solomon codes," *IEEE Transactions on Information Theory*, vol. 49, no. 11, pp. 2809–2825, 2003.