Evading Subspaces Over Large Fields and Explicit List-decodable Rank-metric Codes*

Venkatesan Guruswami and Carol Wang

Computer Science Department, Carnegie Mellon University
Pittsburgh, PA
{venkatg,wangc}@cs.cmu.edu

Abstract

We construct an *explicit* family of *linear* rank-metric codes over any field $\mathbb{F}_h$ that enables efficient list decoding up to a fraction ρ of errors in the rank metric with a rate of $1-\rho-\varepsilon$, for any desired $\rho \in (0,1)$ and $\varepsilon > 0$. Previously, a Monte Carlo construction of such codes was known, but this is in fact the first explicit construction of positive rate rank-metric codes for list decoding beyond the unique decoding radius.

Our codes are explicit subcodes of the well-known Gabidulin codes, which encode linearized polynomials of low degree via their values at a collection of linearly independent points. The subcode is picked by restricting the message polynomials to an $\mathbb{F}_h$-subspace that evades certain structured subspaces over an extension field $\mathbb{F}_{h^t}$. These structured spaces arise from the linear-algebraic list decoder for Gabidulin codes due to Guruswami and Xing (STOC'13). Our construction is obtained by combining subspace designs constructed by Guruswami and Kopparty (FOCS'13) with subspace-evasive varieties due to Dvir and Lovett (STOC'12).

We establish a similar result for subspace codes, which are a collection of subspaces, every pair of which have low-dimensional intersection, and which have received much attention recently in the context of network coding. We also give explicit subcodes of folded Reed-Solomon (RS) codes with small folding order that are list-decodable (in the Hamming metric) with optimal redundancy, motivated by the fact that list decoding RS codes reduces to list decoding such folded RS codes. However, as we only list decode a *subcode* of these codes, the Johnson radius continues to be the best known error fraction for list decoding RS codes.

1998 ACM Subject Classification E.4 Coding and Information Theory

Keywords and phrases list-decoding, pseudorandomness, algebraic coding, explicit constructions

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2014.748

1 Introduction

This paper considers the problem of constructing explicit list-decodable rank-metric codes. A *rank-metric code* is a collection of matrices $M \in \mathbb{F}_h^{n\times t}$ over a finite field $\mathbb{F}_h$ for fixed n, t. The rate of a rank-metric code is $\log_h|\mathcal{C}|/(nt)$, and the distance measure between two codewords is the rank over $\mathbb{F}_h$ of their difference; that is, $\mathrm{dist}(M_1, M_2) = \mathrm{rank}_{\mathbb{F}_h}(M_1 - M_2)$. We will be interested in *linear* rank-metric codes, where $\mathcal{C}$ is a subspace over $\mathbb{F}_h$.

Rank-metric codes have found applications in network coding [23] and public-key cryptography [8, 17], among other areas. They can also be thought of as space-time codes over finite fields, and conversely can be used to construct space-time codes, eg. in [19, 18]. Unique decoding algorithms for rank-metric codes were shown in [5] to be closely related to

* Research supported in part by NSF CCF-0963975.

17th Int'l Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'14) / 18th Int'l Workshop on Randomization and Computation (RANDOM'14).
Editors: Klaus Jansen, José Rolim, Nikhil Devanur, and Cristopher Moore; pp. 748–761

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

the so-called Low-rank Recovery problem, in which the task is to recover a matrix M from few inner products $\langle M, H\rangle$. The authors of [5] use their low-rank recovery techniques to construct rank-metric codes over any field, and show that they can be efficiently decoded.

In this work, we will consider subcodes of *Gabidulin codes*, which are analogues of Reed-Solomon codes for the rank-metric. A Gabidulin code (denoted $\mathcal{C}_G(h; n, t, k)$) encodes h-linearized polynomials over $\mathbb{F}_{h^t}$ of h-degree less than k by $\big(f(\alpha_1), \ldots, f(\alpha_n)\big)^T$, where the $\alpha_i \in \mathbb{F}_{h^t}$ are linearly independent over $\mathbb{F}_h$, and $f(\alpha_j)$ is thought of as a column vector in $\mathbb{F}_h^t$ under a fixed basis of $\mathbb{F}_{h^t}$ over $\mathbb{F}_h$. This is a rank-metric code of rate k/n and minimum distance $n - k + 1$.

We say that a rank-metric code $\mathcal{C}$ can be decoded from up to e *rank errors* if any codeword $M \in \mathcal{C}$ can be recovered from $M + E$ whenever $E \in \mathbb{F}_h^{n\times t}$ has rank at most e. Gabidulin codes can be uniquely decoded from $(n-k)/2$ rank errors by adapting algorithms for Reed-Solomon decoding, as in [6, 7, 22], among others, but it is still open whether they can be *list-decoded* from a larger fraction of errors. We recall that in the list-decoding problem the decoder must output all codewords within the stipulated radius from the noisy codeword it is given as input. It is known that Gabidulin codes *cannot* be list-decoded with a polynomial list size from an error fraction exceeding $1 - \sqrt{R}$ [4, 24]. However, as we show in this work, we can explicitly pick a good subcode of the Gabidulin code, with only a minor loss in rate, that enables efficient list-decoding all the way up to a fraction $(1 - R)$ of errors.

The primary difficulty in previous work on list-decoding Gabidulin codes has been the fact that in contrast to Reed-Solomon codes, where the field size grows with the dimension of the code, for Gabidulin codes, the *dimension* of the ambient space grows with the dimension of the code. This forces us to work over fields whose size can be exponential in the code dimension.

To address this, we show how to find linear list-decodable subcodes of certain Gabidulin codes by adapting the subspace designs of [9] for use over large fields. The key observation, first made in [14], is that although applying a linear-algebraic list-decoder gives a subspace over a field which is too large, the subspace has additional structure which can then be "evaded" using *pseudorandom* subcodes, yielding a polynomial list size.

We combine recent constructions of *subspace designs* [9] and *subspace-evasive sets* [1] in order to give an explicit construction of a subcode (in fact, subspace) of the Gabidulin code which has small intersection with the output of the linear-algebraic list-decoder of [14]. In particular, we show (Theorem 12):

▶ **Theorem** (Main). *For every field $\mathbb{F}_h$, $\varepsilon > 0$ and integer $s > 0$, there exists an explicit $\mathbb{F}_h$-linear subcode of the Gabidulin code $\mathcal{C}_G(h; n, t, k)$ with evaluation points $\alpha_1, \ldots, \alpha_n$ spanning a subfield $\mathbb{F}_{h^n}$ that has (i) rate $(1 - 2\varepsilon)k/n$, and (ii) is list-decodable from $s(n - k)/(s + 1)$ rank errors. The final list is contained in an $\mathbb{F}_h$-subspace of dimension $O(s^2/\varepsilon^2)$.*

Note that the fraction of errors corrected approaches the information-theoretic limit of $(1 - R)$ (where $R = k/n$ is the rate) as the parameter s grows. The authors of [14] give a *Monte Carlo* construction of a subcode of the same Gabidulin code satisfying these guarantees, in fact with a better list size of $O(1/\varepsilon)$. We give an *explicit* subcode, with a worse guarantee on the list size (which, however, is still bounded by a constant depending only on ε).

We also note that the above theorem gives the *first explicit construction* of positive rate rank-metric codes even for list-decoding from a number of errors which is more than half the distance (and in particular for list decoding beyond a fraction $(1-R)/2$ of errors). Previous explicit codes only achieved polynomially small rate [10].

Our techniques also imply analogous results for *subspace codes*, which can be thought of as a basis-independent form of rank-metric codes. They were defined in [16] to address the problem of non-coherent linear network coding in the presence of errors, and have received much attention lately ([2, 20, 3], etc). The authors of [16] also define the Kötter-Kschischang (KK) codes, which, like Gabidulin codes, are linearized variants of Reed-Solomon codes. List-decoding of a folded variant of the KK code was considered in [10] and [21]. However, both of these papers could only guarantee a polynomial list size when the rate of the code was polynomially small, and the question of constructing constant rate list-decodable subspace codes remained open. Note that [14] was able, similarly to the case of rank-metric codes, to give a Monte Carlo construction of a constant rate list-decodable subcode.

In this work, we give the first explicit construction of high-rate subspace codes which are list-decodable past the unique decoding radius (stated in Theorem 20). Our construction does not use folding, but instead takes subcodes of certain KK codes.

Additionally, we use our ideas to list-decode a subcode of the folded Reed-Solomon code where the folding parameter is of low order (see Corollary 16 for a formal statement). List-decoding of the folded Reed-Solomon code up to list-decoding capacity where the folding parameter is primitive was first shown in [11]. In [12], the authors use the linear-algebraic method to list-decode folded Reed-Solomon codes when the folding parameter has order at least the dimension of the code.

Paper Organization. In Section 2, we collect notation and definitions which will be used throughout the paper. In Section 3, we define and construct "(s, A, t)-subspace designs," which is the new twist on the subspace designs of [9] that drives our results. In Section 4, we show how these subspace designs can be used to construct list-decodable rank-metric codes. In Section 5, we give a list-decodable subcode of folded Reed-Solomon codes with low folding order. The construction of list-decodable subspace codes appears as Appendix A.

We conclude in Section 6 with some open problems.

2 Notation and Definitions

Throughout the presentation of rank-metric codes, $\mathbb{F}_h$ is a finite field of constant size. $\mathbb{F}_q := \mathbb{F}_{h^t}$ extends $\mathbb{F}_h$, and we will think of $\mathbb{F}_q$ as a vector space over $\mathbb{F}_h$ by fixing a basis. We will also have $n = mt$, and the field $\mathbb{F}_{h^n} := \mathbb{F}_{q^m} = \mathbb{F}_{h^{mt}}$ extending $\mathbb{F}_q$.

In our final applications, s will be $\approx 1/\varepsilon$, m will be $\approx s/\varepsilon$, where for rate R, we will be list decoding up to error fraction $(1 - R - \varepsilon)$, and t will grow.

We will be talking about subspaces over a field and its extension, so to avoid any confusion about the underlying field, we will usually refer to a subspace over a field $\mathbb{F}$ as an $\mathbb{F}$-subspace.

We recall some of the definitions of the pseudorandom objects concerning subspaces that we require.

▶ **Definition 1** (Strong subspace designs, [14])**.** A collection S of $\mathbb{F}_q$-subspaces $H_1, \ldots, H_M \subseteq \mathbb{F}_q^m$ is called a (s, A) *subspace design* if for every $\mathbb{F}_q$-linear space $W \subset \mathbb{F}_q^m$ of dimension s,

$$\sum_{i=1}^{M} \dim_{\mathbb{F}_q}(H_i \cap W) \leqslant A.$$

▶ **Definition 2** (Subspace-evasive sets, [12])**.** A subset $\mathcal{V} \subseteq \mathbb{F}_q^k$ is (s, L) *subspace-evasive* if for every $\mathbb{F}_q$-subspace $S \subset \mathbb{F}_q^k$ of dimension s, $|S \cap \mathcal{V}| \leqslant L$.

3 Subspace Designs

Throughout this section q and h will be prime powers with $q = h^t$. In what follows, we will think of subspaces $W \subseteq \mathbb{F}_q^m$ as $\mathbb{F}_h$-subspaces of $\mathbb{F}_h^{mt}$ via some fixed basis embedding.

▶ **Definition 3.** A collection S of $\mathbb{F}_h$-subspaces $H_1, \ldots, H_M \subseteq \mathbb{F}_h^{tm}$ is called a (s, A, t) $\mathbb{F}_h$-*subspace design* if for every $\mathbb{F}_{h^t}$-linear space $W \subset \mathbb{F}_{h^t}^m$ of dimension s,

$$\sum_{i=1}^{M} \dim_{\mathbb{F}_h}(H_i \cap W) \leqslant A.$$

Note that in the above definition the dimension of the input W is measured as a subspace over $\mathbb{F}_{h^t}$ whereas for the intersection, which is an $\mathbb{F}_h$-subspace, the dimension is over $\mathbb{F}_h$.

▶ Remark. When $t = 1$, these are the (strong) subspace designs of [9]. We will be interested in settings where $t = \omega(1)$, so that considering W as a subspace of dimension st over $\mathbb{F}_h$ will generally not give strong enough bounds.

3.1 Existential Bounds

The following proposition shows that good subspace designs exist; indeed, a random collection of subspaces works with high probability. The case $t = 1$ was established in [9].

▶ **Proposition 4.** *Let $\varepsilon > 0$. Let S consist of $M = h^{\varepsilon tm/8}$ $\mathbb{F}_h$-subspaces of codimension εtm in $\mathbb{F}_h^{mt}$, chosen independently at random. Then for any $s < m\varepsilon/2$, with probability at least $1 - q^{-ms}$, S is a $(s, 8s/\varepsilon, t)$ $\mathbb{F}_h$-subspace design. (Here $q = h^t$.)*

Proof. Set $\ell = 8s/\varepsilon$, and let $S = \{H_1, \ldots, H_M\}$. For a fixed $\mathbb{F}_{h^t}$ subspace W of dimension s and any j, the probability that $\dim_{\mathbb{F}_h}(W \cap H_j) \geqslant a$ at most $q^{sa} \cdot q^{-\varepsilon ma} \leqslant q^{-\varepsilon ma/2}$, by assumption on s.

Since the H_i are independent, for a fixed tuple $(a_1, \ldots, a_M)$ of nonnegative integers summing to $\ell = 8s/\varepsilon$, the probability that $\dim(W \cap H_j) \geqslant a_j$ for each j is at most $q^{-\varepsilon m\ell/2} = q^{-4ms}$. Union bounding over the at most q^{ms} choices of W and $\binom{\ell+M}{\ell} \leqslant M^{2\ell}$ choices of $(a_1, \ldots, a_M)$, the probability S is *not* a $(s, 8s/\varepsilon, t)$ $\mathbb{F}_h$-subspace design is at most

$$q^{ms} M^{2\ell} \cdot q^{-4ms} = q^{ms} \cdot q^{2ms} \cdot q^{-4ms} \leqslant q^{-ms} \ . \qquad \blacktriangleleft$$

3.2 Constructive Bounds

In this section, we show how to construct an explicit large $\big(s, 2(m-1)s/\varepsilon, t\big)$ $\mathbb{F}_h$-subspace design consisting of $\mathbb{F}_h$-subspaces of $\mathbb{F}_h^{tm}$ of codimension $2\varepsilon tm$.

The idea, which is natural in hindsight, is to first use a subspace design over $\mathbb{F}_{h^t}$ to ensure that the intersection with any $\mathbb{F}_{h^t}$-subspace of dimension s has low dimension over $\mathbb{F}_{h^t}$, and then to use a subspace-evasive set to reduce the dimension further over $\mathbb{F}_h$. The final construction appears as Theorem 8.

3.2.1 Explicit Subspace-evasive Sets

We first describe the construction of explicit subspace-evasive sets which we will be using.

Let $q > h^{m-1}$, and let $\gamma_1, \ldots, \gamma_m$ be distinct elements of $(\mathbb{F}_q)^*$. Let A be the $s \times m$ matrix with $A_{ij} = \gamma_j^i$. Then Dvir and Lovett [1] showed the following:

▶ **Theorem 5.** *Let $1 \leqslant s \leqslant m$. Let $d_1 > d_2 > \cdots > d_m \geqslant 1$ be integers. Define $f_1, \ldots, f_s \in \mathbb{F}_q[X_1, \ldots, X_m]$ as follows:*

$$f_i(x_1, \ldots, x_m) = \sum_{j=1}^{m} A_{ij} x_j^{d_j} \ . \tag{1}$$

Then:

- *The variety $\mathbf{V} = \{x \in \overline{F}_q^m \mid f_1(x) = \cdots = f_s(x) = 0\}$ satisfies $|\mathbf{V} \cap H| \leqslant (d_1)^s$ for all s-dimensional affine subspaces $H \subset \overline{\mathbb{F}}_q^m$.*
- *If at least s of the degrees d_i are relatively prime to $q-1$, then $|\mathbf{V} \cap \mathbb{F}_q^m| = q^{m-s}$.*

Additionally, the product set $(\mathbf{V} \cap \mathbb{F}_q^m)^{n/m} \subseteq \mathbb{F}^n$ is $(k, (d_1)^k)$-subspace evasive for all $k \leqslant s$.

The below statement follows immediately from Theorem 5 and the fact that when the d_j's are powers of h, the polynomials f_i defined in (1) are $\mathbb{F}_h$-linear functions on $\mathbb{F}_q^m$.

▶ **Corollary 6.** *Setting $d_1 = h^{m-1}, d_2 = h^{m-2}, \ldots, d_m = 1$, we obtain an explicit $\mathbb{F}_h$-linear set S of size $q^{(m-s)n/m}$ over $\mathbb{F}_q^n$ which is $(k, h^{(m-1)k})$ subspace-evasive for all $1 \leqslant k \leqslant s$.*

▶ Remark. One can improve on the degree bounds and therefore the final intersection size via a standard subspace-evasive set without the $\mathbb{F}_h$-linearity requirement. For example, [1] gives a construction of a (non-linear) $\left(s, (s/\varepsilon)^s\right)$ subspace-evasive set over $\mathbb{F}^n$ of size $|\mathbb{F}|^{(1-\varepsilon)n}$.

However, especially in applications for rank-metric codes, linearity is a property which is desirable and often necessary.

3.2.2 Combining with Subspace Designs

The following theorem shows how to achieve our initial goal of ensuring small intersection dimension over the larger field $\mathbb{F}_{h^t}$.

▶ **Theorem 7** ([9])**.** *For $\varepsilon \in (0,1)$, positive integers s, m with $s \leqslant \varepsilon m/4$, and $q > m$, there is an explicit collection of $M = q^{\Omega(\varepsilon m/s)}$ subspaces in $\mathbb{F}_q^m$, each of codimension at most εm, which form a $(s, 2s/\varepsilon, 1)$ $\mathbb{F}_q$-subspace design.*

Combined with Corollary 6, we now have a construction of a $(s, 2(m-1)s/\varepsilon, t)$ $\mathbb{F}_h$-subspace design, summarized in the following statement.

▶ **Theorem 8.** *For integers $s \leqslant \varepsilon m/4$ and $q > m$, there exists an explicit set of $q^{\Omega(\varepsilon m/s)}$ $\mathbb{F}_h$-subspaces in $\mathbb{F}_h^{tm}$ of codimension at most $2\varepsilon tm$ forming a $(s, 2(m-1)s/\varepsilon, t)$ $\mathbb{F}_h$-subspace design.*

Proof. Let $V_1, \ldots, V_M \subseteq \mathbb{F}_q^m$ be the elements of the $(s, 2s/\varepsilon, 1)$ $\mathbb{F}_q$-subspace design of Theorem 7. For each i, define $H_i = V_i \cap S$, where $S \subseteq \mathbb{F}_q^m$ is the $(s, h^{(m-1)s})$ subspace-evasive set of Corollary 6. As S and the V_i's are $\mathbb{F}_h$-linear subspaces, H_i is as well. We claim that the H_i's form the desired $\mathbb{F}_h$-subspace design.

For each i, V_i has codimension εtm, and S has codimension $ts \leqslant \varepsilon tm/4$, so the codimension of H_i is at most $2\varepsilon tm$.

Now let W be an $\mathbb{F}_q$-subspace of dimension s. By the $\mathbb{F}_q$-subspace design property of the V_i's we have

$$\sum_{i=1}^{M} \dim_{\mathbb{F}_q}(V_i \cap W) \leqslant 2s/\varepsilon \ . \tag{2}$$

For each i, we also have that $\dim_{\mathbb{F}_q}(W \cap V_i) = s_i \leqslant s$, so by the subsace evasive property of S from Corollary 6, $W \cap H_i = (W \cap V_i) \cap S$ has at most $h^{(m-1)s_i}$ elements. As $W \cap H_i$ is $\mathbb{F}_h$-linear, we have

$$\dim_{\mathbb{F}_h}(W \cap H_i) \leqslant (m-1)\dim_{\mathbb{F}_q}(W \cap V_i) \ . \tag{3}$$

Combining (2) and (3) we have

$$\sum_i \dim_{\mathbb{F}_h}(W \cap H_i) \leqslant \sum_i (m-1)\dim_{\mathbb{F}_q}(W \cap V_i) \leqslant (m-1)\cdot 2s/\varepsilon \ . \qquad \blacktriangleleft$$

The motivation for constructing the above subspace design is that they yield a subspace that has small intersection with so-called periodic subspaces arising in certain linear-algebraic list decoding algorithms. We recall the definition from [14]. Below, for a string $\mathbf{x} = (x_1, x_2, \ldots, x_\ell)$, we denote by $\mathrm{proj}_{[a,b]}(\mathbf{x})$ the substring $(x_a, x_{a+1}, \ldots, x_b)$.

▶ **Definition 9** (Periodic subspaces)**.** For positive integers s, m, k and $\kappa := mk$, an affine subspace $H \subset \mathbb{F}_q^\kappa$ is said to be (s,m,k)-**periodic** if there exists a subspace $W \subseteq \mathbb{F}_q^m$ of dimension at most s such that for every $j = 1, 2, \ldots, k$, and every prefix $\mathbf{a} \in \mathbb{F}_q^{(j-1)m}$, the projected affine subspace of $\mathbb{F}_q^m$ defined by

$$\{\mathrm{proj}_{[(j-1)m+1,jm]}(\mathbf{x}) \mid \mathbf{x} \in H \text{ and } \mathrm{proj}_{[1,(j-1)m]}(\mathbf{x}) = \mathbf{a}\}$$

is contained in an affine subspace of $\mathbb{F}_q^m$ given by $W + \mathbf{v_a}$ for some vector $\mathbf{v_a} \in \mathbb{F}^m$ dependent on $\mathbf{a}$.

▶ **Proposition 10.** *Let H be a (s,m,k)-periodic affine suspace of $\mathbb{F}_q^{mk}$, and $H_1, H_2, \ldots, H_k \subseteq \mathbb{F}_h^{mt}$ be distinct subspaces from a (s, A, t) $\mathbb{F}_h$-subspace design. Then $H \cap (H_1 \times \cdots \times H_k)$ is an affine subspace over $\mathbb{F}_h$ of dimension at most A.*

Proof. It is clear that $H \cap (H_1 \times \cdots \times H_k)$ is an affine subspace over $\mathbb{F}_h$. Let W be the subspace associated to H as in Definition 9. We will show by induction that $|\mathrm{proj}_{[1,im]}(H) \cap (H_1 \times \cdots \times H_i)| \leqslant h^{\sum_{j=1}^{i} \dim_{\mathbb{F}_h}(W \cap H_j)}$.

In the base case, since H_1 is a subspace, $\mathrm{proj}_{[1,m]}(H) \cap H_1 = (W + v_{\mathbf{0}}) \cap H_1$ is an affine subspace whose underlying subspace lies in $W \cap H_1$. In particular, its size is at most $h^{\dim(W \cap H_1)}$.

Continuing, fix an element $\mathbf{a} \in \mathrm{proj}_{[1,im]}(H) \cap (H_1 \times \cdots \times H_i)$. Because H is periodic and H_{i+1} is linear, the possible extensions of $\mathbf{a}$ in $\mathrm{proj}_{[im+1,(i+1)m]}(H) \cap H_{i+1}$ are given by a coset of $W \cap H_{i+1}$. Thus, there are at most $h^{\dim(W \cap H_{i+1})}$ such extensions. Since by induction there were $h^{\sum_{j=1}^{i} \dim_{\mathbb{F}_h}(W \cap H_j)}$ possibilities for the prefix $\mathbf{a}$, the result follows.

In particular, $H \cap (H_1 \times \cdots \times H_k)$ has dimension over $\mathbb{F}_h$ which is at most $\sum_{i=1}^{k} \dim(W \cap H_i) \leqslant A$, by the subspace design property. $\blacktriangleleft$

4 Explicit List-decodable Rank-metric Codes

In this section, we show how to use the subspace designs of Theorem 8 in order to get explicit list-decodable rank-metric codes of optimal rate for any desired error correction radius.

We first review rank-metric codes, and in particular the Gabidulin code [6], which is the starting point of our construction.

Let h be a prime power, and let $\mathbb{M}_{n\times t}(\mathbb{F}_h)$ be the set of $n \times t$ matrices over $\mathbb{F}_h$. The *rank distance* between $A, B \in \mathbb{M}_{n\times t}(\mathbb{F}_h)$ is $d(A, B) = \operatorname{rank}(A - B)$. A rank-metric code $\mathcal{C}$ is a subset of $\mathbb{M}_{n\times t}(\mathbb{F}_h)$, with rate and distance given by

$$R(\mathcal{C}) = \frac{\log_h|\mathcal{C}|}{nt} \quad \text{and} \quad d(\mathcal{C}) = \min_{A\neq B\in\mathcal{C}}\{d(A, B)\}.$$

The *Gabidulin code* encodes h-linearized polynomials by their evaluations at linearly independent points. Recall that an h-linearized polynomial f over $\mathbb{F}_{h^t}$ is a polynomial of the form $\sum_{i=0}^{\ell} a_i X^{h^i}$, with $a_i \in \mathbb{F}_{h^t}$. If $a_\ell \neq 0$, then ℓ is called the *h-degree* of f. We write $\mathcal{L}_h(t)$ for the set of h-linearized polynomials over $\mathbb{F}_{h^t}$.

Let $0 < k \leqslant n \leqslant t$ be integers, and choose $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_{h^t}$ to be linearly independent over $\mathbb{F}_h$. For every h-linearized polynomial $f \in \mathbb{F}_{h^t}[X]$ of h-degree at most $k - 1$, we can encode f by the column vector $M_f = \big(f(\alpha_1), \ldots, f(\alpha_n)\big)^T$ over $\mathbb{F}_{h^t}$. By fixing a basis of $\mathbb{F}_{h^t}$ over $\mathbb{F}_h$, we can also think of M_f as an $n \times t$ matrix over $\mathbb{F}_h$. This yields the Gabidulin code

$$\mathcal{C}_G(h; n, t, k) := \{M_f \in \mathbb{M}_{n\times t}(\mathbb{F}_h) \mid f \in \mathcal{L}_h(t),\ h\text{-degree}(f) \leqslant k - 1\}.$$

If a rank-metric codeword X is transmitted, and a matrix Y is received, we say that $\operatorname{rank}(Y - X)$ *rank errors* have occurred.

Suppose that $t = nm$ for some integer m, so that $\mathbb{F}_{h^t}$ has a subfield $\mathbb{F}_{h^n} =: \mathbb{F}_q$. In the case when the evaluation points $\alpha_1, \ldots, \alpha_n$ of the Gabidulin code span $\mathbb{F}_{h^n}$, Guruswami and Xing [14] show the following:

▶ **Theorem 11** ([14]). *Let $f \in \mathbb{F}_{h^t}[X]$ be an h-linearized polynomial with h-degree at most $k-1$. Suppose that a codeword $M_f = \big(f(\alpha_1), \ldots, f(\alpha_n)\big)^T$ is transmitted and $Y = (y_1, \ldots, y_n)^T$ is received with at most e rank errors. If $e \leqslant s(n - k)/(s + 1)$, then there is an algorithm running in time $\operatorname{poly}(n, m, \log q)$ outputting a $(s - 1, m, k)$-periodic subspace containing all candidate messages f.*

By Proposition 10, by restricting the message polynomials $f = \sum_i f_i X^{q^i}$ to have coefficients $f_i \in H_{i+1}$ for $0 \leqslant i < k$, where $H_1, H_2, \ldots, H_k$ are distinct elements of the subspace design in Theorem 8, the final list of candidate messages will have dimension at most $2(m - 1)s/\varepsilon$ over $\mathbb{F}_h$, or size at most $h^{2(m-1)s/\varepsilon}$. As one can take $m = O(s/\varepsilon)$ for the necessary subspace design guaranteed by Theorem 8, we can conclude the following theorem, which is our main result.

▶ **Theorem 12.** *For every $\varepsilon > 0$ and integer $s > 0$, there exists an explicit $\mathbb{F}_h$-linear subcode of the Gabidulin code $\mathcal{C}_G(h; n, t, k)$ with evaluation points spanning $\mathbb{F}_{h^n}$ of rate $(1 - 2\varepsilon)k/n$ which is list-decodable from $\frac{s}{s+1} \cdot (n - k)$ rank errors. The final list is contained in an $\mathbb{F}_h$-subspace of dimension at most $O(s^2/\varepsilon^2)$.*

5 Application to Low-order Folding of Reed-Solomon Codes

In this section, we show how the idea of only evading subspaces over an extension field can be used to give an algorithm for list-decoding (subcodes of) folded Reed-Solomon codes in the case when the folding parameter has low ($O(1)$) order.

As in the case of KK codes, our decoding algorithm follows the framework of interpolating a linear polynomial and then solving a linear system for candidate polynomials. Fix γ generating $\mathbb{F}_q^*$. Let $N = \frac{q-1}{\ell}$, and let $\zeta = \gamma^N$, which has order ℓ in $\mathbb{F}_q$. Then the **low-order**

folded Reed-Solomon code encodes a polynomial f of degree $< k$ by

$$f \mapsto \begin{bmatrix} f(1) & f(\gamma) & \cdots & f(\gamma^{N-1}) \\ f(\zeta) & f(\zeta\gamma) & \cdots & f(\zeta\gamma^{N-1}) \\ \vdots & \vdots & \ddots & \vdots \\ f(\zeta^{\ell-1}) & f(\zeta^{\ell-1}\gamma) & \cdots & f(\zeta^{\ell-1}\gamma^{N-1}) \end{bmatrix}.$$

Similarly to folded Reed-Solomon codes, this is a code of rate $\frac{k}{\ell N}$ and distance $N - (k-1)/\ell$.

5.1 Interpolation

Given a received word

$$\begin{pmatrix} y_{00} & y_{01} & \cdots & y_{0(N-1)} \\ y_{10} & y_{11} & \cdots & y_{1(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ y_{(\ell-1)0} & y_{(\ell-1)1} & \cdots & y_{(\ell-1)(N-1)} \end{pmatrix},$$

we would like to interpolate a (nonzero) polynomial

$$Q(X, Y_1, \ldots, Y_s) = A_0(X) + A_1(X)Y_1 + \cdots + A_s(X)Y_s$$

such that

$$Q\big(\gamma^{iN+j}, y_{ij}, y_{(i+1)j}, \ldots, y_{(i+s-1)j}\big) = 0 \qquad i \in \{0, \ldots, \ell-1\},\ j \in \{0, \ldots, N-1\}, \tag{4}$$

where all indices are taken modulo ℓ.

We will require $\deg(A_0) \leqslant D + k - 1$, and $\deg(A_i) \leqslant D$ for $i > 0$.

▶ **Lemma 13.** *Let*

$$D = \left\lfloor \frac{\ell N - k + 1}{s+1} \right\rfloor.$$

Then a nonzero polynomial Q satisfying (4) *exists (and can be found by solving a linear system).*

Proof. The number of interpolation conditions is ℓN. The quantity $(D+1)(s+1) + k - 1 > \ell N$ is the number of degrees of freedom for the interpolation, and the conditions are homogeneous, so a nonzero solution exists. ◀

▶ **Lemma 14.** *If the number of agreements t is greater than $\frac{D+k-1}{\ell}$, then*

$$Q\big(X, f(X), f(\zeta X), \ldots, f(\zeta^{s-1}X)\big) = 0. \tag{5}$$

Proof. $Q\big(X, f(x), \ldots, f(\zeta^{s-1}X)\big)$ is a univariate polynomial of degree $D + k - 1$, and each correct column j yields ℓ distinct roots γ^{iN+j} for $i \in \{0, \ldots, \ell-1\}$. Thus if $t\ell > \deg D + k - 1 \geqslant \deg Q$, Q is the zero polynomial. ◀

For our choice of D, the requirement on t in Lemma 14 is met if t satisfies

$$\frac{t}{N} \geqslant \frac{1}{s+1} + \frac{s}{s+1}R. \tag{6}$$

▶ Remark. In ordinary folded Reed-Solomon codes, where the folding parameter is primitive of order $q-1$, the agreement fraction required to satisfy (5) is

$$\frac{t}{N} \geqslant \frac{1}{s+1} + \frac{s}{s+1}\frac{\ell R}{\ell - s + 1},$$

which is higher than (6). In our case, because ζ has low order, we are able to use interpolation conditions that "wrap around," allowing us to impose ℓ conditions per coordinate rather than $\ell - s + 1$. Therefore we can satisfy Equation (5) with lower agreement. On the other hand, it is known how to list-decode folded Reed-Solomon codes themselves, whereas we are only able to list-decode a subcode.

5.2 Decoding

In this section, we describe how to solve the system

$$Q\big(X, f(X), f(\zeta X), \ldots, f(\zeta^{s-1}X)\big) = 0 \tag{5}$$

for candidate polynomials f.

▶ **Proposition 15.** *Given an irreducible polynomial $R(X) \in \mathbb{F}_q[X]$ such that*

- $\deg R \geqslant k$, *and*
- *for some a, $\zeta X \equiv X^{q^a} \pmod R$.*

Then the set of f of degree $< k$ satisfying (5) *is an $\mathbb{F}_{q^a}$-affine subspace of dimension at most $s-1$.*

Proof. The condition (5) says

$$0 = A_0(X) + A_1(X)f(X) + A_2(X)f(\zeta X) + \cdots + A_s(X)f(\zeta^{s-1}X).$$

Then we have

$$A_0(X) + A_1(X)f(X) + A_2(X)f(X)^{q^a} + \cdots + A_s(X)f(X)^{q^{(s-1)a}} \equiv 0 \pmod R.$$

By dividing out the highest power of R which divides every A_i, Equation (5) is still satisfied and we may assume that this equation is nonzero mod R.

In particular, this equation has at most $q^{(s-1)a}$ solutions for f mod R. When $\deg f < k \leqslant \deg R$, f is uniquely determined by its residue mod R and there are at most $q^{(s-1)a}$ solutions for f.

The fact that the solution space is $\mathbb{F}_{q^a}$-affine follows from the fact that the terms in which $f(X)$ appears all have degree q^{ai} for some i. ◀

Because the output space is a subspace (over the large field $\mathbb{F}_{q^a}$), by picking the message polynomials f to come from a subspace-evasive set, we can reduce the list size bound. More specifically, if ℓ is at least s/ε, [1] gives a construction of a $(s, (s/\varepsilon)^s)$ subspace-evasive set S over $(\mathbb{F}_{q^a})^{k/a}$ of size $q^{(1-\varepsilon)k}$. By precoding the messages to come from this set S, we are able to both encode and compute the intersection of the code with the output subspace of Proposition 15 in polynomial time.

Setting $s = O(1/\varepsilon)$ and $\ell = O(s/\varepsilon)$, we obtain the following.

▶ **Corollary 16.** *For every $\varepsilon > 0$ and $R \in (0,1)$, there is an explicit rate R subcode of a low-order folded Reed-Solomon code which is list-decodable from a $1 - R - \varepsilon$ fraction of errors with list size $(1/\varepsilon)^{O(1/\varepsilon)}$, given an irreducible polynomial satisfying the conditions of Proposition 15.*

▶ Remark. By using Corollary 6 instead of the results of [1], we can give a similar guarantee which yields a *linear* subcode, but with a larger list size guarantee of $q^{\mathrm{poly}(1/\varepsilon)}$.

The techniques of [14] using subspace designs could also be applied directly to the case of low-order folding, with a resulting list size of $n^{\mathrm{poly}(1/\varepsilon)}$. We are able to get an improvement using the observation that the space of candidates is actually a low-dimensional subspace over a much larger field.

5.3 Constructing High-degree Irreducibles

The decoding algorithm of the previous section relied on working modulo a high-degree irreducible factor of $X^{q^a} - \zeta X$. In what follows, we consider the problem of finding such a factor efficiently.

▶ **Proposition 17.** *For $\zeta \in \mathbb{F}_q$ of order ℓ, the irreducible factors over $\mathbb{F}_q[X]$ of*

$$X^{q^a-1} - \zeta$$

have degree dividing $a\ell$. In particular, all roots of $X^{q^a-1} - \zeta$ lie in $\mathbb{F}_{q^{a\ell}}$.

Proof. As $X^{(q^a-1)\ell} \equiv 1 \pmod{X^{q^a-1} - \zeta}$, it is enough to see that $(q^a - 1)\ell$ divides $q^{a\ell} - 1$. This implies that $X^{q^a-1} - \zeta$, and thus all of its irreducible factors, divides $X^{q^{a\ell}} - X$.

As $\ell \mid q - 1$, we have

$$\frac{q^{a\ell} - 1}{q^a - 1} = q^{a(\ell-1)} + q^{a(\ell-2)} + \cdots + q^a + 1 \equiv 0 \pmod{\ell} . \qquad \blacktriangleleft$$

▶ **Corollary 18.** *If a and ℓ with $a > 2\ell$ are distinct primes, at least half of the roots of $X^{q^a-1} - \zeta$ have irreducible polynomials of degree $a\ell$.*

Proof. By Proposition 17, all of the irreducible factors of $X^{q^a-1} - \zeta$ have degrees in the set $\{1, a, \ell, a\ell\}$. No irreducible factor has degree 1 or a, because any irreducible of degree 1 or a divides $X^{q^a-1} - 1$ and therefore does not divide $X^{q^a-1} - \zeta$ for $\zeta \neq 1$.

Because $X^{q^a-1} - \zeta$ has no repeated factors, it has at most q^ℓ roots which lie in $\mathbb{F}_{q^\ell}$ (and hence have irreducible polynomials of degree ℓ.

Thus, under the assumptions on a and ℓ, $X^{q^a-1} - \zeta$ has at least $(q^a - q^\ell - 1) \geqslant q^\ell$ roots of degree $a\ell$. Thus at least half of of $X^{q^a-1} - \zeta$'s roots have irreducible polynomials of degree $a\ell$. $\blacktriangleleft$

In particular, by choosing a to be a prime in the range $[k/\ell, 2k/\ell]$, we have $k \leqslant a\ell \leqslant 2k$, so that an irreducible factor of $X^{q^a-1} - \zeta$ will satisfy the conditions of Proposition 15. The next section will show that we cannot hope to improve much on the value of a.

Given a value for a for which $X^{q^a-1} - \zeta$ has many degree $a\ell$ factors, the problem remains to compute one. In what follows, we describe one randomized approach.

Recall that a and ℓ are primes, and that we are trying to find a degree $a\ell$ factor of $X^{q^a-1} - \zeta$. The idea is to sample a root of $X^{(q^a-1)\ell} - 1$. Consider the following procedure:

1. Sample $\beta \in (\mathbb{F}_{q^a})^*$ uniformly at random.
2. Compute the roots $\rho_1, \ldots, \rho_\ell$ of $X^\ell - \beta$, which lie in $\mathbb{F}_{q^{a\ell}}$ by Proposition 17. This can be done in time $\tilde{O}(n^2 \log(q^a) \log^{-1} \varepsilon)$ with failure probability ε using a variant of Berlekamp's algorithm (see, for example, [15]).
3. Compute $\rho_i^{q^a-1}$ for each i and output the minimal polynomial of ρ_i over $\mathbb{F}_q$ if $\rho_i^{q^a-1} = \zeta$.

First note that steps 1–2 sample each root of $X^{(q^a-1)\ell}-1$ uniformly. Each ρ_i computed in step 2 satisfies $\rho_i^\ell \in (\mathbb{F}_{q^a})^*$, so ρ_i is a root of $X^{(q^a-1)\ell}-1$. Conversely, each nonzero β yields ℓ distinct roots of $X^\ell - \beta$, which are distinct for distinct β, yielding $(q^a-1)\ell$ roots.

Therefore, with probability $1/\ell$, we will find a root ρ of $X^{q^a-1}-\zeta$. By Corollary 18, ρ's minimal polynomial has degree $a\ell$ with probability at least $1/2$.

We can thus conclude that, with probability at least $\frac{1}{2\ell}-\varepsilon$, we find an irreducible factor of $X^{q^a-1}-\zeta$ of degree $a\ell$.

5.4 Relationship to Reed-Solomon List-decoding

The original motivation for studying low-order folding was the following reduction from Reed-Solomon codes.

Given a polynomial f of degree $< k/\ell$ evaluated at distinct points $1, \gamma^\ell, \gamma^{2\ell}, \ldots, \gamma^{N\ell}$, we can think of it as a degree $< k$ polynomial $g(X) = f(X^\ell)$. For ζ of order ℓ, we have that $g(\zeta^i X) = g(X)$ for every i. In particular, the associated low-order folded Reed-Solomon codeword encoding $g(X)$ is simply

$$\begin{bmatrix} f(1) & f(\gamma^\ell) & \ldots & f(\gamma^{N\ell}) \\ f(1) & f(\gamma^\ell) & \ldots & f(\gamma^{N\ell}) \\ \vdots & \vdots & \ddots & \vdots \\ f(1) & f(\gamma^\ell) & \ldots & f(\gamma^{N\ell}) \end{bmatrix}. \tag{7}$$

Notice that if $f(\gamma^{i\ell})$ is correct, then the entire ith column is correct, so an algorithm to list-decode the low-order folded RS code from an η fraction of errors will also list-decode the Reed-Solomon code with evaluation points $(1, \gamma^\ell, \ldots, \gamma^{N\ell})$ from the same error fraction.

This reduction also helps to show that the precoding used to conclude Corollary 16 is necessary for a polynomial list size. To see this, consider the behavior of the algorithm on a transmitted codeword as in Equation (7). If there is enough agreement, the algorithm will interpolate polynomials $A_i(X)$ satisfying

$$0 = A_0 + A_1(X)g(X) + A_2(X)g(\zeta X) + \cdots + A_s(X)g(\zeta^{s-1}X) \tag{8}$$

$$= A_0(X) + g(X)\sum_{i=1}^{s} A_i(X). \tag{9}$$

If $\sum_{i>0} A_i(X) \neq 0$, then $g(X)$, and thus $f(X)$, can be recovered *uniquely* by computing $A_0(X)/\sum_{i>0} A_i(X)$; however, this will not be possible in general outside of the unique decoding radius. If $\sum_{i>0} A_i(X)$ is 0, then $A_0(X) = 0$ as well and *any* function which is a polynomial of X^ℓ satisfies Equation (9), and in particular the output list must have size at least $q^{k/\ell}$. Recall that ℓ is a constant in our application.

This implies that without precoding, the dimension of the list output by Proposition 15 over $\mathbb{F}_q$ must be $\Omega(k/\ell)$. Note that for the value $a = \theta(k/\ell)$ found in Section 5.3, the list size before precoding would be $O(ks/\ell)$.

6 Conclusion and Open Questions

We have given an explicit construction of list-decodable rank-metric and subspace codes, which were obtained by restricting known codes to carefully chosen subcodes. However, our results give no insight into whether the Gabidulin and KK codes can be themselves list-decoded beyond half the minimum distance. We close with the following natural open problems.

- Is it combinatorially feasible to list-decode Gabidulin codes *themselves* beyond half the distance? We note that it was recently shown that there is no analog of the classical Hamming-metric Johnson bound in the world of rank-metric codes always guaranteeing list-decodability beyond half the minimum distance [24]. Therefore, a proof of list-decodability past the unique decoding radius (say for the Gabidulin code) must account for the code structure beyond just the minimum distance.
- Assuming it is combinatorially feasible, can we give an efficient algorithm to list-decode Gabidulin codes without using subcodes or special evaluation points?
- Currently, for rate R codes, we do not know where in the range $(1-\sqrt{R}, 1-R)$ the list-decoding radius of Reed-Solomon codes lies, and where in the range $[(1-R)/2, 1-\sqrt{R}]$ the list-decoding radius of Gabidulin codes lies. Is there a relationship between these questions?
- Can one construct better subspace-evasive sets to give an *explicit* code that is list-decodable from a fraction $1-R-\varepsilon$ of errors with $\mathrm{poly}(1/\varepsilon)$ list-size? We only known a list-size upper bound that is exponential in $1/\varepsilon$ for current explicit constructions, whereas a list-size of $O(1/\varepsilon)$ can be obtained with a Monte Carlo construction [12, 13, 14]. This question is open for errors in the usual Hamming metric also.

Acknowledgment. We thank Antonia Wachter-Zeh for bringing to our attention the lack of a Johnson-type bound for list decoding rank-metric codes [24].

References

1 Z. Dvir and S. Lovett. Subspace evasive sets. In *Proceedings of the 44th ACM Symposium on Theory of Computing*, pages 351–358, 2012.

2 T. Etzion and N. Silberstein. Error-correcting codes in projective spaces via rank-metric codes and ferrers diagrams. *IEEE Transactions on Information Theory*, 55:2909–2919, 2009.

3 T. Etzion and A. Vardy. Error-correcting codes in projective space. *IEEE Transactions on Information Theory*, 57:1165–1173, 2011.

4 C. Faure. Average number of Gabidulin codewords within a sphere. In *Int. Workshop on Alg. Combin. Coding Theory (ACCT)*, pages 86–89, 2006.

5 M. A. Forbes and A. Shpilka. On identity testing of tensors, low-rank recovery and compressed sensing. In *Proceedings of the 44th ACM Symposium on Theory of Computing*, pages 163–172, 2012.

6 E. M. Gabidulin. Theory of codes with maximal rank distance. *Problems of Information Transmission*, 21(7):1–12, 1985.

7 E. M. Gabidulin. A fast matrix decoding algorithm for rank-error-correcting codes. In G. D. Cohen, S. Litsyn, A. Lobstein, and G. Zémor, editors, *Algebraic Coding*, volume 573 of *Lecture Notes in Computer Science*, pages 126–133. Springer, 1991.

8 E. M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov. Ideals over a non-commutative ring and their applications in cryptology. In D. W. Davies, editor, *EUROCRYPT*, volume 547 of *Lecture Notes in Computer Science*, pages 482–489. Springer, 1991.

9 V. Guruswami and S. Kopparty. Explicit subspace designs. In *Proceedings of the 54th IEEE Symposium on Foundations of Computer Science*, 2013.

10 V. Guruswami, S. Narayanan, and C. Wang. List decoding subspace codes from insertions and deletions. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 183–189, January 2012.

11 V. Guruswami and A. Rudra. Explicit codes achieving list decoding capacity: Error-correction with optimal redundancy. *IEEE Transactions on Information Theory*, 54(1):135–150, 2008.

12 V. Guruswami and C. Wang. Linear-algebraic list decoding for variants of Reed-Solomon codes. *IEEE Transactions on Information Theory*, 59(6):3257–3268, 2013.

13 V. Guruswami and C. Xing. Folded codes from function field towers and improved optimal rate list decoding. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:36, 2012. Extended abstract appeared in the *Proceedings of the 44th ACM Symposium on Theory of Computing (STOC'12)*.

14 V. Guruswami and C. Xing. List decoding Reed-Solomon, Algebraic-Geometric, and Gabidulin subcodes up to the Singleton bound. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:146, 2012. Extended abstract appeared in the *Proceedings of the 45th ACM Symposium on Theory of Computing (STOC'13)*.

15 E. Kaltofen. Polynomial factorization 1987–1991. *Proceedings of LATIN '92, LNCS*, 583:294–313, 1992.

16 R. Koetter and F. R. Kschischang. Coding for errors and erasures in random network coding. *IEEE Transactions on Information Theory*, 54(8):3579–3591, 2008.

17 P. Loidreau. Designing a rank metric based McEliece cryptosystem. In N. Sendrier, editor, *PQCrypto*, volume 6061 of *Lecture Notes in Computer Science*, pages 142–152. Springer, 2010.

18 H. Lu and P. V. Kumar. A unified construction of space-time codes with optimal rate-diversity tradeoff. *IEEE Transactions on Information Theory*, 51(5):1709–1730, 2005.

19 P. Lusina, E. M. Gabidulin, and M. Bossert. Maximum rank distance codes as space-time codes. *IEEE Transactions on Information Theory*, 49(10):2757–2760, 2003.

20 H. Mahdavifar and A. Vardy. Algebraic list-decoding on the operator channel. In *Proceedings of the IEEE International Symposium on Information Theory*, pages 1193–1197, 2010.

21 H. Mahdavifar and A. Vardy. List-decoding of subspace codes and rank-metric codes up to Singleton bound. *CoRR*, abs/1202.0866, 2012.

22 R. M. Roth. Maximum-rank array codes and their application to crisscross error correction. *IEEE Transactions on Information Theory*, 37(2):328–336, 1991.

23 D. Silva, F. R. Kschischang, and R. Koetter. A rank-metric approach to error control in random network coding. *IEEE Transactions on Information Theory*, 54(9):3951–3967, 2008.

24 A. Wachter-Zeh. Bounds on list decoding of rank-metric codes. *IEEE Transactions on Information Theory*, 59(11):7268–7277, 2013.

A Explicit List-decodable Subspace Codes

A.1 The Operator Channel and Subspace Codes

For a vector space W, let $\mathcal{P}(W)$ denote the set of all subspaces of W, and $\mathcal{P}_n(W)$ the set of all n-dimensional subspaces of W.

We recall the definition of the operator channel from [16].

▶ **Definition 19.** An *operator channel* C associated with the *ambient space* W is a channel with input and output alphabet $\mathcal{P}(W)$. The channel input V and output U are related by

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

where $k = \dim(U \cap V)$, E is an error subspace (wlog E may be taken such that $E \cap V = \{0\}$), and $\mathcal{H}_k(V)$ is an operator returning an arbitrary k-dimensional subspace of V.

In transforming V to U, we say that operator channel commits $r = \dim(V) - k$ *deletions* and $t = \dim(E)$ *insertions*.

A *subspace code* C is a subset of $\mathcal{P}_n(\mathbb{F}_q^t)$ for some n. We define the *rate* of a subspace code to be

$$R(C) = \frac{\log_q |C|}{nt}.$$

A.2 The Kötter-Kschischang (KK) Code

Our constructions will be subcodes of the KK code (as introduced in [16]), which we now define.

For n dividing t, let $\mathbb{F}_{h^t}$ extend $\mathbb{F}_h$, and let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_{h^t}$ generate the subfield $\mathbb{F}_{h^n} := \mathbb{F}_q$.

Set $m = t/n$. Then the (n, k, t) **KK code** encodes an $\mathbb{F}_h$-linearized polynomial over $\mathbb{F}_{q^m} = \mathbb{F}_{h^t}$ of q-degree $< k$ by

$$f(X) \mapsto \text{span}\{(\alpha_i, f(\alpha_i)\}_{i=1}^n.$$

The encoding of f is an n-dimensional vector space in the ambient space of dimension $n + t$ over $\mathbb{F}_h$.

When $k < n$, this code has distance $2(n - k + 1)$ and rate

$$\frac{\log_h q^{mk}}{n(n+t)} = \frac{k}{n}\left(\frac{1}{1+n/t}\right) \approx \frac{k}{n} \quad (\text{when } n \ll t).$$

If the channel commits $\leqslant \mu$ deletions and $\leqslant \rho$ insertions, where $s\mu + \rho < s(n - k + 1)$, Guruswami and Xing [14] give a list-decoding algorithm which outputs a $(s-1, m, k)$-periodic subspace in $\mathbb{F}_q^{mk}$ containing all candidate messages.

A.3 List-decodable Subcodes

By restricting the coefficients of the message polynomial f to come from distinct $H_1, \ldots, H_k$ from the $\big(s, 2(m-1)s/\varepsilon, t\big)$-subspace design of Theorem 8, and setting $m \approx s/\varepsilon$, we can prune the list down to a $\mathbb{F}_h$-subspace of dimension $O(s^2/\varepsilon^2)$.

Notice that the H_i's are $\mathbb{F}_h$-linear subspaces, so the restricted subcode is linear. In summary, we have:

▶ **Theorem 20.** *For every $\varepsilon > 0$ and integer $s > 0$, there exists an explicit linear subcode of the $\big(n, k, sn/\varepsilon\big)$ KK code of rate $(1 - \varepsilon)k/n$ which is list-decodable from ρ insertions and μ deletions, provided $\rho + s\mu < s(n - k + 1)$.*

Moreover, the output list is contained in an $\mathbb{F}_h$-subspace of dimension $O(s^2/\varepsilon^2)$.