Construction of optimal insdel codes from linearized polynomials

Rakhi Pratihar
Project team - GRACE, Inria Saclay Centre

(joint work with Vaneet Aggarwal,
Purdue University)

NonCommutative Rings and their Applications (NCRA) VIII
Lens, Aug 28 - 31, 2023

Overview

Part I
Preliminaries

Insertion-deletion metric

$\mathbb{F}_q$ - finite field with q elements, q a prime power.

- The **insdel distance** $d_{insdel}(\mathbf{a}, \mathbf{b})$ between two words $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^n$ is the smallest number of insertions and deletions of coordinates required to get one from the other.

Insertion-deletion metric

$\mathbb{F}_q$ - finite field with q elements, q a prime power.

- The **insdel distance** $d_{insdel}(\mathbf{a}, \mathbf{b})$ between two words $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^n$ is the smallest number of insertions and deletions of coordinates required to get one from the other.
- A **common subsequence** of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^n$: a sequence $\mathbf{u}$ of length r $(0 \le r \le n)$ such that there are indices $1 \le i_1 < i_2 < \ldots < i_r \le n$ and $1 \le j_1 < j_2 < \ldots < j_r \le n$ satisfying

$$(a_{i_1}, \ldots, a_{i_r}) = u = (b_{j_1}, \ldots, b_{j_r}).$$

Lemma

Let $LCS(\mathbf{a}, \mathbf{b})$ be a ***largest common subsequence*** *of $\mathbf{a}$ and $\mathbf{b}$. Then*

$$d_{insdel}(\mathbf{a}, \mathbf{b}) = 2(n - \ell), \quad \textit{where } \ell = |LCS(\mathbf{a}, \mathbf{b})|.$$

Insertion-deletion (or insdel) codes

- To deal with this synchronization errors, the class of codes, called insertion and deletion (insdel for short) are introduced in the 1960s by Varshamov, Tenengolts, and Levenshtein.

Insertion-deletion (or insdel) codes

- To deal with this synchronization errors, the class of codes, called insertion and deletion (insdel for short) are introduced in the 1960s by Varshamov, Tenengolts, and Levenshtein.
- An $(n, M, d)_q$-insdel code $\mathcal{C}$ is a subset of $\mathbb{F}_q^n$ of size M and minimum insdel distance d, i.e., $d = \min\{d_{insdel}(\mathbf{a}, \mathbf{b}) : \mathbf{a}, \mathbf{b} \in \mathcal{C},\ \mathbf{a} \neq \mathbf{b}\}$.
- d_{insdel} is indeed a metric on $\mathbb{F}_q^n$. Also, note that $d_{insdel}(\mathbf{a}, \mathbf{b}) \leq 2d_H(\mathbf{a}, \mathbf{b})$ for any $\mathbf{a},\ \mathbf{b} \in \mathbb{F}_{q^n}$, where d_H is the Hamming distance.

Insertion-deletion (or insdel) codes

- To deal with this synchronization errors, the class of codes, called insertion and deletion (insdel for short) are introduced in the 1960s by Varshamov, Tenengolts, and Levenshtein.
- An $(n, M, d)_q$-insdel code $\mathcal{C}$ is a subset of $\mathbb{F}_q^n$ of size M and minimum insdel distance d, i.e., $d = \min\{d_{insdel}(\mathbf{a}, \mathbf{b}) : \mathbf{a}, \mathbf{b} \in \mathcal{C},\ \mathbf{a} \neq \mathbf{b}\}$.
- d_{insdel} is indeed a metric on $\mathbb{F}_q^n$. Also, note that $d_{insdel}(\mathbf{a}, \mathbf{b}) \leq 2d_H(\mathbf{a}, \mathbf{b})$ for any $\mathbf{a},\ \mathbf{b} \in \mathbb{F}_{q^n}$, where d_H is the Hamming distance.

Example

For a normal basis $\{\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}\}$ of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, $d_{insdel}(\mathbf{a}, \mathbf{a^q}) = 2$, where $\mathbf{a} = (\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}})$. But $d_H(\mathbf{a}, \mathbf{a^q}) = m$.

Optimal (non-)linear insdel codes

$\mathcal{C} \subseteq \mathbb{F}_q^n$ - an $(n, d_{insdel})_q$ insdel code. Then

Lemma (Singleton-like bound)

$$|\mathcal{C}| \leq q^{n - \frac{d_{insdel}}{2} + 1}. \tag{1}$$

A code achieving the bound (1) *is called insdel-metric Singleton-optimal.*

Optimal (non-)linear insdel codes

$\mathcal{C} \subseteq \mathbb{F}_q^n$ - an $(n, d_{insdel})_q$ insdel code. Then

Lemma (Singleton-like bound)

$$|\mathcal{C}| \leq q^{n - \frac{d_{insdel}}{2} + 1}. \tag{1}$$

A code achieving the bound (1) *is called insdel-metric Singleton-optimal.*

- If rate $R := \frac{log_q(\mathcal{C})}{n}$, relative distance $\delta := \frac{d_{insdel}}{2n}$, then the Singleton bound implies $R + \delta \leq 1$.

Optimal (non-)linear insdel codes

$\mathcal{C} \subseteq \mathbb{F}_q^n$ - an $(n, d_{insdel})_q$ insdel code. Then

Lemma (Singleton-like bound)

$$|\mathcal{C}| \leq q^{n-\frac{d_{insdel}}{2}+1}. \tag{1}$$

A code achieving the bound (1) *is called insdel-metric Singleton-optimal.*

- If rate $R := \frac{log_q(\mathcal{C})}{n}$, relative distance $\delta := \frac{d_{insdel}}{2n}$, then the Singleton bound implies $R + \delta \leq 1$.

Theorem (Con, Shpilka, and Tamo, 2023)

Every linear insdel code that is capable of correcting a δ fraction of deletions has rate at most $\frac{1-\delta}{2} + o(1)$.

Part II

Construction of insdel codes from subspace codes

Construction from subspace codes: why is it natural?

- $\mathcal{P}_q(n)$ - the set of all $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^n$.
- Subspace codes ($\mathcal{C} \subseteq \mathcal{P}_q(n)$) were introduced for error-control in network coding through operator channel.

Definition (Koetter and Kschischang, 2008)

An operator channel associated with $\mathbb{F}_q^n$ is a channel with input and output alphabet $\mathcal{P}_q(n)$. A channel input U is related to the corresponding output V as

$$V = (U \cap V) \oplus E,$$

where $E \in \mathcal{P}_q(n)$ is an error space. In this case, the channel commits $t = \dim U - \dim(U \cap V)$ erasures and $\rho = \dim E$ errors.

Note that the errors and erasures an operator channel commits are essentially measured by insertion and deletion of dimension, respectively

Insdel codes from subspace codes

Construction (Chen, 2021)

$\mathcal{C} \subseteq \mathcal{G}_q(n,k)$ - a constant-dimension $[n, k, \log_q|\mathcal{C}|, d]$-type subspace code. The induced insdel code from $\mathcal{C}$:

$$Span(\mathcal{C}) := \{\underbrace{(\beta_1, \dots, \beta_k)}_{c_U} : \{\beta_i : i = 1, \dots, k\} \text{ is a basis of } U \text{ for } U \in \mathcal{C}\}.$$

Insdel codes from subspace codes

Construction (Chen, 2021)

$\mathcal{C} \subseteq \mathcal{G}_q(n,k)$ - a constant-dimension $[n, k, \log_q|\mathcal{C}|, d]$-type subspace code. The induced insdel code from $\mathcal{C}$:

$$Span(\mathcal{C}) := \{\underbrace{(\beta_1, \ldots, \beta_k)}_{c_U} : \{\beta_i : i = 1, \ldots, k\} \text{ is a basis of } U \text{ for } U \in \mathcal{C}\}.$$

- **Subspace distance**: For $U, V \in \mathcal{P}_q(n)$, $d_S(U,V) = \dim(U+V) - \dim(U \cap V)$. It defines a metric on $\mathcal{P}_q(n)$.
- $Span(\mathcal{C})$ is a nonlinear insdel code over $\mathbb{F}_{q^n}$ of length k and insdel distance $d_{insdel}(Span(\mathcal{C})) \geq d_S(\mathcal{C})$ as follows: for $U, V \in \mathcal{C}$ with $l = |LCS(c_U, c_V)|$,

$$d_S(U,V) = 2(\dim(U) - \dim(U \cap V)) \leq 2(k-l) = d_{insdel}(c_U, c_V).$$

Optimal non-linear insdel codes from subspace codes

- (Koetter and Kschischang, 2008) $\mathcal{C} \subseteq \mathcal{G}_q(n,k)$ - a constant dimension subspace code with subspace distance d. Then asymptotic Singleton bound in terms of rate $R = \frac{log_q(|\mathcal{C}|)}{nk}$, normalized weight $\lambda = \frac{k}{n}$, relative distance $\delta = \frac{d}{2k}$ is

$$R \leq (1-\delta)(1-\lambda) + \frac{1}{\lambda n}(1-\lambda+o(1)). \tag{2}$$

Optimal non-linear insdel codes from subspace codes

- (Koetter and Kschischang, 2008) $\mathcal{C} \subseteq \mathcal{G}_q(n,k)$ - a constant dimension subspace code with subspace distance d. Then asymptotic Singleton bound in terms of rate $R = \frac{log_q(|\mathcal{C}|)}{nk}$, normalized weight $\lambda = \frac{k}{n}$, relative distance $\delta = \frac{d}{2k}$ is

$$R \leq (1-\delta)(1-\lambda) + \frac{1}{\lambda n}(1-\lambda+o(1)). \tag{2}$$

Definition (Linearized polynomials)

A linearized polynomial over $\mathbb{F}_{q^m}$ - $\sum_i f_i X^{q^i}$ where $f_i \in \mathbb{F}_{q^m}$ and only finitely many f_i's are nonzero. The largest i with f_i nonzero is called its q-degree.

We denote by $\mathcal{L}_k[X]_{q^m} := \{f_0 X + f_1 X^q + \ldots + f_{k-1} X^{q^{k-1}} : f_i \in \mathbb{F}_{q^m}\}$.

Optimal insdel codes from interleaved subspace codes

- $\mathcal{A} = \{\alpha_1, \ldots, \alpha_{n_t}\} \subseteq \mathbb{F}_{q^m}$ - a set of $\mathbb{F}_q$-linearly independent elements with $n_t \leq m$, $\langle \mathcal{A} \rangle_q$ - the $\mathbb{F}_q$-space generated by the elements in $\mathcal{A}$.
- Let $W_s = \langle \mathcal{A} \rangle_q \oplus \underbrace{\mathbb{F}_{q^m} \oplus \cdots \oplus \mathbb{F}_{q^m}}_{s \text{ times}}$, a $\mathbb{F}_q$-space of dimension $n_t + sm$.

Optimal insdel codes from interleaved subspace codes

- $\mathcal{A} = \{\alpha_1, \ldots, \alpha_{n_t}\} \subseteq \mathbb{F}_{q^m}$ - a set of $\mathbb{F}_q$-linearly independent elements with $n_t \leq m$, $\langle \mathcal{A} \rangle_q$ - the $\mathbb{F}_q$-space generated by the elements in $\mathcal{A}$.
- Let $W_s = \langle \mathcal{A} \rangle_q \oplus \underbrace{\mathbb{F}_{q^m} \oplus \cdots \oplus \mathbb{F}_{q^m}}_{s \text{ times}}$, a $\mathbb{F}_q$-space of dimension $n_t + sm$.
- For fixed integers $k_1, \cdots, k_s < n_t$, an interleaved subspace code $\mathcal{C}^{(s)}$ is the collection of n_t-dimensional subspace of W_s
$\left\langle \{(\alpha_i, f^{(1)}(\alpha_i), \cdots, f^{(s)}(\alpha_i)) : i = 1, \ldots, n_t\} \right\rangle_q$, where $f^{(j)}(x) \in \mathcal{L}_{k_j}[x]_{q^m}$ for $j = 1, \ldots, s$.

Optimal insdel codes from interleaved subspace codes

- $\mathcal{A} = \{\alpha_1, \ldots, \alpha_{n_t}\} \subseteq \mathbb{F}_{q^m}$ - a set of $\mathbb{F}_q$-linearly independent elements with $n_t \leq m$, $\langle \mathcal{A} \rangle_q$ - the $\mathbb{F}_q$-space generated by the elements in $\mathcal{A}$.
- Let $W_s = \langle \mathcal{A} \rangle_q \oplus \underbrace{\mathbb{F}_{q^m} \oplus \cdots \oplus \mathbb{F}_{q^m}}_{s \textit{ times}}$, a $\mathbb{F}_q$-space of dimension $n_t + sm$.
- For fixed integers $k_1, \cdots, k_s < n_t$, an interleaved subspace code $\mathcal{C}^{(s)}$ is the collection of n_t-dimensional subspace of W_s
  $\left\langle \{(\alpha_i, f^{(1)}(\alpha_i), \cdots, f^{(s)}(\alpha_i)) : i = 1, \ldots, n_t\} \right\rangle_q$, where $f^{(j)}(x) \in \mathcal{L}_{k_j}[x]_{q^m}$ for $j = 1, \ldots, s$.
- $|\mathcal{C}^{(s)}| = q^{m(\sum_{i=1}^{s} k_i)}$ and $d_S(\mathcal{C}^{(s)}) = 2(n_t - \max_j k_j + 1)$
- Set $n_t = m$ and $k_i = m/2$ for $i = 1, \ldots, s$, then $\mathrm{Span}(\mathcal{C}^{(s)})$ has

$$R = \frac{s}{2(s+1)} \text{ and } \delta = 1/2 + 1/m, \text{ i.e., } R + \delta \to 1 \text{ when } s \to \infty$$

Improved construction

- We can take more basis vectors for each subspace in a constant dimensional subspace code $\mathcal{C}$ to get a larger code than $Span(\mathcal{C})$.

Question: Let U be an n-dimensional subspace of $\mathbb{F}_q^m$. Any ordered basis $\{\alpha_1, \ldots, \alpha_n\}$ is considered as an n-tuple $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_{q^m}^n$. Let S be a collection of $\mathbb{F}_q$-basis vectors of U such that any two vectors $\alpha, \beta \in S$ has a largest common sequence of length at most l. What is the largest possible size of S?

Improved construction

- We can take more basis vectors for each subspace in a constant dimensional subspace code $\mathcal{C}$ to get a larger code than $Span(\mathcal{C})$.

Question: Let U be an n-dimensional subspace of $\mathbb{F}_q^m$. Any ordered basis $\{\alpha_1, \ldots, \alpha_n\}$ is considered as an n-tuple $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_{q^m}^n$. Let S be a collection of $\mathbb{F}_q$-basis vectors of U such that any two vectors $\alpha, \beta \in S$ has a largest common sequence of length at most l. What is the largest possible size of S?

Proposition (Aggarwal and P., 2023)

For $\alpha = (\alpha_1, \cdots, \alpha_n) \in S$, we denote by $\sigma(\alpha)$ the permuted vector $(\alpha_{\sigma(1)}, \cdots, \alpha_{\sigma(n)})$ where $\sigma \in S_n$. Let $S_{n,i}$ denote the set of $(123 \cdots i)$-avoiding permutations in S_n. Then we can have a larger collection of S with $|S| = (q-1) + \Sigma_{i=1}^{l} \binom{n}{i} |S_{n,i}|(q-1)(q^i - 1)$.

Part III

Construction of linear insdel codes from rank metric codes

Rank metric codes

- A (linear) vector rank-metric code over the finite extension $\mathbb{F}_{q^m}/\mathbb{F}_q$ of length n and dimension k is an $\mathbb{F}_{q^m}$-subspace of $\mathbb{F}_{q^m}^n$ of dimension k.
- The *rank* of an element $\alpha = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{F}_{q^m}^n$ is defined by
$$rank(\alpha) := \dim \langle \alpha_1, \ldots, \alpha_n \rangle_{\mathbb{F}_q}.$$
The rank function induces a metric $\mathbf{d}_r$, called *rank metric*, on $\mathbb{F}_{q^m}^n$ where $\mathbf{d}_r(\alpha, \alpha') := rank(\alpha - \alpha')$ for α, α' in $\mathbb{F}_{q^m}^n$.

Rank metric codes

- A (linear) vector rank-metric code over the finite extension $\mathbb{F}_{q^m}/\mathbb{F}_q$ of length n and dimension k is an $\mathbb{F}_{q^m}$-subspace of $\mathbb{F}_{q^m}^n$ of dimension k.
- The *rank* of an element $\alpha = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{F}_{q^m}^n$ is defined by
$$rank(\alpha) := \dim \langle \alpha_1, \ldots, \alpha_n \rangle_{\mathbb{F}_q} .$$
The rank function induces a metric $\mathbf{d}_r$, called *rank metric*, on $\mathbb{F}_{q^m}^n$ where $\mathbf{d}_r(\alpha, \alpha') := rank(\alpha - \alpha')$ for α, α' in $\mathbb{F}_{q^m}^n$.

Definition (Delsarte, '78, and Gabidulin, '85)

The Gabidulin code over $\mathbb{F}_{q^m}/\mathbb{F}_q$ of length n and dimension k by evaluation at a $\mathbb{F}_q$-linearly independent set $\alpha = \{\alpha_1, \cdots, \alpha_n\}$ is defined as
$$Gab(q; m, n, k, \alpha) := \{(f(\alpha_1), \ldots, f(\alpha_n)) : f(X) \in \mathcal{L}_k[X]_{q^m}\}.$$

Algebraic condition for optimal linear insdel code

For two vectors $I = (1 \leq I_1 < \cdots < I_{2k-1} \leq n)$ and $J = (1 \leq J_1 < \cdots < J_{2k-1} \leq n)$, consider the matrix

$$V_{I,J,q}(\mathbf{X}) = \begin{bmatrix} X_{I_1} & X_{I_1}^{q} & \cdots & X_{I_1}^{q^{k-1}} & X_{J_1}^{q} & \cdots & X_{J_1}^{q^{k-1}} \\ X_{I_2} & X_{I_2}^{q} & \cdots & X_{I_2}^{q^{k-1}} & X_{J_2}^{q} & \cdots & X_{J_2}^{q^{k-1}} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ X_{I_{2k-1}} & X_{I_{2k-1}}^{q} & \cdots & X_{I_{2k-1}}^{q^{k-1}} & X_{J_{2k-1}}^{q} & \cdots & X_{J_{2k-1}}^{q^{k-1}} \end{bmatrix}.$$

Proposition (Aggarwal, and P., 2023)

Consider the $[n, k]$ linearized RS code or Gabidulin code $Gab(q; m, n, k, \alpha)$ over $\mathbb{F}_{q^m}/\mathbb{F}_q$ with evaluation vector $\alpha = (\alpha_1, \ldots, \alpha_n)$. If for any two increasing vectors $I, J \in [n]^{2k-1}$ that agree on at most $k-1$ coordinates, it holds that $\det(V_{I,J,q}(\alpha)) \neq 0$, then the code can correct any $n - 2k + 1$ insdel errors.

Optimal linear insdel codes from Gabidulin codes

The algebraic condition is an adaptation of the algebraic condition for Reed-Solomon codes to be optimal linear insdel codes given by Con, Shpilka, and Tamo.

Optimal linear insdel codes from Gabidulin codes

The algebraic condition is an adaptation of the algebraic condition for Reed-Solomon codes to be optimal linear insdel codes given by Con, Shpilka, and Tamo.

Theorem (Aggarwal, and P. , 2023)

Let k, n be positive integers such that $2k-1 \leq n$ and $q \geq 3$ be any prime power. For $m = O(n^{4k-2})$ there exists an $[n, k]$ linearized Reed-Solomon code or Gabidulin code over $\mathbb{F}_{q^m}/\mathbb{F}_q$ obtained by evaluating linearized polynomials over $\mathbb{F}_{q^m}/\mathbb{F}_q$ of degree at most $k-1$ at n $\mathbb{F}_q$-linearly independent elements $\alpha_1, \ldots, \alpha_n$ of $\mathbb{F}_{q^m}$ that can recover from $n-2k+1$ adversarial insertion-deletion errors.

Optimal linear insdel codes from Gabidulin codes

The algebraic condition is an adaptation of the algebraic condition for Reed-Solomon codes to be optimal linear insdel codes given by Con, Shpilka, and Tamo.

Theorem (Aggarwal, and P. , 2023)

Let k, n be positive integers such that $2k-1 \leq n$ and $q \geq 3$ be any prime power. For $m = O(n^{4k-2})$ there exists an $[n, k]$ linearized Reed-Solomon code or Gabidulin code over $\mathbb{F}_{q^m}/\mathbb{F}_q$ obtained by evaluating linearized polynomials over $\mathbb{F}_{q^m}/\mathbb{F}_q$ of degree at most $k-1$ at n $\mathbb{F}_q$-linearly independent elements $\alpha_1, \ldots, \alpha_n$ of $\mathbb{F}_{q^m}$ that can recover from $n-2k+1$ adversarial insertion-deletion errors.

But finding an explicit Gabidulin code satisfying the algebraic condition is still open...

Part IV

Nonlinear codes by combining Sidon spaces

Sidon spaces

$\mathcal{G}_q(n,k)$ - set of all k-dimensional $\mathbb{F}_q$-subspaces of $\mathbb{F}_{q^n}$

- $U \in \mathcal{G}_q(n,k)$ is a **Sidon space** if for all $a,b,c,d \in U \setminus \{0\}$, $ab = cd$ implies $\{a\mathbb{F}_q, b\mathbb{F}_q\} = \{c\mathbb{F}_q, d\mathbb{F}_q\}$.
- $\mathcal{C} \subseteq \mathcal{G}_q(n,k)$ is a **cyclic subspace code** if $\alpha U := \{\alpha u \colon u \in U\} \in \mathcal{C}$ for all $\alpha \in \mathbb{F}_{q^n}^*$ and $U \in \mathcal{C}$.
- cyclic subspace codes can be obtained as orbits of the group action

$$\begin{aligned} \mathbb{F}_{q^n}^* \times \mathcal{G}_q(n,k) &\to \mathcal{G}_q(n,k) \\ (\alpha, U) &\mapsto \alpha U = \{\alpha u \colon u \in U\}. \end{aligned}$$

Lemma (Roth, Raviv, Tamo, 2018)

For $U \in \mathcal{G}_q(n,k)$, $Orb(U)$ is cyclic subspace code of size $\frac{q^n-1}{q-1}$ and minimum distance $2k-2$ if and only if U is a Sidon space.

Linear insdel codes from Sidon spaces

Lemma

Any Sidon space $U \in \mathcal{G}_q(n,k)$ gives an one dimensional linear insdel code over $\mathbb{F}_{q^n}$ with length k and minimum distance $2k-2$.

Linear insdel codes from Sidon spaces

Lemma

Any Sidon space $U \in \mathcal{G}_q(n,k)$ gives an one dimensional linear insdel code over $\mathbb{F}_{q^n}$ with length k and minimum distance $2k-2$.

Theorem (Niu, Xiao, Gao, 2022)

For $k \geq 2$ and $n = 3k$, let ξ be a primitive element in $\mathbb{F}_{q^k}$ and γ be the root of an irreducible polynomial of degree n/k over $\mathbb{F}_{q^k}$. Set $\gamma_i = \xi^i\gamma$ and $\gamma_j = \xi^j\gamma$ for $0 \leq i,j \leq q^k-2$. Then for $0 \leq i,\, j \leq q^k - 2$,

$$U_i = \{u + (u^q - u)\gamma_i : u \in \mathbb{F}_{q^k}\} \text{ and } V_j = \{v + v^q\gamma_j : v \in \mathbb{F}_{q^k}\}$$

are Sidon spaces of dimension k. Moreover,

$$\dim(U_i \cap \alpha_3 V_j) \leq 1,\ \dim(\alpha_1 U_i \cap \mathbb{F}_{q^k}) \leq 1,\ \text{and}\ \dim(\alpha_2 V_j \cap \mathbb{F}_{q^k}) \leq 1$$

for all $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_{q^n}^*$ and for every $i, j \in \{0, \cdots, q^k-2\}$.

Nonlinear insdel codes

- For each element of the set S of $(2q^k - 1)$ Sidon spaces of dimension k in $\mathbb{F}_{q^n}$, we consider the corresponding linear 2-dimensional RS codes as follows.

Nonlinear insdel codes

- For each element of the set S of $(2q^k - 1)$ Sidon spaces of dimension k in $\mathbb{F}_{q^n}$, we consider the corresponding linear 2-dimensional RS codes as follows.
- For $q = 3$ and $m \in \mathbb{N}$, let $U \subseteq \mathbb{F}_{3^{6m}}$ be a $2m$-dimensional Sidon space over $\mathbb{F}_3$. Let $u_1, \ldots, u_{2m}$ be a basis of U. [Roth, Raviv, Tamo, 2018]
- Let $H = (h_{i,j})$ be a $2m \times ((3^m + 1)/2)$ parity check matrix of an $[(3^m + 1)/2, (3^m + 1)/2 - 2m]_3$ linear code with minimum distance at least 5. [Gashkov and Sidelnikov, 1986]
- Then our $[n, 2]_{3^{6m}}$ RS codes $\mathcal{C}_U$ of length $n = (3^m + 1)/2$, defined by the evaluation points $\alpha_j = \Sigma_{i=1}^{2m} u_i h_{i,j}$ for $1 \leq j \leq (3^m + 1)/2$ can correct from $n - 3$ insdel errors.

Nonlinear insdel codes

- For each element of the set S of $(2q^k - 1)$ Sidon spaces of dimension k in $\mathbb{F}_{q^n}$, we consider the corresponding linear 2-dimensional RS codes as follows.
- For $q = 3$ and $m \in \mathbb{N}$, let $U \subseteq \mathbb{F}_{3^{6m}}$ be a $2m$-dimensional Sidon space over $\mathbb{F}_3$. Let $u_1, \ldots, u_{2m}$ be a basis of U. [Roth, Raviv, Tamo, 2018]
- Let $H = (h_{i,j})$ be a $2m \times ((3^m + 1)/2)$ parity check matrix of an $[(3^m + 1)/2, (3^m + 1)/2 - 2m]_3$ linear code with minimum distance at least 5. [Gashkov and Sidelnikov, 1986]
- Then our $[n, 2]_{3^{6m}}$ RS codes $\mathcal{C}_U$ of length $n = (3^m + 1)/2$, defined by the evaluation points $\alpha_j = \Sigma_{i=1}^{2m} u_i h_{i,j}$ for $1 \leq j \leq (3^m + 1)/2$ can correct from $n - 3$ insdel errors.
- $\mathcal{C} := \bigcup_{U \in S} \mathcal{C}_U$ is a larger nonlinear code with same error correcting capacity.

References

- V. Aggarwal, and R. Pratihar, *Insdel codes from subspace and rank-metric codes*, Discrete Mathematics, 2023.
- H. Chen, *Explicit good subspace-metric codes and subset-metric codes*, arXiv:2108.12334, 2021.
- R. Con, A. Shpilka, and I. Tamo, *Reed-Solomon codes against adversarial insertions and deletions*, IEEE Trans. Inf. Theory, 2023.
- R. Kötter and F. R. Kschischang, *Coding for errors and erasures in random network coding*, IEEE Trans. Inform. Theory, 2008.
- V. I. Levenshtein, *Binary codes capable of correcting deletions, insertions and reversals*, Doklady Akademii Nauk SSSR, 1965.
- M. Niu, J. Xiao, Y. Gao, *New constructions of large cyclic subspace codes via Sidon spaces*, Advances in Mathematics of Communications, 2022.