Code-Based Cryptography in Different Metrics I

Anna-Lena Horlemann

Institute of Computer Science, University of St.Gallen, Switzerland

attacc, DLR Oberpfaffenhofen, 2024

Disclaimer

- This is an overview of many results out there, which is certainly not complete.
- More focus on algebraic structures, less on probabilistic algorithms.
- For most references I refer to
  - V. Weger, N. Gassner, and J. Rosenthal. A survey on code-based cryptography. https://arxiv.org/abs/2201.07119, 2022.

  For references about the sum-rank metric or the subspace metric please contact me.
- There are two parts to this:
  - Part 1: Hamming, rank, and sum-rank metric
  - Part 2: subspace and Lee metric, and connection to lattices

1. General Setup for CBC
2. The Hardness of the Syndrome Decoding Problem
3. Hamming Metric
4. Rank Metric
5. Sum-Rank Metric
6. Comparison and Overview

Main idea of code-based cryptosystems

- Decoding a random linear code is a hard problem.
- **Public key/information**: the parity check matrix of a random (looking) linear code, and a syndrome
- **Secret**: the solution to the corresponding syndrome decoding problem: usually a low-weight error vector (and/or the corresponding message/codeword)

$$\underbrace{\mathbf{s}}_{\text{syndrome}} = \underbrace{\mathbf{e}}_{\text{error vector}} \cdot \underbrace{\mathbf{H}^\top}_{\text{PC matrix}}$$

Main idea of code-based cryptosystems

- Decoding a random linear code is a hard problem.
- **Public key/information**: the parity check matrix of a random (looking) linear code, and a syndrome
- **Secret**: the solution to the corresponding syndrome decoding problem: usually a low-weight error vector (and/or the corresponding message/codeword)

$$\underbrace{\mathbf{s}}_{\text{syndrome}} = \underbrace{\mathbf{e}}_{\text{error vector}} \cdot \underbrace{\mathbf{H}^\top}_{\text{PC matrix}}$$

Various weights:

- Hamming weight
- rank weight
- Lee weight
- etc. (homogeneous weight, sum rank weight)

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

- an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
- in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

- an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
- in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

$\Longrightarrow$ we can relax or generalize both of the above

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

- an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
- in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

$\Longrightarrow$ we can relax or generalize both of the above

But what do we really need?

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

- an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
- in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

$\Longrightarrow$ we can relax or generalize both of the above

But what do we really need?

- We need an efficient representation of the code (e.g. by linearity).
- If we do not think about weights/distances any more, it is not code-based crypto.
- For PKE we need an efficient decoding algorithm.
- For identification schemes decoding is not necessary. But we need transitive "linear" maps on the spheres (in the existing schemes), and identifiers of the cosets (e.g. syndromes).

A generalized McEliece framework

- The receiver chooses a code C with generator matrix G and an efficient decoding algorithm. Moreover, they need a disguising function ϕ that is a (near-)isometry.
- **Private key:** G and the decoding algorithm
- **Public key:** $\phi(G)$ together with the error correction capability of the code generated by $\phi(G)$, say $\hat{t}$
- **Encryption:** Choose a random error vector e of weight at most $\hat{t}$ and encrypt the message m as

$$c = m\,\phi(G) + e.$$

- **Decryption:** Compute $\phi^{-1}(c)$ and decode this in the secret code C to recover m.

A generalized Niederreiter framework

- The receiver chooses a code C with parity check matrix H and an efficient decoding algorithm. Moreover, they need a disguising function ϕ that is a (near-)isometry.
- **Private key:** H and the decoding algorithm
- **Public key:** $\phi(H)$ together with the error correction capability of the code which is the kernel of $\phi(H)$, say $\hat{t}$
- **Encryption:** Represent the message as an error vector e of weight at most $\hat{t}$ and encrypt it as

$$c = e\,\phi(H)^\top.$$

- **Decryption:** Compute $\phi^{-1}(c)$ and decode this in the secret code C to recover e and hence the message.

Necessary features for security and efficiency

- We need a code C that has **efficient encoding and decoding** algorithms.
- We assume that the code family used is known, hence we require this family to have enough elements to prevent brute force attacks.
- The error correction capability needs to be large enough such that a brute force attack on the possible errors can be prevented.
- We need a **disguising function** such that the original code cannot be found from the public generator matrix.
- Any **generic decoding** should be **infeasible**.

The syndrome decoding problem (SDP)

Definition (Syndrome Decoding Problem)

Given some $t \in \mathbb{N}$, a parity check matrix $H \in \mathbb{F}_q^{r \times n}$ and a syndrome $s \in \mathbb{F}_q^r$, find a vector $e \in \mathbb{F}_q^n$ of weight at most t that fulfills $eH^\top = s$.

The syndrome decoding problem (SDP)

Definition (Syndrome Decoding Problem)

Given some $t \in \mathbb{N}$, a parity check matrix $H \in \mathbb{F}_q^{r \times n}$ and a syndrome $s \in \mathbb{F}_q^r$, find a vector $e \in \mathbb{F}_q^n$ of weight at most t that fulfills $eH^\top = s$.

Theorem

Let $\mathrm{wt} : \mathbb{F}_q \to \mathbb{R}_{\geq 0}$ *be a function that satisfies the following properties:*

1. $\mathrm{wt}(0) = 0$,
2. $\mathrm{wt}(1) = 1$ *and* $\mathrm{wt}(x) \geq 1$ *for all* $x \neq 0$.

If wt *is extended additively to* $\mathbb{F}_q^n$ *(i.e., by adding the weights of the coordinates) then the SDP with respect to* wt *is NP-complete.*

This includes the **Hamming, Lee**, and the **homogeneous weight**, however not the rank weight (for the sum-rank weight it depends). For the rank weight there is a probabilistic reduction from NP to the SDP.

Hamming metric and isometries

Definition (Hamming distance)

$$d_H((u_1, \ldots, u_n), (v_1, \ldots, v_n)) := |\{i \mid u_i \neq v_i\}|$$

Theorem

The linear isometries for the metric space $(\mathbb{F}_q^n, d_H)$ are represented by the monomial matrices, i.e., matrices of the form

$$\begin{pmatrix} v_1 & 0 & \ldots & 0 \\ 0 & v_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & v_n \end{pmatrix} P$$

for some $P \in S_n$ and non-zero $v_1, \ldots, v_n \in \mathbb{F}_q$.
These linear isometries can be extended to semi-linear isometries by combining them (coordinate-wise) with the Galois group of $\mathbb{F}_q$.

The original McEliece scheme

- The receiver chooses a binary Goppa code $C \subseteq \mathbb{F}_2^n$ of dimension $k = n - mt$ with generator matrix G and minimum distance $2t + 1$. Moreover, they choose a random $A \in \mathrm{GL}_k(2)$ and $P \in S_n$.
- **Private key:** G and (A, P) (decoder for C given by G)
- **Public key:** $G' = AGP$
- **Encryption:** Choose a random error vector e of weight at most t and encrypt the message m as

$$c = m\,G' + e.$$

- **Decryption:** Compute $c' = cP^{-1}$ and decode this in C to recover mA. Recover m by multiplying with A^{-1}.

Other variants

Use other codes like

- Reed-Solomon codes,
- Reed-Muller codes,
- polar codes,
- algebraic geometry codes,
- etc.

Use other disguising functions like

- larger weight monomial matrices,
- concatenating with columns,
- etc.

McEliece with LDPC/MDPC codes

- Choose an efficiently decodable LDPC/MDPC code with sparse parity check matrix $H \subseteq \mathbb{F}_q^{r \times n}$.
- **Private key:** H (decoder for C given by H)
- **Public key:** G any generator matrix for the code defined by H
- **Encryption:** Choose a random error vector e of weight at most t and encrypt the message m as

$$c = m\,G + e.$$

- **Decryption:** Decode c in C (with the help of H) to recover m.

McEliece with LDPC/MDPC codes

- Choose an efficiently decodable LDPC/MDPC code with sparse parity check matrix $H \subseteq \mathbb{F}_q^{r \times n}$.
- **Private key:** H (decoder for C given by H)
- **Public key:** G any generator matrix for the code defined by H
- **Encryption:** Choose a random error vector e of weight at most t and encrypt the message m as

$$c = m\,G + e.$$

- **Decryption:** Decode c in C (with the help of H) to recover m.

Possible attack: Find low weight codewords in the dual code to set up a low density parity check matrix.

Digital signatures

- Hash and sign
  - given a message m, find a random m' such that $Hash(m, H_{pub}, m')$ is a decodable syndrome (error vector is main part of signature)
  - Problem: finding this m' is inefficient
  - Example: CFS (Courtois-Finiasz-Sendrier)

Digital signatures

- Hash and sign
  - given a message m, find a random m' such that $Hash(m, H_{pub}, m')$ is a decodable syndrome (error vector is main part of signature)
  - Problem: finding this m' is inefficient
  - Example: CFS (Courtois-Finiasz-Sendrier)
- Fiat-Shamir transform of zero-knowledge identification (ZKID) schemes
  - public information: random parity check matrix
  - secret: decodable error vector
  - interactive (5-pass) scheme, where one reveals either the weight of the error vector, or that it is a solution to the syndrome equation
  - Problem: need many rounds to get a security
  - Example: CVE (Cayrel-Véron-Alaoui)

Digital signatures

- Hash and sign
  - given a message m, find a random m' such that $Hash(m, H_{pub}, m')$ is a decodable syndrome (error vector is main part of signature)
  - Problem: finding this m' is inefficient
  - Example: CFS (Courtois-Finiasz-Sendrier)
- Fiat-Shamir transform of zero-knowledge identification (ZKID) schemes
  - public information: random parity check matrix
  - secret: decodable error vector
  - interactive (5-pass) scheme, where one reveals either the weight of the error vector, or that it is a solution to the syndrome equation
  - Problem: need many rounds to get a security
  - Example: CVE (Cayrel-Véron-Alaoui)

Philosophical question: Is this still CBC? No decoding is needed.

Digital signatures

- Hash and sign
  - given a message m, find a random m' such that $Hash(m, H_{pub}, m')$ is a decodable syndrome (error vector is main part of signature)
  - Problem: finding this m' is inefficient
  - Example: CFS (Courtois-Finiasz-Sendrier)
- Fiat-Shamir transform of zero-knowledge identification (ZKID) schemes
  - public information: random parity check matrix
  - secret: decodable error vector
  - interactive (5-pass) scheme, where one reveals either the weight of the error vector, or that it is a solution to the syndrome equation
  - Problem: need many rounds to get a security
  - Example: CVE (Cayrel-Véron-Alaoui)

Philosophical question: Is this still CBC? No decoding is needed.
BUT: We need results about the minimum distance of a random linear code.

Minimum distance of random codes

Theorem

- *For $q \to \infty$ a random linear code in $\mathbb{F}_q^n$ is MDS by probability 1.*
- *For $n \to \infty$ a random linear code in $\mathbb{F}_q^n$ gets arbitrarily close to the Gilbert-Varshamov bound by probability 1.*

$\Longrightarrow$ Use these results as an estimate for the minimum distance of the random linear code in an ZKID scheme.

— Attacks —

Generic decoding (message recovery attacks)

Information set decoding (ISD), birthday/meet-in-the-middle decoders

(see presentations of Andre Esser)

Distinguishers (key recovery attacks)

Definition

Schur/star/coordinate-wise product:

$$u \star v := \left(u_1 v_1, \ldots, u_n v_n\right),$$

$$C^{\star 2} := \langle \{ c \star d \mid c, d \in C \} \rangle.$$

Theorem

If $C \subseteq \mathbb{F}_q^n$ is a random linear code then by high probability

$$\dim(C^{\star 2}) = \min\left\{ \binom{k+1}{2}, n \right\},$$

whereas if C is a (generalized) Reed-Solomon code, then

$$\dim(C^{\star 2}) = \min\{2k-1, n\}.$$

Example of a distinguisher attack

Wieschebrink cryptosystem

- The receiver chooses a RS code $C \subseteq \mathbb{F}_q^n$ of dimension k with generator matrix G. They also choose $C_1, \ldots, C_r \in \mathbb{F}_q^k$, $S \in \mathrm{GL}_k(q)$ and $P \in S_{n+r}$ uniformly at random. Let $\bar{G}$ be the matrix obtained by concatenating G and the columns $C_1, \ldots, C_r$.
- **Private key:** G and (S, P)
- **Public key:** $G' = S^{-1}\bar{G}P^{-1}$
- **Encryption:** Choose a random error vector e of weight at most t and encrypt the message m as

$$c = m\,G' + e.$$

- **Decryption:** Compute $c' = cP^{-1}$, erase the last r coordinates and decode in C to recover mS^{-1}. Recover m by multiplying with S.

Example of a distinguisher attack

Attack

- Choose a random $1 \leq i \leq n + r$ and shorten G' in the ith position (i.e., erase the ith column in G').
- Compute the dimension of the square code of this new generator matrix.
- If the dimension of the square is

$$\begin{cases} 2k + r - 2 & \text{then the column was prob. from } \{C_1, \ldots, C_r\}, \\ 2k + r - 1 & \text{then the column was prob. from the GRS code.} \end{cases}$$

- Identify all random columns like above to recover a generator matrix for the original GRS code.
- Apply Sidelnikov-Shestakov to recover the code locators and column multipliers to be able to decode.

Distinguishers (key recovery attacks)

The star square code distinguisher works for:

- low-codimensional subcodes of GRS codes,
- Reed-Muller codes,
- Polar codes,
- special types of Goppa codes,
- high rate alternant codes,
- algebraic geometry codes.

The rank metric and isometries

Definition

The *rank distance* d_R for $A, B \in \mathbb{F}_q^{m\times n}$:

$$d_R(A, B) := \operatorname{rank}(A - B).$$

Can also be represented via a vector space isomorphism $\mathbb{F}_q^{m\times n} \cong \mathbb{F}_{q^m}^n$.

Theorem

The linear isometries for the metric space $(\mathbb{F}_{q^m}^n, d_R)$ are represented by the elements of $\mathrm{GL}_n(q)$.
These linear isometries can be extended to semi-linear isometries by combining them (coordinate-wise) with the Galois group of $\mathbb{F}_{q^m}$.

GPT (Gabidulin-Paramonov-Tretjakov) cryptosystem

- The receiver chooses a Gabidulin code of minimum rank distance $2t+1$ with generator matrix $G \subseteq \mathbb{F}_{q^m}^{k \times n}$. Moreover, they randomly choose $S \in \mathrm{GL}_k(q^m)$ and $X \in \mathbb{F}_{q^m}^{k \times n}$ of column rank at most $s < t$.
- **Private key:** G and S
- **Public key:** $G' = SG + X$
- **Encryption:** Choose a random error vector e of rank weight at most $t - s$ and encrypt the message m as

$$c = m\, G' + e.$$

- **Decryption:** Decode c in the Gabidulin code to recover mS. Multiply with S^{-1} to recover m.

Digital signatures

Analog to the Hamming case...

- Hash and sign
  - given a message m, find a random m' such that $Hash(m, H_{pub}, m')$ is a decodable syndrome (error vector is main part of signature)
  - Problem: finding this m' is inefficient
  - Example: RankSign
- Fiat-Shamir transform of zero-knowledge identification (ZKID) schemes
  - interactive (5-pass) scheme, where one reveals either the weight of the error vector, or that it is a solution to the syndrome equation
  - Problem: need many rounds to get a security
  - Example: RYDE

Minimum distance of random codes

Theorem

- *For $m \to \infty$ a random linear code in $\mathbb{F}_{q^m}^n$ is MRD by probability 1.*
- *For $n \to \infty$ a random linear code in $\mathbb{F}_{q^m}^n$ gets arbitrarily close to the Gilbert-Varshamov bound by probability 1.*

$\Longrightarrow$ Use these results as an estimate for the minimum distance of the random linear code in an ZKID scheme.

— Attacks —

Generic decoding (message recovery attacks)

1. Basic idea of ISD does not work in the rank metric.
2. Most generic decoders make use of the rank decomposition of the error

$$\underbrace{e}_{\in \mathbb{F}_{q^m}^n, wt_R = t} = \underbrace{t}_{\in \mathbb{F}_{q^m}^t, wt_R = t} \cdot \underbrace{T}_{\in \mathbb{F}_q^{t \times n}}$$

   and try to find one of the two parts on the right side first.

$\Longrightarrow$ Generally much slower than in Hamming metric
(for comparable parameters)

Distinguishers (key recovery attacks)

Theorem

If $C \subseteq \mathbb{F}_{q^m}^n$ is a random linear code of dimension $0 < k < n$ then by high probability

$$\dim(C \cap C^{(q)}) = \max\{2k - n, 0\},$$

whereas if C is a Gabidulin code then

$$\dim(C \cap C^{(q)}) = k - 1$$

or equivalently

$$\dim(C + C^{(q)}) = k + 1.$$

Note that this behavior differs as soon as $1 < k < n - 1$, i.e., for any non-trivial Gabidulin code.

Example of a distinguisher attack

Lemma

Let $G \in \mathbb{F}_{q^m}^{k \times n}$ and X have column rank s. Then all elements of rank one in

$$\sum_{i=0}^{s} \langle G + X \rangle^{(q^i)}$$

exactly span the $\mathbb{F}_q$-space of X.

Idea: $\sum_{i=0}^{s} \langle X \rangle^{(q^i)}$ is the whole space, whereas $\sum_{i=0}^{s} \langle G \rangle^{(q^i)}$ is Gabidulin code of minimum rank distance > 1.

Rank one elements can easily be found as solutions to $x^q - x = 0$.

The sum-rank metric

... is a concatenation of the Hamming and the rank metric

Definition

Let $x = (x_1 | \dots | x_t) \in \mathbb{F}_{q^m}^n$. The ($t$-)*sum-rank distance* is defined as

$$d_{SR}(x, y) := \sum_{i=1}^{t} d_R(x_i, y_i).$$

The sum-rank metric

... is a concatenation of the Hamming and the rank metric

Definition

Let $x = (x_1 | \dots | x_t) \in \mathbb{F}_{q^m}^n$. The ($t$-)*sum-rank distance* is defined as

$$d_{SR}(x, y) := \sum_{i=1}^{t} d_R(x_i, y_i).$$

- The isometries are a (block-wise) combination of the isometries of the Hamming and the rank metric.
- There is a Singleton-type bound, which can be attained by a linearized Reed-Solomon code construction (MSRD codes).

Minimum distance of random codes

Theorem

- *For $m \to \infty$ a random linear code in $\mathbb{F}_{q^m}^n$ is MSRD by prob.* 1.
- *For $n \to \infty$ a random linear code in $\mathbb{F}_{q^m}^n$ gets arbitrarily close to the Gilbert-Varshamov bound by probability* 1.

— Attacks —

Distinguishers (key recovery attacks)

- **Linearized Reed-Solomon codes** are a generalization of concatenating RS with Gabidulin codes:
  - **Outer code:** (generalized) RS-code of length t.
  - **Inner code:** Each coordinate is replaced with a block and filled with a Gabidulin code.
- Generalization ideas are via skew polynomials and an operator evaluation with flexible field automorphisms and derivations.

Distinguishers (key recovery attacks)

- **Linearized Reed-Solomon codes** are a generalization of concatenating RS with Gabidulin codes:
  - **Outer code:** (generalized) RS-code of length t.
  - **Inner code:** Each coordinate is replaced with a block and filled with a Gabidulin code.
- Generalization ideas are via skew polynomials and an operator evaluation with flexible field automorphisms and derivations.
- In the trivial automorphism case we can use the **star square code** distinguisher.
- If the outer code is known we can use (a generalization of) the **Frobenius** distinguisher.
- We can use the known **key recovery attacks**
  - first to recover the parameters of the outer RS code
  - then to recover the parameters of the inner Gabidulin code

  only in the case where we have trivial automorphism and trivial derivation.

Current ideas from there....

Great: We cannot use the known distinguisher attacks for non-trivial automorphisms...

Current ideas from there....

Great: We cannot use the known distinguisher attacks for non-trivial automorphisms...

... let's use this for cryptography!

Current ideas from there....

Great: We cannot use the known distinguisher attacks for non-trivial automorphisms...

... let's use this for cryptography!

But then: We can also go one step back and only use the outer code (skew RS code)...

... let's see what comes out.

Comparison and overview

	Hamming	rank	sum-rank
NP-hard	YES	probabilistic	YES/probl.
good random codes	YES	YES	YES
distinguishers	star product	Frobenius intersection/sum	star + Frobenius (in trivial case)
generic decoding	x	$> x$	$\geq x$