CRYPTO READING GROUP
HAMMING QUASI CYCLIC (HQC)

Spring 2026

UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

Contents

About

This document contains a detailed plan for the code-based cryptography study group.

Code-Based Cryptography

This semester, the reading group will focus on *code-based cryptography*. The idea is to explore a topic that lies somewhat outside the usual comfort zone of most of us, and to use the reading group as an opportunity to learn together.

Our main objective will be to build enough background to understand the design, security, and implementation aspects of the HQC cryptosystem. More broadly, the reading group will also serve as an introduction to some of the central ideas and techniques in code-based cryptography.

Structure

Each week, two participants will be assigned to prepare the material for that session together. The goal is to encourage collaboration, make the preparation more manageable, and allow for discussion before the presentation itself.

The speakers may then decide how they would like to present the material. For instance, they may choose to split the talk into two parts and each present one half, or they may decide that only one of them will deliver the presentation while both contribute to the preparation. Some coordination with the speakers from previous and following weeks may occasionally be helpful, but we expect this to remain fairly light.

Useful Information

Talks. The talks will be held in the **Mathematics & Computer Science (MC)** building at the University of Waterloo, in room **MC 547**, located on the fifth floor.

Lunches. We are also planning to organize group lunches every other week, alternating between the lounge and the plaza.

Organizers

Leonardo Colò	University of Waterloo
Seunghoon Lee	University of Waterloo
Bruno Sterner	University of Waterloo

Schedule

Hamming Quasi-Cyclic

	Spring 2026		
May 15, 2026	1	**Bruno Sterner** **Seunghoon Lee**	Introduction to coding theory
May 22, 2026	2	**Maher Mamah** **Elle Wen**	Introduction to code based cryptography
May 22, 2026	**Lunch in the Plaza**		
May 29, 2026	3	**Pranshu Kumar** **John Premkumar** **Karaneh Keypoor**	Quasi-cyclic codes
June 1, 2026	4	**Roman Langrehr** **Sam Jaques**	Information Set Decoding
June 1, 2026	**Lunch in the Lounge**		
June 12, 2026	5	**Camryn Steckel** **Speaker 10**	Decoding for quasi-cyclic codes
June 19, 2026	6	**Mojtaba Fadavi** **Speaker 12**	HQC PKE/KEM
June 19, 2026	**Lunch in the Plaza**		
June 26, 2026	7	**Bruce Xu** **Maggie Simmons**	HQC implementation/optimization
July 3, 2026	8	**Speaker 15** **Speaker 16**	Optional
July 3, 2026	**End of the term lunch**		

Plan

Introduction to Coding Theory I

Seunghoon Lee & Bruno Sterner 1

This first session introduces the basic language of coding theory and explains why coding-theoretic objects became relevant to cryptography. We will discuss the main definitions and parameters of codes, present linear codes as the fundamental setting of the theory, and study some classical examples such as Reed–Solomon codes and Goppa codes. The goal is to provide the algebraic and combinatorial background needed for the rest of the reading group.
References: [§2.1-2.2, 12] and [9]

Introduction to Code-Based Cryptography

Maher Mamah & Elle Wen 2

This session serves as an introduction to code-based cryptography. We will explain the decoding problem and its central role as a hardness assumption, then present the general McEliece framework and its security intuition. We will also give a first overview of the Goppa-code instantiation, with the aim of understanding how structured codes can be used to build public-key encryption schemes.
References: [§3.1, 12]

Quasi-Cyclic Codes

Pranshu Kumar & John Premkumar & Karaneh Keypoor 3

This session is devoted to quasi-cyclic codes, one of the main structured code families used in modern code-based cryptography. We will introduce their definition and main properties, and explain why their additional algebraic structure is both useful for efficiency and delicate from a security perspective. This week will provide the background needed to understand HQC and related constructions.
References: [7] and [§3.4, 12]

Information Set Decoding

Roman Langrehr & Sam Jaques 4

In this session, we study information set decoding (ISD), one of the main generic approaches for attacking code-based cryptosystems. We will present the basic ideas behind Prange's algorithm and Stern's algorithm, together with the general philosophy of decoding attacks in the random-code setting. The aim is to understand both the algorithmic framework and its importance in concrete security estimates.
References: [2] and [§5.3, 12]

Decoding for Quasi-Cyclic Codes

Camryn Steckel & 5

This session focuses on decoding questions specific to quasi-cyclic codes. We will discuss syndrome decoding in the quasi-cyclic setting and compare generic ISD methods with approaches that exploit additional structure. The goal is to better understand the tension between efficiency and security, and to prepare the ground for the study of the HQC scheme.
References: [§6.3, 4], [§3, 6], and [§5, 10]

HQC PKE/KEM

Mojtaba Fadavi & 6

This session is devoted to the HQC cryptosystem itself, in both its public-key encryption and key-encapsulation forms. We will explain how the scheme works, describe its main components and design choices, and discuss the corresponding security analysis, including comments on the post-quantum setting. By this stage, the reading group should have enough background to appreciate both the structure and the rationale of HQC.
References: [1] and [4]

HQC Implementation and Optimisation

Bruce Xu & Maggie Simmons 7

This week will examine practical aspects of HQC, with an emphasis on implementation and optimisation. The purpose of this session is to complement the theoretical study of HQC with a more concrete understanding of how the scheme behaves in practice.
References: [3] and [4]

Optional Session: Further Directions in Code-Based Cryptography

2 Speakers 8

This optional final session may be used to explore further directions in code-based cryptography. Possible topics include code-based signature schemes currently submitted to NIST, such as CROSS or LESS, or a more advanced topic building on the study of HQC and quasi-cyclic codes. Exact contentto be adjusted.
References: [5] and [11]

Week 1

Introduction to coding theory

Talk by Seunghoon and Bruno.

Motivation

Coding theory studies how to encode messages so that they can be recovered even after errors occur during transmission or storage. Informally, we start with a message m, add redundancy, and obtain an encoded message $m' = m + (\text{redundancy})$. If some errors occur, the redundancy may allow us to recover the original message.

$$m' + (\text{errors}) \longmapsto m.$$

One of the classical motivations comes from the work of Richard Hamming. Coding theory is also an important background for code-based cryptography, one of the main families of post-quantum cryptography.

Linear Codes

Let q be a prime power, and let $\mathbb{F}_q$ be the finite field with q elements.

Definition ($[n,k]$-linear code)**.** *Let $k, n \in \mathbb{Z}_{>0}$ with $k \leq n$. An $[n,k]$-*linear code *over $\mathbb{F}_q$ is a k-dimensional linear subspace $\mathcal{C} \subseteq \mathbb{F}_q^n$. An element $\mathbf{c} \in \mathcal{C}$ is called a codeword. The rate of $\mathcal{C}$ is $R = \frac{k}{n}$.*

Thus, n is the length of the code, while k is the dimension, or the number of information symbols (message).

Example (Repetition code)**.** *The binary repetition code of length 3 is $\mathcal{C} = \{000, 111\} \subseteq \mathbb{F}_2^3$. This is a $[3,1]$-linear code. It encodes one bit $m \in \mathbb{F}_2$ as $m \longmapsto (m,m,m)$.*

Generator and Parity-Check Matrices

Definition (Generator matrix)**.** *A matrix $G \in \mathbb{F}_q^{k \times n}$ is a generator matrix of an $[n,k]$-linear code $\mathcal{C}$ if $\mathcal{C} = \left\{\mathbf{x}G : \mathbf{x} \in \mathbb{F}_q^k\right\} = \langle G \rangle$, where $\langle G \rangle$ denotes the row span of G.*

Equivalently, the rows of G form a basis of $\mathcal{C}$.

Example. *For the binary repetition code $\mathcal{C} = \{000, 111\}$, a generator matrix is $G = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{1 \times 3}$. For $m \in \mathbb{F}_2$, the corresponding codeword is $\mathbf{c} = mG = (m,m,m)$.*

Definition (Parity-check matrix). *A matrix $H \in \mathbb{F}_q^{(n-k)\times n}$ is a parity-check matrix of $\mathcal{C}$ if $\mathcal{C} = \left\{\mathbf{y} \in \mathbb{F}_q^n : H\mathbf{y}^T = 0\right\}$.*

Equivalently, H gives linear equations that characterize membership in $\mathcal{C}$.

Example. *For the repetition code $\mathcal{C} = \{000, 111\}$, a vector $\mathbf{y} = (y_1, y_2, y_3)$ is a codeword precisely when $y_1 + y_2 = 0 \quad$ and $\quad y_1 + y_3 = 0$. Hence, we may take $H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. One can check that $GH^T = 0$. Note that the parity-check matrix may not be unique. For our example, $H = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ is also a parity-check matrix, which comes from the equations $y_1 + y_3 = 0$ and $y_2 + y_3 = 0$.*

Definition (Syndrome). *For any received word $\mathbf{r} \in \mathbb{F}_q^n$, the vector $H\mathbf{r}^T$ is called the syndrome of $\mathbf{r}$.*

If $\mathbf{r} = \mathbf{c} + \mathbf{e}$ for some codeword $\mathbf{c} \in \mathcal{C}$ and error vector $\mathbf{e}$, then $H\mathbf{r}^T = H(\mathbf{c} + \mathbf{e})^T = H\mathbf{c}^T + H\mathbf{e}^T = H\mathbf{e}^T$. Thus, the syndrome depends only on the error vector $\mathbf{e}$, not on the transmitted codeword $\mathbf{c}$.

Example. *For the repetition code with $H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, consider the single-bit errors $\mathbf{e}_1 = 100, \mathbf{e}_2 = 010, \mathbf{e}_3 = 001$. Their syndromes are $H\mathbf{e}_1^T = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, H\mathbf{e}_2^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, H\mathbf{e}_3^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Since these syndromes are distinct, the syndrome tells us the location of a single error.*

Hamming Weight and Hamming Distance

Definition (Hamming weight). *The Hamming weight of $\mathbf{x} \in \mathbb{F}_q^n$ is $\mathrm{wt}(\mathbf{x}) = |\{i \in [n] : x_i \neq 0\}|$.*

Definition (Hamming distance). *The Hamming distance between $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$ is $d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}) = |\{i \in [n] : x_i \neq y_i\}|$.*

Definition (Minimum Hamming distance). *The minimum Hamming distance of a code $\mathcal{C}$ is $d_{\mathrm{H}}(\mathcal{C}) = \min\{d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{y}\}$.*

Example. *For the repetition code $\mathcal{C} = \{000, 111\}$, we have $d_{\mathrm{H}}(\mathcal{C}) = d_{\mathrm{H}}(000, 111) = 3$.*

Exercise. *Let $\mathcal{C}$ be a linear code. Show that $d_{\mathrm{H}}(\mathcal{C}) = \min\{\mathrm{wt}(\mathbf{x}) : \mathbf{x} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{0}\}$.*

Proof. Since $\mathcal{C}$ is linear, if $\mathbf{x}, \mathbf{y} \in \mathcal{C}$, then $\mathbf{x} - \mathbf{y} \in \mathcal{C}$. Also, $d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}) = \mathrm{wt}(\mathbf{x} - \mathbf{y})$. Therefore, $d_{\mathrm{H}}(\mathcal{C}) = \min\{d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{y}\} = \min\{\mathrm{wt}(\mathbf{x} - \mathbf{y}) : \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{y}\}$. As $\mathbf{x} - \mathbf{y}$ ranges over nonzero codewords of $\mathcal{C}$, this becomes $d_{\mathrm{H}}(\mathcal{C}) = \min\{\mathrm{wt}(\mathbf{z}) : \mathbf{z} \in \mathcal{C},\ \mathbf{z} \neq \mathbf{0}\}$. □

Example. *The previous statement can fail if $\mathcal{C}$ is not linear. For example, let $\mathcal{C} = \{000, 111, 110\} \subseteq \mathbb{F}_2^3$. This code is not linear, since $111 + 110 = 001 \notin \mathcal{C}$. Here, $d_{\mathrm{H}}(\mathcal{C}) = d_{\mathrm{H}}(111, 110) = 1$, but $\min\{\mathrm{wt}(\mathbf{x}) : \mathbf{x} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{0}\} = \mathrm{wt}(110) = 2$.*

Error Correction

Definition (Correction up to t errors)**.** *We say that a code $\mathcal{C}$ can correct up to t errors if, for every received word $\mathbf{r} \in \mathbb{F}_q^n$ with $d_{\mathrm{H}}(\mathbf{r}, \mathcal{C}) \leq t$, there exists a unique codeword $\mathbf{c} \in \mathcal{C}$ such that $d_{\mathrm{H}}(\mathbf{r}, \mathbf{c}) \leq t$.*

Here, $d_{\mathrm{H}}(\mathbf{r}, \mathcal{C}) = \min_{\mathbf{c} \in \mathcal{C}} d_{\mathrm{H}}(\mathbf{r}, \mathbf{c})$. A decoding algorithm is then expected to map such a received word $\mathbf{r}$ to the unique closest codeword $\mathbf{c}$.

Exercise. *Let $\mathcal{C}$ be a linear code of length n over $\mathbb{F}_q$ with $d_{\mathrm{H}}(\mathcal{C}) = d$. Show that $\mathcal{C}$ can correct up to $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ errors.*

Proof. First, we show existence. Suppose $\mathbf{r}$ is obtained from some codeword $\mathbf{c} \in \mathcal{C}$ by adding an error vector $\mathbf{e}$ with $\mathrm{wt}(\mathbf{e}) \leq t$. Then $\mathbf{r} = \mathbf{c} + \mathbf{e}$. Thus, $d_{\mathrm{H}}(\mathbf{r}, \mathbf{c}) = \mathrm{wt}(\mathbf{e}) \leq t$. So there exists at least one codeword within distance t of $\mathbf{r}$.

Next, we show uniqueness. Suppose $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ are two codewords such that $d_{\mathrm{H}}(\mathbf{r}, \mathbf{c}_1) \leq t$ and $d_{\mathrm{H}}(\mathbf{r}, \mathbf{c}_2) \leq t$. By the triangle inequality, $d_{\mathrm{H}}(\mathbf{c}_1, \mathbf{c}_2) \leq d_{\mathrm{H}}(\mathbf{c}_1, \mathbf{r}) + d_{\mathrm{H}}(\mathbf{r}, \mathbf{c}_2) \leq 2t$. Since $t = \lfloor (d-1)/2 \rfloor$, we have $2t \leq d - 1$. Therefore, $d_{\mathrm{H}}(\mathbf{c}_1, \mathbf{c}_2) \leq d - 1$. But distinct codewords in $\mathcal{C}$ have distance at least d. Hence $\mathbf{c}_1$ and $\mathbf{c}_2$ cannot be distinct. Therefore, $\mathbf{c}_1 = \mathbf{c}_2$. □

$[n, k, d]$-Linear Codes and the Singleton Bound

Definition. *An $[n, k, d]$-linear code is an $[n, k]$-linear code $\mathcal{C}$ such that $d = d_{\mathrm{H}}(\mathcal{C})$.*

There is a tradeoff between rate and minimum distance. We want the rate $R = \frac{k}{n}$ to be large, but we also want the relative distance $\delta = \frac{d}{n}$ to be large.

Theorem (Singleton bound)**.** *Let $\mathcal{C}$ be an $[n, k]$-linear code over $\mathbb{F}_q$ with minimum distance $d = d_{\mathrm{H}}(\mathcal{C})$. Then $d \leq n - k + 1$.*

Proof. Define a map $\varphi : \mathcal{C} \to \mathbb{F}_q^{n-d+1}$ by deleting the last $d-1$ coordinates: $\varphi(x_1, \ldots, x_n) = (x_1, \ldots, x_{n-d+1})$. We claim that φ is injective. Suppose not. Then there exist distinct codewords $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ such that $\varphi(\mathbf{c}_1) = \varphi(\mathbf{c}_2)$. This means that $\mathbf{c}_1$ and $\mathbf{c}_2$ agree on the first $n - d + 1$ coordinates, so they can differ only in the remaining $d - 1$ coordinates. Hence $d_{\mathrm{H}}(\mathbf{c}_1, \mathbf{c}_2) \leq d - 1$, which contradicts the definition of d.

Therefore, φ is injective. Hence $|\mathcal{C}| \leq |\mathbb{F}_q^{n-d+1}| = q^{n-d+1}$. Since $\mathcal{C}$ has dimension k, we also have $|\mathcal{C}| = q^k$. Therefore, $q^k \leq q^{n-d+1}$, so $k \leq n - d + 1$. Rearranging gives $d \leq n - k + 1$. □

Intuition. The minimum distance d means that any two distinct codewords differ in at least d positions. Therefore, even after deleting $d - 1$ positions, two distinct codewords cannot become equal. This is the key reason why the deletion map in the proof is injective.

Definition (Maximum distance separable code)**.** *A code $\mathcal{C}$ satisfying $d = n - k + 1$ is called a maximum distance separable code, or MDS code.*

MDS codes achieve the maximum possible minimum distance for fixed n and k. Therefore, they can correct the maximum possible number of errors for those parameters. Reed–Solomon codes are a central example of MDS codes.

Hamming Balls and the Gilbert–Varshamov Bound

For $r \geq 0$, define the Hamming ball of radius r in $\mathbb{F}_q^n$ by $B_{\mathrm{H}}(r, n, q) = \left\{\mathbf{x} \in \mathbb{F}_q^n : \mathrm{wt}(\mathbf{x}) \leq r\right\}$. Its size is $|B_{\mathrm{H}}(r, n, q)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i$.

The Gilbert–Varshamov bound gives a sufficient condition for the existence of good linear codes.

Theorem (Gilbert–Varshamov bound, informal form)**.** *If $|B_{\mathrm{H}}(d-2, n-1, q)| < q^{n-k}$, then there exists an $[n, k]$-linear code over $\mathbb{F}_q$ with minimum distance at least d.*

The main message is that good linear codes exist. For suitable parameters, one can have both nontrivial rate and nontrivial relative minimum distance. This is part of the background reason why random-looking good-rate linear codes are meaningful objects in code-based cryptography.

Nontrivial Rate and Nontrivial Minimum Distance

For a family of $[n, k, d]$-linear codes, nontrivial rate and nontrivial relative distance mean that there exist constants $R_0 > 0$ and $\delta_0 > 0$ such that $R = \frac{k}{n} \geq R_0$ and $\delta = \frac{d}{n} \geq \delta_0$. That is, both k and d are proportional to n.

Example (Repetition codes)**.** *A binary repetition code of length n has parameters $[n, 1, n]$. Thus its relative distance is 1, but its rate tends to 0 as n grows.*

Example (Hamming codes)**.** *A binary Hamming code has parameters $[2^r - 1,\ 2^r - r - 1,\ 3]$. Thus its rate tends to 1, but its relative distance tends to 0 as r grows.*

Reed-Solomon Codes

We have already seen one example of an MDS code in the $[n, 1]$ repetition code. This can be checked directly as its minimum Hamming distance is n. However this is has a low rate and a lot of reducancy is needed for correcting errors. The briefly mentioned Hamming code has a significantly better rate but is not an MDS code. We will give an overview of the *Reed Solomon code* which is an MDS code that has a very good rate.

Definition. *Fixed some distinct and non-zero $\alpha_1, \cdots, \alpha_n \in \mathbb{F}_q$ and write $\alpha = (\alpha_1, \cdots, \alpha_n)$. For an integer $k \leq n$, the* Reed Solomon code, $\mathsf{RS}_{n,k}(\alpha)$, *is defined as*

$$\mathsf{RS}_{n,k}(\alpha) \coloneqq \{(f(\alpha_1), \cdots, f(\alpha_n)) : f \in \mathbb{F}_q[x] \text{ with } \deg(f) \leq k-1\}.$$

Lemma. $\mathsf{RS}_{n,k}(\alpha)$ *is an (n,k)-linear code with the following generator and parity check matrices:*

Proof (sketch).

1. The linearity follows from the fact that $\mathbb{F}_q[x]$ is a ring.
2. For any $f(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} \in \mathbb{F}_q[x]$, the codeword $c_f = (f(\alpha_1), \cdots, f(\alpha_n))$ can be written as
$$c_f = (a_0\ a_1\ \cdots\ a_{k-1})G_{\mathsf{RS}}.$$
3. More involved and omitted (can be done by looking at the dual code). □

One can transmit a message $m = (m_0, \cdots, m_{k-1}) \in (\mathbb{F}_q)^k$ as a codeword

$$c_m = (f_m(\alpha_1), \cdots, f_m(\alpha_n)) \in \mathsf{RS}_{n,k}(\alpha),$$

where $f_m(x) = m_0 + m_1 x + \cdots + m_{k-1}x^{k-1} \in \mathbb{F}_q[x]$. The extra $n-k$ entries in this codeword compared to the message acts as the redundancy.

Question: Given some transmition $y = c_f \oplus e$ of a codeword $c_f \in \mathsf{RS}_{n,k}(\alpha)$, how many errors in c_f can we correct? Are there efficient decoding algorithms to actually correct such errors.

Proposition. $\mathsf{RS}_{n,k}(\alpha)$ *is an MDS code, i.e.* $d = d_H(\mathsf{RS}_{n,k}(\alpha)) = n - k + 1$

Proof (sketch). Reformulate the minimal Hamming distance to

$$d_H(\mathcal{C}) = \max\left\{ d' \in \mathbb{N} : \begin{array}{c}\text{any } (d'-1) \text{ columns of the parity-check} \\ \text{matrix are linearly independent}\end{array} \right\}$$

(See Roth's book [9, Theorem 2.2]).

Any $(n-k) \times (n-k)$ submatrix of H_{RS} is non-singular (since it has a Vandemonde form). So any $(n-k)$ columns of H_{RS} are linearly independent and hence $d \geq n - k + 1$. With the singleton bound we get an equality: $d = n - k + 1$. □

So one can correct at most $\lfloor (d-1)/2 \rfloor = \lfloor (n-k)/2$ errors (which is maximal among all $[n,k]$-linear codes. Remarkably the answer to the second part of the above question is yes – with this maximal number of corrections.

Theorem (informal)**.** *There are efficient decoding algorithms for* $\mathsf{RS}_{n,k}[\alpha]$ *that can correct* $\lfloor (n-k)/2$ *errors.*

One of these algorithms is based on a truncated version of the extended Euclidean algorithm.

Example. *The Reed Solomon code with $n = 255$ and $k = 223$ commonly used in many applications (e.g. communication between NASA and Voyager missions). It can correct up to 16 errors in a transmitted codeword.*

The q-ary entropy function

The q-ary entropy function is $h_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x)$. For binary codes, the Gilbert–Varshamov bound gives the asymptotic guarantee $R \gtrsim 1 - h_2(\delta)$.

Some rough tradeoffs are: $R \approx 0.8 \implies \delta \approx 0.03$, $R \approx 0.5 \implies \delta \approx 0.11$, and $R \approx 0.28 \implies \delta \approx 0.20$.

For McEliece-type parameter examples, one often sees roughly $R \approx 0.7$–0.8, $t/n \approx 0.015$–0.02, and $d/n \approx 0.03$–0.04. Here t is the number of correctable errors, and t/n is the error rate.

Week 2

Introduction to code-based cryptography

Talk by Maher and Elle.

Goppa Codes

Definition

Fix parameters:

- $m \geq 1$ positive integer, prime $q \geq 2$, $\mathbb{F}_{q^m}$ finite field
- Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_{q^m}^n$ such that α_i are distinct, $n \leq q^m$
- Pick $g(x) \in \mathbb{F}_{q^m}[x]$ *Goppa polynomial* such that $g(\alpha_i) \neq 0 \, \forall \, i$ and $\deg g(x) \leq t$

$$\Gamma(\boldsymbol{\alpha}, g) = \left\{ c \in \mathbb{F}_q^n \;\middle|\; \sum_{i=1}^{n} \frac{c_i}{x - \alpha_i} = 0 \pmod{g(x)} \right\}$$

Easy:

$$\frac{1}{x - \alpha_i} = -\frac{g(x) - g(\alpha_i)}{x - \alpha_i} \, g^{-1}(\alpha_i) \pmod{g(x)}$$

Parity-Check Matrix H

Rewrite $g(x) = \sum_{i=0}^{t} g_i x^i, \quad g_i \in \mathbb{F}_{q^m}$.

$$\frac{g(x) - g(\alpha_i)}{x - \alpha_i} = \frac{g_t(x^t - \alpha_i^t) + \cdots + g_1(x - \alpha_i) + g_0 \cdot 0}{x - \alpha_i}$$
$$= g_t\big(x^{t-1} + \alpha_i x^{t-2} + \cdots + \alpha_i^{t-1}\big) + \cdots + g_2(x + \alpha_i) + g_1$$

Look at the coefficients of a codeword $\displaystyle\sum_{i=1}^{n} c_i \cdot \frac{g(x) - g(\alpha_i)}{x - \alpha_i} \, g^{-1}(\alpha_i)$:

$$x^{t-1}: \quad g_t \cdot g^{-1}(\alpha_1)\, c_1 + \cdots + g_t \cdot g^{-1}(\alpha_n)\, c_n$$

$$\vdots$$

$$x^0: \quad \left(g_1 + \cdots + g_t \alpha_1^{t-1}\right) g^{-1}(\alpha_1)\, c_1 + \cdots + \left(g_1 + \cdots + g_t \alpha_n^{t-1}\right) g^{-1}(\alpha_n)\, c_n$$

Observe: $\deg(P) \leq t-1, \deg(g) \leq t \Rightarrow P$ *and also g(x) divides P(x). So polynomial P(x) has 0 on its coefficients*

$$c \in \Gamma(\boldsymbol{\alpha}, g) \iff \bar{H}\, c = 0 \quad \text{for } \bar{H} \in \mathbb{F}_{q^m}^{t \times n}$$

$$\mathbb{F}_{q^m}^{t \times n} \ni \bar{H} = \begin{pmatrix} g_t \cdot g^{-1}(\alpha_1) & \cdots & g_t \cdot g^{-1}(\alpha_n) \\ \vdots & & \vdots \\ \left(g_1 + \cdots + g_t \alpha_1^{t-1}\right) g^{-1}(\alpha_1) & \cdots & \left(g_1 + \cdots + g_t \alpha_n^{t-1}\right) g^{-1}(\alpha_n) \end{pmatrix}$$

Applying a bijection via a fixed basis gives:

$$\mathbb{F}_q^{tm \times n} \ni H = \begin{pmatrix} g^{-1}(\alpha_1) & \cdots & g^{-1}(\alpha_n) \\ \alpha_1\, g^{-1}(\alpha_1) & & \alpha_n\, g^{-1}(\alpha_n) \\ \vdots & & \vdots \\ \alpha_1^{t-1}\, g^{-1}(\alpha_1) & \cdots & \alpha_n^{t-1}\, g^{-1}(\alpha_n) \end{pmatrix}$$

Classic McEliece Key Generation

Set $q = 2$. Pick a monic irreducible $g(x) \in \mathbb{F}_{2^m}[x]$ of degree t, and n random distinct $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_{2^m}$. Convert to systematic form:

$$\tilde{H}(g, \boldsymbol{\alpha}) \xrightarrow{\text{systematic form}} H = (I \mid T) \in \mathbb{F}_2^{tm \times n}, \qquad \boldsymbol{\alpha} \longrightarrow \boldsymbol{\alpha}'$$

- **Secret key (sk):** $g(x),\ \boldsymbol{\alpha}'$
- **Public key (pk):** $T,\ t$ — assumes T is indistinguishable from a random binary matrix, and that the public key is useless for making decoding efficient.

$+$ Safe for large parameters (other GRS codes are broken).

$-$ Key size is too large: final-round pk size $= 261$ kilobytes (HQC is 2.2).

Decoding Goppa Codes

Received word $y = c + e$ with minimum distance $d \geq t + 1$ ($2t$ for binary), and

$$w(e) \leq \left\lfloor \frac{d-1}{2} \right\rfloor, \qquad B = \{i \mid e_i \neq 0\} \quad \text{(error positions)}.$$

The *syndrome* of y (which depends only on e) is:

$$S(x) = \sum_i \frac{y_i}{x - \alpha_i} = \sum_i \frac{e_i}{x - \alpha_i} \pmod{g(x)}.$$

Define two helper polynomials:

$$\text{error locator:} \quad \sigma(x) = \prod_{i \in B} (x - \alpha_i),$$

$$\text{error evaluator:} \quad \omega(x) = \sum_{i \in B} e_i \prod_{j \in B} (x - \alpha_j).$$

These are found via polynomial interpolation or Patterson Algorithm for binary Goppa codes(using Euclidean Algorithm). There is a polynomial-time procedure to correct t errors using degree-t polynomials.

Week 3

Quasi-cyclic codes

Week 4

Information Set Decoding

Week 5

Decoding for quasi-cyclic codes

Week 6

Hamming Quasi Cyclic (HQC)

Week 7

HQC implementations/optimizations

References

Bibliography:

[1] C. Aguilar-Melchor, O. Blazy, J. -C. Deneuville, P. Gaborit and G. Zémor. Efficient Encryption From Random Quasi-Cyclic Codes. In *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3927–3943, 2018.

[2] T. Debris-Alazard. *Code-based Cryptography Lecture Notes*, 2023.

[3] J. Dong, Y. Hou, S. Wang, L. Sha, F. Xiao, Z. Dong, and J. Lin. HIGH: Harnessing GPU Parallelism for Optimized HQC Performance. In *IACR Cryptology ePrint Archive*, 2026.

[4] HQC Team. *Hamming Quasi-Cyclic (HQC)*, NIST Submission, 2025.

[5] L. Huguenin-Dumittan and S. Vaudenay. FO-like Combiners and Hybrid Post-Quantum Cryptography. In *Cryptology and Network Security: 20th International Conference*, *CANS 2021*, pp. 225–244, 2021.

[6] C. Löndahl, T. Johansson, M. Koochak Shooshtari, M. Ahmadian-Attari, and M. Reza Aref. Squaring attacks on McEliece public-key cryptosystems using quasi-cyclic codes of even dimension. *Designs, Codes and Cryptography* , vol. 80, pp. 359–377, 2016.

[7] R. Misoczki, J.-P. Tillich, N. Sendrier, and P. S. L. M. Barreto. *MDPC-McEliece: New McEliece Variants from Moderate Density Parity-Check Codes*. 2013 IEEE International Symposium on Information Theory, 2013.

[8] NIST. *Post-Quantum Cryptography: Additional Digital Signature Schemes*.

[9] R. Roth. *Introduction to Coding Theory*. Cambridge University Press, 2006.

[10] N. Sendrier. Decoding One Out of Many. *Post-Quantum Cryptography. PQCrypto 2011. Lecture Notes in Computer Science*, vol. 7071, Springer, 2011.

[11] T. Wang, Q. Teng, A. Wang, J. Zhang, B. Pang, C. Zhao, S. Hu, and X, Wang. HARE: Compact HQC via Distance-Informed Erasure Decoding. In *Cryptology ePrint Archive*, *Paper 2026/544*, 2026.

[12] V. Weger, N. Gassner, and J. Rosenthal. *A Survey on Code-based Cryptography*. arXiv, 2024.