



Multiplicative stability properties of dual Goppa codes and cryptanalytic applications

Hugues Randriam

ANSSI, Laboratoire de cryptographie
& Télécom Paris, C²

AM-PQC, Ohrid, 2026-08-14

Preliminaries on coding theory

Preliminaries — algebra

- ▶ $\mathbb{F}$ field $\rightarrow \mathbb{F}^n$ product algebra (componentwise multiplication):

$$\mathbf{a} = (a_1, \dots, a_n), \mathbf{b} = (b_1, \dots, b_n) \rightarrow \mathbf{ab} = (a_1 b_1, \dots, a_n b_n)$$

- ▶ extend to product of linear subspaces taking the linear span:

$$A, B \subseteq \mathbb{F}^n \rightarrow AB = \langle \mathbf{ab} : \mathbf{a} \in A, \mathbf{b} \in B \rangle_{\mathbb{F}}$$

- ▶ powers: $A^{\langle 2 \rangle} = AA$, $A^{\langle r+1 \rangle} = A^{\langle r \rangle} A$
- ▶ trace linear form: $\tau(\mathbf{a}) = \text{tr}(m_{\mathbf{a}} : \mathbf{x} \mapsto \mathbf{ax}) = a_1 + \dots + a_n$
- ▶ trace bilinear form $\leftrightarrow$ standard scalar product (nondegenerate):

$$\langle \mathbf{a}, \mathbf{b} \rangle = \tau(\mathbf{ab}) = a_1 b_1 + \dots + a_n b_n$$

- ▶ “dual” $\leftrightarrow$ orthogonal:

$$A^{\perp} = \{\mathbf{b} \in \mathbb{F}^n : \forall \mathbf{a} \in A, \langle \mathbf{a}, \mathbf{b} \rangle = 0\} = (\ker \tau : A)$$

- ▶ “colon subspace”:

$$(B : A) = \{\mathbf{x} \in \mathbb{F}^n : \mathbf{x}A \subseteq B\} = (AB^{\perp})^{\perp}$$

Preliminaries — analysis

- ▶ $\mathbb{F}$ endowed with the trivial absolute value: $|0_{\mathbb{F}}| = 0$, $|x| = 1$ if $x \neq 0_{\mathbb{F}}$
- ▶ Hamming norm/weight $\leftrightarrow$ the associated L^1 norm on $\mathbb{F}^n$:

$$w_H(\mathbf{x}) = \|\mathbf{x}\| = |x_1| + \cdots + |x_n| = \#\{i : x_i \neq 0_{\mathbb{F}}\}$$

- ▶ Hamming distance/metric:

$$d_H(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \#\{i : x_i \neq y_i\}$$

- ▶ a $[n, k, d]_{\mathbb{F}}$ -code is a k -dimensional linear subspace $C \subseteq \mathbb{F}^n$ that has **minimum distance** $d(C) = d$, where

$$d(C) = d_{\min}(C) = \min_{\mathbf{c} \in C \setminus \{\mathbf{0}\}} \|\mathbf{c}\|$$

we will often drop certain of these parameters from our notation if they are implicit or unspecified $\rightarrow [n, k]$ -code

Preliminaries — geometry

- ▶ $\mathbf{G} \in \mathbb{F}^{k \times n}$ generator matrix of $C \leftrightarrow$ its rows $\mathbf{c}_1, \dots, \mathbf{c}_k$ form a basis of C
- ▶ for any polynomial $f(\mathbf{X}) = f(X_1, \dots, X_k)$ in k variables we set

$$\text{ev}_{\mathbf{G}}(f) = f(\mathbf{c}_1, \dots, \mathbf{c}_k) = (f(\mathbf{p}_1), \dots, f(\mathbf{p}_n))$$

where $\mathbf{p}_1, \dots, \mathbf{p}_n$ are the columns of $\mathbf{G}$

- ▶ this defines an evaluation map $\text{ev}_{\mathbf{G}} : \mathbb{F}[X_1, \dots, X_k] \longrightarrow \mathbb{F}^n$, and we have

$$C^{\langle r \rangle} = \text{ev}_{\mathbf{G}}(\mathbb{F}[X_1, \dots, X_k]_r)$$

- ▶ $C \sim C'$ linearly isometric $\leftrightarrow C' = C^\sigma \mathbf{a}$ for some $\sigma \in \mathfrak{S}_n$, $\mathbf{a} \in (\mathbb{F}^n)^\times$
- ▶ lin. isometry class of $C \leftrightarrow$ eq. class of multiset $\{\bar{\mathbf{p}}_1, \dots, \bar{\mathbf{p}}_n\} \subseteq \mathbb{P}^{k-1}(\mathbb{F})$
mod. projective automorphisms
- ▶ moduli space:

$$\text{GL}(\mathbb{F}^k) \setminus \mathbb{F}^{k \times n} / \mathfrak{S}_n \ltimes (\mathbb{F}^\times)^n$$

$\rightarrow$ geometric invariants allow to characterize/distinguish certain classes of codes

Preliminaries — algorithmics

Decoding problem

Given $\mathbf{G} \in \mathbb{F}^{k \times n}$ generating matrix, $\mathbf{y} \in \mathbb{F}^n$ “received vector”, $t \in \mathbb{Z}_{>0}$

→ find $\mathbf{x} \in \mathbb{F}^k$ s.t. $\|\mathbf{y} - \mathbf{x}\mathbf{G}\| \leq t$

“Syndrome decoding”

Given $\mathbf{H} \in \mathbb{F}^{(n-k) \times n}$ parity matrix, $\mathbf{s} \in \mathbb{F}^{n-k}$ syndrome, $t \in \mathbb{Z}_{>0}$

→ find $\mathbf{e} \in \mathbb{F}^n$ with $\|\mathbf{e}\| \leq t$ s.t. $\mathbf{e}\mathbf{H}^\top = \mathbf{s}$

- ▶ These two problems are equivalent.
- ▶ For generic codes they are NP-hard, and assumed to be quantum-resistant
→ foundational problem of code-based crypto.
- ▶ Best algorithms (e.g. Prange-like information set decoding) have complexity

$$(c + o(1))^t$$

Decomposition into geometric progressions

Basic setting

Let $C \subseteq \mathbb{F}^n$ be a $[n, k]$ -code, and fix (once and for all) some $\mathbf{x} \in \mathbb{F}^n$.

Definition

- ▶ $(\mathbf{c}_0, \dots, \mathbf{c}_\ell) \in C^{\ell+1}$ is a **geometric progression in C** of length ℓ (and ratio $\mathbf{x}$) if $\mathbf{c}_i = \mathbf{x}^i \mathbf{c}_0$ for all $i = 1, \dots, \ell$
- ▶ **initial codewords** of such length ℓ progressions form a subcode

$$C_\ell = \{\mathbf{c} \in C : \mathbf{x}^i \mathbf{c} \in C \text{ for all } i = 1, \dots, \ell\} = \bigcap_{0 \leq i \leq \ell} (C : \mathbf{x}^i)$$

so we get a **filtration**

$$C = C_0 \supseteq C_1 \supseteq C_2 \supseteq C_3 \supseteq \dots$$

Simplifying assumptions for this talk:

- ▶ $\mathbf{x} \in (\mathbb{F}^n)^\times$ is invertible $\rightarrow$ space of final codewords $\mathbf{x}^\ell C_\ell = \bigcap_{0 \leq i \leq \ell} \mathbf{x}^i C$
- ▶ C has no nonzero $\mathbf{x}$ -stable subcode, i.e. $C_r = 0$ for some r .

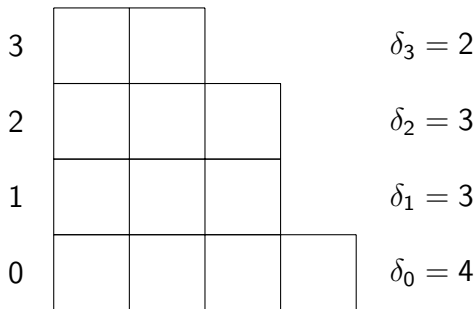
The associated diagram

Set $\delta_i = \dim C_i - \dim C_{i+1}$, and consider the diagram made of left justified rows of boxes, with row number i of length δ_i .

This diagram encodes many properties of C , e.g.

- ▶ $\dim C_i$ is equal to the total area of all rows from i and above
- ▶ columns can be filled with geometric progressions that form a basis of C

Example: $\dim C_i = 12, 8, 5, 2, 0$ for $i = 0, \dots, 4$ give



The associated diagram

Set $\delta_i = \dim C_i - \dim C_{i+1}$, and consider the diagram made of left justified rows of boxes, with row number i of length δ_i .

This diagram encodes many properties of C , e.g.

- ▶ $\dim C_i$ is equal to the total area of all rows from i and above
- ▶ columns can be filled with geometric progressions that form a basis of C

Example: $\dim C_i = 12, 8, 5, 2, 0$ for $i = 0, \dots, 4$ give

3	c	c'			$\delta_3 = 2$
2	xc	xc'	c''		$\delta_2 = 3$
1	x²c	x²c'	xc''		$\delta_1 = 3$
0	x³c	x³c'	x²c''	c'''	$\delta_0 = 4$

Shortening

Definition

$$\text{Short}_a(C) = \{(c_1, \dots, \widehat{c}_a, \dots, c_n) \in \mathbb{F}^{n-1} : (c_1, \dots, 0_a, \dots, c_n) \in C\}$$

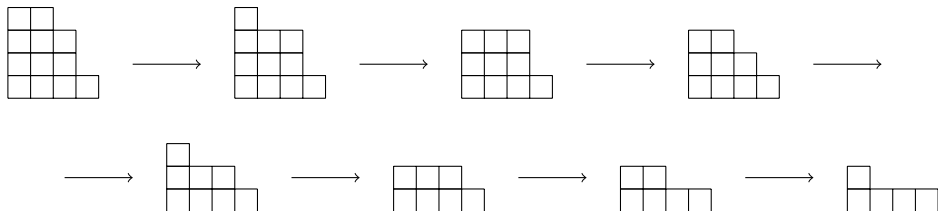
effect on parameters: $[n, k, d] \rightarrow [n-1, k-1, \geq d]$ (as long as $a \in \text{Supp}(C)$)

Lemma

$$\text{Short}_a(C)_\ell = \text{Short}_a(C_\ell)$$

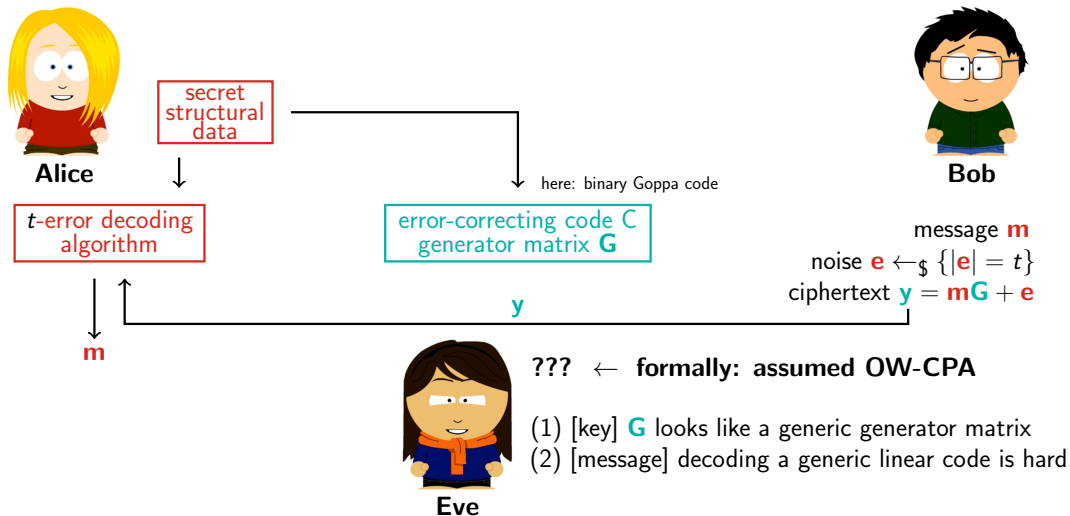
with $\mathbf{x} \leftarrow \text{Punct}_a(\mathbf{x})$

Thus for convenient a , $\dim \text{Short}_a(C)_\ell = \dim C_\ell - 1$ for all $\ell \rightarrow$ shortening one position removes the **top right** box of the diagram.



McEliece

The 1978 McEliece cryptosystem



Note:

- (1) ad hoc problem, trapdoor similar to some in today's multivariate cryptography
- (2) very classical and well-studied problem

The trapdoor

We will work with the **dual** codes (so our $\mathbf{H}$'s will act as $\mathbf{G}$'s):

- ▶ $\mathbf{x} = (x_1, \dots, x_n) \in (\mathbb{F}_{2^m})^n$ and $g(X) \in \mathbb{F}_{2^m}[X]$ irreducible of degree t
 - ▶ $m =$ **extension degree**, $n \lesssim 2^m$
 - ▶ $t =$ **Goppa polynomial degree** = **error-correcting capability**
- ▶ construct the generalized Vandermonde matrix

$$\mathbf{H}_{\text{priv}} = \begin{pmatrix} 1/g(x_1) & 1/g(x_2) & \dots & 1/g(x_n) \\ x_1/g(x_1) & x_2/g(x_2) & \dots & x_n/g(x_n) \\ \vdots & \vdots & & \vdots \\ x_1^{t-1}/g(x_1) & x_2^{t-1}/g(x_2) & \dots & x_n^{t-1}/g(x_n) \end{pmatrix} \in (\mathbb{F}_{2^m})^{t \times n}$$

- ▶ identify $\mathbb{F}_{2^m} \simeq (\mathbb{F}_2)^m$ (columns), put in reduced row echelon form, get

$$\mathbf{H}_{\text{pub}} \in (\mathbb{F}_2)^{mt \times n}$$

- ▶ **Goppa structure recovery**: find some $(\mathbf{x}, g)$ that give $\mathbf{H}_{\text{pub}}$
- ▶ **Goppa distinguisher**: decide if a given $\mathbf{H}$ comes from some $(\mathbf{x}, g)$

Quadratic relations

Recall from the preliminaries

- ▶ $C \subseteq \mathbb{F}^n$ with basis $\mathbf{c}_1, \dots, \mathbf{c}_k$
- ▶ evaluation map: $\text{ev}: \mathbb{F}[X_1, \dots, X_k] \rightarrow \mathbb{F}^n, X_i \mapsto \mathbf{c}_i$
- ▶ space of quadratic relations: $I_2(C) = \ker(\text{ev}_2: \mathbb{F}[X_1, \dots, X_k]_2 \rightarrow C^{\langle 2 \rangle})$

Example

- ▶ $C = \text{rowspan}_{\mathbb{F}_{2^m}}(\mathbf{H}_{\text{priv}}) = \text{GRS}_t(\mathbf{x}, \mathbf{y})$ where $\mathbf{y} = g(\mathbf{x})^{-1}$
- ▶ basis $\mathbf{c}_i = \mathbf{y}\mathbf{x}^{i-1}$ for $1 \leq i \leq t$ in geometric progression

then for $a + d = b + c$:

$$\mathbf{c}_a \mathbf{c}_d = \mathbf{c}_b \mathbf{c}_c$$

$$X_a X_d - X_b X_c \in I_2(C)$$

The [FGOPT11] distinguisher (a.k.a. “square distinguisher”)

Theorem

Let $C = \text{rowspan}_{\mathbb{F}_2}(\mathbf{H}_{\text{pub}})$ be a *dual Goppa code*. Then we have a lower bound

$$\dim_{\mathbb{F}_2}(I_2(C)) \geq T$$

for some explicit $T = T(m, n, t)$.

On the other hand for a *random* $[n, k]_2$ -code (where $k = mt$) w.h.p. [CCMZ15]

$$\dim_{\mathbb{F}_2}(I_2(C)) = \left(\frac{k(k+1)}{2} - n \right)^+$$

→ can distinguish when: $n > \frac{k(k+1)}{2} - T$

or equivalently ([MP12] reinterpretation) when: $C^{\langle 2 \rangle} \subsetneq \mathbb{F}^n$

very restrictive!

When these conditions fail, the “structural” quadratic relations *are still there*, but they are diluted in a larger space. Can we still detect them?

The syzygy distinguisher

Linear syzygies

The syzygy distinguisher is one generalization (among at least 3) of [FGOPT11].

Definition

If $\mathcal{Q} = \langle Q_1, \dots, Q_M \rangle$ is a space of **quadratic forms**, a linear syzygy on $\mathcal{Q}$ is a **tuple of linear forms** $R = (\ell_1, \dots, \ell_M)$ such that

$$\ell_1 Q_1 + \dots + \ell_M Q_M = 0.$$

Definition

If $\mathcal{R} = \langle R_1, \dots, R_N \rangle$ is a space of **tuples of linear forms**, a linear syzygy on $\mathcal{R}$ is a **tuple of linear forms** $S = (\ell_1, \dots, \ell_N)$ such that

$$\ell_1 R_1 + \dots + \ell_N R_N = 0.$$

This last construction iterates!

Effective computation: kernel of **Macaulay matrix** on the previous kernel...

High-level overview of the distinguisher:

- ▶ Start from $\mathcal{Q} = I_2(\mathbb{C})$, and at some point expect to get more iterated linear syzygies for algebraic codes than for random codes.
- ▶ A crucial ingredient to make everything work better will be **shortening**.

Dual Goppa codes have very high syzygies (+ shortening)

Let $C = \text{rowspan}_{\mathbb{F}_{2^m}}(\mathbf{H}_{\text{pub}})$ be a dual Goppa code with scalars extended.

- ▶ In my Eurocrypt paper I constructed an explicit $2 \times f$ matrix of linear forms

$$\Phi = \begin{pmatrix} \ell_1 & \ell_2 & \cdots & \ell_f \\ \ell'_1 & \ell'_2 & \cdots & \ell'_f \end{pmatrix} \text{ such that } l_2(C) \text{ contains the } 2 \times 2 \text{ minors of } \Phi$$

- ▶ with a technical argument I showed that this also holds for s -shortenings of C , with a matrix Φ_s of size $2 \times (f - s)$
- ▶ the Eagon-Northcott complex of $\Phi_s : F = S^{f-s} \longrightarrow G = S^2$

$$0 \rightarrow (S^{f-s-2} G)^\vee \otimes \bigwedge^{f-s} F \rightarrow \cdots \rightarrow G^\vee \otimes \bigwedge^3 F \rightarrow \bigwedge^2 F \xrightarrow{\wedge^2 \Phi} \bigwedge^2 G \simeq S$$

then shows that these shortenings admit syzygies up to degree $f - s$

- ▶ moreover this explicit f is very close to $k = mt$, namely $f \approx \left(1 - \frac{\log \log n}{\log n}\right) k$.

Are we happy with that? Actually, this f is still not optimal, and moreover, the proof of the shortening property looks unnecessarily complicated.

Reinterpretation with length 1 progressions

Idea: $I_2(C)$ contains the 2×2 minors of $\Phi \leftrightarrow 2 \times 2$ minors of $\text{ev}(\Phi)$ vanish ... and this is ensured when $\text{ev}(\Phi)$ has “rank 1” i.e. its rows are proportional

- ▶ Let $\mathbf{x}$ be a support vector of C , and recall $C_1 = \{\mathbf{c} \in C : \mathbf{x}\mathbf{c} \in C\}$
- ▶ set $f = \dim C_1$ and $\mathbf{c}_1, \dots, \mathbf{c}_f$ basis of C_1
- ▶ let the linear forms $\ell_1, \dots, \ell_f$ evaluate to $\mathbf{c}_1, \dots, \mathbf{c}_f$ and $\ell'_1, \dots, \ell'_f$ evaluate to $\mathbf{x}\mathbf{c}_1, \dots, \mathbf{x}\mathbf{c}_f$
- ▶ then take $\Phi = \begin{pmatrix} \ell_1 & \ell_2 & \dots & \ell_f \\ \ell'_1 & \ell'_2 & \dots & \ell'_f \end{pmatrix}$, so $\text{ev}(\Phi) = \begin{pmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_f \\ \mathbf{x}\mathbf{c}_1 & \mathbf{x}\mathbf{c}_2 & \dots & \mathbf{x}\mathbf{c}_f \end{pmatrix}$

Theorem 1

$$\dim C_1 \geq (a+2)t - 2^a \quad \text{with } a = \lceil \log_2 t \rceil.$$

Experiments show that this lower bound is an equality w.h.p., and that this new f is **optimal** i.e. it gives the degree of the highest nonzero syzygy space of C .

Slight disappointment: **asymptotics unchanged** $f \approx \left(1 - \frac{\log \log n}{\log n}\right) k$.

Bonus: the **shortening property** follows readily from $\text{Short}(C)_1 = \text{Short}(C_1)$.

Remark: one long progression $\rightarrow$ many length 1 progressions

1. C contains $\text{GRS}_t(\mathbf{x}, \mathbf{y}) \rightarrow$ readily have Φ with $\text{ev}(\Phi) = \begin{pmatrix} \mathbf{y} & \mathbf{y}\mathbf{x} & \dots & \mathbf{y}\mathbf{x}^{t-2} \\ \mathbf{y}\mathbf{x} & \mathbf{y}\mathbf{x}^2 & \dots & \mathbf{y}\mathbf{x}^{t-1} \end{pmatrix}$
but it has $t - 1$ columns, which is only a **fraction** of $k = mt$

2. C also contains all the Frobenius conjugates $\text{GRS}_t(\mathbf{x}^{1/2^i}, \mathbf{y}^{1/2^i}) \rightarrow$ get Φ with

$$\text{ev}(\Phi) = \left(\begin{array}{cccc} \mathbf{y} & \mathbf{y}\mathbf{x} & \dots & \mathbf{y}\mathbf{x}^{t-2} \\ \mathbf{y}\mathbf{x} & \mathbf{y}\mathbf{x}^2 & \dots & \mathbf{y}\mathbf{x}^{t-1} \end{array} \middle| \begin{array}{cccc} \mathbf{y}^{1/2} & \mathbf{y}^{1/2}\mathbf{x}^{1/2} & \dots & \mathbf{y}^{1/2}(\mathbf{x}^{1/2})^{t-3} \\ \mathbf{y}^{1/2}(\mathbf{x}^{1/2})^2 & \mathbf{y}^{1/2}(\mathbf{x}^{1/2})^3 & \dots & \mathbf{y}^{1/2}(\mathbf{x}^{1/2})^{t-1} \end{array} \right. \\ \left. \begin{array}{cccc} \mathbf{y}^{1/4} & \mathbf{y}^{1/4}\mathbf{x}^{1/4} & \dots & \mathbf{y}^{1/4}(\mathbf{x}^{1/4})^{t-5} \\ \mathbf{y}^{1/4}(\mathbf{x}^{1/4})^4 & \mathbf{y}^{1/4}(\mathbf{x}^{1/4})^5 & \dots & \mathbf{y}^{1/4}(\mathbf{x}^{1/4})^{t-1} \end{array} \middle| \dots \right)$$

so its first row gives explicit elements in C_1 (but not a whole basis)

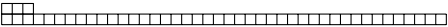
3. Theorem 1 reaches $\dim C_1$ but is **not constructive**: its proof characterizes (a space of rational fractions whose univariate evaluation at $\mathbf{x}$ gives) C_1 as a kernel and then use the rank-nullity theorem.

The distinguisher

- ▶ For random codes, [CCMZ15] says $l_2(C) = 0$ when $\frac{k(k+1)}{n} \leq 2$
- ▶ generalization: under technical hypotheses (that will be met), we heuristically do not expect syzygies of degree higher than $\frac{k(k+1)}{n}$.
- ▶ Now we shorten s times, as long as

$$f - s > \frac{(k - s)(k - s + 1)}{n - s}$$

→ shortened dual Goppa codes admit syzygies of degree $f - s$, while corresponding random codes do not

- ▶ $f = \dim C_1$ is very close to k → can shorten a lot → end up with shortened code of dimension $k_s \approx \frac{\log \log n}{\log n} k$ and rate $R_s = k_s/n_s \approx \frac{R}{1-R} \frac{\log \log n}{\log n}$
- ▶ diagram of the filtration after shortening: 
- ▶ complexity of syzygy computation is polynomial in $\binom{k_s}{R_s k_s} = 2^{\Theta(\frac{(\log \log n)^3}{(\log n)^2} n)}$
 → **subexponential** in $\frac{n}{\log n} = \Theta(t)$, while generic decoding is exponential.

Further remarks

Our invariants actually are the **graded Betti numbers** $\beta_{r-1,r}$ where more generally

$$\beta_{i,j} = \dim_{\mathbb{F}} \operatorname{Tor}_i^S(C^{\langle \cdot \rangle}, \mathbb{F})_j$$

with

- ▶ $S = \mathbb{F}[X_1, \dots, X_k]$ homogeneous coordinate ring of $\mathbb{P}^{k-1}$
- ▶ $C^{\langle \cdot \rangle} = \bigoplus_{u \geq 0} C^{\langle u \rangle}$ homogeneous coordinate ring of $\{\bar{\mathbf{p}}_1, \dots, \bar{\mathbf{p}}_n\}$
- ▶ Tor = derived functor of $\otimes$ → use a resolution on the left or on the right.

1. Resolving $C^{\langle \cdot \rangle}$ gives the “higher syzygy” interpretation
2. alternatively, resolving $\mathbb{F}$ computes $\beta_{r-1,r}$ through **Koszul cohomology**:

$$\bigwedge^r S_1 \longrightarrow \bigwedge^{r-1} S_1 \otimes C \longrightarrow \bigwedge^{r-2} S_1 \otimes C^{\langle 2 \rangle}$$

more direct, slightly more efficient, same “expressive power”, e.g.:

- ▶ **Lars’ suggestion**: for $\mathbf{c}_1, \dots, \mathbf{c}_r \in C_1$, look at

$$\sum_{i=1}^r (-1)^i \ell_1 \wedge \dots \wedge \widehat{\ell_i} \wedge \dots \wedge \ell_r \otimes \mathbf{x} \mathbf{c}_i$$

- ▶ M. Boutros: alternative proof of $\beta_{r-1,r} > 0$ for $r = k + 1 - d_{\min}$

Key recovery

(joint work with P. Briaud, A. Lemoine, and J.-P. Tillich)

The Pfaffian distinguisher [CMT23]

- ▶ I'm somehow a free-rider in this work so my point of view might differ from that of my co-authors. Anyway, all started with the Pfaffian distinguisher, another generalization (among at least 3) of [FGOPT11].
- ▶ Idea: even when [FGOPT11] fails, $I_2(C)$ is not a generic space of quadratic forms: it contains **short** quadratic relations (e.g. 2×2 minors).
- ▶ **Detecting** their existence readily gives a distinguisher, and with luck, **finding** these short relations could even give structural recovery.
- ▶ That was also my plan when I first got interested in McEliece (~ 2021 -2022). But I wanted to measure shortness with the Chow/decomposable/split rank of a quadratic form, which was not very practical. I failed **miserably**.
- ▶ Instead [CMT23] passed to the associated polar form and used the usual matrix rank.
- ▶ Moreover they observed that in char 2, these polar forms are skew-symmetric $\rightarrow$ their rank can be characterized by the vanishing of Pfaffians.
- ▶ I was **scandalized** by what I perceived as a lack of canonicity...
- ▶ ... but it worked! how could it be?

Reworking the Pfaffian approach

- ▶ First one can rephrase:

quadratic form $\rightarrow$ polar form $\rightarrow$ skew-symmetric matrix

as:

in char 2, $\bigwedge^2 L$ is a quotient of $S^2 L$

(nice, but looks like an accident — I was sure **never** to hear about it again)

- ▶ [CMT23] take the image of $I_2(C) \subseteq \mathbb{F}[\mathbf{X}]_2 = S^2 \mathbb{F}[\mathbf{X}]_1$ in $\bigwedge^2 \mathbb{F}[\mathbf{X}]_1$
- ▶ rank is minimized by quadratic forms like $UW - V^2$, sent to the decomposable 2-form $U \wedge W$
- ▶ $UW - V^2 \in I_2(C) \leftrightarrow$ quadratic relation $\mathbf{u}\mathbf{w} = \mathbf{v}^2$ between codewords, i.e. $\mathbf{u}, \mathbf{v}, \mathbf{w}$ form a length 2 geometric progression
- ▶ $U \wedge W$ does not involve $V \leftrightarrow$ this **has to** be expressible in $\mathbf{u}, \mathbf{w}$ only
- ▶ use that C is stable under Frobenius! we want:

$$\mathbf{u}, \mathbf{w} \in C \quad \text{such that} \quad \mathbf{u}\mathbf{w} \in C$$

and then they determine $\mathbf{v} = (\mathbf{u}\mathbf{w})^{1/2}$.

Comparison with [BLT26]

We want

$$\mathbf{u}, \mathbf{w} \in C \quad \text{such that} \quad \mathbf{uw} \in C.$$

This gives a quadratic system (actually bilinear), whose number of variables is $\approx 2 \dim C$.

[BLT26] remain closer to [CMT23] and look for $\lambda = U \wedge W \in \bigwedge^2 \mathbb{F}[\mathbf{X}]_1$:

- ▶ in the image of $I_2(C) \rightarrow$ linear equations in the Plücker coordinates
- ▶ of rank 1 $\rightarrow$ quadratic equations (Plücker relations).

This is **equivalent**, however the number of independent variables left after the linear equations is $\dim I_2(C)$, **quadratic in $\dim C$** .

The Plücker embedding is a powerful tool to work with subspaces... but **sometimes** directly working with a basis can be more effective.

However, if you insist on a neat characterization of subcodes $D = \langle \mathbf{u}, \mathbf{w} \rangle$ associated with length 2 geometric progressions, without choosing a basis, here is one: they are the 2-dimensional $D \subseteq C$ such that $D^{(2)} \subseteq C$.

Constraining the system: (1) shortening

Two problems:

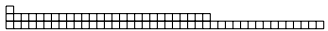
1. in life we **always** have too many variables; please remove as many as you can!
2. we aim at geometric progressions $\mathbf{u}, \mathbf{x}\mathbf{u}, \mathbf{x}^2\mathbf{u}$ with ratio $\mathbf{x} = a$ defining support vector for $C \rightarrow$ directly get $\mathbf{x}$, and easily derive a Goppa polynomial; however there are others, e.g. $\mathbf{u}, f(\mathbf{x})\mathbf{u}, f(\mathbf{x})^2\mathbf{u}$; please remove them!

Theorem 2

Let C be a dual Goppa code with scalars extended, and $\mathbf{x}$ a defining support vector for C . Then the second subspace in its filtration has dimension

$$\dim C_2 \geq (a+1)t - 2^a \quad \text{with } a = \lceil \log_2 t \rceil.$$

Experiments suggest this is an equality w.h.p.

We shorten $s = \dim C_2 - 1$ times $\rightarrow$ filtration diagram  $\rightarrow$ for each defining support vector $\mathbf{x}$ there remains **only one** length 2 progression. For other ratios we expect C_2 smaller $\rightarrow$ all length 2 progressions killed.

Asymptotics: $s = (1 - \frac{\log \log n}{\log n})k$ again (where $k = mt$)

Constraining the system: (2) fixing coordinates

Update notation: $C = s$ -shortened code, $\mathbf{H} = (\mathbf{1}_{mt-s} | \mathbf{H}') \in \mathbb{F}_2^{mt-s \times n-s}$ a generator matrix in systematic form, $\mathbf{G} \in \mathbb{F}_2^{n-mt \times n-s}$ a parity matrix

► $\mathbf{u}, \mathbf{w} \in C \iff \mathbf{u} = \mathbf{aH}, \mathbf{w} = \mathbf{bH}$ for $\mathbf{a}, \mathbf{b} \in (\mathbb{F}_{2^m})^{mt-s}$

Two support vectors are related by homography and Frobenius: $\mathbf{x}' = \left(\frac{\alpha\mathbf{x} + \beta}{\gamma\mathbf{x} + \delta} \right)^{2^e}$.
Conversely [BMT23] we can force $\mathbf{x} = (0, \infty, 1, *, \dots, *)$ using the **Cauchy code trick** $\rightarrow$ fix the homography, only remain the m choices of Frobenius.

1. We can fix $\mathbf{a} = (1, 0, *, \dots, *)$ and $\mathbf{b} = (0, 1, *, \dots, *) \rightarrow$
 $\mathbf{u} = (1, 0, *, \dots, *)$ and $\mathbf{w} = (0, 1, *, \dots, *)$ linearly independent, and
 $\mathbf{x} = (0, \infty, *, \dots, *)$
2. fix $a_3 = b_3 \rightarrow u_3 = w_3 \rightarrow x_3 = 1$ (rem: $a_3 = b_3 \neq 0$ otherwise shorten)

$\rightarrow$ the system has $2(mt - s) - 5$ variables

► $\mathbf{uw} \in C \iff (\mathbf{uw})\mathbf{G}^\top = \mathbf{0}$ in $(\mathbb{F}_{2^m})^{n-mt}$

$\rightarrow$ it has $n - mt$ equations

Two final remarks on this attack

(1) Solving the shortened system, we recover only a **punctured** $\mathbf{x}$.

However we can restart with another set of shortened positions, while keeping positions $\{1, 2, 3, 4\}$ in the support and fixing one more coordinate to ensure the new $\mathbf{x}'$ is of the form $(0, \infty, 1, x_4, *, \dots, *)$.

This ensures $\mathbf{x}$ and $\mathbf{x}'$ coincide on the intersection of their supports (it fixes the homography and the Frobenius) hence they can be glued.

We have $s \leq k$ hence $n - s = \Theta(n)$, so we only need $\Theta(1)$ iterations to recover the full $\mathbf{x}$.

(2) By very construction, our bilinear map

$$(\mathbf{a}, \mathbf{b}) \mapsto ((\mathbf{a}\mathbf{H})(\mathbf{b}\mathbf{H}))\mathbf{G}^\top$$

is alternating.

This implies that our bilinear system is actually linear in the minors $a_i b_j - a_j b_i$.

(Alternative proof: expand $((\sum_i a_i \mathbf{h}_i)(\sum_i b_i \mathbf{h}_i))\mathbf{G}^\top$ and use the fact that the rows $\mathbf{h}_i$ of $\mathbf{H}$ satisfy $\mathbf{h}_i^2 = \mathbf{h}_i$)

As explained in Pierre's talk, this is an essential starting point for the analysis of the system.

Further considerations

General result on the filtrations

Let C be a dual Goppa code with scalars extended, of extension degree m and Goppa polynomial degree t .

Let $\mathbf{x}$ be a defining support vector for C , and $Z \in \mathbb{F}_{2^m}(X)$ a univariate rational function of degree d

Let then $\mathbf{z} \in (\mathbb{F}_{2^m}^\times)^n$ coincide with $Z(\mathbf{x})$ except at a set of e positions, and up to some Frobenius.

Theorem 3

The filtration $C = C_0 \supseteq C_1 \supseteq C_2 \supseteq \dots$ associated to $\mathbf{z}$ satisfies

$$\dim C_\ell \geq (a+2)t - \ell d 2^a \quad \text{with } a = \lceil \log_2 \frac{t}{\ell d} \rceil.$$

Most of the time this is tight. Examples where the inequality is strict can be constructed, but I consider they are rare.

The higher order vanishing distinguisher and attack

The higher order vanishing distinguisher [HW25] and the higher order vanishing attack [Hemmert26] rely on the fact that **catalecticant matrices** of the form

$$\begin{pmatrix} \ell_0 & \ell_1 & \ell_2 & \cdots & \ell_p \\ \ell_1 & \ell_2 & \ddots & \ell_p & \vdots \\ \ell_2 & \ddots & \ell_p & \ddots & \vdots \\ \vdots & \ell_p & \ddots & \ddots & \vdots \\ \ell_p & \cdots & \cdots & \cdots & \ell_{2p} \end{pmatrix}$$

come from geometric progressions of length $2p$ ($\rightarrow$ exist when $\dim C_{2p} \geq 1$) and have determinant vanishing at order $p - 1$ on $\mathbf{H}$.

This could help understand the complexity.

Actually I guess (did not check all details) that this higher order vanishing property only comes from the fact that a progression of length $2p$ induces p progressions of length p ($\rightarrow \dim C_p \geq p$). If correct, improve complexity.

Bak to the syzygy distinguisher: what remains of it?

Two main ingredients:

1. the cool math stuff (Betti numbers, syzygies, etc.) are very specialized tools
→ did not catch that much (yet?)
2. determinantal quadratic relations, compatibility with shortening (a lot!) → very versatile.

The theory presented here reinterpret and generalize point 2. in an elementary and handy way.

- ▶ In the end, the syzygy distinguisher uses Betti numbers only to detect the existence of a large subcode C_1 such that $(C : C_1)$ is nontrivial (they can do other things, but this is really what we use them for)
- ▶ but we can imagine plenty of other approaches to detect this C_1 , or even to find it → key-recovery of $\mathbf{x} \in (C : C_1)$
- ▶ actually I've been working for some time on a quantum algorithm for that (still not finalized); this was my primary motivation for not estimating only the size of C_1 (Theorem 1) but establishing the full theory (Theorem 3)
- ▶ nothing is published, but I promise I'll finish the job... even if it will not be that impressive anymore now that 2026/1630 has appeared.