

Properties of the Quadratic Relations Code and Its Application to the Decoding Problem of Linear Codes

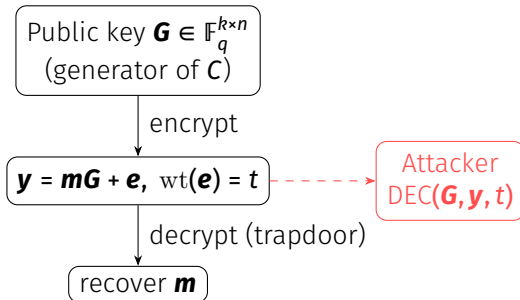
Ivan Chizhov

June 4, 2026

1. Introduction and motivation
2. The code of quadratic relationships
3. Existence of reducible forms in code of quadratic relations
4. Application: a decoding algorithm
5. Conclusions

[1/5] Introduction and motivation

McEliece scheme and the basic objects



Trapdoor = hidden algebraic structure of $\mathcal{C}$.

Notation

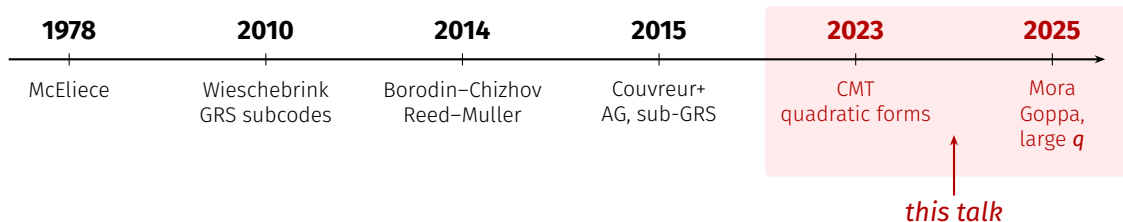
- ▶ $[n, k]_q$ -code $\mathcal{C} \subseteq \mathbb{F}_q^n$;
- ▶ $\mathbf{G}$: generator;
- ▶ $\text{wt}(\mathbf{x})$: Hamming weight
- ▶ $\text{DEC}(\mathbf{G}, \mathbf{y}, t)$: find $\mathbf{c} \in \mathcal{C}$ with $\text{wt}(\mathbf{y} - \mathbf{c}) = t$

Schur–Hadamard product: 15 years of structural attacks

- ▶ Componentwise product $\mathcal{C} \circ \mathcal{C} = \langle \mathbf{c} \circ \mathbf{c}' : \mathbf{c}, \mathbf{c}' \in \mathcal{C} \subseteq \mathbb{F}_q^n \rangle$, where

$$\mathbf{c} \circ \mathbf{c}' \stackrel{\text{def}}{=} (c_1 \cdot c'_1, \dots, c_n \cdot c'_n)$$

- ▶ Random codes saturate $\dim \mathcal{C}^2 \approx \min(n, k(k+1)/2)$; structured codes do not — exploitable gap



From Schur product to quadratic relationships (2023)

- ▶ Couvreur–Mora–Tillich, ASIACRYPT 2023:
code of quadratic relationships Q_G
- ▶ Mora, AMC 2025:
attack on Goppa-McEliece for large q
- ▶ **Key ingredient**: existence of a *low-rank*
quadratic form in $\ker \text{eval}_G^{(2)}$
- ▶ **Our question**: when do such forms exist?

A new approach based on quadratic forms to attack the McEliece cryptosystem

Alain Couvreur¹, Rocco Mora², and Jean-Pierre Tillich²

¹ Inria Saclay, LIX, CNRS UMR 7161, École Polytechnique, 1 rue Honoré d'Estienne d'Orves, 91120 Palaiseau Cedex

² Inria Paris, 2 rue Simone Iff, 75012 Paris, France
{alain.couvreur, rocco.mora, jean-pierre.tillich}@inria.fr

Abstract. We introduce a novel algebraic approach for attacking the McEliece cryptosystem which is currently at the 4-th round of the NIST competition. The contributions of the article are twofold. (1) We present a new distinguisher on alternant and Goppa codes working in a much broader range of parameters than [FGO⁺11]. (2) With this approach we also provide a polynomial-time key recovery attack on alternant codes which are distinguishable with the distinguisher [FGO⁺11].

[2/5] The code of quadratic relationships

Evaluation operator $\text{eval}_{\mathbf{G}}$

- ▶ $\mathbf{G} \in \mathbb{F}_q^{k \times n}$ with columns $\mathbf{g}_1^T, \dots, \mathbf{g}_n^T \in \mathbb{F}_q^k$.
- ▶ $\mathbb{F}_q[\mathbf{x}] \stackrel{\text{def}}{=} \mathbb{F}_q[x_1, x_2, \dots, x_n]$ — polynomial ring.
- ▶ $\mathbb{F}_q[\mathbf{x}]^{(d)}$ — vector space of homogeneous polynomials degree d .
- ▶ $\text{eval}_{\mathbf{G}} : \mathbb{F}_q[\mathbf{x}] \rightarrow \mathbb{F}_q^n$; $\text{eval}_{\mathbf{G}}(f) = (f(\mathbf{g}_1), \dots, f(\mathbf{g}_n)) \in \mathbb{F}_q^n$.
- ▶ Two useful restrictions:
 - ▶ $\text{eval}_{\mathbf{G}}^{(1)} : \mathbb{F}_q[\mathbf{x}]^{(1)} \rightarrow \mathbb{F}_q^n$ — bijection onto $\mathbb{C}$; $\text{Im } \text{eval}_{\mathbf{G}}^{(1)} = \mathbb{C}$.
 - ▶ $\text{eval}_{\mathbf{G}}^{(2)} : \mathbb{F}_q[\mathbf{x}]^{(2)} \rightarrow \mathbb{F}_q^n$ — image is $\mathbb{C}^2$

Kernel and the code Q_G

- ▶ $\ker \text{eval}_G^{(2)}$: quadratic forms vanishing on all columns of G
- ▶ Depends on G , unlike the image
- ▶ Q_G = coefficient vectors of these forms a $[k(k+1)/2, k(k+1)/2 - \dim C^2]_q$ -code (Couvreur–Mora–Tillich, 2023)

When is the kernel nontrivial?

$\dim C^2 \leq \min(n, k(k+1)/2)$, so

$$\dim \ker \text{eval}_G^{(2)} \geq \frac{k(k+1)}{2} - n.$$

For $n < k(k+1)/2$ the kernel is automatically nontrivial — but its *rank structure* is not (*random case*).

Matrix view: forms as symmetric matrices

For $Q = \sum_{i \leq j} a_{ij} x_i x_j$, set $\mathbf{M}_Q = (a_{ij}) \in \mathbb{F}_q^{k \times k}$ (upper-triangular) and the *symmetric matrix*

$$\mathbf{B}_Q = \mathbf{M}_Q + \mathbf{M}_Q^T.$$

- ▶ char $\neq 2$: $\mathbf{B}_Q$ is the matrix of the associated bilinear form
- ▶ rank $\mathbf{B}_Q$: a *geometric* invariant of Q ;
- ▶ Small rank $\Leftrightarrow Q$ has algebraic structure.

Matrix code of quadratic relationships (CMT 2023)

$$M_{\mathbf{G}} = \{ \mathbf{B}_Q : Q \in \ker \text{eval}_{\mathbf{G}}^{(2)} \} \subseteq \text{Sym}_k(\mathbb{F}_q)$$

Cryptanalysis question: does $M_{\mathbf{G}}$ contain a **low-rank** matrix?

Low-rank forms: a cryptanalytic ladder

rank $\mathbf{B}_Q$	Algebraic form	Geometric meaning	Cryptanalytic use
1	$Q = L^2$	perfect square	not directly used
2	$Q = L \cdot K$	reducible	this talk
3–4	Q irreducible	low geometric rank	Mora 2025 (Goppa, large q)
high	generic	random-like	no known attack

- ▶ Each rank tier needs different techniques (decoder, structural recovery, ...)
- ▶ This talk: tier-2 — when does $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contain reducible forms, and how to exploit them?

1. **Characterize** codes whose $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contains reducible forms.
2. **Find parameter regimes** where this is guaranteed (or forbidden).
3. **Exploit** (if possible) the existence of such forms to design a decoding algorithm; benchmark against the ISD family.

[3/5] Existence of reducible forms in code of quadratic relations

A combinatorial criterion via information sets

$\mathbf{G}|_I$: columns of $\mathbf{G}$ indexed by $I \subseteq \llbracket n \rrbracket = \{1, 2, \dots, n\}$.

Recall: $I \subseteq \llbracket n \rrbracket$ is an **information set** if $|I| = k$ and $\mathbf{G}|_I$ is invertible.

Theorem

Let $\mathbf{G}$ generate C . Then $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contains **no** reducible form iff the support of every nonzero codeword of C contains an information set.

- ▶ Algebraic question $\iff$ combinatorial question on supports
- ▶ All results in this section are corollaries of this criterion, applied in different regimes

Sufficient condition via low-weight codewords

Theorem

If C has a codeword of weight $< k$, then $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contains a reducible form.

- ▶ *Proof idea: a support of size $< k$ cannot contain a k -element information set; apply the criterion.*
- ▶ *Covers most code families used in practice (Goppa, Reed–Solomon with $n > k$, Reed–Muller, ...)*

Definition

- ▶ If $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{F}_q^n$, then $\text{supp}(\mathbf{x}) = \{i \in \llbracket n \rrbracket \mid x_i \neq 0\}$ — support of $\mathbf{x}$.
- ▶ $C \subseteq \mathbb{F}_q^n$ is **minimal** if for any codewords $\mathbf{c}, \mathbf{b} \in C$ holds $\text{supp}(\mathbf{b}) \subseteq \text{supp}(\mathbf{c}) \Rightarrow \mathbf{b} = a \cdot \mathbf{c}$ for some $a \in \mathbb{F}_q$.

The binary case: equivalence with non-minimal codes

Theorem

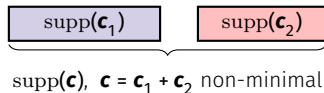
For full-rank $\mathbf{G} \in \mathbb{F}_2^{k \times n}$: $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contains no reducible form iff C is a minimal binary code.

For any q : if $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contains a reducible form then C is not minimal.

Proof idea (Ashikhmin–Barg lemma): a non-minimal $\mathbf{c}$ splits as

$$\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2, \quad \text{supp}(\mathbf{c}_1) \cap \text{supp}(\mathbf{c}_2) = \emptyset.$$

Pulling back via $\text{eval}_{\mathbf{G}}^{(1)}$ gives nonzero L_1, L_2 with $L_1 \cdot L_2 \in \ker \text{eval}_{\mathbf{G}}^{(2)}$.



Parametric guarantee: $n < 2k - 1$

Singleton bound: $d \leq n - k + 1$ for any $[n, k]_q$ -code.

Combine the previous corollary with the Singleton bound:

Theorem

*If $n < 2k - 1$, then for **any** $[n, k]_q$ -code C and any generator matrix $\mathbf{G}$, $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contains a reducible form.*

- ▶ *Proof idea: Singleton gives $d \leq n - k + 1 < k$, so a weight- $< k$ codeword exists; previous corollary applies.*
- ▶ ***Compared with CMT 2023:***

*CMT 2023: rank-2 form with **non-negligible probability** for random codes*

*This work: reducible form with **probability 1** for every code*

MDS codes: no reducible forms

Theorem (Theorem)

If C is MDS, then $\ker \text{eval}_{\mathbf{G}}^{(2)}$ contains **no** reducible form.

Proof idea: MDS gives $d^\perp \geq k + 1$; this forces every kernel form to be irreducible.

Code class	Reducible forms in $\ker \text{eval}_{\mathbf{G}}^{(2)}$?
$d < k$	yes
$n < 2k - 1$ (any code)	yes
binary, non-minimal	yes
binary, minimal	no
MDS	no

Remark about minimal codewords

- ▶ Ashikhmin, A. and Barg, A. Minimal vectors in linear codes, IEEE Transactions on Information Theory : Institute of Electrical and Electronics Engineers (IEEE), 1998, vol. 44, no. 5, pp. 2010–2017.
- ▶ $\mathcal{M}$ — set of minimal codewords.

Corollary 2.5: Let $n \rightarrow \infty$, $k = Rn$, $0 < R < 1$. Then

$$\frac{1}{n} \log_q \mathbb{E}|\mathcal{M}| = \begin{cases} H_q(1-R) - (1-R), & 0 < 1-R < \frac{q-1}{q} \\ R, & \frac{q-1}{q} \leq 1-R < 1. \end{cases}$$

Here $H_q(\cdot)$ is the entropy function.

- ▶ If $2k \leq n$ then in random $[n, k]_q$ -code almost all codewords are minimal.

Space of divisors $Z_{\mathbf{G}}(L)$

For a nonzero linear form $L \in \mathbb{F}_q[\mathbf{x}]^{(1)}$, the space of divisors of L in the kernel:

$$Z_{\mathbf{G}}(L) = \{ K \in \mathbb{F}_q[\mathbf{x}]^{(1)} \mid L \cdot K \in \ker \text{eval}_{\mathbf{G}}^{(2)} \}.$$

- ▶ $Z_{\mathbf{G}}(L)$ is a *linear subspace* of $\mathbb{F}_q[\mathbf{x}]^{(1)}$;
- ▶ $\dim Z_{\mathbf{G}}(L) \geq 1 \Leftrightarrow L$ divides *some* nonzero kernel form
 $\Leftrightarrow \text{supp}(\mathbf{c})$ does not contain an information set ($\mathbf{c} = \text{eval}_{\mathbf{G}}^{(1)}(L)$)

Isomorphism with a shortened code

Let $T = \text{supp}(\mathbf{c})$. Then $Z_{\mathbf{G}}(L) \cong C^T$ (the code C shortened on T), and

$$\dim Z_{\mathbf{G}}(L) = \dim C^T = k - \text{rank } \mathbf{G}|_T.$$

Large divisor space and low-weight words

Theorem (Wei's generalized Hamming weights, 1991)

If $\dim Z_{\mathbf{G}}(L) \geq r$, then $\text{wt}(\mathbf{c}) \leq n - d_r(C)$,

where $\mathbf{c} = \text{eval}_{\mathbf{G}}^{(1)}(L)$ and $d_r(C) = \min\{|\text{supp } D| : D \subseteq C, \dim D = r\}$.

- ▶ *Proof idea:* $Z_{\mathbf{G}}(L) \cong C^T$ gives an r -dim subcode of C^T ; Wei's bound forces its support $\geq d_r$; combine with $|\text{supp } C^T| \leq n - |T|$ and $|T| = \text{wt}(\mathbf{c})$.
- ▶ *Special case* $r = 1$: $\text{wt}(\mathbf{c}) \leq n - d$ — every L dividing any kernel form gives a codeword bounded by the minimum distance.
- ▶ *Take-away:* the more kernel forms L divides, the **lighter** the codeword $\mathbf{c}$ becomes.

Cryptanalytic consequence

Rank-2 forms in $\ker \text{eval}_{\mathbf{G}}^{(2)}$ are **hard to find in general**: they force L to give a low-weight $\mathbf{c}$, and low-weight codewords are themselves rare.

[4/5] Application: a decoding algorithm

Reduction of decoding to factor search

Input: $\mathbf{G} \in \mathbb{F}_q^{k \times n}$, $\mathbf{y} \in \mathbb{F}_q^n$, $t < k$

Extended matrix $(\mathbf{G}; \mathbf{y}) \in \mathbb{F}_q^{(k+1) \times n}$

$(k+1)$ rows, $t+1 < k+1$

Existence theorem $\Rightarrow \exists L, K \neq 0 :$
 $L \cdot K \in \ker \text{eval}_{(\mathbf{G}; \mathbf{y})}^{(2)}, \quad \mathbf{e} = \text{eval}_{(\mathbf{G}; \mathbf{y})}^{(1)}(L)$

Decoding $\equiv$ enumerate L dividing a kernel element

- ▶ Either $\mathbf{e} = \text{eval}_{(\mathbf{G}; \mathbf{y})}^{(1)}(L)$ or $\mathbf{e} = \text{eval}_{(\mathbf{G}; \mathbf{y})}^{(1)}(K)$ — both candidates cheap to test
- ▶ One iteration $\Rightarrow$ one random information set $\Rightarrow t+1$ candidate forms

Algorithm

Input: $\mathbf{G} \in \mathbb{F}_q^{k \times n}$, $\mathbf{y} \in \mathbb{F}_q^n$, error weight t

Output: $\mathbf{e}$ with $\text{wt}(\mathbf{e}) = t$ and $\mathbf{y} - \mathbf{e} \in \mathcal{C}$

```
1 while true do
2    $I \leftarrow$  random  $k$ -subset of  $\llbracket n \rrbracket$ ;
3    $(\mathbf{R}, \mathbf{A}) \leftarrow \text{rref}(\mathbf{G}|_I)$  //  $\mathbf{R} = \mathbf{A} \cdot \mathbf{G}|_I$ ;
4   for  $i = 1, \dots, t + 1$  do
5      $L \leftarrow$  linear form with coefficients  $a_{(i)}$  (row  $i$  of  $\mathbf{A}$ );
6      $\mathbf{e} \leftarrow \text{testLF}(\mathbf{G}, \mathbf{y}, t, L)$ ;
7     if  $\mathbf{e} \neq \perp$  then return  $\mathbf{e}$ ;
8   end
9 end
```

- ▶ At most t such rows can be nonzero on $\text{supp}(\mathbf{e}) \cap I$ — first $t + 1$ must contain a useful one
- ▶ $\text{testLF}(\mathbf{G}, \mathbf{y}, t, L)$: tries to recover K such that $L \cdot K \in \ker \text{eval}_{(\mathbf{G}; \mathbf{y})}^{(2)}$ and the error has weight t

Complexity of one iteration

Theorem (Theorem)

For a fixed error vector $\mathbf{e}$ of weight t ,

$$\mathbb{E}_{\mathbf{G}}[P_{\text{success per iter.}}] \geq \binom{n}{k}^{-1} \cdot \max_{0 \leq s \leq t} q^{-(t-s)} \binom{t}{s} \binom{n-t}{k-s}.$$

- ▶ Reading the formula: s = how many of the t error positions land in I ; $\binom{t}{s} \binom{n-t}{k-s}$ counts such subsets; $q^{-(t-s)}$ is the “discount” for the remaining $t-s$ unhit positions
- ▶ Prange-style structure: same binomial machinery, with extra $q^{-(t-s)}$ from the kernel-form trick
- ▶ Hidden cost factor $q^{z(c)}$ in per-iteration work.

Probability of success per iteration

Algorithm	Parameters (q, n, k, t)		
	$(2, 1536, 640, 150)$	$(2, 2700, 1280, 320)$	$(2^8, 208, 104, 78)$
Prange	-125	-325	-114
Stern	-101	-232	-94
Dumer	-94	-225	—
Lee–Brickell	—	—	-105
BJMM / BJMM+	-58	-154	—
MayOzerov	-54	-168	—
BothMay	-62	-214	—
BallCollision	-94	-232	—
This work	-80	-149	-113

Probability of success per iteration ($\log_2$). Estimated with `CryptographicEstimator`.

Numerical comparison with the ISD family

Algorithm	Parameters (q, n, k, t)		
	(2, 1536, 640, 150)	(2, 2700, 1280, 320)	(2^8 , 208, 104, 78)
Prange	153	355	137
Stern	134	327	127
Dumer	134	327	—
Lee–Brickell	—	—	129
BJMM / BJMM+	129	306	—
MayOzerov	128	308	—
BothMay	129	314	—
BallCollision	134	327	—
This work	116	189	149

Bit-complexity ($\log_2$ of $\mathbb{F}_q$ -operations). Estimated with `CryptographicEstimator`.

- **Wins:** binary, $n \approx 2k$, moderate t/n ; **Loses:** large q , small ratio k/n , large t

[5/5] Conclusions

Structural

Reducible forms exist in $\ker \text{eval}_{\mathbf{G}}^{(2)} \iff$ some nonzero codeword's support fails to contain an information set.

Over $\mathbb{F}_2$: $\iff \mathcal{C}$ is non-minimal.

Parametric

$n < 2k - 1$: reducible form exists with **probability 1**, every code.

MDS: reducible forms never exist.

Algorithmic

Decoder beats best ISD by ≈ 12 bits on $[1536, 640]_2$ and ≈ 117 bits on $[2700, 1280]_2$.

Loses ground over large fields (lost by 22 bits on $[208, 104]_{2^8}$).

Most pressing

- ▶ Rank > 2 : efficient search for forms of small but > 2 rank (Mora 2025 uses rank up to 4)

Algorithmic

- ▶ Closing the gap for large q — is the binary structure essential, or just convenient?
- ▶ Tighter estimates of $\mathbf{z}(\mathbf{C})$ for structured code families (Goppa, RM, AG, Quadratic residue codes, self-dual and self-orthogonal codes, ...)

Cross-pollination

- ▶ Leverage the full Ashikhmin–Barg minimal-code toolbox in cryptanalysis
- ▶ ISD-like decoding for codes with high minimal distance.