The McEliece–type Cryptosystem based on D–codes

Yu. V. Kosolapov **E. A. Lelyuk**

Southern Federal University, Russia,
e-mail: yvkosolapov@sfedu.ru, lelukevgeniy@mail.ru

The 12th Workshop on Current Trends in Cryptology
(CTCrypt 2023), Volgograd, June 6-9, 2023

Content

Object of Research

Post–Quantum Cryptography

Competitions for post–quantum algorithms

- NIST PQC (USA, announced in 2016, now 4th round)
- KpqC Competition (South Korea, announced in 2022, now 1st round)

Competitions for post–quantum algorithms

- NIST PQC (USA, announced in 2016, now 4th round)
- KpqC Competition (South Korea, announced in 2022, now 1st round)

McEliece scheme

- Classic McEliece (NIST PQC)
- PALOMA (KpqC Competition)

McEliece Cryptosystem $\mathrm{McE}(C)$

Keys

- $K_{sec} = (S, G_C, P)$, $S \in GL_k(\mathbb{F}_2)$, G_C — generating matrix of $[n, k, d]_2$ Goppa code C, P — permutation $(n \times n)$–matrix;
- $K_{pub} = (\tilde{G} = SG_C P, t = \lfloor (d-1)/2 \rfloor)$.

McEliece Cryptosystem $\mathrm{McE}(C)$

Keys

- $K_{sec} = (S, G_C, P)$, $S \in GL_k(\mathbb{F}_2)$, G_C — generating matrix of $[n, k, d]_2$ Goppa code C, P — permutation $(n \times n)$–matrix;
- $K_{pub} = (\tilde{G} = SG_C P, t = \lfloor (d-1)/2 \rfloor)$.

Advantages

- resistance to structural attacks and attacks on the ciphertext
- fast encryption and decryption

McEliece Cryptosystem $\mathrm{McE}(C)$

Keys

- $K_{sec} = (S, G_C, P)$, $S \in GL_k(\mathbb{F}_2)$, G_C — generating matrix of $[n, k, d]_2$ Goppa code C, P — permutation $(n \times n)$–matrix;
- $K_{pub} = (\tilde{G} = SG_C P, t = \lfloor (d-1)/2 \rfloor)$.

Advantages

- resistance to structural attacks and attacks on the ciphertext
- fast encryption and decryption

Disadvantages

- public key size

McEliece Cryptosystem $\mathrm{McE}(C)$

Attempts to reduce the key size

- Reed-Solomon codes
  - proposed (Niederreiter, 1986)
  - attacked by (Sidelnikov & Shestakov, 1992)
- Reed-Muller codes
  - proposed (Sidelnikov, 1994)
  - attacked by (Minder & Shokrollahi, 2007), (Borodin & Chizhov, 2014)
- algebro-geometric codes
  - proposed (Janwa & Moreno, 1996)
  - attacked by (Couvreur et. al., 2017)
- low-density parity-check codes
  - proposed (Baldi et. al., 2013)
  - attacked by (Fabšič et. al., 2017)

Remarks on the Structural Security of the McEliece System

- Goppa codes — alternate codes
- Attacks on some classes of alternate codes:
  - **Wild-Goppa codes:** "Polynomial time attack on wild McEliece over quadratic extensions" (Couvreur A., Otmani A., Tillich J. P., 2016)
  - **subspace subcodes of Reed-Solomon codes:** "On the Security of Subspace Subcodes of Reed–Solomon Codes for Public Key Encryption" (Couvreur A., Lequesne M., 2021)
  - **BCH codes:** "An Algebraic Attack Against McEliece-like Cryptosystems Based on BCH Codes" (Elbro F., Majenz C., 2022)
- The problem of investigation of other codes in a McEliece-type systems is relevant.

Combining Codes

Advantages

- the ability to evaluate the characteristics of the new code through the base codes
- simplification of building a decoder
- the new code belongs to another class
- simplification of the security analysis of a new cryptosystem

Combinations examples:

- direct sum of codes
- combination of codes (repetition codes)
- transition from field extensions to basic fields
- code concatinating
- $(U, U + V)$–construction and its generalizations
- tensor product of codes and its generalizations

Object of Research

- McEliece-type cryptosystem based on D–code construction of Reed-Muller codes
- Security analysis of a cryptosystem based on the properties of the Schur-Hadamard degrees of D–codes on Reed-Muller codes.

D-codes

Tensor Product of Codes

Tensor product of matrices

- Let $A = (a_{i,j})$ — $(k_1 \times n_1)$–matrix, B — $(k_2 \times n_2)$–matrix
- $A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,n_1}B \\ \vdots & \ddots & \vdots \\ a_{k_1,1}B & \cdots & a_{k_1,n_1}B \end{pmatrix}$ — $(k_1k_2 \times n_1n_2)$–matrix.

Tensor product of codes

- Let C_i — $[n_i, k_i, d_i]_q$–code, $i \in \{1, 2\}$.
- $C_1 \otimes C_2 = \mathcal{L}(G_{C_1} \otimes G_{C_2})$ — $[n_1n_2, k_1k_2, d_1d_2]_q$–code.

D–codes (Kasami & Lin, 1971)

Let

- $J_1, J_2 \in \mathbb{N}$,
- $\mathcal{S}_t = \{C_t(0), \ \ldots, \ C_t(J_t)\}, \ C_t(i) \subseteq \mathbb{F}_q^{n_t}, \ t = 1, 2,$
- $D_0 = \{(i,j) | i = 0, \ldots, J_1, j = 0, \ldots, J_2\},$
- $D \subseteq D_0,$
- $C(D) = \mathcal{L}\left(\bigcup_{(i,j)\in D} C_1(i) \otimes C_2(j)\right).$

Then D–code is a code $\overline{C(D)}(\subseteq \mathbb{F}_q^{n_1 n_2})$ dual to $C(D)$.

D–codes (another representation)

Let

- $C_t(0) \supset C_t(1) \supset ... \supset C_t(J_t),\ C_t(i) \in \mathcal{S}_t,\ t = 1, 2,$
- $\overline{C_t(0)} \subset \overline{C_t(1)} \subset ... \subset \overline{C_t(J_t)},\ t = 1, 2,$
- $k_1 < k_2 < ... < k_s,\ k_i \in \{0, ..., J_1\},$
- $l_1 > l_2 > ... > l_s,\ l_i \in \{0, ..., J_2\}.$

Then

$$\overline{C(D)} = \sum_{i=1}^{s} \overline{C_1(k_i)} \otimes \overline{C_2(l_i)}.$$

$\overline{C(D)}$ — $[n, k, d]_q$–code, where

- $n = n_1 n_2,$
- $d = \mathrm{d}(\overline{C(D)}) = \min\{\mathrm{d}(\overline{C_1(k_i)})\mathrm{d}(\overline{C_2(l_i)}) \mid i = 1, ..., s\}$

The Example of $\overline{C(D)}$ Based on Binary Reed-Muller Codes

Let

$$\mathcal{S}_1 = \{C_1(0) = \mathrm{RM}(8,8), C_1(1) = \mathrm{RM}(7,8), \ldots, \\ C_1(8) = \mathrm{RM}(0,8), C_1(9) = \{\bar{0}\}\},$$
$$\mathcal{S}_2 = \{C_2(0) = \mathrm{RM}(8,8), C_2(1) = \mathrm{RM}(7,8), \ldots, \\ C_2(8) = \mathrm{RM}(0,8), C_2(9) = \{\bar{0}\}\},$$

then

$$\overline{C(D)} = \overline{C_1(3)} \otimes \overline{C_2(6)} + \overline{C_1(5)} \otimes \overline{C_2(4)} = \\ = \mathrm{RM}(2,8) \otimes \mathrm{RM}(5,8) + \mathrm{RM}(4,8) \otimes \mathrm{RM}(3,8),$$

$\overline{C(D)}$ — $[65536, 19821, 512]_2$–code.

McEliece–type Cryptosystem Based on D–codes

Keys

- $K_{sec} = (S, G_{\overline{C(D)}}, P)$, $S \in GL_k(\mathbb{F}_2)$, $G_{\overline{C(D)}}$ — generating matrix of D–code $\overline{C(D)}$ with parameters $[n, k, d]_2$, P — permutation $(n \times n)$–matrix;
- $K_{pub} = (\tilde{G} = SG_{\overline{C(D)}}P, t = \lfloor (d-1)/2 \rfloor)$.

Encryption $\mathbf{m}(\in \mathbb{F}_q^k)$:

- $\mathbf{z} = \mathbf{m}\tilde{G} + \mathbf{e}$, $\mathrm{wt}(\mathbf{e}) = t$.

Decryption $\mathbf{z}(\in \mathbb{F}_q^n)$:

- $\mathbf{m} = S^{-1}\tau(G)^{-1}\tau(\mathrm{Dec}_{\overline{C(D)}}(\mathbf{z}P^{-1}))$, where $\mathrm{Dec}_{\overline{C(D)}} : \mathbb{F}_q^n \to \overline{C(D)}$ — efficient decoder for code $\overline{C(D)}$ and τ — any information set.

Security Analysis of the Cryptosystem

Schur–Hadamard Product

- Definition for codes $C, D \subseteq \mathbb{F}_q^n$:

$$C \star D = \mathcal{L}(\{\mathbf{x} \star \mathbf{y} | \mathbf{x} \in C, \mathbf{y} \in D\}), \mathbf{x} \star \mathbf{y} = (x_1 y_1, ..., x_n y_n)$$

- Product properties for some codes:
  - $\mathrm{RM}(r_1, m) \star \mathrm{RM}(r_2, m) = \mathrm{RM}(r_1 + r_2, m)$
  - $\mathrm{GRS}_{k_1} \star \mathrm{GRS}_{k_2} = \mathrm{GRS}_{k_1 + k_2 - 1}$
- Used as code distinguisher

Structural Attack

Theorem 1

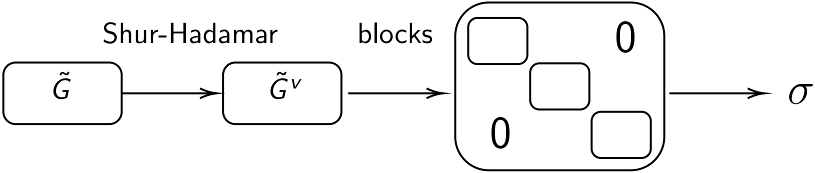
Let $n_1, n_2 \in \mathbb{N}$, $n = n_1 n_2$, C *be a* $[n, k, d]$*–code satisfying the following conditions:*

- $C \subset \mathbb{F}_q^{n_1} \otimes C_2$, C_2 — $[n_2, k_2, d_2]$*–code,*
- $\mathrm{rank}(\tau_i(G_C)) = k_2, \tau_i = \{(i-1)n_2 + 1, ..., in_2\}, i = 1, ..., n_1$
- *Attack — efficient algorithm for structural attack on* $McE(C_2)$,
- $C^v = \mathbb{F}_q^{n_1} \otimes C_2^v$ *for some* $v (\in \mathbb{N})$,
- C_2^v *— indecomposable code.*

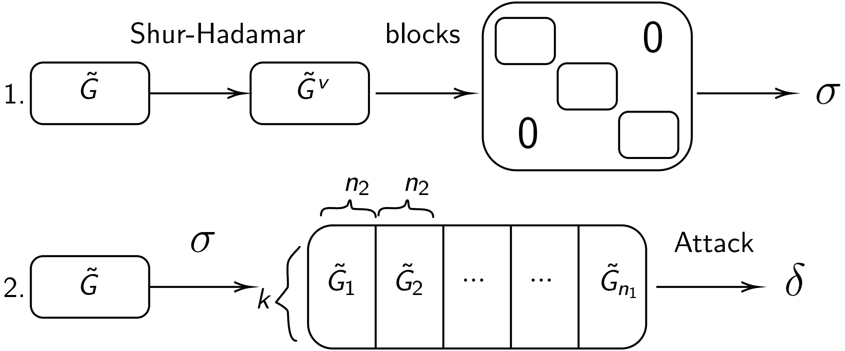
Then there is an efficient algorithm AttackDkey that, given $\tilde{G}$, *finds a permutation* π *such that*

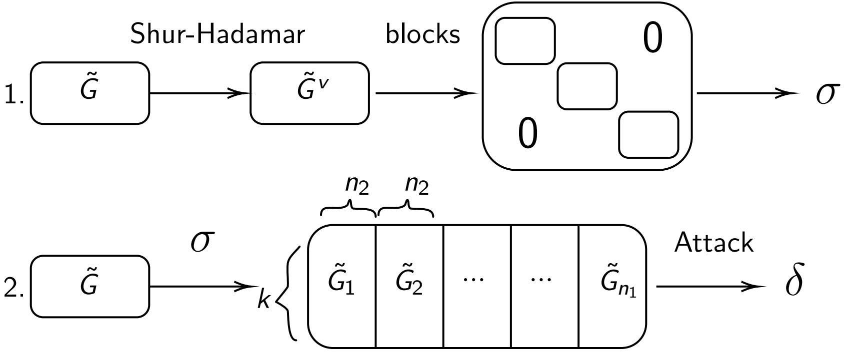
$$\pi(\mathcal{L}(\tilde{G})) \subseteq \mathbb{F}_q^{n_1} \otimes C_2.$$

Structural Attack

1. 

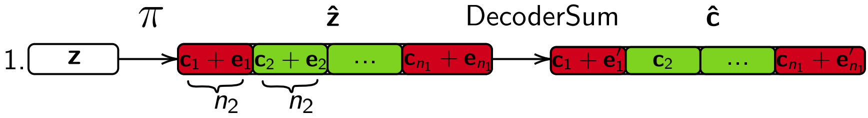
Structural Attack





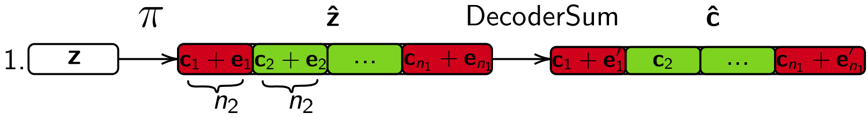
3. $\pi = \delta \circ \sigma$

Combined Attack

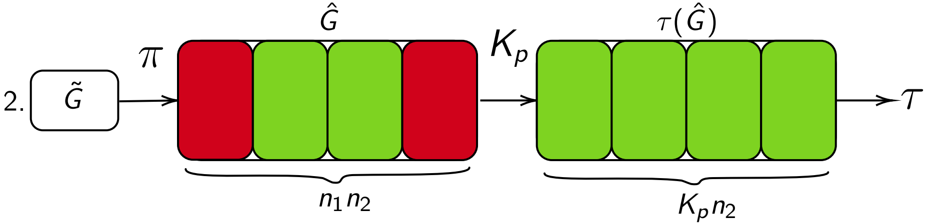


$$\mathbf{c}_i \in C_2,\ 0 \leq \mathrm{wt}(\mathbf{e}_i) \leq n_2, \sum_{i=1}^{n_1} \mathrm{wt}(\mathbf{e}_i) = t$$

Combined Attack



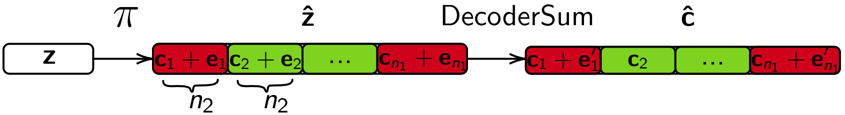
$$\mathbf{c}_i \in C_2,\ 0 \leq \mathrm{wt}(\mathbf{e}_i) \leq n_2, \sum_{i=1}^{n_1} \mathrm{wt}(\mathbf{e}_i) = t$$



$$\mathrm{rank}(\tau(\hat{G})) = k$$

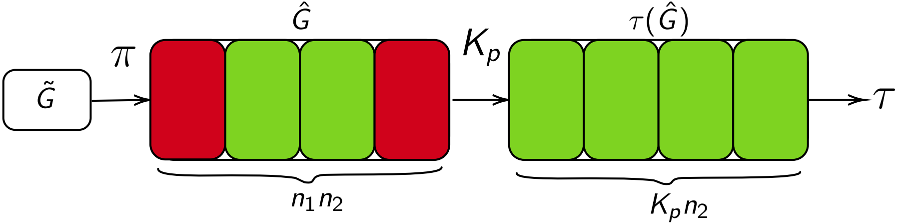
Combined Attack

1.



$$\mathbf{c}_i \in C_2,\ 0 \leq \mathrm{wt}(\mathbf{e}_i) \leq n_2, \sum_{i=1}^{n_1} \mathrm{wt}(\mathbf{e}_i) = t$$

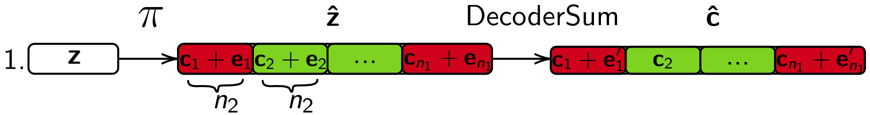
2.



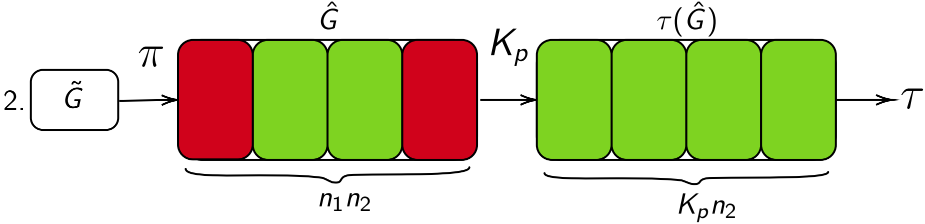
$$\mathrm{rank}(\tau(\hat{G})) = k$$

3. $\hat{\mathbf{m}} = \tau(\hat{\mathbf{c}})\ \tau(\hat{G})^{-1}$

Combined Attack



$$\mathbf{c}_i \in C_2,\ 0 \leq \mathrm{wt}(\mathbf{e}_i) \leq n_2, \sum_{i=1}^{n_1} \mathrm{wt}(\mathbf{e}_i) = t$$



$$\mathrm{rank}(\tau(\hat{G})) = k$$

3. $\hat{\mathbf{m}} = \tau(\hat{\mathbf{c}})\ \tau(\hat{G})^{-1}$

4. $\mathrm{wt}(\hat{\mathbf{z}} - \hat{\mathbf{m}}\hat{G}) \leqslant t$

Success Probability of the Attack

Let $\overline{C(D)}$ — $[n, k, d]_2$–code, $n = n_1 n_2$, C_2 — $[n_2, k_2, d_2]_2$–code, $\overline{C(D)} \subset \mathbb{F}_q^{n_1} \otimes C_2$.

Model for generating error vector e of weight t[1]

- $\Pr(\mathbf{e}_i = 1) = t/n, \ i = 1, ..., n$
- $\Pr(\mathbf{e}_i = 0) = 1 - t/n, \ i = 1, ..., n$

[1] By analogy with "A CCA secure variant of the McEliece cryptosystem" (Dottling N. et. al., 2012)

Success Probability of the Attack

Number of "good" blocks

- **Minimum**:
$$N_g^{min} = n_1 - \lfloor (d-1)/(d_2+1) \rfloor.$$
- **Average** (by the inclusion–exclusion formula):
$$N_g^{avg} = \lfloor n_1 - \sum_{r=0}^{n_1} r \cdot Q_r \rfloor,$$

where $Q_r = \frac{C_r(n_1,n_2,t_1,t_2,t)}{\binom{n}{t}}$, $C_r(n_1,n_2,t_1,t_2,t) = \sum_{k=r}^{n_1} (-1)^{k-r} \binom{k}{r} S_k$,
$S_k = \binom{n_1}{k} \left(\binom{n_2}{t_2+1} \right)^k \binom{(n_1-k)n_2}{t-(t_2+1)k}$.

Success Probability of the Attack

Then the probability P_{attack} of the success of the combined attack is estimated as follows:

- $P_{attack} \geqslant P_1 \cdot P_2$
- $P_1 = \binom{N_g}{K_p} / \binom{n_1}{K_p}$ — probability of choosing K_p "good" blocks
- P_2 — probability that $K_p n_2$ of selected columns form a matrix of rank k
- $P_1^{min} = \binom{N_g^{min}}{K_p} / \binom{n_1}{K_p}$, $P_1^{avg} = \binom{N_g^{avg}}{K_p} / \binom{n_1}{K_p}$
- $P_{attack}^{min} \geqslant P_1^{min} \cdot P_2$, $P_{attack}^{avg} \geqslant P_1^{avg} \cdot P_2$

Properties of Degrees of a Tensor Product and a D–code

Tensor product of Reed–Muller codes

- "On some properties of the Schur–Hadamard product for linear codes and their applications" (Deundyak V. M., Kosolapov Yu. V., 2020)

D–**codes based on Reed–Muller codes**

- "On the structural security of a McEliece–type cryptosystem based on the sum of tensor products of binary Reed–Muller codes" (Kosolapov Yu. V., Lelyuk E. A., 2022)

Examples of Estimating the Attack Success Probability

Table 1: Attack success probability for the tensor product of Reed–Muller codes

$\mathrm{RM}(r_1, m_1) \otimes \mathrm{RM}(r_2, m_2)$	p	K_p	P_{ISD}	P^{min}_{attack}	P^{avg}_{attack}
$\mathrm{RM}(4,7) \otimes \mathrm{RM}(3,7)$	0.1	99	2.379E-14	2.883E-06	3.956E-02
$\mathrm{RM}(5,7) \otimes \mathrm{RM}(3,7)$	0.3	120	1.577E-09	5.945E-05	2.265E-02
$\mathrm{RM}(6,7) \otimes \mathrm{RM}(3,7)$	0.9	127	1.716E-05	7.812E-03	7.812E-03
$\mathrm{RM}(4,8) \otimes \mathrm{RM}(3,8)$	0.2	163	4.868E-30	2.603E-08	8.101E-02
$\mathrm{RM}(5,8) \otimes \mathrm{RM}(3,8)$	0.1	219	1.931E-21	1.488E-07	2.749E-02
$\mathrm{RM}(6,8) \otimes \mathrm{RM}(3,8)$	0.3	247	9.941E-13	1.019E-05	1.179E-02
$\mathrm{RM}(4,8) \otimes \mathrm{RM}(3,7)$	0.2	163	4.412E-22	2.575E-08	8.014E-02
$\mathrm{RM}(4,8) \otimes \mathrm{RM}(2,8)$	0.2	163	2.788E-22	2.551E-08	7.938E-02

Examples of Estimating the Attack Success Probability

Table 2: Attack success probability for D–codes based on Reed–Muller codes

D–code	p	K_p	P_{ISD}	P^{min}_{attack}	P^{avg}_{attack}
[[4, 3], [5, 2]]	0.01	219	1.013E-34	1.117E-16	0.145
[[4, 3], [5, 2], [6, 1]]	0.01	247	2.646E-35	0	0.011
[[4, 3], [5, 2], [6, 1], [7, 0]]	0.01	255	2.536E-35	0	0.004

$$[[r_1^1, r_1^2],[r_2^1, r_2^2], \ldots] =$$
$$= \mathrm{RM}(r_1^1, 8) \otimes \mathrm{RM}(r_1^2, 8) + \mathrm{RM}(r_2^1, 8) \otimes \mathrm{RM}(r_2^2, 8) + \ldots$$

Resistant D–codes

Table 3: Resistant D-codes

#	D–code	k	d	P_{ISD}
1[2]	$[[0,8],[1,7],[2,6],[3,5],[4,4],[5,3],[6,2],[7,1],[8,0]]$	39203	256	1.715E-51
2[3]	$[[0,8],[1,7],[2,6],[3,5],[4,4],[5,3],[6,2],[7,1]]$	39202	256	1.723E-51
3	$[[1,7],[2,6],[3,5],[4,4],[5,3],[6,2],[7,1]]$	39201	256	1.732E-51
4	$[[1,7],[2,6],[3,5],[4,4],[5,3],[6,2]]$	39129	256	2.458E-51
5	$[[2,6],[3,5],[4,4],[5,3],[6,2]]$	39057	256	3.486E-51
6	$[[2,6],[3,5],[4,4],[5,3]]$	38021	256	4.796E-49
7	$[[3,5],[4,4],[5,3]]$	36985	256	5.498E-47
8	$[[2,5],[4,3]]$	19821	512	7.279E-41
9	$[[4,4]]$	26569	256	1.156E-29

[2] "Effective attack on the McEliece cryptosystem based on Reed-Muller codes" (Borodin M. A., Chizhov I. V., 2014)

[3] "Classification of Hadamard products of subcodes of codimension 1 of Reed-Muller codes" (Borodin M. A., Chizhov I. V., 2020)

Comparison of the McEliece–type Cryptosystems

Table 4: Comparison of the characteristics of McEliece–type cryptosystems

Code	Goppa code	Reed–Muller code		D–code	
$[n, k, d]$	$[3488, 2720, \geqslant 129]$	$[65536, 14893, 1024]$	$[65536, 39203, 256]$	$[65536, 19821, 512]$	$[65536, 39201, 256]$
Size of publ. key	1.13Mb	116.35Mb	306.27Mb	154.85Mb	306.25Mb
$R = k/n$	≈ 0.78	≈ 0.23	≈ 0.6	≈ 0.3	≈ 0.6
Decoder	Patterson decoding	Reed decoding		majority–logical decoding	
t	64	511	127	255	127
P_{ISD}	$2^{-142.8}$	$2^{-192.62}$	$2^{-169.37}$	$2^{-136.16}$	$2^{-169.37}$
Structural attacks	$-$	$+$		$-$	

Conclusion and Further Research

- D–codes based on Reed–Muller codes can be subcodes of Reed–Muller codes $\Rightarrow$ decoders for Reed–Muller codes can be used.
- Decoders for Reed–Muller codes:
  - **Sidelnikov–Pershakov decoder and its modifications:** "Decoding Reed—Muller Codes with a Large Number of Errors" (Sidelnikov V. M., Pershakov A. S., 1992),
  - **Dumer's list decoder:** "Recursive decoding and its performance for low–rate Reed–Muller codes" (Dumer I., 2004),
  - **permutation decoder:** "A new permutation decoding method for Reed–Muller codes" (Kamenev M. et. al., 2019),
  - **decoder for low–density codes:** "Iterative Reed–Muller Decoding" (Geiselhart M. et. al., 2021).
- Reed–Muller codes are now being actively investigated due to their connection with polar codes.