

CODE-BASED PUBLIC-KEY ENCRYPTION SCHEMES

Vlad-Florin Dragoi

vlad.dragoi@uav.ro

Universitatea "Aurel Vlaicu" Arad, Romania
LITIS EA 4108 Université de Rouen Normandie, France



Insights into the theory of error correcting codes

ERROR CORRECTING CODES

DEFINITION 1

A binary linear code \mathcal{C} defined over \mathbb{F}_2 is a k dimension sub-vector space of \mathbb{F}_2^n . $\mathbf{G} \in \mathbb{F}_2^{k \times n}$ a basis, and $\mathbf{H} \in \mathbb{F}_2^{(n-k) \times n}$ a basis for the dual.

$$\mathcal{C} = \langle \mathbf{G} \rangle = \{ \mathbf{c} = \mathbf{m}\mathbf{G} \mid \mathbf{m} \in \mathbb{F}_2^k \} \quad \mathcal{C} = \langle \mathbf{H} \rangle^\perp = \{ \mathbf{H}\mathbf{c} = 0 \mid \mathbf{c} \in \mathbb{F}_2^n \}$$

ERROR CORRECTING CODES

DEFINITION 1

A binary linear code \mathcal{C} defined over \mathbb{F}_2 is a k dimension sub-vector space of \mathbb{F}_2^n . $\mathbf{G} \in \mathbb{F}_2^{k \times n}$ a basis, and $\mathbf{H} \in \mathbb{F}_2^{(n-k) \times n}$ a basis for the dual.

$$\mathcal{C} = \langle \mathbf{G} \rangle = \{ \mathbf{c} = \mathbf{m}\mathbf{G} \mid \mathbf{m} \in \mathbb{F}_2^k \} \quad \mathcal{C} = \langle \mathbf{H} \rangle^\perp = \{ \mathbf{H}\mathbf{c} = 0 \mid \mathbf{c} \in \mathbb{F}_2^n \}$$

REMARK

For any $\mathbf{x} \in \mathbb{F}_2^n$ denote $\text{supp}(\mathbf{x}) = \{i \mid x_i \neq 0\}$.

Any $\mathbf{x} \in \mathbb{F}_2^n$ with $|\text{supp}(\mathbf{x})| = 0 \pmod 2$ is self-orthogonal.

$$\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^n x_i \pmod 2 = 0.$$

A LINEAR CODE IS A METRIC SPACE

DEFINITION 2 (HAMMING WEIGHT AND DISTANCE)

Let $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{F}_2^n$

$$\|\mathbf{x}\| \stackrel{\text{def}}{=} |\{i \mid x_i \neq 0\}| \quad d_H(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} |\{i \mid x_i \neq y_i\}|$$

A LINEAR CODE IS A METRIC SPACE

DEFINITION 2 (HAMMING WEIGHT AND DISTANCE)

Let $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{F}_2^n$

$$\|\mathbf{x}\| \stackrel{\text{def}}{=} |\{i \mid x_i \neq 0\}| \quad d_H(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} |\{i \mid x_i \neq y_i\}|$$

$$\begin{aligned} d_{\min}(\mathcal{C}) &= \min_{\substack{(\mathbf{c}, \mathbf{c}^*) \in \mathcal{C} \times \mathcal{C} \\ \mathbf{c} \neq \mathbf{c}^*}} d_H(\mathbf{c}, \mathbf{c}^*) \\ &= \min_{\mathbf{c} \in \mathcal{C}, \mathbf{c} \neq \mathbf{0}} \|\mathbf{c}\| \\ &= \min_{\mathbf{c} \in \mathcal{C}, \mathbf{c} \neq \mathbf{0}} |\text{supp}(\mathbf{c})|. \end{aligned}$$

CODE PARAMETERS

- \mathcal{C} is a $[n, k, d]$ code : n -length, k -dimension, d -minimum distance

CODE PARAMETERS

- \mathcal{C} is a $[n, k, d]$ code : n -length, k -dimension, d -minimum distance
- n, k are easy to determine (basis)

CODE PARAMETERS

- \mathcal{C} is a $[n, k, d]$ code : n -length, k -dimension, d -minimum distance
- d depends on the family of codes

CODE PARAMETERS

- \mathcal{C} is a $[n, k, d]$ code : n -length, k -dimension, d -minimum distance
- d depends on the family of codes
 - ▶ Codes with particular underlying structure could have an easy computable d

CODE PARAMETERS

- \mathcal{C} is a $[n, k, d]$ code : n -length, k -dimension, d -minimum distance
- d depends on the family of codes
 - ▶ In general computing d given \mathbf{G} or \mathbf{H} is a difficult problem¹

1. A. Vardy, "The intractability of computing the minimum distance of a code," in IEEE Transactions on Information Theory, vol. 43, no. 6, pp. 1757-1766, Nov. 1997

CODE PARAMETERS

- \mathcal{C} is a $[n, k, d]$ code : n -length, k -dimension, d -minimum distance
- d depends on the family of codes
 - ▶ In general computing d given \mathbf{G} or \mathbf{H} is a difficult problem
 - ▶ The Gilbert-Varshamov bound, d_{GV} is the smallest d s.t.

$$\sum_{i=0}^{d-1} \binom{n}{i} \geq 2^{(n-k)}$$

CODE PARAMETERS

- \mathcal{C} is a $[n, k, d]$ code : n -length, k -dimension, d -minimum distance
- d depends on the family of codes
 - ▶ In general computing d given \mathbf{G} or \mathbf{H} is a difficult problem
 - ▶ The Gilbert-Varshamov bound, d_{GV} is the smallest d s.t.

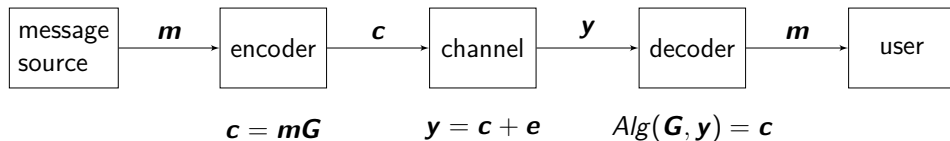
$$\sum_{i=0}^{d-1} \binom{n}{i} \geq 2^{(n-k)}$$

- ▶ In the asymptotics : The minimum distance of a $[n, k]$ linear code meets the Gilbert-Varshamov bound¹ $d_{GV} = n\delta_{GV}$

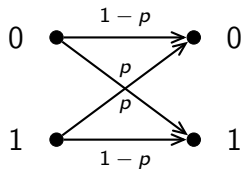
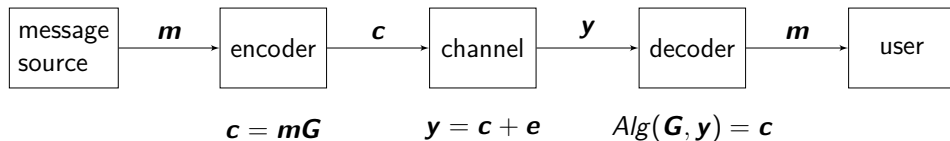
$$1 - k/n = H(\delta_{GV})$$

1. A. Barg and G. D. Forney, "Random codes : minimum distances and error exponents," in IEEE Transactions on Information Theory, vol. 48, no. 9, pp. 2568-2573, Sept. 2002

ENCODING-DECODING



ENCODING-DECODING



DECODING

DEFINITION 1 (DISCRETE CHANNEL)

Let k and m be two strictly positive integers. Then a discrete channel W is defined by

- A finite input alphabet $\mathcal{X} = \{x_1, \dots, x_k\}$.
- A finite output alphabet $\mathcal{Y} = \{y_1, \dots, y_m\}$.
- The transition probability matrix $\mathbf{P} = (p_{i,j})_{1 \leq i \leq k, 1 \leq j \leq m}$ with $p_{i,j} = W(y_j|x_i)$ is the probability that y_j is received knowing that x_i was sent over the channel.

DECODING

DEFINITION 2

A decoder for \mathcal{C} with respect to W is a function $\mathcal{D} : \mathcal{Y}^n \rightarrow \mathcal{C}$.

The probability that a codeword \mathbf{c} is decoded erroneously, given that \mathbf{c} was transmitted

$$P_{\text{err}}(\mathbf{c}) \stackrel{\text{def}}{=} \sum_{\substack{\mathbf{y} \in \mathcal{Y}^n \\ \mathcal{D}(\mathbf{y}) \neq \mathbf{c}}} W(\mathbf{y} | \mathbf{c}).$$

The *error probability* of \mathcal{D} is

$$P_{\text{err}} = \max_{\mathbf{c} \in \mathcal{C}} P_{\text{err}}(\mathbf{c}).$$

DECODING

DEFINITION 3 (MAXIMUM-LIKELIHOOD DECODER)

Given a $[n, k, d]$ linear code \mathcal{C} over \mathbb{F}_2 and a channel $W = (\mathbb{F}_2, \mathcal{Y}, \mathbf{P})$ a *maximum-likelihood decoder* (MLD) for \mathcal{C} with respect to W is the function $\mathcal{D}_{\text{MLD}} : \mathcal{Y}^n \rightarrow \mathcal{C}$ defined as :

$$\text{for every } \mathbf{y} \in \mathcal{Y}^n, \mathcal{D}_{\text{MLD}}(\mathbf{y}) \stackrel{\text{def}}{=} \operatorname{argmax}_{\mathbf{c} \in \mathcal{C}} W(\mathbf{y} | \mathbf{c}).$$

DECODING

DEFINITION 3 (MAXIMUM-LIKELIHOOD DECODER)

Given a $[n, k, d]$ linear code \mathcal{C} over \mathbb{F}_2 and a channel $W = (\mathbb{F}_2, \mathcal{Y}, \mathbf{P})$ a *maximum-likelihood decoder* (MLD) for \mathcal{C} with respect to W is the function $\mathcal{D}_{\text{MLD}} : \mathcal{Y}^n \rightarrow \mathcal{C}$ defined as :

$$\text{for every } \mathbf{y} \in \mathcal{Y}^n, \mathcal{D}_{\text{MLD}}(\mathbf{y}) \stackrel{\text{def}}{=} \operatorname{argmax}_{\mathbf{c} \in \mathcal{C}} W(\mathbf{y} | \mathbf{c}).$$

Ex. BSC(p) with crossover probability $0 < p < 1/2$

$$\begin{aligned} W(\mathbf{y} | \mathbf{c}) &= \prod_{i=1}^n W(y_i | c_i) \\ &= p^{\text{d}_H(\mathbf{y}, \mathbf{c})} (1-p)^{n-\text{d}_H(\mathbf{y}, \mathbf{c})} \\ &= (1-p)^n \left(\frac{p}{1-p} \right)^{\text{d}_H(\mathbf{y}, \mathbf{c})}. \end{aligned}$$

DECODING

DEFINITION 3 (MAXIMUM-LIKELIHOOD DECODER)

Given a $[n, k, d]$ linear code \mathcal{C} over \mathbb{F}_2 and a channel $W = (\mathbb{F}_2, \mathcal{Y}, \mathbf{P})$ a *maximum-likelihood decoder* (MLD) for \mathcal{C} with respect to W is the function $\mathcal{D}_{\text{MLD}} : \mathcal{Y}^n \rightarrow \mathcal{C}$ defined as :

$$\text{for every } \mathbf{y} \in \mathcal{Y}^n, \mathcal{D}_{\text{MLD}}(\mathbf{y}) \stackrel{\text{def}}{=} \operatorname{argmax}_{\mathbf{c} \in \mathcal{C}} W(\mathbf{y} | \mathbf{c}).$$

Ex. BSC(p) with crossover probability $0 < p < 1/2$

$$W(\mathbf{y} | \mathbf{c}) = (1 - p)^n \left(\frac{p}{1 - p} \right)^{d_{\text{H}}(\mathbf{y}, \mathbf{c})}.$$

$\mathcal{D}_{\text{MLD}}(\mathbf{y})$ is the codeword \mathbf{c} which minimize $d_{\text{H}}(\mathbf{y}, \mathbf{c})$
 \mathbf{c} is the closest codeword of \mathcal{C} to \mathbf{y} .

NEAREST CODEWORD PROBLEM

DEFINITION 4 (NEAREST CODEWORD PROBLEM FOR BSC)

Given : $[n, k, d]$ linear code \mathcal{C} over \mathbb{F}_2 and a vector $\mathbf{y} \in \mathbb{F}_2^n$.

Find : $\mathbf{e} \in \mathbb{F}_2^n$ of minimum Hamming weight such that $\mathbf{y} - \mathbf{e} \in \mathcal{C}$.

NEAREST CODEWORD PROBLEM

DEFINITION 4 (NEAREST CODEWORD PROBLEM FOR BSC)

Given : $[n, k, d]$ linear code \mathcal{C} over \mathbb{F}_2 and a vector $\mathbf{y} \in \mathbb{F}_2^n$.

Find : $\mathbf{e} \in \mathbb{F}_2^n$ of minimum Hamming weight such that $\mathbf{y} - \mathbf{e} \in \mathcal{C}$.

A possible solution is to use the dual code

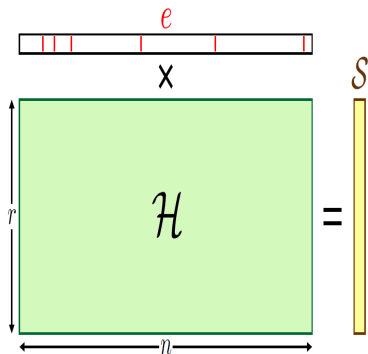
- ▶ $\mathbf{y} - \mathbf{e} \in \mathcal{C} \Leftrightarrow \mathbf{H}(\mathbf{y} - \mathbf{e}) = \mathbf{0}$
- ▶ let $\mathbf{s} = \mathbf{H}\mathbf{y}$ be a syndrome (associated to a vector, with respect to a matrix)
- ▶ We have

$$\mathbf{y} - \mathbf{e} \in \mathcal{C} \Leftrightarrow \mathbf{H}\mathbf{e} = \mathbf{s}$$

SYNDROME DECODING ²

Given : A parity-check matrix \mathbf{H} for a $[n, k, d]$ binary linear code
a syndrome vector $\mathbf{s} \in \mathbb{F}_2^{n-k}$ and $t \in \mathbb{N}$

Find : $\mathbf{e} \in \mathbb{F}_2^n$ of weight at most t such that $\mathbf{H}\mathbf{e} = \mathbf{s}$.



2. 1978. Berlekamp E., McEliece R.J., Van Tilborg "On the inherent intractability of certain coding problems."

BOUNDED DECODING³

If there is a codeword \mathbf{c} s.t. $d_H(\mathbf{c}, \mathbf{y}) \leq \lfloor \frac{d-1}{2} \rfloor$ we talk about unique solution (bounded decoding).

Given : A parity-check matrix \mathbf{H} for a $[n, k]$ binary linear code
a syndrome vector $\mathbf{s} \in \mathbb{F}_2^{n-k}$ and $t \leq \lfloor \frac{d-1}{2} \rfloor$

Promise : any $d - 1$ columns of \mathbf{H} are linearly independent

Find : $\mathbf{e} \in \mathbb{F}_2^n$ of weight at most t such that $\mathbf{H}\mathbf{e} = \mathbf{s}$.

3. A Barg. Complexity issues in coding theory. *Handbook of Coding Theory, Elsevier Science*, 1998.

BOUNDED DECODING³

If there is a codeword \mathbf{c} s.t. $d_H(\mathbf{c}, \mathbf{y}) \leq \lfloor \frac{d-1}{2} \rfloor$ we talk about unique solution (bounded decoding).

Given : A parity-check matrix \mathbf{H} for a $[n, k]$ binary linear code
a syndrome vector $\mathbf{s} \in \mathbb{F}_2^{n-k}$ and $t \leq \lfloor \frac{d-1}{2} \rfloor$

Promise : any $d - 1$ columns of \mathbf{H} are linearly independent

Find : $\mathbf{e} \in \mathbb{F}_2^n$ of weight at most t such that $\mathbf{H}\mathbf{e} = \mathbf{s}$.

Verifying the promise condition is NP-complete. Bounded Decoding was conjectured NP-hard for random linear codes.

3. A Barg. Complexity issues in coding theory. *Handbook of Coding Theory, Elsevier Science*, 1998.

DECODING

- Random linear codes :
 - ▶ maximum likelihood decoding (NP-complete)

DECODING

- Random linear codes :
 - ▶ maximum likelihood decoding (NP-complete)
 - ▶ bounded decoding (Conjectured NP-hard)

DECODING

- Random linear codes :
 - ▶ maximum likelihood decoding (NP-complete)
 - ▶ bounded decoding (Conjectured NP-hard)
- Codes with a particular structure :
 - ▶ maximum likelihood decoding is NP-complete even for Reed-Solomon, concatenated codes.

DECODING

- Random linear codes :
 - ▶ maximum likelihood decoding (NP-complete)
 - ▶ bounded decoding (Conjectured NP-hard)

- Codes with a particular structure :
 - ▶ maximum likelihood decoding is NP-complete even for Reed-Solomon, concatenated codes.
 - ▶ efficient algorithms for bounded decoding exist :
 - ★ Patterson/Berlekamp-Massey algorithm - Goppa codes
 - ★ Extended Euclidean Algorithm - Alternant codes, Reed-Solomon codes, BCH codes
 - ★ Bit flipping algorithm - LDPC/MDPC codes,
 - ★ Reed algorithm, Berlekamp-Welsh algorithm - Reed-Muller codes

SOME USEFUL PROBLEMS

- Given a random linear code \mathcal{C} specified by \mathbf{G} and an erroneous codeword, retrieve the initial codeword.

$$\psi(\mathbf{Gm} + \mathbf{e}, \mathbf{G}) = m$$

- Given a random linear code \mathcal{C} specified by \mathbf{H} and a syndrome vector, retrieve the error vector.

$$\psi(\mathbf{He}, \mathbf{H}) = \mathbf{e}$$

- Given a random linear code \mathcal{C} and a vector, distinguish between random vectors and erroneous codewords.

$$\varphi(\mathbf{G}, \mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{y} = \text{random} \\ 1 & \text{if } \mathbf{y} = m\mathbf{G} + \mathbf{e} \end{cases}$$

Public-key encryption schemes from codes

PUBLIC-KEY ENCRYPTION FROM CODES

- Choose a family of codes that admits an efficient decoding algorithm

PUBLIC-KEY ENCRYPTION FROM CODES

- Choose a family of codes that admits an efficient decoding algorithm
- Intentionally add errors to a codeword – **Encryption**

(McEliece) $\mathbf{z} = \mathbf{mG} + \mathbf{e}$ or $\mathbf{m} \rightarrow \mathbf{e}$, $\mathbf{z} = \mathbf{He}^t$ (Niederreiter)

PUBLIC-KEY ENCRYPTION FROM CODES

- Choose a family of codes that admits an efficient decoding algorithm
- Intentionally add errors to a codeword – **Encryption**

(McEliece) $\mathbf{z} = \mathbf{mG} + \mathbf{e}$ or $\mathbf{m} \rightarrow \mathbf{e}$, $\mathbf{z} = \mathbf{He}^t$ (Niederreiter)

- Mask the structure of the underlying code – **Key generation**

$$\mathbf{G}_{pub} = \mathbf{SGP} \quad , \quad \mathbf{H}_{pub} = \mathbf{SHP}$$

McEliece PKE	Niederreiter PKE
$\text{KeyGen}(n, k, t) = (\text{pk}, \text{sk})$	
\mathbf{G} -generator matrix of \mathcal{C}	\mathbf{H} -parity-check of \mathcal{C}
$\backslash \backslash \mathcal{C}$ an $[n, k]$ that corrects t errors An $n \times n$ permutation matrix \mathbf{P}	
A $k \times k$ invertible matrix \mathbf{S}	An $(n - k) \times (n - k)$ invertible matrix \mathbf{S}
Compute $\mathbf{G}_{pub} = \mathbf{SGP}$	Compute $\mathbf{H}_{pub} = \mathbf{SHP}$
$\text{pk} = (\mathbf{G}_{pub}, t)$	$\text{pk} = (\mathbf{H}_{pub}, t)$
$\text{sk} = (\mathbf{S}, \mathbf{G}, \mathbf{P})$	$\text{sk} = (\mathbf{S}, \mathbf{H}, \mathbf{P})$
$\text{Encrypt}(\mathbf{m}, \text{pk}) = \mathbf{z}$	
Encode $\mathbf{m} \rightarrow \mathbf{c} = \mathbf{mG}_{pub}$	Encode $\mathbf{m} \rightarrow \mathbf{e}$
Choose \mathbf{e}	
$\backslash \backslash \mathbf{e}$ a vector of weight t	
$\mathbf{z} = \mathbf{c} + \mathbf{e}$	$\mathbf{z} = \mathbf{H}_{pub}\mathbf{e}^t$
$\text{Decrypt}(\mathbf{z}, \text{sk}) = \mathbf{m}$	
Compute $\mathbf{z}^* = \mathbf{zP}^{-1}$	Compute $\mathbf{z}^* = \mathbf{S}^{-1}\mathbf{z}$
$\mathbf{z}^* = \mathbf{mSG} + \mathbf{eP}^{-1}$	$\mathbf{z}^* = \mathbf{HPe}$
$\mathbf{m}^* = \text{Decode}(\mathbf{z}^*, \mathbf{G})$	$\mathbf{e}^* = \text{Decode}(\mathbf{z}^*, \mathbf{H})$
Retrieve \mathbf{m} from $\mathbf{m}^*\mathbf{S}^{-1}$	Retrieve \mathbf{m} from $\mathbf{P}^{-1}\mathbf{e}^*$

SEMANTIC SECURITY

ONE-WAY FUNCTION

- Assumptions

- ▶ Indistinguishability : The public code is computationally indistinguishable from a uniformly chosen code of the same size (n,k) .
- ▶ Decoding hardness : Decoding a random linear code with parameters n, k, t is hard.

4. B. Biswas, N. Sendrier. McEliece Cryptosystem Implementation : Theory and Practice. PQCrypto. pp. 47-62. 2008.

SEMANTIC SECURITY

ONE-WAY FUNCTION

- Assumptions
 - ▶ Indistinguishability : The public code is computationally indistinguishable from a uniformly chosen code of the same size (n,k) .
 - ▶ Decoding hardness : Decoding a random linear code with parameters n, k, t is hard.
- Given that both the above assumptions hold, the McEliece cryptosystem is one-way secure under passive attacks.⁴

4. B. Biswas, N. Sendrier. McEliece Cryptosystem Implementation : Theory and Practice. PQCrypto. pp. 47-62. 2008.

DECODING HARDNESS IN THE McELIECE SCHEME⁵

The binary Goppa code is a $[2^m, 2^m - mt, 2t + 1]$

5. Finiasz, Matthieu. “Nouvelles constructions utilisant des codes correcteurs d’erreurs en cryptographie à clef publique.” (2004).

DECODING HARDNESS IN THE McELIECE SCHEME⁵

The binary Goppa code is a $[2^m, 2^m - mt, 2t + 1]$

Given : A parity-check matrix \mathbf{H} for a $[n, n - k]$ binary linear code
a syndrome vector $\mathbf{s} \in \mathbb{F}_2^{n-k}$ and $t \in \mathbb{N}$ ($n = 2^m$)

Find : $\mathbf{e} \in \mathbb{F}_2^n$ of weight $t \leq (n - k) / \log_2(n)$ such that $\mathbf{H}\mathbf{e} = \mathbf{s}$.

5. Finiasz, Matthieu. "Nouvelles constructions utilisant des codes correcteurs d'erreurs en cryptographie à clef publique." (2004).

DISTINGUISHER ASSUMPTION FOR GOPPA CODES⁶

- Pseudo-randomness assumption

Input : A generator matrix \mathbf{G} for a $[2^m, 2^m - mt]$ binary linear code

Output : \mathbf{G} generates a Goppa code?

6. Jean-Charles Faugère, Valérie Gauthier-Umana, Ayoub Otmani, Ludovic Perret, Jean-Pierre Tillich. A Distinguisher for High Rate McEliece Cryptosystems. IEEE Transactions on Information Theory 2013.

SEMANTIC SECURITY

CRITICAL ATTACKS

- McEliece PKE does not satisfy Non-Malleability (linearity)

given a McEliece criptogram $\mathbf{y} = \mathbf{m}\mathbf{G}_{pub} + \mathbf{e}$

compute a well-choose criptogram $\mathbf{y}^* = \mathbf{m}^*\mathbf{G}_{pub}$

as the oracle to decrypt $\mathbf{y} + \mathbf{y}^* = (\mathbf{m} + \mathbf{m}^*)\mathbf{G}_{pub} + \mathbf{e}$

SEMANTIC SECURITY

CRITICAL ATTACKS

- McEliece PKE does not satisfy Non-Malleability (linearity)

given a McEliece criptogram $\mathbf{y} = \mathbf{m}\mathbf{G}_{pub} + \mathbf{e}$
compute a well-choose criptogram $\mathbf{y}^* = \mathbf{m}^*\mathbf{G}_{pub}$
as the oracle to decrypt $\mathbf{y} + \mathbf{y}^* = (\mathbf{m} + \mathbf{m}^*)\mathbf{G}_{pub} + \mathbf{e}$

- Reaction attacks in the CCA model

given a McEliece criptogram $\mathbf{y} = \mathbf{m}\mathbf{G}_{pub} + \mathbf{e}$
flip a bit $\mathbf{y}' = \mathbf{y} + (1, 0, \dots, 0)$
 $\mathbf{y}' = \mathbf{m}\mathbf{G}_{pub} + \mathbf{e} + (1, 0, \dots, 0)$
if the decoder reaction is invalid ciphertext $e_i = 0$
if the decoder reaction is valid ciphertext $e_i = 1$

SEMANTIC SECURITY

CRITICAL ATTACKS

- Resend-message attacks : the same message was encrypted several times

intercept $\mathbf{y}_1 = m\mathbf{G}_{pub} + \mathbf{e}_1$

intercept $\mathbf{y}_2 = m\mathbf{G}_{pub} + \mathbf{e}_2$

notice that $d_H(\mathbf{y}_1, \mathbf{y}_2) = d_H(\mathbf{e}_1, \mathbf{e}_2) = 2t - 2\delta$

if the messages were different $d_H(\mathbf{y}_1, \mathbf{y}_2) \sim n/2$

select the set $I = \text{supp}(\mathbf{y}_1 - \mathbf{y}_2)$

Gaussian elimination on I $\mathbf{H}_{pub}\mathbf{e}_1 = \mathbf{s}_1$.

SEMANTIC SECURITY

CONVERSIONS

- For a McEliece IND-CPA without random oracles simply randomize the message $\mathbf{m}^* = (\mathbf{m}|\mathbf{r})^7$

7. Nojima, R., Imai, H., Kobara, K. et al. Semantic security for the McEliece cryptosystem without random oracles. *Des. Codes Cryptogr.* 49, 289–305 (2008)

8. K. Kobara and H. Imai. *Semantically Secure McEliece Public-Key Cryptosystems Conversions for McEliece PKC*, LNCS Springer, 2001

SEMANTIC SECURITY

CONVERSIONS

- For a McEliece IND-CPA without random oracles simply randomize the message $m^* = (m|r)^7$
- For random oracles model - convert the one way trap door function into an IND-CCA2 PKC
 - ▶ simple OAEP conversion not working because of reaction attacks
 - ▶ Kobara, Imai conversion to obtain an IND-CCA2⁸

7. Nojima, R., Imai, H., Kobara, K. et al. Semantic security for the McEliece cryptosystem without random oracles. *Des. Codes Cryptogr.* 49, 289–305 (2008)

8. K. Kobara and H. Imai. *Semantically Secure McEliece Public-Key Cryptosystems Conversions for McEliece PKC*, LNCS Springer, 2001

McEliece AND Niederreiter

MRA, KRA, DISTINGUISHER

	McEliece	Niederreiter
pk	\mathbf{G}_{pub}	\mathbf{H}_{pub}
MRA	Generic decoding $Alg(\mathbf{mG}_{pub} + \mathbf{e}, \mathbf{G}_{pub}) = \mathbf{m}$ $\ \mathbf{e}\ $ small	Syndrome decoding $Alg(\mathbf{H}_{pub}\mathbf{e}, \mathbf{H}_{pub}) = \mathbf{e}$ $\ \mathbf{e}\ $ small
KRA	Code Equivalence Problem	
	$Alg(\mathbf{G}_{pub}, \mathbf{G}) = \mathbf{P}^*$	$Alg(\mathbf{H}_{pub}, \mathbf{H}) = \mathbf{P}^*$
	$\mathcal{C} \stackrel{P.E.}{\sim} \mathcal{C}_{pub} \Leftrightarrow \mathcal{C}^\perp \stackrel{P.E.}{\sim} \mathcal{C}_{pub}^\perp$	
Distinguisher	$D(\mathbf{G}_{pub}) = \begin{cases} 0 & \text{if } \delta = \delta_{Goppa} \\ 1 & \text{if } \delta = \delta_{Random} \end{cases}$	$D(\mathbf{H}_{pub}) = \begin{cases} 0 & \text{if } \delta = \delta_{Reed-Solomon} \\ 1 & \text{if } \delta = \delta_{Random} \end{cases}$

Distinguish a public code from a random
code

EFFICIENT DISTINGUISHER FOR SOME FAMILIES OF CODES

$$\mathbf{x} \star \mathbf{y} = (x_1y_1, \dots, x_ny_n).$$

EFFICIENT DISTINGUISHER FOR SOME FAMILIES OF CODES

$$\mathbf{x} \star \mathbf{y} = (x_1y_1, \dots, x_ny_n).$$

THEOREM 5 (CASCUDO, CRAMER, MIRANDOLA, ZEMOR -2015)

Let $\mathcal{C}_1 = [n, k_1]$ and $\mathcal{C}_2 = [n, k_2]$. Then w.h.p. we have

$$\text{Dim}(\mathcal{C}_1 \star \mathcal{C}_2) = \min \left\{ n, k_1k_2 - \binom{\text{Dim}(\mathcal{C}_1 \cap \mathcal{C}_2)}{2} \right\}.$$

EFFICIENT DISTINGUISHER FOR SOME FAMILIES OF CODES

$$\mathbf{x} \star \mathbf{y} = (x_1 y_1, \dots, x_n y_n).$$

THEOREM 5 (CASCUDO, CRAMER, MIRANDOLA, ZEMOR -2015)

Let $\mathcal{C}_1 = [n, k_1]$ and $\mathcal{C}_2 = [n, k_2]$. Then w.h.p. we have

$$\text{Dim}(\mathcal{C}_1 \star \mathcal{C}_2) = \min \left\{ n, k_1 k_2 - \binom{\text{Dim}(\mathcal{C}_1 \cap \mathcal{C}_2)}{2} \right\}.$$

In particular, for $\mathcal{C} = [n, k]$ random binary code we have

$$\text{Dim}(\mathcal{C}^2) = \min \left\{ n, \binom{k+1}{2} \right\}. \quad (1)$$

REED-SOLOMON CODES

DEFINITION 6 (GENERALIZED REED-SOLOMON CODES)

Let $(\mathbf{x}, \mathbf{y}) \in \mathbb{F}_{2^m}^n \times \mathbb{F}_{2^m}^n$ be a pair such that $\forall i, y_i \neq 0$ and $\forall i \neq j, x_i \neq x_j$.

$$\mathbf{GRS}_k(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \{(y_1 f(x_1), \dots, y_n f(x_n)) \mid f \in \mathbb{F}_q[x], \deg(f) < k\}.$$

REED-SOLOMON CODES

DEFINITION 6 (GENERALIZED REED-SOLOMON CODES)

Let $(\mathbf{x}, \mathbf{y}) \in \mathbb{F}_{2^m}^n \times \mathbb{F}_{2^m}^n$ be a pair such that $\forall i, y_i \neq 0$ and $\forall i \neq j, x_i \neq x_j$.

$$\mathbf{GRS}_k(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \{(y_1 f(x_1), \dots, y_n f(x_n)) \mid f \in \mathbb{F}_q[x], \deg(f) < k\}.$$

$$\mathbf{G}_{\mathbf{GRS}_k(\mathbf{x}, \mathbf{y})} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{k-1} & x_2^{k-1} & \dots & x_n^{k-1} \end{pmatrix} \begin{pmatrix} y_1 & & & \\ & y_2 & & 0 \\ & & \ddots & \\ 0 & & & y_n \end{pmatrix}.$$

REED-SOLOMON CODES

$$\mathbf{GRS}_k(\mathbf{x}, \mathbf{y})^\perp = \mathbf{GRS}_{n-k}(\mathbf{x}, \mathbf{z}), \quad \mathbf{H}_{\mathbf{GRS}_{n-1}(\mathbf{x}, \mathbf{y})} \mathbf{z}^T = 0$$

REED-SOLOMON CODES

$$\mathbf{GRS}_k(\mathbf{x}, \mathbf{y})^\perp = \mathbf{GRS}_{n-k}(\mathbf{x}, \mathbf{z}), \quad \mathbf{H}_{\mathbf{GRS}_{n-1}(\mathbf{x}, \mathbf{y})} \mathbf{z}^T = 0$$

$$\mathbf{GRS}_k(\mathbf{x}, \mathbf{y})^2 = \mathbf{GRS}_{2k-1}(\mathbf{x}, \mathbf{y}^2)$$

REED-SOLOMON CODES

$$\mathbf{GRS}_k(\mathbf{x}, \mathbf{y})^\perp = \mathbf{GRS}_{n-k}(\mathbf{x}, \mathbf{z}), \quad \mathbf{H}_{\mathbf{GRS}_{n-1}(\mathbf{x}, \mathbf{y})} \mathbf{z}^T = 0$$

$$\mathbf{GRS}_k(\mathbf{x}, \mathbf{y})^2 = \mathbf{GRS}_{2k-1}(\mathbf{x}, \mathbf{y}^2)$$

$$3 \leq k \leq \frac{n+1}{2}, \quad \text{Dim}(\mathbf{GRS}_k(\mathbf{x}, \mathbf{y})^2) = 2k - 1 < \binom{k+1}{2}$$

REED-MULLER CODES

$$\mathcal{R}(r, m) \stackrel{\text{def}}{=} \left\{ (g(v_1, \dots, v_m))_{(v_1, \dots, v_m) \in \mathbb{F}_2^m} \mid g \in \mathbb{F}_2[x_1, \dots, x_m], \deg g \leq r \right\}.$$

REED-MULLER CODES

$$\mathcal{R}(r, m) \stackrel{\text{def}}{=} \left\{ (g(v_1, \dots, v_m))_{(v_1, \dots, v_m) \in \mathbb{F}_2^m} \mid g \in \mathbb{F}_2[x_1, \dots, x_m], \deg g \leq r \right\}.$$

$$\text{Dim}(\mathcal{R}(r, m)) = \sum_{i=0}^r \binom{m}{i}$$

REED-MULLER

$$\mathcal{R}(r, m)^\perp = \mathcal{R}(m - r - 1, m)$$

REED-MULLER

$$\mathcal{R}(r, m)^\perp = \mathcal{R}(m - r - 1, m)$$

$$\mathcal{R}(r, m)^2 = \mathcal{R}(2r, m)$$

REED-MULLER

$$\mathcal{R}(r, m)^\perp = \mathcal{R}(m - r - 1, m)$$

$$\mathcal{R}(r, m)^2 = \mathcal{R}(2r, m)$$

$$\text{Dim}(\mathcal{R}(r, m)^2) = \sum_{i=0}^{2r} \binom{m}{i} < \binom{\sum_{i=0}^r \binom{m}{i} + 1}{2}.$$

ALTERNANT AND GOPPA CODES

$$\mathbf{Alt}_r(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \mathbf{GRS}_r(\mathbf{x}, \mathbf{y})^\perp \cap \mathbb{F}_2^n.$$

ALTERNANT AND GOPPA CODES

$$\mathbf{Alt}_r(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \mathbf{GRS}_r(\mathbf{x}, \mathbf{y})^\perp \cap \mathbb{F}_2^n.$$

$$\Gamma(\mathbf{x}, g) \stackrel{\text{def}}{=} \mathbf{Alt}_t(\mathbf{x}, \mathbf{y}), \text{ where } y_i = \frac{1}{g(x_i)}, g \in \mathbb{F}_{2^m}[x], \deg g = t$$

ALTERNANT AND GOPPA CODES

$$\mathbf{Alt}_r(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \mathbf{GRS}_r(\mathbf{x}, \mathbf{y})^\perp \cap \mathbb{F}_2^n.$$

$$\Gamma(\mathbf{x}, g) \stackrel{\text{def}}{=} \mathbf{Alt}_t(\mathbf{x}, \mathbf{y}), \text{ where } y_i = \frac{1}{g(x_i)}, g \in \mathbb{F}_{2^m}[x], \deg g = t$$

$$\mathbf{Alt}_r(\mathbf{x}, \mathbf{y})^2 = ???^9$$

BINARY GOPPA CODES

WANTED FOR CRYPTO RESILIENCE

DEFINITION 7 (BINARY GOPPA CODES)

Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}_{2^m}^n$ with $x_i \neq x_j$,

$g \in \mathbb{F}_{2^m}[x]$ with $\deg(g) = t$ s.t. $\forall 1 \leq i \leq n, g(x_i) \neq 0$.

$\forall \mathbf{c} \in \mathbb{F}_2^n$ define the rational function $s_{\mathbf{c}}(x) \stackrel{\text{def}}{=} \sum_{i=1}^n \frac{c_i}{x-x_i}$.

The binary Goppa code is

$$\Gamma(\mathbf{x}, g) \stackrel{\text{def}}{=} \{\mathbf{c} \in \mathbb{F}_2^n \mid s_{\mathbf{c}}(x) \equiv 0 \pmod{g(x)}\}.$$

PATTERSON ALGORITHM

- If $\mathbf{y} = \mathbf{c} + \mathbf{e}$ then

$$s_{\mathbf{y}}(x) = \sum_{i=0}^n \frac{c_i + e_i}{x - x_i} \equiv s_{\mathbf{e}}(x) \pmod{g(x)}$$

PATTERSON ALGORITHM

- If $\mathbf{y} = \mathbf{c} + \mathbf{e}$ then

$$s_{\mathbf{y}}(x) = \sum_{i=0}^n \frac{c_i + e_i}{x - x_i} \equiv s_{\mathbf{e}}(x) \pmod{g(x)}$$

- This implies

$$s_{\mathbf{y}}(x) \equiv \sum_{i \in \text{supp}(\mathbf{e})} \frac{1}{x - x_i} \pmod{g(x)}$$

- $\sigma(x)$ is called the error-locator polynomial : $\sigma(x) = \prod_{i \in \text{supp}(\mathbf{e})} (x - x_i)$.

PATTERSON ALGORITHM

$$\begin{aligned}\sigma(x)' &= \sum_{i \in \text{supp}(e)}^n \prod_{j \in \text{supp}(e), j \neq i} (x - x_j) \\ &= \sum_{i \in \text{supp}(e)}^n \frac{1}{x - x_i} \prod_{i \in \text{supp}(e)} (x - x_i) \\ &= \sigma(x) \sum_{i \in \text{supp}(e)}^n \frac{1}{x - x_i} \\ \sigma'(x) &\equiv \sigma(x) s_y(x) \pmod{g(x)}.\end{aligned}$$

PATTERSON ALGORITHM

- Let $\sigma(x) = a(x)^2 + xb(x)^2$ ($\deg(a) \leq (t-1)/2$, $\deg(b) \leq t/2$).
- This implies $\sigma(x)' = b(x)^2$ (over \mathbb{F}_2), which makes

$$b^2 = \sigma' = \sigma s_y = (a^2 + xb^2)s_y \pmod{g}$$

- Since s_y, g coprime, we have

$$a^2 = b^2 \sqrt{x + s_y^{-1}} \pmod{g}.$$

- Find $a(x), b(x)$ using Extended Euclidean Algorithm and compute $\sigma(x)$.

PATTERSON ALGORITHM

Input : The syndrome polynomial $s_s(x)$ and the Goppa code $g(x)$.

Output : The error vector \mathbf{e}

① $s_s(x)^{-1} \leftarrow \text{EEA}(g(x), s_s(x))$

② $\tau(x) \leftarrow \sqrt{x + s_s(x)^{-1}}$

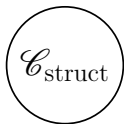
③ $a(x), b(x) \leftarrow \text{EEA}(g(x), \tau(x))$ s.t. $b(x)\tau(x) \equiv a(x) \pmod{g(x)}$

④ $\sigma(x) \leftarrow a^2(x) + xb^2(x)$

⑤ $\mathbf{e} \leftarrow (\sigma(x_1), \dots, \sigma(x_n)) \oplus (1, \dots, 1);$

McEliece and Niederreiter Summary Perspectives

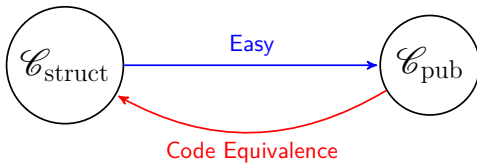
SUMMARY



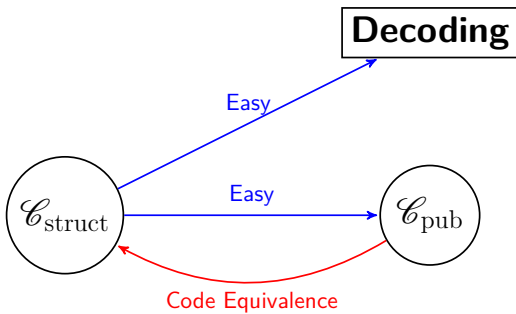
SUMMARY



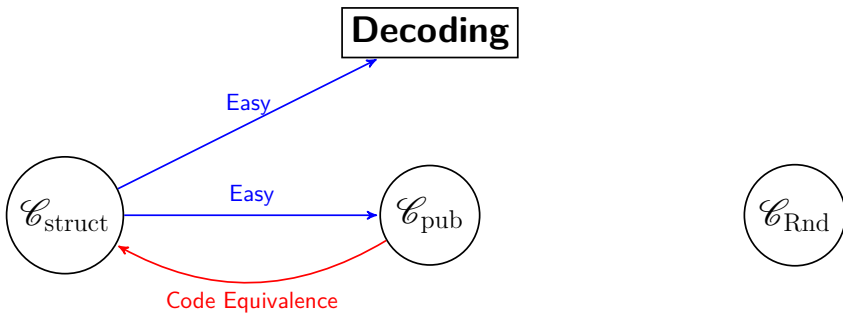
SUMMARY



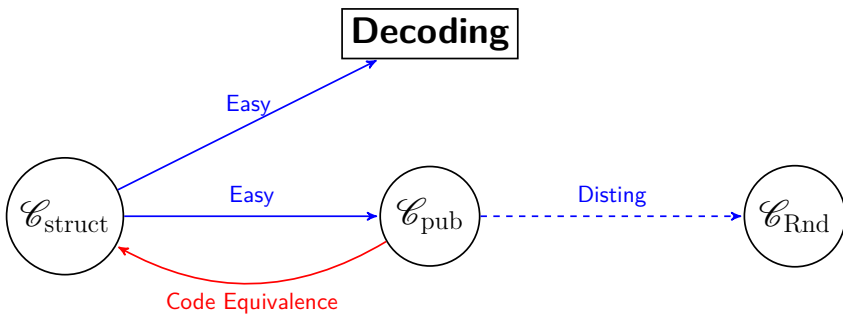
SUMMARY



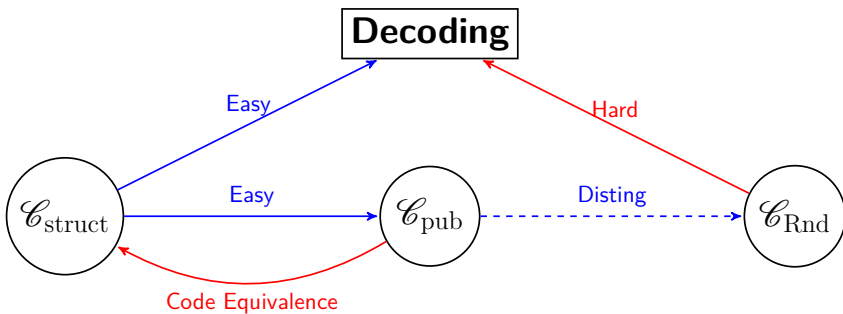
SUMMARY



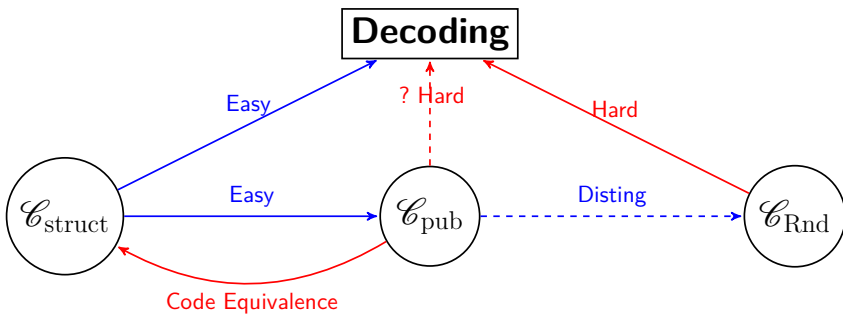
SUMMARY



SUMMARY



SUMMARY



OTHER CONSTRUCTIONS

ALEKHNOVICH'S CRYPTOSYSTEMS

- Underlying problem : distinguish a random vector from an erroneous codeword of a random code \mathcal{C} .

OTHER CONSTRUCTIONS

ALEKHNOVICH'S CRYPTOSYSTEMS

- Underlying problem : distinguish a random vector from an erroneous codeword of a random code \mathcal{C} .
- The public key is a random code while the private key is an error vector.

OTHER CONSTRUCTIONS

ALEKHNOVICH'S CRYPTOSYSTEMS

- Underlying problem : distinguish a random vector from an erroneous codeword of a random code \mathcal{C} .
- The public key is a random code while the private key is an error vector.
- Decryption is probabilistic

ALEKHNOVICH'S CRYPTOSYSTEMS

- *Key Generation*

ALEKHNOVICH'S CRYPTOSYSTEMS

- *Key Generation*

- 1 Chose a random matrix $\mathbf{A} \in \mathcal{M}_{k,n}(\mathbb{F}_2)$

ALEKHNOVICH'S CRYPTOSYSTEMS

- *Key Generation*

- 1 Chose a random matrix $\mathbf{A} \in \mathcal{M}_{k,n}(\mathbb{F}_2)$
- 2 Choose $\mathbf{e} \in \mathbb{F}_2^n$ at random of weight t

ALEKHNOVICH'S CRYPTOSYSTEMS

- *Key Generation*

- ① Choose a random matrix $\mathbf{A} \in \mathcal{M}_{k,n}(\mathbb{F}_2)$
- ② Choose $\mathbf{e} \in \mathbb{F}_2^n$ at random of weight t
- ③ Choose $\mathbf{x} \in \mathbb{F}_2^k$ at random

ALEKHNOVICH'S CRYPTOSYSTEMS

- *Key Generation*

- 1 Choose a random matrix $\mathbf{A} \in \mathcal{M}_{k,n}(\mathbb{F}_2)$

- 2 Choose $\mathbf{e} \in \mathbb{F}_2^n$ at random of weight t

- 3 Choose $\mathbf{x} \in \mathbb{F}_2^k$ at random

- 4 Compute $\mathbf{y} = \mathbf{x}\mathbf{A} + \mathbf{e}$ and $\mathbf{H} = \begin{pmatrix} \mathbf{A} \\ \mathbf{y} \end{pmatrix}$

ALEKHNOVICH'S CRYPTOSYSTEMS

- *Key Generation*

- 1 Choose a random matrix $\mathbf{A} \in \mathcal{M}_{k,n}(\mathbb{F}_2)$
- 2 Choose $\mathbf{e} \in \mathbb{F}_2^n$ at random of weight t
- 3 Choose $\mathbf{x} \in \mathbb{F}_2^k$ at random
- 4 Compute $\mathbf{y} = \mathbf{x}\mathbf{A} + \mathbf{e}$ and $\mathbf{H} = \begin{pmatrix} \mathbf{A} \\ \mathbf{y} \end{pmatrix}$
- 5 Choose \mathbf{G} a generator matrix for $\mathcal{C} = \ker(\mathbf{H})$.

ALEKHNOVICH'S CRYPTOSYSTEMS

- *Key Generation*

- 1 Choose a random matrix $\mathbf{A} \in \mathcal{M}_{k,n}(\mathbb{F}_2)$

- 2 Choose $\mathbf{e} \in \mathbb{F}_2^n$ at random of weight t

- 3 Choose $\mathbf{x} \in \mathbb{F}_2^k$ at random

- 4 Compute $\mathbf{y} = \mathbf{x}\mathbf{A} + \mathbf{e}$ and $\mathbf{H} = \begin{pmatrix} \mathbf{A} \\ \mathbf{y} \end{pmatrix}$

- 5 Choose \mathbf{G} a generator matrix for $\mathcal{C} = \ker(\mathbf{H})$.

- 6 The private key $\text{sk} = (\mathbf{e})$ and the public key $\text{pk} = (\mathbf{G}, t)$

ALEKHNOVICH'S CRYPTOSYSTEM

ENCRYPTION

Let $\mathbf{m} \in \mathbb{F}_2$,

- 1 If $\mathbf{m} = 0$ then
 - ▶ choose $\mathbf{a} \in \mathbb{F}_2^{n-k}$
 - ▶ choose $\mathbf{e}' \in \mathbb{F}_2^n$ of weight t
 - ▶ send $\mathbf{c} = \mathbf{a}\mathbf{G} + \mathbf{e}'$
- 2 If $\mathbf{m} = 1$ then send a random vector $\mathbf{c} \in \mathbb{F}_2^n$

ALEKHNOVICH'S CRYPTOSYSTEM

ENCRYPTION

Let $\mathbf{m} \in \mathbb{F}_2$,

- 1 If $\mathbf{m} = 0$ then
 - ▶ choose $\mathbf{a} \in \mathbb{F}_2^{n-k}$
 - ▶ choose $\mathbf{e}' \in \mathbb{F}_2^n$ of weight t
 - ▶ send $\mathbf{c} = \mathbf{aG} + \mathbf{e}'$
- 2 If $\mathbf{m} = 1$ then send a random vector $\mathbf{c} \in \mathbb{F}_2^n$

DECRYPTION

- 1 Compute $\mathbf{b} = \langle \mathbf{e}, \mathbf{c} \rangle$
- 2 If $\mathbf{m} = 0$ then $\mathbf{b} = 0$ w.h.p.
- 3 If $\mathbf{m} = 1$ then $\mathbf{b} = 1$ w.p. $1/2$

$$\mathbf{b} = \langle \mathbf{e}, \mathbf{aG} \rangle + \langle \mathbf{e}, \mathbf{e}' \rangle = \langle \mathbf{e}, \mathbf{e}' \rangle$$

Questions