Where did my RAM go? Using algebraic cryptanalysis in practice (modelling exercises)

Lars Ran

Monika Trimoska

Summer school on RWC and privacy July 1, Dubrovnik, Croatia



Algebraic cryptanalysis



A type of cryptanalytic methods where the problem of finding the secret key (or any attack goal) is reduced to the problem of finding a solution to a nonlinear multivariate polynomial system of equations.



Trivium

Initialisation:

$$(s_1, ..., s_{93}) \leftarrow (K_1, ..., K_{80}, 0, ..., 0)$$

 $(s_{94}, ..., s_{177}) \leftarrow (IV_1, ..., IV_{80}, 0, ..., 0)$
 $(s_{178}, ..., s_{288}) \leftarrow (0, ..., 0, 1, 1, 1)$

Algorithm 8.1 Trivium's iterative function for keystream generation.

Input: The number of bits to be generated, denoted Z.

Output: Keystream vector z.

```
1: for i = 1 to Z do
           t_1 \leftarrow s_{66} + s_{93}
        t_2 \leftarrow s_{162} + s_{177}
        t_3 \leftarrow s_{243} + s_{288}
        z_i \leftarrow t_1 + t_2 + t_3
                                                                                                                                     Iterate for 1155 rounds
           t_1 \leftarrow t_1 + s_{91} \cdot s_{92} + s_{171}
                                                                                                                                  without producing any
           t_2 \leftarrow t_2 + s_{175} \cdot s_{176} + s_{264}
                                                                                                                                     output
           t_3 \leftarrow t_3 + s_{286} \cdot s_{287} + s_{69}
            (s_1, s_2, \ldots, s_{93}) \leftarrow (t_3, s_1, \ldots, s_{92})
            (s_{94}, s_{95}, \ldots, s_{177}) \leftarrow (t_1, s_{94}, \ldots, s_{176})
10:
            (s_{178}, s_{179}, \ldots, s_{288}) \leftarrow (t_2, s_{178}, \ldots, s_{287})
12: end for
```

Trivium

Keystream generation:

```
Algorithm 8.1 Trivium's iterative function for keystream generation.
```

Input: The number of bits to be generated, denoted Z.

Output: Keystream vector z.

```
1: for i = 1 to Z do

2: t_1 \leftarrow s_{66} + s_{93}

3: t_2 \leftarrow s_{162} + s_{177}

4: t_3 \leftarrow s_{243} + s_{288}

5: z_i \leftarrow t_1 + t_2 + t_3

6: t_1 \leftarrow t_1 + s_{91} \cdot s_{92} + s_{171}

7: t_2 \leftarrow t_2 + s_{175} \cdot s_{176} + s_{264}

8: t_3 \leftarrow t_3 + s_{286} \cdot s_{287} + s_{69}

9: (s_1, s_2 \dots, s_{93}) \leftarrow (t_3, s_1, \dots, s_{92})

10: (s_{94}, s_{95} \dots, s_{177}) \leftarrow (t_1, s_{94}, \dots, s_{176})

11: (s_{178}, s_{179} \dots, s_{288}) \leftarrow (t_2, s_{178}, \dots, s_{287})

12: end for
```





• Central map:

$$f:(x_1,...,x_n) \in \mathbb{F}_q^n \to (f^{(1)}(x_1,...,x_n),...,f^{(m)}(x_1,...,x_n)) \in \mathbb{F}_q^m$$

- Two bijective linear (or affine) transformations: $\mathbf{S} \in \mathrm{GL}_n(\mathbb{F}_q)$ and $\mathbf{T} \in \mathrm{GL}_m(\mathbb{F}_q)$
- Public map: $p = \mathbf{T} \circ f \circ \mathbf{S}$

• Central map:

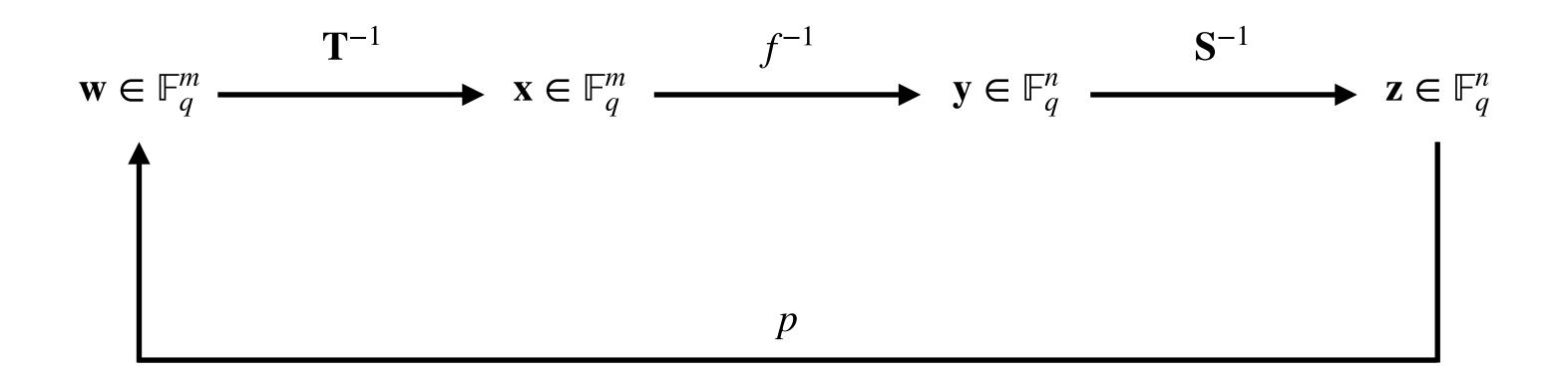
$$f: (x_1, ..., x_n) \in \mathbb{F}_q^n \to (f^{(1)}(x_1, ..., x_n), ..., f^{(m)}(x_1, ..., x_n)) \in \mathbb{F}_q^m$$

- Two bijective linear (or affine) transformations: $\mathbf{S} \in \mathrm{GL}_n(\mathbb{F}_q)$ and $\mathbf{T} \in \mathrm{GL}_m(\mathbb{F}_q)$
- Public map: $p = \mathbf{T} \circ f \circ \mathbf{S}$



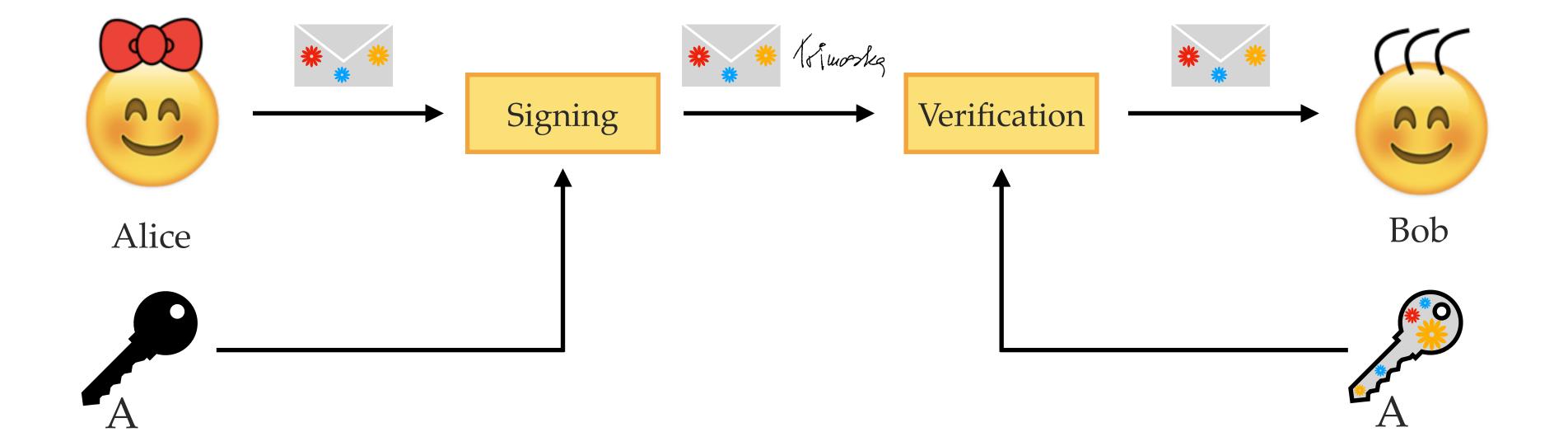
Main idea:

- The central map has a structure such that it is easy to find preimages: it is easy (polynomial time) to compute $f^{-1}(\mathbf{x})$ for a target vector \mathbf{x} .
- The linear transformations hide the structure of the central map.

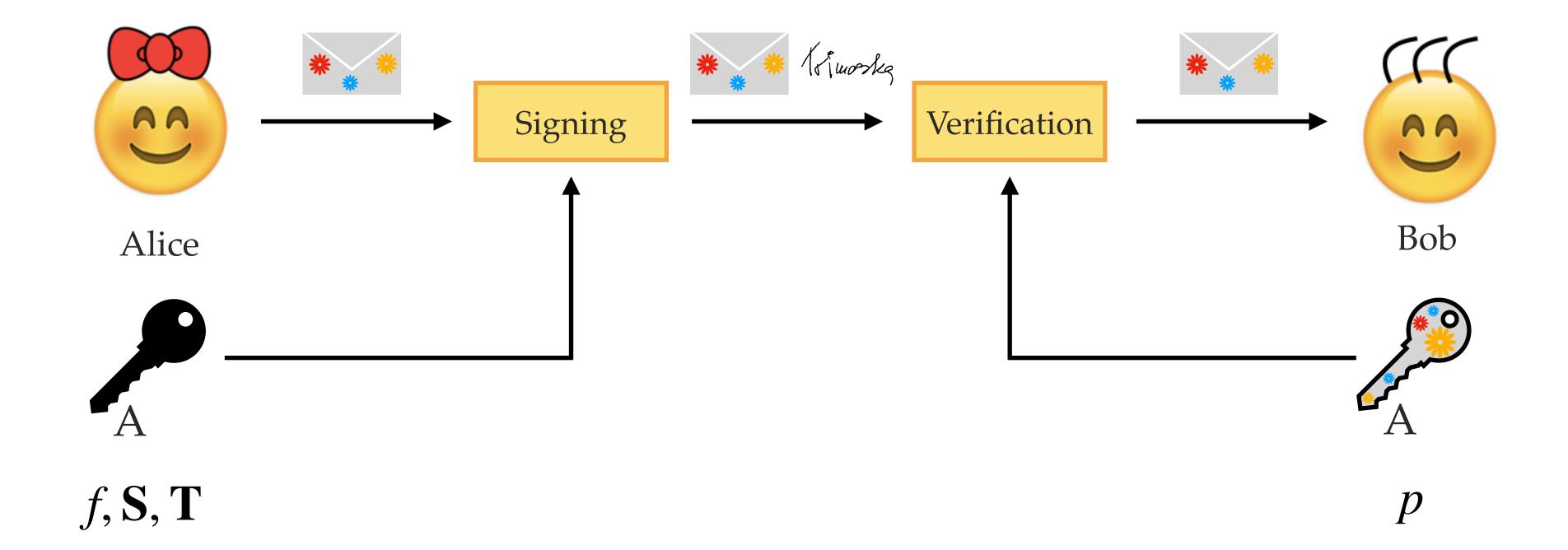


General workflow

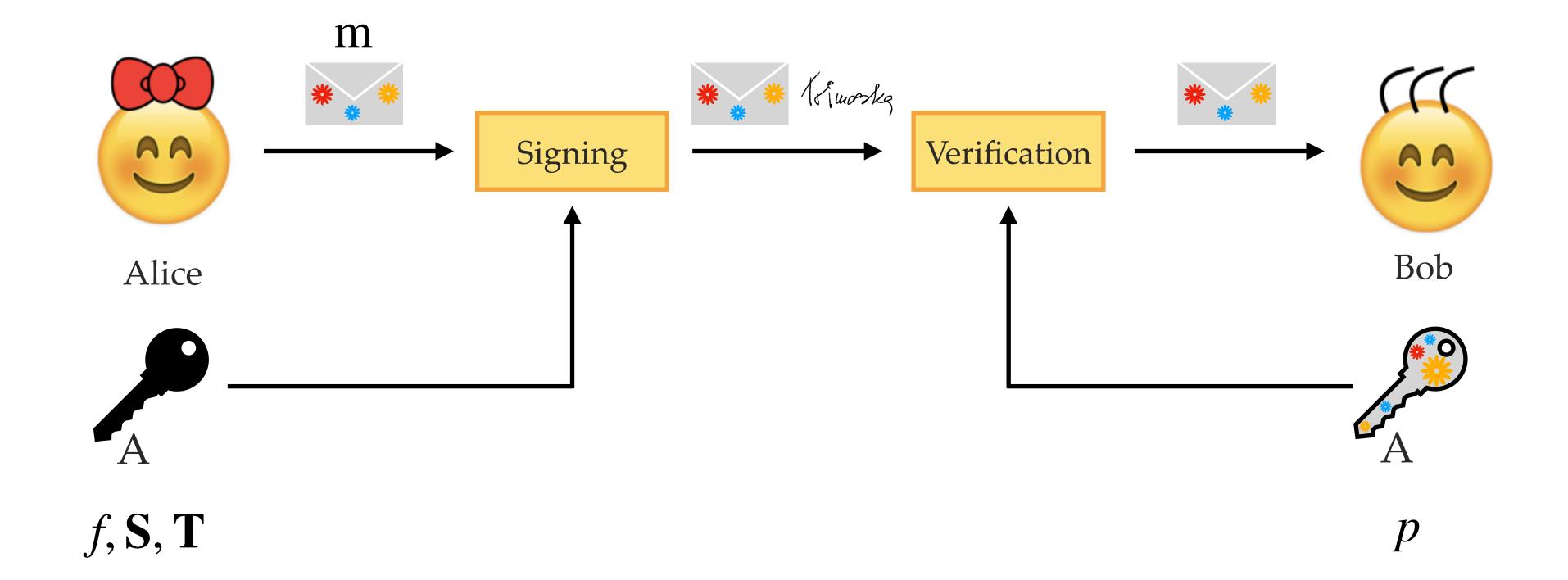




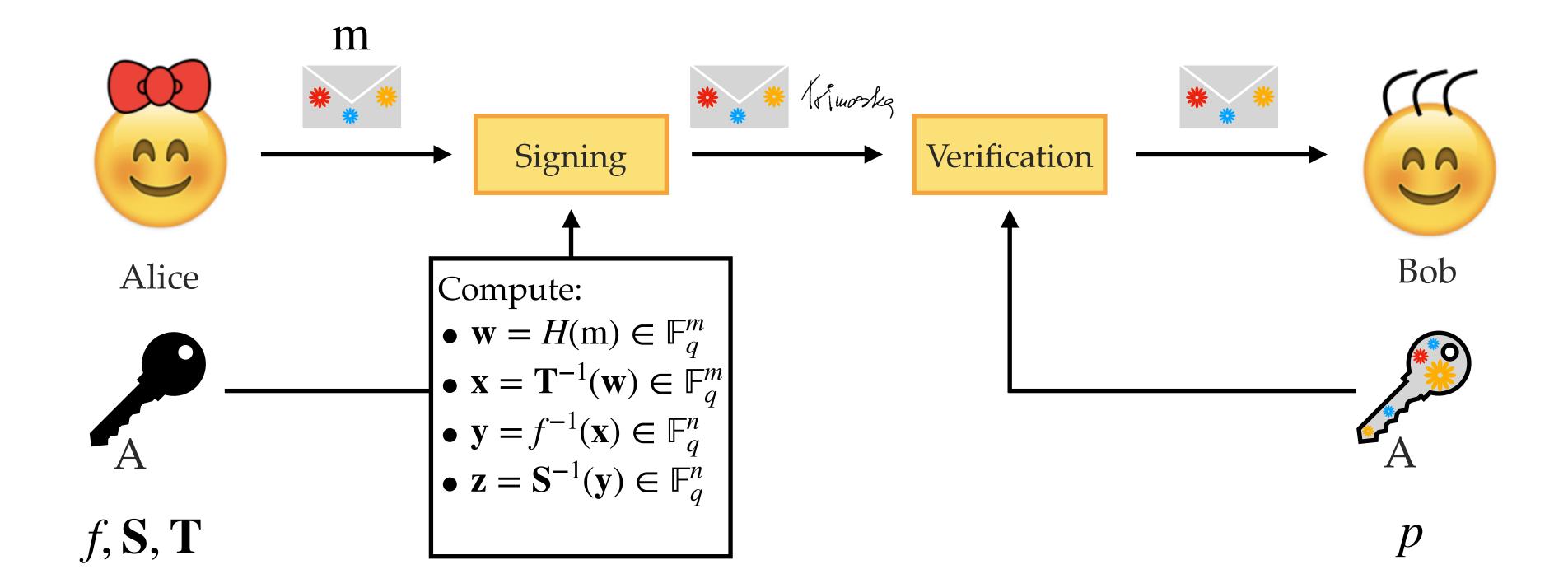




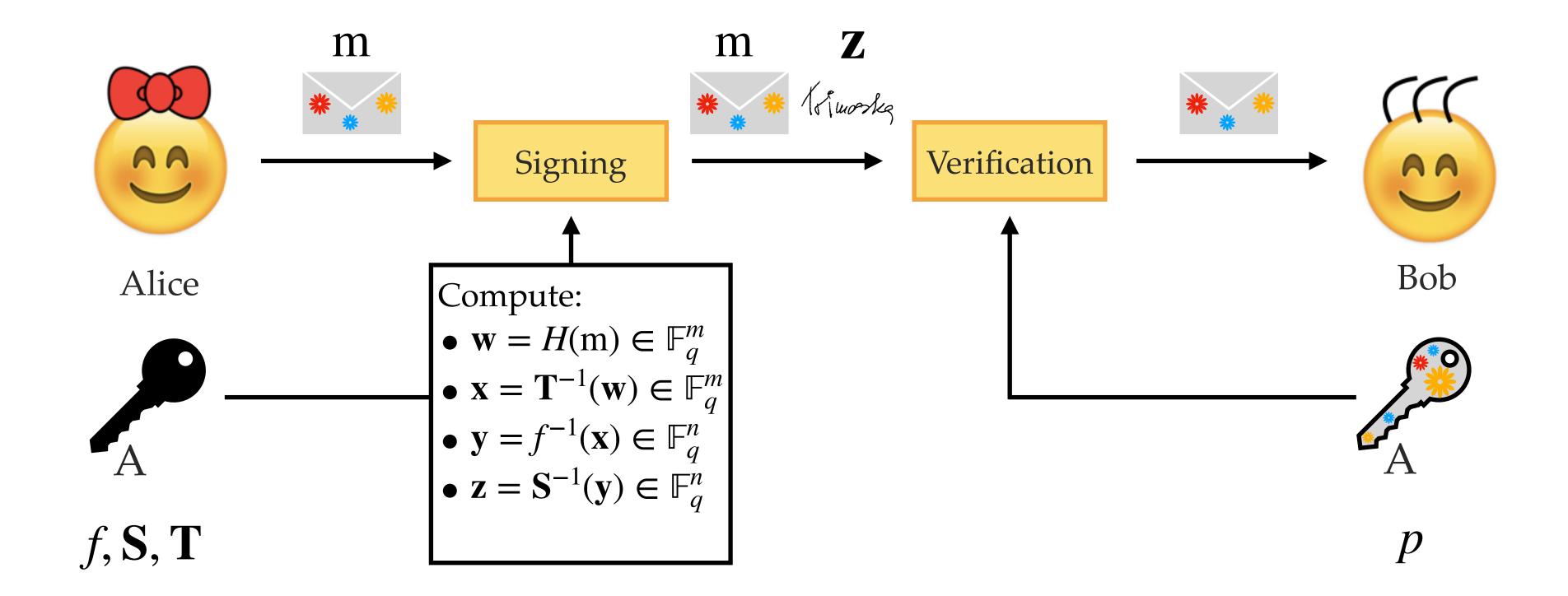




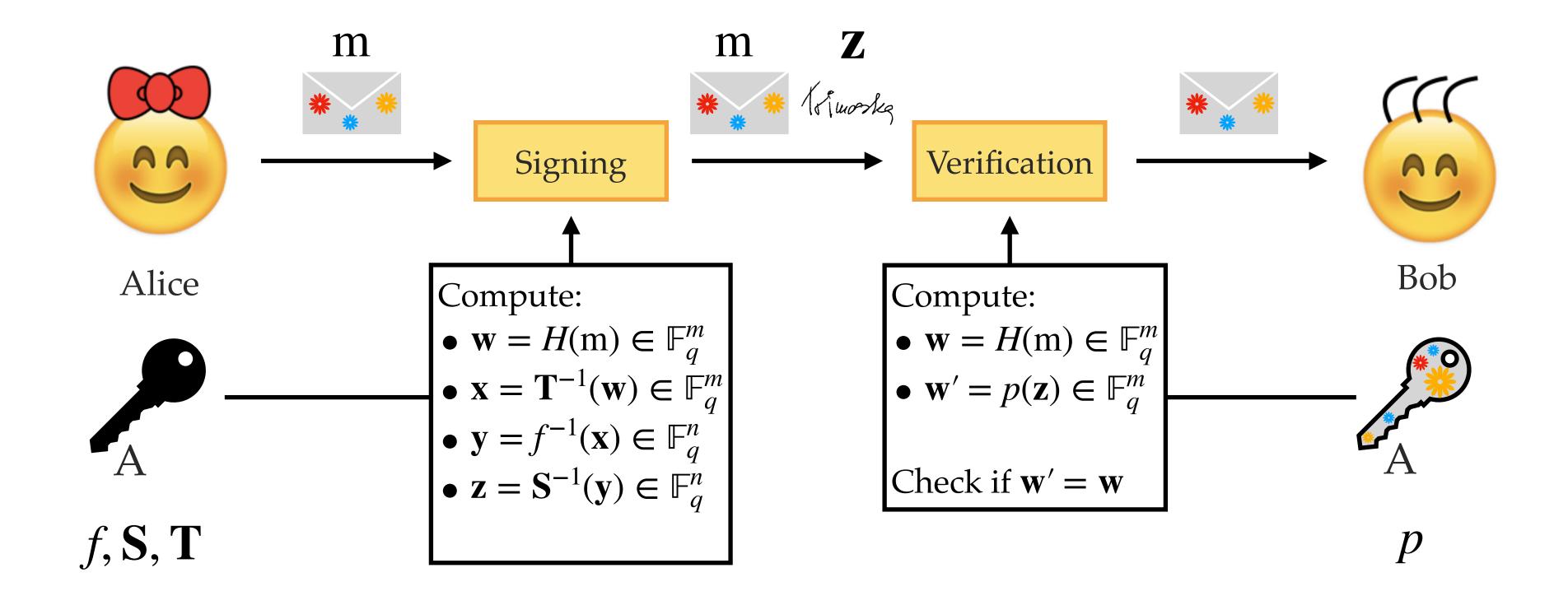




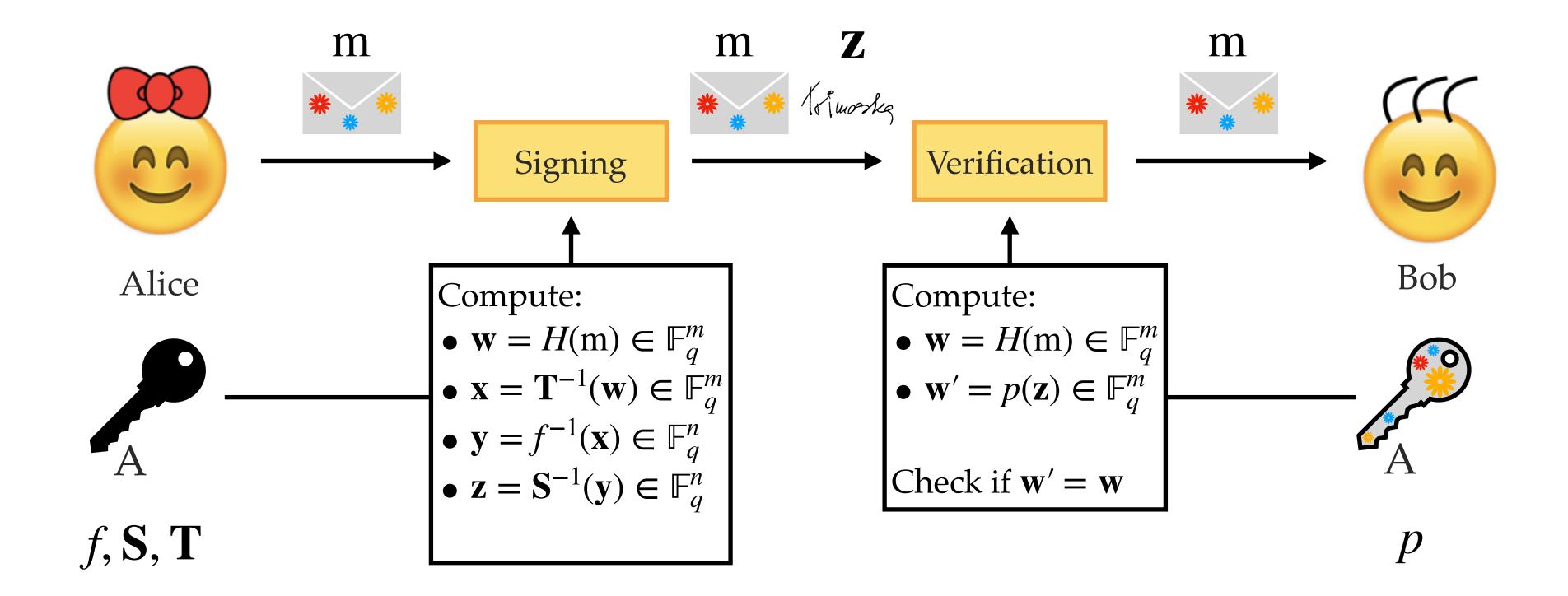














The UOV central map



Unbalanced Oil and Vinegar [Kipnis, Patarin, Goubin, '99]

$$f^{(k)}(x_1, \dots, x_n) = \sum_{i \in V, j \in V} \gamma_{ij}^{(k)} x_i x_j + \sum_{i \in V, j \in O} \gamma_{ij}^{(k)} x_i x_j + \sum_{i=1}^n \beta_i^{(k)} x_i + \alpha^{(k)}$$

Index set of vinegar variables: $V = \{1, ..., v\}$

Index set of oil variables: $O = \{v + 1, ..., n\}$



The UOV central map



Unbalanced Oil and Vinegar [Kipnis, Patarin, Goubin, '99]

Index set of vinegar variables: $V = \{1, ..., v\}$

$$f^{(k)}(x_1, \dots, x_n) = \sum_{i \in V, j \in V} \gamma_{ij}^{(k)} x_i x_j + \sum_{i \in V, j \in O} \gamma_{ij}^{(k)} x_i x_j + \sum_{i=1}^n \beta_i^{(k)} x_i + \alpha^{(k)}$$

Index set of oil variables: $O = \{v + 1, ..., n\}$

The central map is constructed in such a way that enumerating all of the vinegar variables leaves us with a linear system in the oil variables (oil does not mix with oil).



The UOV central map



Unbalanced Oil and Vinegar [Kipnis, Patarin, Goubin, '99]

$$f^{(k)}(x_1, \dots, x_n) = \sum_{i \in V, j \in V} \gamma_{ij}^{(k)} x_i x_j + \sum_{i \in V, j \in O} \gamma_{ij}^{(k)} x_i x_j + \sum_{i=1}^n \beta_i^{(k)} x_i + \alpha^{(k)}$$

Index set of vinegar variables: $V = \{1, ..., v\}$

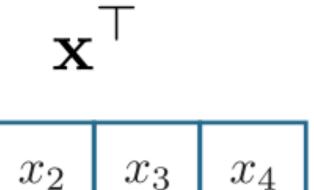
Index set of oil variables: $O = \{v + 1, ..., n\}$

- The central map is constructed in such a way that enumerating all of the vinegar variables leaves us with a linear system in the oil variables (oil does not mix with oil).
- Everything is as described in the previous slides, except that we do not have a linear transformation on the output: T = I.



Matrix representation of quadratic forms

Quadratic form: $f(\mathbf{x}) = \sum \gamma_{ij} x_i x_j$



 x_1

 ${f F}$

$${f X}$$

 x_1 x_2

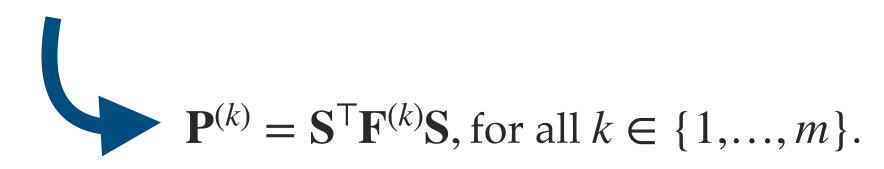
 x_4

 x_3

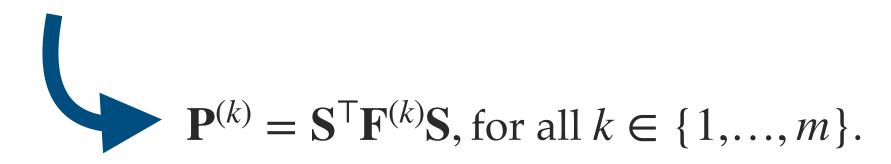
so with $\mathbf{x} = (x_1, ..., x_n)$, we get $\mathbf{x}^T \mathbf{F} \mathbf{x}$.



In matrix representation

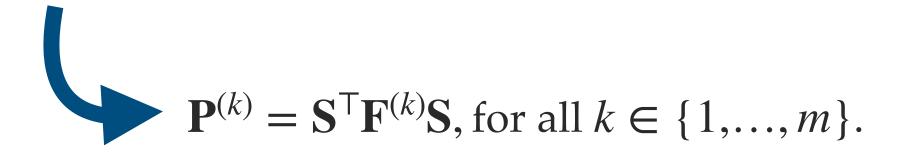


In matrix representation



Why?

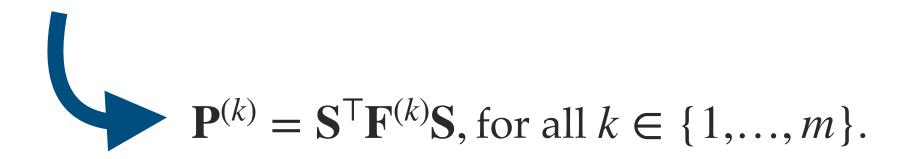
In matrix representation



Why?



In matrix representation



Why?

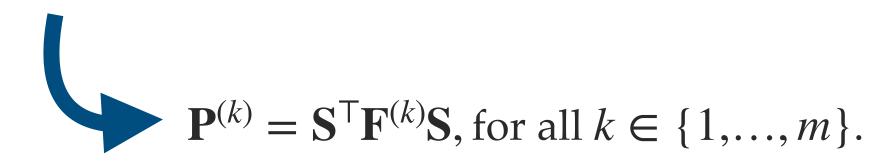


By definition, $p = f \circ S$.

In matrix representation, we need:

$$\mathbf{x}^{\mathsf{T}}\mathbf{P}^{(k)}\mathbf{x} = (\mathbf{S}\mathbf{x})^{\mathsf{T}}\mathbf{F}^{(k)}(\mathbf{S}\mathbf{x})$$

In matrix representation



Why?



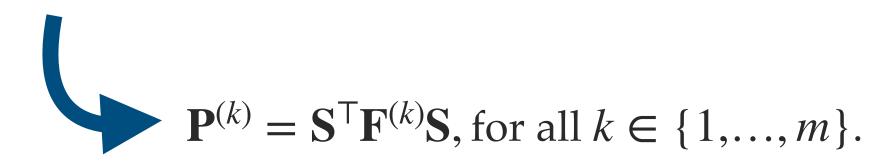
By definition, $p = f \circ S$.

In matrix representation, we need:

$$\mathbf{x}^{\mathsf{T}}\mathbf{P}^{(k)}\mathbf{x} = (\mathbf{S}\mathbf{x})^{\mathsf{T}}\mathbf{F}^{(k)}(\mathbf{S}\mathbf{x})$$

$$\mathbf{x}^{\mathsf{T}}\mathbf{P}^{(k)}\mathbf{x} = \mathbf{x}^{\mathsf{T}}\mathbf{S}^{\mathsf{T}}\mathbf{F}^{(k)}\mathbf{S}\mathbf{x}$$

In matrix representation



Why?

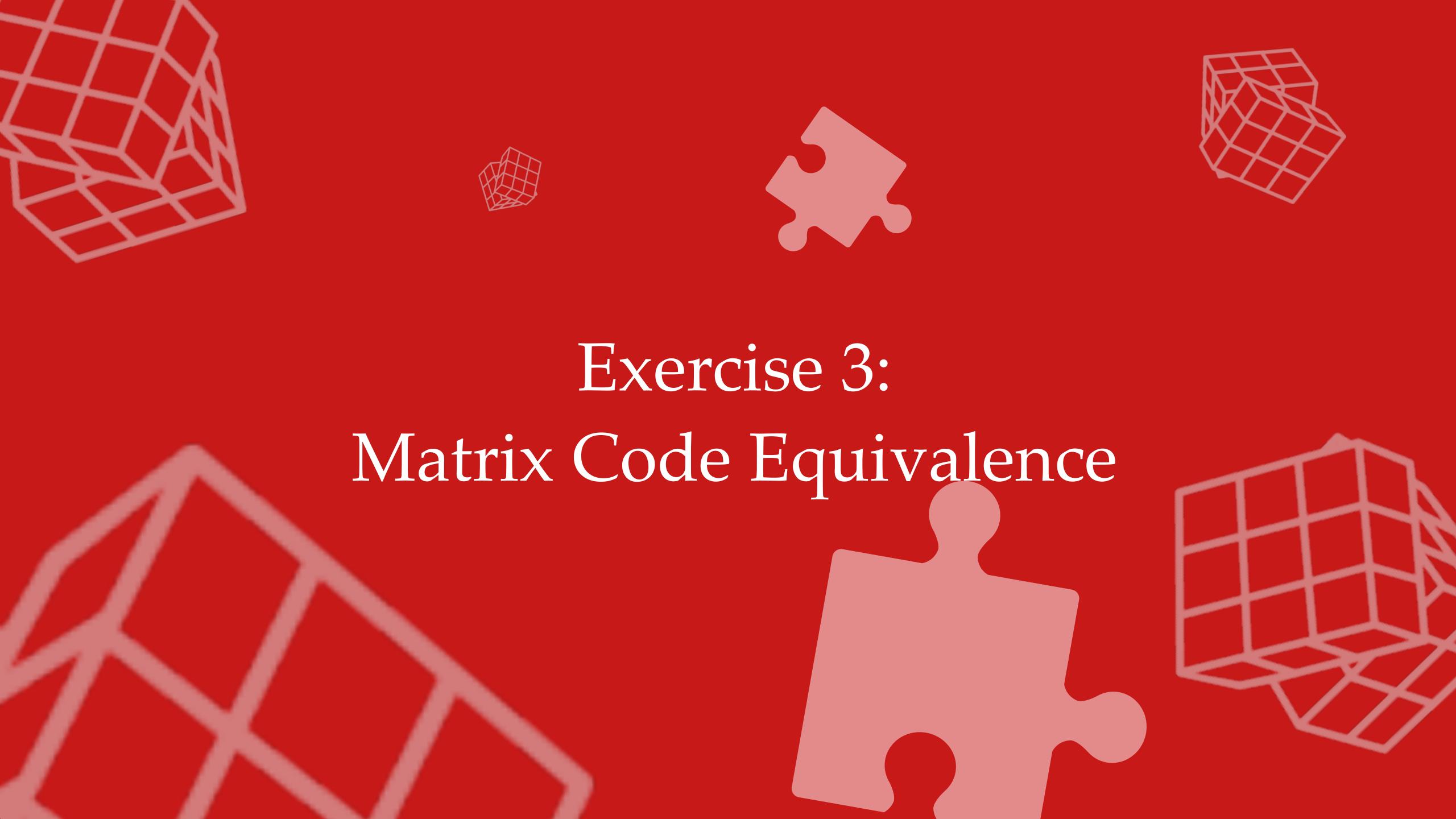


By definition, $p = f \circ S$.

In matrix representation, we need:

$$\mathbf{x}^{\mathsf{T}}\mathbf{P}^{(k)}\mathbf{x} = (\mathbf{S}\mathbf{x})^{\mathsf{T}}\mathbf{F}^{(k)}(\mathbf{S}\mathbf{x})$$

$$\mathbf{x}^{\mathsf{T}}\mathbf{P}^{(k)}\mathbf{x} = \mathbf{x}^{\mathsf{T}}\mathbf{S}^{\mathsf{T}}\mathbf{F}^{(k)}\mathbf{S}\mathbf{x}$$



Matrix code

A matrix code \mathscr{C} over \mathbb{F}_q is a k-dimensional \mathbb{F}_q -linear subspace of $\mathbb{F}_q^{m \times n}$.



Matrix code

A matrix code \mathscr{C} over \mathbb{F}_q is a k-dimensional \mathbb{F}_q -linear subspace of $\mathbb{F}_q^{m \times n}$.

Basis of a matrix code ----------

The basis of a matrix code \mathscr{C} is given by the k-tuple ($\mathbb{C}^{(1)},...,\mathbb{C}^{(k)}$).



Matrix code

A matrix code \mathscr{C} over \mathbb{F}_q is a k-dimensional \mathbb{F}_q -linear subspace of $\mathbb{F}_q^{m \times n}$.

Basis of a matrix code -----

The basis of a matrix code \mathscr{C} is given by the k-tuple ($\mathbb{C}^{(1)},...,\mathbb{C}^{(k)}$).

Rank metric

For $C \in \mathbb{F}_q^{m \times n}$, the rank weight of C is given by the rank of C, aka.

$$\operatorname{wt}(\mathbf{C}) = \operatorname{rk}(\mathbf{C}).$$

Example. q = 13, m = 4, n = 6, k = 5

$$\mathbf{C} = \lambda_1 \cdot \begin{pmatrix} 2 & 8 & 10 & 4 & 5 & 7 \\ 1 & 11 & 7 & 9 & 6 & 12 \\ 3 & 0 & 13 & 5 & 4 & 8 \\ 9 & 6 & 3 & 2 & 10 & 11 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 12 & 0 & 4 & 11 & 9 & 3 \\ 5 & 6 & 8 & 13 & 2 & 1 \\ 10 & 7 & 3 & 9 & 4 & 6 \\ 2 & 5 & 11 & 8 & 1 & 10 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 5 & 2 & 9 & 11 & 4 & 8 \\ 3 & 7 & 1 & 10 & 12 & 0 \\ 6 & 9 & 2 & 13 & 11 & 8 \\ 1 & 5 & 6 & 3 & 10 & 7 \end{pmatrix} + \lambda_4 \cdot \begin{pmatrix} 9 & 4 & 6 & 1 & 13 & 2 \\ 8 & 0 & 5 & 12 & 6 & 11 \\ 3 & 7 & 10 & 9 & 4 & 5 \\ 2 & 8 & 11 & 3 & 7 & 1 \end{pmatrix} + \lambda_5 \cdot \begin{pmatrix} 7 & 10 & 4 & 6 & 8 & 3 \\ 1 & 5 & 2 & 11 & 9 & 0 \\ 13 & 7 & 6 & 4 & 12 & 2 \\ 8 & 3 & 1 & 9 & 5 & 10 \end{pmatrix} \quad \lambda_i \in \mathbb{F}_q$$



Isometry -

An isometry (for our purposes) between two codes $\mathscr C$ and $\mathscr D$ is a linear map $\mu:\mathscr C\to\mathscr D$ that preserves the metric.



Isometry

An isometry (for our purposes) between two codes $\mathscr C$ and $\mathscr D$ is a linear map $\mu:\mathscr C\to\mathscr D$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.



Isometry -

An isometry (for our purposes) between two codes \mathscr{C} and \mathscr{D} is a linear map $\mu:\mathscr{C}\to\mathscr{D}$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

Which linear transformations preserve the rank?



Isometry

An isometry (for our purposes) between two codes $\mathscr C$ and $\mathscr D$ is a linear map $\mu:\mathscr C\to\mathscr D$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

Which linear transformations preserve the rank?

 \longrightarrow Multiply a codeword on the right by any $\mathbf{M} \in \mathbb{F}_q^{n \times r}$

Isometry

An isometry (for our purposes) between two codes \mathscr{C} and \mathscr{D} is a linear map $\mu:\mathscr{C}\to\mathscr{D}$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

Which linear transformations preserve the rank?

 \longrightarrow Multiply a codeword on the right by any $\mathbf{M} \in \mathbb{F}_q^{n \times r}$



Isometry

An isometry (for our purposes) between two codes \mathscr{C} and \mathscr{D} is a linear map $\mu:\mathscr{C}\to\mathscr{D}$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

Which linear transformations preserve the rank?

 \longrightarrow Multiply a codeword on the right by any $\mathbf{M} \in \mathbb{F}_q^{n \times r}$



▶ Multiply a codeword on the right by $\mathbf{B} \in \mathrm{GL}_n$

Isometry

An isometry (for our purposes) between two codes \mathscr{C} and \mathscr{D} is a linear map $\mu:\mathscr{C}\to\mathscr{D}$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

- \longrightarrow Multiply a codeword on the right by any $\mathbf{M} \in \mathbb{F}_q^{n \times r}$
- \longrightarrow Multiply a codeword on the right by $\mathbf{B} \in \mathrm{GL}_n$



Isometry

An isometry (for our purposes) between two codes \mathscr{C} and \mathscr{D} is a linear map $\mu:\mathscr{C}\to\mathscr{D}$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

- \longrightarrow Multiply a codeword on the right by any $\mathbf{M} \in \mathbb{F}_q^{n \times r}$
- \longrightarrow Multiply a codeword on the right by $\mathbf{B} \in \mathrm{GL}_n$
- \longrightarrow Multiply a codeword on the left by $\mathbf{A} \in \mathrm{GL}_m$



Isometry

An isometry (for our purposes) between two codes \mathscr{C} and \mathscr{D} is a linear map $\mu:\mathscr{C}\to\mathscr{D}$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

- \longrightarrow Multiply a codeword on the right by any $\mathbf{M} \in \mathbb{F}_q^{n \times r}$
- \longrightarrow Multiply a codeword on the right by $\mathbf{B} \in \mathrm{GL}_n$
- \longrightarrow Multiply a codeword on the left by $\mathbf{A} \in \mathrm{GL}_m$



Isometry

An isometry (for our purposes) between two codes \mathscr{C} and \mathscr{D} is a linear map $\mu:\mathscr{C}\to\mathscr{D}$ that preserves the metric.



In this case: an isometry preserves the rank weight of codewords.

- \longrightarrow Multiply a codeword on the right by any $\mathbf{M} \in \mathbb{F}_q^{n \times r}$
- \longrightarrow Multiply a codeword on the right by $\mathbf{B} \in \mathrm{GL}_n$
- \longrightarrow Multiply a codeword on the left by $\mathbf{A} \in \mathrm{GL}_m$
- Take the transposition of a codeword (only when m = n, does not make the equivalence problem harder)





The Matrix Code Equivalence (MCE) problem

Input: Two *k*-dimensional matrix codes \mathscr{C} , $\mathscr{D} \subset \mathbb{F}_q^{m \times n}$ for two matrix codes \mathscr{C} and \mathscr{D} . **Question:** Find - if any - a map (\mathbf{A}, \mathbf{B}) , where $\mathbf{A} \in \mathrm{GL}_m(\mathbb{F}_q)$ and $\mathbf{B} \in \mathrm{GL}_n(\mathbb{F}_q)$ such that for all $\mathbf{C} \in \mathscr{C}$, it holds that $\mathbf{ACB} \in \mathscr{D}$.



The MCE problem in matrix form

Let $(\mathbf{C}^{(1)},...,\mathbf{C}^{(k)})$ be a basis of code $\mathscr C$ and let $(\mathbf{D}^{(1)},...,\mathbf{D}^{(k)})$ be a basis of code $\mathscr D$. Find $\mathbf{A} \in \mathrm{GL}_m(\mathbb F_q)$, $\mathbf{B} \in \mathrm{GL}_n(\mathbb F_q)$ and $\mathbf{T} \in \mathrm{GL}_k(\mathbb F_q)$ such that

$$\mathbf{D}^{(i)} = \sum_{1 \le j \le k} t_{j,i} \mathbf{A} \mathbf{C}^{(j)} \mathbf{B}, \quad \forall 1 \le i \le k$$



The MCE problem in matrix form

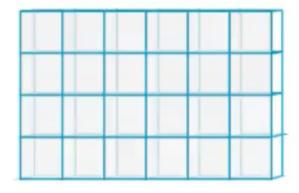
Let $(\mathbf{C}^{(1)},...,\mathbf{C}^{(k)})$ be a basis of code $\mathscr C$ and let $(\mathbf{D}^{(1)},...,\mathbf{D}^{(k)})$ be a basis of code $\mathscr D$. Find $\mathbf{A}\in \mathrm{GL}_m(\mathbb F_q)$, $\mathbf{B}\in \mathrm{GL}_n(\mathbb F_q)$ and $\mathbf{T}\in \mathrm{GL}_k(\mathbb F_q)$ such that

$$\mathbf{D}^{(i)} = \sum_{1 \le i \le k} t_{j,i} \mathbf{A} \mathbf{C}^{(j)} \mathbf{B}, \quad \forall 1 \le i \le k$$

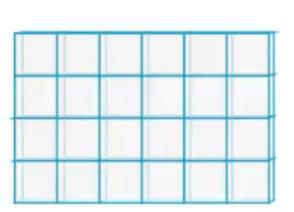
change of basis



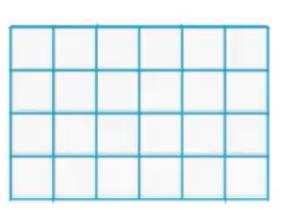
We can think of a matrix code as a 3-tensor over \mathbb{F}_q .



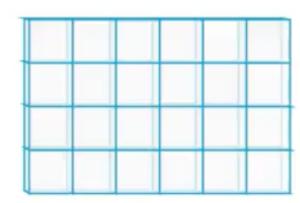




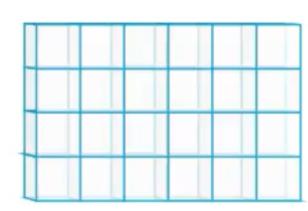
 \mathbb{C}_2



 \mathbb{J}_3



 \mathbb{C}_4



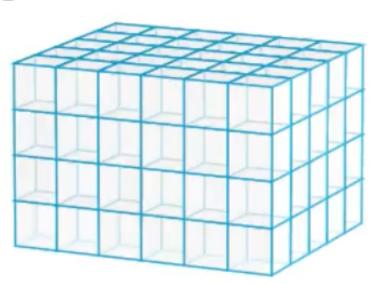
 \mathbf{C}_5





We can think of a matrix code as a 3-tensor over \mathbb{F}_q .

$$\mathcal{C} \subseteq \mathbb{F}_q^{m \times n \times k}$$

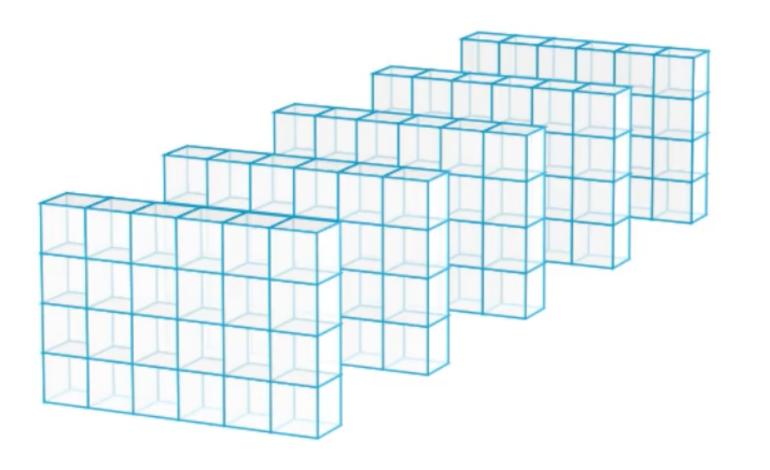






Viewed as a 3-tensor, we can see $\mathscr C$ from three directions

- a *k*-dimensional code in $\mathbb{F}_q^{m \times n}$
- an m-dimensional code in $\mathbb{F}_q^{n \times k}$
- an *n*-dimensional code in $\mathbb{F}_q^{m \times k}$

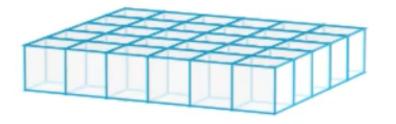


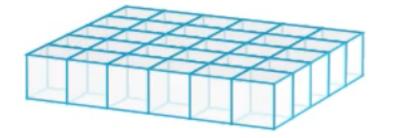


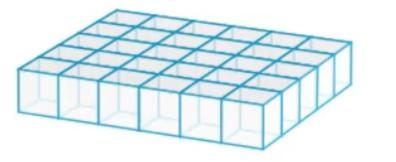
Viewed as a 3-tensor, we can see $\mathscr C$ from three directions

- a *k*-dimensional code in $\mathbb{F}_q^{m \times n}$
- an *m*-dimensional code in $\mathbb{F}_q^{n \times k}$
- an *n*-dimensional code in $\mathbb{F}_q^{m \times k}$







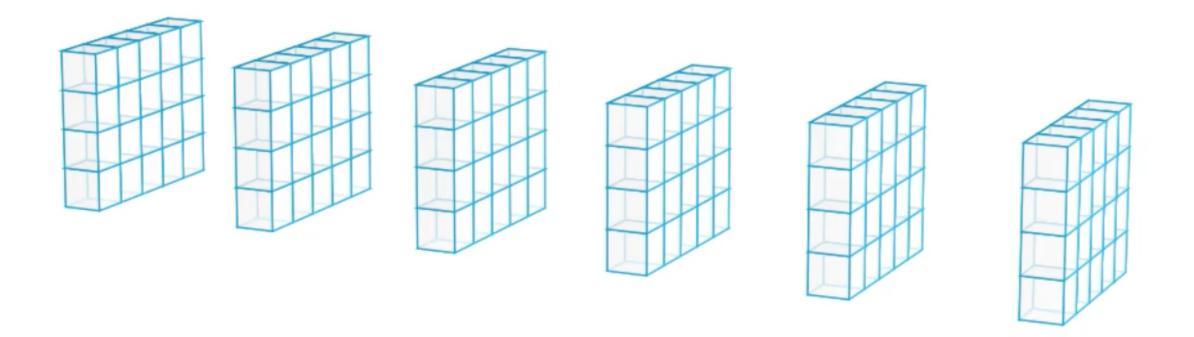




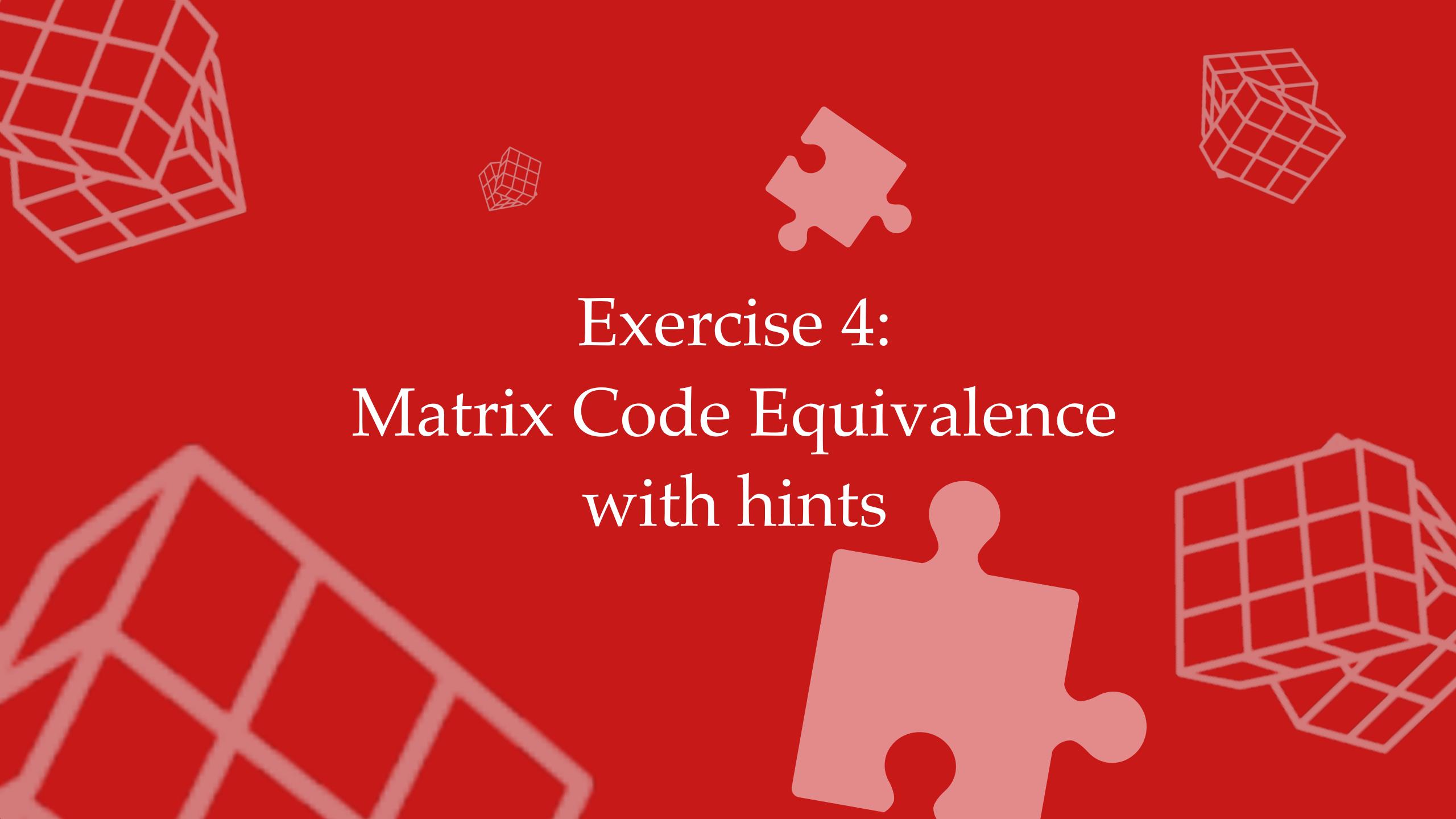


Viewed as a 3-tensor, we can see $\mathscr C$ from three directions

- a *k*-dimensional code in $\mathbb{F}_q^{m \times n}$
- an *m*-dimensional code in $\mathbb{F}_q^{n \times k}$
- an *n*-dimensional code in $\mathbb{F}_q^{m \times k}$







Collision



We have a collision when we know a codeword ${\bf C}$ in ${\mathscr C}$ that maps to a codeword ${\bf D}$ in ${\mathscr D}$.

$$\mathbf{D} = \mathbf{ACB}$$

Collision



We have a collision when we know a codeword ${\bf C}$ in ${\mathscr C}$ that maps to a codeword ${\bf D}$ in ${\mathscr D}$.

$$D = ACB$$

Recall how we can represent codewords with their coordinate vectors

$$\mathbf{C} = \lambda_1 \cdot \begin{pmatrix} 2 & 8 & 10 & 4 & 5 & 7 \\ 1 & 11 & 7 & 9 & 6 & 12 \\ 3 & 0 & 13 & 5 & 4 & 8 \\ 9 & 6 & 3 & 2 & 10 & 11 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 12 & 0 & 4 & 11 & 9 & 3 \\ 5 & 6 & 8 & 13 & 2 & 1 \\ 10 & 7 & 3 & 9 & 4 & 6 \\ 2 & 5 & 11 & 8 & 1 & 10 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 5 & 2 & 9 & 11 & 4 & 8 \\ 3 & 7 & 1 & 10 & 12 & 0 \\ 6 & 9 & 2 & 13 & 11 & 8 \\ 1 & 5 & 6 & 3 & 10 & 7 \end{pmatrix} + \lambda_4 \cdot \begin{pmatrix} 9 & 4 & 6 & 1 & 13 & 2 \\ 8 & 0 & 5 & 12 & 6 & 11 \\ 3 & 7 & 10 & 9 & 4 & 5 \\ 2 & 8 & 11 & 3 & 7 & 1 \end{pmatrix} + \lambda_5 \cdot \begin{pmatrix} 7 & 10 & 4 & 6 & 8 & 3 \\ 1 & 5 & 2 & 11 & 9 & 0 \\ 13 & 7 & 6 & 4 & 12 & 2 \\ 8 & 3 & 1 & 9 & 5 & 10 \end{pmatrix} \quad \lambda_i \in \mathbb{F}_q$$

$$(q = 13, m = 4, n = 6, k = 5)$$