Key-Recovery Attacks on ASASA

Brice Minaud[1], Patrick Derbez[2], Pierre-Alain Fouque[3], Pierre Karpman[4]

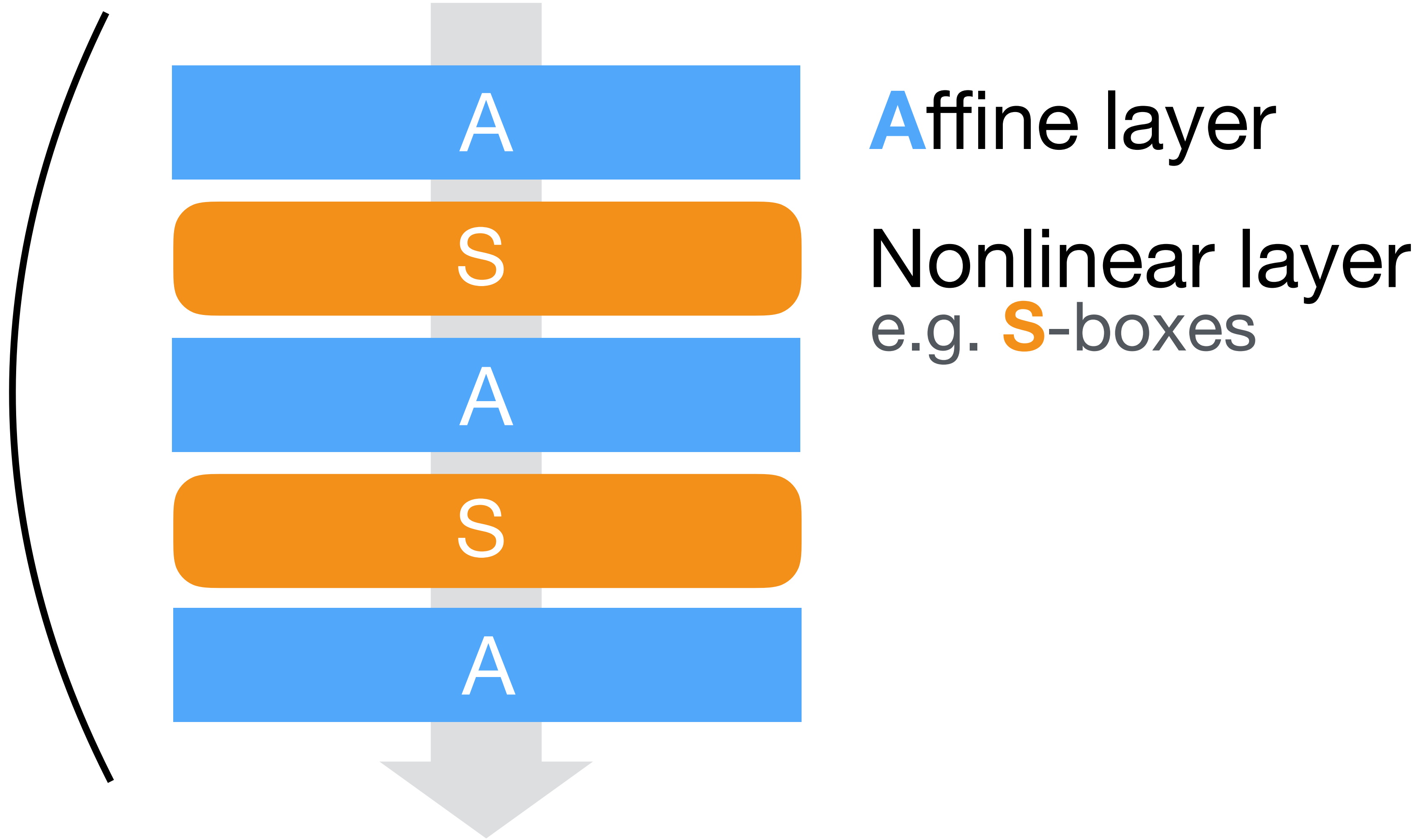
[1] Université Rennes 1
[2] Université du Luxembourg
[3] Université Rennes 1 et Institut Universitaire de France
[4] Inria et Nanyang Technological University, Singapour

ENS Lyon, May 2017

ASASA Structure

At Asiacrypt 2014, Biryukov, Bouillaguet and Khovratovich considered various applications of the ASASA structure.

$F = A \circ S \circ A \circ S \circ A$



ASASA

Three uses cases were proposed in [BBK14]:

same attack!

- 1 “black-box” scheme ≈ block cipher ✘ this paper
- 2 “strong whitebox” schemes ≈ public-key encryption scheme
  - “Expanding S-box” scheme ✘ Crypto’15 [GPT15]
  - “χ-based” scheme ✘ this paper
- 1 “weak whitebox” scheme ✘ this paper & [DDKL15]

Plan

1. Public-key **ASASA**.
2. Cryptanalysis.
3. Secret-key **ASASA**.
4. White-Box **ASASA**.

Public-key **ASASA**

Multivariate Cryptography

Hard problem: solving a system of random, say, quadratic, equations over some finite field.

→ How to get an encryption scheme $\mathbb{F}_q^n \to \mathbb{F}_q^n$:

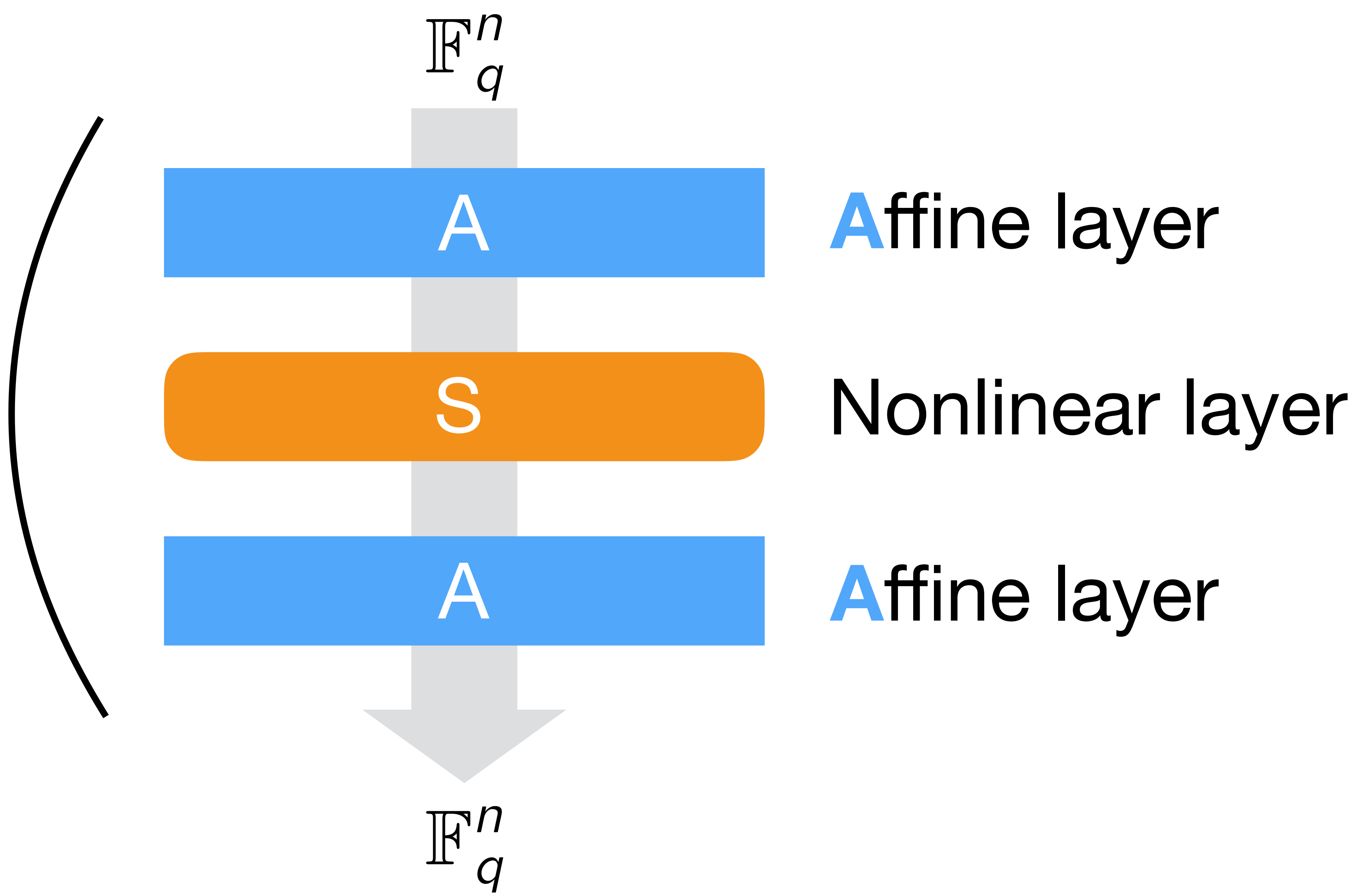
Public key: encryption function **F** given as sequence of *n* quadratic polynomials in *n* variables.

Private key: hidden structure (decomposition) of **F** that makes it easy to invert.

+: small message space, fast with private key.
-: slow public-key operations, large key, no reduction.

ASA

$F = A \circ S \circ A$

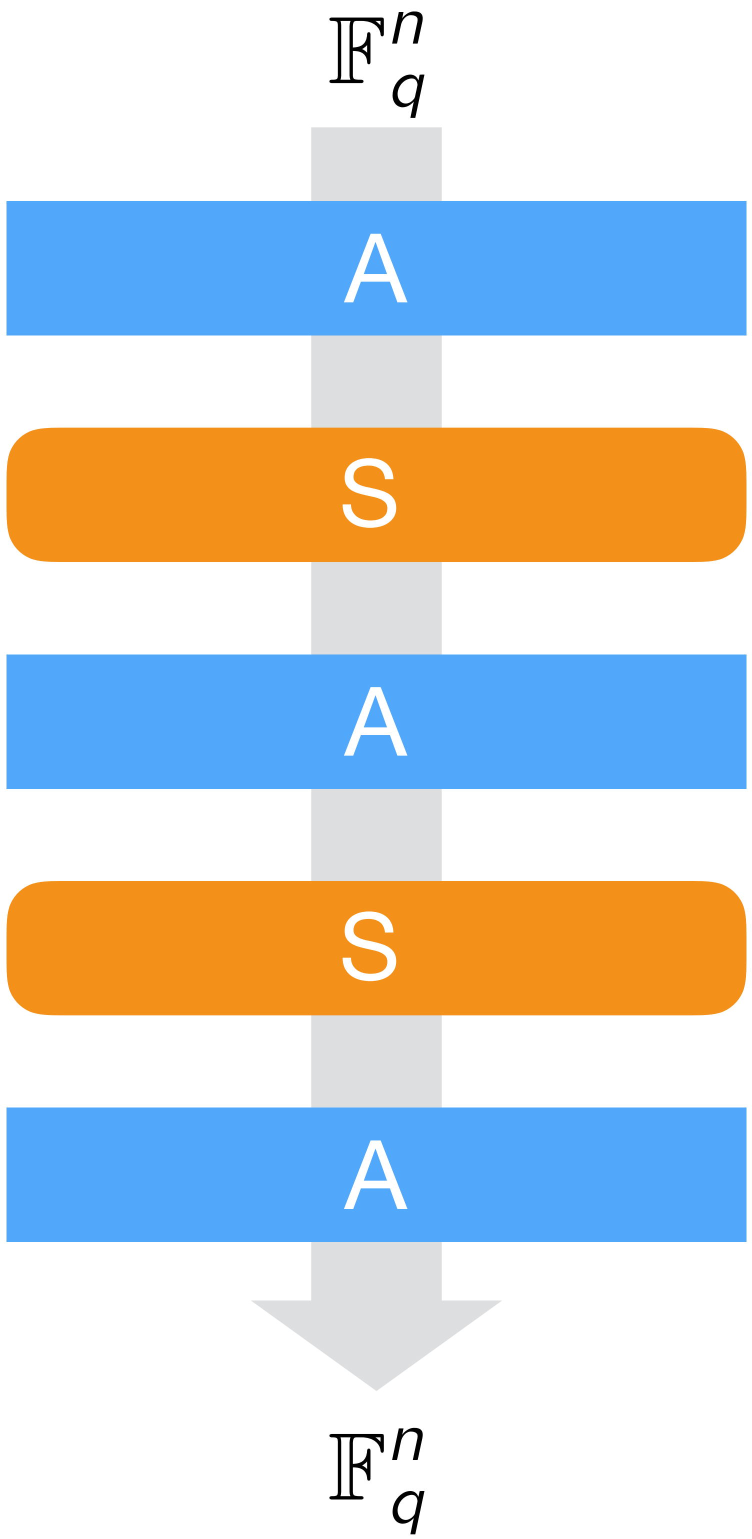


Many proposed scheme follow an ASA structure.

Matsumoto-Imai, Hidden Field Equations, Oil and Vinegar…

Almost all have been broken.

ASASA



History of ASASA

Idea already proposed by Goubin and Patarin: “2R” scheme (ICICS’97).

Broken by **decomposition** attacks.

- Introduced by Ding-Feng, Lam Kwok-Yan, and Dai Zong-Duo.
- Developped in a general setting by Faugère et al.

Decomposition attack

Problem: Let **f**, **g** be quadratic polynomials over $x_1, \ldots, x_n$. Let **h** = **g**∘**f**. Recover **f**, **g** knowing **h**.

Attack: $h_\ell = \sum \alpha_{i,j} f_i f_j$

degree 1

$$\frac{\partial h_\ell}{\partial x_k} = \sum \alpha_{i,j} \left(\frac{\partial f_i}{\partial x_k} f_j + \frac{\partial f_j}{\partial x_k} f_i \right)$$

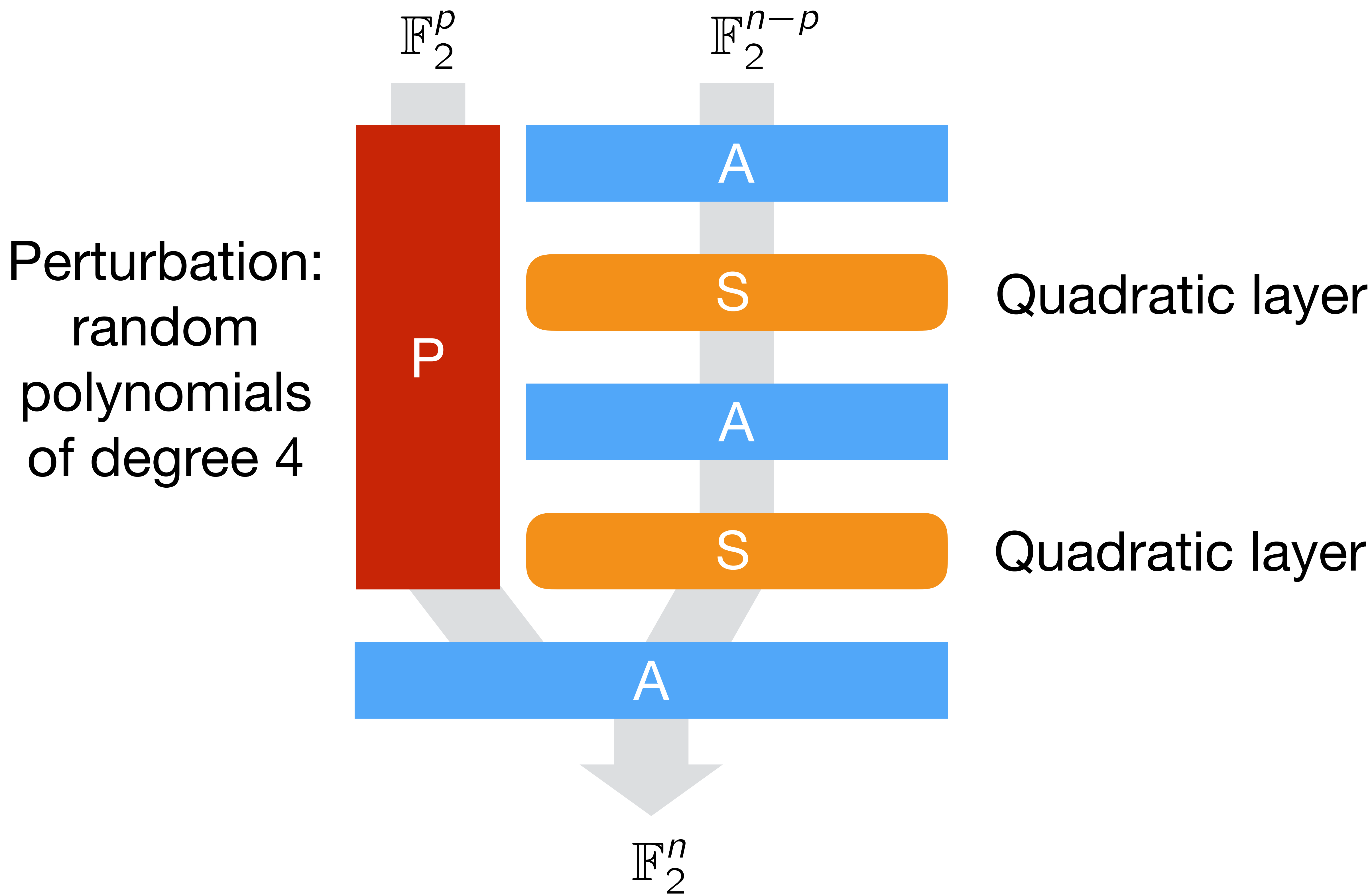
$$\in \operatorname{span}\{x_i f_j : i, j \leq n\}$$

→ We get:

$$\operatorname{span}\left\{\frac{\partial h_\ell}{\partial x_k}\right\} = \operatorname{span}\{x_i f_j : i, j \leq n\}$$

→ Yields $\operatorname{span}\{f_j : i, j \leq n\}$.

Structure ASASA + P [BBK14]



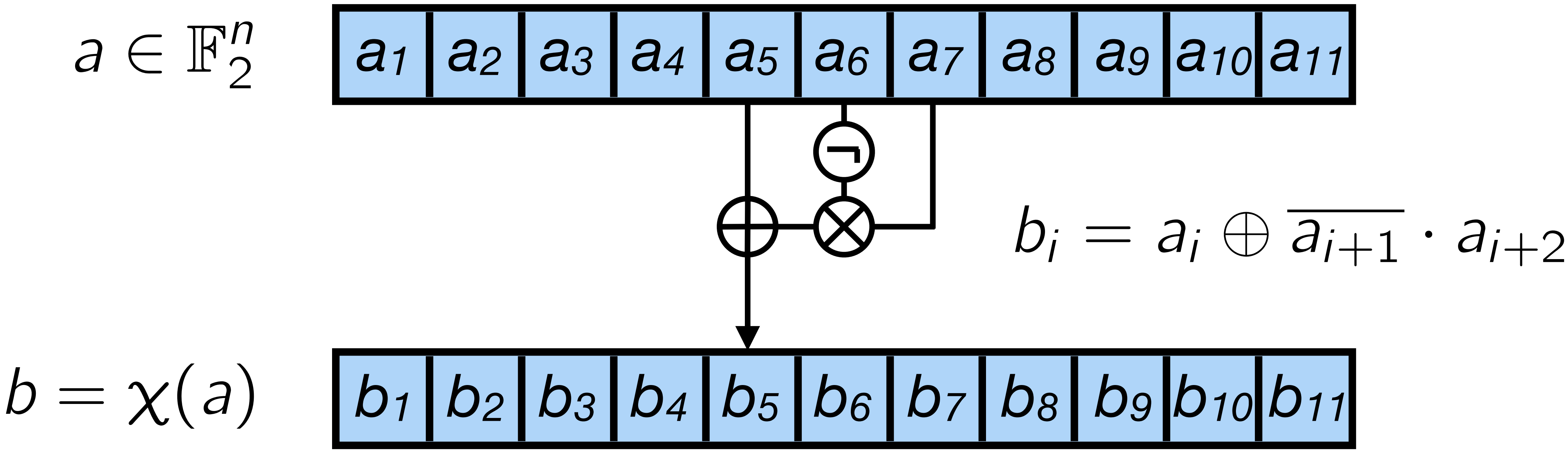
Note : this is slightly different from BBK14.

Instances of ASASA + P

Two instances were proposed in BBK14 :

- "Expanding S-boxes" : decomposition attack by Gilbert, Plût and Treger, Crypto'15.
- χ-based scheme: using the χ function of Keccak.

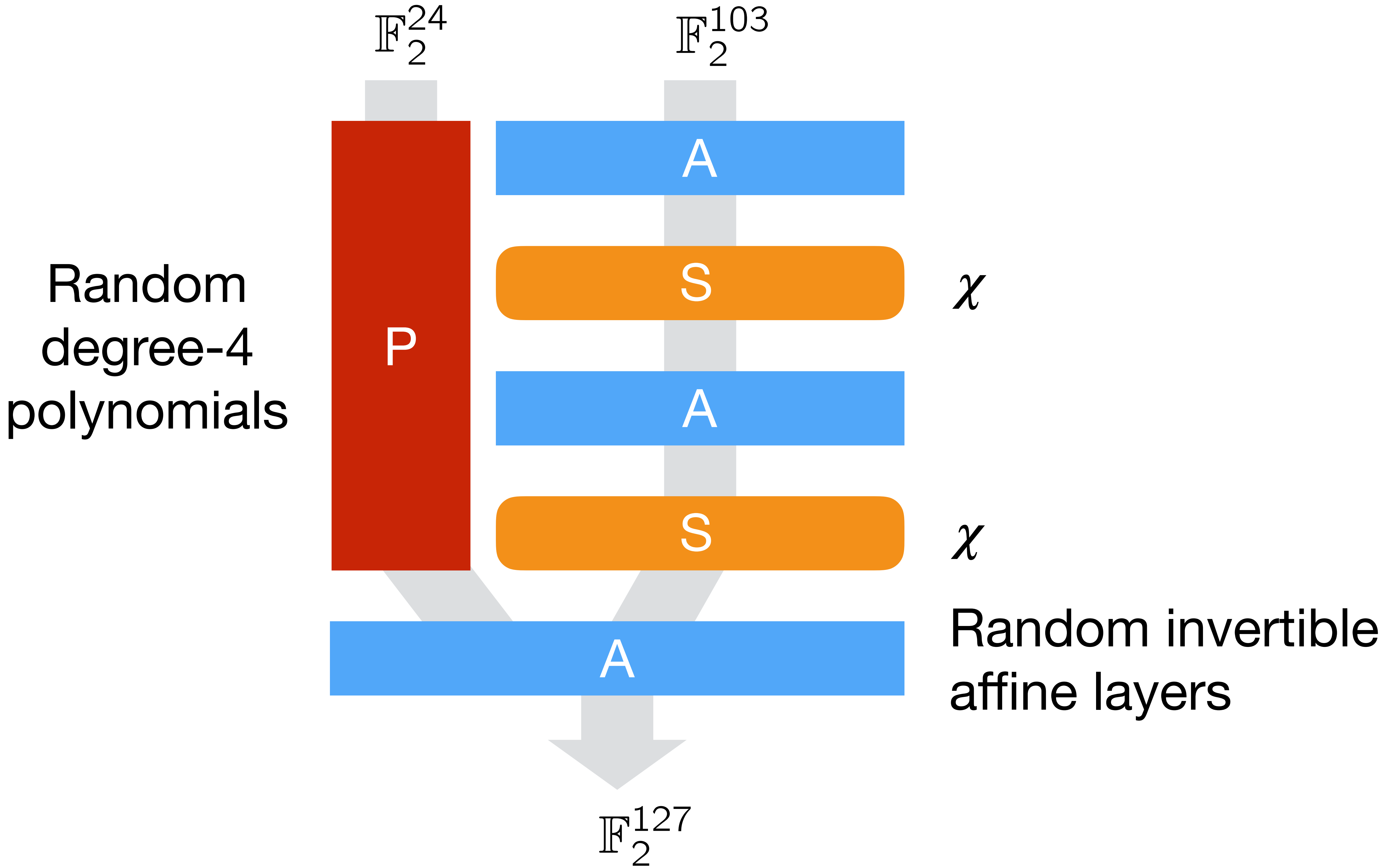
χ function of Keccak



Introduced by Daemen in 1995, known for its use in Keccak (SHA-3).

Invertible for odd number of bits.

χ-based instance



Attack!

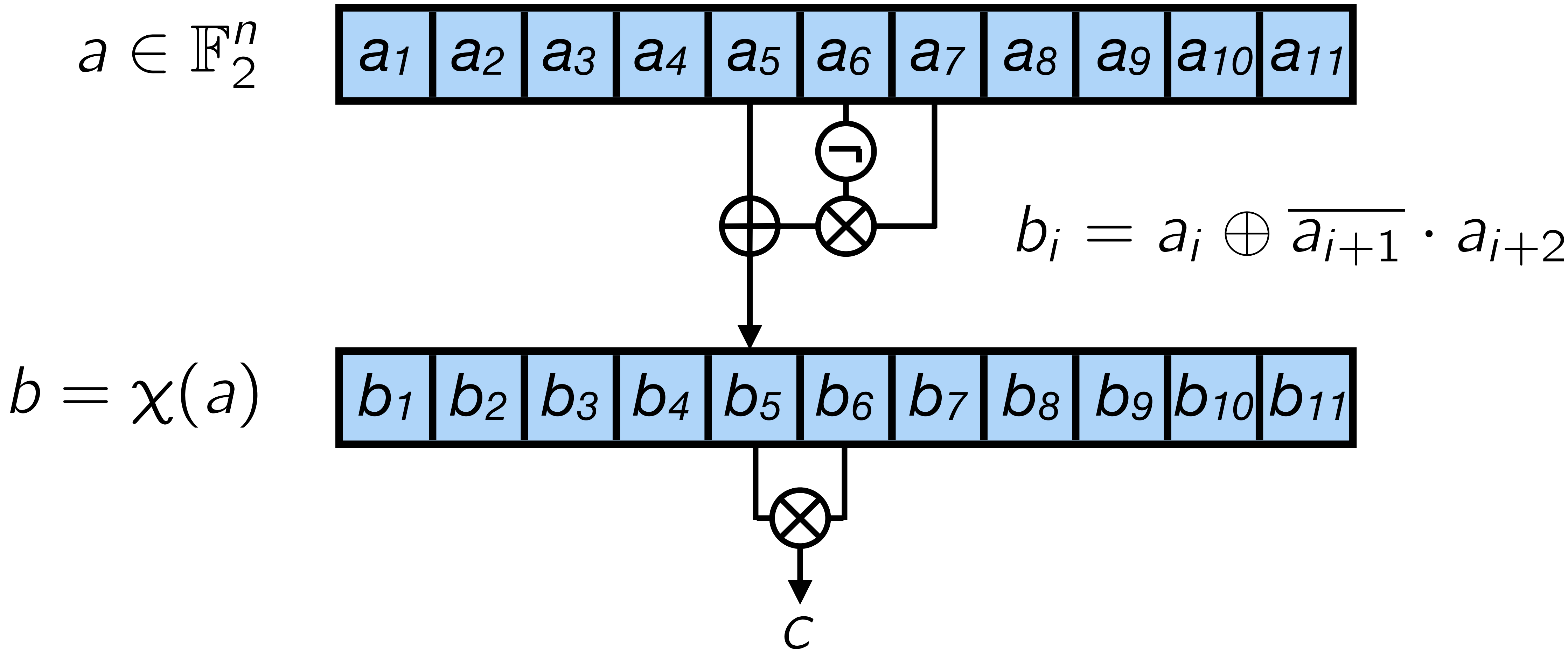
Cubes

A **cube** is an affine subspace [DS08].

Property : Let *f* be a degree-*d* polynomial over binary variables. If *C* is a cube of dimension *d+1*, then :

$$\sum_{c \in C} f(c) = 0$$

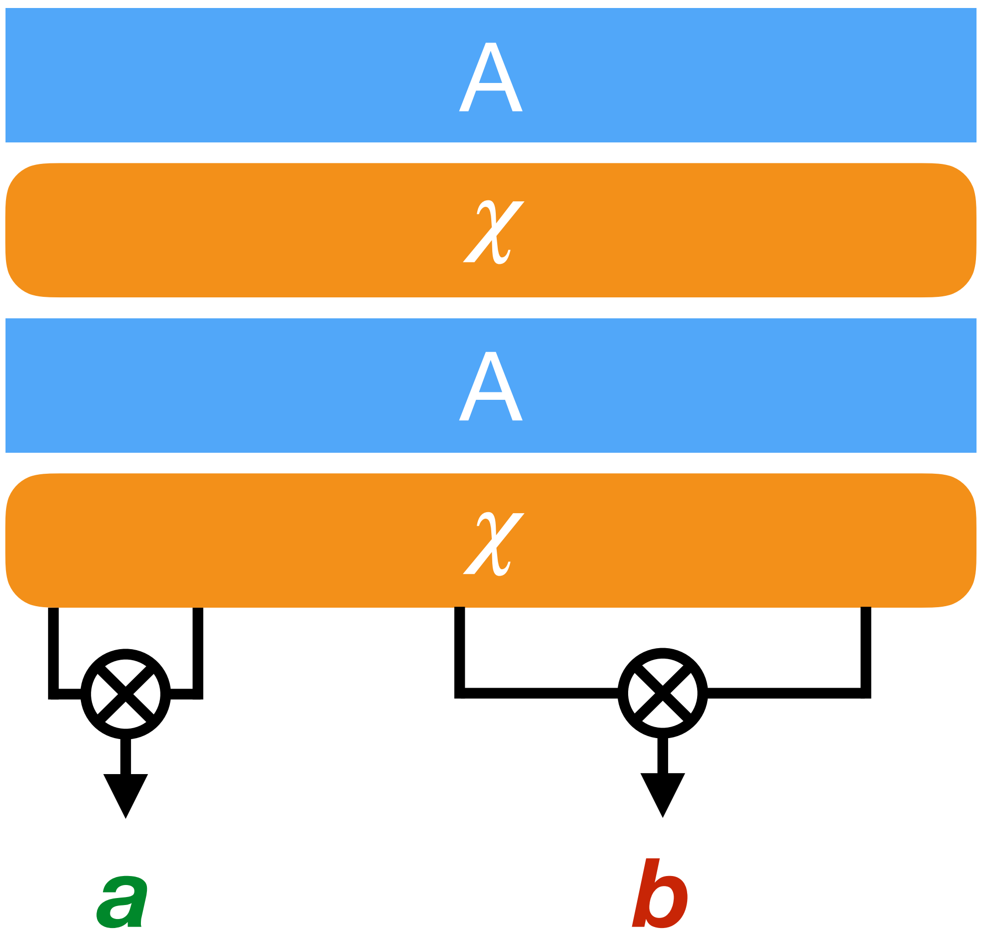
Degree deficiency



$$
\begin{aligned}
c &= b_i \cdot b_{i+1} \\
&= (a_i \oplus \overline{a_{i+1}} \cdot a_{i+2}) \cdot (a_{i+1} \oplus \overline{a_{i+2}} \cdot a_{i+3})
\end{aligned}
$$

➞ c has degree 3. Sums up to 0 over cube of dim 4.

ASASA Cryptanalysis



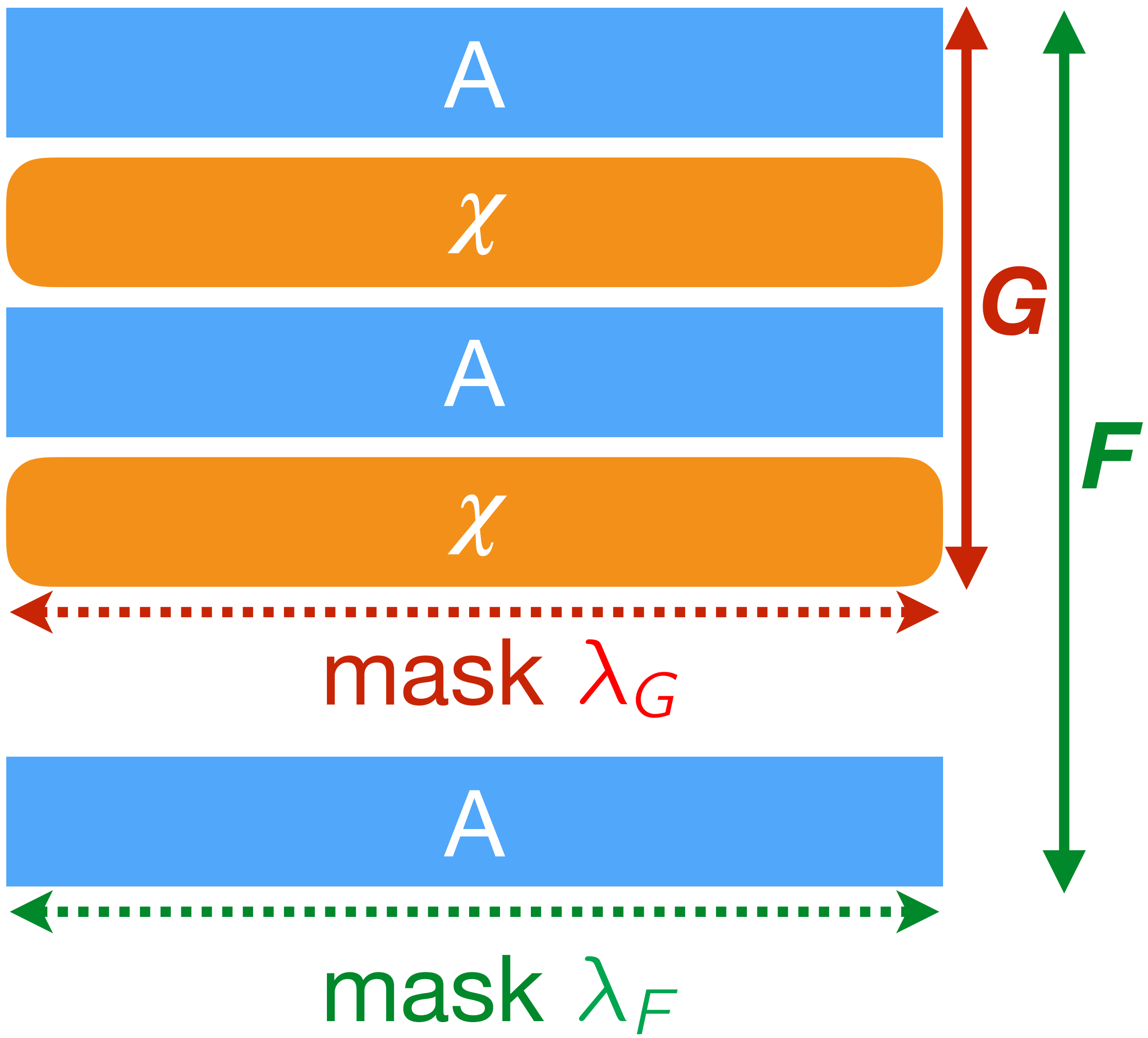
- Let ***a*** = product of 2 **adjacent** bits at the output of χ.

  Then ***a*** has degree 6.

- Let ***b*** = product of 2 **non-adjacent** bits at the output of χ.

  Then ***b*** has degree 8.

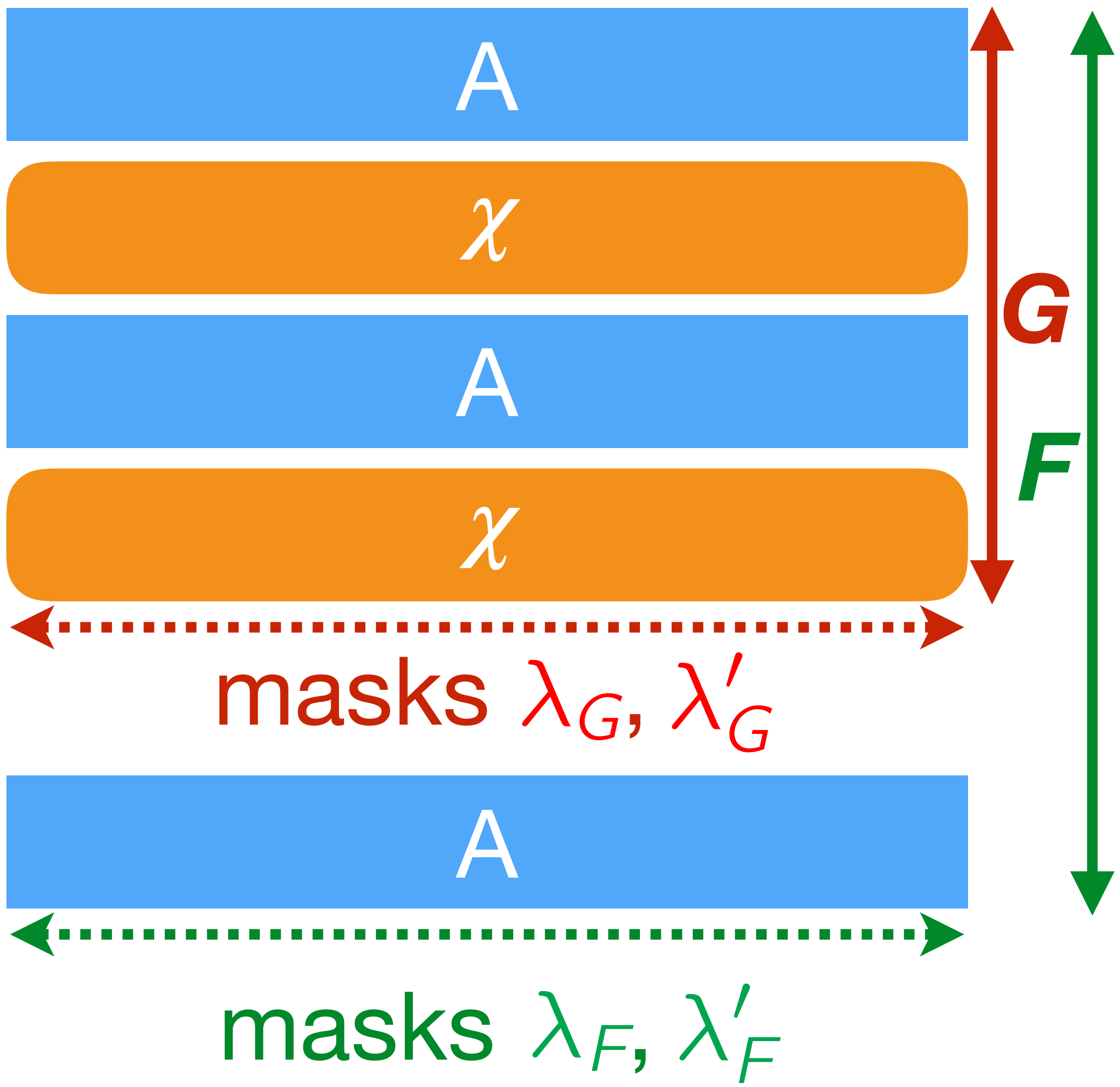
ASASA Cryptanalysis



Let λ_F be an output mask, i.e. we look at $\langle F | \lambda_F \rangle = x \mapsto \langle F(x) | \lambda_F \rangle$.

Then there exists a mask λ_G s.t. $\langle F | \lambda_F \rangle = \langle G | \lambda_G \rangle$.

ASASA Cryptanalysis

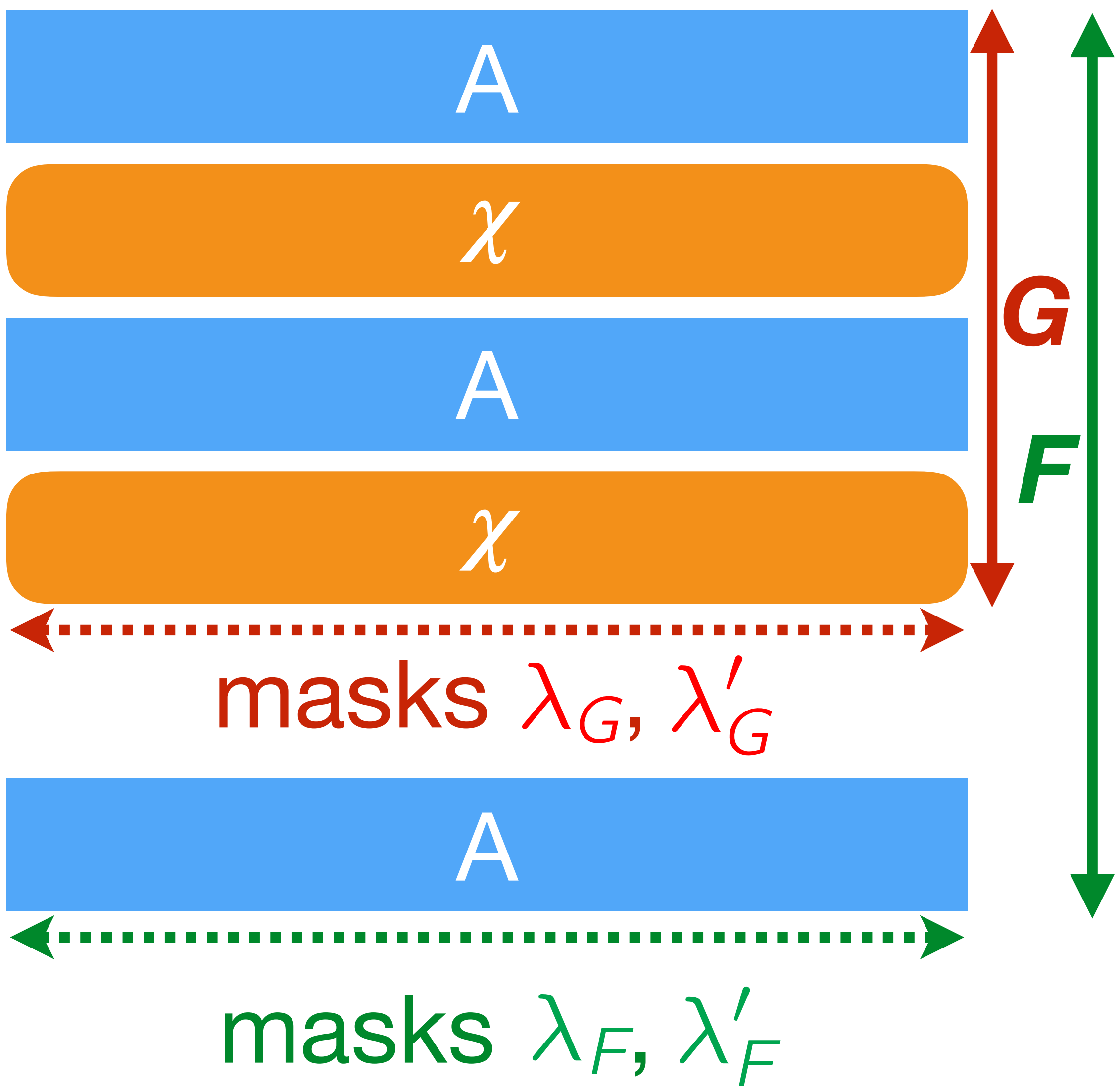


Let λ_F, λ'_F be two output masks, and λ_G, λ'_G the associated masks.

- If λ_G and λ'_G activate **single adjacent** bits, $\langle F|\lambda_F\rangle \cdot \langle F|\lambda'_F\rangle$ has degree 6.

- Otherwise $\langle F|\lambda_F\rangle \cdot \langle F|\lambda'_F\rangle$ has degree 8.

ASASA Cryptanalysis



Goal : Find λ_F, λ'_F such that

$$\deg(\langle F|\lambda_F\rangle \cdot \langle F|\lambda'_F\rangle) = 6$$

Let C be a dimension-7 cube. Then :

$$\Sigma_{c\in C}\langle F(c)|\lambda_F\rangle \cdot \langle F(c)|\lambda'_F\rangle = 0$$

→ we get an equation on λ_F, λ'_F.

ASASA Cryptanalysis

View λ_F, λ'_F as two vectors of n binary unknowns: $(\lambda_0, \ldots, \lambda_{n-1})$ and $(\lambda'_0, \ldots, \lambda'_{n-1})$. Then:

$$\begin{aligned}\sum_{c \in C} \langle F(c)|\lambda\rangle\langle F(c)|\lambda'\rangle &= \sum_{c \in C} \sum_{i<n} \lambda_i F_i(c) \sum_{j<n} \lambda'_j F_j(c) \\ &= \sum_{i,j<n} \Big(\sum_{c \in C} F_i(c) F_j(c) \Big) \lambda_i \lambda'_j \\ &= 0\end{aligned}$$

$\Rightarrow$ We get a quadratic equation on the λ_i, λ'_i's.

ASASA Cryptanalysis

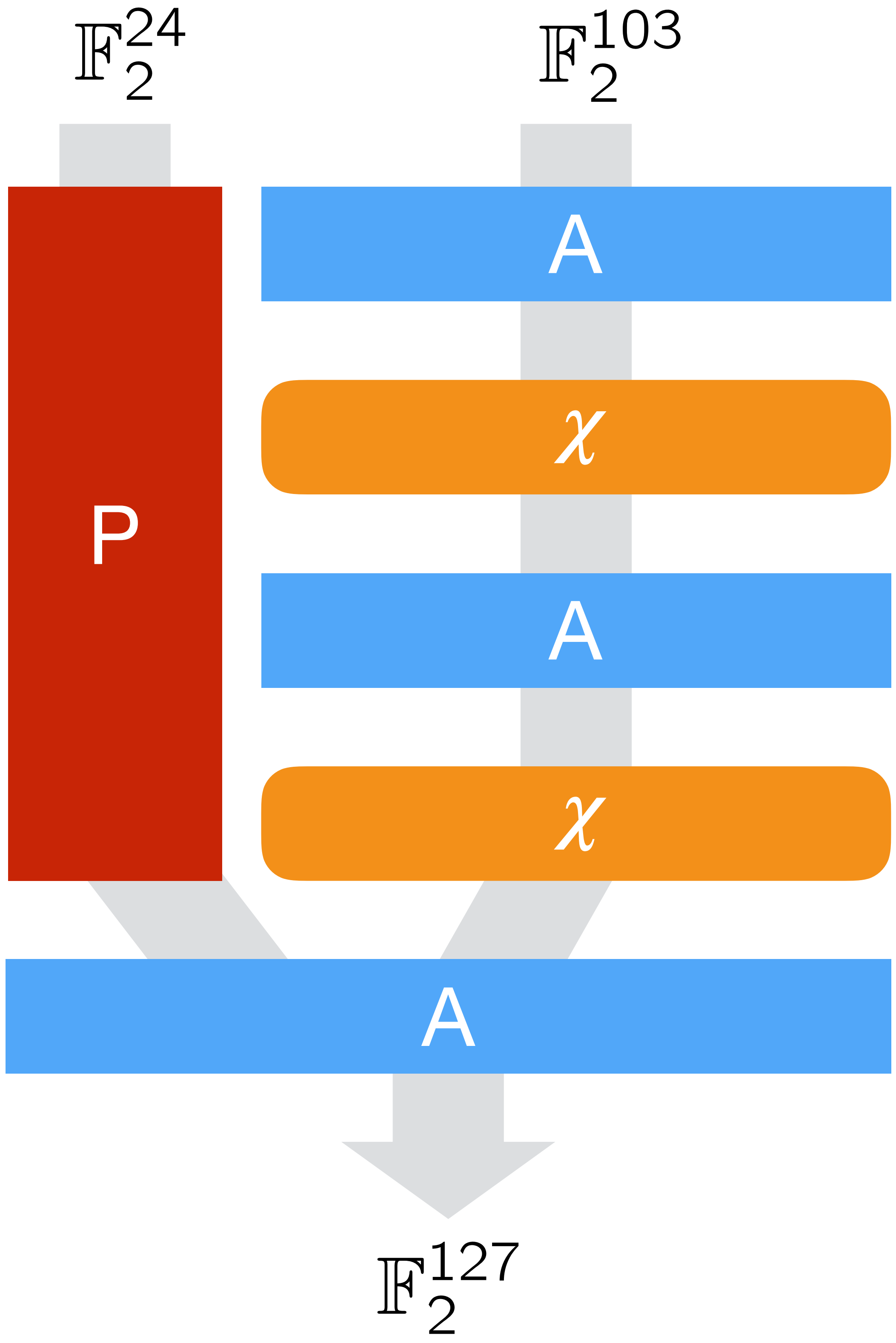
Each cube yields 1 quadratic equation on the λ_i, λ'_i's.

Using relinearization, there are $127^2 \approx 2^{14}$ terms $\lambda_i \lambda'_j$
→ we need 2^{14} cubes of dimension 7.

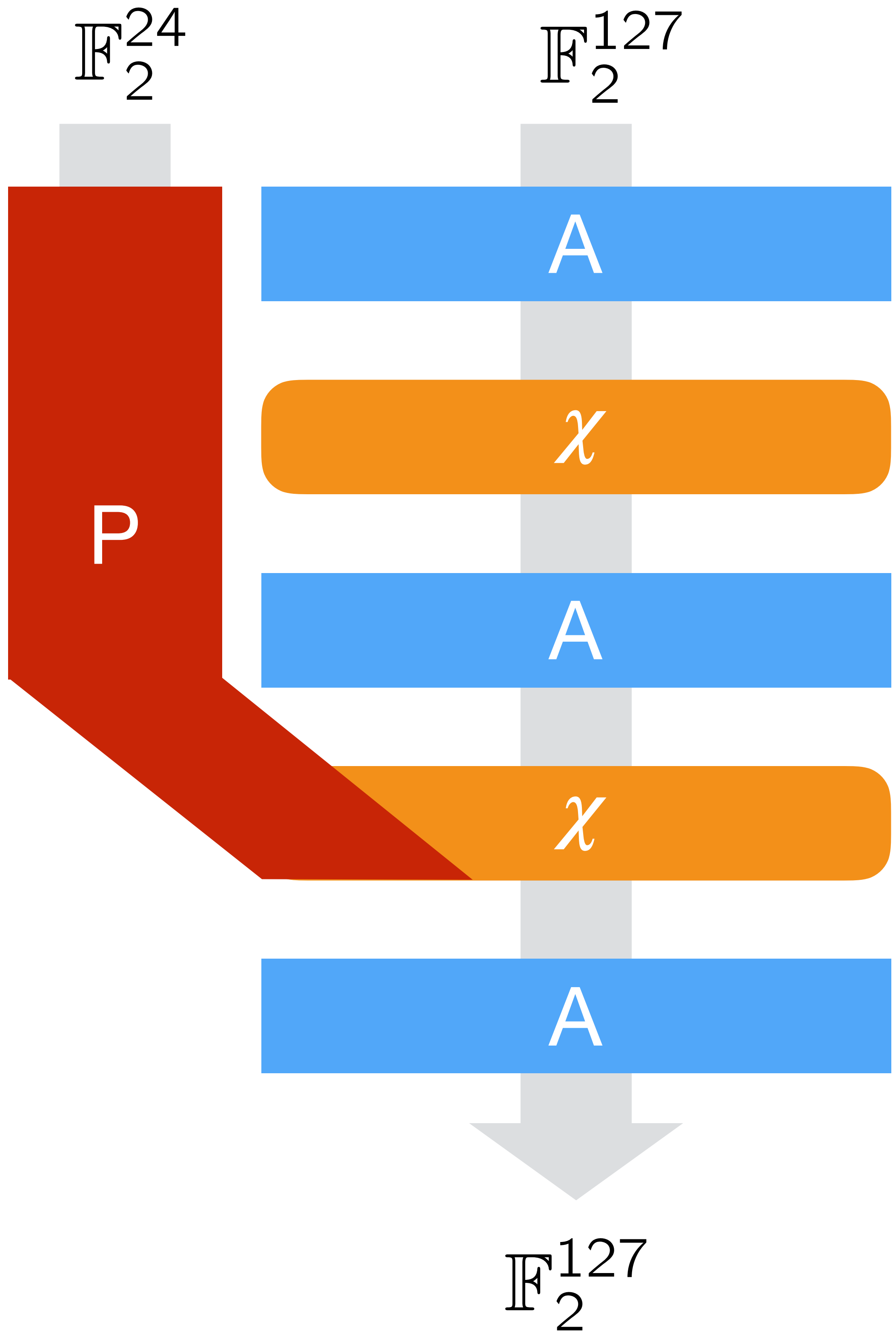
- Step 1: Solve linear system. Yields linear span *L* of solutions.
- Step 2: Recover vectors of the form $\lambda_i \lambda'_j$ within *L*.

Conclusion: the last layer is recovered using 2^{21} CP, with time complexity ≈ 2^{39} (for inverting a binary matrix of size 2^{13}).
(In general: $n^6/4$ time and $7n^2/2$ data.)

Remaining layers

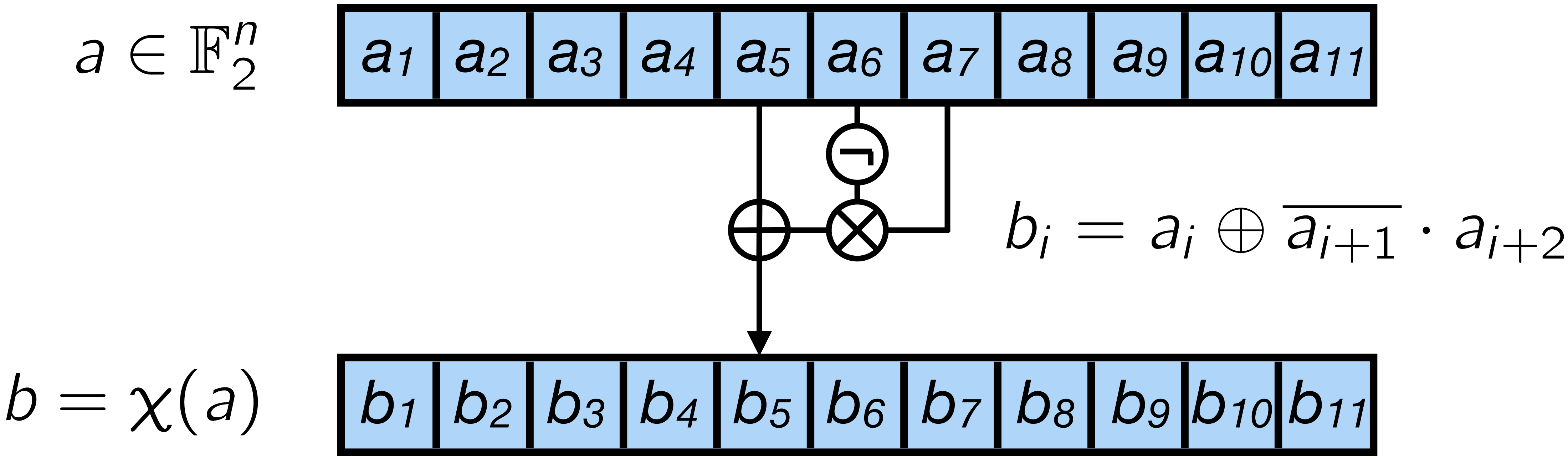


Remaining layers



Due to the perturbation, it is not possible to simply invert the last χ layer.

χ function of Keccak

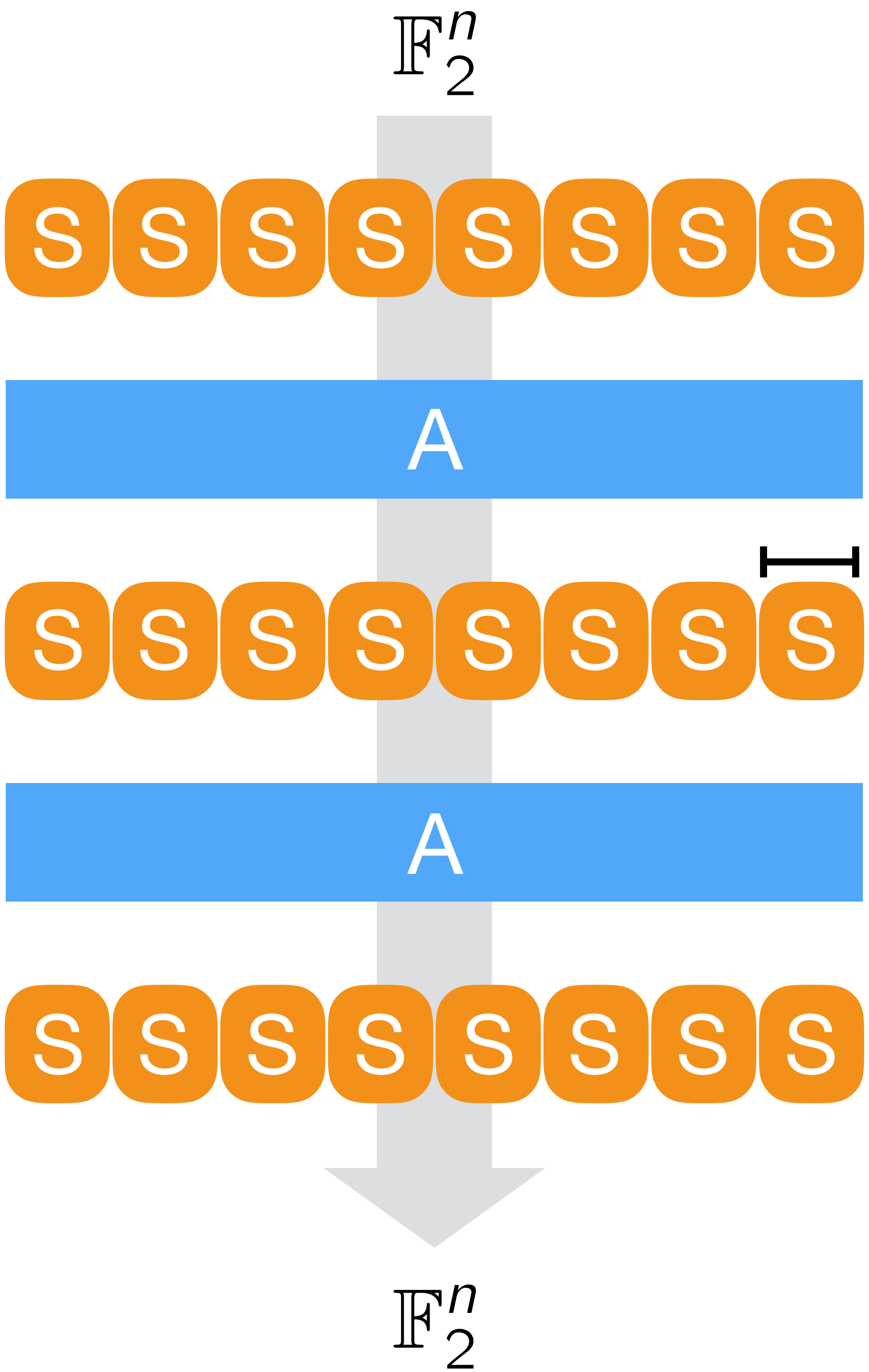


Problem 1: Given $P = A{\cdot}B \oplus C$ for quadratic A, B, C in $\mathbb{F}_2[X_1, \ldots, X_n]/\langle X_i^2 - X_i \rangle$, find A, B, C.

- There exists an efficient (heuristic) quadratic algorithm.

Black-box **ASASA**

SASAS structure



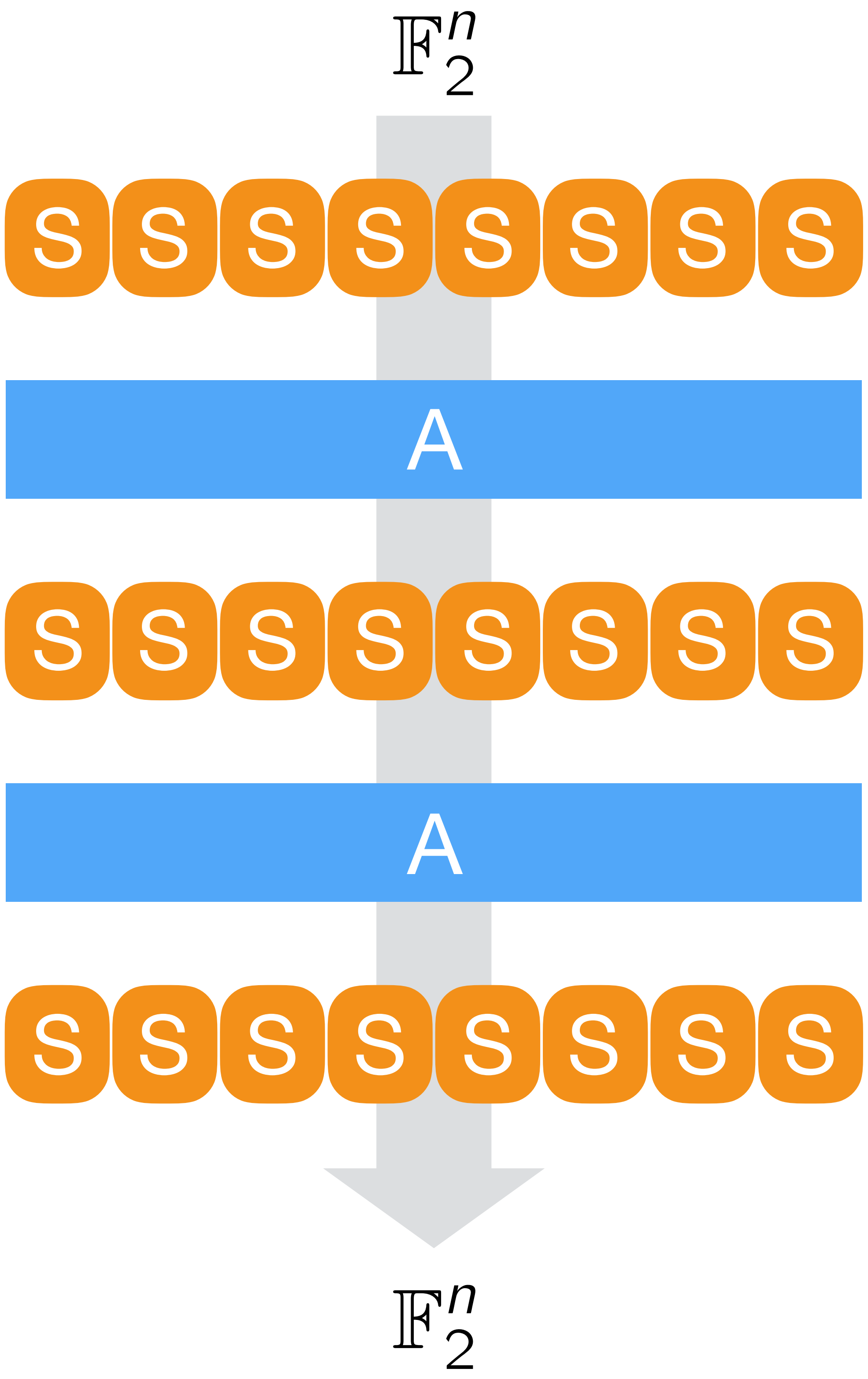
Independent random m-bit **S**-boxes

Random **A**ffine layer on n bits

m bits

→ «Large» permutation over n bits from «small» permutations over k bits.

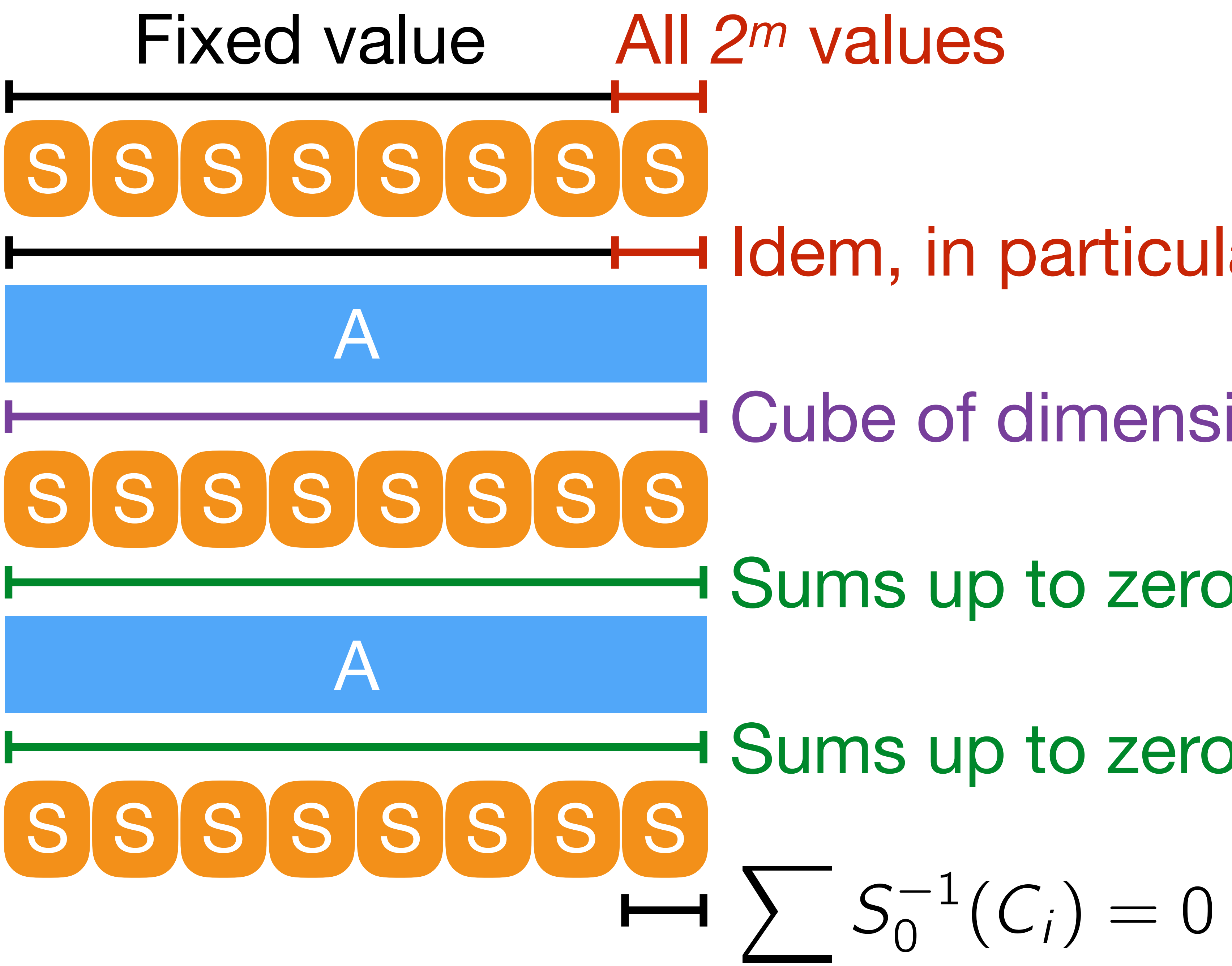
SASAS structure



Analyzed by Biryukov and Shamir at Eurocrypt 2001.

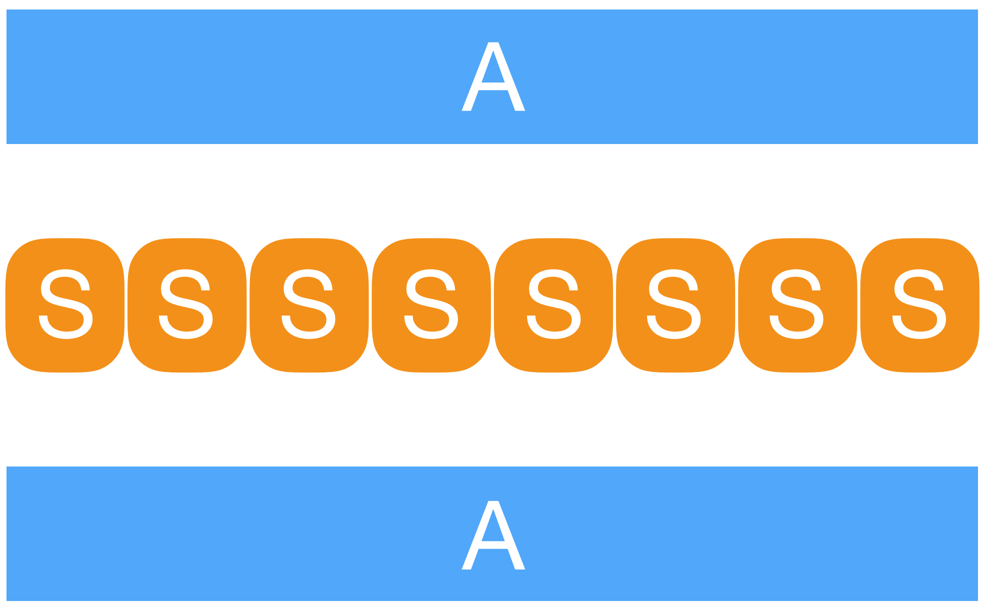
Goal: recover all internal components (affine layers **A** and **S**-boxes) with only "black-box" access (KP/CP/CC).

Cryptanalysis of SASAS



→ linear equations with unknowns $x_i = S_0^{-1}(i)$

Cryptanalysis of SASAS



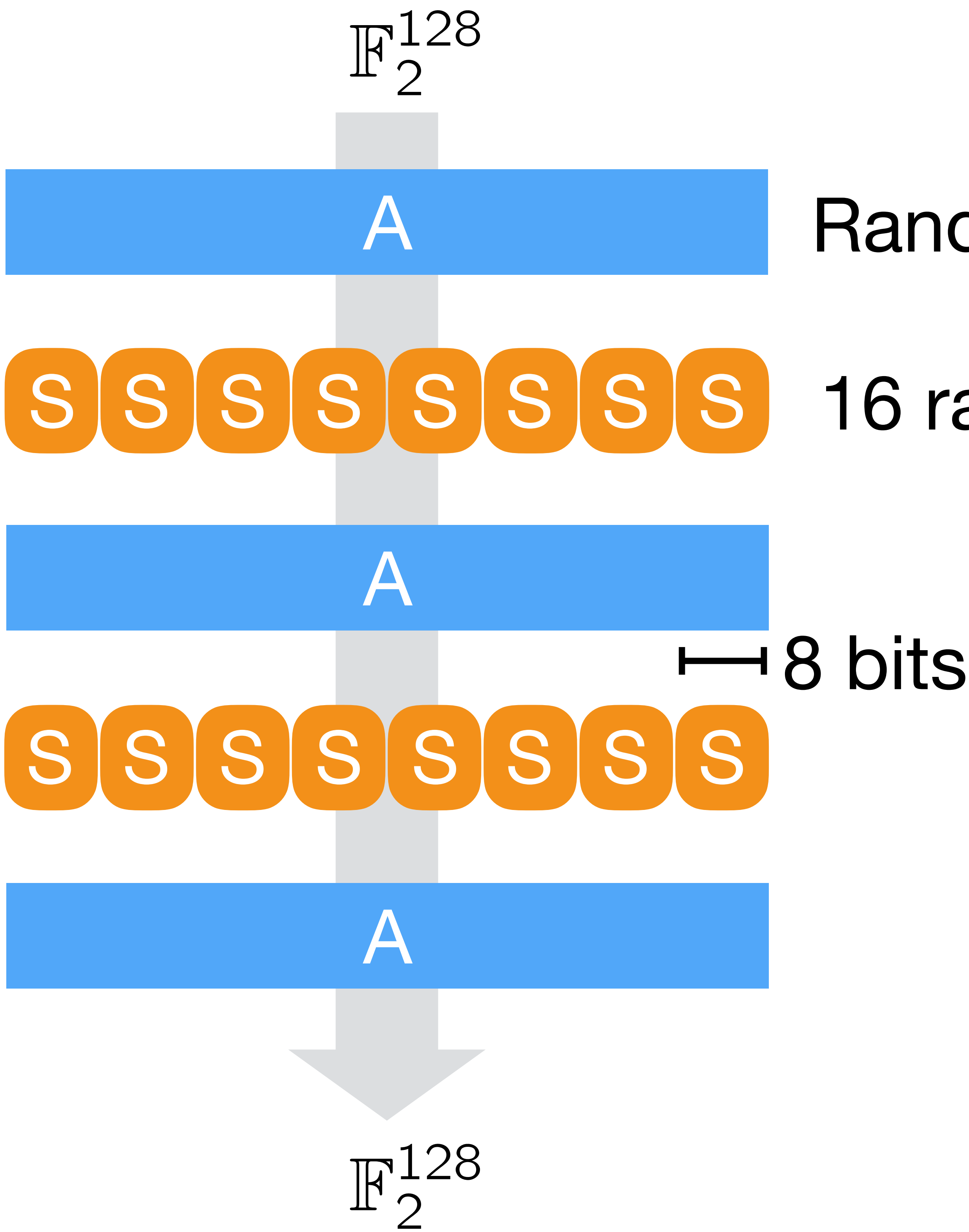
- Repeat until enough equations are gathered.
- Solve linear system of dim. 2^m to recover the final **S** layer.

By symmetry, we can do the same for the first layer.

Cost: time $k \cdot 2^{3m}$, data $3 \cdot 2^{2m}$, with m = n/k = #*S*-boxes.

Then **ASA** can be decomposed by a simple differential attack.

Black-box ASASA [BBK14]



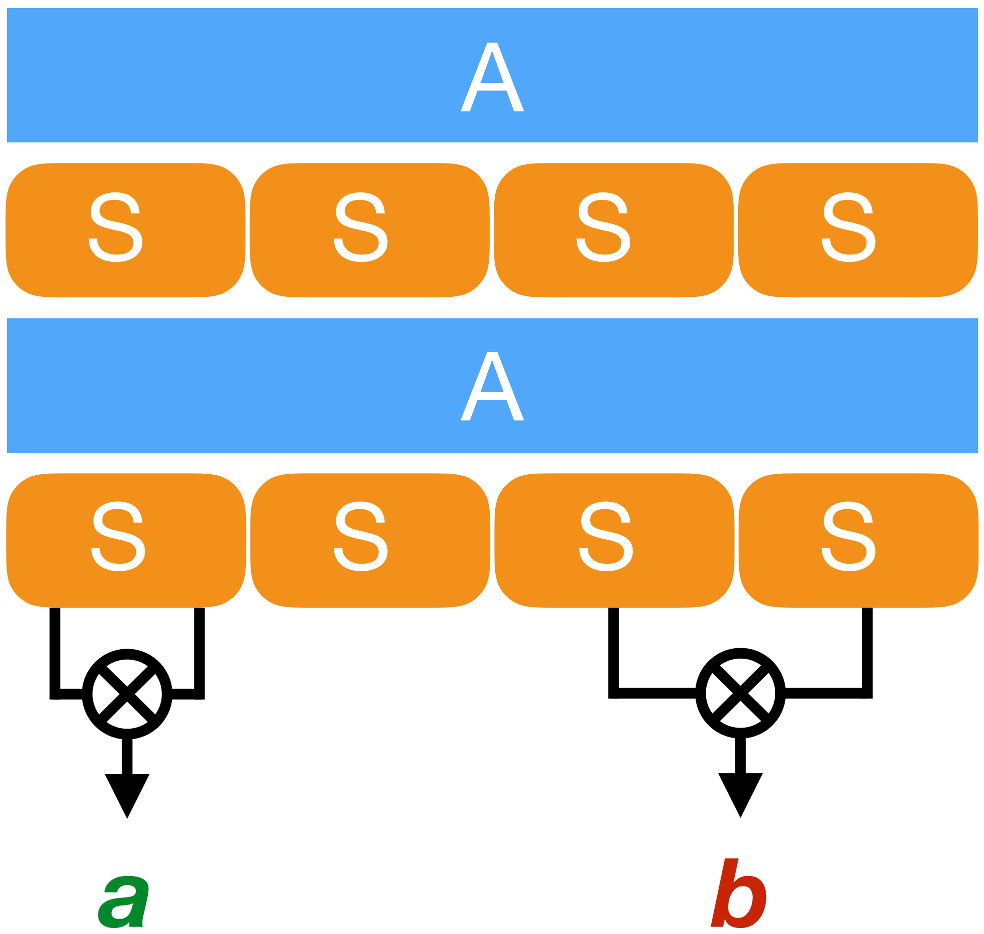
Random Affine layer over 128 bits.

16 random independent S-boxes

Goal : recover all internal components.

Note: degree ≤ 49
⇒ distinguisher w. 2^{50} CP

ASASA cryptanalysis



Degree of an S-box = 7.

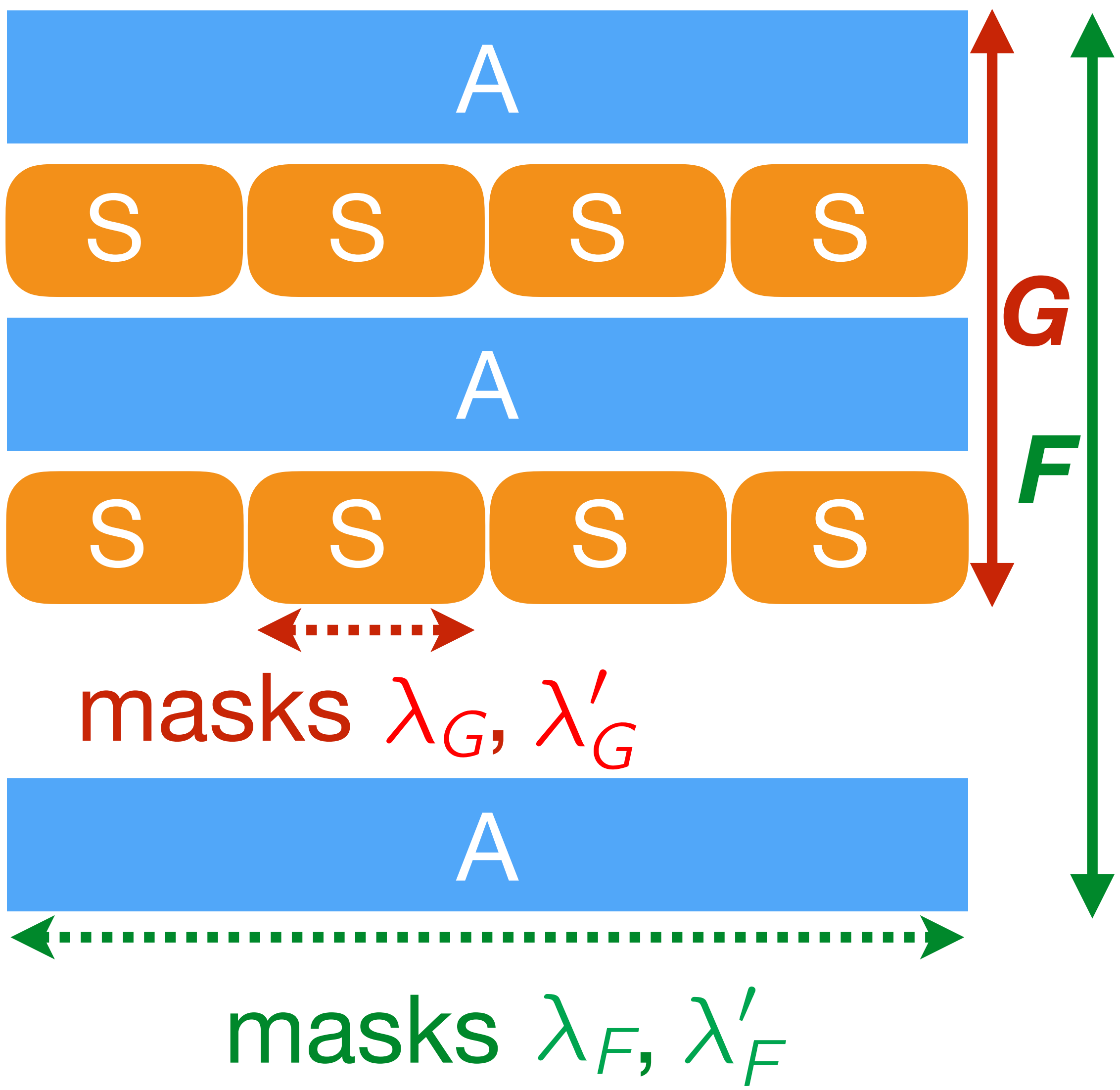
- Let ***a*** = product of 2 output bits of a **single common** S-box.

  Then ***a*** has degree 7x7 = 49.

- Let ***b*** = product of 2 output bits of two **distinct** S-boxes.

  Then ***b*** has max degree (127).

ASASA Cryptanalysis



Goal : Find λ_F, λ'_F such that

$$\deg(\langle F|\lambda_F\rangle \cdot \langle F|\lambda'_F\rangle) = 49$$

Let C be a dimension-50 cube. Then:

$$\Sigma_{c \in C}\langle F(c)|\lambda_F\rangle \cdot \langle F(c)|\lambda'_F\rangle = 0$$

→ we get an equation on λ_F, λ'_F.

Conclusion : All internal components are recovered in time and data complexity 2^{63}. In general: $n^2 2^{(m-1)^2}$.

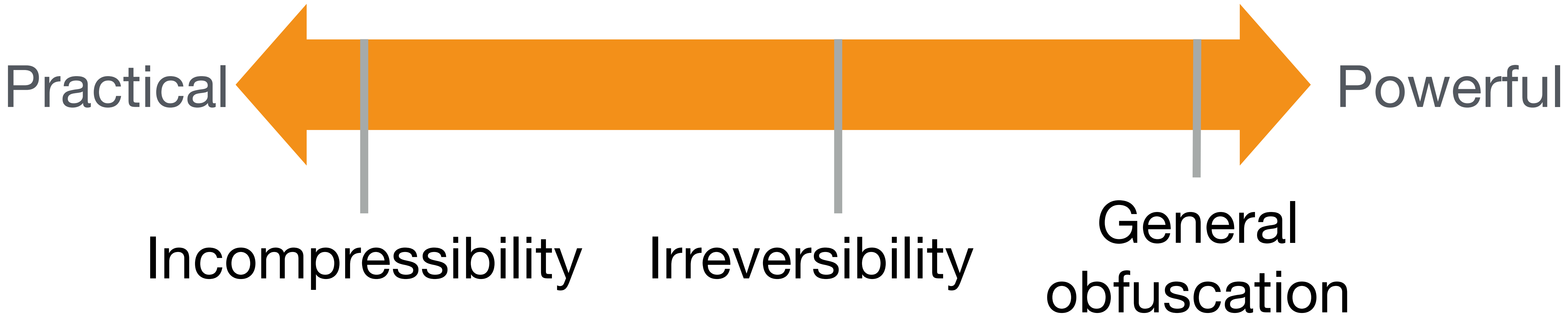
For comparison: the distinguisher is in 2^{50}. In general $2^{(m-1)^2+1}$.

Small-block **ASASA**

White-Box Cryptogaphy

White-Box Cryptography: protection against adversaries having complete access to the implementation of a cipher.

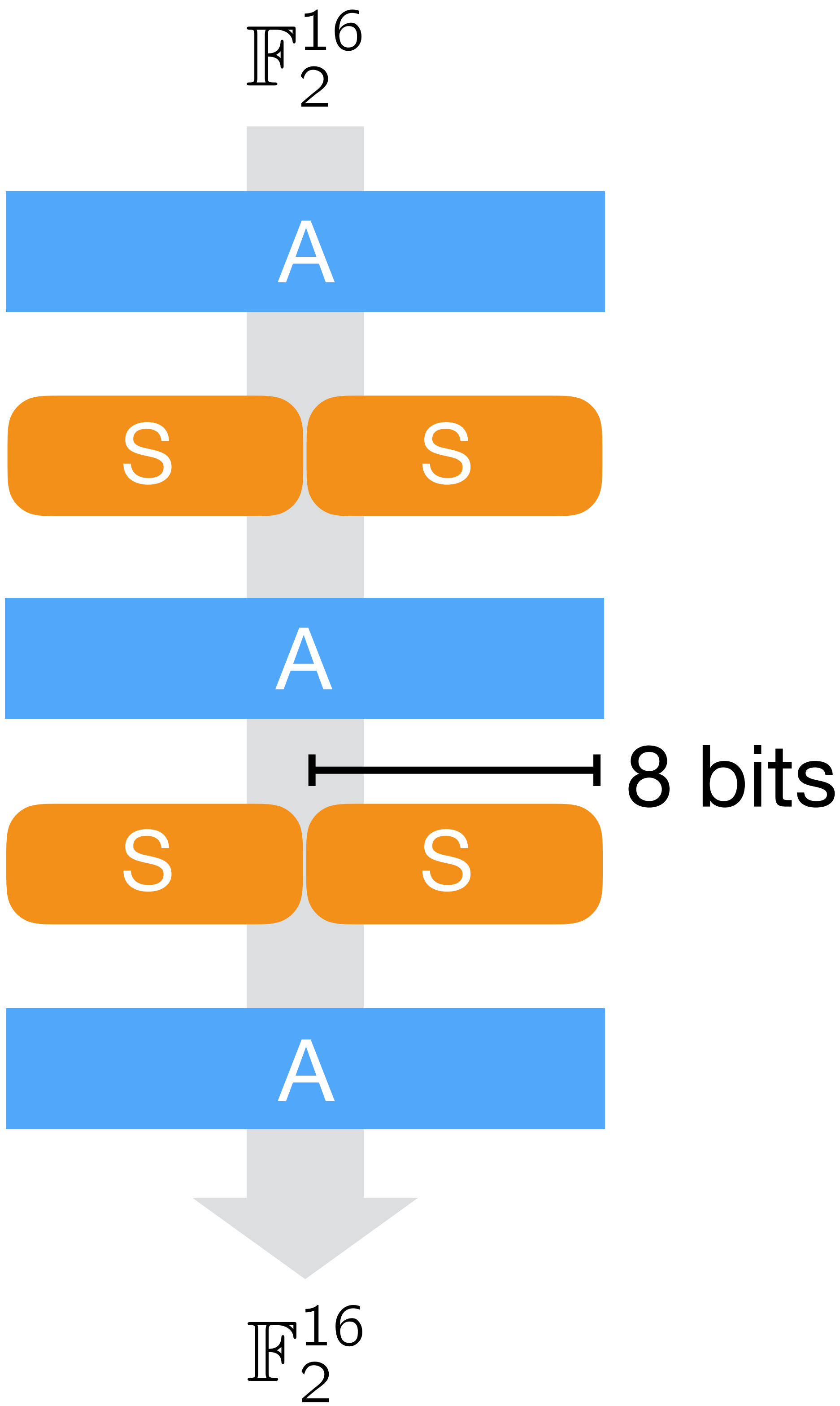
Important topic within industry. No complete solution. Various trade-offs → different models.



Incompressible cipher: block cipher with large description.

Goal: impede code lifting and code distribution.

White-box ASASA [BBK14]



Idea: use large **S**-boxes with secret structure within conventional design.

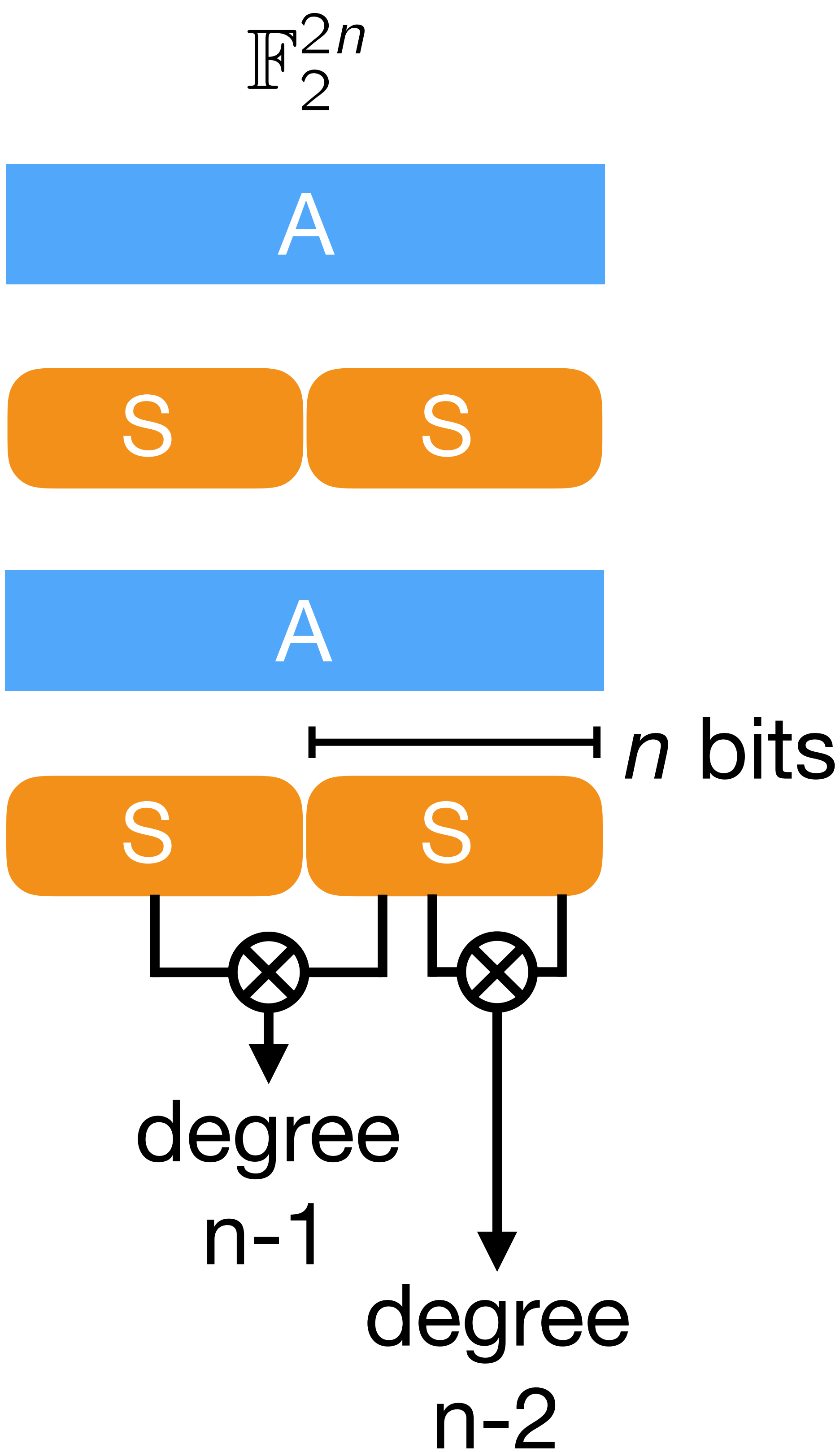
It may seem that our attack fails because $\deg(S)^2 = 49 > 15$.

[DDKL15] (from [BC13]):

$$\deg(F) < n - (k-1)\left(1 - \frac{1}{m-1}\right)$$

with n: #input bits, k: #S-boxes,
$m = n/k$: #input bits per S-box.

Small-block ASASA



Idea: use large **S**-boxes with secret structure within conventional design.

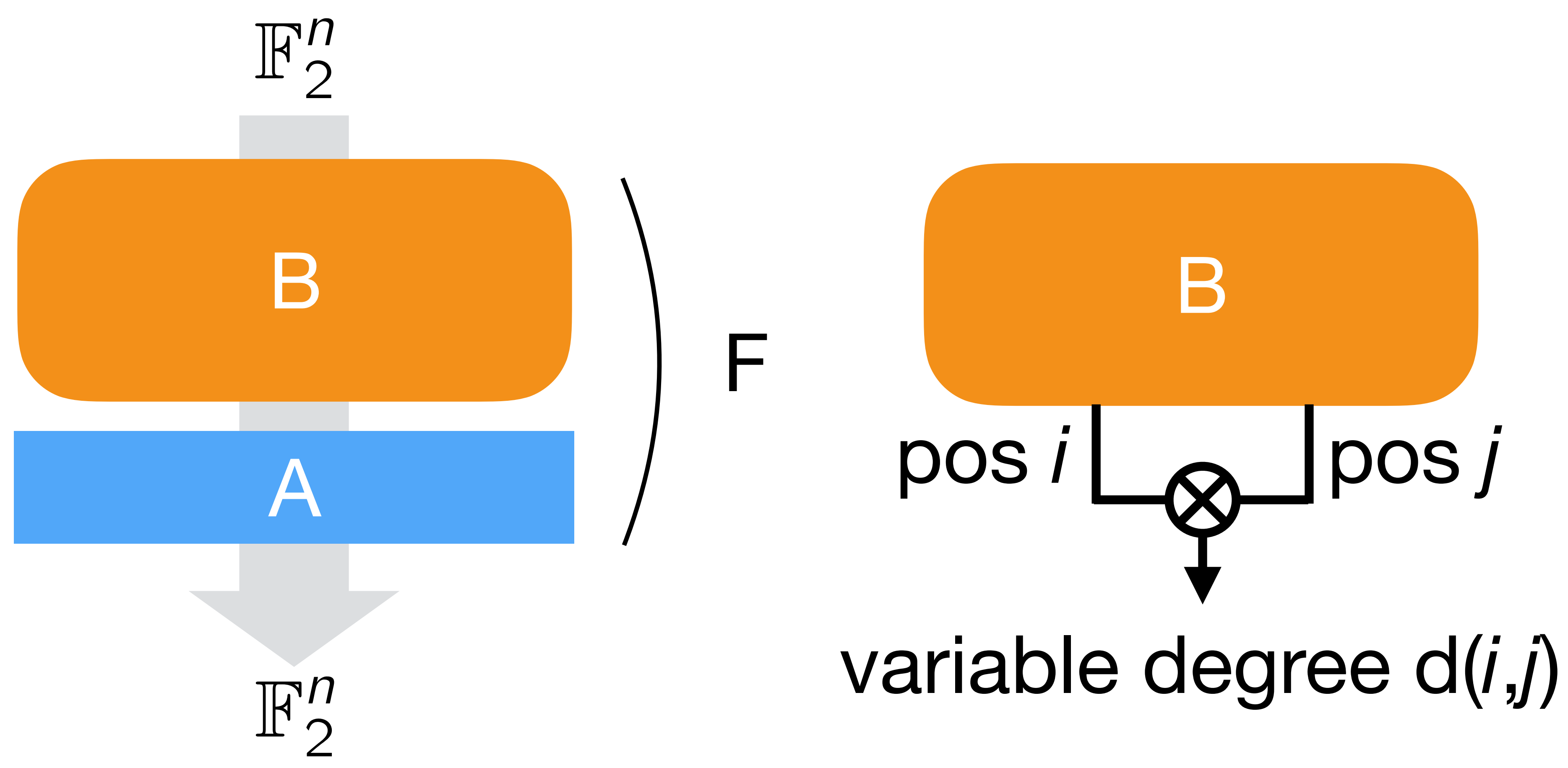
It may seem that our attack fails because $\deg(S)^2 = 49 > 15$.

[DDKL15] (from [BC13]):

$$\deg(F) < n - (k-1)\left(1 - \frac{1}{m-1}\right)$$

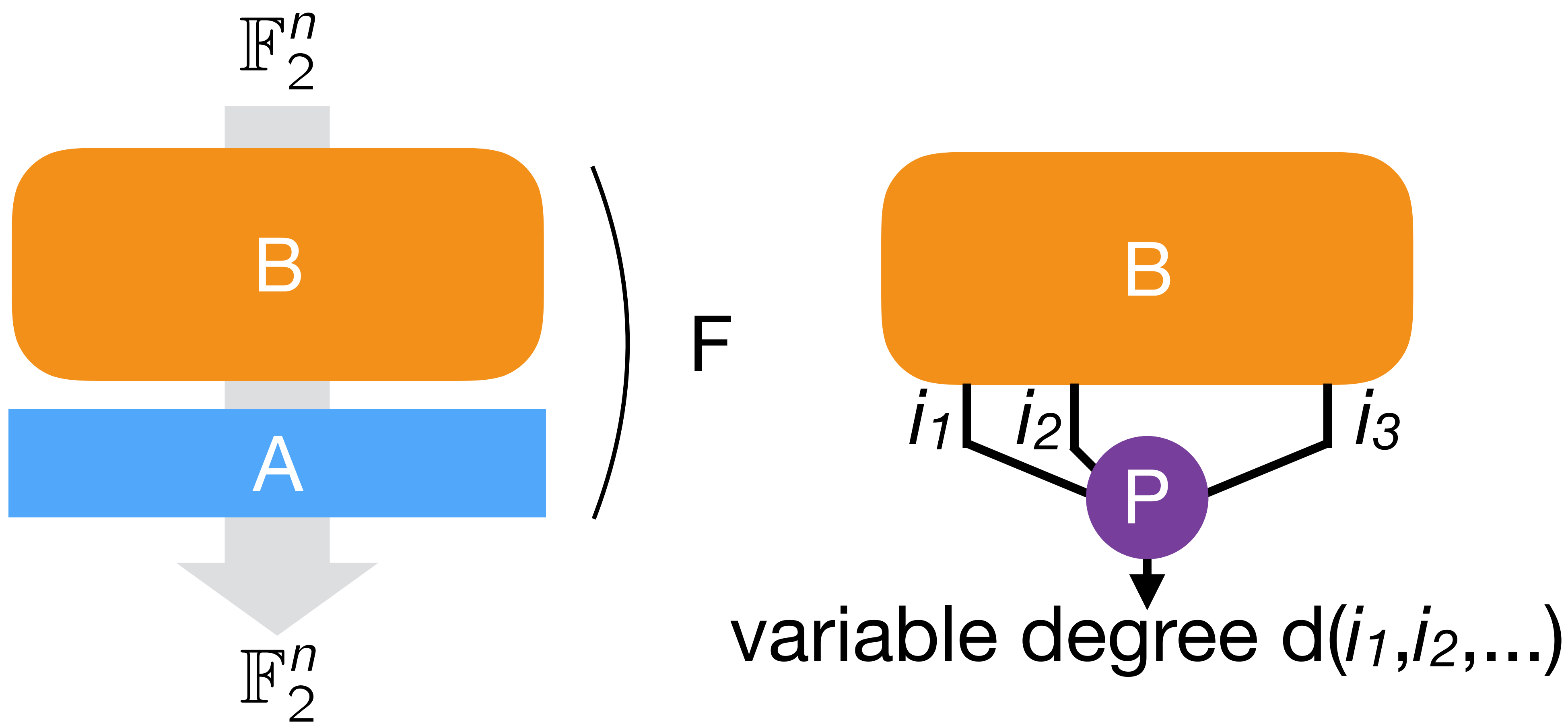
with n: #input bits, k: #S-boxes,
$m = n/k$: #input bits per S-box.

The attack in general



In general, all that matters is that the degree of bit products before the last linear layer depend on bit positions.

The attack in general



More generally still, any low-degree polynomial will do.

Cryptanalysis of SASASASAS

Short article by Biryukov et Khovratovich: the same attack extends ASASASA and even SASASASAS [BK15].

Indeed the main obstacle is that the overall function must not be full degree.

Conclusion

- A new generic attack on ASASA-type structures.

- Not presented: LPN-based attack on the χ-based scheme, heuristic attacks on white-box scheme.

- Regarding multivariate ASASA proposals, [GPT15] and our result are somewhat complementary.

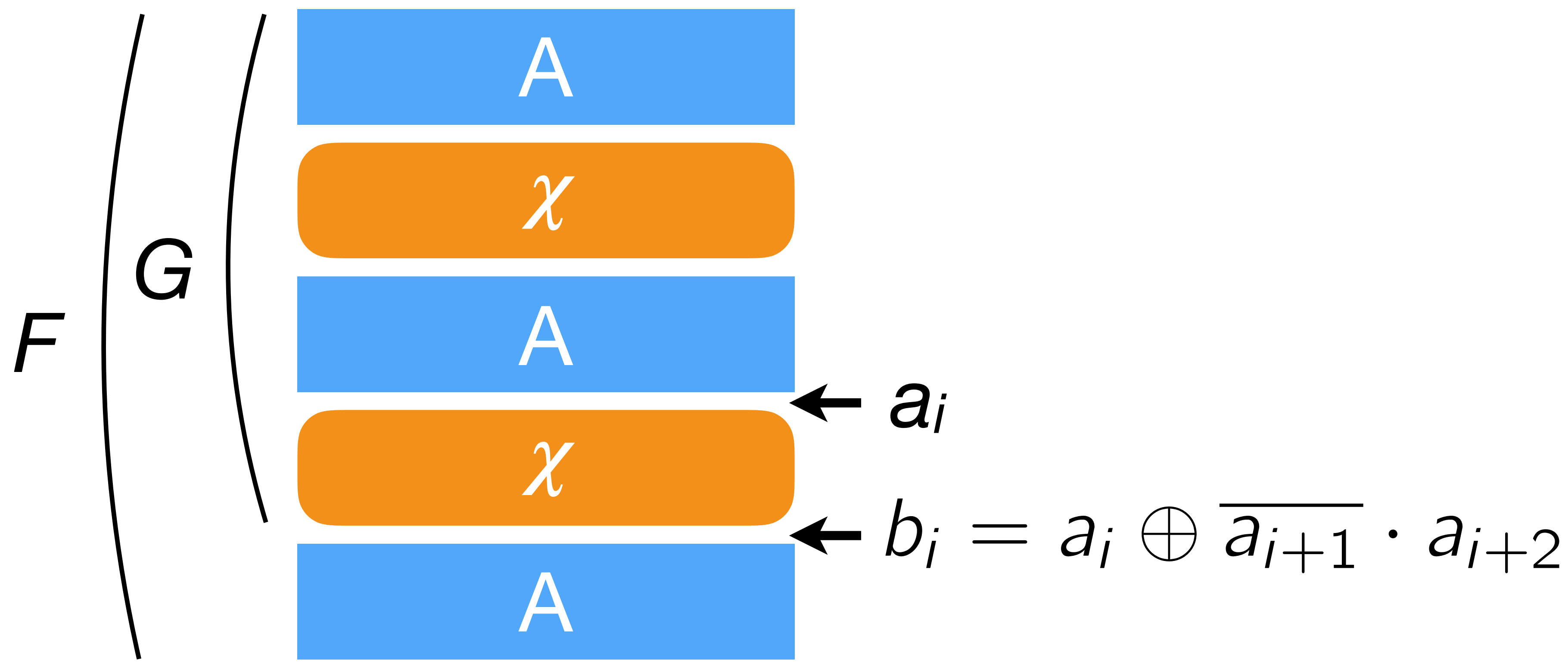
- Open problems:

  Other applications of this type of attack.

  Secure white-box scheme.

Thank you for your attention!

LPN-based attack



If we differentiate G twice along two arbitrary vectors d_1, d_2:

$$\begin{aligned} G_i''(x) &= a_i''(x) \oplus (\overline{a_{i+1}} \cdot a_{i+2})''(x) \\ &= C \oplus P_i(x) \oplus P_i(x \oplus d_1) \oplus P_i(x \oplus d_2) \oplus P_i(x \oplus d_1 \oplus d_2) \\ &\quad \text{with } P_i = \overline{a_{i+1}} \cdot a_{i+2} \end{aligned}$$

LPN-based attack

G'' is a constant + four products.

- Each bit of G'' has bias 2^{-4} (heuristically).

- Each computation of F''(x) yields a fresh sample of a binary vector *a* s.t. there exist *n* (fixed) values *s* s.t. *a*·*s* has bias 2^{-4}.

→ Can be (heuristically) solved by BKW.
(est. 2^{56} time, 2^{50} data).