PHILIPS

Read-Proof Hardware from Protective Coatings

CHES 2006, Tokyo-Yokohama

Pim Tuyls
G.J. Schrijen, B. Skoric, J. van Geloven, N.Verhaegh, R. Wolters

Philips Research Eindhoven
The Netherlands
Pim.tuyls@philips.com

PHILIPS

Read-Proof Hardware from Coating PUFs

CHES 2006, Tokyo-Yokohama

Pim Tuyls
G.J. Schrijen, B. Skoric, J. van Geloven, N.Verhaegh, R. Wolters

Philips Research Eindhoven
The Netherlands
Pim.tuyls@philips.com

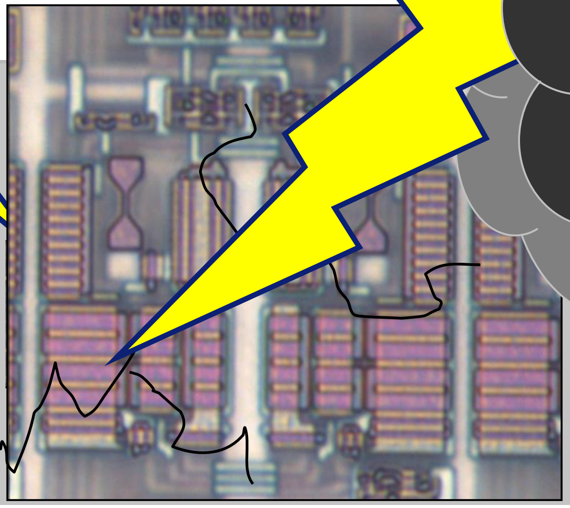
Contents

- Limitations of the Black-Box Model
- Brief Overview of Physical Attacks
- Security in a Physical World
- Methods and Requirements
- Components
  - Coating PUFs
  - Fuzzy Extractors/Helper data
- Secure Key Storage Device

Limitations of the Black-Box Model

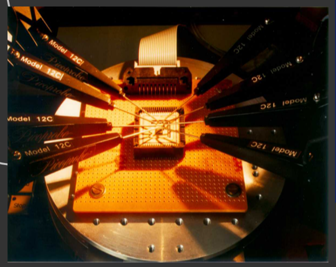
Mathematical Attacks
Protocol Attacks

Assumption:
IC: Black-Box
↓
Crypto guarantees
Security level

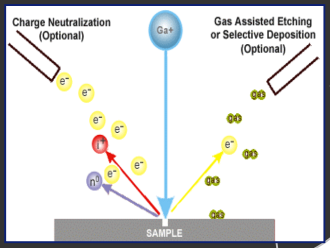


Secret Key: 001011101011

Physical Attacks



Micro Probes

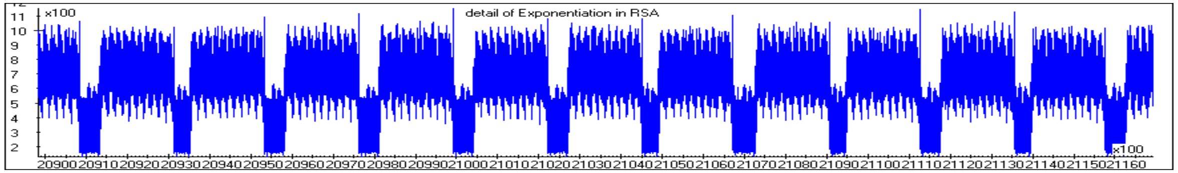


Focused Ion Beam

Security not guaranteed by cryptography

Brief Overview Physical Attacks

- Invasive Attacks
  - Micro Probing
  - Focused Ion Beams
  - Chemical
  - Mechanical
  - Etching
- Side Channel Attacks
  - Timing Analysis
  - Power Analysis
  - Electromagnetic Radiation
- Fault Induction (light, X-ray, power glitch)
- Optical Inspection



Security in a Physical World

Big Challenge: Develop theory and practical components for security in the presence of physical leakage: **No Black-Boxes!**

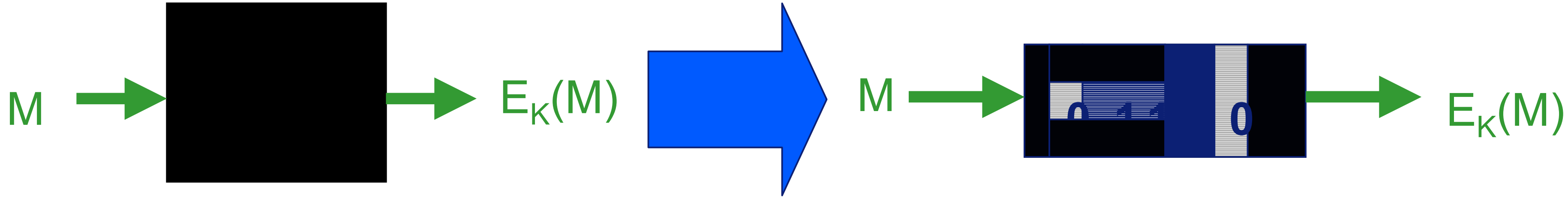
Components

1. Read-Proof Hardware:
   Enemy can not read the data stored in it
2. Tamper-Proof Hardware:
   Enemy can not change the data stored in it
3. Self Destruction Capability

Algorithmic Tamper Proof Security can be achieved [Gennaro et al]

Goal

Practical Methods



Focus: Read-Proof Hardware

Read-Proof Hardware is hardware where the attacker can not read any information on the data stored in it

Practical Meaning?!

Should be resistant against:

- Invasive Physical Attacks
- Side-Channel Attacks
- Fault Attacks
- Optical Inspection

Invasive vs Non-Invasive Attacks

Invasive Physical Attacks

Definition

An *invasive* physical attack is an attack where the attacker physically breaks into the device by modifying its structure

Examples:

- Chemical etching
- Drilling a hole
- Focused Ion Beam attack

Non - Invasive Physical Attacks

Definition

An non-*invasive* physical attack is an attack where the attacker physically breaks into the device without modifying its structure

Examples:

- Optical inspection of the memory
- Side-Channel attacks (Time, EMA, DPA, …)

Methods and Requirements

In order to protect keys against physical attacks:

1. Do **not** store a key in digital form in a device
2. Generate the key only when needed
(extract it from a physical source on the IC)
3. Delete the key

Components

Two components are needed:

1. **Hardware component (Physics)**
    1. Physical Source
2. **Cryptographic component**
    1. Fuzzy Extractor/Helper data algorithm

Hardware Requirements

Security Requirements:

1. Physical Inscrutability (opaqueness)
2. Unclonability
    1. Physical Unclonability
    2. Mathematical Unclonability
3. Tamper evident: key is destroyed upon damage

Practicality Requirements:

1. Easy to challenge the source
2. Cheap and easy integratable on an IC
3. Excellent mechanical and chemical properties

Components: Physical Source

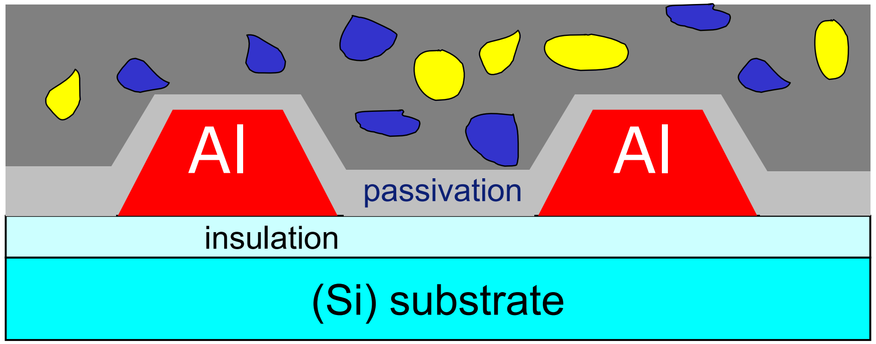
Physical Unclonable Function (PUF):

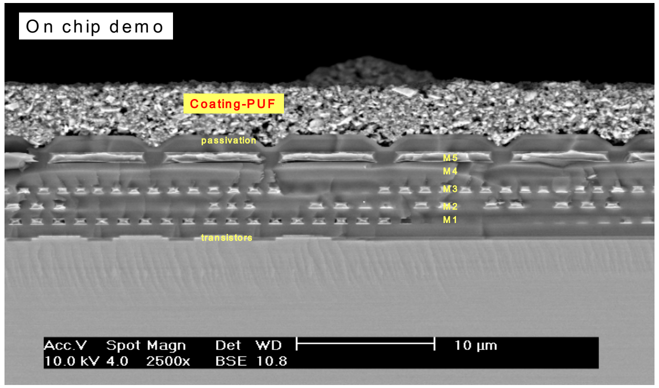
Inherently unclonable Physical Structure (consisting of many random/uncontrollable components) satisfying:

- Easy to evaluate: Challenges-Responses
- Responses are unpredictable
- Inherently tamper evident
- Manufacturer not-reproducable
- Extract keys from measurements

Coating PUF

- An IC is covered with an opaque coating containing random particles with high ε_r
- Array of capacitive sensors in upper metal layer detects local coating properties.
- Inhomogeneous coating → random capacitive properties





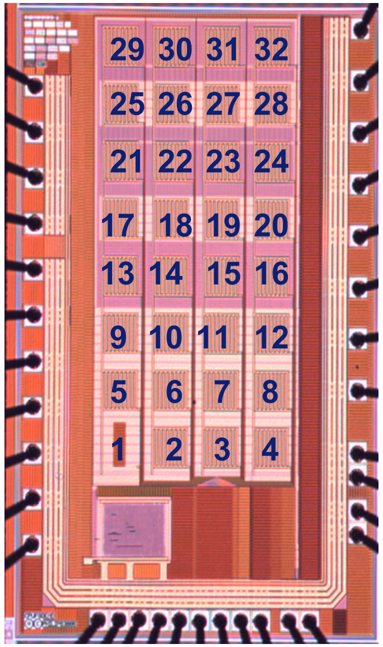
- PUF is used as a source of secret random information which are derived from the local coating capacitances (secure key storage).

Information Content of a Coating PUF (Response)

Coating PUF [JAP06]

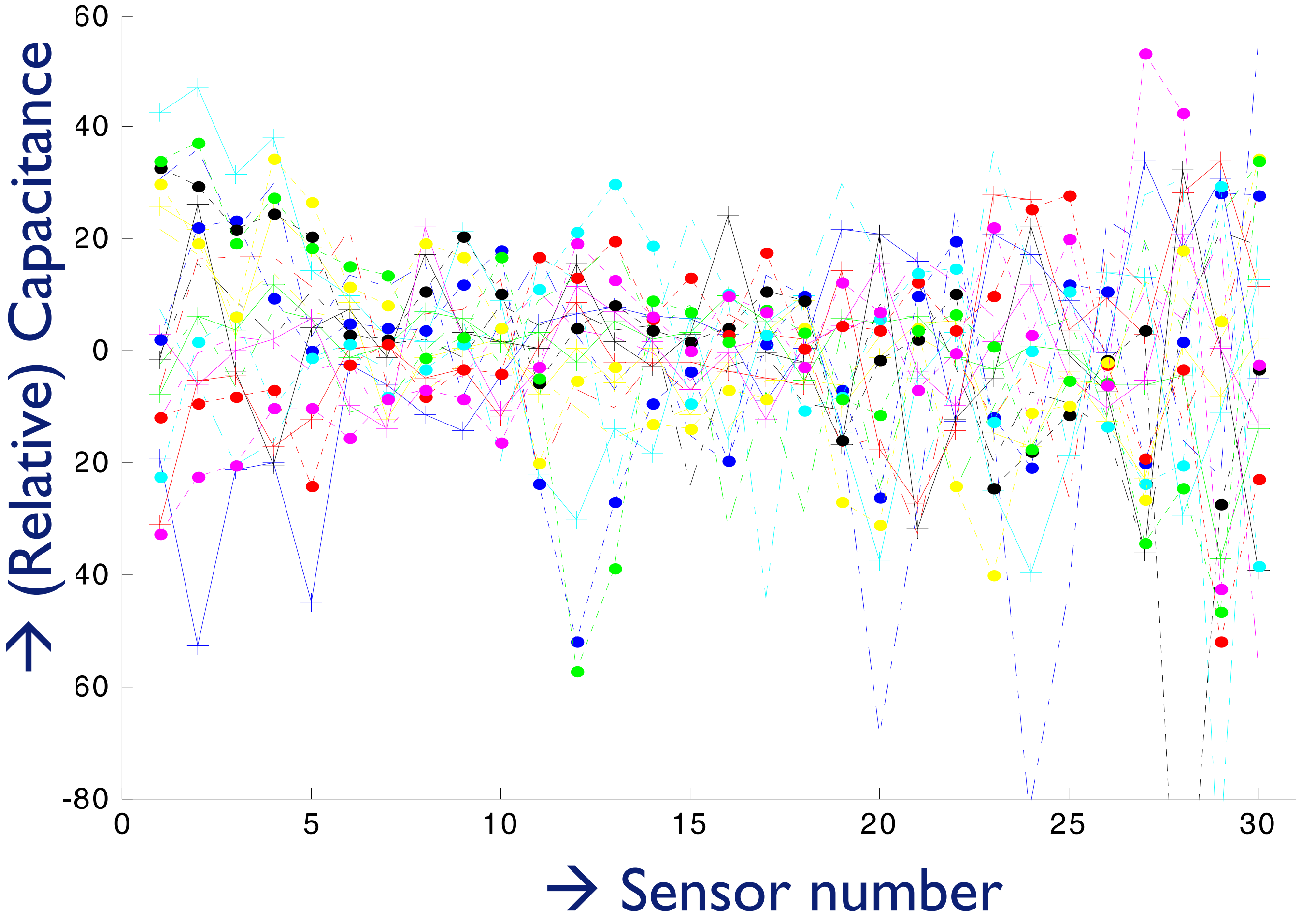
$$H = \log\left[\frac{\sqrt{2\pi e}}{\sigma_{\mathrm{N}}}\frac{A\varepsilon_0}{d}\sqrt{\frac{q(1-q)}{Ad/s^3}}\frac{|\varepsilon_1^{-1}-\varepsilon_2^{-1}|}{[(1-q)\varepsilon_1^{-1}+q\varepsilon_2^{-1}]^2}\right]$$

≈ 6.6 bits/sensor



Capacitance values of 21 ICs

$(C_x-C_0)/(C_{ref}-C_0)$



Fuzzy Extractor/Helper Data Algorithm

- Information present in the PUF has to be extracted
    - Measurements (Challenges - Responses)
- Measurements on Physical Systems are noisy
- Noisy values can not be used as keys in cryptography
- A Fuzzy Extractor/Helper Data Algorithm is needed

Key Extraction from PUFs: Fuzzy Extractor

- Grid points represent ECC Code words

Enrollment

- Random codeword C(S) is chosen
- Response X is measured
- Helper data **W** is generated (difference between X and C) and stored in EEPROM
- Key **K** is generated and its public key **P(K)** is output and the Key **K** is destroyed

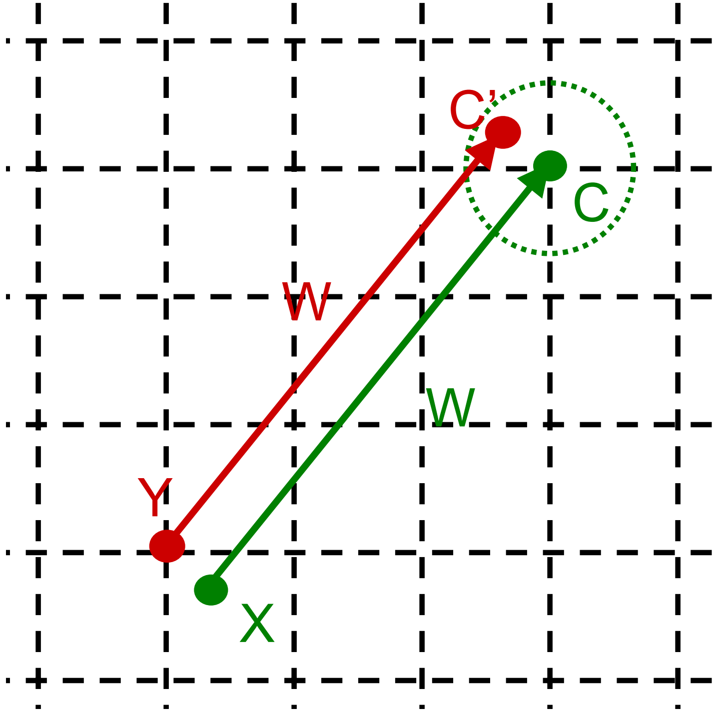
Key Reconstruction

- Y is noisy response
- Y+W=C'
- S'=DEC(C')

Security Condition

- $I(K;W) \leq \varepsilon$

Assumption: Response X uniformly random



Properties

- The parameter ε can be made negligible in the security parameter

- The maximal length of a secret key is given by

$$\mathbf{I(X;Y)}$$

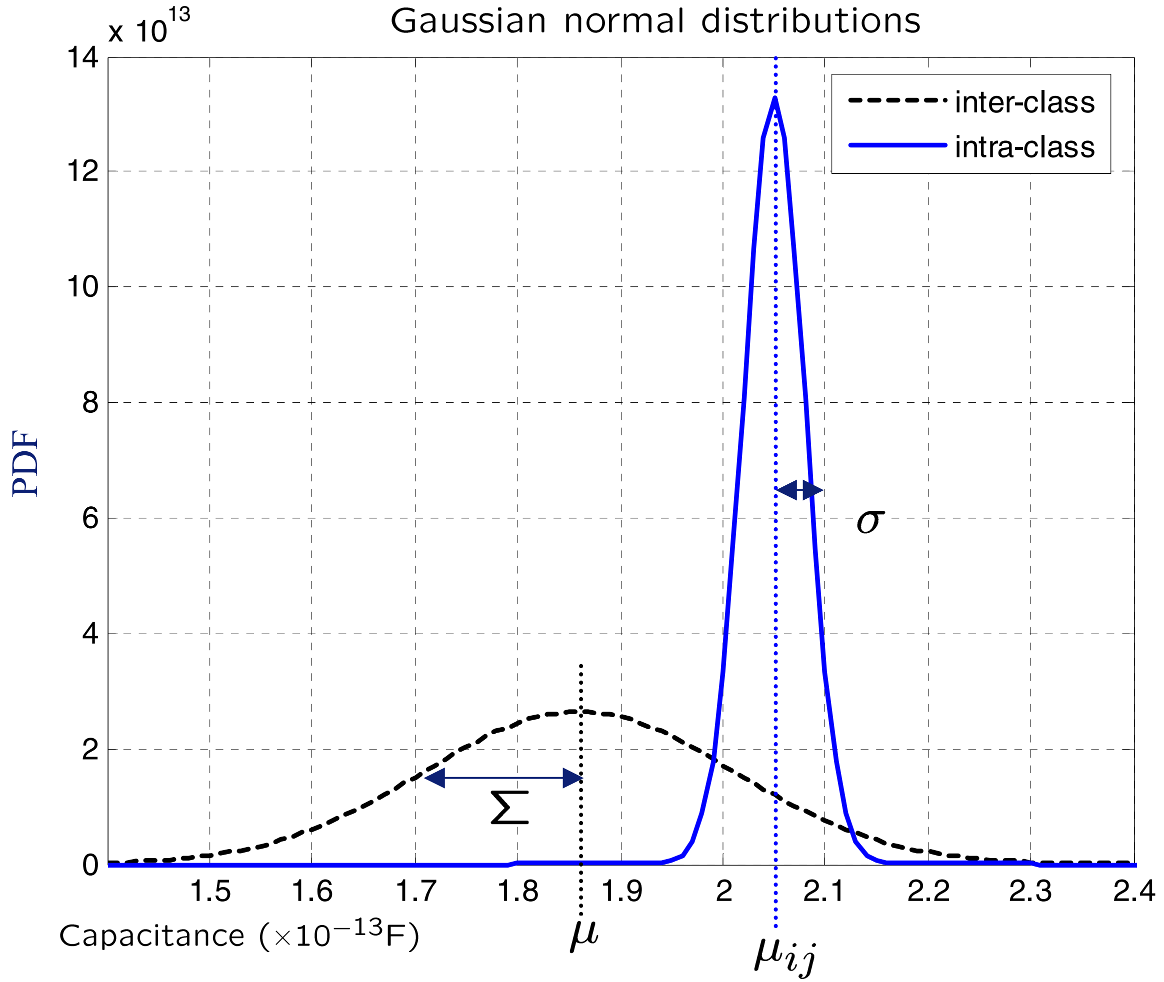
where I(X;Y) is the mutual information between

Enrollment: X → Key Reconstruction: Y

Practical Key extraction requirements

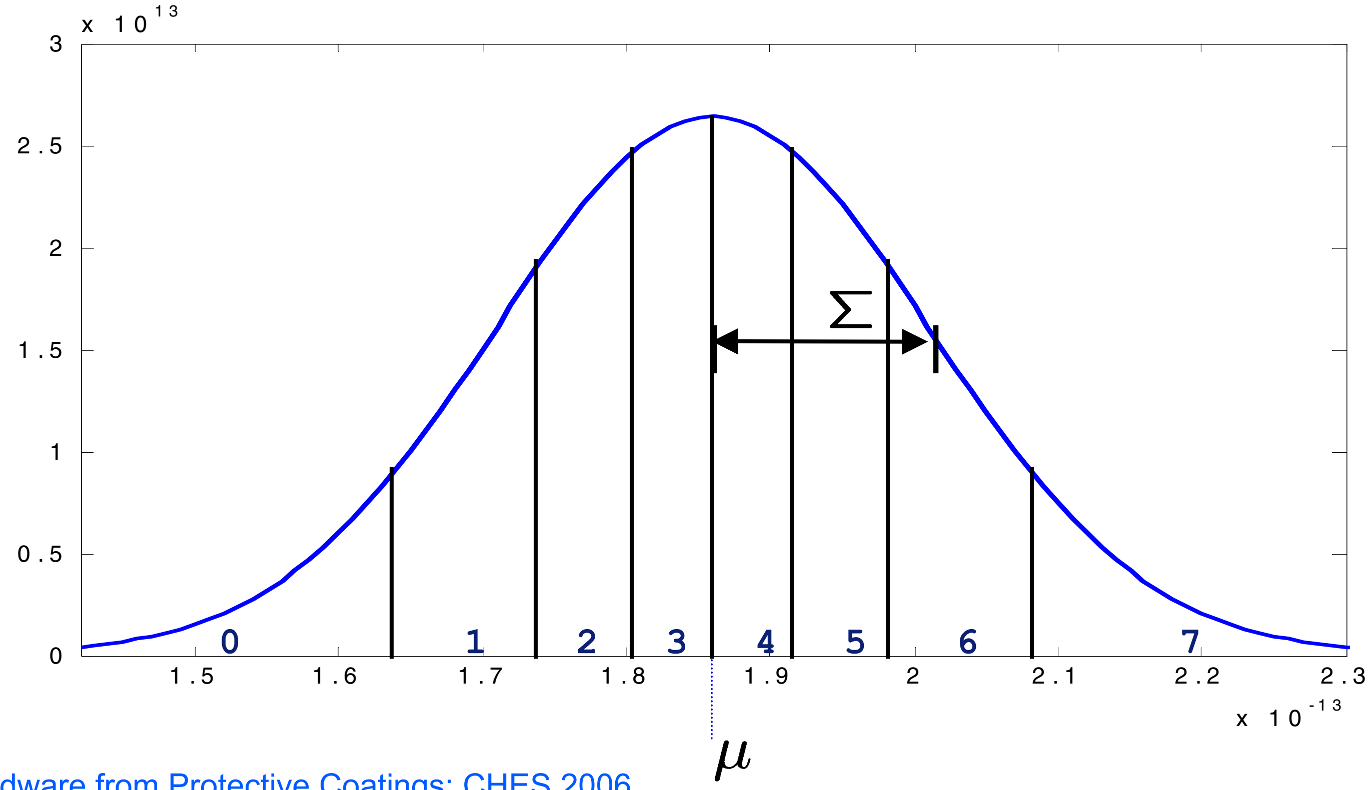
- **Measured Data are continuous, not discrete!**
- **Uniformly Distributed Keys**: All possible *n*-bit keys must be equally probable.
- **Robustness**: key extraction must be reproducible, regardless of measurement noise.

Statistics


Gaussian normal distributions
x 10^13
14
12
10
8
6
4
2
0
PDF
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4
Capacitance (×10^−13F)
inter-class
intra-class
Σ
σ
μ
μ_ij

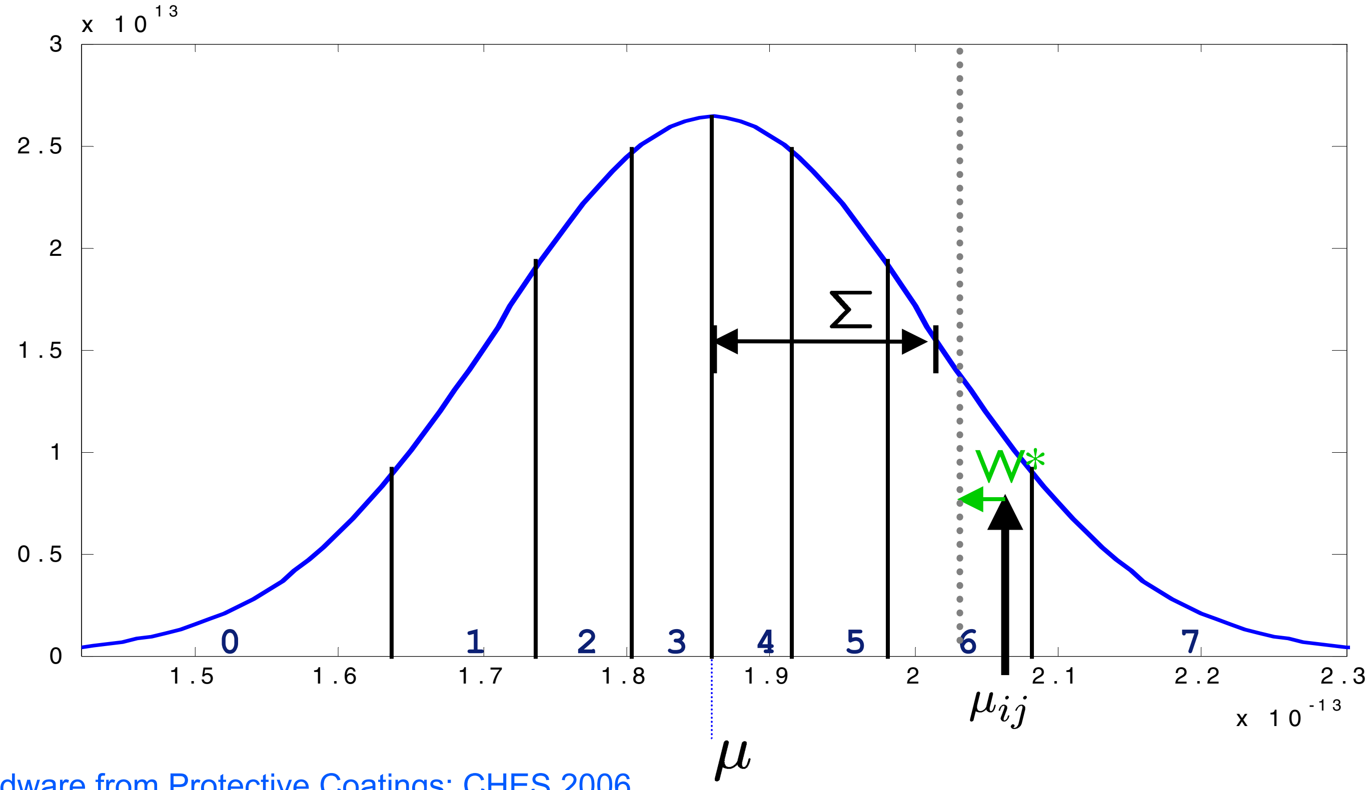
Uniformly Distributed Keys

- Quantization with equiprobable intervals



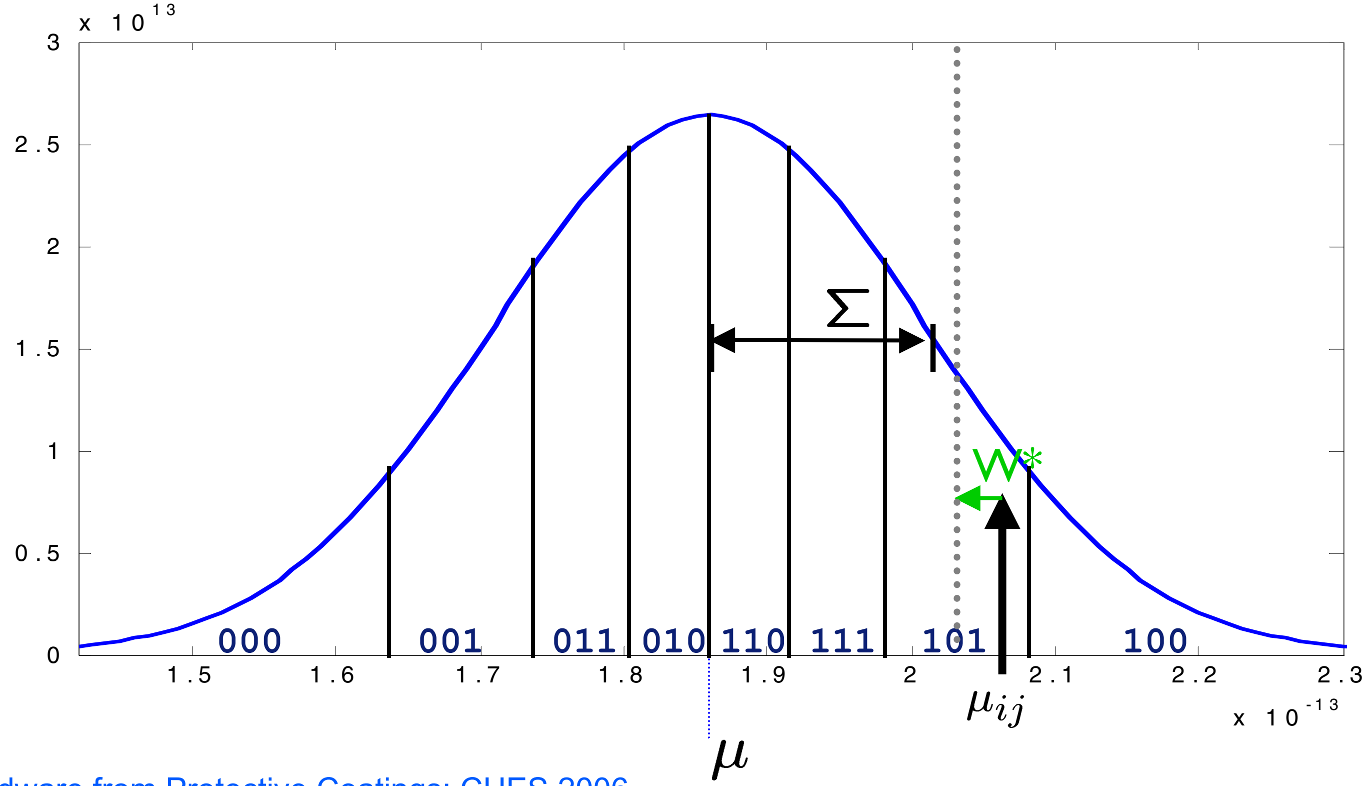
Achieving Robustness (I)

- Define helper-data *W** that shifts measurements to the center of a quantization interval.



Achieving Robustness (II)

- Assign bits to quantization intervals according to a Gray-code.



Achieving Robustness (III)

- Concatenate bits from multiple sensors to construct a key of length *n*.
- Use an Error Correcting Code (ECC) and the XOR-Fuzzy Extractor:

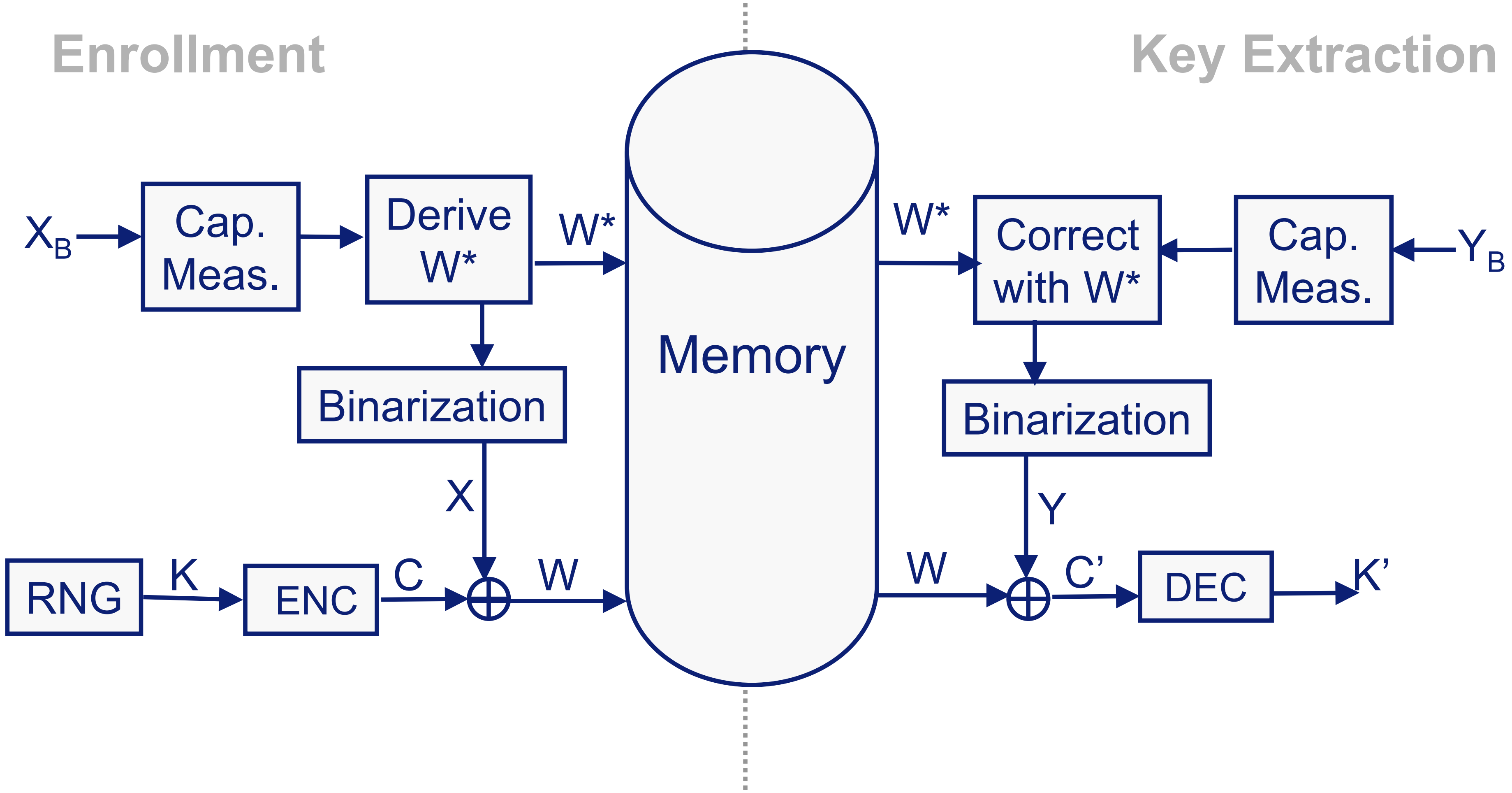
Enrollment: $K, W = X \oplus C_K$

Key Reconstruction: $\mathrm{Dec}(Y \oplus W)$

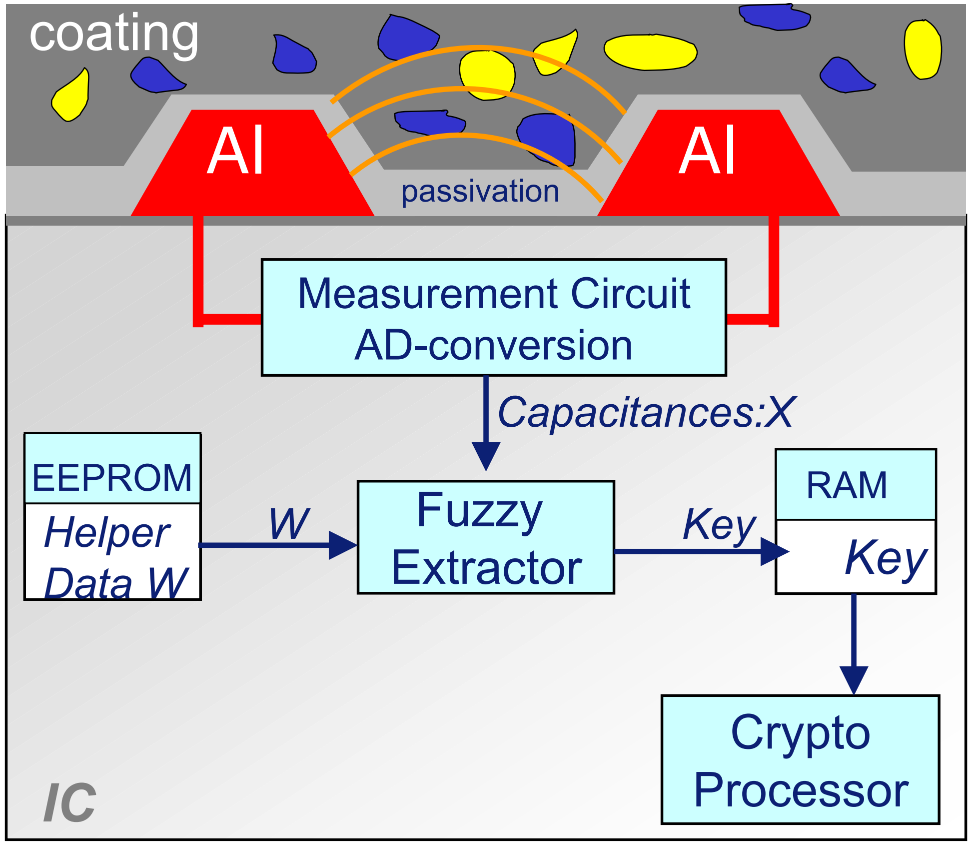
$= \mathrm{Dec}(Y \oplus X \oplus C_K)$

$= C_K$ iff $d(X,Y) < T$

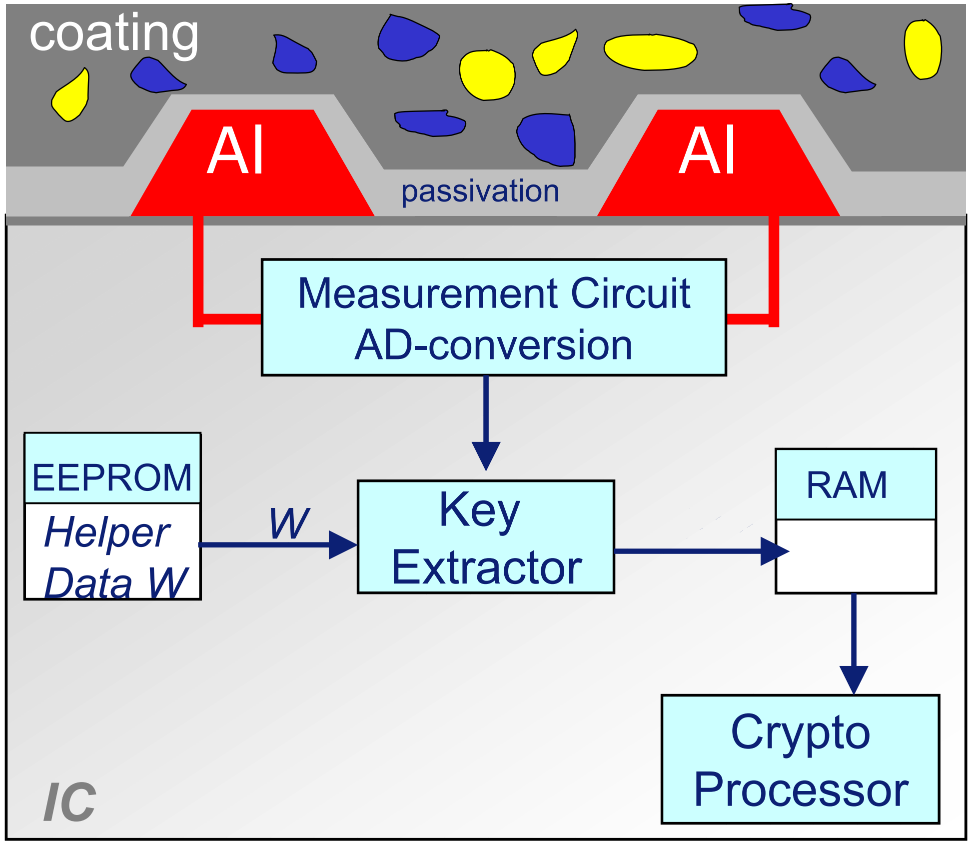
Key Extraction, helperdata scheme



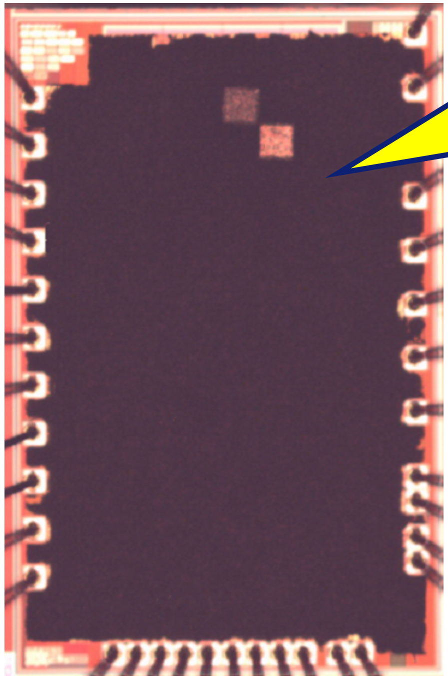
Store key temporarily in Volatile Memory



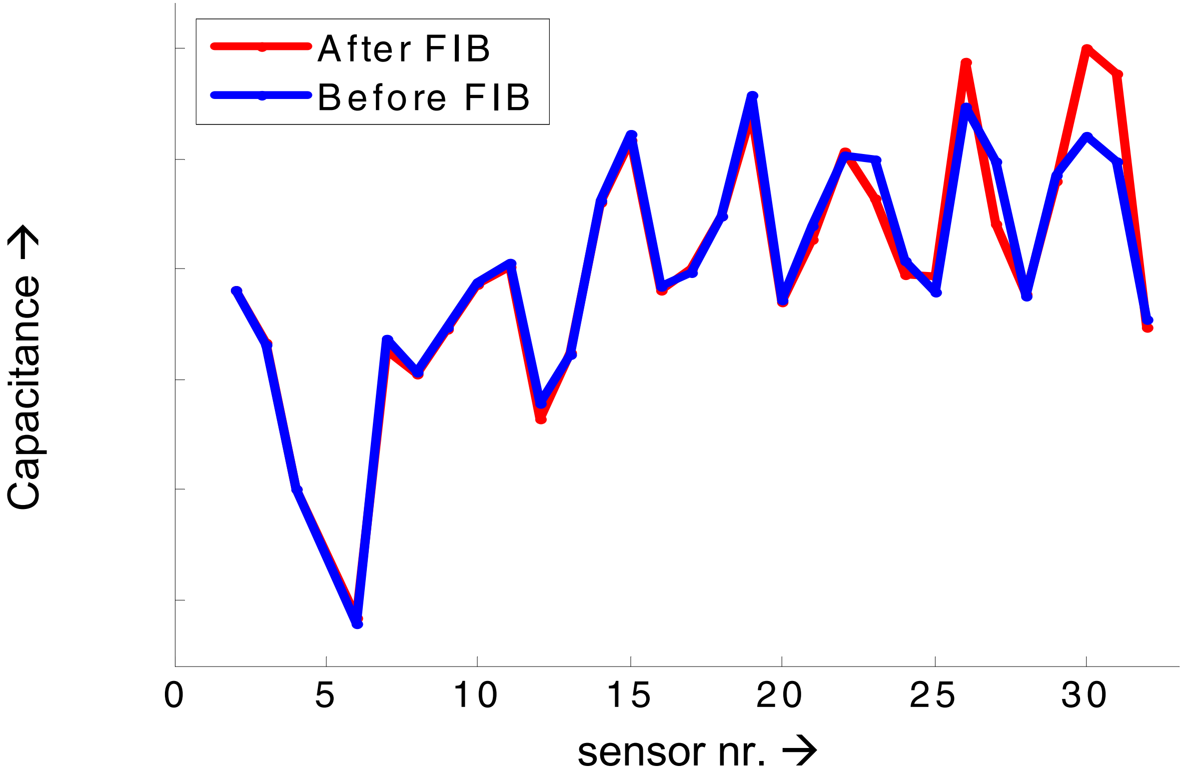
Delete key afterwards



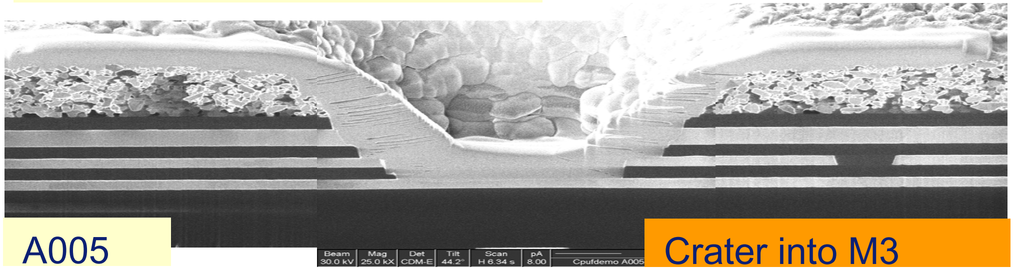
Attack Detection



Focused Ion Beam Attack



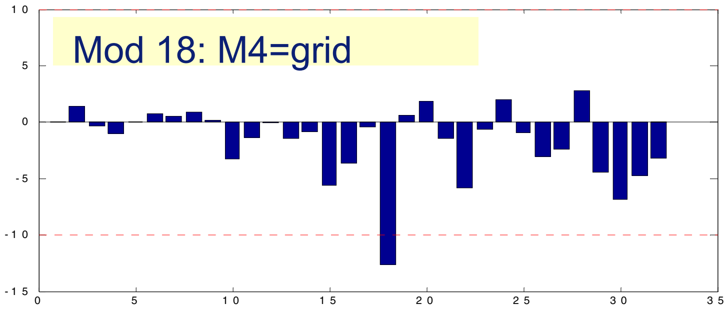
Craters: 10 µm x10 µm

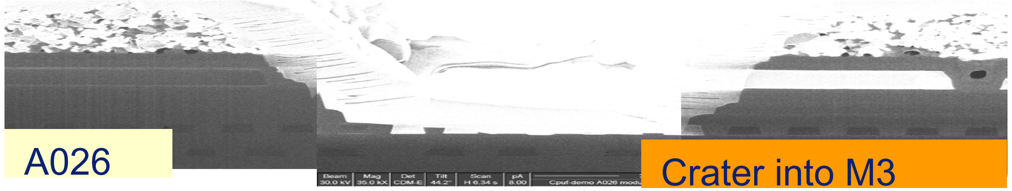


A005

Crater into M3

3.0-3.5 coating

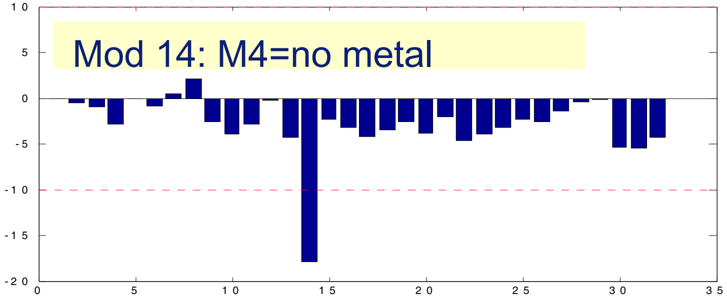


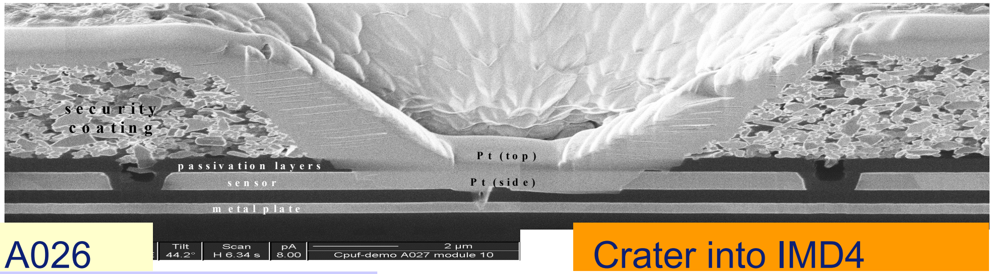


A026

Crater into M3

3.0-3.5 coating

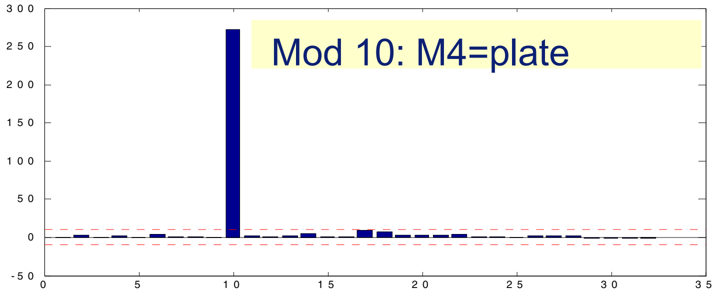




A026

Crater into IMD4

6.0-6.5 coating



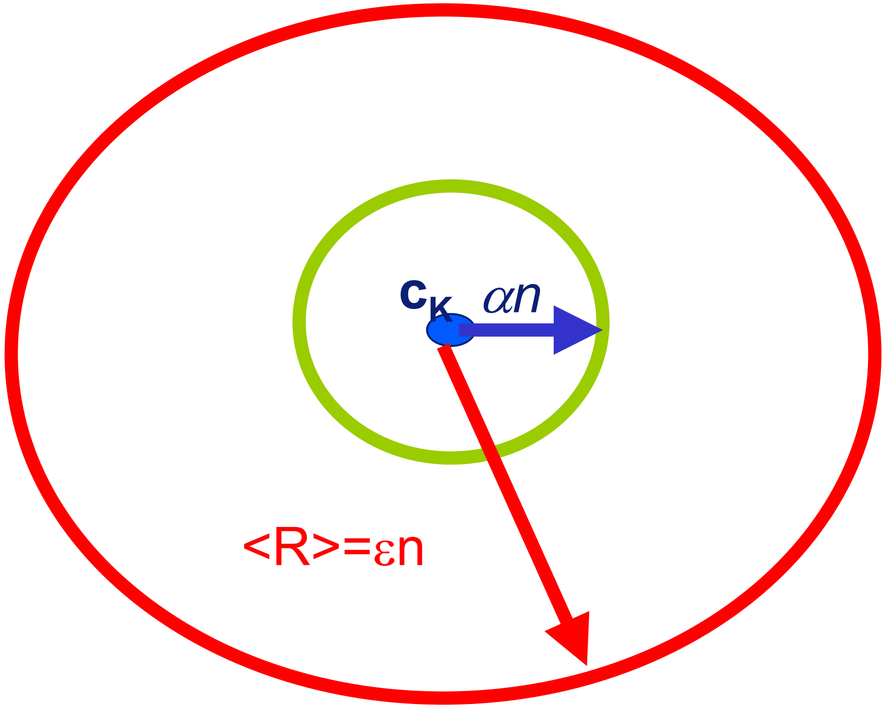
Next: craters of 5x5 mu

Model of Key Damage

Unattacked Device: Measurement Channel: $X \rightarrow Y$ Model BSC: Error Rate: α

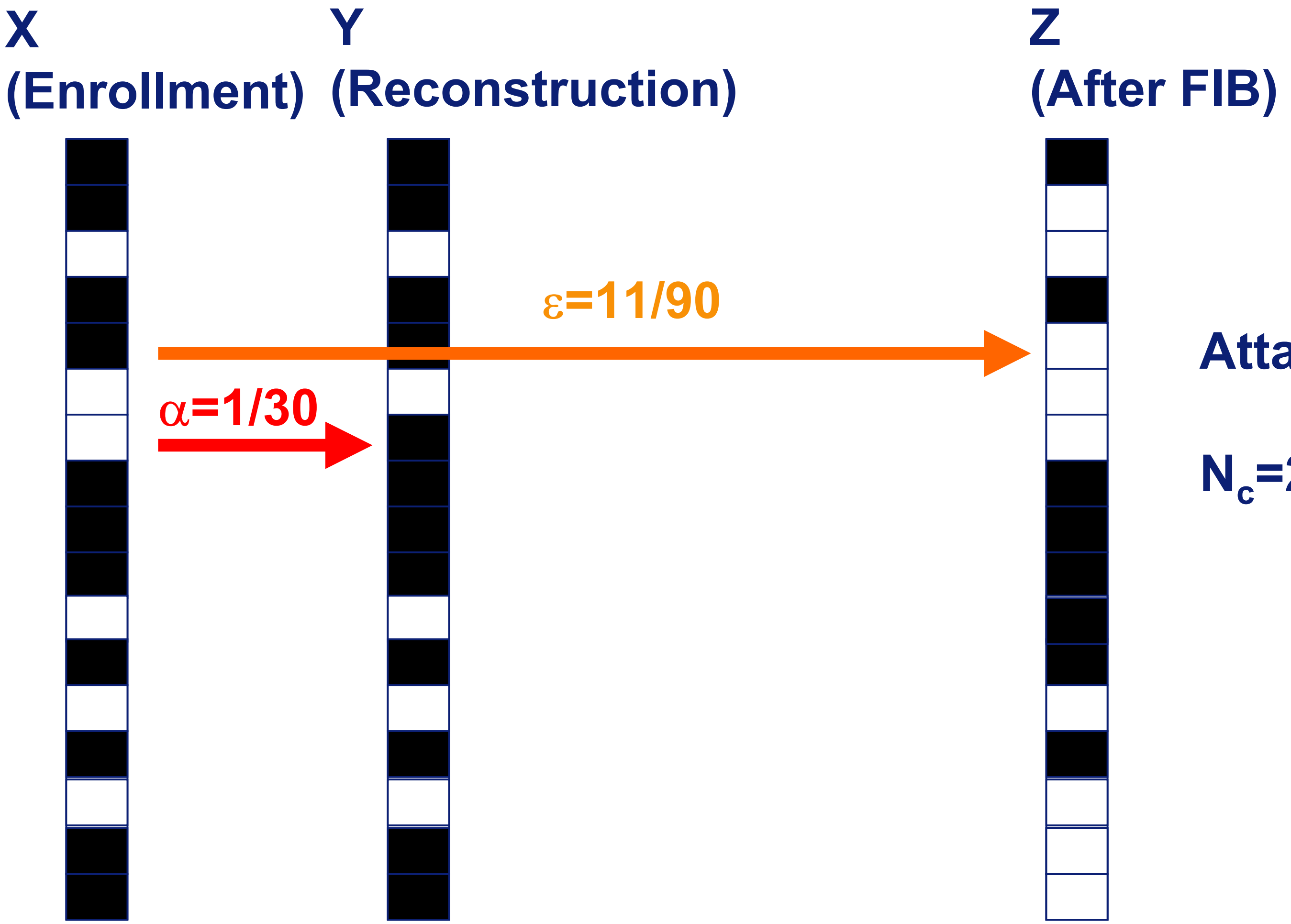
Attacked Device: Measurement Channel: $X \rightarrow Z$ Model BSC: Error Rate: ε

Fuzzy Extractor corrects αn errors



N_c= density of codewords x volume ball = $2^{n(h(\varepsilon)-h(\alpha))}$

Key Damage: Experiments



X
(Enrollment)

Y
(Reconstruction)

Z
(After FIB)

$\varepsilon=11/90$

$\alpha=1/30$

Attack Complexity:

$N_c=2^{51}$ for 128 bit keys

Summary of Results

- Test ICs with 30 sensors per IC
- Deriving 3 bits per sensor → 90 bits per IC
- Limit error correction: 4 of the 90 bits
  - Depends on the coarseness of the quantisation
- Temperature compensation
- No humidity influence

Conclusions

- Developed Read-Proof Hardware (Invasive Attacks)
  - Coating PUF
  - Fuzzy Extractor
- Made a demonstrator
  - Attacks can be detected
  - Key Damage is shown
- Next Steps
  - Further investigate side-channel leakages
  - Investigate the impact of smaller holes

?