NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

pace Physical Analysis and Cryptographic Engineering

Fault Injection Attacks and Countermeasures

Brněnské bezpečnostní setkávání, FEKT VUT Brno

Jakub Breier

28 March 2018

Physical Analysis and Cryptographic Engineering
Nanyang Technological University
Singapore

PACE Laboratory

Our team:

- Principal investigator
- 6 researchers
- 1 PhD student

Our main focus:

- Side-channel attacks
- Fault attacks
- Hardware trojans
- Countermeasures

Outline

1. Physical Attacks on Cryptographic Systems
2. Fault Attacks
3. Laser Fault Attacks
4. Fault Attack on AES
5. Countermeasures
6. Conclusion

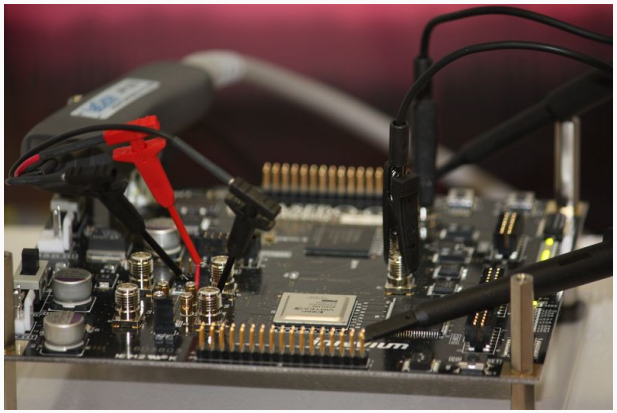
Physical Attacks on Cryptographic Systems

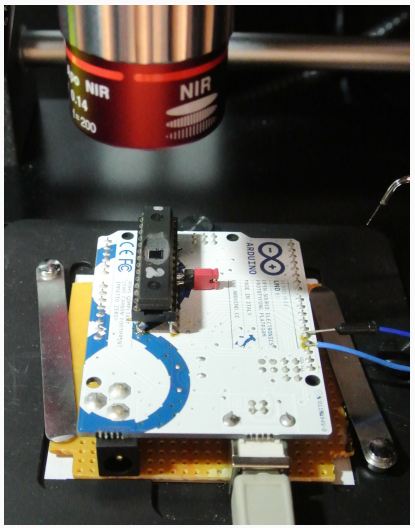
Why Physical Attacks?

- Cryptography provides algorithms that enable secure communication in theory
- In real world, these algorithms have to be implemented on real devices:
  - software implementations - general-purpose devices
  - hardware implementations - dedicated secure hardware devices
- To evaluate security level of cryptographic implementations, it is necessary to include physical security assessment

Classification

- **Fault attacks**
  - Optical fault injection
  - Electromagnetic fault injection
  - Clock/voltage glitch
- **Side-channel attacks**
  - Power analysis
  - Timing analysis
  - Electromagnetic analysis
  - Acoustic analysis
- Hardware Trojans
- Probing

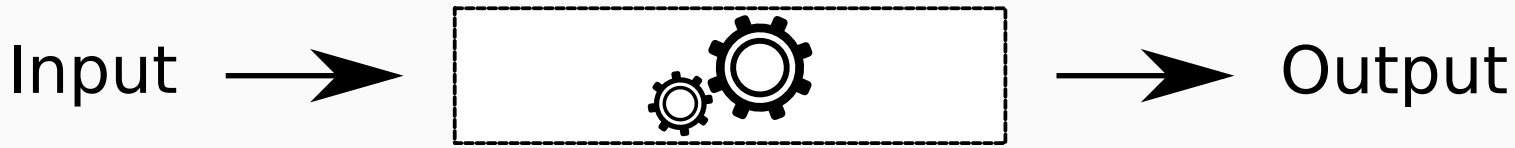




Why Fault Attacks?

- [1]The best cryptanalysis of AES needs complexity of $2^{126.1}$

- [2,3]The best fault attack on AES needs just one faulty and correct plaintext/ciphertext pair



[1]A. Bogdanov et al. Biclique cryptanalysis of the full AES. ASIACRYPT 2011.

[2]D. Saha et al. A diagonal fault attack on the advanced encryption standard. Cryptology ePrint Archive 2009/581.

[3]J. Breier et al. Laser Profiling for the Back-Side Fault Attacks: With a Practical Laser Skip Instruction Attack on AES. CPSS 2015.

Fault Attacks

Single Event Upsets

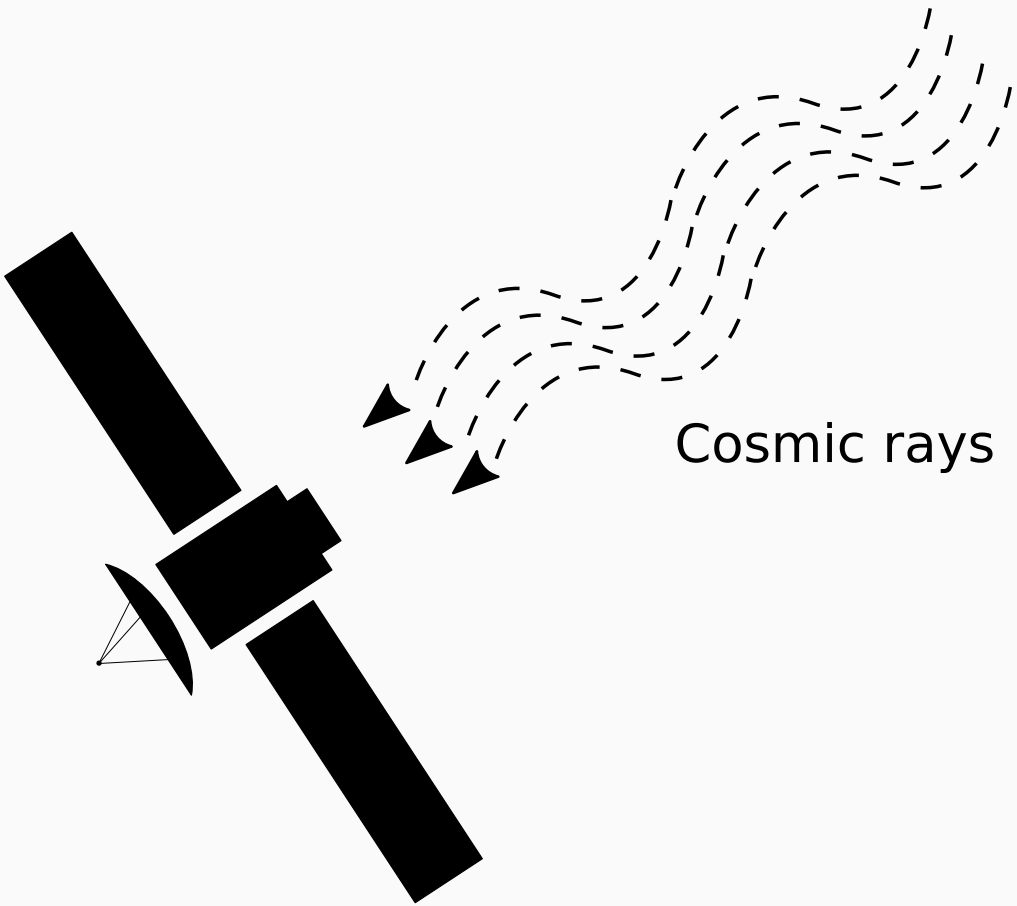


Figure 1: Cosmic rays and satellites[4].

[4] D. Binder et al. Satellite anomalies from galactic cosmic rays. IEEE Transactions on Nuclear Science, 1975.

Fault Attacks

- Fault attacks exploit the possibility to insert a fault in the process of the algorithm execution in a way that could help to reveal the key.
- The idea of fault attacks was introduced by Boneh, DeMillo and Lipton in 1996[5].
- The first practical attack was implemented by Biham and Shamir, introducing a Differential Fault Analysis on DES[6].

---

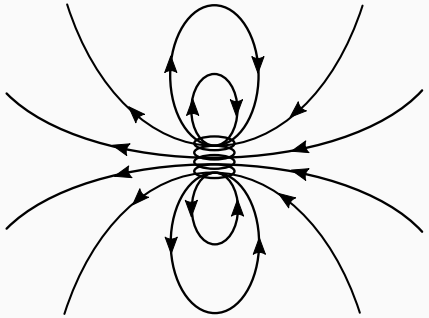
[5] D. Boneh, R. A. DeMillo, and R. J. Lipton. On the Importance of Checking Cryptographic Protocols for Faults, EUROCRYPT'97.

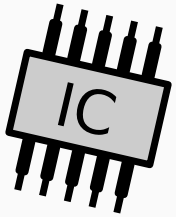
[6] E. Biham and A. Shamir. Differential Fault Analysis of Secret Key Cryptosystems, CRYPTO'97.

Fault Injection Techniques

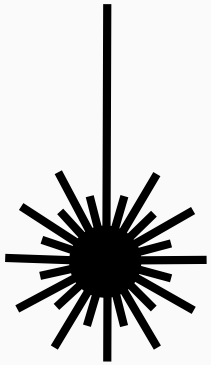
Voltage/clock glitch

EM field

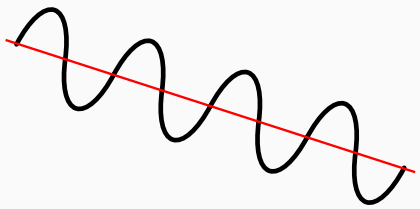




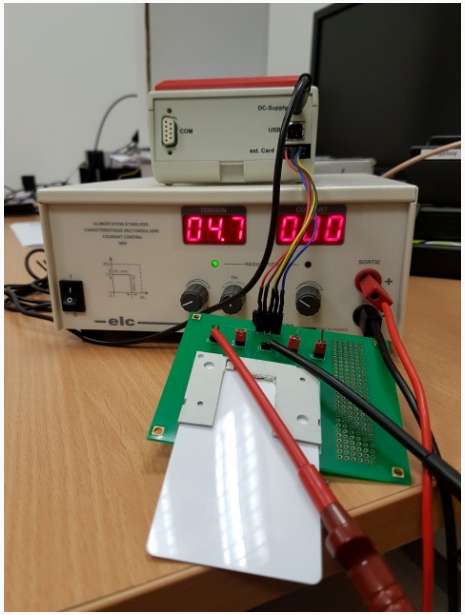
Laser



FIB/X-ray



Fault Injection Techniques in Practice

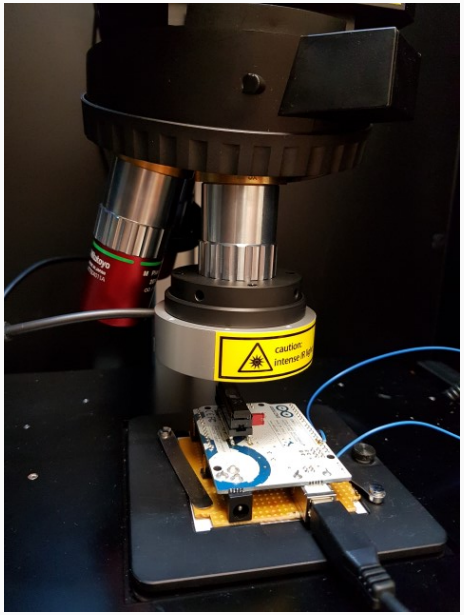


Voltage glitching

$

EM injection

$$



Laser fault injection

$$$

Fault Models

1. Precise bit errors
    - Attacker can cause a single bit fault.
    - Full control over the timing and location.
2. Precise byte errors
    - Attacker can cause a single byte fault.
    - Full control over the timing, partial control over the location.
3. Unknown byte errors
    - Attacker can cause a single byte fault.
    - Partial control over the timing and location.
4. Random byte errors
    - Partial control over the timing and no control over the location.

Fault Types

- Permanent faults
  - destructive faults
  - fault changing the value of a cell permanently
- Transient faults
  - circuit recovers its original behavior after reset or after fault's stimulus ceases
  - data or instruction is perturbed

Fault Attack Methods 1/3

- **Differential Fault Analysis** attacker injects a fault in a chosen round of the algorithm to get the desired fault propagation in the end of an encryption. The secret key can then be determined by examining the differences between a correct and a faulty ciphertext.
- **Collision Fault Analysis**[7] attacker invokes a fault in the beginning of the algorithm and then he tries to find a plaintext, which encrypts into the same ciphertext as the faulty ciphertext in the previous case, by using the same key.

[7] J. Blömer and J.-P. Seifert: Fault based cryptanalysis of the Advanced Encryption Standard. Cryptology ePrint Archive, Report 2002/075, 2002.

Fault Attack Methods 2/3

- **Ineffective Fault Analysis**[8] the goal is to find such fault that does not change the intermediate result, therefore it leads into a correct ciphertext. The attacker gains information which faults do not locally modify intermediate values.
- **Safe-Error Analysis**[9] also exploits a situation when ciphertexts are equal, but it changes the intermediate result. It utilizes a state when the data is changed but it is not used.

---

[8] J. Blömer and J.-P. Seifert: Fault based cryptanalysis of the Advanced Encryption Standard. Cryptology ePrint Archive, Report 2002/075, 2002.

[9] Yen, S.M., Joye, M.: Checking before output may not be enough against fault-based cryptanalysis. IEEE Transac. Comput. 49(9), 2000.

Fault Attack Methods 3/3

- **Fault Sensitivity Analysis**[10] exploits the side-channel information, such as sensitivity of a device to faults and uses this information to retrieve the secret key. It does not use values of faulty ciphertexts.
- **Differential Fault Intensity Analysis**[11] similarly to FSA, it tests the system responses under different fault intensity and takes advantage of a non-uniform distribution of the faults – a biased fault model. Unlike FSA, it does not require a fault sensitivity profile of the attacked device.

[10] Y. Li et al.: Fault Sensitivity Analysis, CHES 2010.
[11] N. F. Ghalaty et al.: Differential Fault Intensity Analysis. FDTC 2014.

Laser Fault Attacks

Advantages and Disadvantages of a Laser Fault Injection

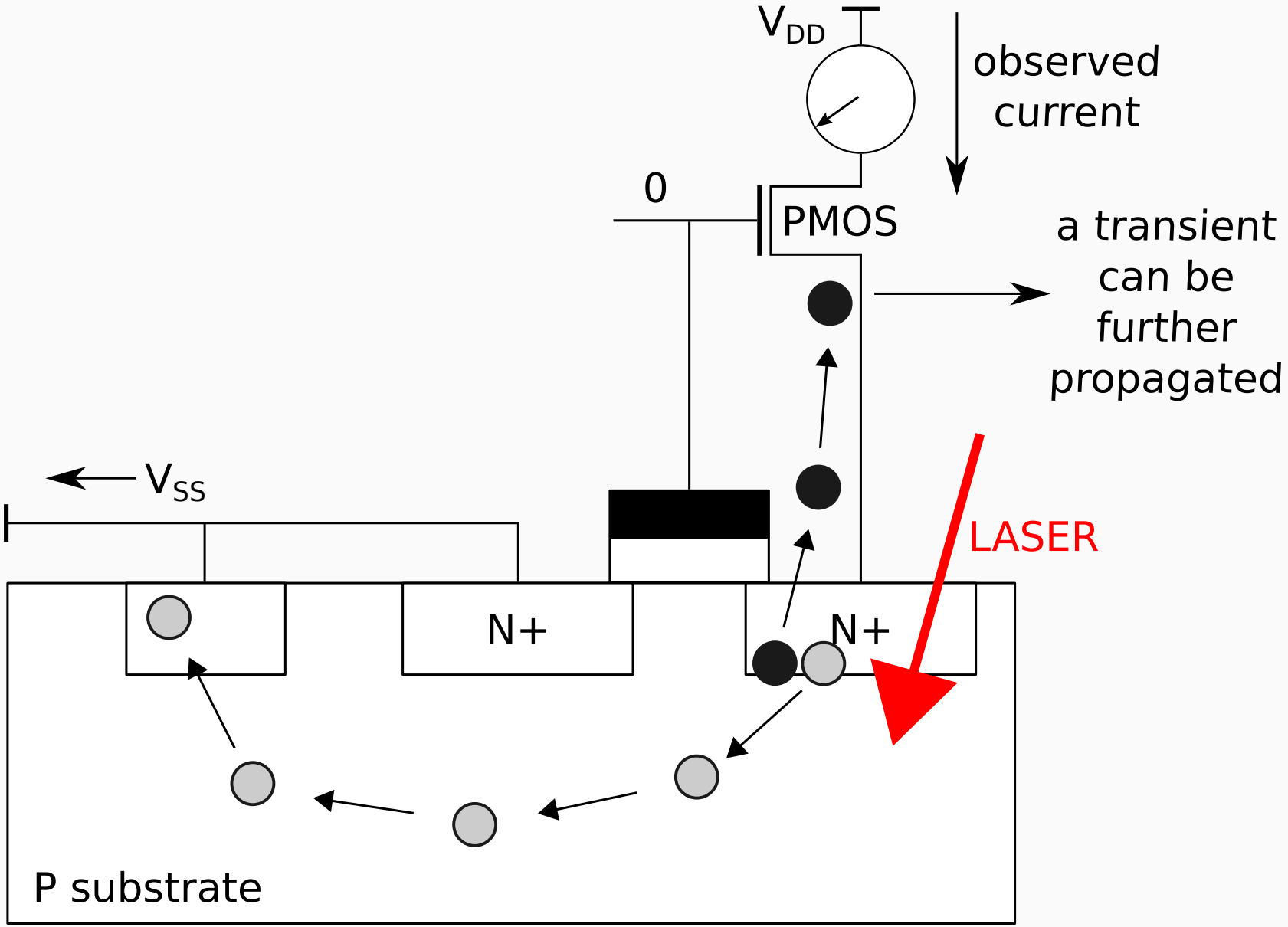
Advantages:

- Precision - beam diameter is usually few micrometers large.
- Reproducibility - identical faults can be repeated with same laser parameters.

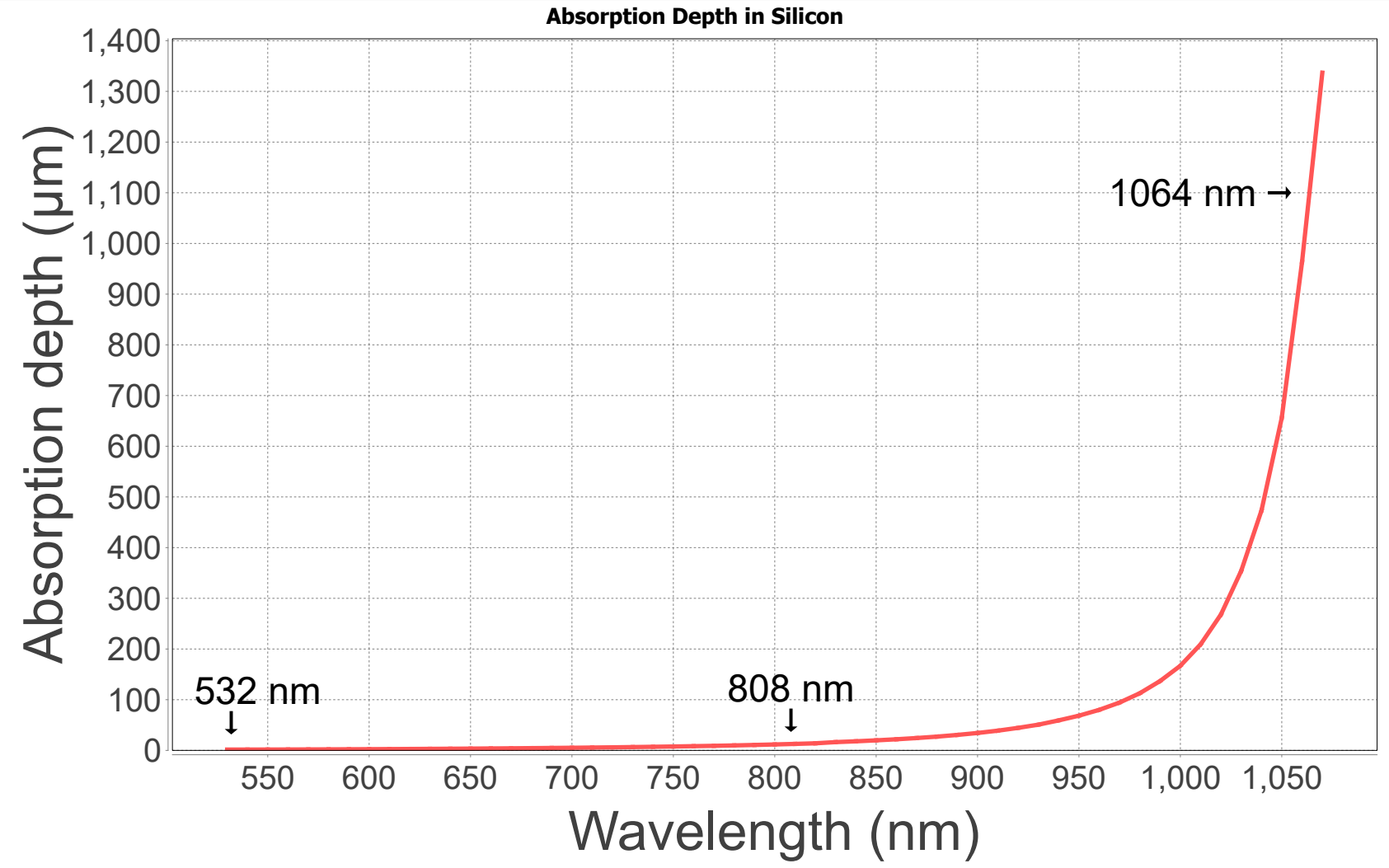
Disadvantages:

- Chip surface has to be accessible by the laser beam - need of de-packaging.
- Cost of the laser equipment is high.
- IC can be destroyed by large number of repetitions or by a high laser power.

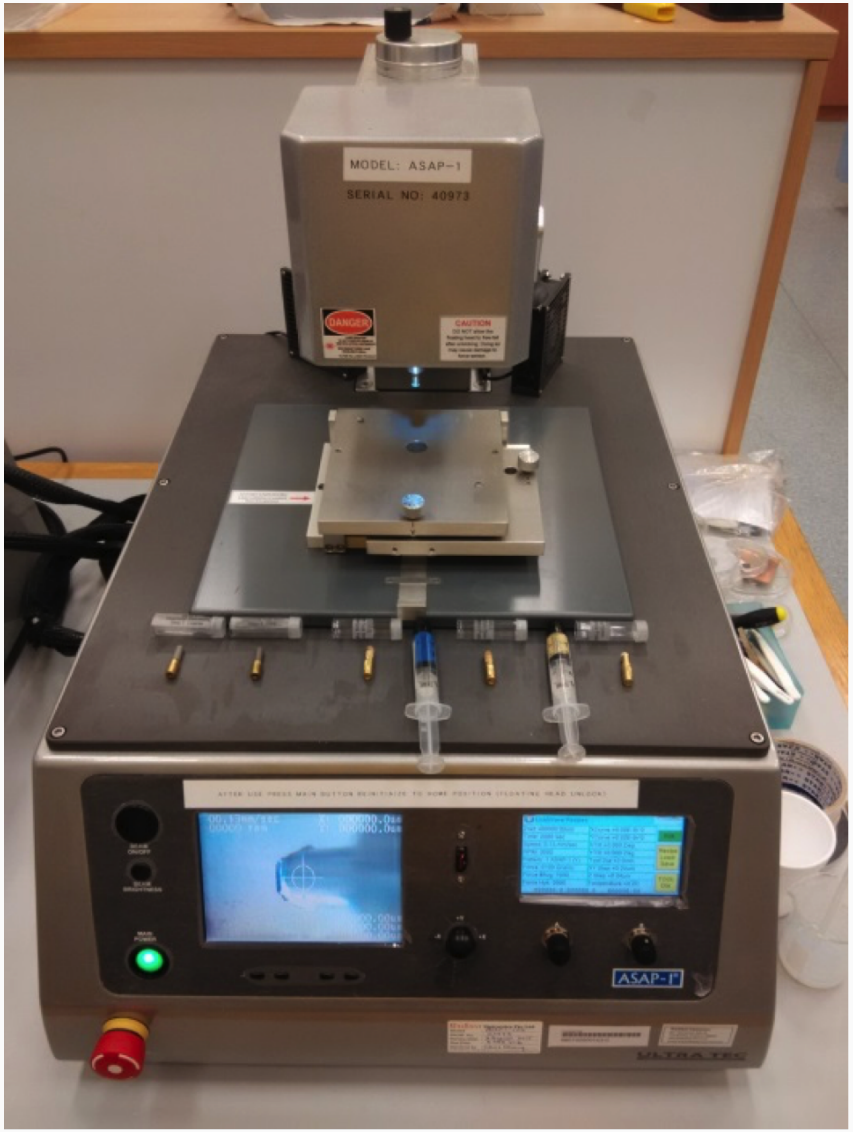
Irradiation Effect on Transistor



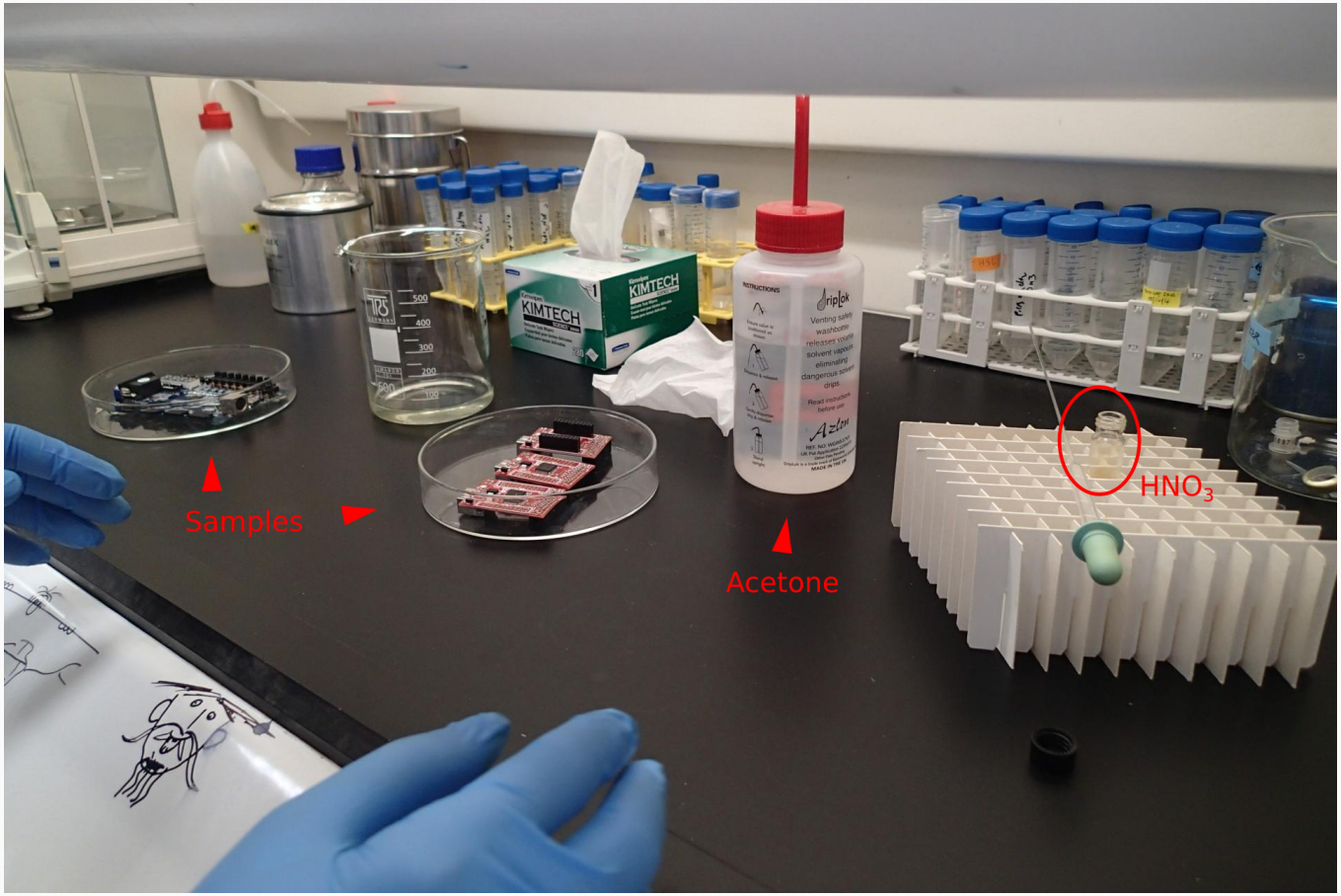
Absorption Depth in Silicon



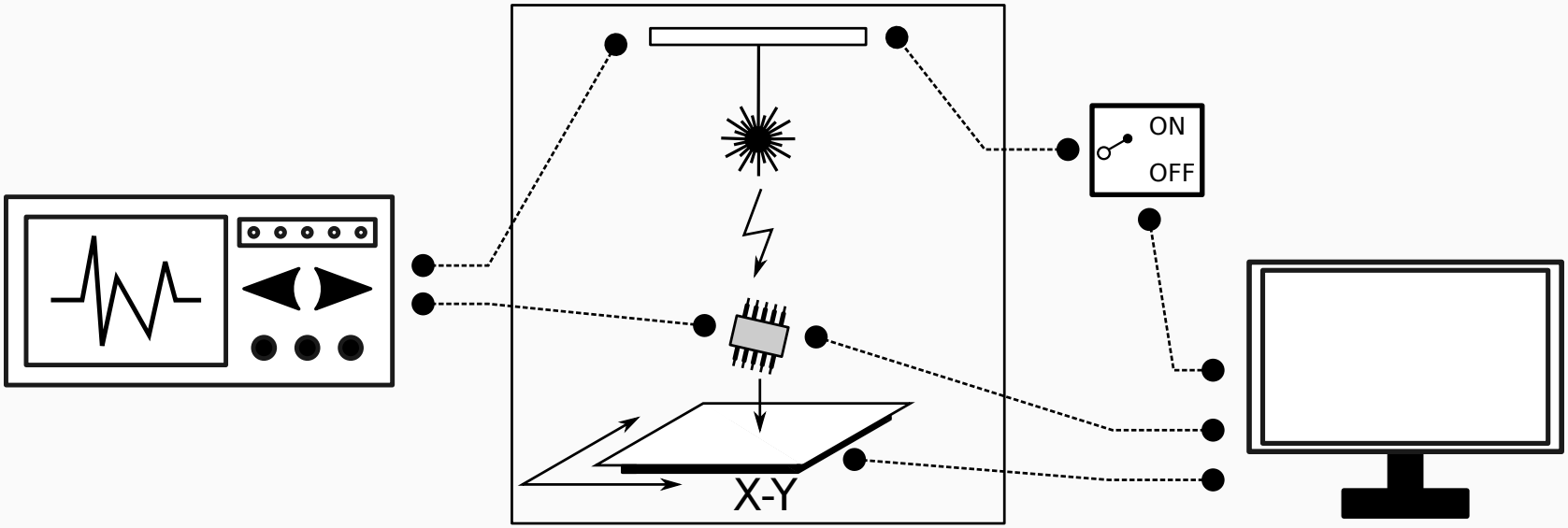
Decapsulation Techniques – Mechanical



Decapsulation Techniques – Chemical



LFI Setup



Our Setup

Near infrared diode pulse laser:

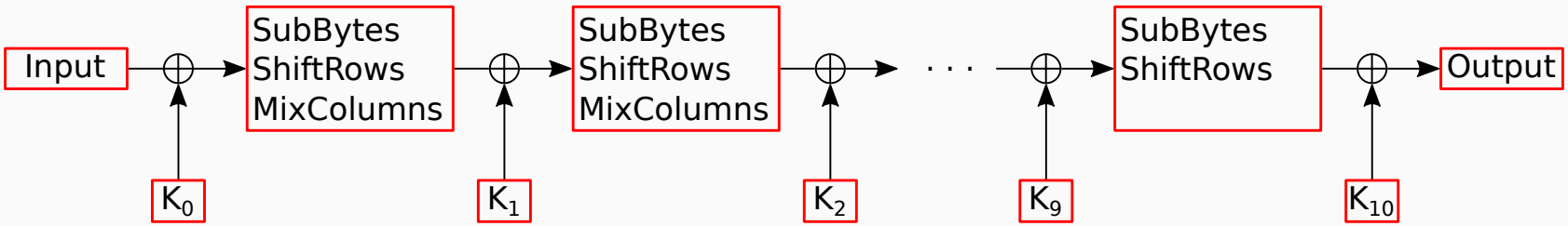
- Pulse power: 20 W (reduced to 8 W with 20× objective and 7 W with 50× objective)
- Pulse repetition: 10 MHz
- Spot size: 30×12 μm^2 (15×3.5 μm^2 with 20× objective and 6×1.4 μm^2 with 50× objective)
- Response to trigger pulse: $\leq$60 ns

Device under test:

- Atmel ATmega328P (8-bit microcontroller)

Fault Attack on AES

AES-128



Key - key expansion generates round keys K_0 - K_{10} from the 16B secret key K

Figure 2: Schematic diagram of AES-128.

- 10 rounds
- 4×4 bytes state matrix
- AES key schedule is reversible

Fault Attacks on AES

- The first attack on AES was proposed by Giraud in 2002 (published in 2003) using DFA technique[12]
- He could reveal the AES-128 key either by using 50 faulty ciphertexts by inducing bit faults or 250 faulty ciphertexts by using the byte fault model

[12] C. Giraud. DFA on AES. Cryptology ePrint Archive, Report 2003/008, 2003.

Diagonal Fault Attack[13]

- Most powerful attack on AES
- Fault is injected in one of the four diagonals of AES state matrix at the input of the eighth round
- Single faulty ciphertext reduces a key search space to 2^{32}
- If the fault corrupts two or three diagonals, 2 and 4 faulty ciphertexts can still recover the key

[13] D. Saha, D. Mukhopadhyay, and D. Roychowdhury: A Diagonal Fault Attack on the Advanced Encryption Standard. Cryptology ePrint Archive, Report 2009/581, 2009.

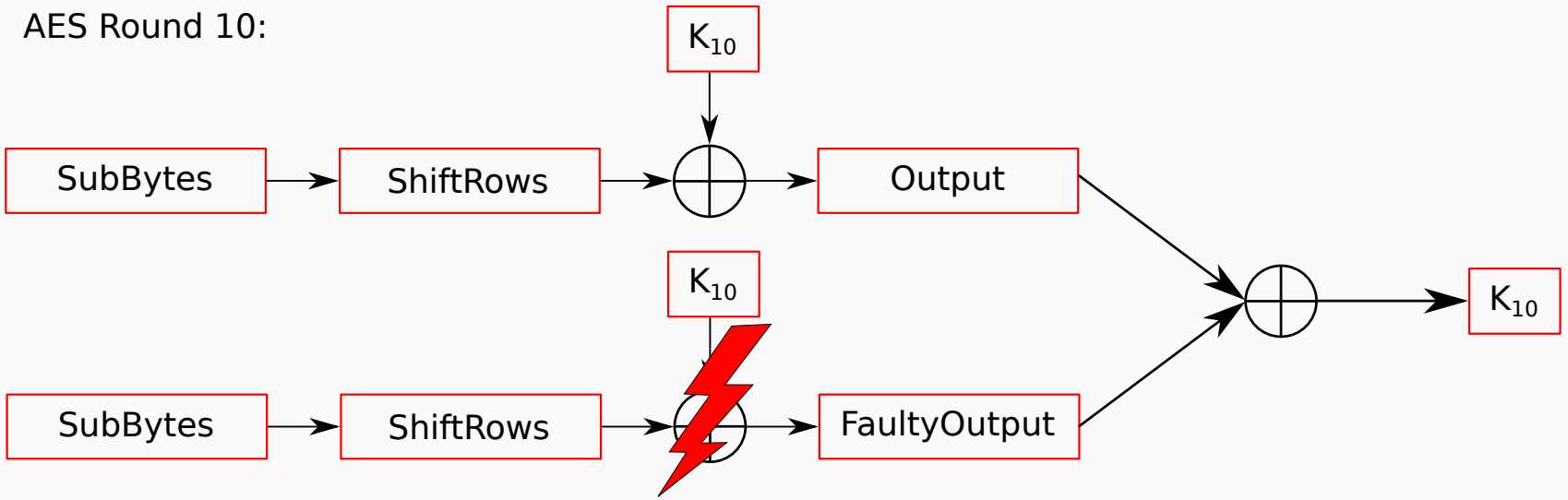
Our Attack Idea

- The main goal of the attack is to show vulnerability of unprotected AES implementation against laser fault injection
- Such attack is powerful - requires only one fault, no need to know the plaintext
- Our experiments show high repeatability
- Instruction skip is easy to perform - laser equipment does not have to be very precise and a chip surface can be unpolished

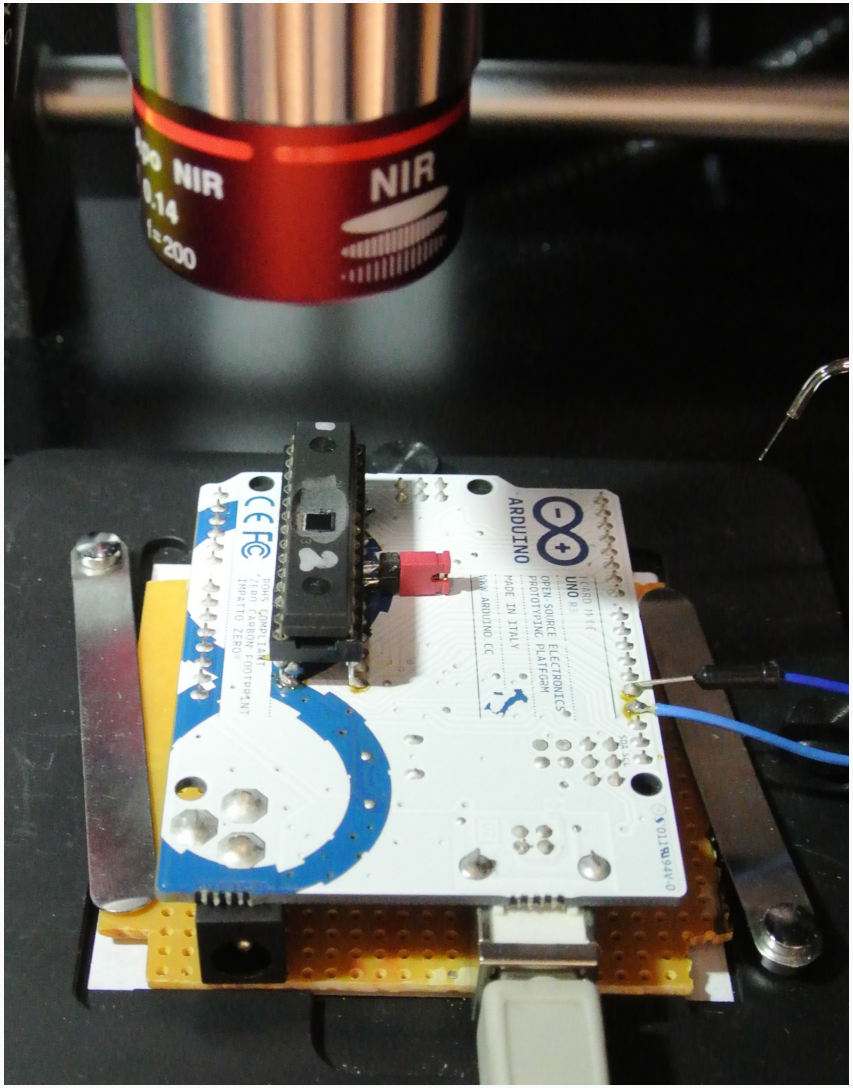
Practical Fault Attack on AES

Attack steps:

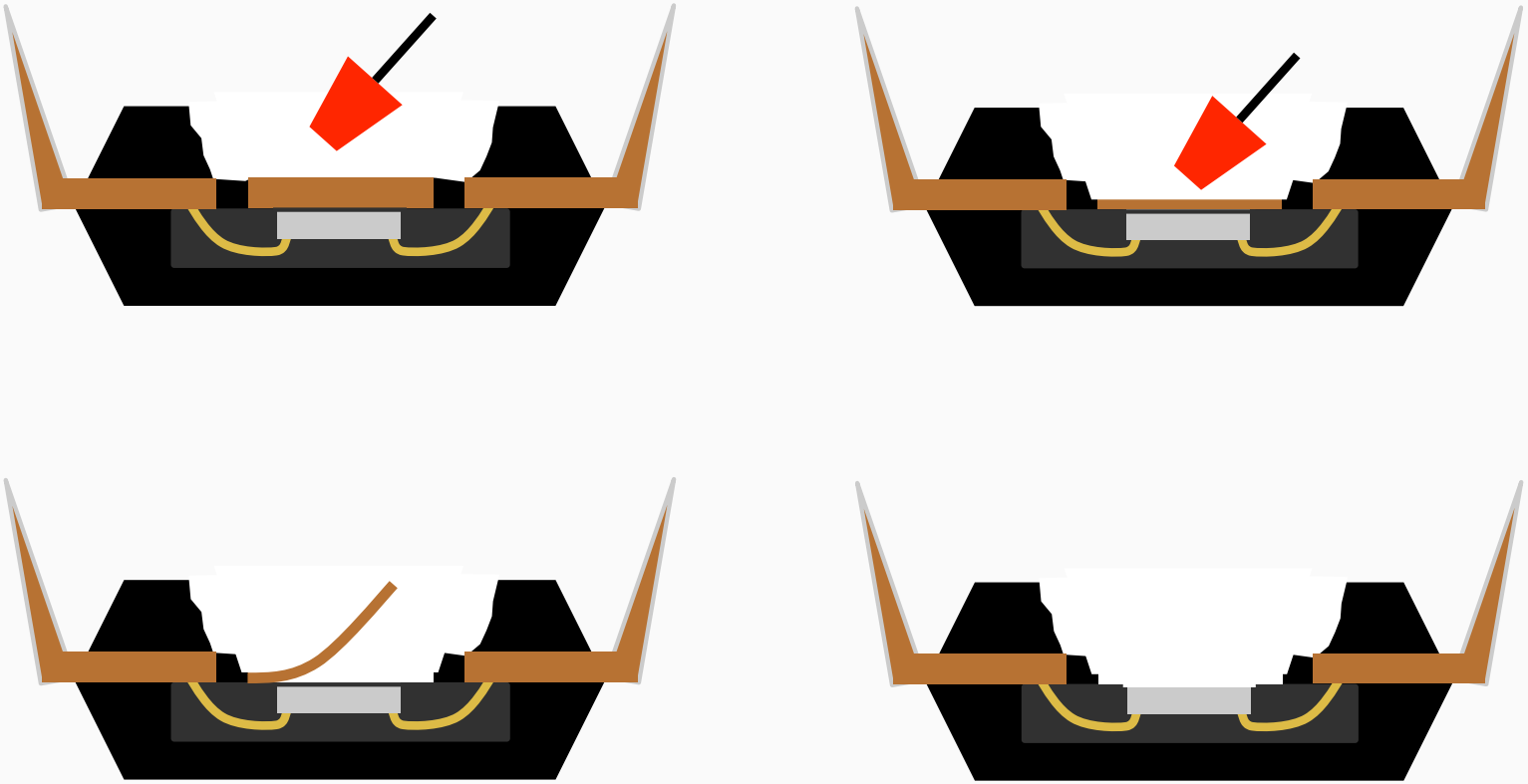
1. Remove the chip package
2. Find a correct position on the chip
3. Determine a correct timing of the last *AddRoundKey*
4. Inject a fault causing instruction skip
5. Compare faulty and correct ciphertext and get K_{10}
6. Get the secret key by inversing a key schedule



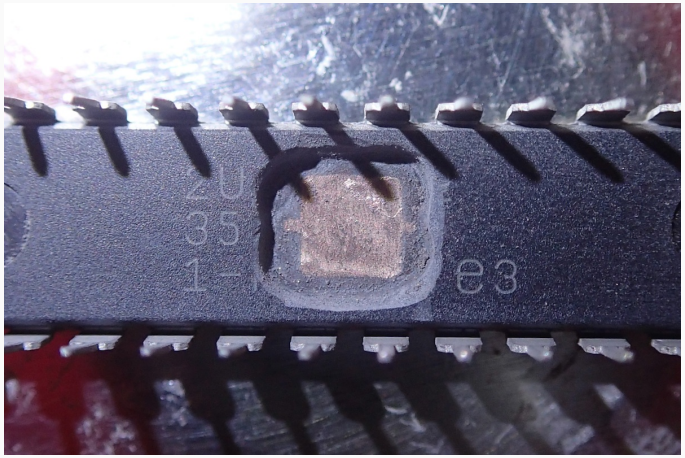
DUT - Arduino Board

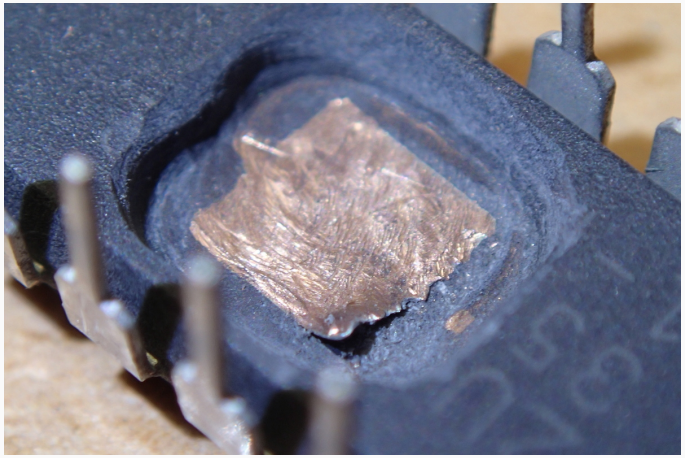


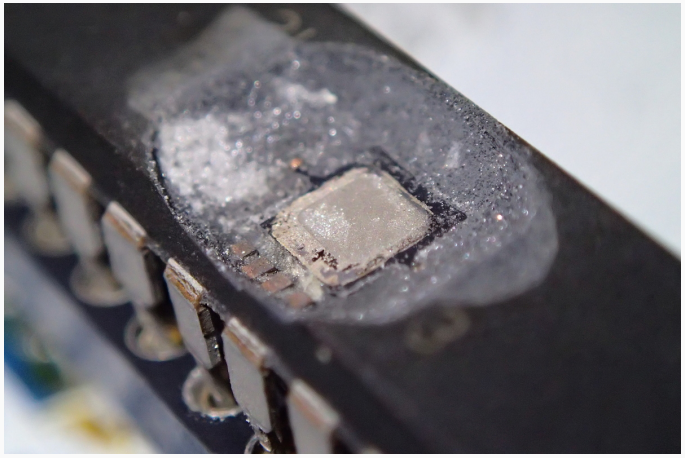
Chip Decapsulation From the Backside 1/2

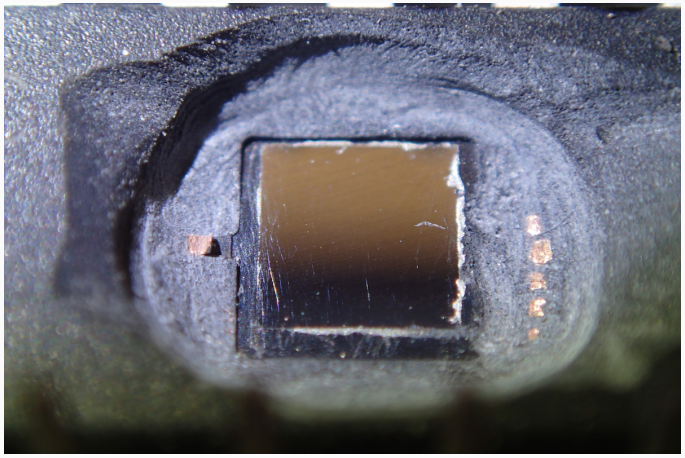


Chip Decapsulation From the Backside 2/2

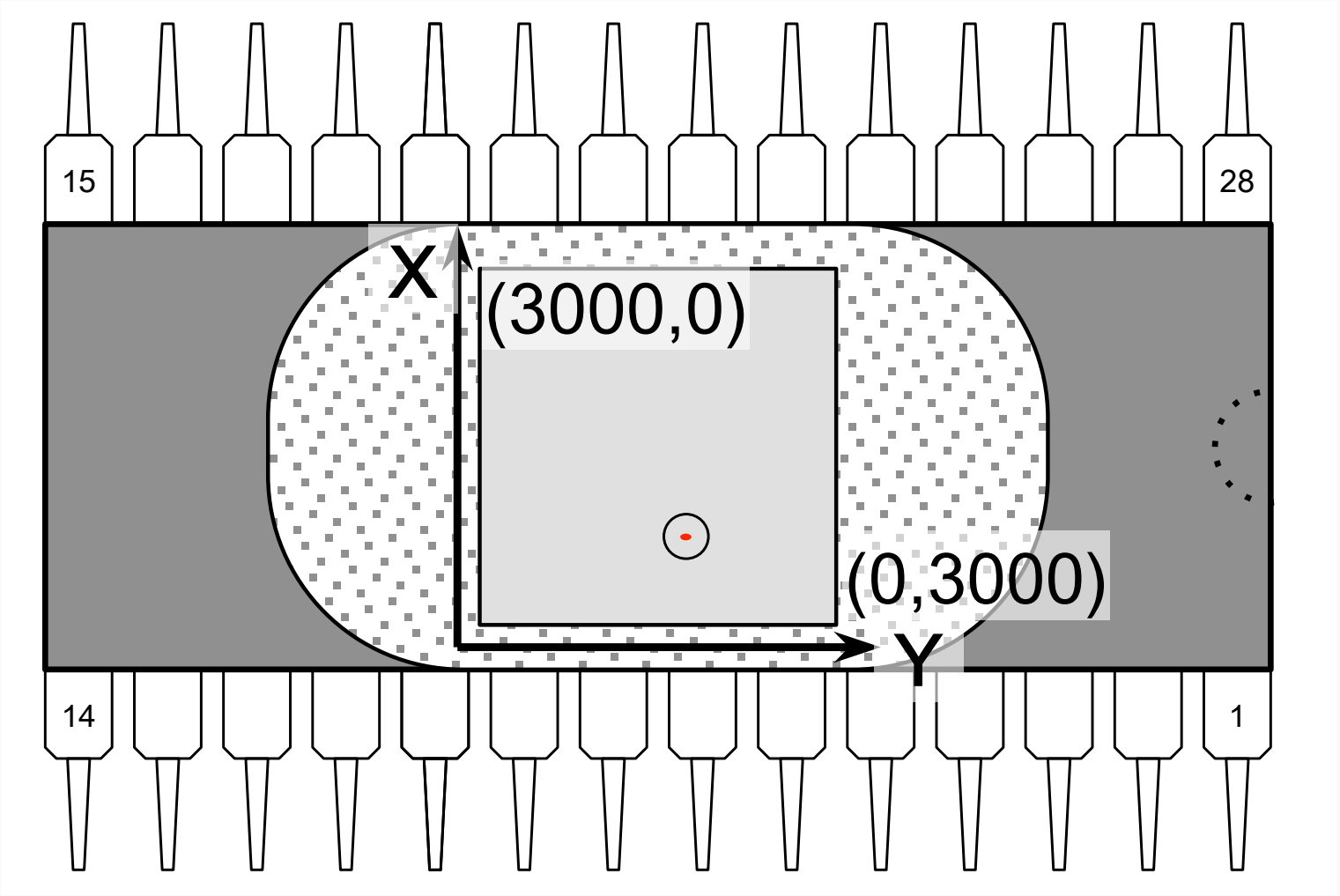




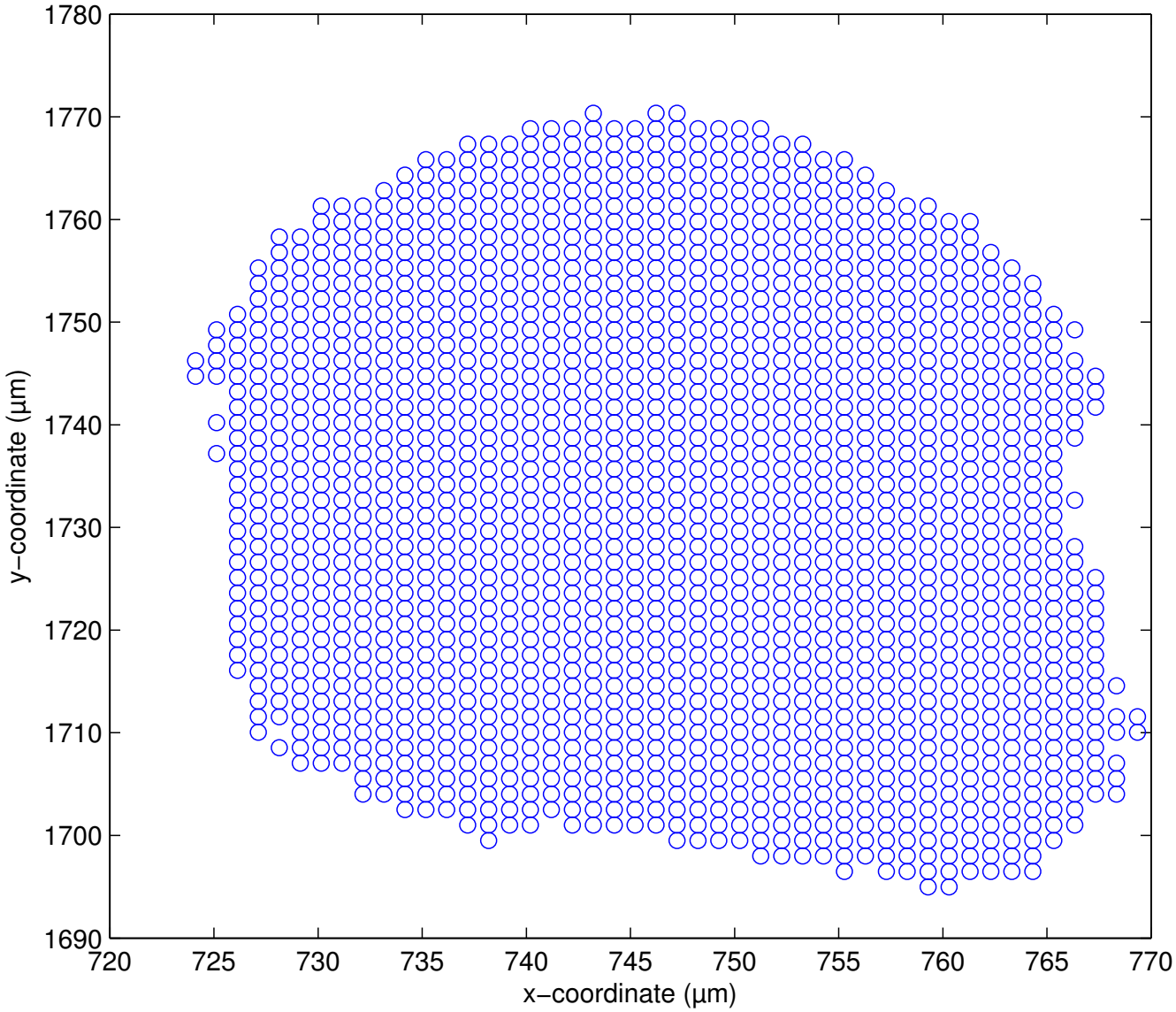




Finding the Correct Spot - Area Size



Finding the Correct Spot - Zoomed



Profiling Phase 1/2

Riscure laser fault injection station was set up to following parameters:

- Glitch length – 150 ns.
- Step size – 15 μm (200 steps in each direction, 40.000 experiments in total).
- Laser power – 1.8%.

Profiling phase took approximately 2 hours.

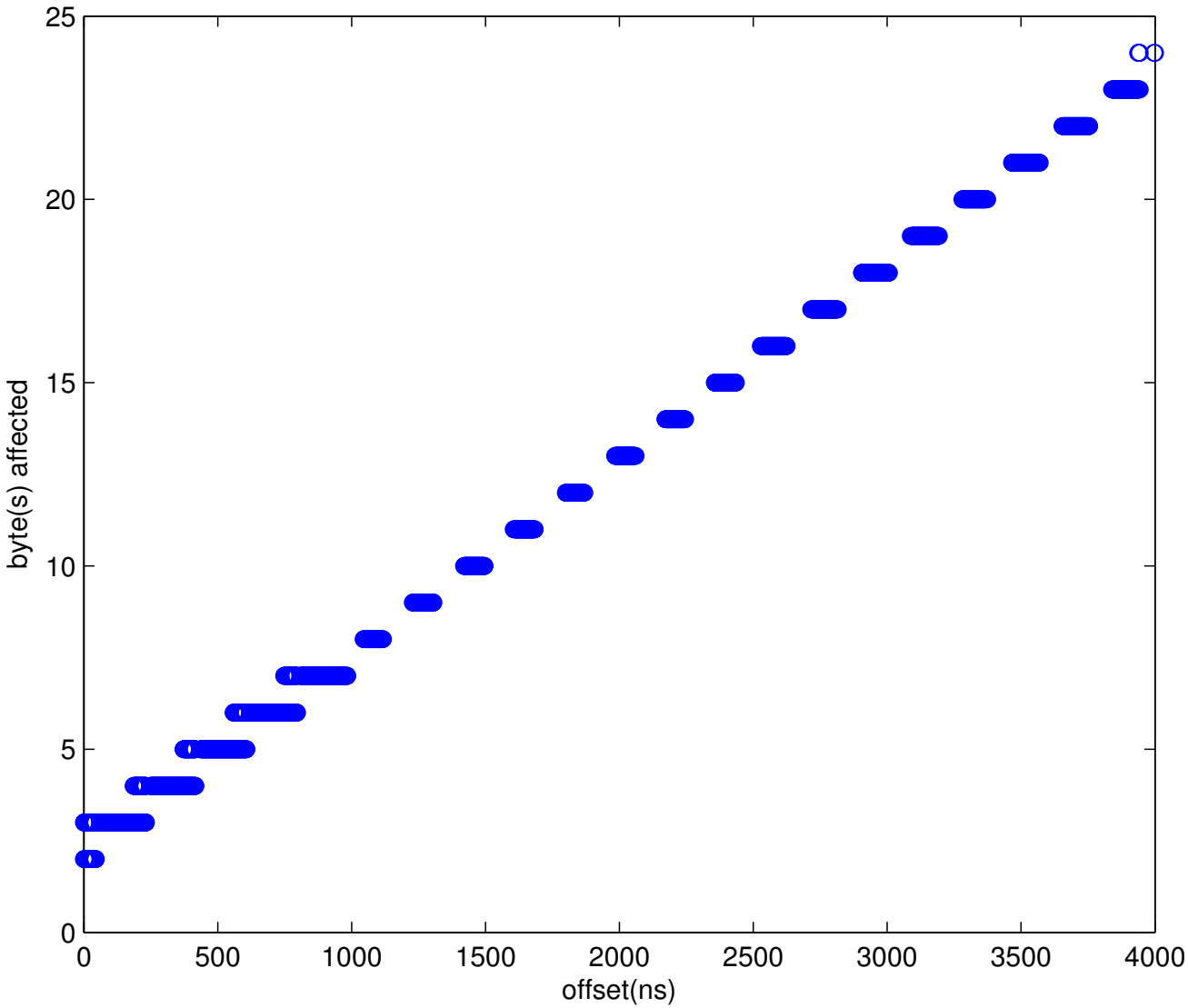
Profiling Phase 2/2

Following code snippet was repeated 25 times in the program with different registers:

```
LD    r0,-Y     (2 clock cycles)
EOR   r0,r25    (1 clock cycle)
ST    Y,r0      (2 clock cycles)
```

- `EOR` instruction was used in order to simulate *AddRoundKey* operation.

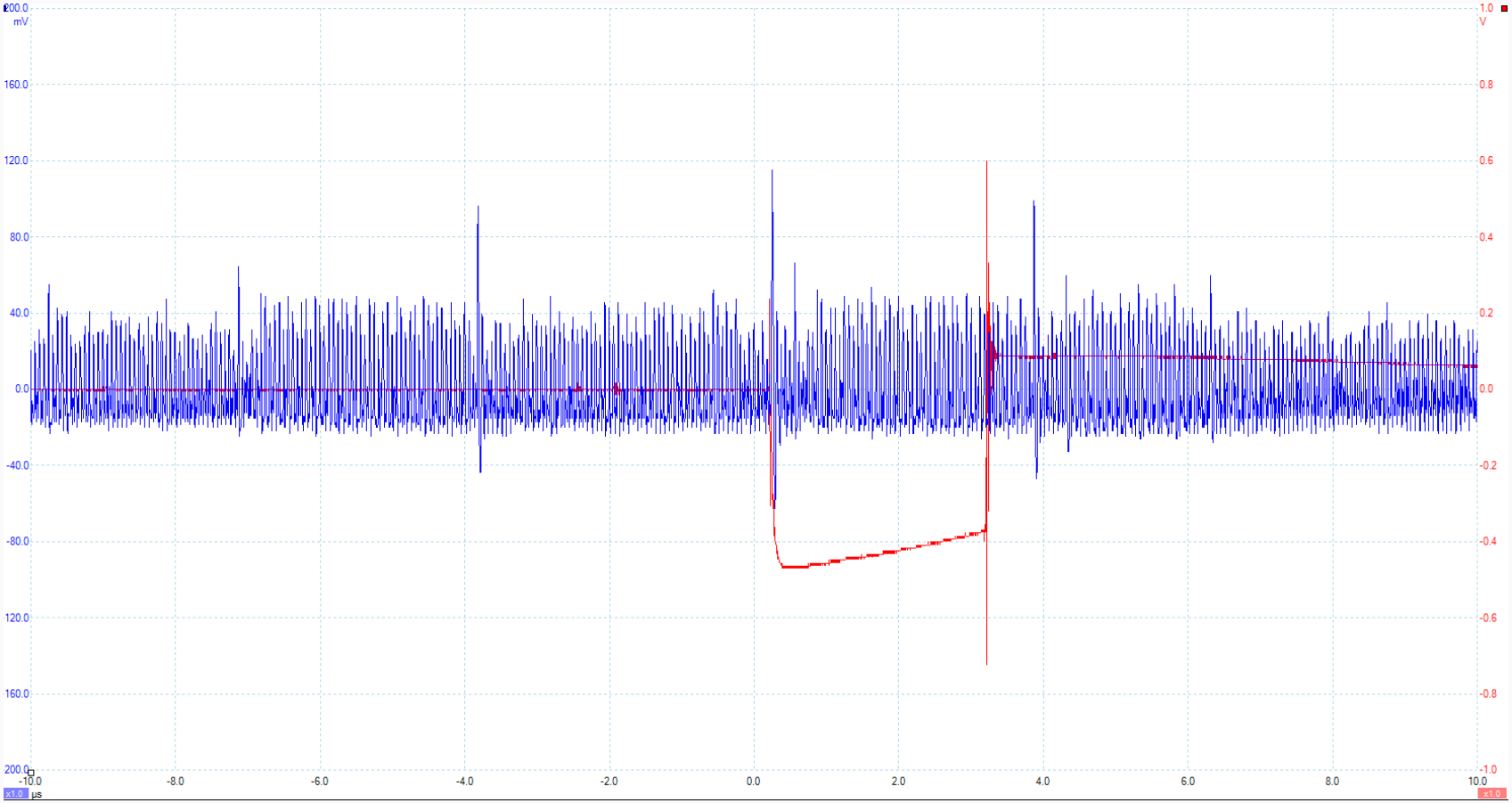
Profiling Phase - Skipping `EOR` Instruction



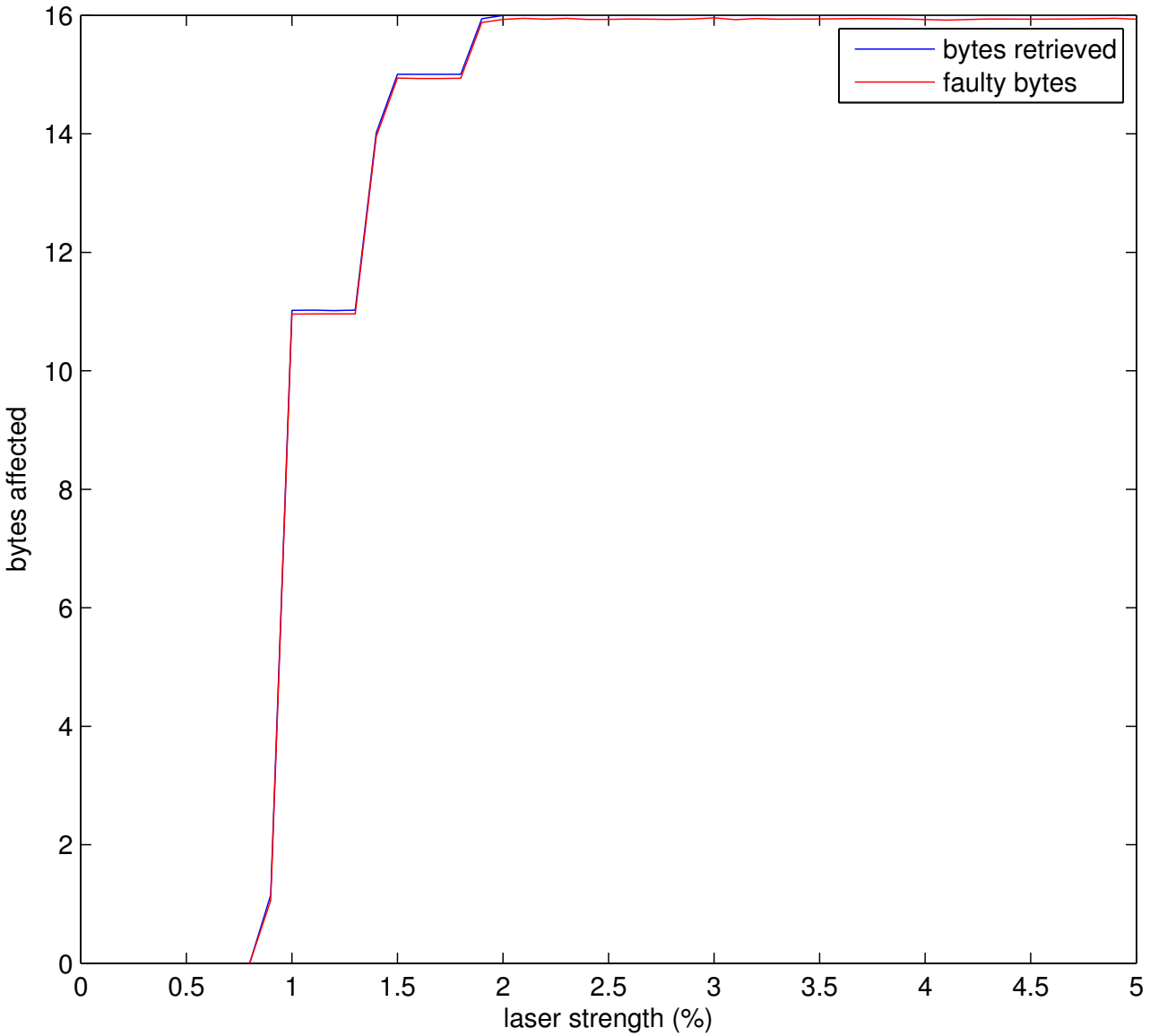
Attack on AES

- After profiling phase we used a software implementation of AES written in assembly language
- Since the `AddRoundKey` lasts 48 clock cycles (16 `load` and 16 `xor` instructions), the laser glitch length in this case was 3 μs
- The area that produces faults in all of 16 bytes is approximately 20x55 μm^2 large ($\sim$0.012% of the whole chip area)

Power Trace and Laser Glitch



Faulty Bytes with Obtained Key Bytes



Attack Results and Discussion

- We were able to perform a simple yet very powerful attack on AES implementation.
- This fault attack requires only one faulty and one correct ciphertext.
- Our experiments show a very high repeatability of such attack.
- It is easy to break implementations with countermeasures which perform encryption, decryption and then compare plaintexts.
- The success rate was 100% when using 2% laser power and 3 μs glitch length, aiming at the correct region on the chip.

Countermeasures

How to Defend the Implementation?

Three main approaches:

- *Fault detection* - error correction/detection codes, sensors, spatial/temporal redundancy, infection
- *Fault prevention* - special packages, sensors, metal layers
- *Analysis prevention* - re-keying, design level protection

IC Package as a Countermeasure

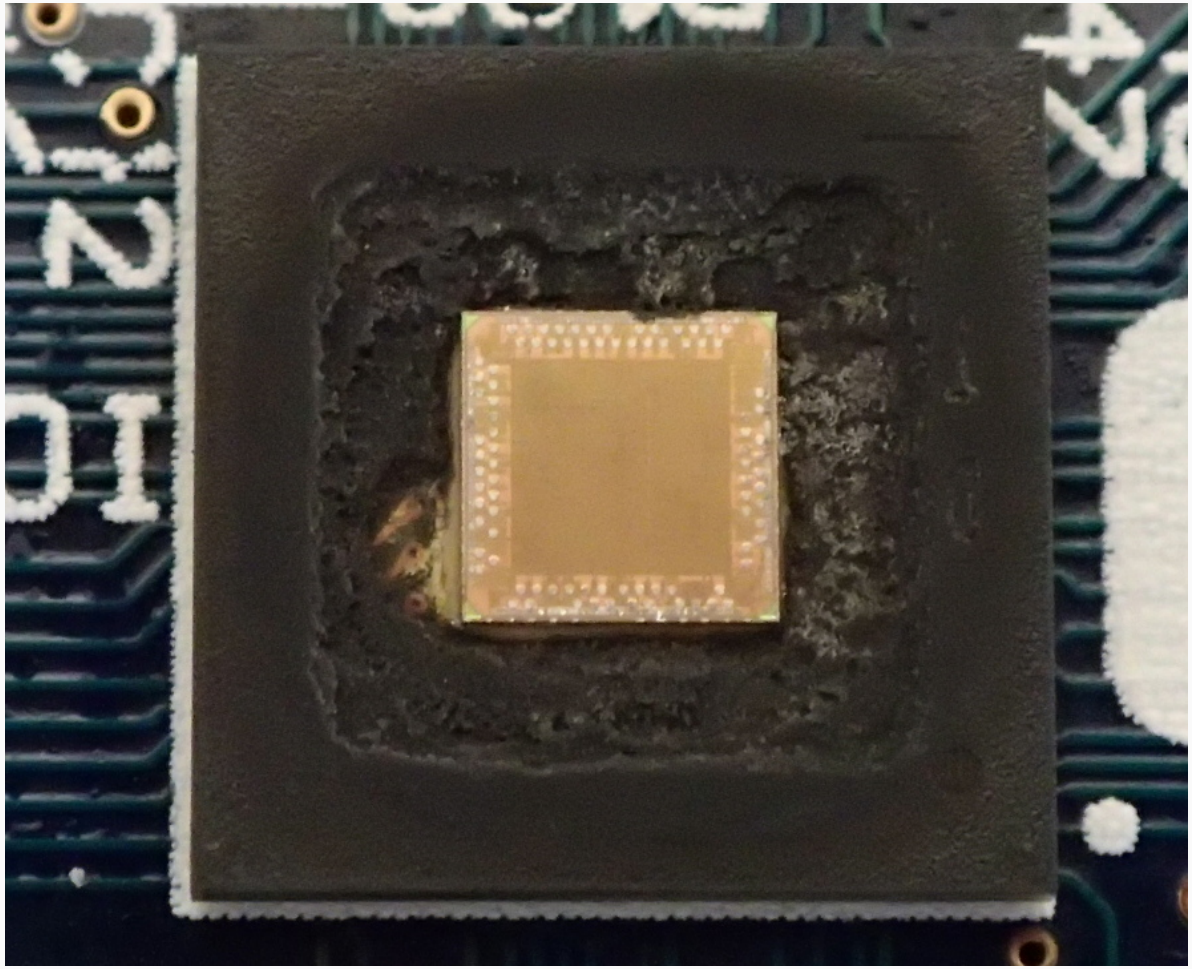


Figure 3: Bonding wires dissolved during the decapsulation process.

Redundancy

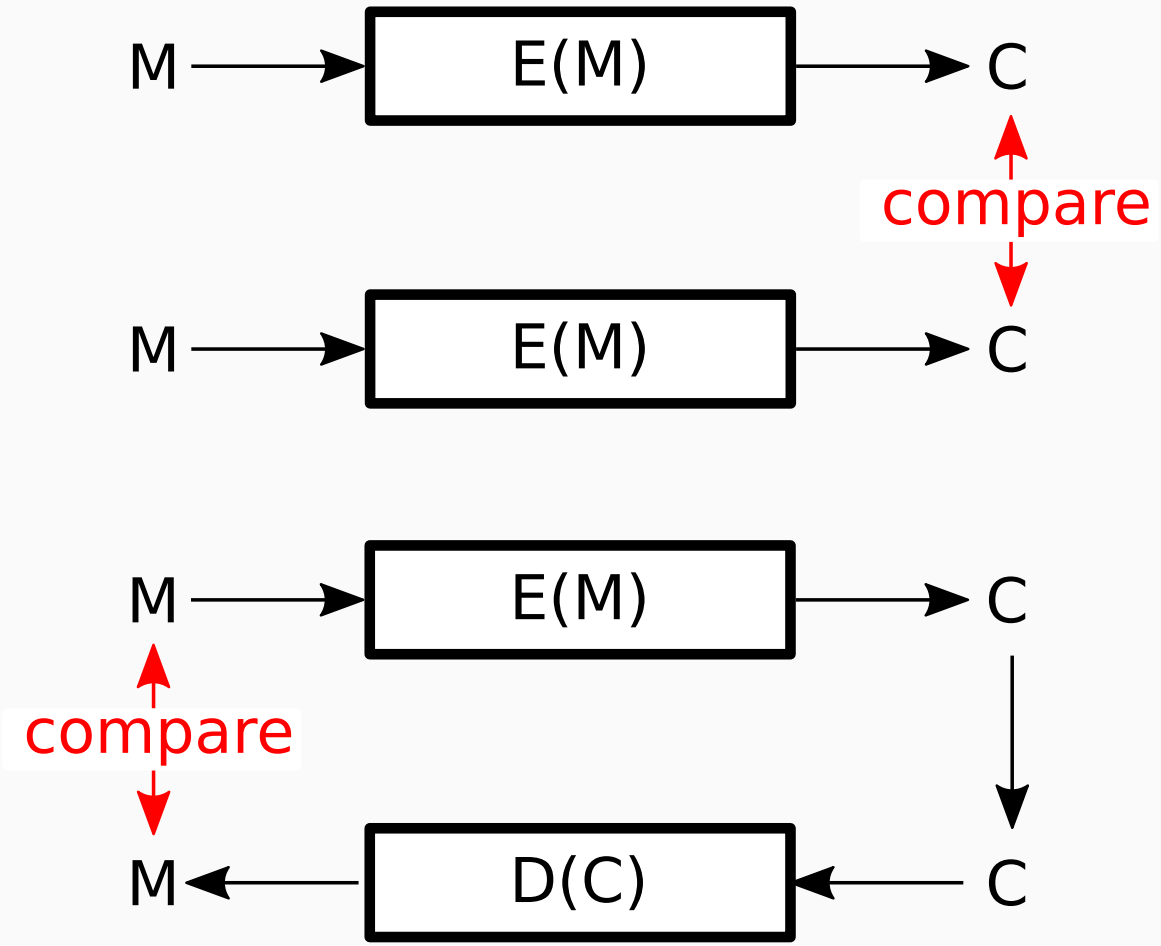


Figure 4: Basic redundancy approaches.

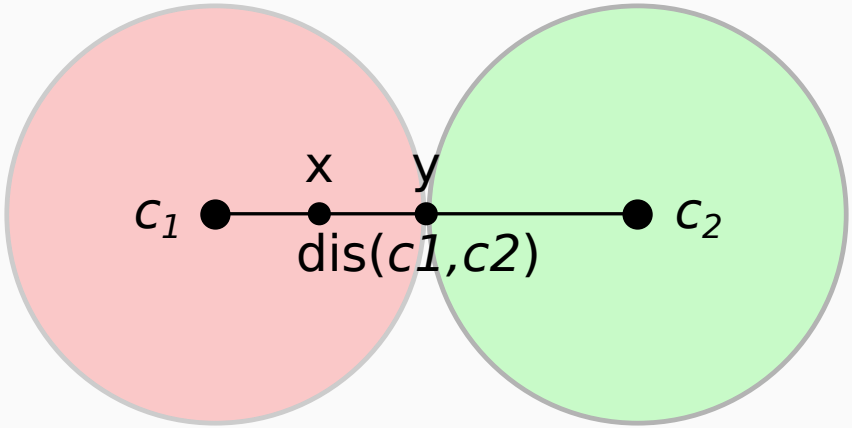
Encoding – Detection and Correction

Definition

For a binary code $\mathcal{C}$ of length n with $\mathrm{dis}\,(\mathcal{C}) = d$, let $M = |\mathcal{C}|$ denote the number of codewords in $\mathcal{C}$. Then $\mathcal{C}$ is called an $(n, M, d)-$binary code.

Definition

For a binary code $\mathcal{C}$ of length n with $\mathrm{dis}\,(\mathcal{C}) = d$, let $M = |\mathcal{C}|$ denote the number of codewords in $\mathcal{C}$. Then $\mathcal{C}$ is called an $(n, M, d)-$binary code.



Encoding – Detection Table

Detection table for $\mathcal{C}_{3,2,min2}$, where $0 \rightarrow 001$ and $1 \rightarrow 100$.

xor	000	001	010	011	100	101	110	111
000	000	000	000	000	000	000	000	000
001	000	001	000	000	100	000	000	000
010	000	000	000	000	000	000	000	000
011	000	000	000	000	000	000	000	000
100	000	100	000	000	001	000	000	000
101	000	000	000	000	000	000	000	000
110	000	000	000	000	000	000	000	000
111	000	000	000	000	000	000	000	000

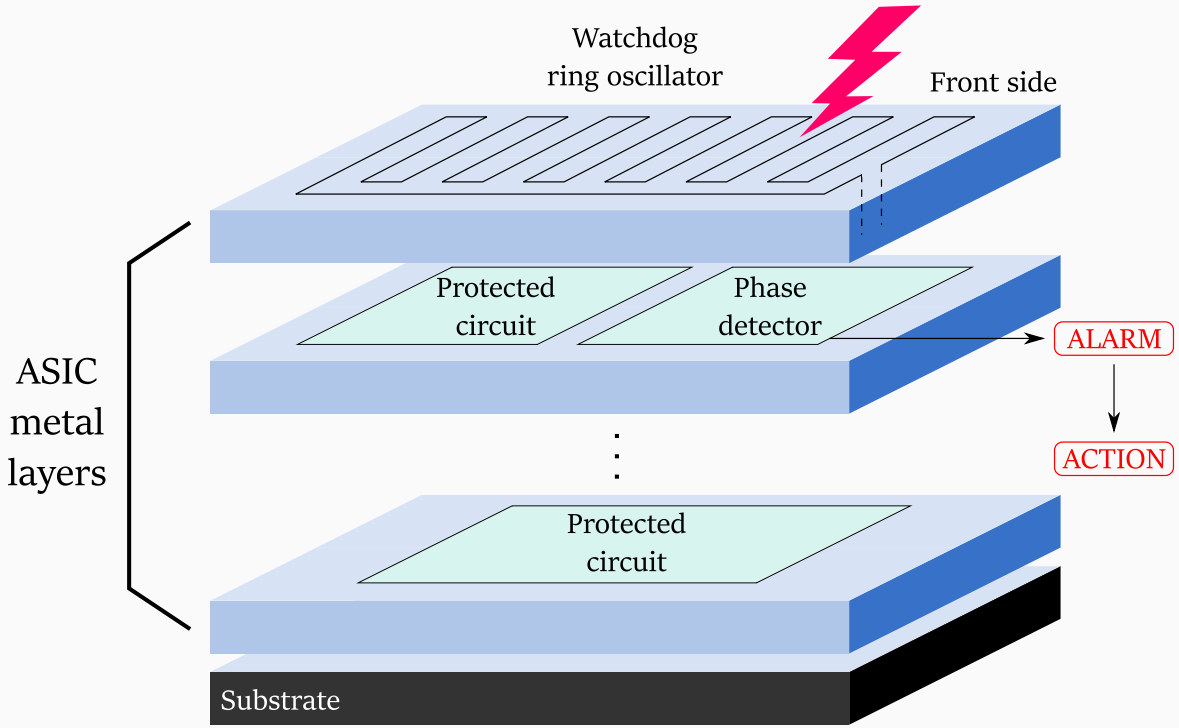
Physical Sensor as a Reactive Countermeasure

- Reactive countermeasures are required for strong attacker models, which assume breaking of a cryptosystem with one fault injection.
- Physical sensors were shown to be effective against these. The ones proposed so far consist of two components:
  - Watchdog Ring Oscillator (WRO)
  - Phase Detection (PD) circuit
- High energy impacts WRO and the resulting phase change is detected by PD to raise an alarm.
- **PACE Sensor 1** uses a phase locked loop as PD[14].
- **PACE Sensor 2** uses an all-digital PD[15].

[14] W. He, J. Breier, S. Bhasin, N. Miura, and M. Nagata. Ring oscillator under laser: Potential of PLL-based countermeasure against laser fault injection, FDTC 2016.

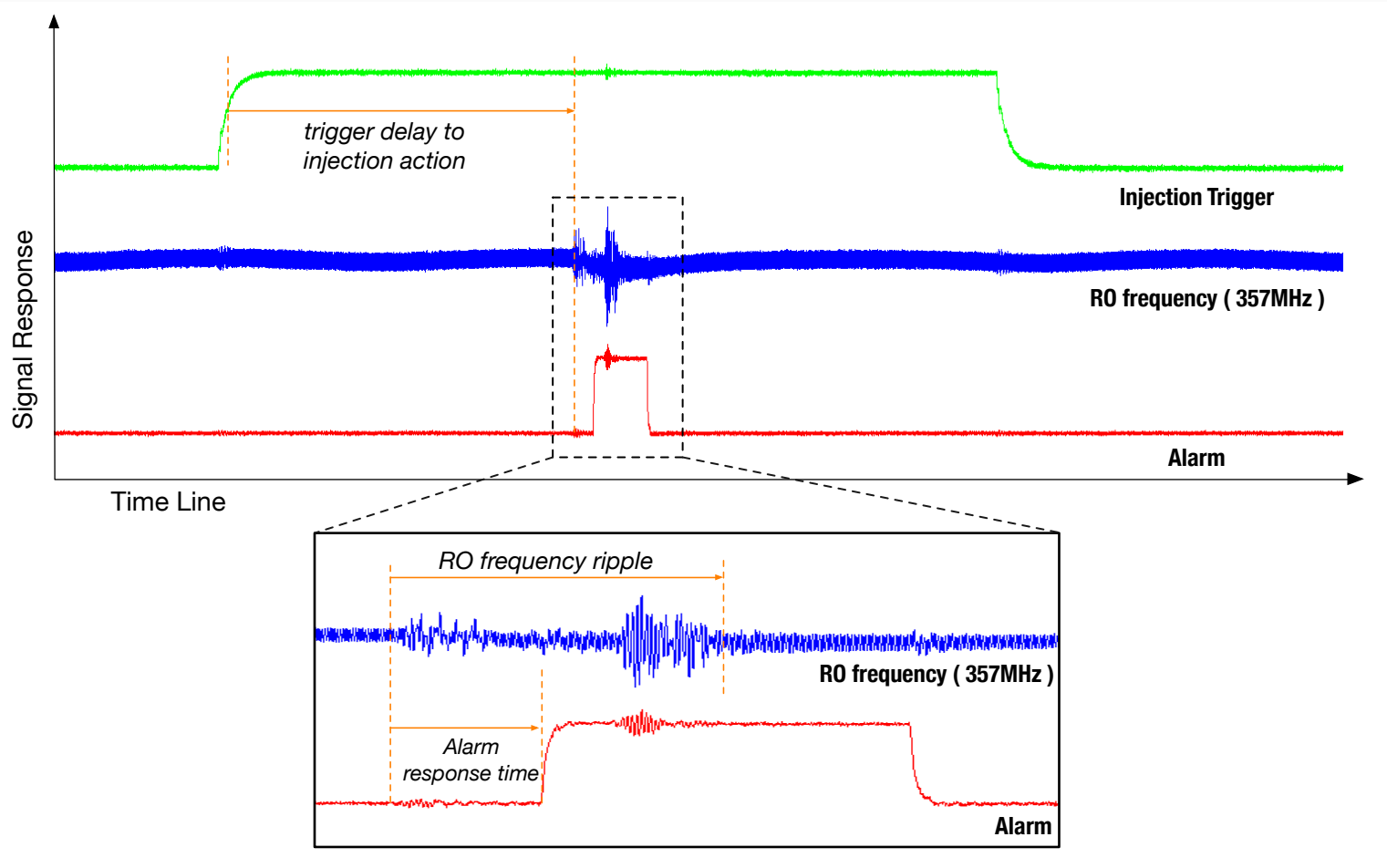
[15] W. He, J. Breier and S. Bhasin. Cheap and Cheerful: A Low-Cost Digital Sensor for Detecting Laser Fault Injection Attacks, SPACE 2016.

Sensor Deployment



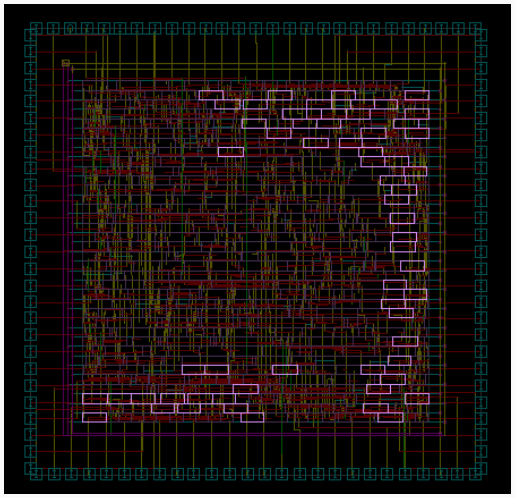
- WRO detecting laser/EM injection can be routed in the top-metal layers.
- Sensitivity of the sensor should be such that it is disturbed at lower laser/EM power than the sensitive circuit.

Detection – PACE Sensor 2

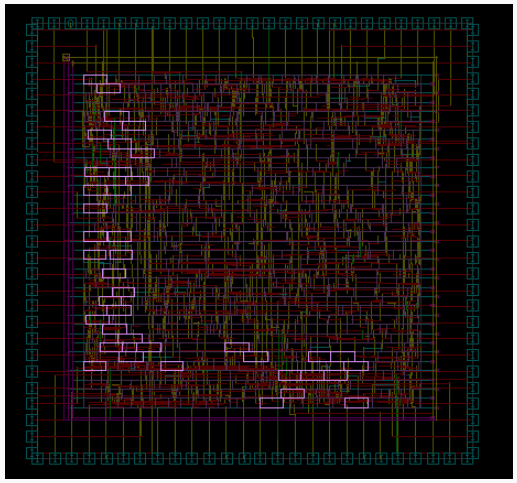


PACE Sensor 1 – Placement Automation

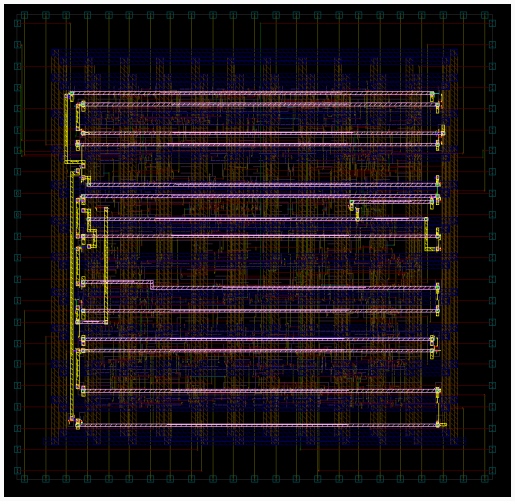
Protecting a stream cipher Plantlet with the detection circuit[16]:



LFSR



NLFSR

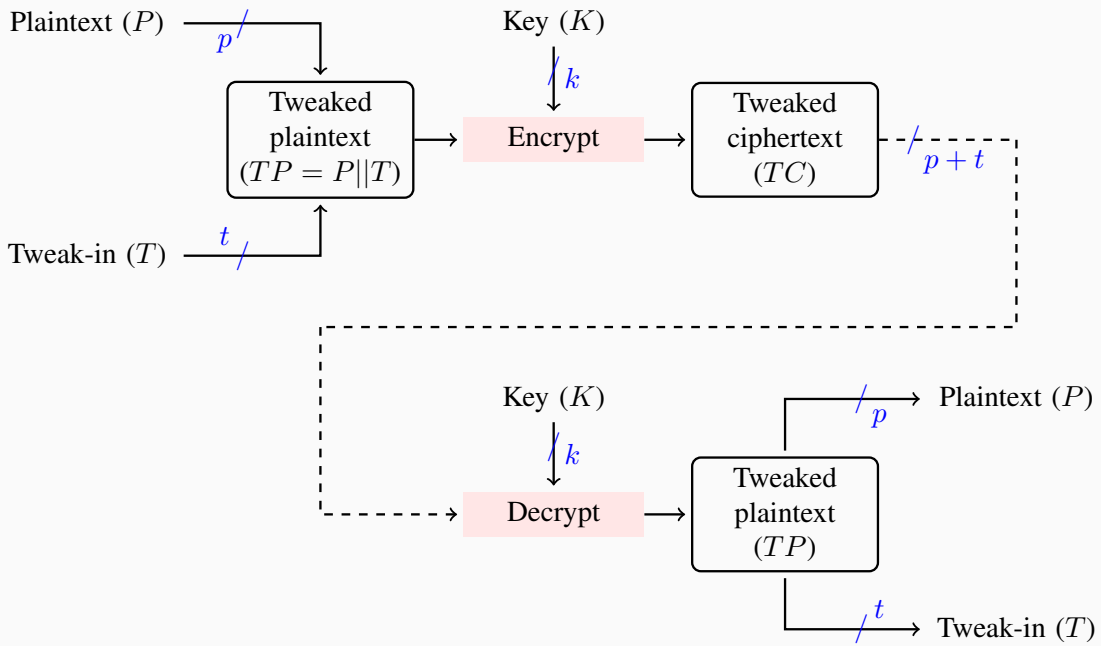


WRO

- Sensitive components are automatically covered by the Watchdog Ring Oscillator circuit on the top layer.

[16] M. Khairallah, R. Sadhukhan, R. Samanta, J. Breier, S. Bhasin, R. S. Chakraborty, A. Chattopadhyay and D. Mukhopadhyay: DFARPA: Differential Fault Attack Resistant Physical Design Automation. DATE 2018.

Protocol Countermeasure Example – Tweak-In[17]



- Tweak-in is a pseudorandom value – larger the size, harder to perform DFA.
- The attacker needs to find collisions in tweak-in to do a successful attack.

[17] A. Baksi, S. Bhasin, J. Breier, M. Khairallah and T. Peyrin: Protecting Block Ciphers against Differential Fault Attacks without Re-keying. HOST 2018.

Conclusion

Conclusion

- Fault attacks are a powerful class of physical attacks
- Powerful equipment, such as LFI or EMFI, is becoming more accessible to attackers
- It is not possible to completely stop the attacker to mount an attack, it can only be made more difficult
- One has to solve the security/cost trade-off before designing a countermeasure

Thank you!

Any questions?

Web: http://jbreier.com

E-mail: jbreier@ntu.edu.sg