Fault Attacks on Cryptographic Devices

Jakub Breier

18 May 2016

Physical Analysis and Cryptographic Engineering
Nanyang Technological University
Singapore

PACE Laboratory

Our team:

- Principal investigator
- 5 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. ATmega328P and Attack on AES
5. Virtex-5 and Bypassing the Parity Protection
6. Countermeasures
7. Conclusion

Physical Attacks on Cryptographic Systems

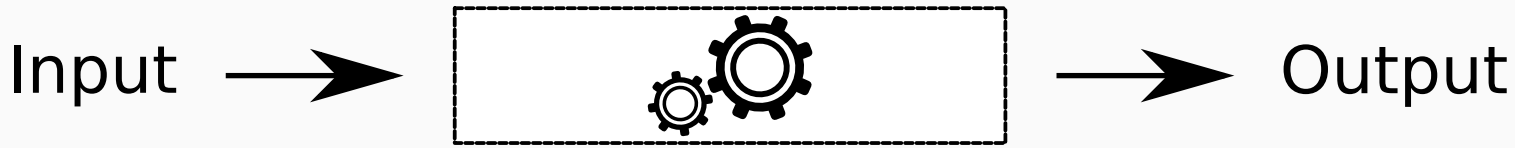
Why Physical Attacks?

- 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

Why Physical Attacks?

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

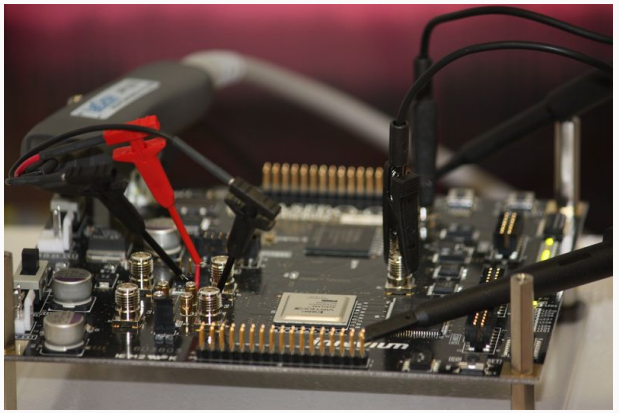
- [2,3]The best fault attack on AES needs just one faulty and correct plaintext/ciphertext pair

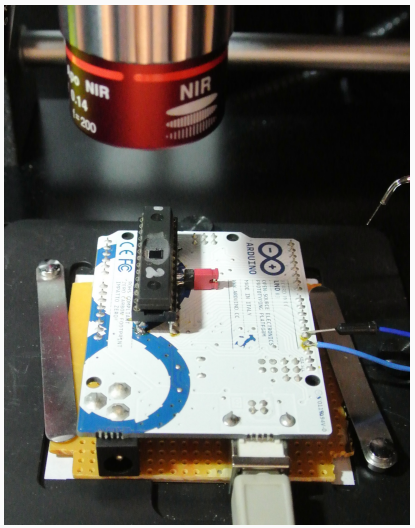


[1]A. Bogdanov et al. Biclique cryptanalysis of the full AES. ASIACRYPT 2011.
[2]D. Saha et al. A diagonal fault attack on the advanced encryption standard. Cryptology ePrint Archive 2009/581.
[3]J. Breier et al. Laser Profiling for the Back-Side Fault Attacks: With a Practical Laser Skip Instruction Attack on AES. CPSS 2015.

Classification

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





Simple Side-Channel Attacks

- Attacker uses information from one measurement to determine parts of the secret key
- Exploit relationship between the executed *operations* and the side-channel information

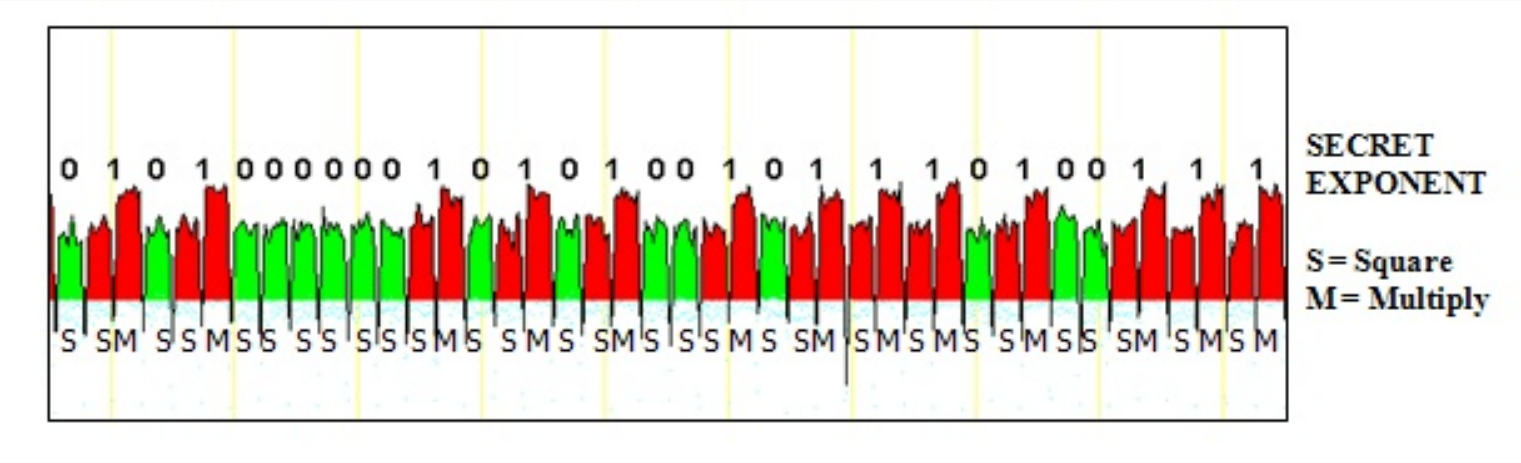


Figure 1: SPA on RSA exponentiation[4].

[4] P. Rohatgi. Protecting FPGAs from power analysis. `http://www.embedded.com/print/4199399`, Cryptography Research, 2010.

Differential Side-Channel Attacks

- Attacker uses multiple measurements to filter out the noise
- Exploit relationship between the processed *data* and the side-channel information

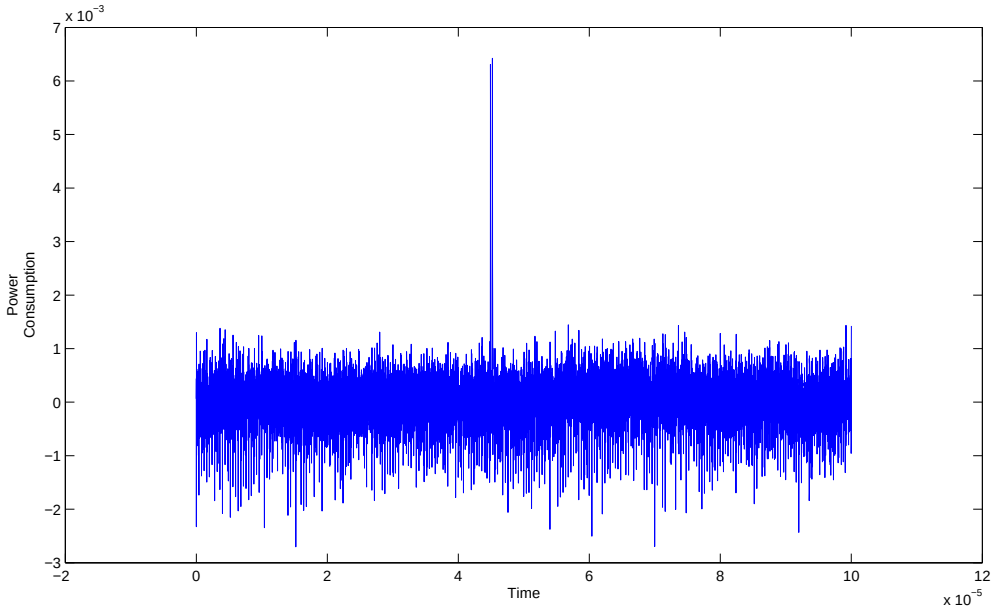


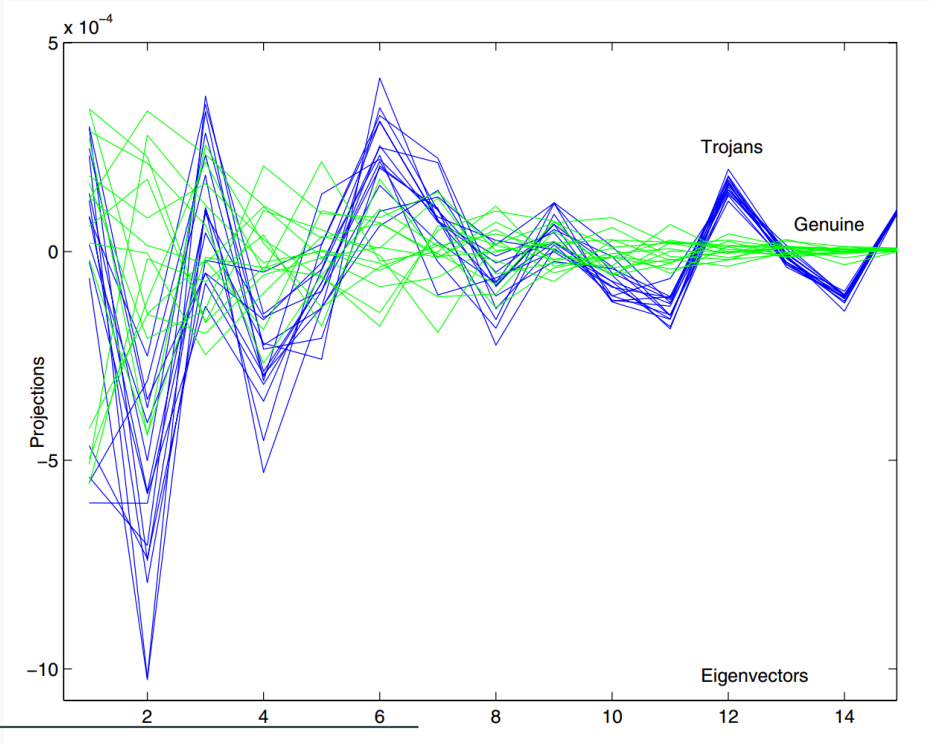
Figure 2: DPA on AES[5].

[5] J. Breier and M. Kleja. On practical results of the differential power analysis. JEEEC, 2012.

SCA to Detect Hardware Trojans

Hardware Trojan:

- Malicious modification of an integrated circuit
- Can leak information, change the behavior, disable or destroy the chip
- SCA can help in detecting hardware trojans[6]



[6] D. Agrawal et al. Trojan Detection using IC Fingerprinting. IEEE S&P 2007.

Fault Attacks

Single Event Upsets

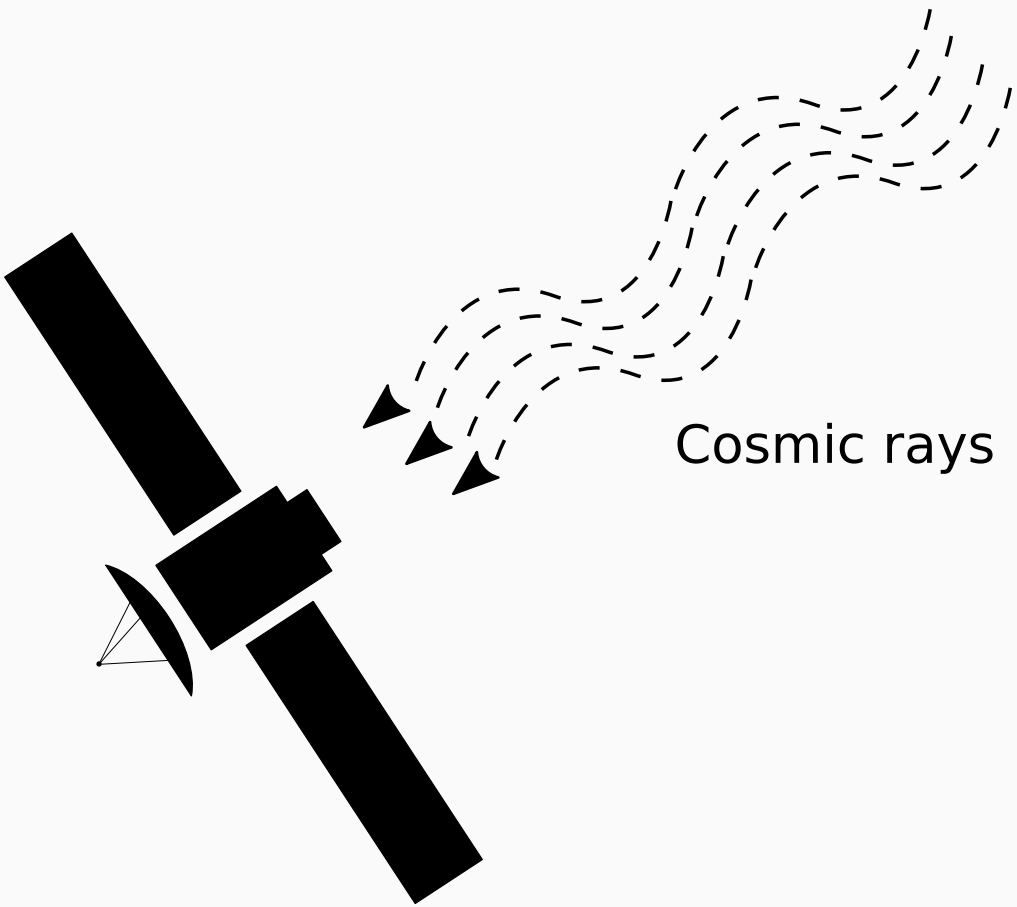


Figure 3: Cosmic rays and satellites[7].

[7] 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[8].
- The first practical attack was implemented by Biham and Shamir, introducing a Differential Fault Analysis on DES[9].

---

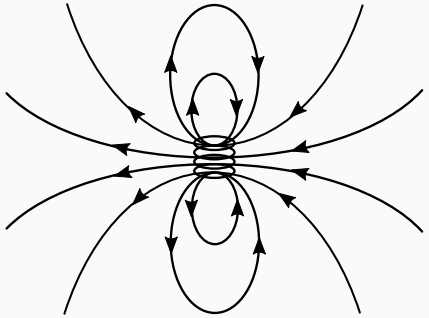
[8] D. Boneh, R. A. DeMillo, and R. J. Lipton. On the Importance of Checking Cryptographic Protocols for Faults, EUROCRYPT97.

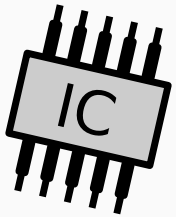
[9] E. Biham and A. Shamir. Differential Fault Analysis of Secret Key Cryptosystems, CRYPTO 97.

Fault Injection Techniques

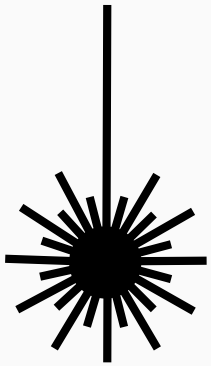
Voltage/clock glitch

EM field

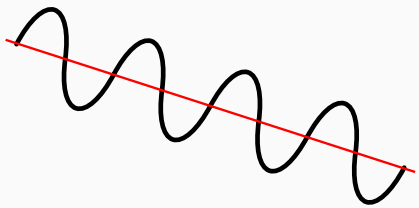




Laser



FIB/X-ray



Fault 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**[10] 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.

[10] 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**[11] 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**[12] 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.

---

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

[12] 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**[13] 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.
- **Linear Fault Analysis**[14] examines linear characteristics for some consecutive rounds of a block cipher. It is a combination of linear cryptanalysis and fault analysis.

---

[13] Y. Li, K. Sakiyama, S. Gomisawa, T. Fukunaga, J. Takahashi, and K. Ohta: Fault Sensitivity Analysis, CHES 2010.

[14] C. H. Kim: Improved Differential Fault Analysis on AES Key Schedule. Information Forensics and Security, IEEE Transactions on, 7(1), 2012.

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.

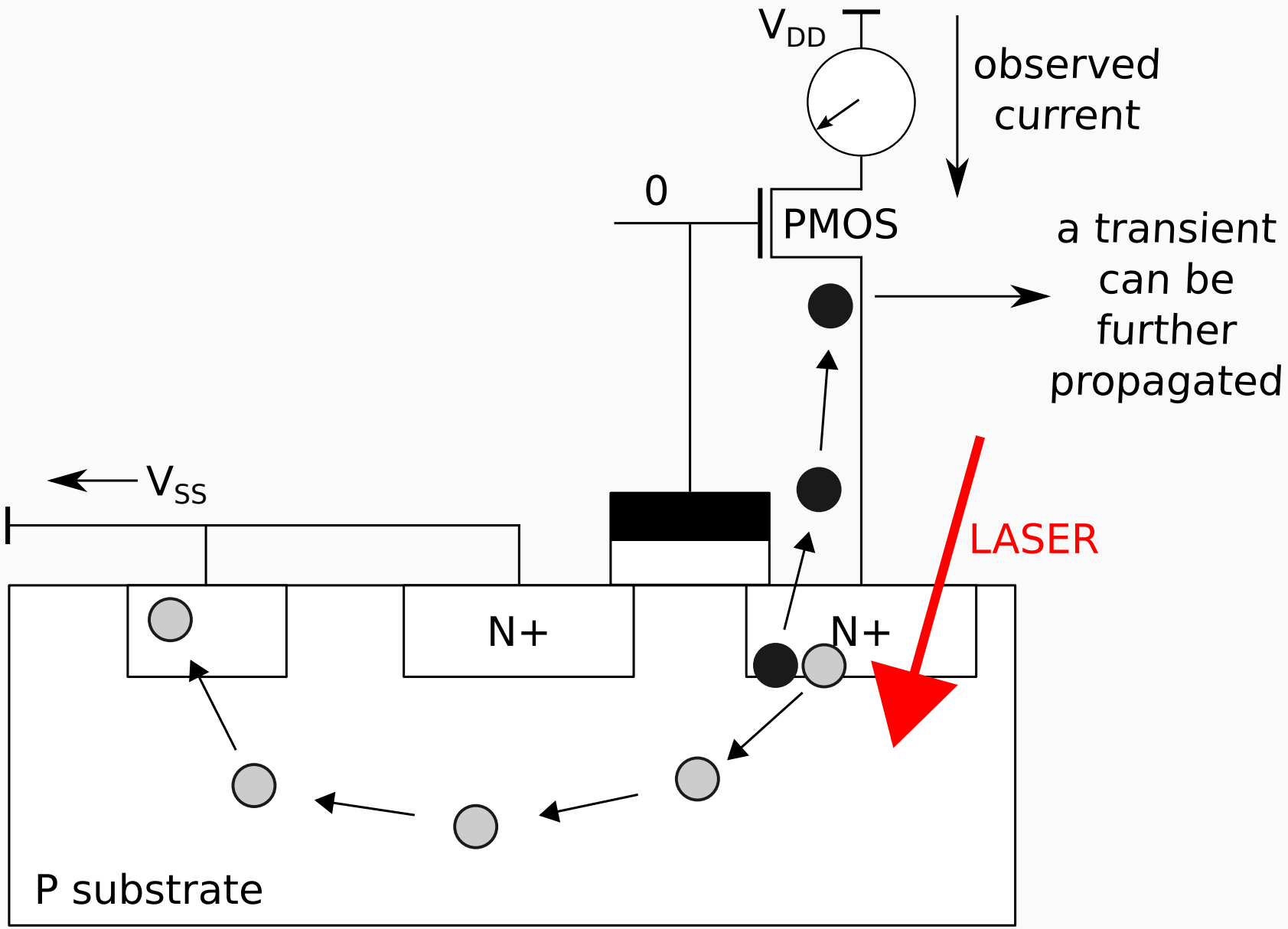
Laser Fault Injection Attacks - Theory 1/2

- Photoelectric effect - when a laser beam with a wavelength corresponding to an energy level higher than the silicon bandgap passes through silicon, it creates electrons-hole pairs along its path.
- If the laser beam passes through the reverse-biased PN junction of a transistor, charge carriers can be drifted in opposite directions and a current pulse is created. This current pulse creates a transient voltage pulse which propagates through the combinatorial logic of the IC.
- This phenomenon is called a Single Event Transient (SET).

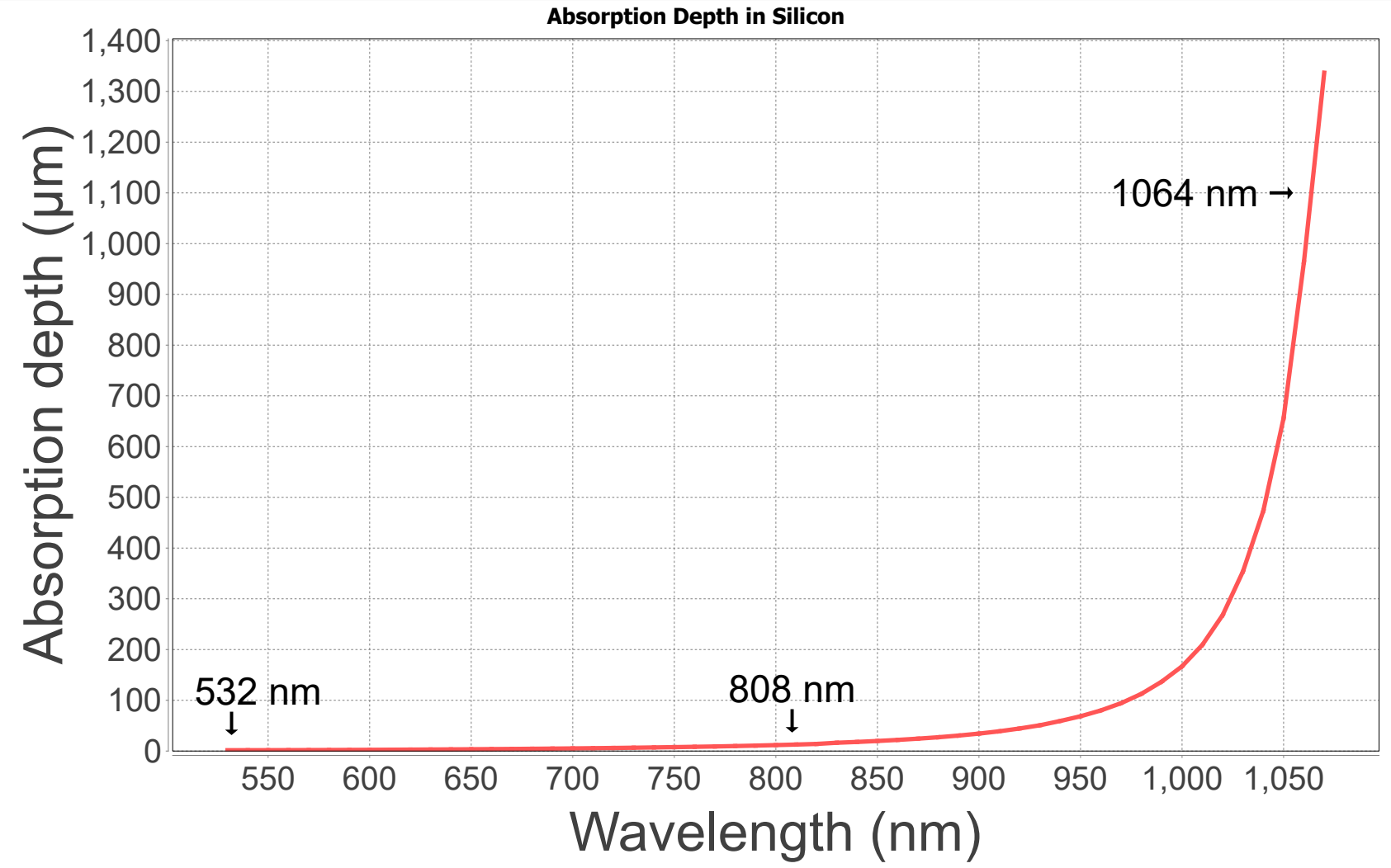
Laser Fault Injection Attacks - Theory 2/2

- Fault is induced if a SET propagates through the logic and is captured by a register.
- Single Event Upset (SEU) occurs when the transient voltage is directly induced into a SRAM or a register: it flips and locks its state to the opposite one.
- By carefully tuning the beam's energy level below a destructive threshold, it is possible to inject faults into a device and it will not suffer any permanent damage.

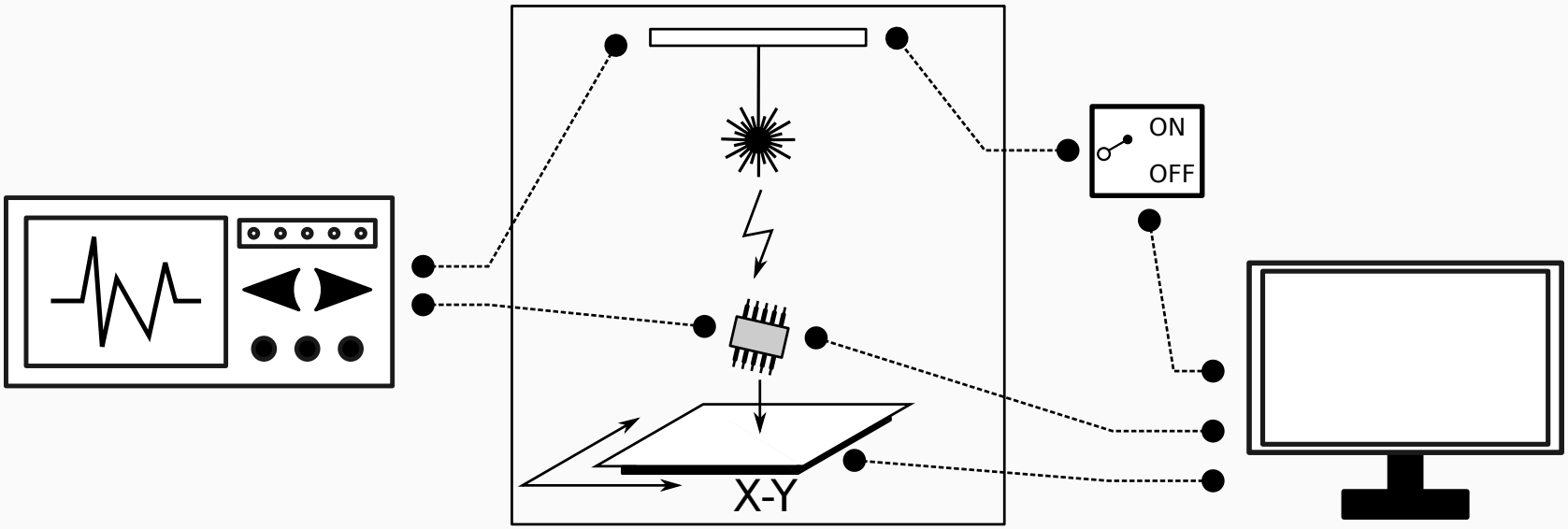
Irradiation Effect on Transistor



Absorption Depth in Silicon



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

Devices under test:

- Atmel ATmega328P (8-bit microcontroller)
- Xilinx Virtex-5 (FPGA)

ATmega328P and Attack on AES

AES-128

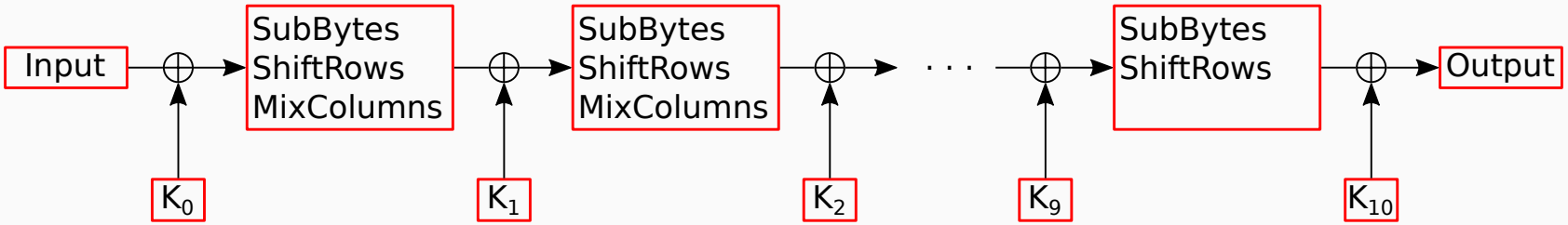


Figure 4: Schematic diagram of AES-128.

- 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[15]
- 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

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

Diagonal Fault Attack[16]

- 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

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

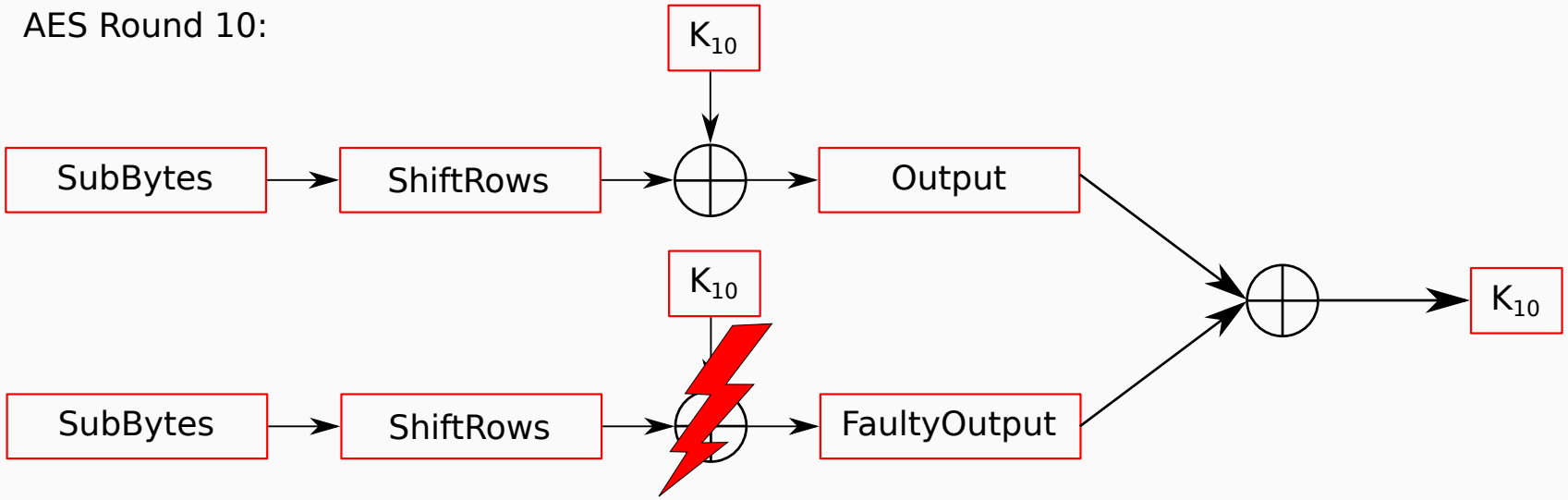
Our Attack Idea

- 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

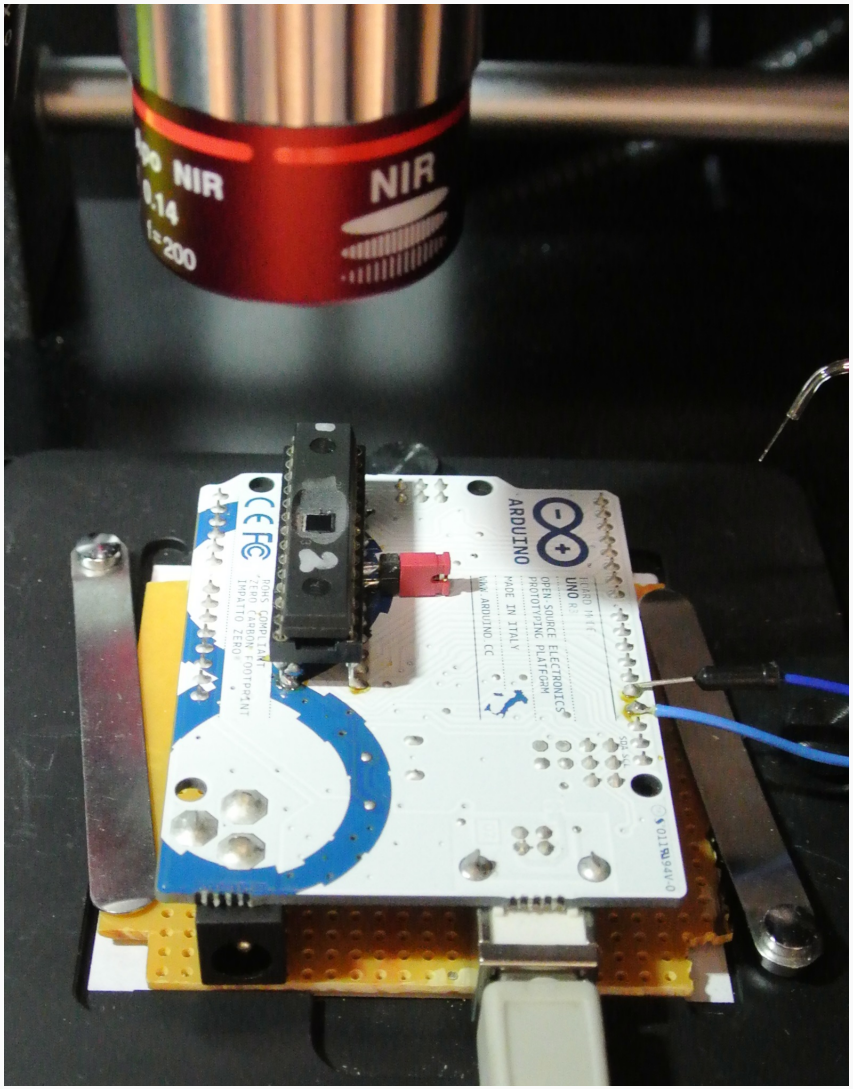
Practical Fault Attack on AES

Attack steps:

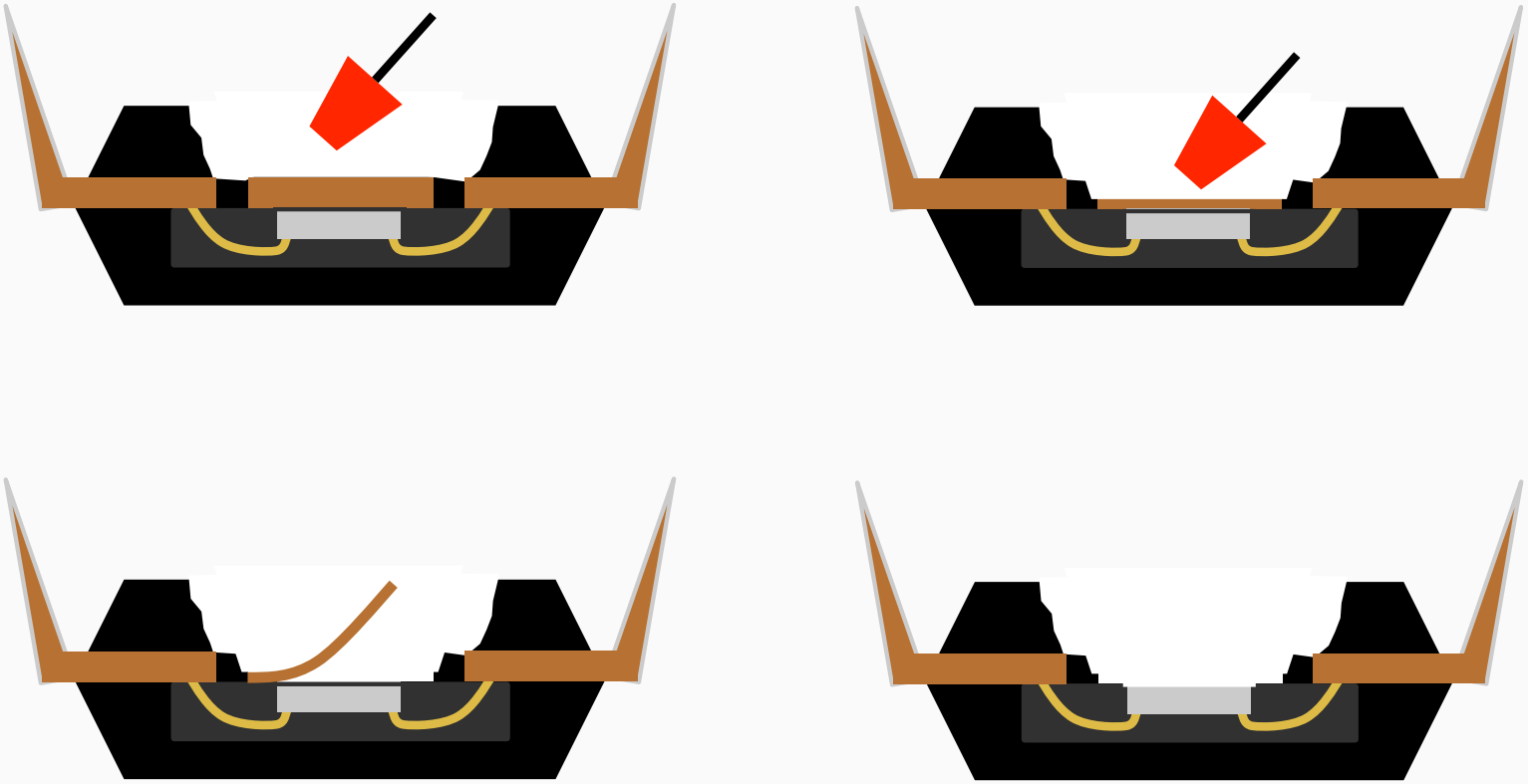
1. Remove the chip package
2. Find a correct position on the chip
3. Determine a correct timing of the last *AddRoundKey*
4. Inject a fault causing instruction skip
5. Compare faulty and correct ciphertext and get K_{10}
6. Get the secret key by inversing a key schedule



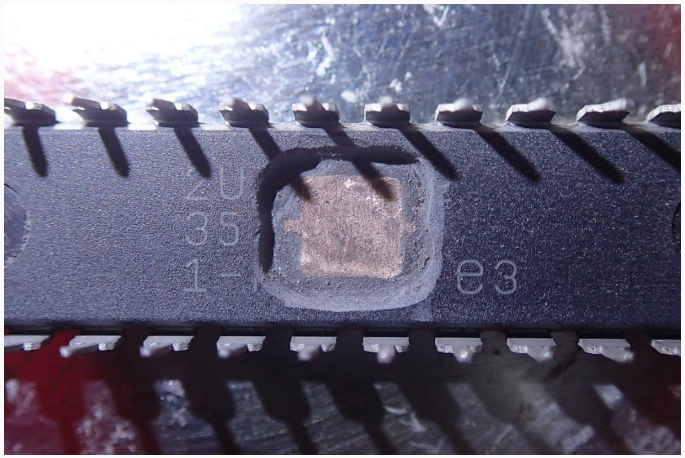
DUT - Arduino Board

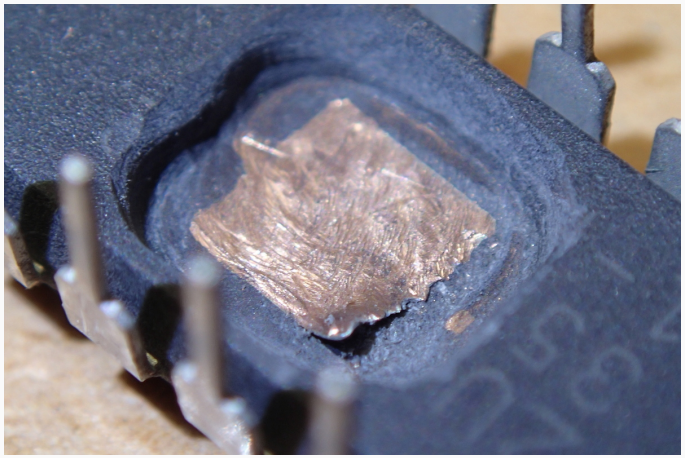


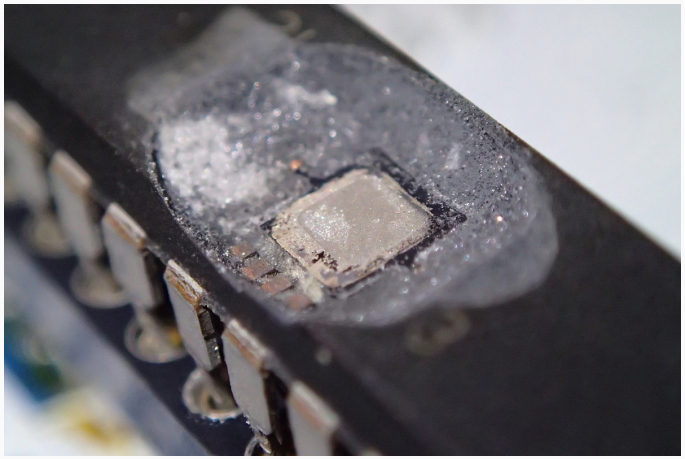
Chip Decapsulation From the Backside 1/2

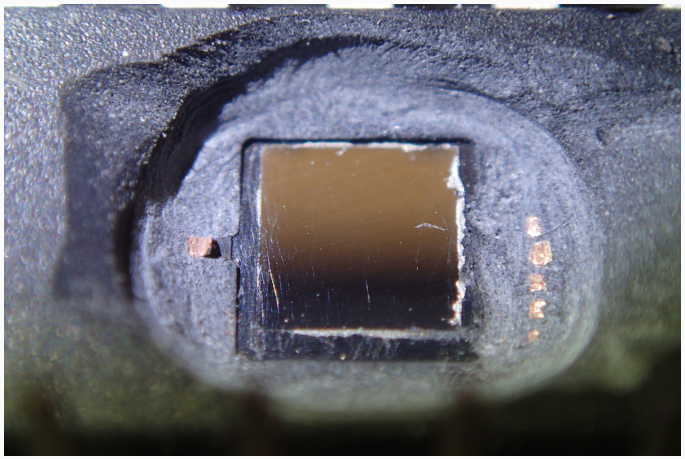


Chip Decapsulation From the Backside 2/2

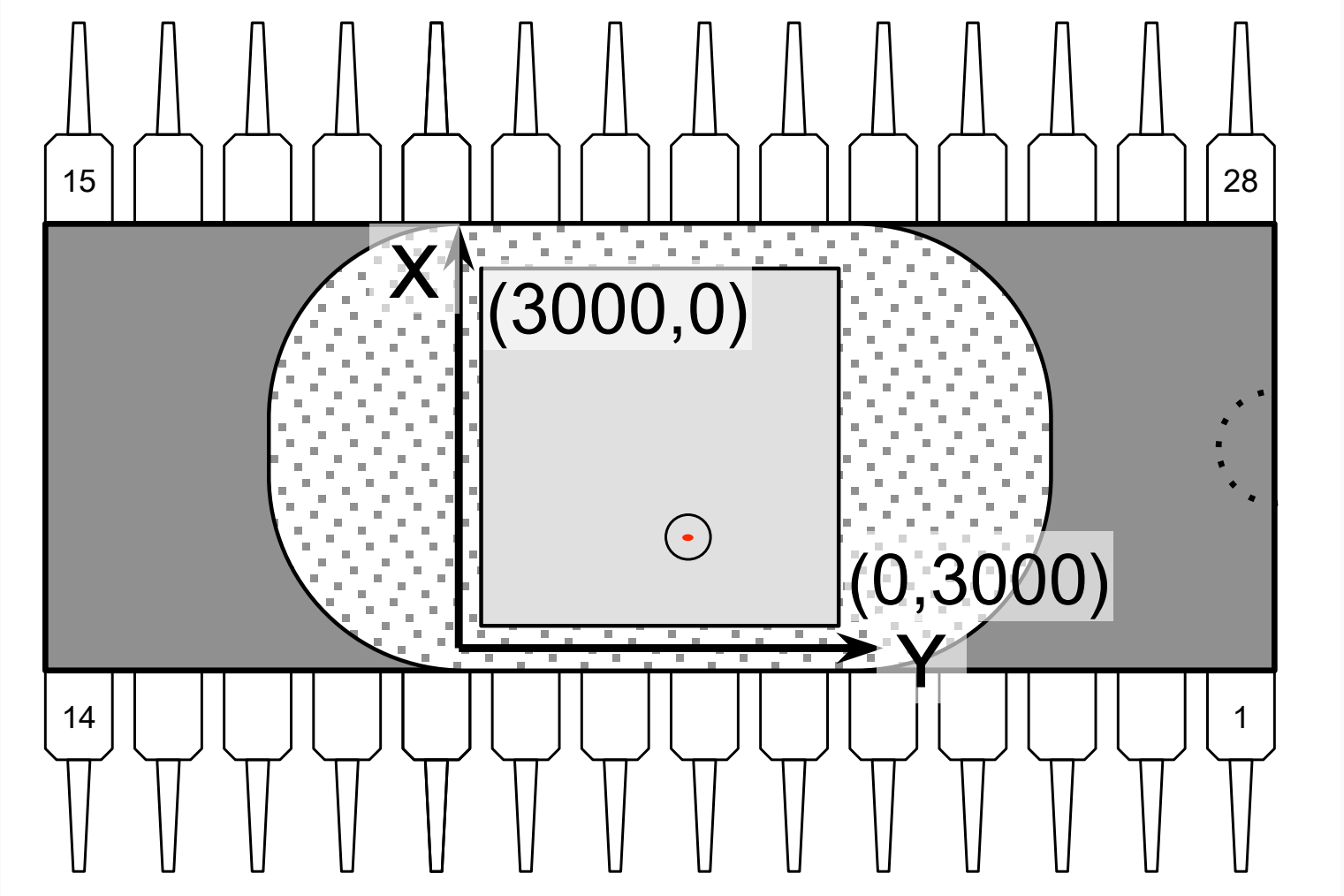




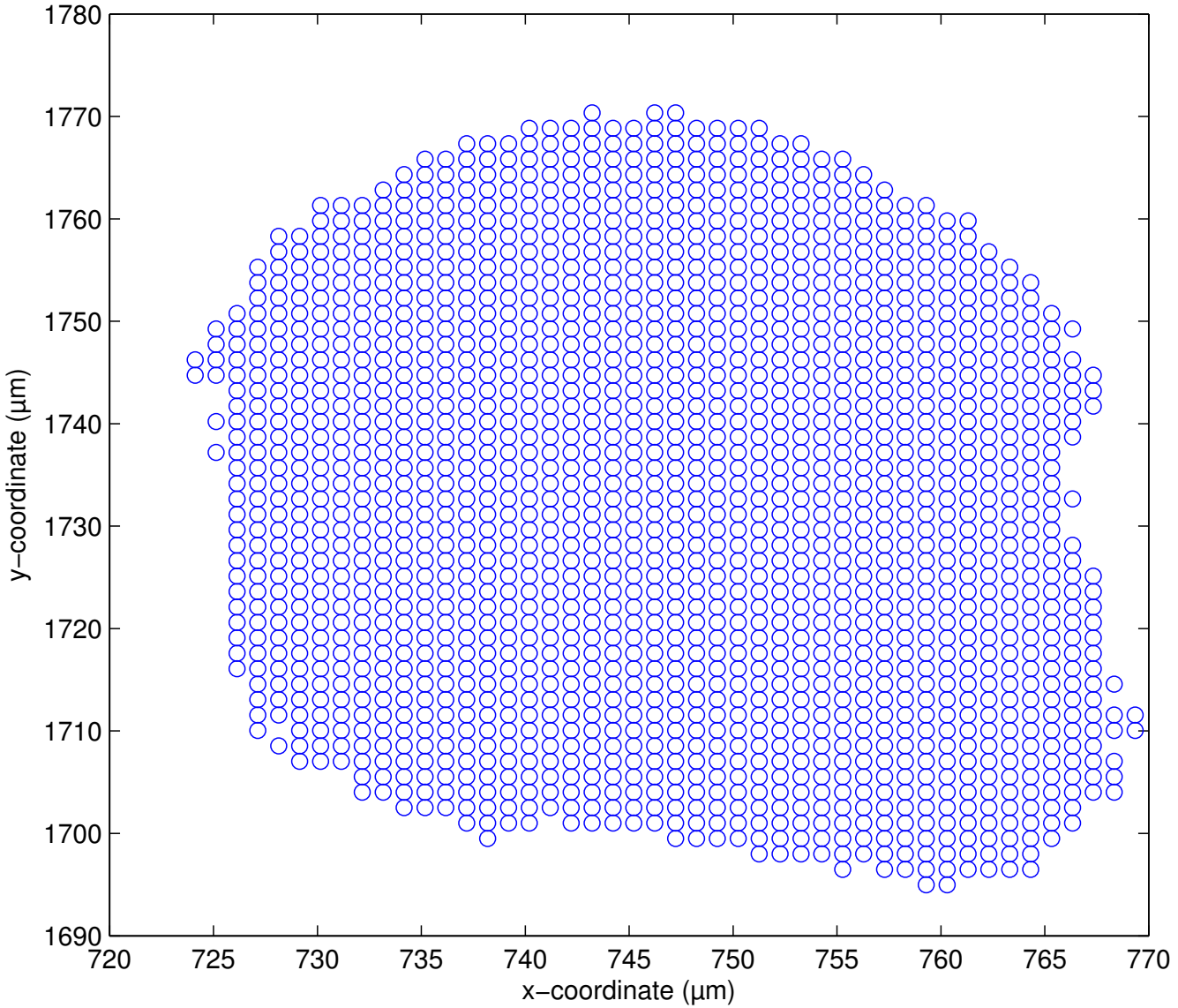




Finding the Correct Spot - Area Size



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.

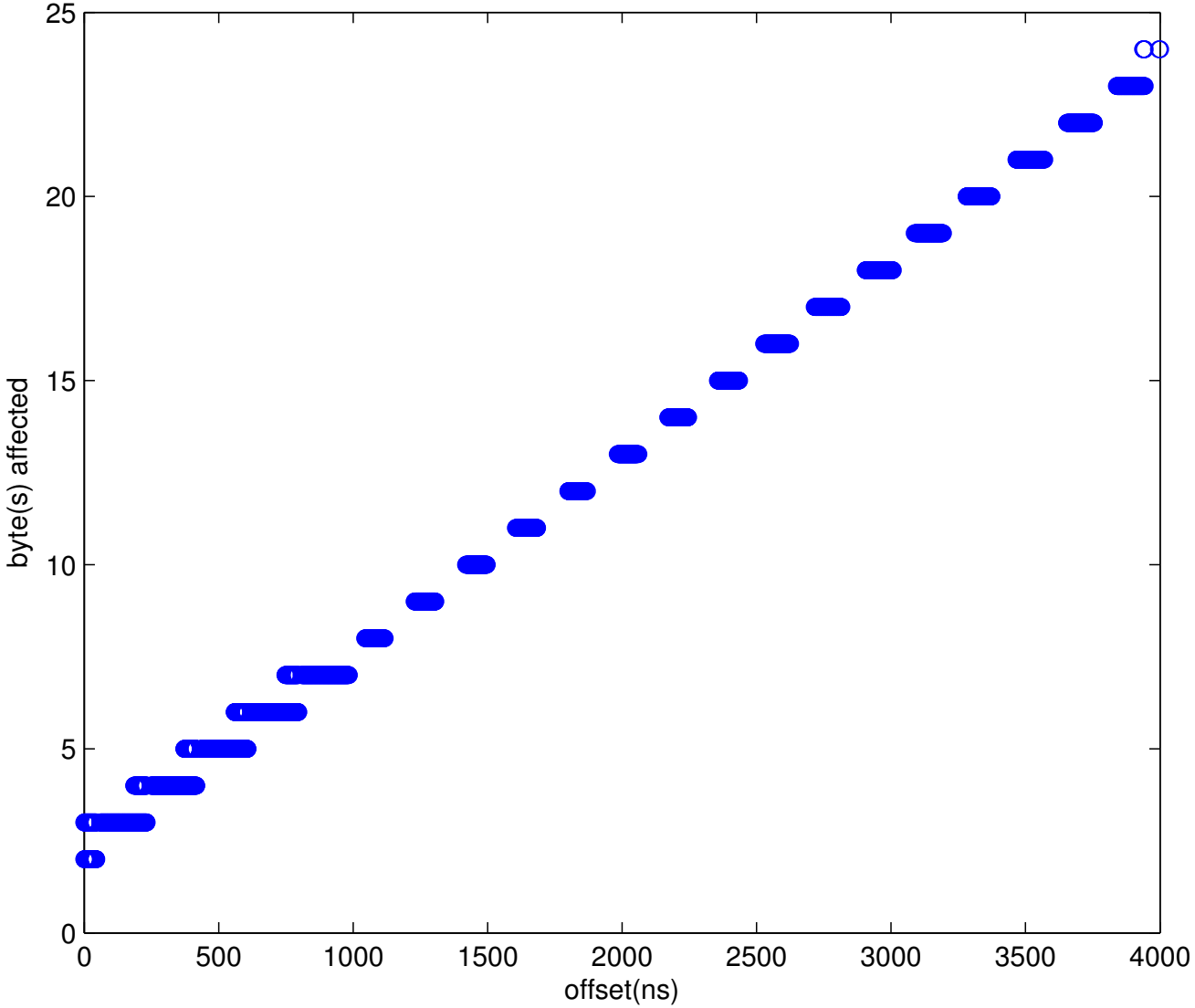
Profiling Phase 2/2

Following code snippet was repeated 25 times in the program with different registers (ATmega328P has 25 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.

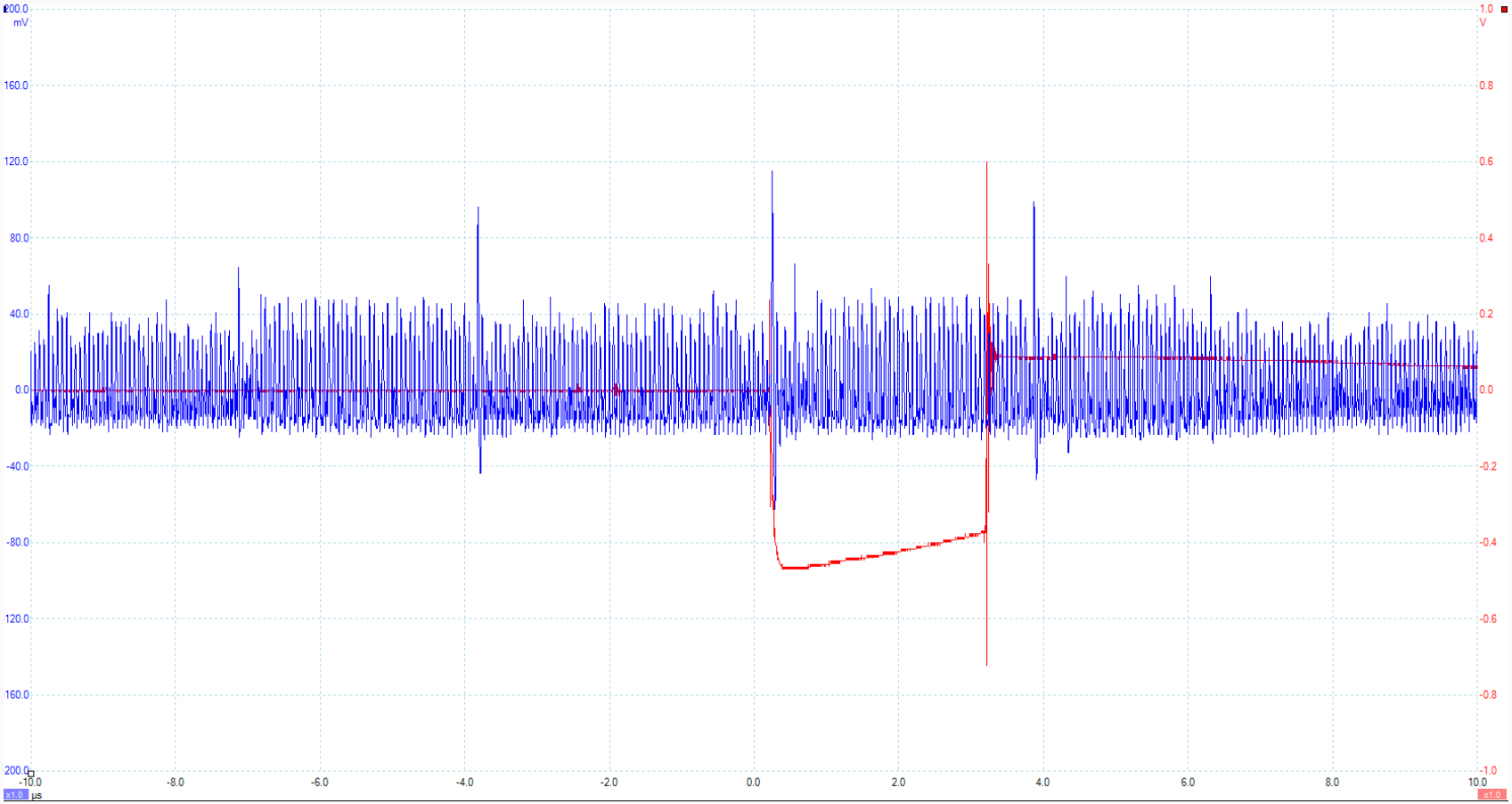
Profiling Phase - Skipping `EOR` Instruction



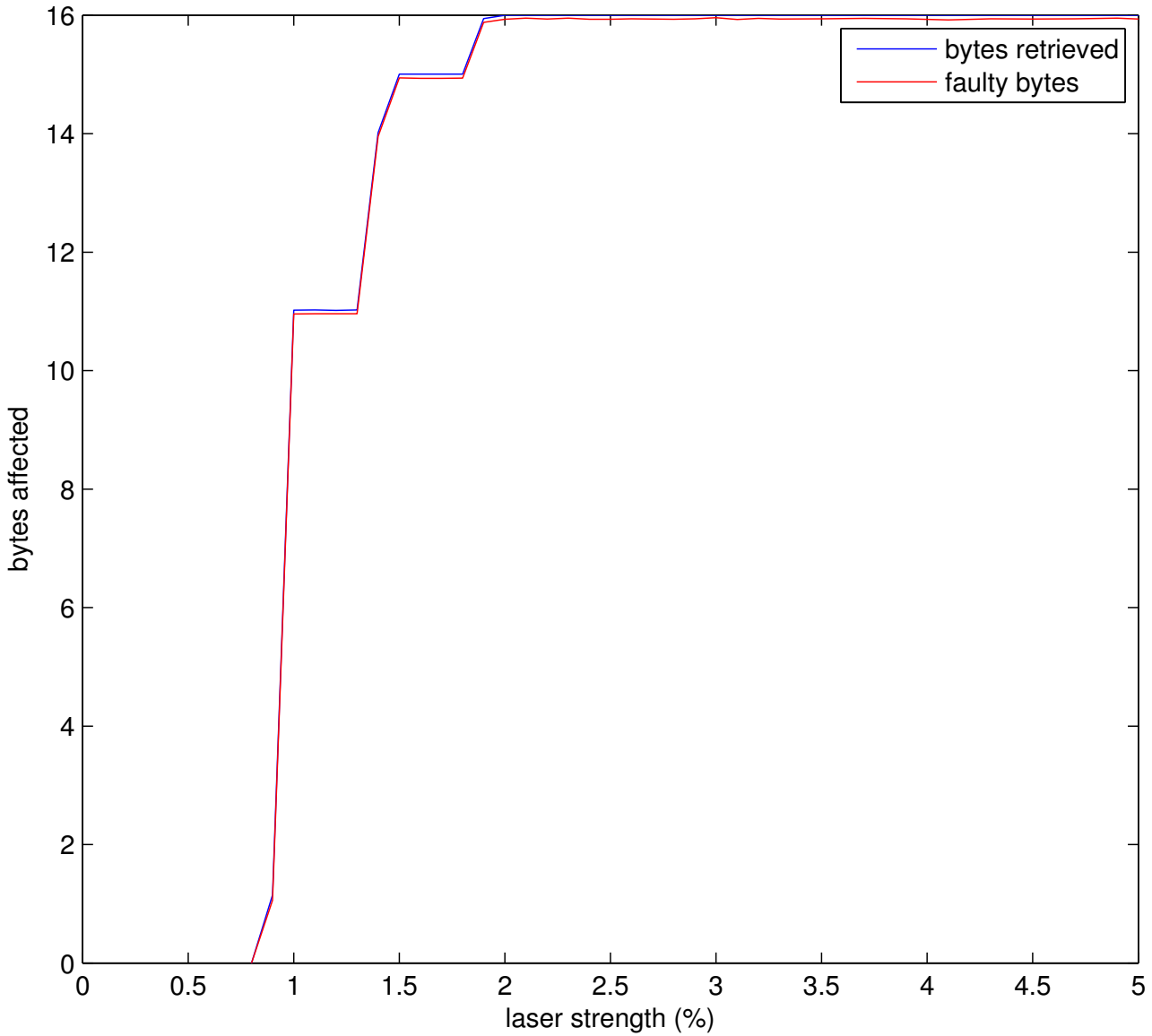
Attack on AES

- 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.

Virtex-5 and Bypassing the Parity Protection

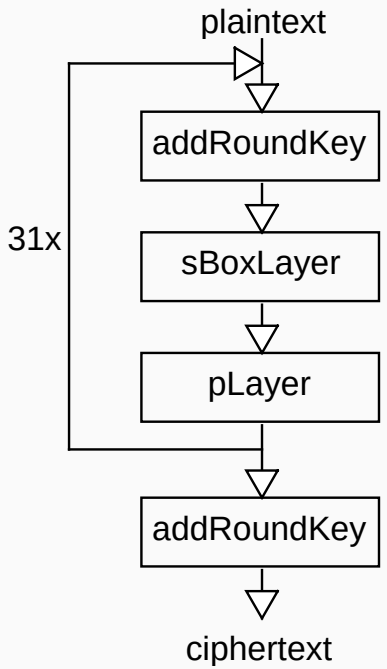
Parity Protection

- Parity is a low-cost fault detection solution for constrained devices
- It uses one redundant bit for fault detection, precomputed concurrently with the crypto algorithm and compared with the true round output
- In case the faults are detected by the parity bit/bits, the cipher execution can be immediately halted

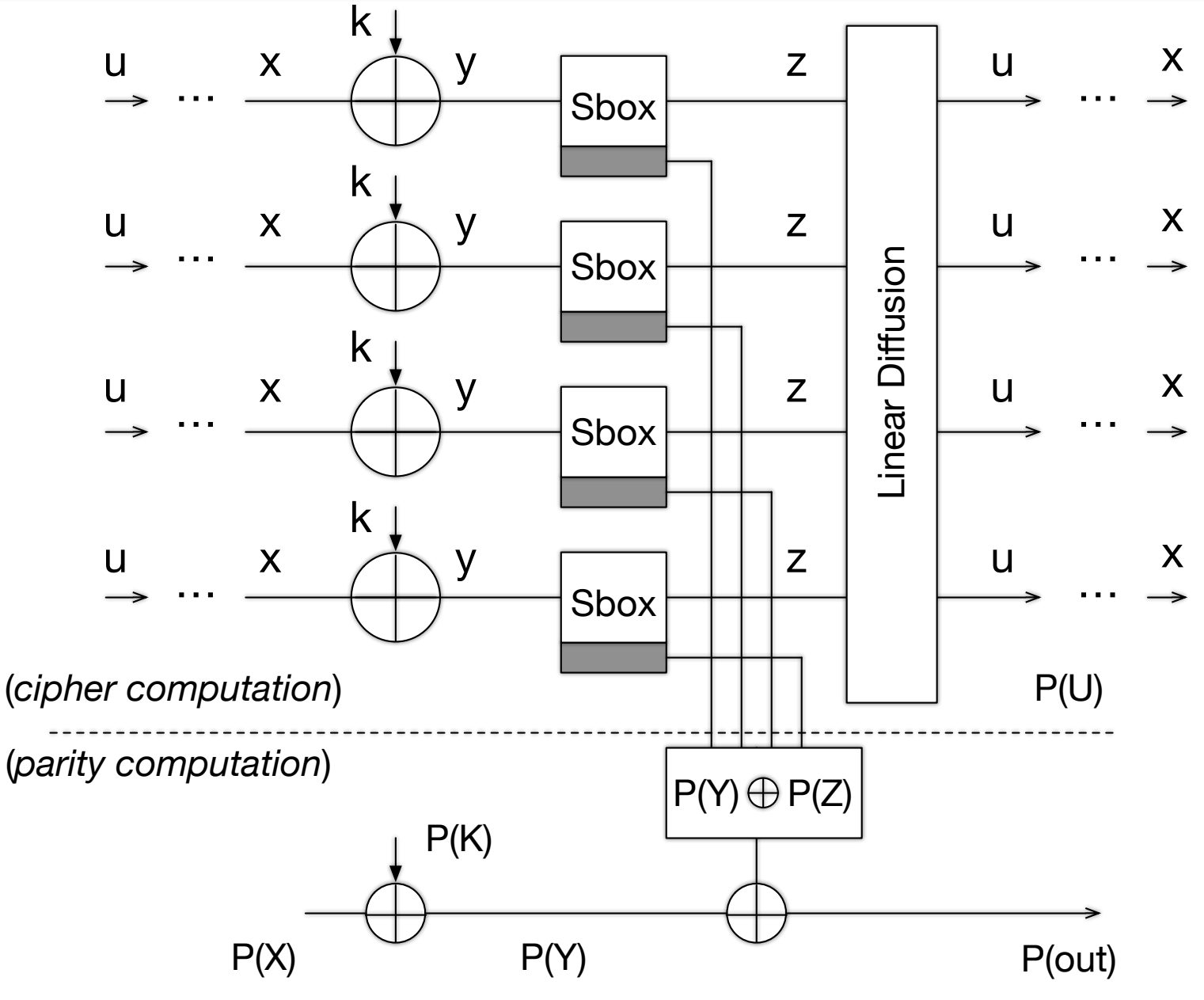
data	parity
00001111	0
00000111	1

PRESENT

- Lightweight block cipher
- Block size is 64b, key sizes 80b and 128b
- 31 rounds, consisting of *addRoundKey*, *sBoxLayer*, *pLayer*
- S-Box size is 4b



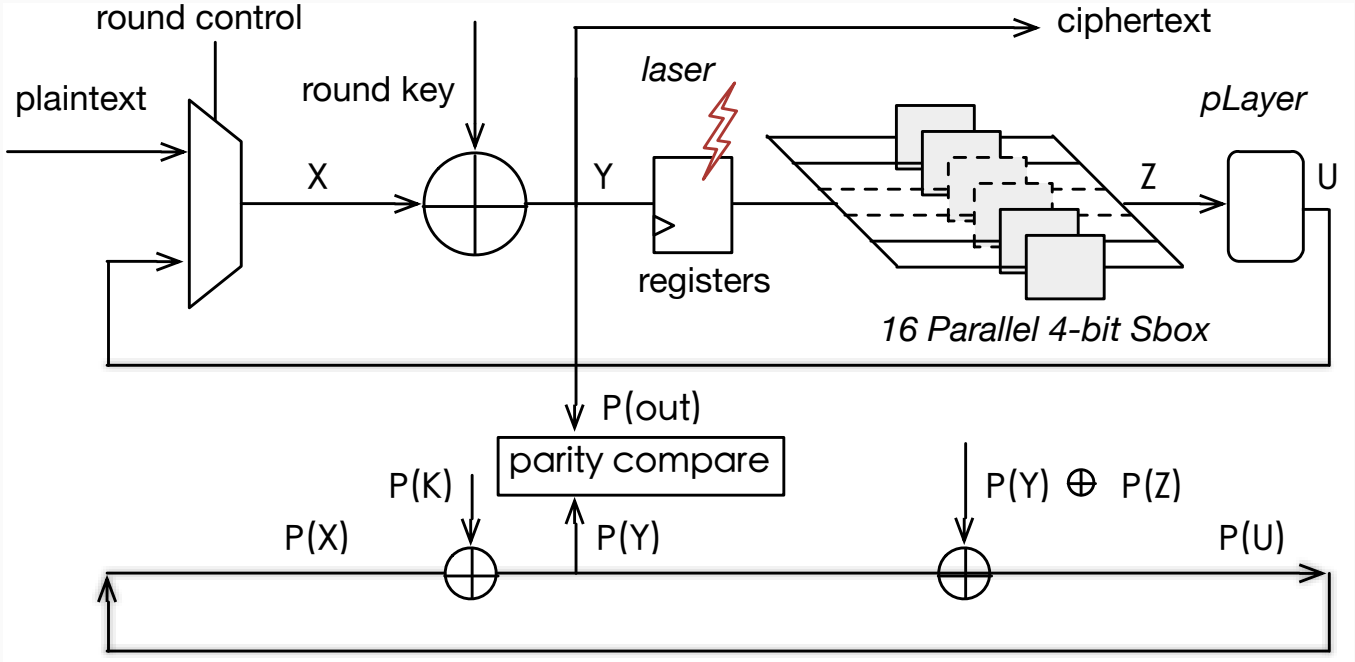
Error Detection Scheme



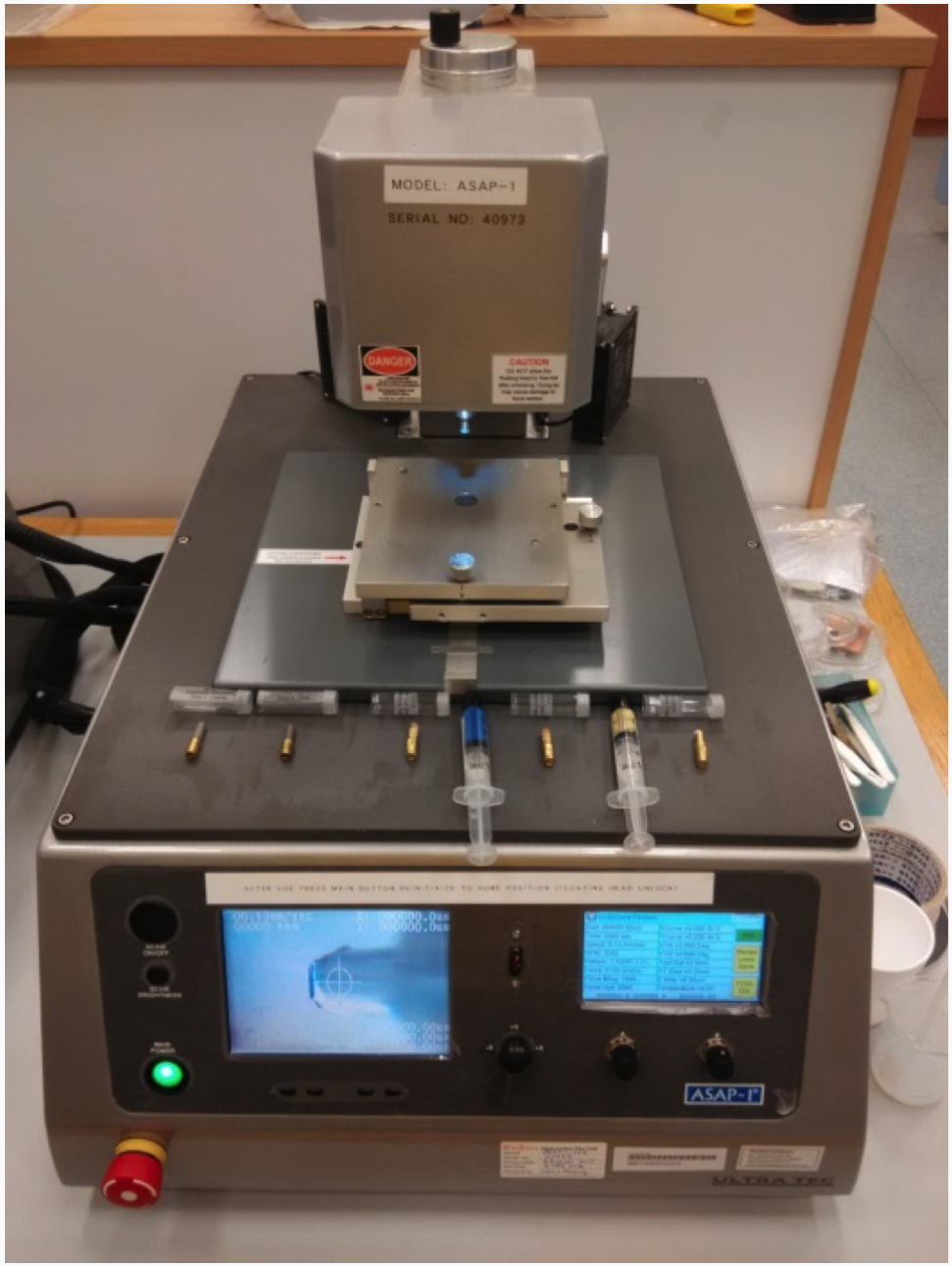
Bypassing the Parity

Attack steps:

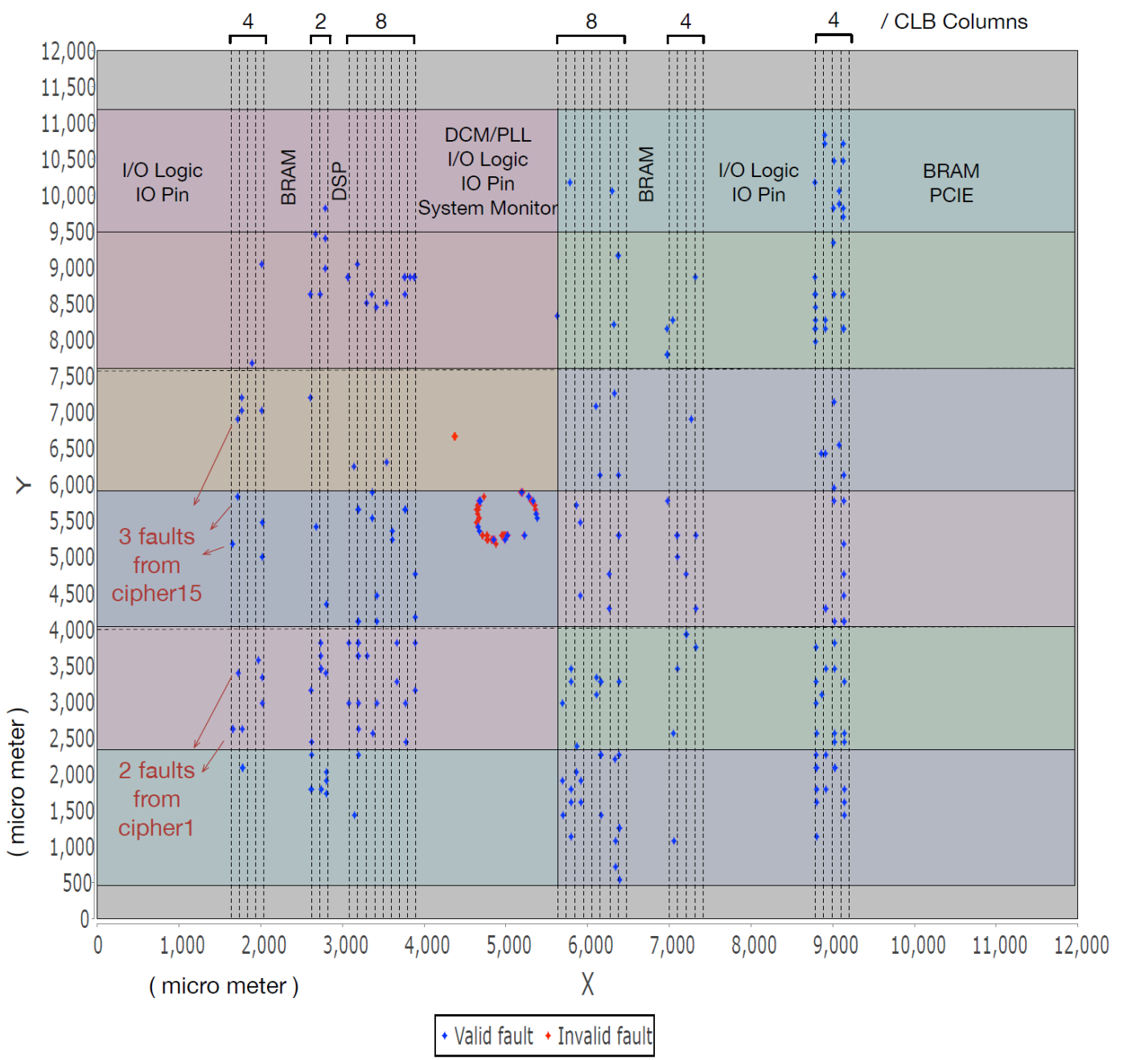
1. Delayer and polish the backside of Virtex-5 FPGA
2. Localize the implementation of PRESENT
3. Find a slice that processes the least significant nibble of the cipher
4. Disturb the four registers with an even fault mask



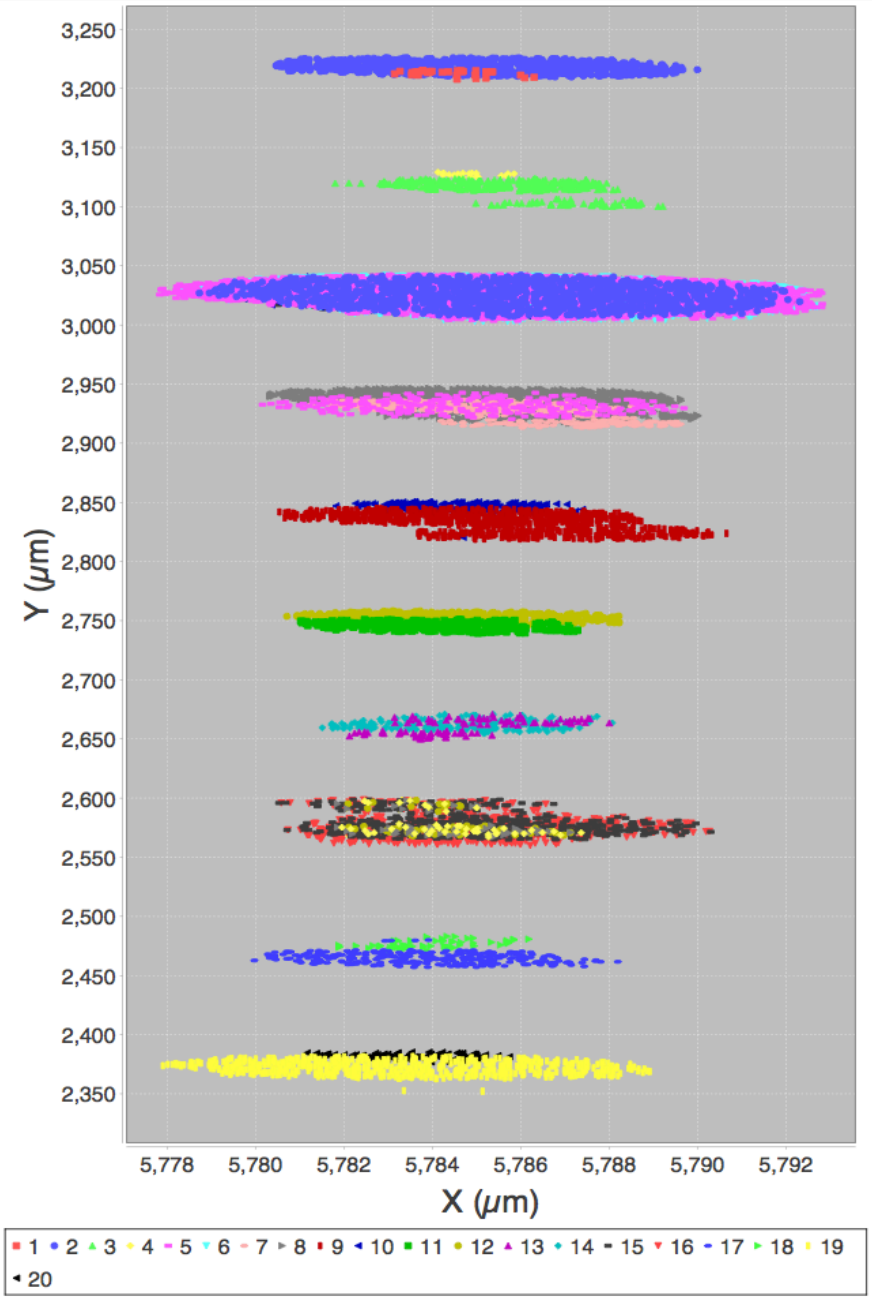
Delayering of Virtex-5 with Ultra Tec ASAP-1



Profiling Phase - Full Chip Scan



Profiling Phase - CLB Column Scan



Fault 1, 2 ⟷ (CLB 1: slice 1, 2)

Fault 3, 4 ⟷ (CLB 2: slice 3, 4)

Fault 5, 6, 2 ⟷ (CLB 3: slice 5, 6)

Fault 7, 8, 5 ⟷ (CLB 4: slice 7, 8)

Fault 9, 10 ⟷ (CLB 5: slice 9, 10)

Fault 11, 12 ⟷ (CLB 6: slice 11, 12)

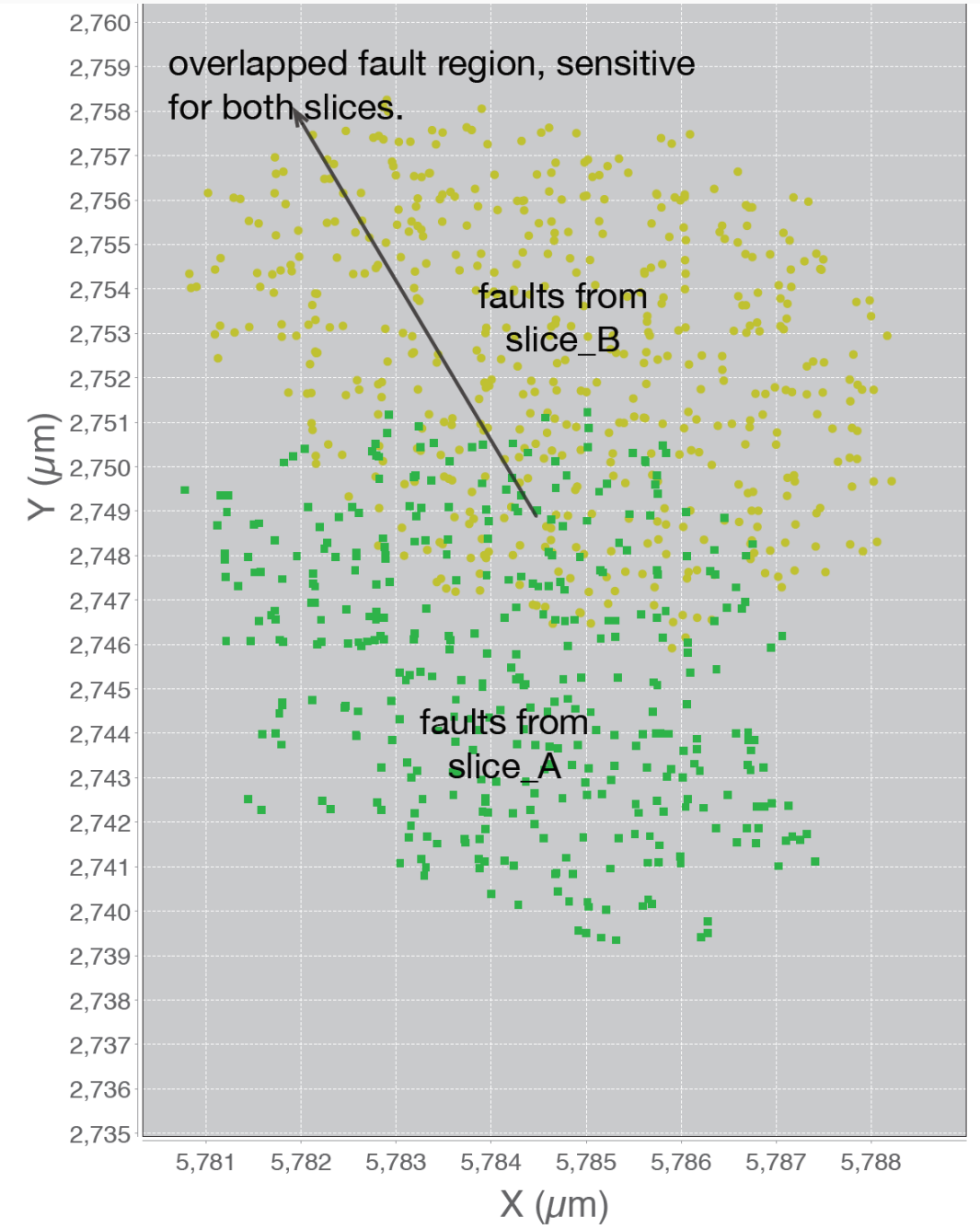
Fault 13, 14, 4 ⟷ (CLB 7: slice 13, 14)

Fault 15, 16 ⟷ (CLB 8: slice 15, 16)

Fault 17, 18 ⟷ (CLB 9: slice 17, 18)

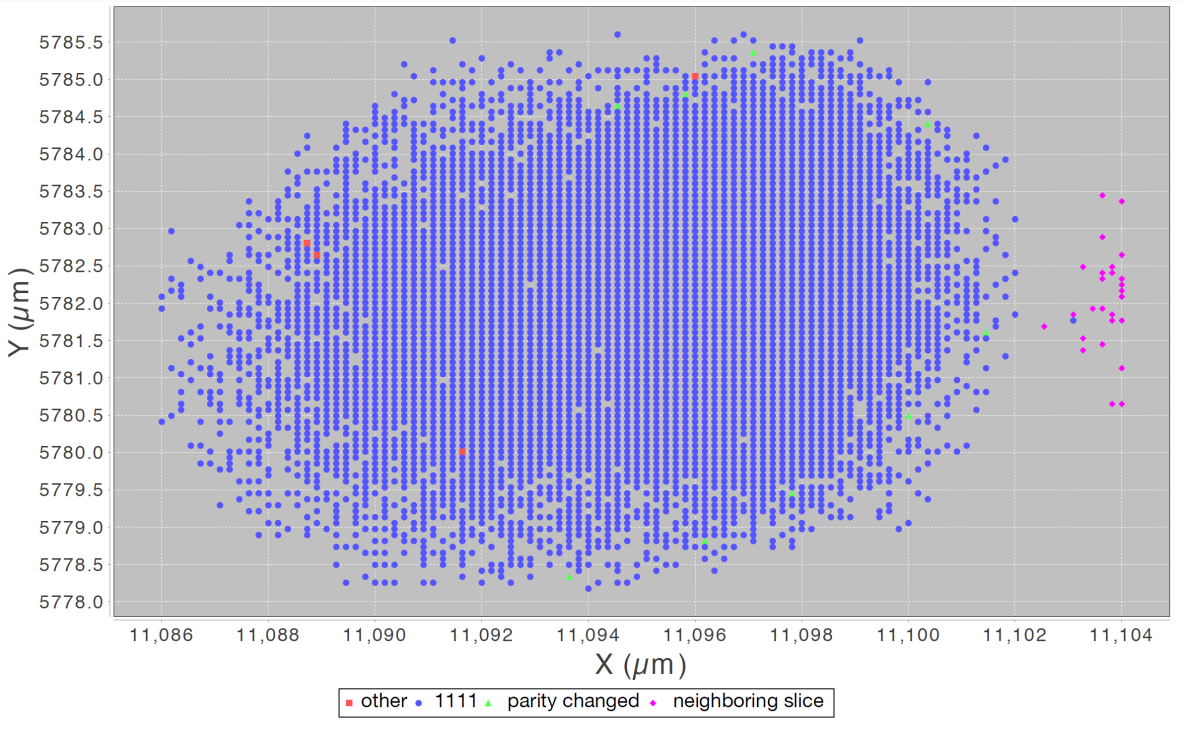
Fault 19, 20 ⟷ (CLB 10: slice 19, 20)

Profiling Phase - Single CLB Scan



Results

- After localizing the correct slice, we scanned the region of 12,000 points, out of which 4,947 resulted to faults
- With correct settings we were able to produce following even byte faults:
  - 1111 with probability of 99.212%
  - 0101 with probability of 0.081%



Countermeasures

How to Defend the Implementation?

Two main approaches

- *Fault detection* - error correction/detection codes, sensors, spatial/temporal redundancy, infection
- *Fault prevention* - special packages, sensors, metal layers

IC Package as a Countermeasure

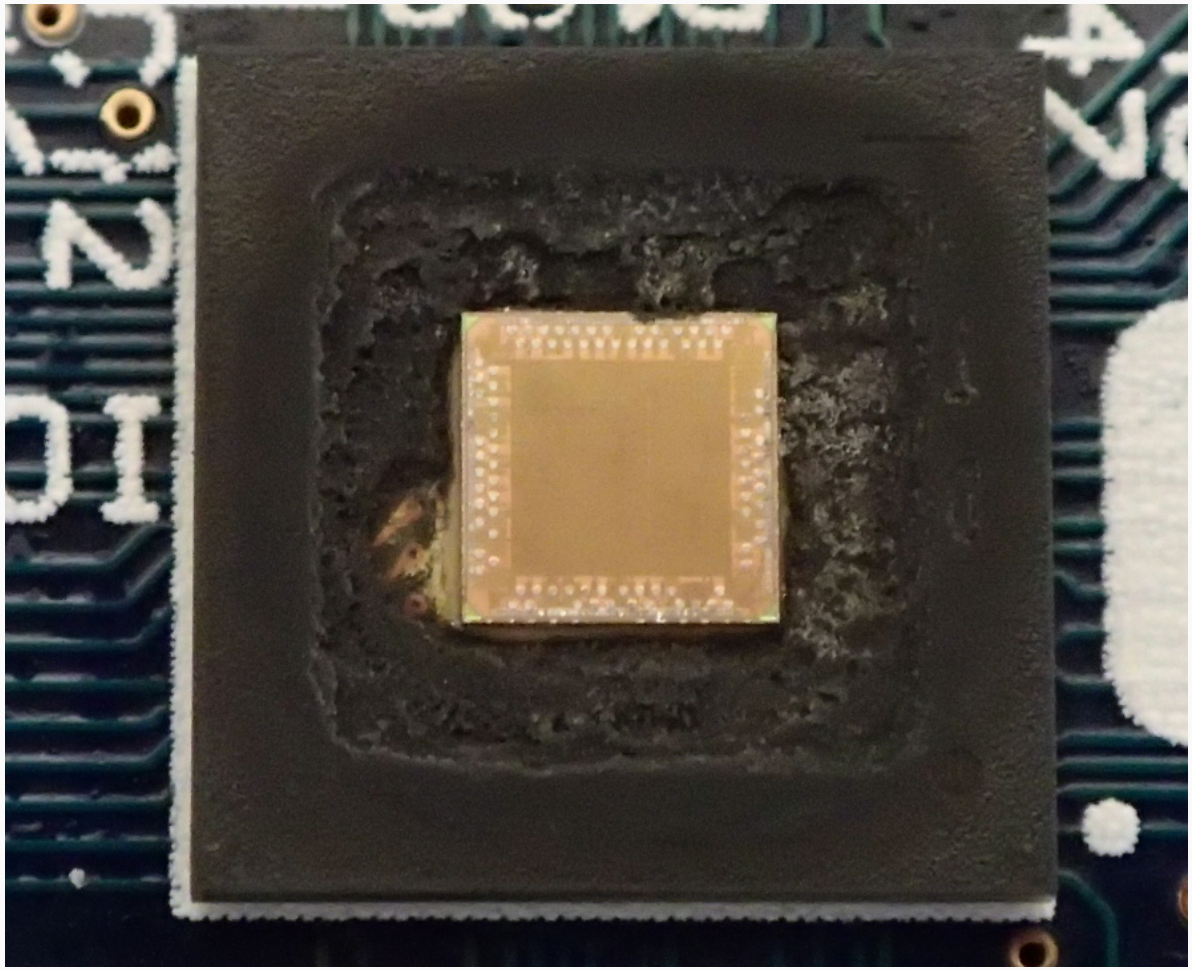


Figure 5: Bonding wires dissolved during the decapsulation process.

Redundancy

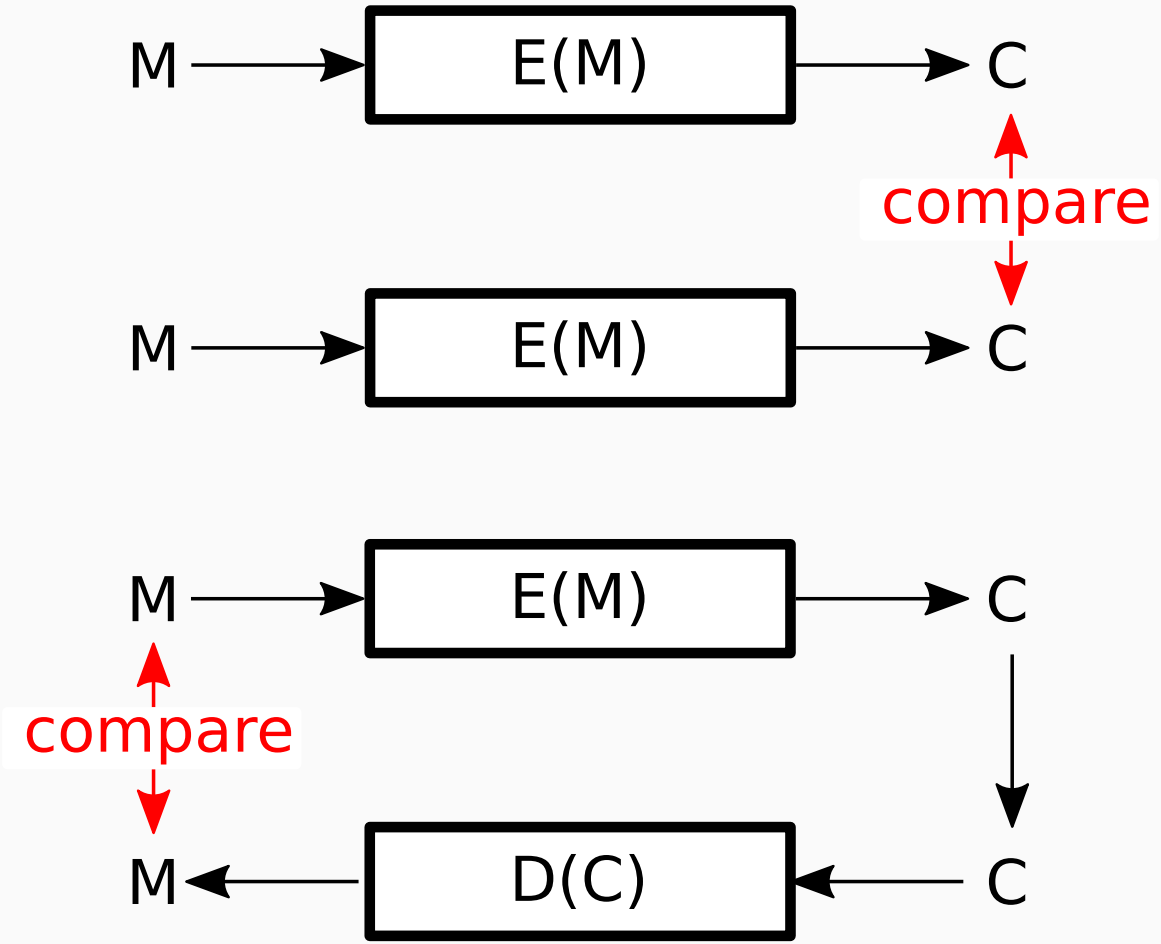


Figure 6: Basic redundancy approaches.

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?