Cryptanalysis of SP Networks with Partial Non-Linear Layers

Achiya Bar-On[1], Itai Dinur[2], Orr Dunkelman[3], Nathan Keller[1], Virginie Lallemand[4], and Boaz Tsaban[1]

[1]Bar-Ilan University, Israel

[2]École normale supérieure, France

[3]University of Haifa, Israel

[4]Inria, France

AES (Advanced Encryption Standard)

- Most **widely** used block cipher today
- **No known attacks** in standard models
- Does not imply that **implementations** are secure
- The attacker could to obtain **side channel** information about secret variables
  - Power analysis
  - Acoustic analysis
  - …

Masking

- A popular method to mitigate side channel attacks
- **Randomizes** the internal state
- Observation of few intermediate values during encryption does not provide information about **sensitive variables**
- Has performance **overhead**
- Particularly for **non-linear** computations

Zorro [GGNS13]

- Proposed at CHES 2013 by Gerard, Grosso, Naya-Plasencia, and Standaert
- **Substitution-permutation** (**SP**) network (as AES)
  - Operations are "local" **non-linear substitutions** (Sboxes) and "global" **linear permutations**
- Design resembles AES, with **less Sboxes**
  - First design of SP network with partial non-linear layers
- Allows more **efficient masking** by **minimizing** non-linear computation

Security of Zorro

- We are mainly interested in security against **differential** and **linear attacks**
- AES was designed to **resist** these attacks
- However, current analysis tools are **inefficient** for SP networks with partial non-linear layers

Security of Zorro

- Designers of Zorro used **heuristic** arguments to claim security against differential attacks
- However, these arguments failed
  - "Differential Cryptanalysis and Linear Distinguisher of Full-Round Zorro" (Wang et al. 2013)
  - "Total Break of Zorro Using Linear and Differential Attacks" (Rasoolzadeh et al. 2014)
- Use techniques based on **specific** properties of Zorro's linear layer
  - Do not apply for slightly modified variants

This Work

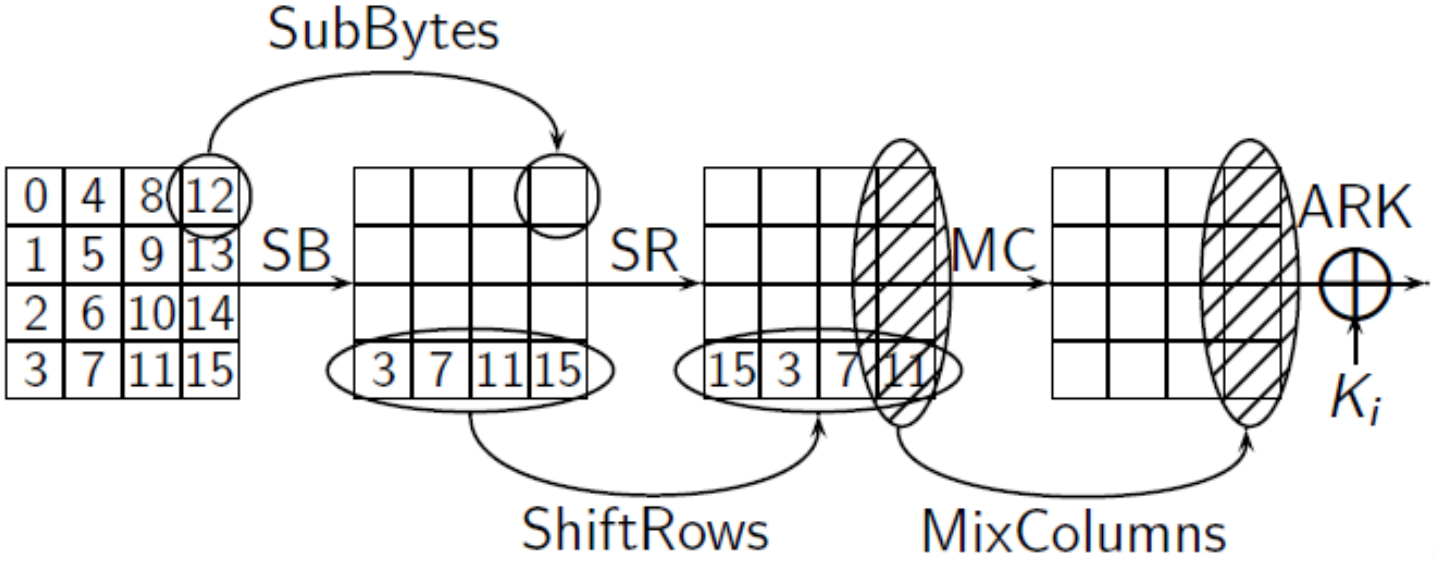
- Develop **generic analysis techniques** for **differential** (and **linear**) cryptanalysis of SP networks with partial non-linear layers
  - Techniques can be used both for **finding** such attacks or **proving resistance** against (the basic form of) these attack
- Applications:
  - A **practical attack** on full Zorro
    - Fully simulated in a few days on a single PC
  - Propose a **light modification** of Zorro that offers much **better resistance** to differential and linear cryptanalysis

Summary

- Description of AES-128 and Zorro
- Differential cryptanalysis of AES and inapplicability of tools to Zorro
- The model for SP-networks with partial non-linear layers
- Our tools
  - Differential characteristic search
  - Key recovery
- Applications
- Conclusions

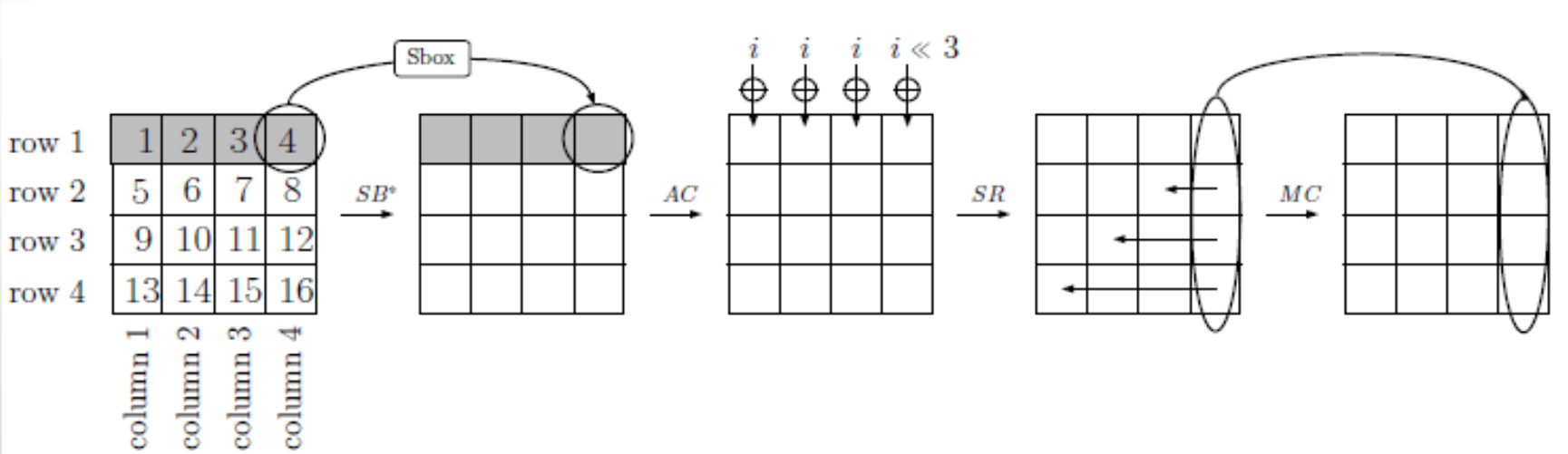
AES-128

- Key size in 128 bits, block size is 128 bits
- Plaintext is encrypted using 10 (almost) identical **rounds**
  - **Round keys** computed from master key according to a key schedule

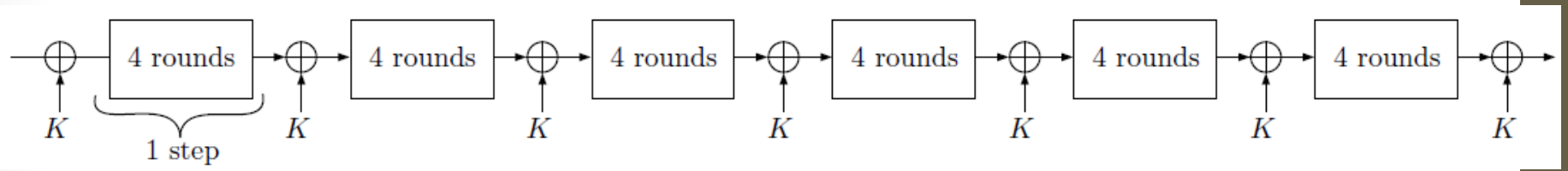


Zorro

- Key size in 128 bits, block size is 128 bits
- Plaintext is encrypted using 24 (almost) identical **rounds**



Zorro Key Schedule



Masking Zorro

- Number of Sboxes computations in AES-128 encryption is $16 \cdot 10 = 160$
- **Reduced** to $4 \cdot 24 = 96$ in Zorro
- Furthermore, the Sbox is different
- Masking Zorro is much **more efficient** compared to AES-128

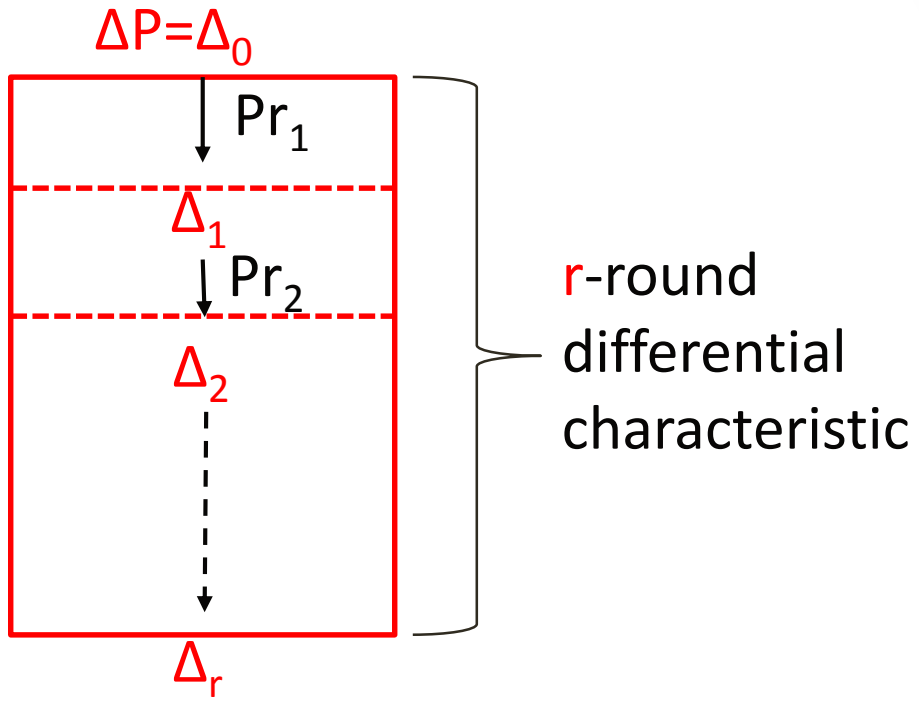
- What about **security**?

Differential Cryptanalysis [BS'90]

- Study the **statistical evolution of differences** in the state of the cipher between encryptions of 2 **plaintexts**
- Cipher is **vulnerable** if: there is an input difference to round 1 Δ_0, and output difference of round r Δ_r such that Δ_0 is mapped to Δ_r with **high probability** (for large r)
  - (almost) independent of the key

Differential Cryptanalysis [BS‘90]

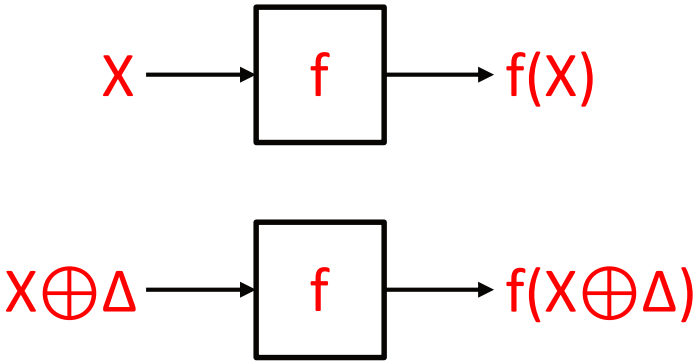
- Analyzed by constructing a **differential characteristic**



- Estimate $Pr=Pr_1 \cdot Pr_2 \ldots \cdot Pr_r$

Differential Cryptanalysis

- Given pair $X, X \oplus \Delta$ input to f, what is $f(X) \oplus f(X \oplus \Delta)$?
  - If f is constant A addition, $(X \oplus A) \oplus (X \oplus \Delta \oplus A) = \Delta$ w.p. 1
  - If f is linear mapping L, $L(X) \oplus L(X \oplus \Delta) = L(\Delta)$ w.p. 1



Differential Cryptanalysis

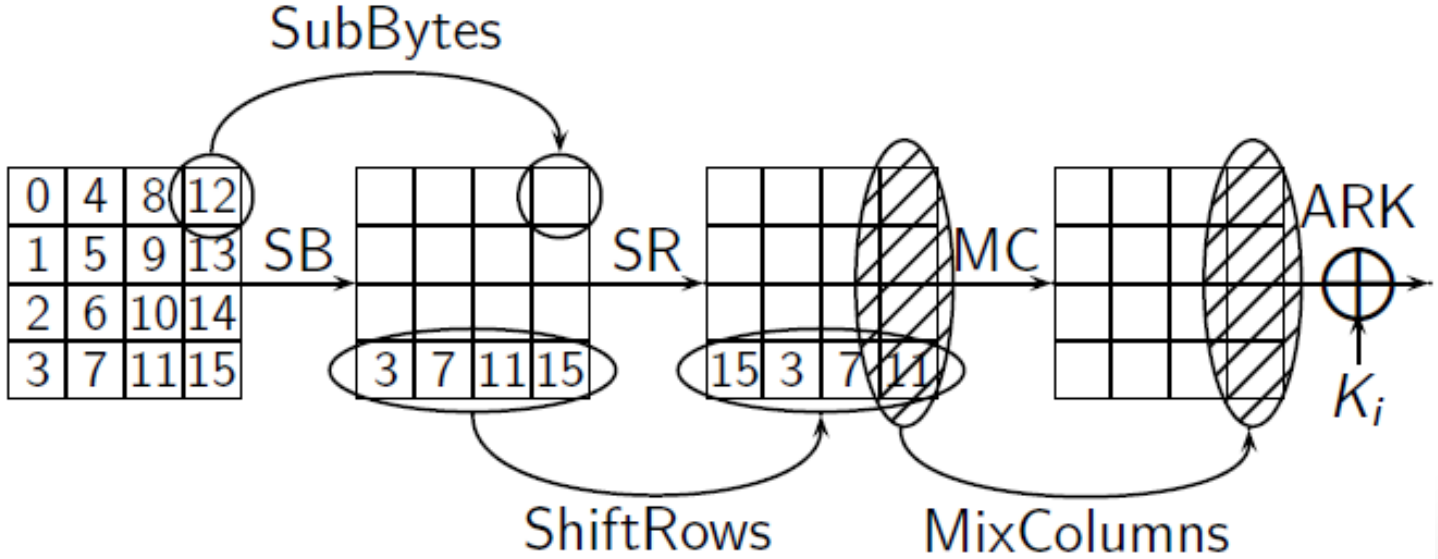
- Given pair $X, X \oplus \Delta$ input to non-linear Sbox S
  - If $\Delta=0$ then $S(X) \oplus S(X \oplus \Delta)=0$ w.p. 1
  - Otherwise, the output difference depends on X
  - Assuming X is uniformly distributed, the output difference is distributed according to the **difference distribution table** (**DDT**) for an Sbox
  - The DDT associates a **probability** to each $(\Delta_{in}, \Delta_{out})$

Resisting Differential Cryptanalysis

- Goal of designer: ensure that the probability of **any** differential characteristic for r rounds is sufficiently small
  - e.g. less than 2^{-128} for a 128-bit cipher

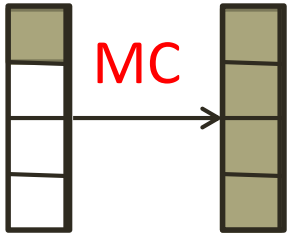
Differential Cryptanalysis of AES

- **ARK** can be ignored
- **SB** is the only non-linear function
- In a differential characteristic, if an Sbox (byte) has **non-zero** input\output difference it is **active**
  - A transition through an active Sbox has $pr=2^{-6}$ or $pr=2^{-7}$
- If an Sbox (byte) has **zero** input difference it is **inactive**



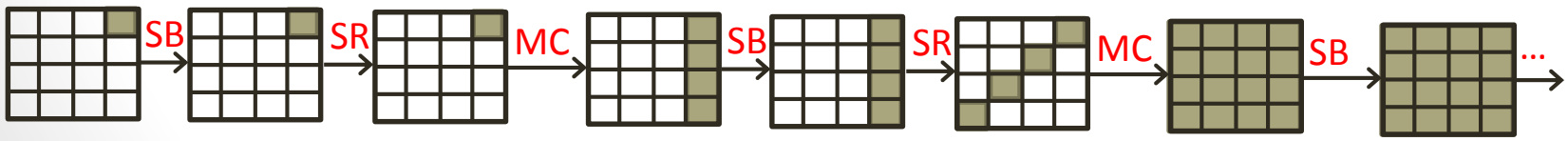
Differential Cryptanalysis of AES

- **MC** has branch number 5 (over **GF(2^8)**)
- For a non-zero input the number of **active** bytes in the input and output is **at least** 5

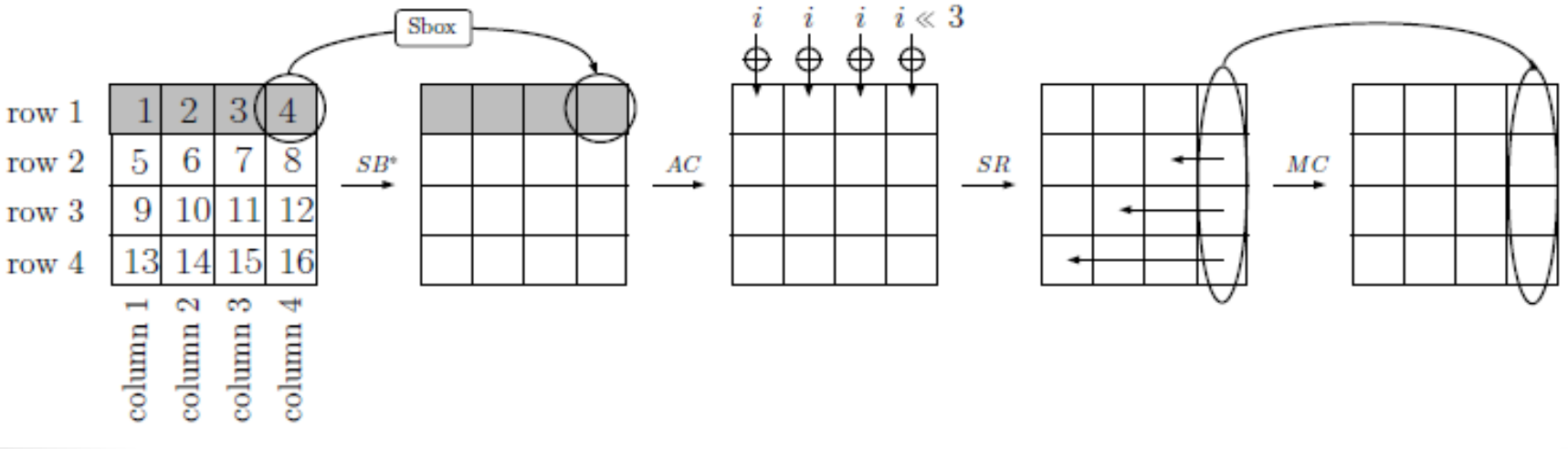


Differential Cryptanalysis of AES

- A characteristic that starts with 1 active byte has 1+4+16=21 active Sboxes for 3 rounds
  - Probability **at most** $2^{-6 \cdot 21} = 2^{-126}$
- Moreover **any** characteristic of 4 rounds has **at least 25** active Sboxes ("wide trail strategy")
  - Probability **at most** $2^{-6 \cdot 25} = 2^{-150}$

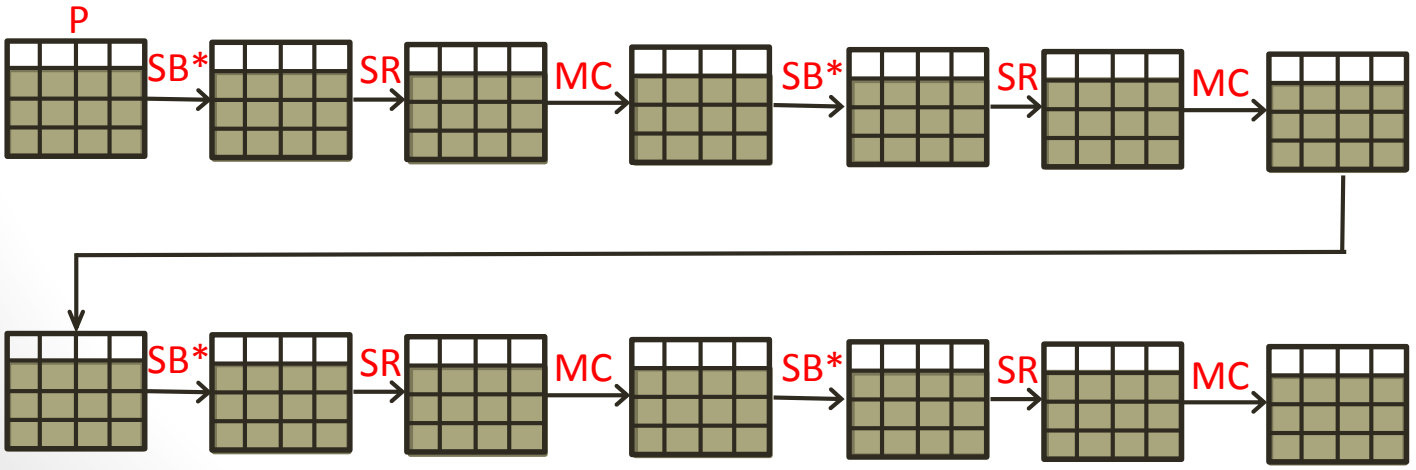


Zorro Round Function



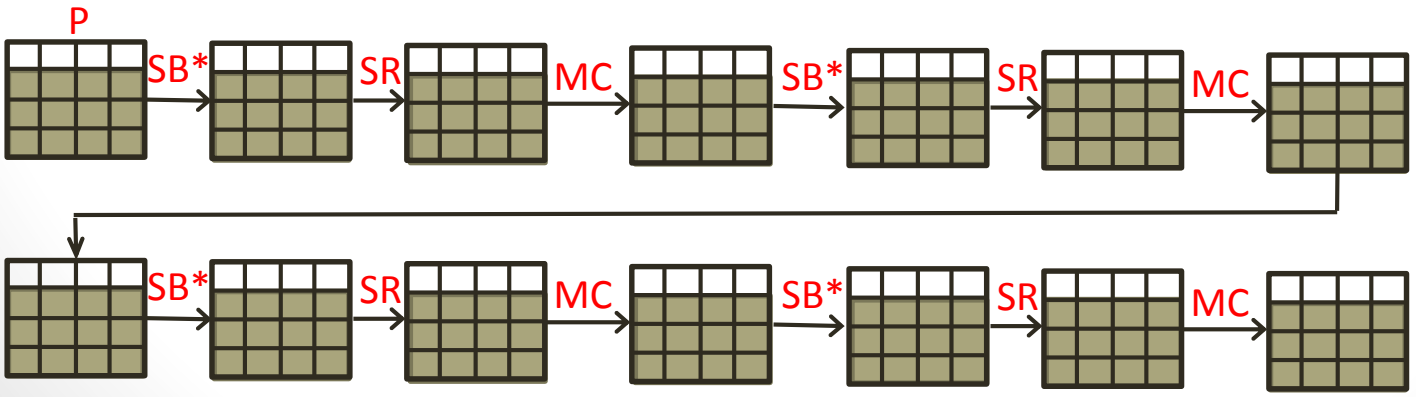
Differential Cryptanalysis of Zorro

- For AES such characteristic has $12 \cdot 4 = 48$ active Sboxes
- For Zorro there are 0 active Sboxes -> the probability of such a characteristic is 1 !
- However, there is no such **valid** characteristic for 4-round Zorro



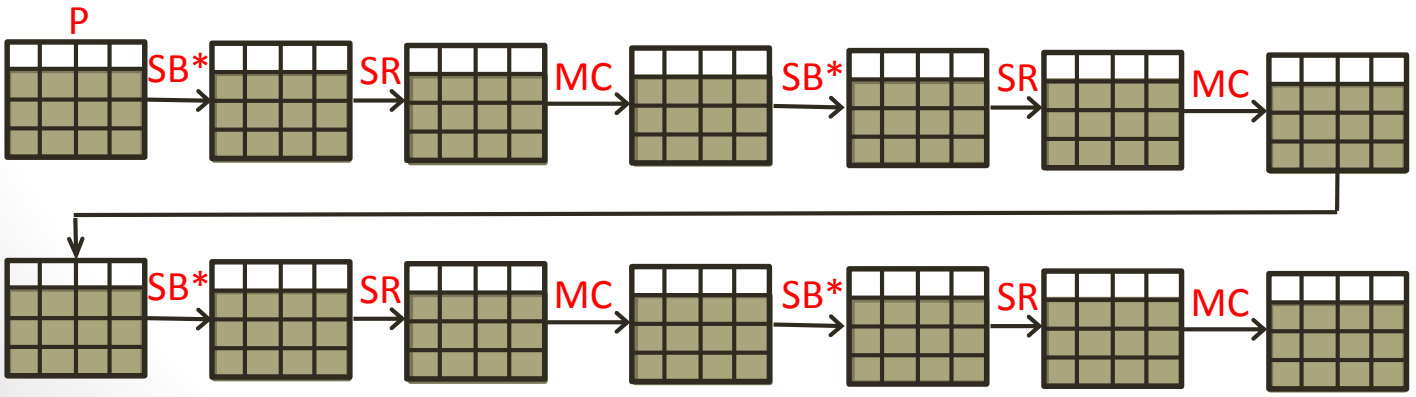
Differential Cryptanalysis of Zorro

- There is no characteristic w.p. 1 for 4-round Zorro
  - There are $2^{12 \cdot 8}=2^{96}$ possible input differences (96 binary variables)
  - Each **MC** operation imposes $4 \cdot 8=32$ constraints
  - After 3 rounds the number of variables and constraints are equal -> with good probability there is **no solution**



Differential Cryptanalysis of Zorro

- Must "keep track" of a **global system of constraints** imposed on a characteristic
- Previously published generic differential search algorithms only check "**local consistencies**"
  - Exhaust a **huge** amount of **invalid** characteristics that appear to have pr=1



Differential Cryptanalysis of Zorro

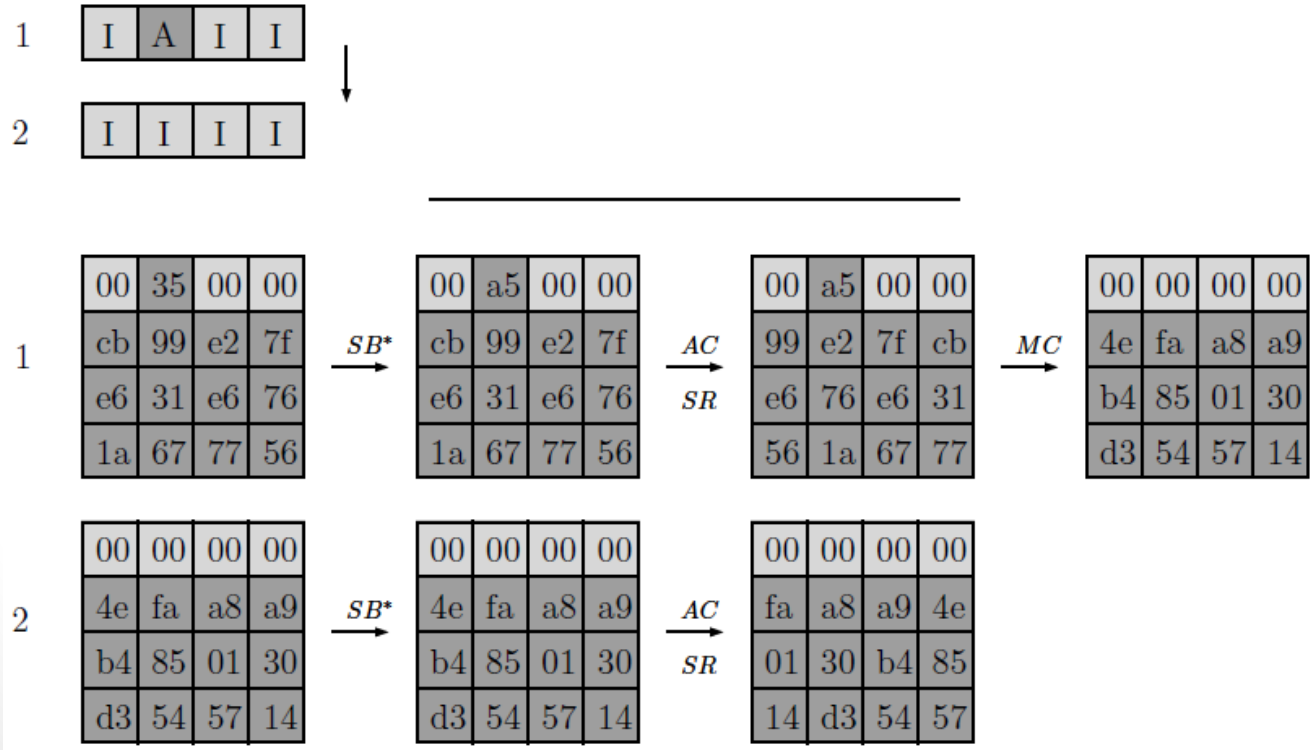
- Previous generic differential search algorithms are **inefficient** for Zorro
- Designers of Zorro used **heuristic** arguments to claim security against differential attacks
- However, these arguments **failed** (Wang et al. 2013), (Rasoolzadeh et al. 2014)
  - Constructed high-probability differential\linear characteristics based on the **linear layer of Zorro**

Model of SP-networks with Partial Non-Linear Layers

- Assume (for simplicity): AES-like state (contains 16 words of 8 bits)
- Consists of 3 layers:
  - Key addition layer: XOR key into state
  - S-box layer S: Apply Sbox to t out of the 16 state bytes (words)
  - Linear layer L: Apply **arbitrary** linear layer to state

Patterns

- **Pattern** for a characteristic: A description of the **activity** for each of its spanned S-boxes
  - Similar concept was defined by Biryukov and Nikolic (Eurocrypt 2010)



New Characteristic Search Algorithm

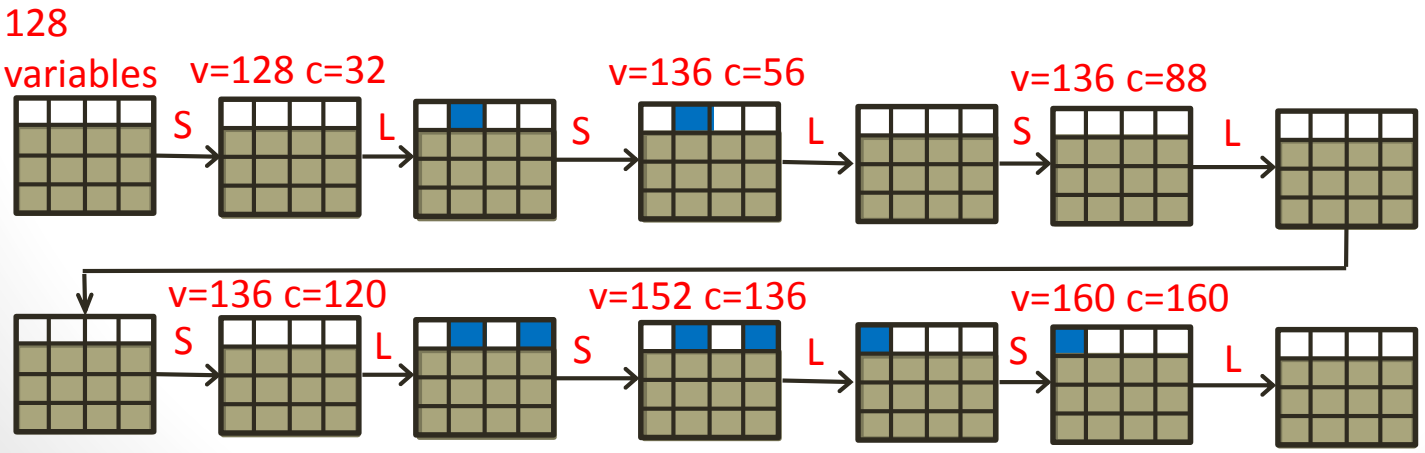
- Assume: t (Sboxes per layer) is small (t=4 for Zorro)
- Goal: find **high probability** differential characteristic for r rounds, or prove that one **does not exist**
- Interested in **patterns** with a **small number of active** Sboxes a
- Observations:
  - Number of possible patterns is **small**: $\binom{tr}{\leq a} = \sum_{i=0}^{a} \binom{tr}{i}$
    - For r=10, t=4, a=10 the number of patterns is smaller than 2^{32}
  - Characteristics following a fixed pattern can be **enumerated** efficiently

Enumerating Characteristics for a Pattern

- Observation: all characteristics that follow a pattern reside in a **restricted linear subspace**
- Computing the subspace for a **given** r-round pattern:
  - For i=0,1,2,…,r compute the **symbolic linear representation** of Δ_i

Enumerating Characteristics for a Pattern

- Example: r=6 pattern with a=4 active Sboxes (4·6-4=20 inactive Sboxes)
- Final system: 128+4·8=160 variables 20·8=160 constraints
- A **solution** to the system gives the actual **characteristic**



New Characteristic Search Algorithm

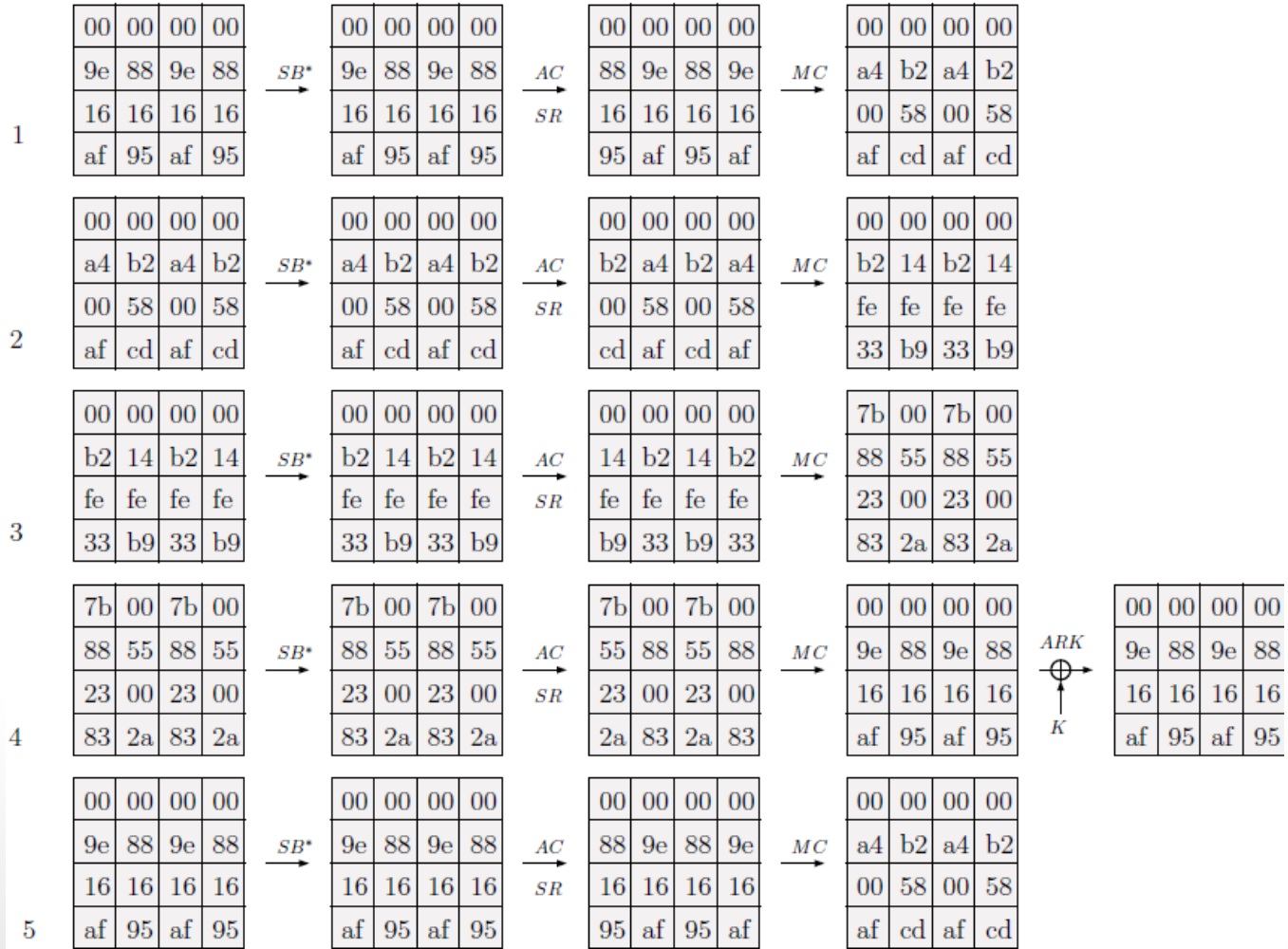
- Enumerating characteristics with **at most** a active Sboxes (or proving non existence)
- For each of the $\binom{tr}{\leq a}$ patterns:
  - Compute the **linear subspace** of all characteristics that follow the pattern
  - **Post-filter** the characteristics according to DDT
- Assuming a is small -> the size of the subspaces is not too big -> the **complexity** is about $\binom{tr}{\leq a}$
- Optimization: combine analysis of patterns with a **common prefix**

Implication of Characteristic Search Algorithm

- For r rounds with a active Sboxes we have tr-a inactive Sboxes
- There are 128+8a variables and 8(tr-a) constraints -> we expect **no solution** when 128+8a < 8(tr-a) or r>(16+2a)/t
- Gives a **lower bound** on the number of rounds needed to **resist** (basic) differential cryptanalysis
  - If we want that all differential characteristics have $pr<2^{-128}$ -> a=128/6≈22 (using AES Sbox) -> r>60/t

Application to Zorro

- Found (iterative) differential characteristic for 19-round Zorro with probability 2^{-43}



Application to Zorro

- Combined with the **key recovery** technique we get an attack with **complexity** 2^{45}
  - The attack was **simulated** several times
- We also devised a **linear attack** with the same complexity

- We propose a **light fix** to Zorro and **formally prove** that it provides **much better resistance** to basic differential\linear cryptanalysis

The Choice of t

- To resist differential attacks:
  - We add $16/t$ rounds at the end of the cipher
  - Recall: we expect **no differential characteristic** with (at most) a active Sboxes when $r \geq (16+2a)/t$
- The number of required Sboxes is $tr \geq r(32+2a)$ which is **independent** of t
- Why not choose $t=16$ as in AES?
- A small value of t (for fixed tr) provides **better resistance** against **structural attacks**

Conclusions

- We developed **new techniques** for **differential** and **linear** cryptanalysis of SP networks with partial non-linear layers
  - Give a **practical attack** on full Zorro
- Our tools will be useful to **design secure** SP networks with **partial non-linear layers** in the future

Thank you for your attention!