Radboud University

Building a permutation: comparing design approaches

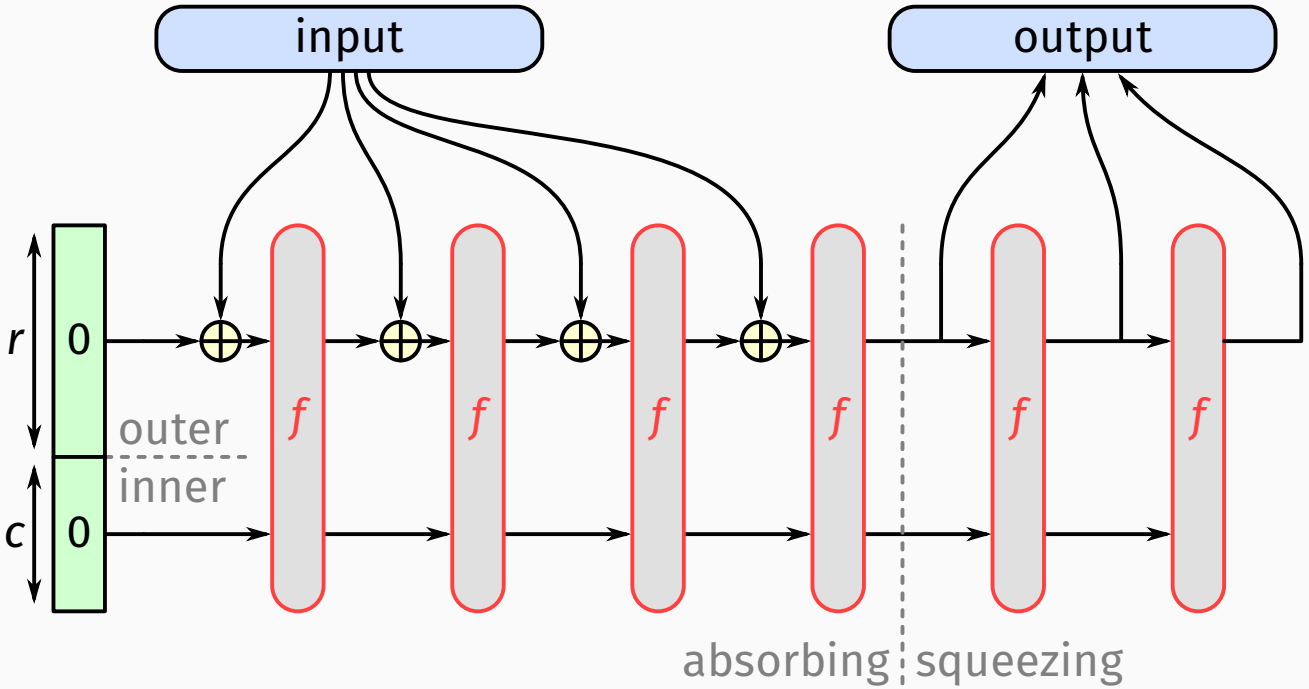
Joan Daemen[1]

based on joint work with Nicolas Bordes[3], Daniël Kuijsters[1] and Gilles Van Assche[2]

Summer School on real-world crypto and privacy, June 17-21, 2019, Šibenik

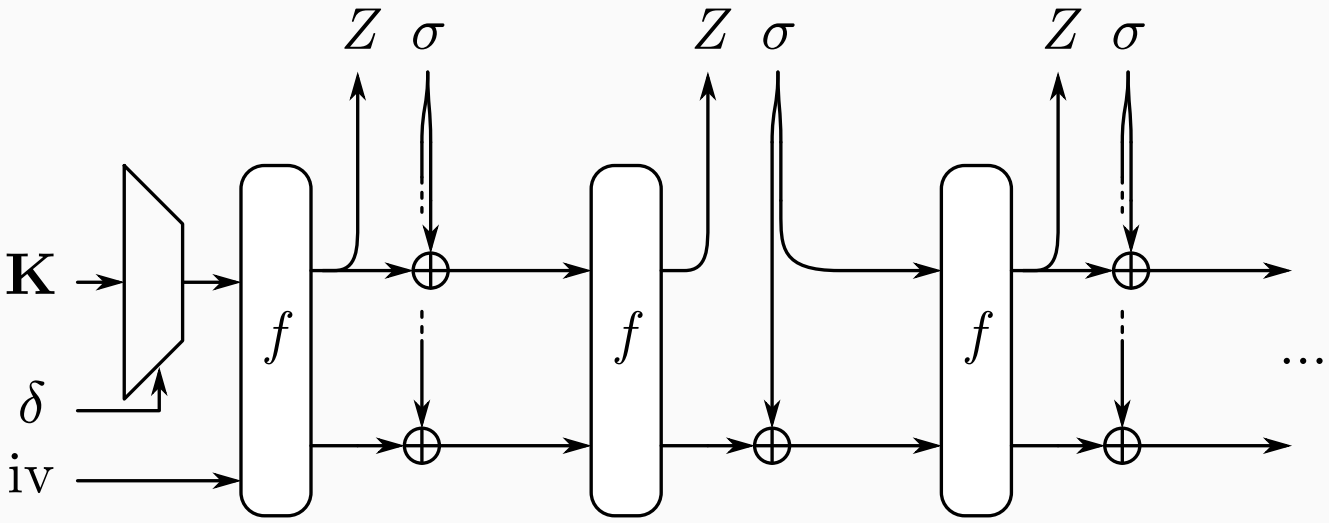
[1]Radboud University [2]STMicroelectronics [3]Université Grenoble Alpes

The sponge construction



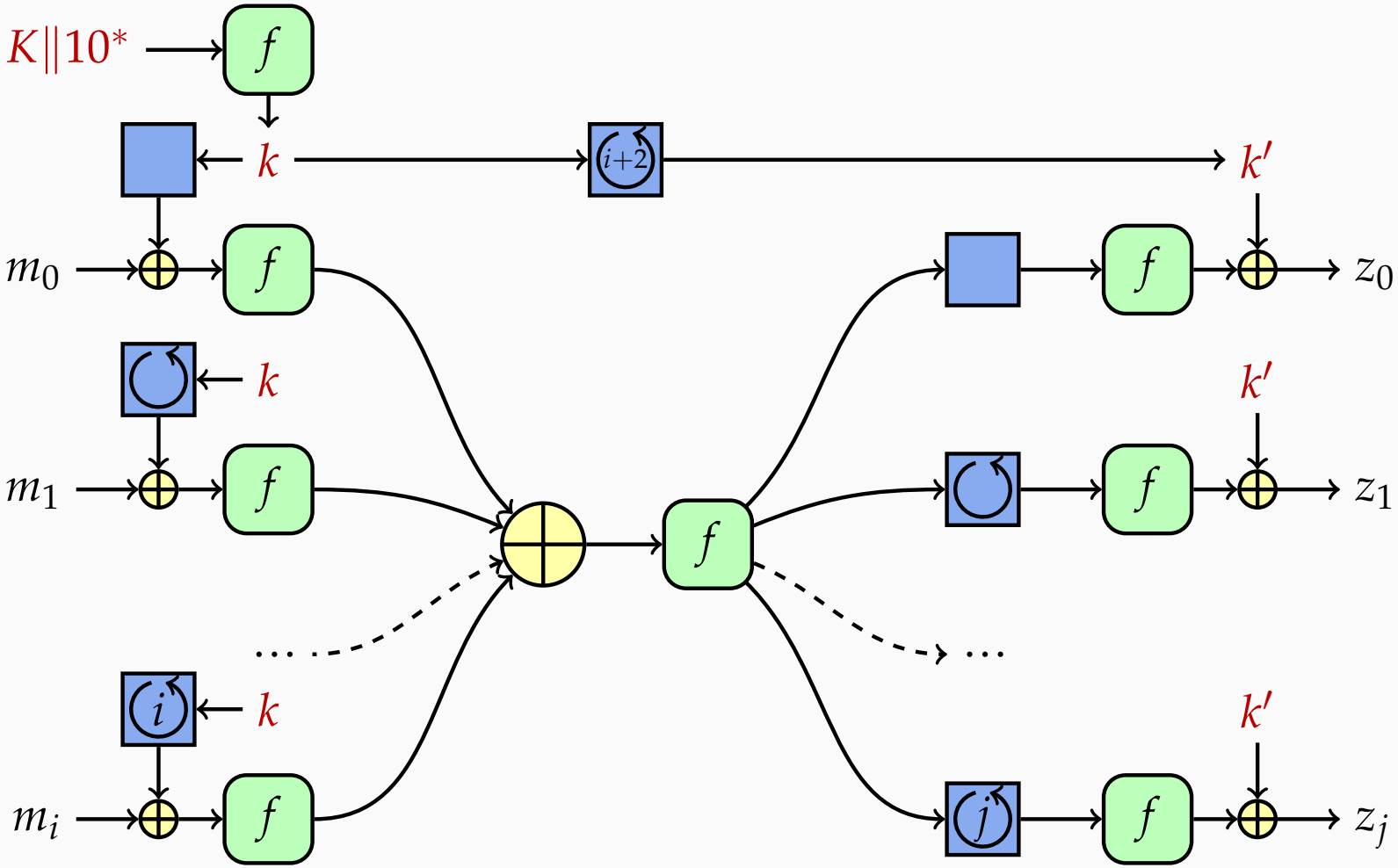
Proven secure if f is an ideal permutation

Keyed duplex



Proven secure if f is an ideal permutation

Farfalle



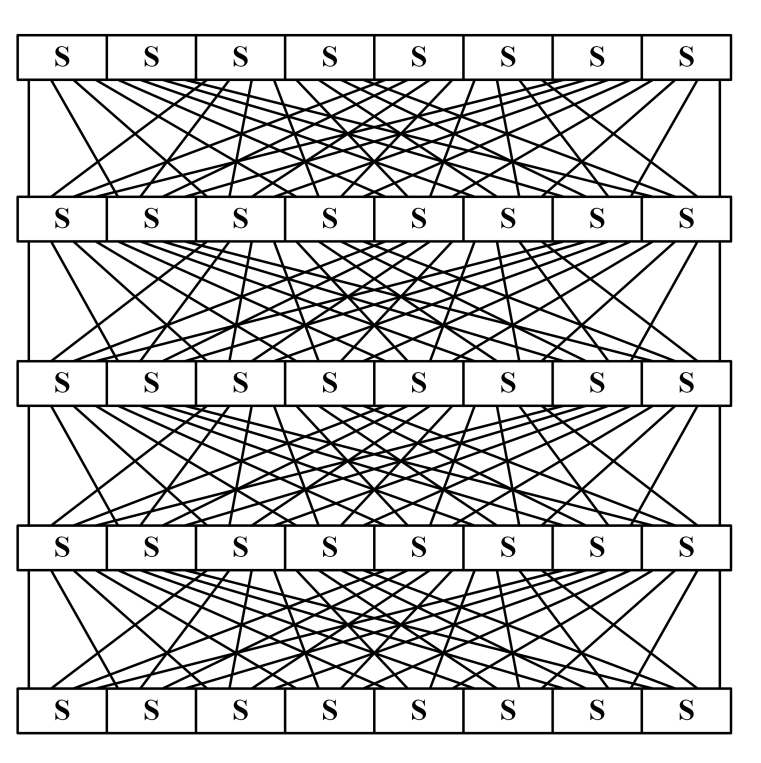
Can likely be proven secure if f is an ideal permutation

Security of these permutation-based constructions

- Build a permutation f that behaves like an ideal permutation!
- This cannot be formalized
- Assurance has to come from cryptanalytic evaluation of f
  - ...inside sponge, duplex or farfalle
- Requirements depend on the construction
- Deck functions (e.g., farfalle) are at level of block ciphers
  - PRP security of AES: distinguish $\mathrm{AES}[K]$ from random permutation
  - PRF security of XOOFFF: distinguish $\mathrm{XOOFFF}[K]$ from random oracle
- Sponge: collision-resistance, preimage resistance, some distinguishing property ...
- This is about where security reductions stop

and cryptanalysis takes over

How to build a permutation? [Claude Shannon, 1949]



Substitution-Permutation Network (SPN): round with 2 layers:

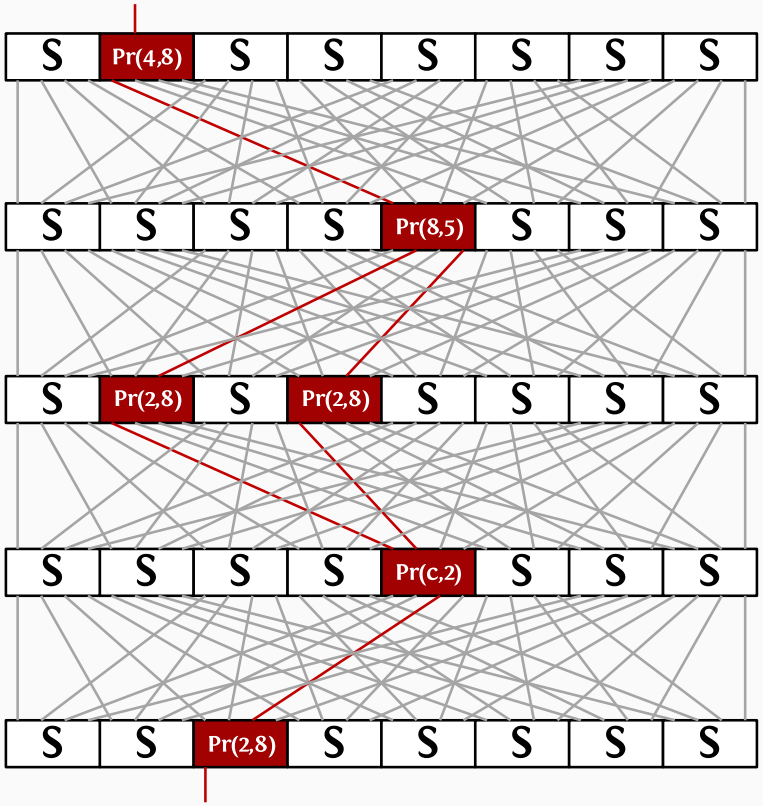
- non-linear substitution layer: S-boxes applied in parallel
- permutation layer: transposes bits to different S-box positions

More rounds gives more security

Difference propagation

- There are many attack vectors in cryptanalysis
- In this lecture: focus on difference propagation
- Relevant in
  - inner collisions: (partial) inputs leading to same state
  - rebound attacks in hashing
  - differential cryptanalysis in keyed constructions
  - ...
- Consider pairs of inputs x and x^* with $\Delta_{\text{in}} = x \oplus x^*$ and evaluate
  - $\mathrm{DP}(\Delta_{\text{in}}, \Delta_{\text{out}})$: probability that $f(x) \oplus f(x^*) = \Delta_{\text{out}}$
  - effort to find a pair that satisfies differential $(\Delta_{\text{in}} \rightarrow \Delta_{\text{out}})$

Difference propagation in an SPN



- Differences follow trails Q with some probability

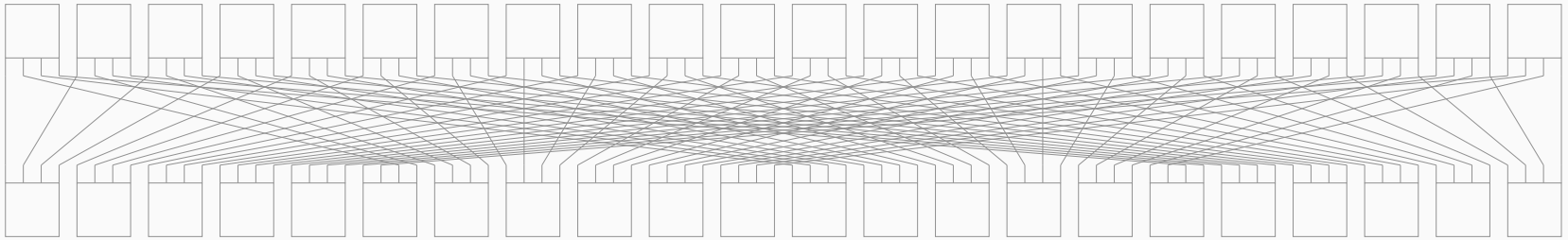
$$\mathrm{DP}(Q) \approx \prod_i \mathrm{DP}(\mathrm{Sbox}_i)$$

- different trails may lead to same difference at output:

$$\mathrm{DP}(\Delta_{\mathrm{in}}, \Delta_{\mathrm{out}}) = \sum_{\Delta_{\mathrm{in}} \to Q \to \Delta_{\mathrm{out}}} \mathrm{DP}(Q)$$

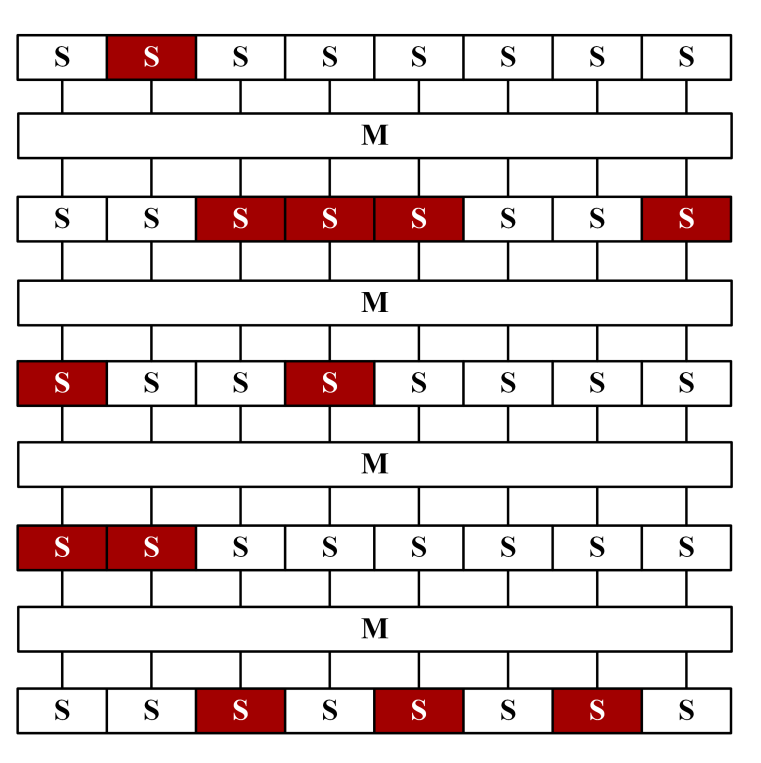
SPN approach 2011 AD: Spongent

[Bogdanov, Knežević, Leander, Toz, Varici, Verbauwhede, 2011]



- Defined for any width b that is a multiple of 4
- Disadvantages:
  - requires many rounds n_{r}: for $b = 384$, $n_{\mathrm{r}} = 195$
  - transposition layer makes it unsuited for software

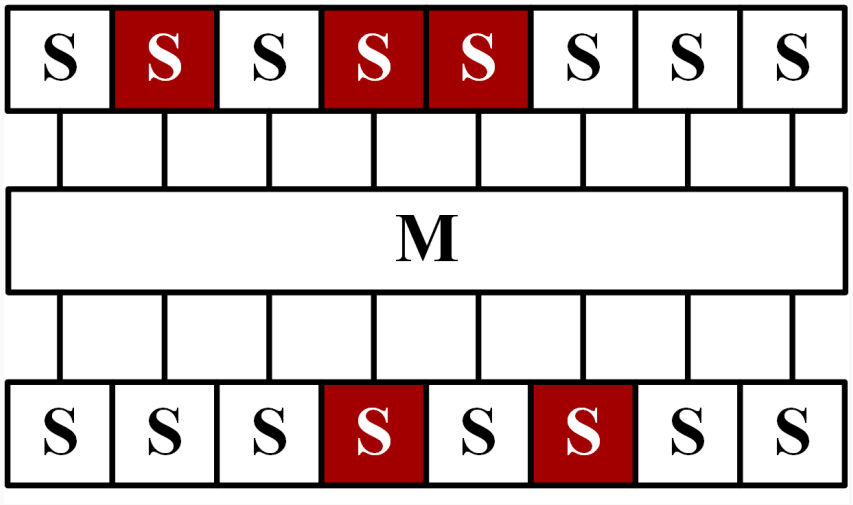
Doing better than an SPN



$$\mathrm{DP}(Q) = \prod_i \mathrm{DP}_{\mathsf{Sbox}}(a_i, b_i)$$

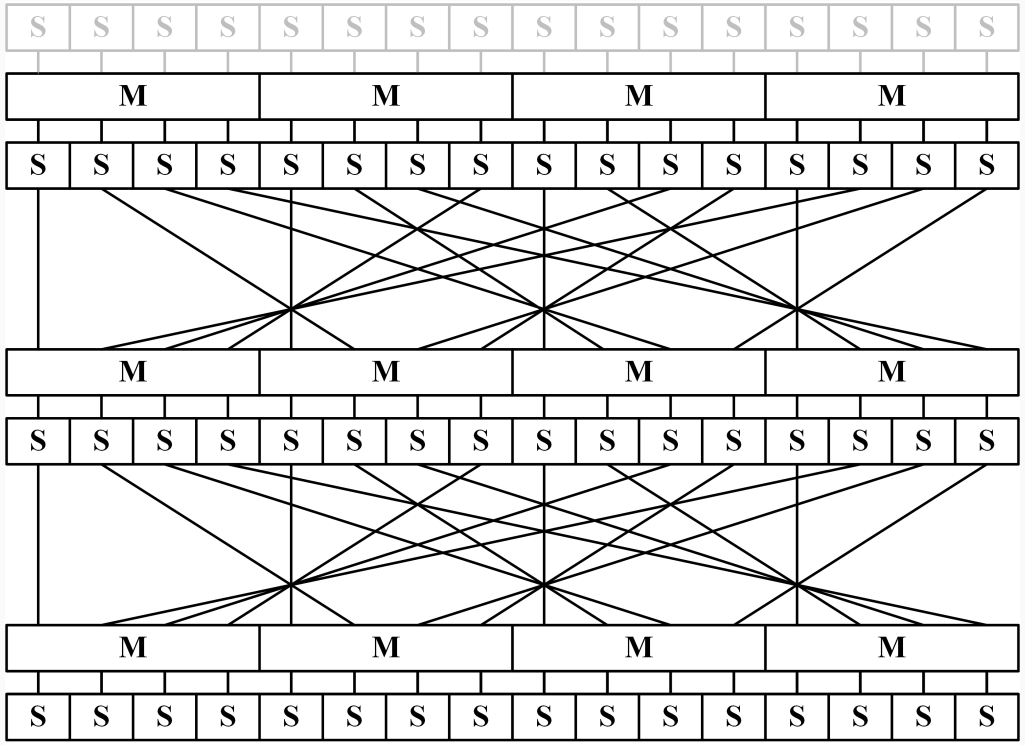
- Design goal: have no trails Q with high
- High DP if trail has few active S-boxes or S-boxes have high DP
- Wide trail strategy: *ensure all trails have many active S-boxes*

Mixing layer criterion: Branch number $\mathcal{B}$



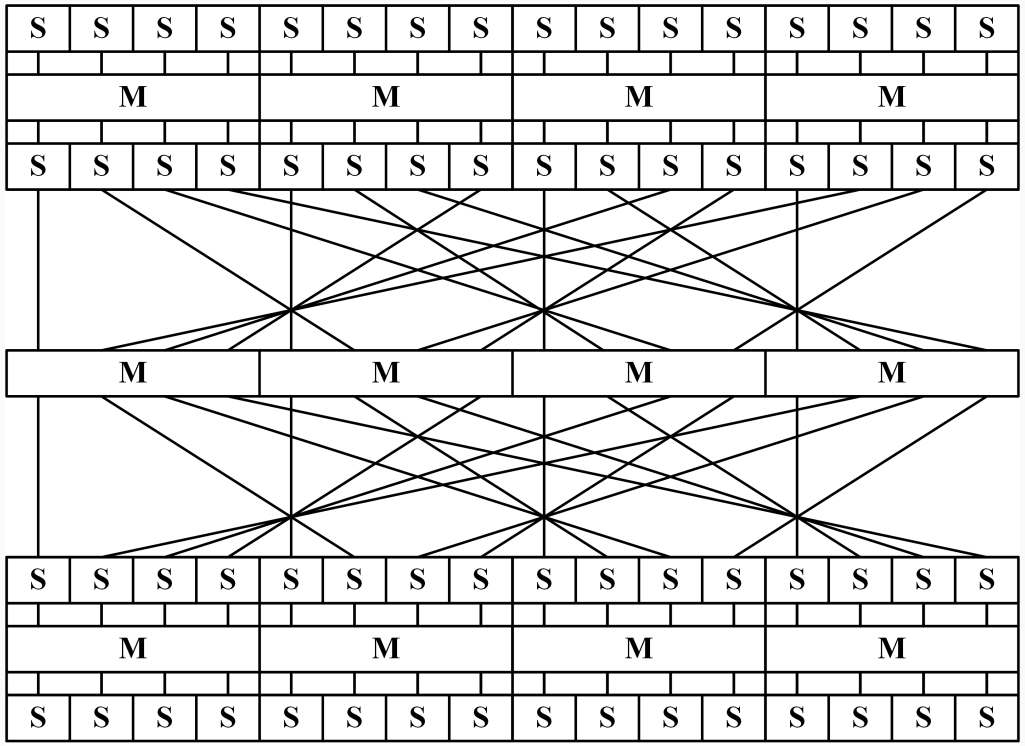
- Choose the mixing layer λ so that
  - few active S-boxes in A give many active S-boxes in $\lambda(A)$
  - few active S-boxes in B give many active S-boxes in $\lambda^{-1}(B)$
- Branch number $\mathcal{B}$: min. over A of nr. of active S-boxes in $A|\lambda(A)$
  - # of active S-boxes per two rounds is at least $\mathcal{B}$
  - $\mathcal{B} \leq n+1$ as an input can have a single active S-box
  - If $\mathcal{B} = n+1$, we call λ maximum-distance separable (MDS)

Rijndael (AES) [Daemen, Rijmen 1998]



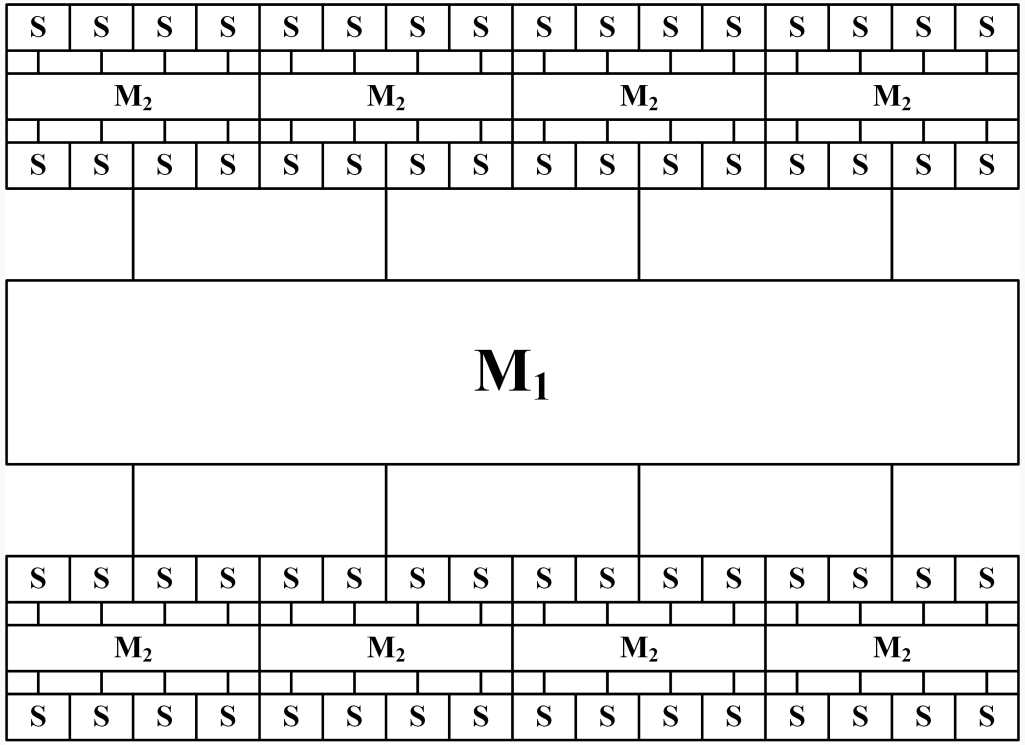
- Strong alignment: operates on bytes instead of bits
- `MixColumns` matrix M is MDS: branch number 5

Rijndael (cont'd)



`ShiftRows` and `SubBytes` commute

Rijndael (some more)



- Recursive structure with 8 **super boxes**
- # active super boxes ≥ 5 so # active S-boxes ≥ 25
- 8-bit S-box with $\mathrm{DP} \leq 2^{-6}$, so for 4R trails $\mathrm{DP} \leq 2^{-6\times 25} = 2^{-150}$

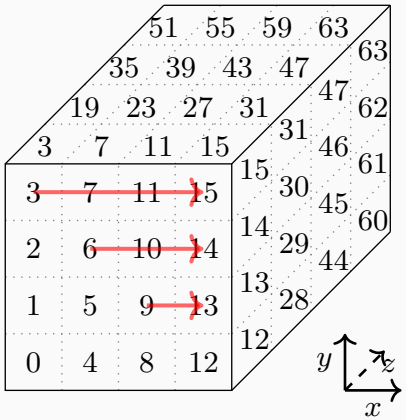
Disadvantages of Rijndael

- Rijndael was software-oriented
  - T-tables: 1 TLU and 32-bit XOR per byte per round
  - for 8-bit CPU: similar but more XORs and smaller tables
- Performance independent of S-box specifics: we chose the best one known
- We did choose `MixColumns` matrix with 8-bit CPU in mind
- Problem: timing attacks based on cache misses
- Countermeasure: dedicated hardware [AES-NI, Intel] or bitsliced software [Käsper, Schwabe 2009]
- *Gate cost*: # binary operations per bit per round: 16 XOR and 4 AND

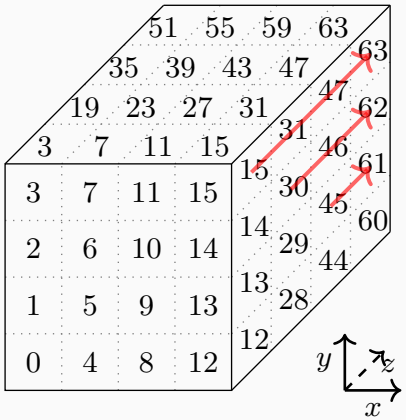
Strongly aligned approach 2019 AD: Saturnin

[Canteaut, Duval, Leurent, Naya-Plasencia, Perrin, Pornin, Schrottenloher]

- Block cipher with 256-bit block length submitted to NIST lightweight
- Gate cost only 3.875 XOR and 1.5 AND/OR
  - 4-bit S-box layer: 1.5 XOR and 1.5 AND/OR
  - MC matrix MDS $\mathcal{B} = 5$ with cost 2.375 XOR
- AES square becomes $4 \times 4 \times 4$ cube



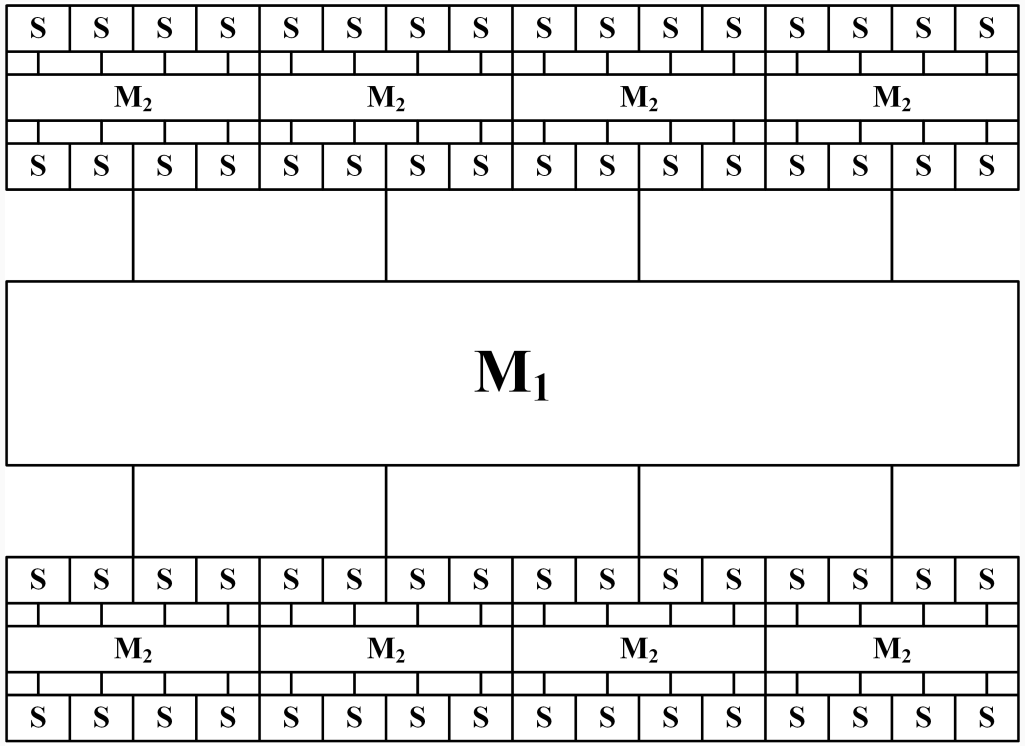
(a) $\mathrm{SR}_{\mathrm{slice}}$ (when $r \equiv 1 \bmod 4$)



(b) $\mathrm{SR}_{\mathrm{sheet}}$ (when $r \equiv 3 \bmod 4$)

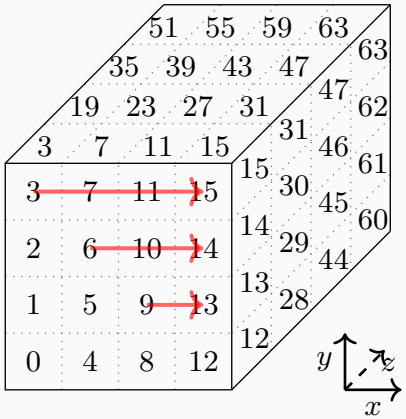
figure courtesy of Saturnin team

Saturnin

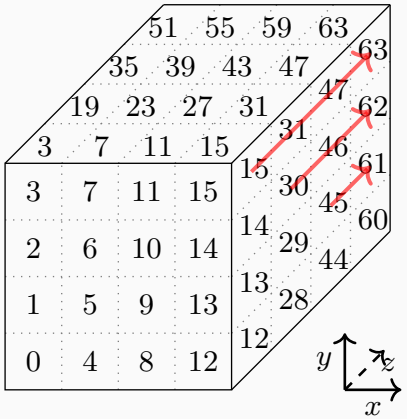


- Recursive structure with 64-bit *mega boxes*
- Mega box has 16-bit super boxes, that have 4-bit S-boxes
- # active S-boxes is 5^3 and the S-boxes have $\mathrm{DP} \leq 2^{-2}$
- 8-round trails have $\mathrm{DP} \leq 2^{-250}$

Disadvantages of Saturnin: ShiftRows



(a) $\mathsf{SR}_{\mathrm{slice}}$ (when $r \equiv 1 \bmod 4$)



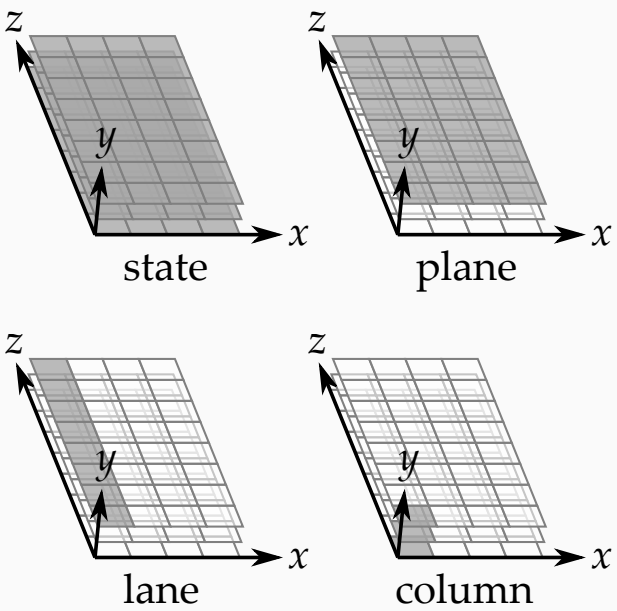
(b) $\mathsf{SR}_{\mathrm{sheet}}$ (when $r \equiv 3 \bmod 4$)

- There are three transposition mappings:
  - Identity in even-indexed rounds
  - $\mathrm{SR}_{\mathsf{slice}}$ if index is 1 modulo 4
  - $\mathrm{SR}_{\mathsf{sheet}}$ if index is 3 modulo 4
- Hardware: gives hassle in single-round combinatorial logic
- Not so efficient in software, e.g., on ARM Cortex M3
  - $\mathrm{SR}_{\mathsf{sheet}}$ costs more than MC step
  - $\mathrm{SR}_{\mathsf{slice}}$ costs more than MC + S-box layer

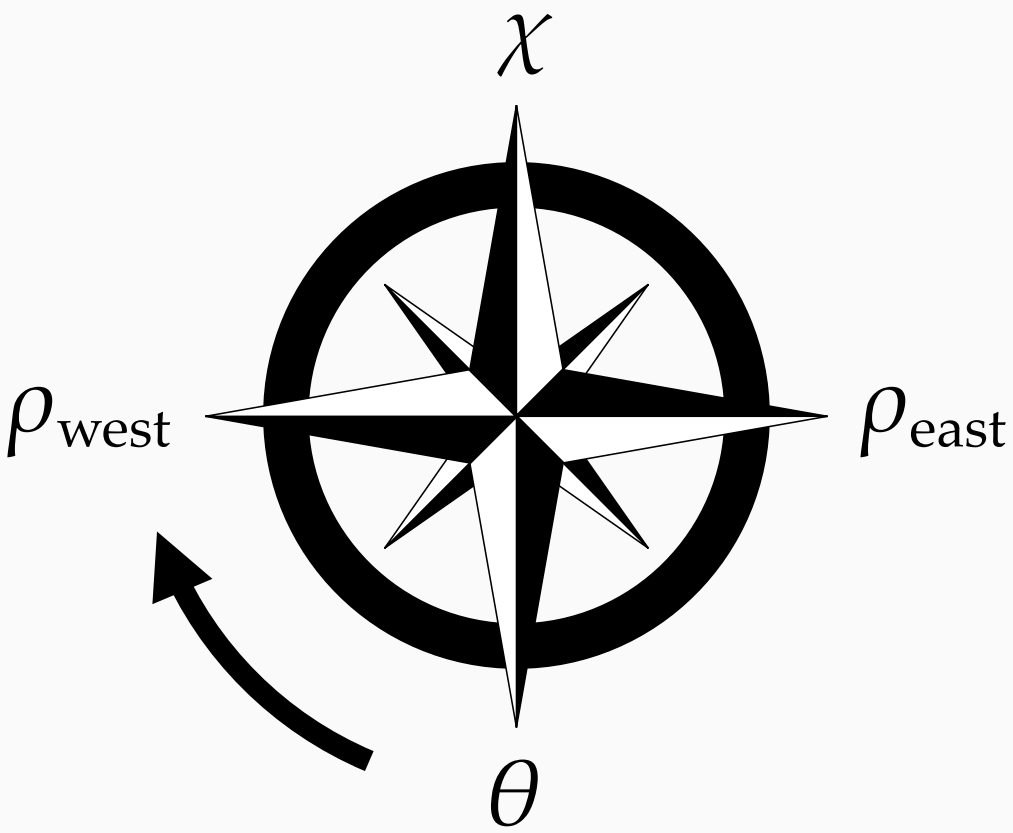
Weakly aligned approach 2017 AD: Xoodoo

[Bertoni, Daemen, Hoffert, Peeters, Van Assche, Van Keer, ToSC 2018]

- 384-bit permutation
- Size and *shape* inspired by Gimli [Bernstein, Kölbl, Lucks, Massolino, Mendel, Nawaz, Schneider, Schwabe, Standaert, Todo, Viguier, 2017]
- Design approach: that of the permutation in Keccak
- **Weak alignment**: operates on bits rather than bytes (or nibbles)



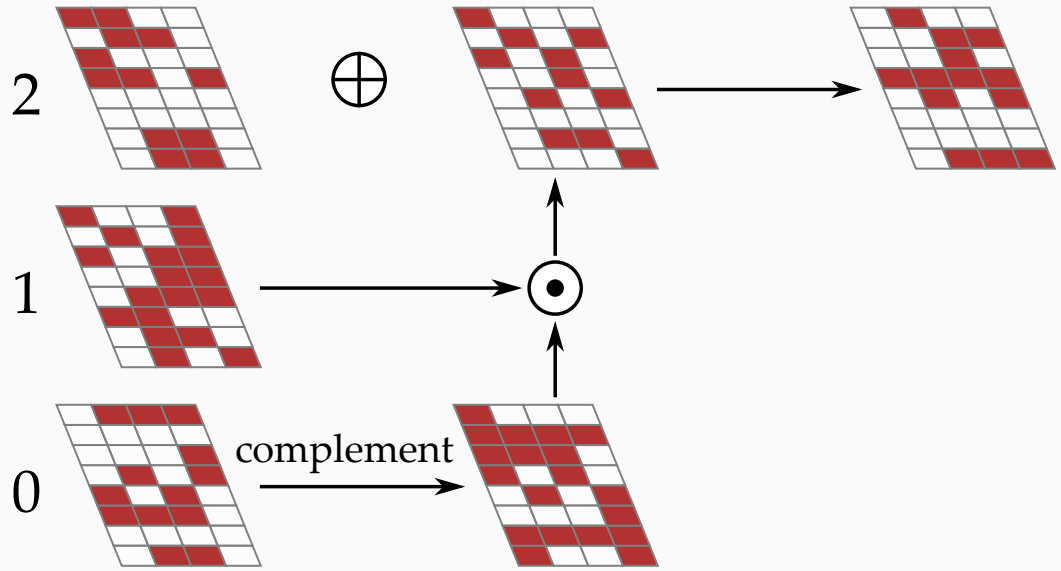
Xoodoo round function



Iterated: n_{r} rounds that differ only by round constant

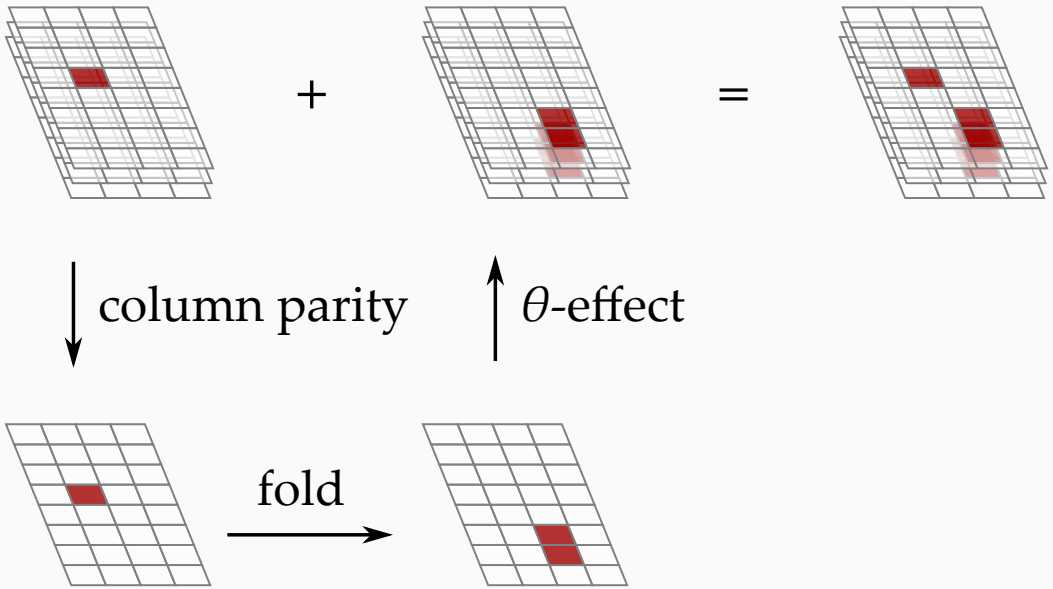
Nonlinear mapping χ

Effect on one plane:



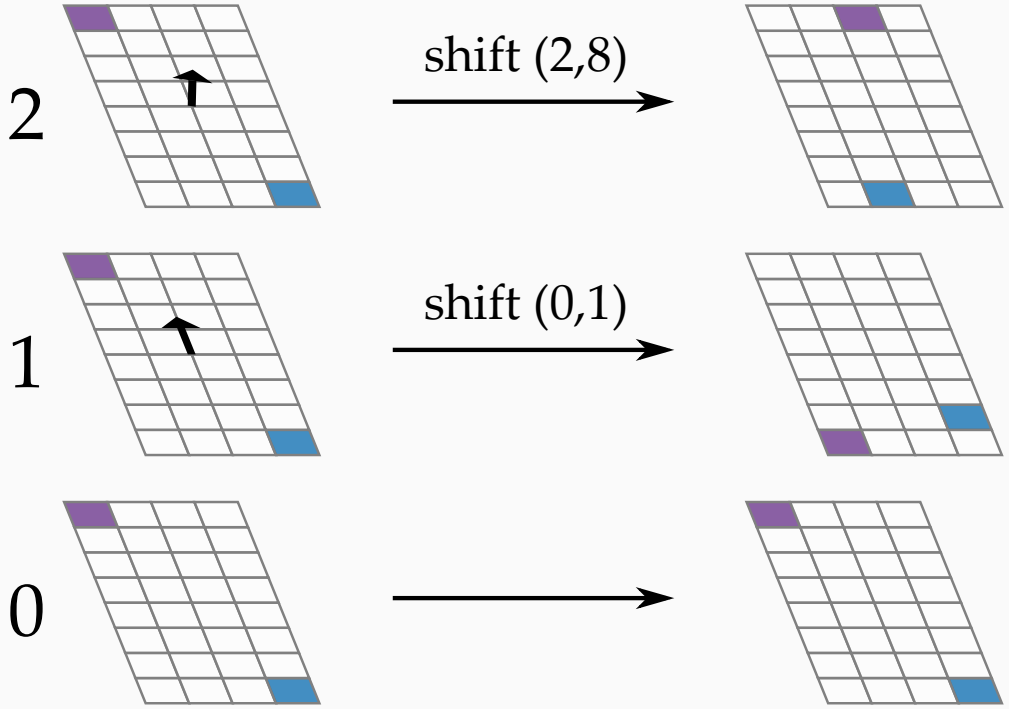
- χ as in KECCAK-p, operating on 3-bit columns
- Involution and same propagation differentially and linearly

Mixing layer θ



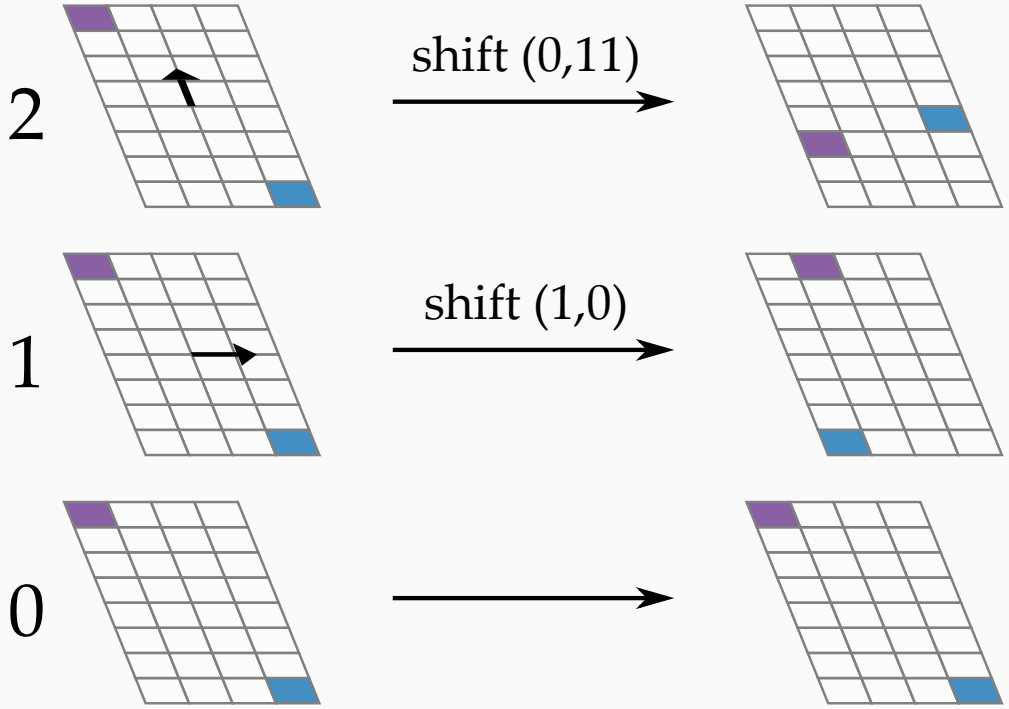
- Column parity mixer: compute parity, fold and add to state
- good average diffusion, identity for states in *kernel*
- heavy inverse

Plane shift ρ_{east}



- After χ and before θ
- Shifts planes $y = 1$ and $y = 2$ over different directions

Plane shift ρ_{west}



- After θ and before χ
- Shifts planes $y = 1$ and $y = 2$ over different directions

Xoodoo implementation aspects

- One single round function
- Gate cost: 3 XOR and 1 AND
  - 5 Xoodoo rounds cost less than 4 Saturnin rounds
- Transposition layer software-friendly, e.g., on ARM Cortex M3
  - ρ steps have zero overhead
  - 7 Xoodoo rounds cost less than 4 Saturnin rounds

Disadvantage of Xoodoo

- No simple proofs for trail DP values
- Because it does not have super-boxes
- Instead computer-assisted bounds
  - Scanning the space of trails with DP below some limit
  - Using techniques of [Mella, Daemen, Van Assche, ToSC 2016]

Comparison of minimum trail weights $\mathrm{w}(Q) = -\log_2 \mathrm{DP}(Q)$

# rounds:	1	2	3	4	6	8	10
Spongent	2	4	8	12	28	?	?
Rijndael	6	30	54	150	≥ 180	≥ 300	≥ 330
Saturnin	2	10	18	50	210	250	260
Xoodoo	2	8	36	≥ 74	≥ 104	≥ 148	≥ 180

A closer look at difference propagation

What we compare

- Permutations we study have a round function with two layers
  - non-linear S-box layer
  - linear layer, that we will denote as λ
- We don't consider other design approaches
  - ARX such as in ChaCha, ARX-like such as in NORX
  - no S-box layer, as Gimli, Subterranean or Friet
  - Feistel networks, ...
- These are harder to describe and compare, so *future work*
- We will limit our comparison to the following

	alignment	diffusion	S-box size
Spongent	weak	weak	4
Rijndael	strong	strong	8
Saturnin	strong	strong	4
Xoodoo	weak	strong	3

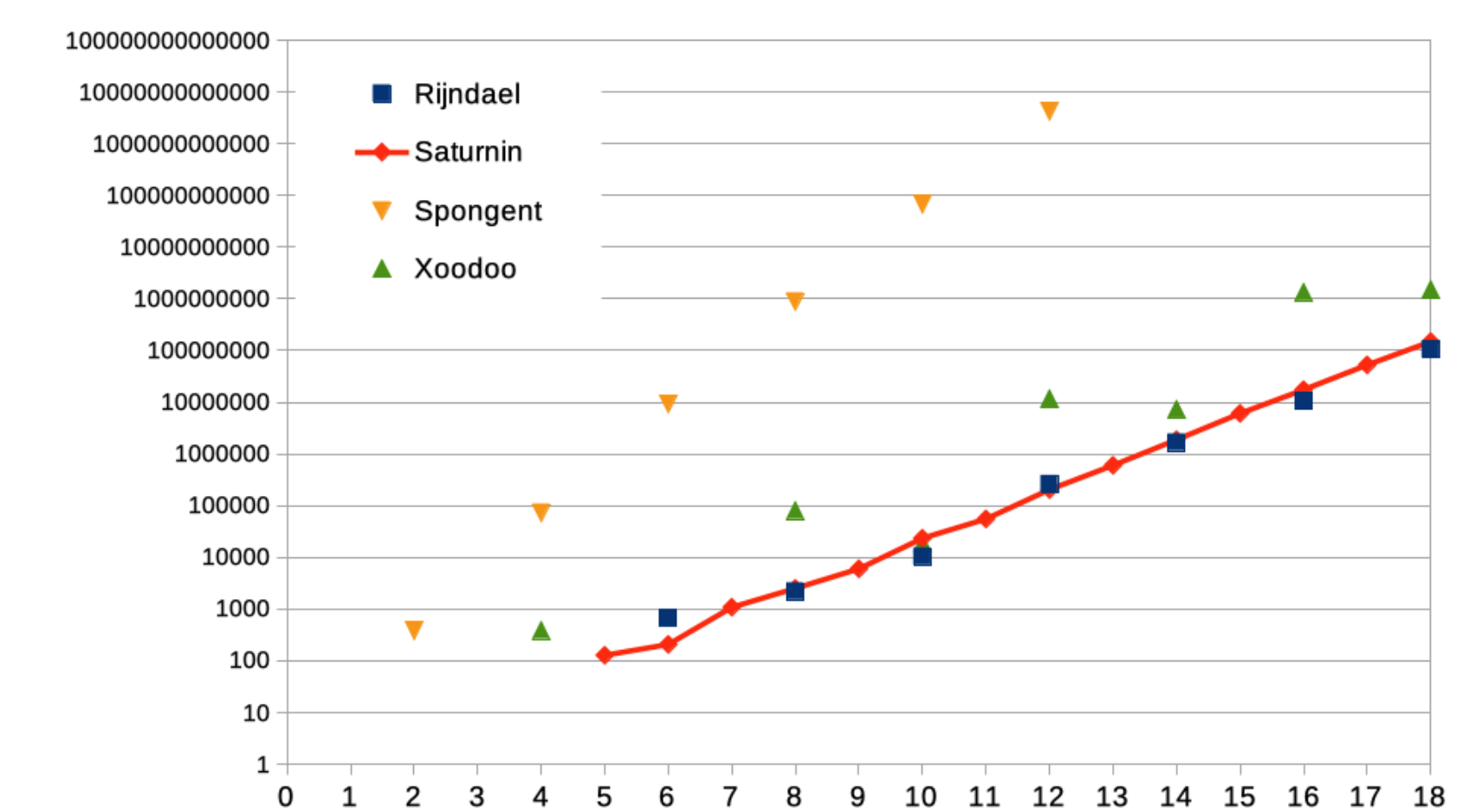
Characterizing diffusion of the linear layer

- Currently mostly considered: branch number $\mathcal{B}$

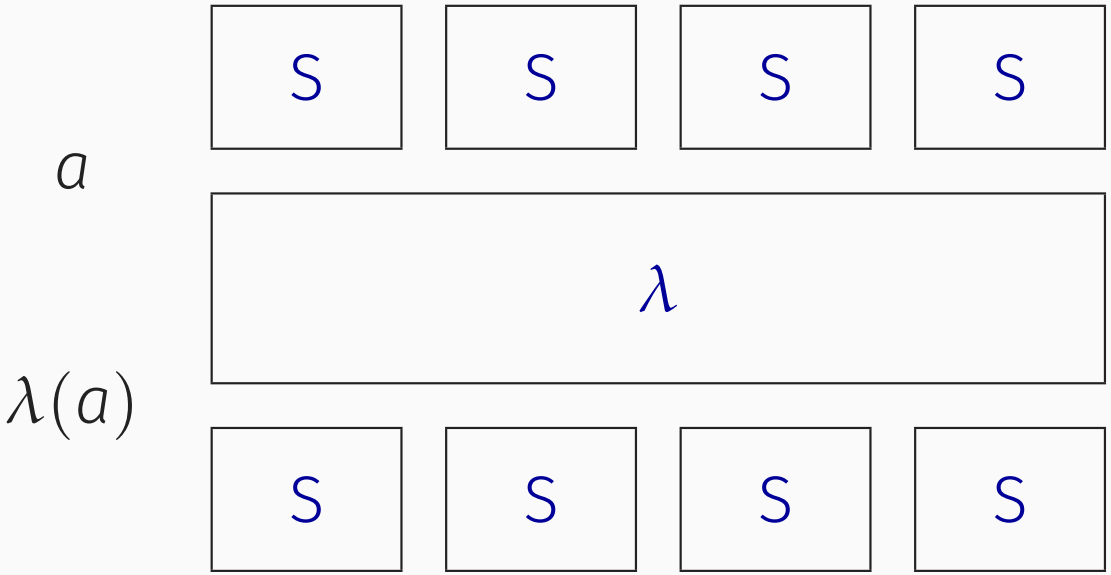
$$\mathcal{B} = \min_{a \neq 0} (\mathrm{w}(a) + \mathrm{w}(\lambda(a)))$$

  with $\mathrm{w}(x)$ the *box weight*: # active S-boxes in x
- More informative: histogram of # states per weight of $(a, \lambda(a))$
  - for all 2^b states a compute $\mathrm{w}(a) + \mathrm{w}(\lambda(a))$
  - list them in a histogram
  - the tail of the histogram at the low-weight end says something about diffusion
- Hamming weight histogram: absolute diffusion power independent of S-box layer
- Box weight histogram: the 2-round differential trails with given # active S-boxes

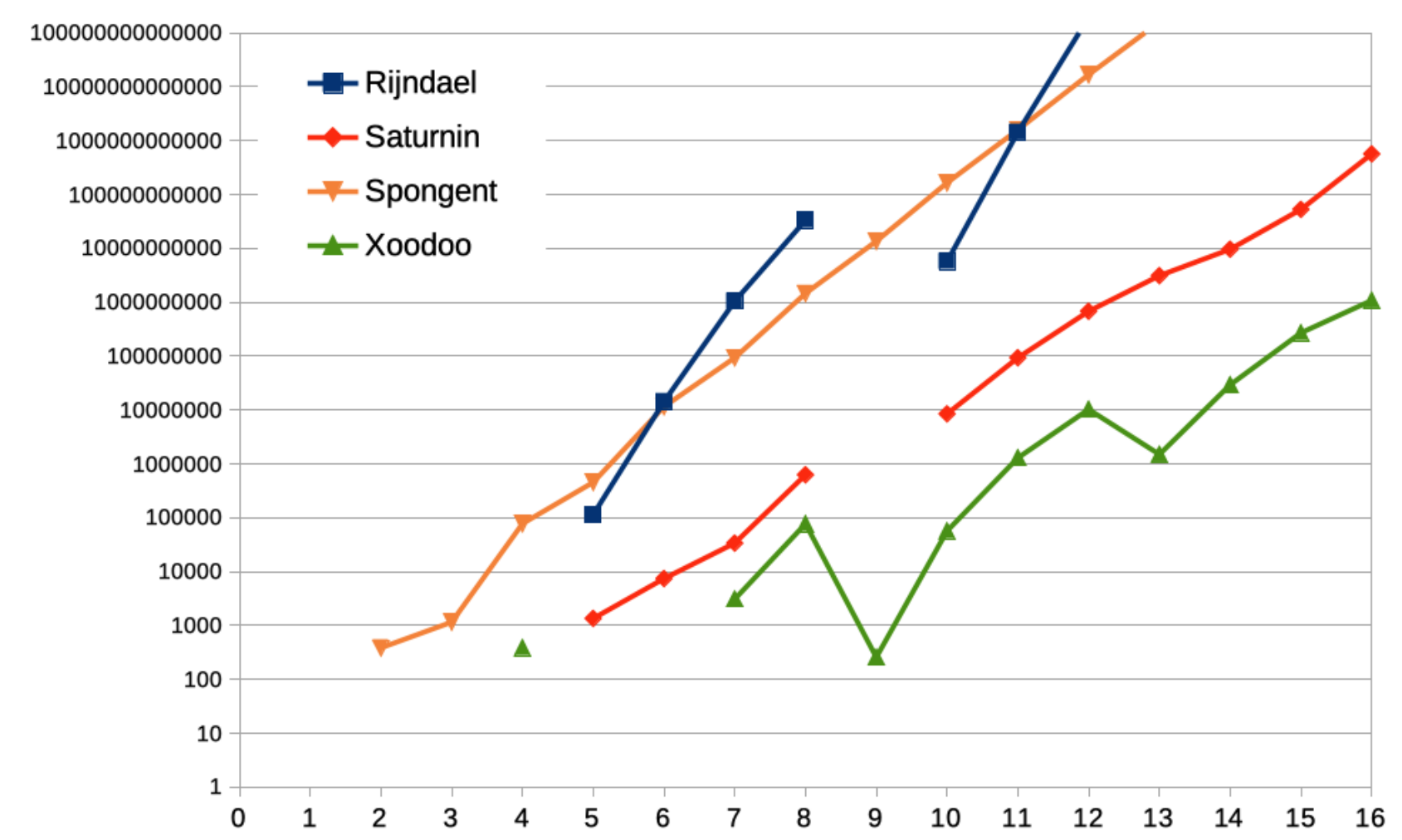
Hamming weight histograms



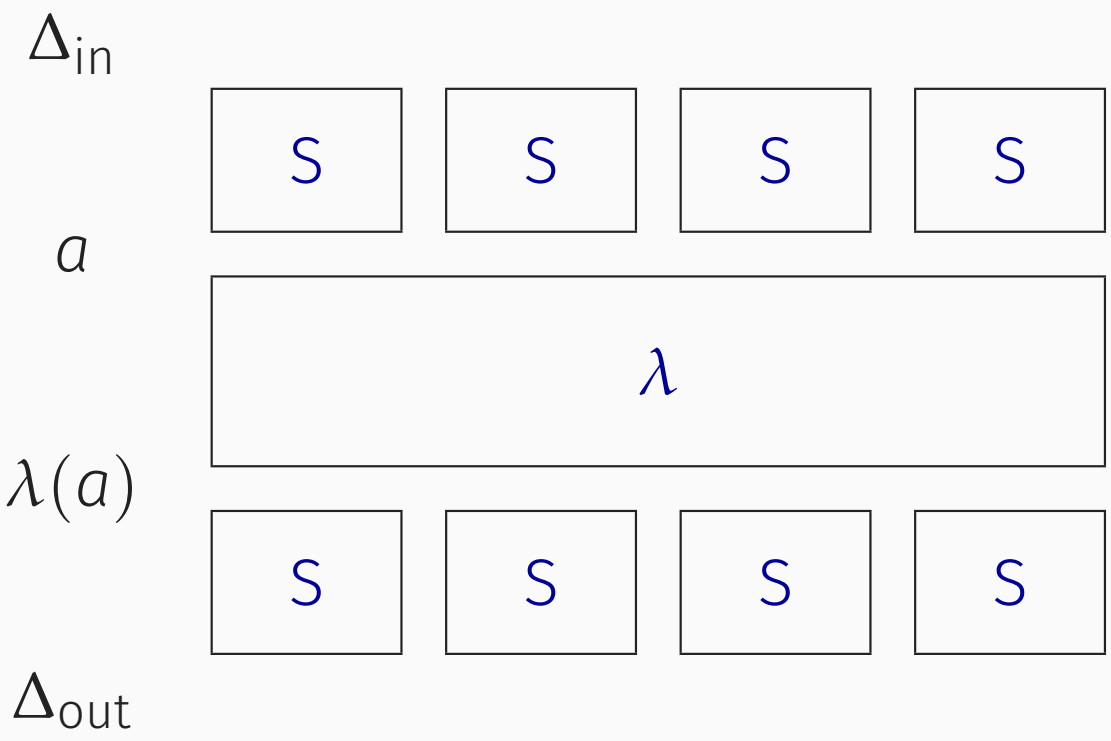
From Hamming weight to box weight



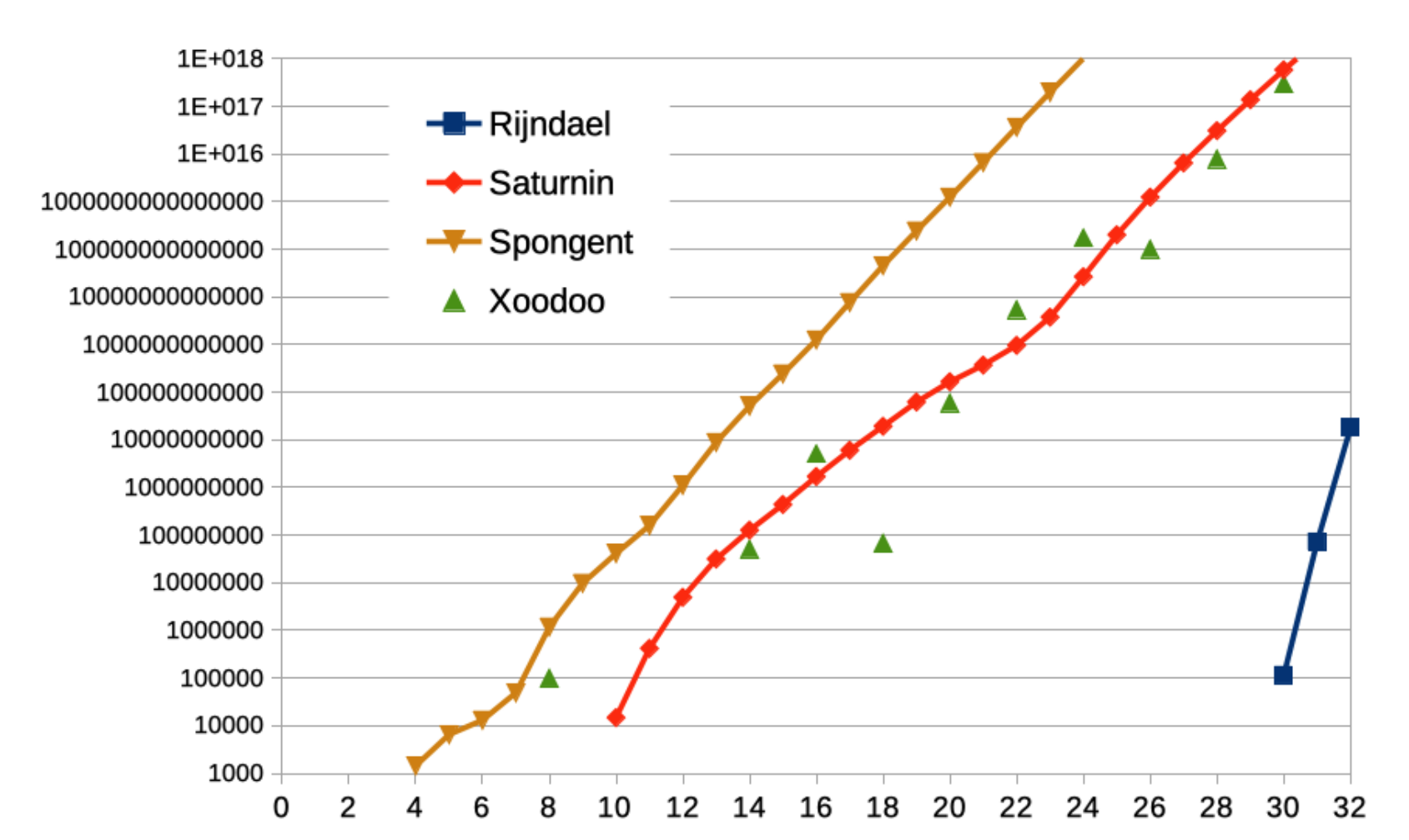
Box weight histogram



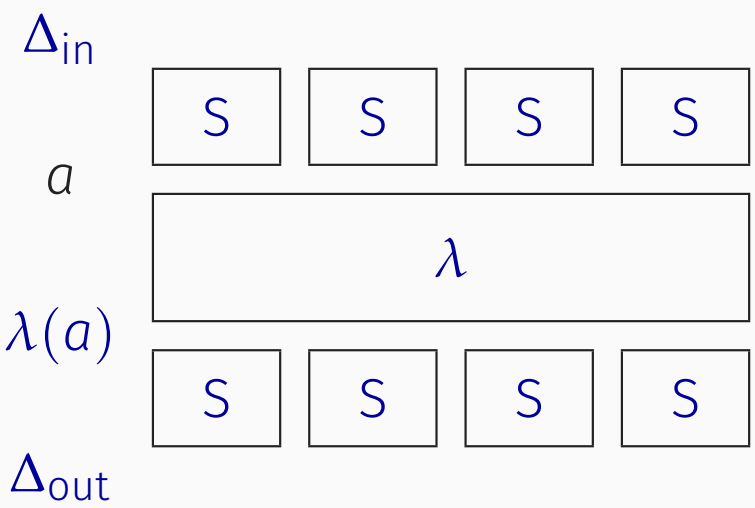
From box weight to trail weight



Trail weight histogram



λ-box partition



- Trail: sequence of states (differences) $(\Delta_{\text{in}}, a, \lambda(a), \Delta_{\text{out}})$
- *box pattern*: set of states with same pattern of passive boxes
- In the trail above
  - Δ_{in} and a are in same box pattern
  - Δ_{out} and $\lambda(a)$ are in same box pattern
- So differential $(\Delta_{\text{in}}, \Delta_{\text{out}})$ fixes box pattern of a and $\lambda(a)$
- **λ-box partition** of states a by *box pattern* of $(a, \lambda(a))$
- Size of these subsets bound the number of trails in a differential

λ-box partition: Rijndael and Saturnin

Rijndael super box:

box weight	# subsets $\times$ size
5	(56×255)
6	(28×64005)
7	(8×16323825)
8	(1×4162570275)

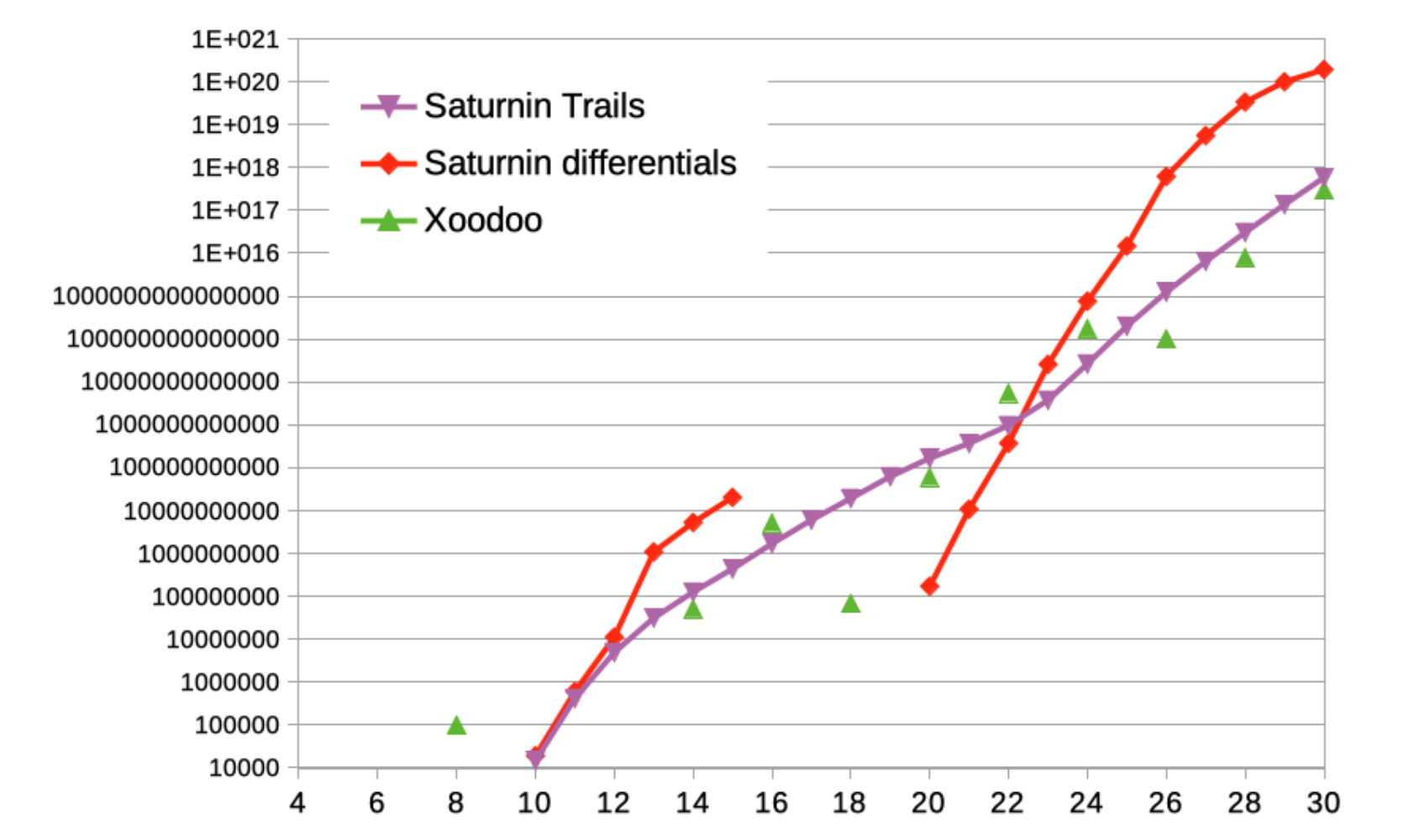
Quasi all differentials have multiple trails [Daemen Rijmen, 2008]

Saturnin super box:

box weight	# subsets $\times$ size
5	(56×15)
6	(28×165)
7	(8×2625)
8	(1×39075)

There is significant clustering of trails in differentials

Differential weight histogram



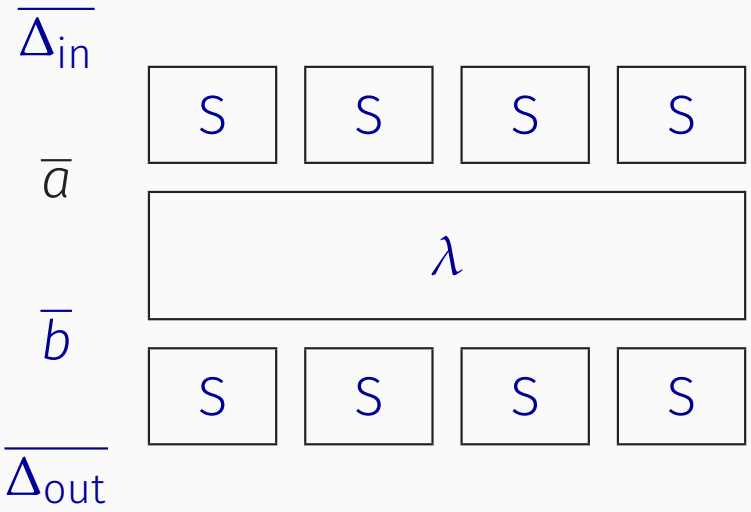
λ-box partition of Xoodoo (modulo horizontal translation)

box weight	# subsets $\times$ size	# subsets $\times$ size
4	(3×1)	
7	(24×1)	
8	(600×1)	
9	(2×1)	
10	(442×1)	
11	(10062×1)	
12	(80218×1)	
13	(11676×1)	
14	(228531×1)	(3×2)
15	(2107864×1)	(90×2)
16	(8447176×1)	(702×2)
17	...	

Differentials with more than a single trail are rare

For 3 rounds: all trails $\mathrm{w}(Q) < 50$ are alone in their differential

Truncated differentials



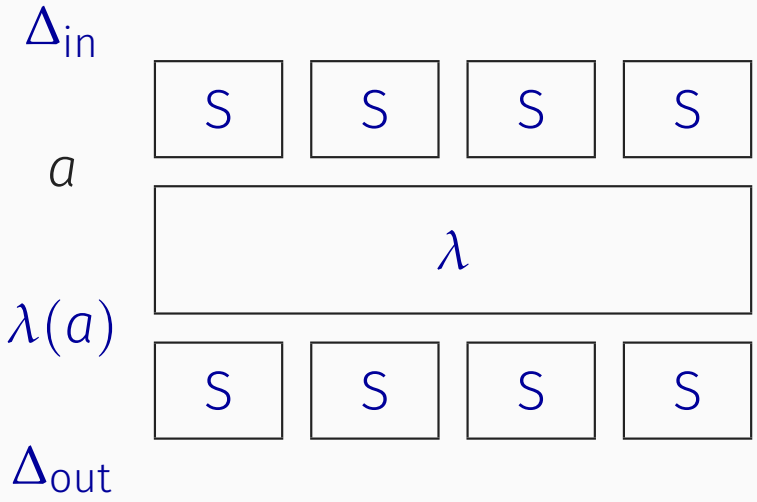
- *Truncated trail*: sequence of **box patterns** $(\overline{\Delta_{\text{in}}}, \overline{a}, \overline{b}, \overline{\Delta_{\text{out}}})$

$$\mathrm{DP}(\overline{\Delta_{\text{in}}}, \overline{\Delta_{\text{out}}}) = \Pr(y \in \overline{\Delta_{\text{out}}} | x \in \overline{\Delta_{\text{in}}})$$

- S-box layers have probability 1 iff $\overline{\Delta_{\text{in}}} = \overline{a}$ and $\overline{b} = \overline{\Delta_{\text{out}}}$, so

$$\mathrm{DP}(\overline{\Delta_{\text{in}}}, \overline{\Delta_{\text{out}}}) = \Pr(\lambda(a) \in \overline{\Delta_{\text{out}}} | a \in \overline{\Delta_{\text{in}}})$$

Truncated differentials in Rijndael and Saturnin



- In Rijndael super box: $\mathrm{DP}(\overline{\Delta_{\text{in}}}, \overline{\Delta_{\text{out}}}) \approx 255^{\mathrm{w}(\overline{\Delta_{\text{out}}})-4}$
  - AES has 4-round super box truncated diff. with $\mathrm{DP} = 1$
  - Saturnin has 8-round mega box truncated diff. with $\mathrm{DP} = 1$
- These can be used as distinguisher in attacks
- Can also be exploited in other attacks, e.g.,
  - impossible differentials (most powerful AES attack)
  - rebound attacks in hashing

Truncated differentials in Xoodoo

- In Xoodoo truncated trails have higher weight than trails

that's it

Please make up your own conclusions

Thanks for listening!

European Research Council

Established by the European Commission

ESCADA

Part of this work was sponsored by ERC grant ESCADA