Exploiting algebraic properties of block ciphers

Anne Canteaut

Exploiting algebraic properties of block ciphers

Anne Canteaut

Anne.Canteaut@inria.fr

https://www.paris.inria.fr/secret/Anne.Canteaut/

COST Training School, Torremolinos, February 2018

Random behaviour of cryptographic primitives

Cryptographic primitives should behave like random functions.

Security proofs of many constructions assume random building blocks

- modes of operation;
- sponge construction...

A distinguishing property may lead to some attacks
e.g., finding the plaintext among a few possibilities.

For iterated constructions
a distinguisher on a round-reduced version may be exploited for recovering the key (e.g., last-round attacks).

Algebraic Normal Form

Algebraic Normal Form of $f : \mathbf{F}_2^n \to \mathbf{F}_2$:
unique polynomial representation in $\mathbf{F}_2[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$.

Monomials of n variables

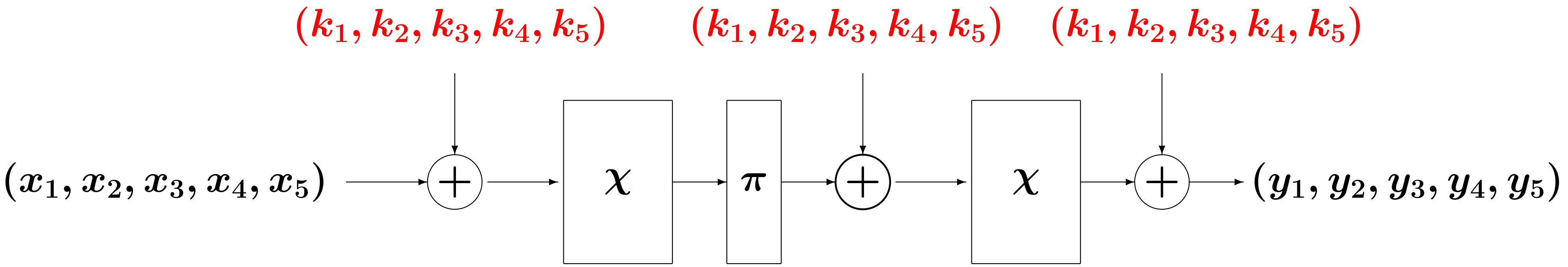
$$\prod_{i=1}^{n} x_i^{u_i} = x^u \text{ with } u \in \mathbf{F}_2^n$$

Example:
For $u = (0101)$,

$$x^u = x_4^0 x_3^1 x_2^0 x_1^1 = x_3 x_1$$

$$f(x_1, \ldots, x_n) = \sum_{u \in \mathbf{F}_2^n} c_u x^u \text{ with } c_u \in \mathbf{F}_2$$

Example



$$
\begin{aligned}
y_1 \quad = \quad & k_2k_3k_4k_5 + k_2k_3k_4x_5 + k_2k_3k_5x_4 + k_2k_3x_4x_5 + k_2k_4k_5x_3 + k_2k_4x_3x_5 + k_2k_5x_3x_4 \\
& + k_2x_3x_4x_5 + k_3k_4k_5x_2 + k_3k_4x_2x_5 + k_3k_5x_2x_4 + k_3x_2x_4x_5 + k_4k_5x_2x_3 + k_4x_2x_3x_5 \\
& + k_5x_2x_3x_4 + x_2x_3x_4x_5 + k_2k_3k_4 + k_2k_3k_5 + k_2k_3x_4 + k_2k_3x_5 + k_2k_4k_5 + k_2k_4x_5 \\
& + k_2k_5x_3 + k_2k_5x_4 + k_2x_3x_5 + k_2x_4x_5 + k_3k_4k_5 + k_3k_4x_2 + k_3k_4x_5 + k_3k_5x_4 + k_3x_2x_4 \\
& + k_3x_4x_5 + k_4k_5x_2 + k_4k_5x_3 + k_4x_2x_5 + k_4x_3x_5 + k_5x_2x_4 + k_5x_3x_4 + x_2x_4x_5 + x_3x_4x_5 \\
& + k_1k_3 + k_1x_3 + k_2k_3 + k_2k_5 + k_2x_5 + k_3k_5 + k_3x_1 + k_3x_3 + k_3x_5 + k_4k_5 \\
& + k_4x_2 + k_4x_3 + k_4x_5 + k_5x_3 + k_5x_4 + x_1x_3 + x_2x_4 + x_3x_4 + x_3x_5 + x_4x_5 \\
& + k_2 + k_3 + k_5 + x_2 + x_3 + x_5
\end{aligned}
$$

ANF of a random function

Uniform distribution over all functions:
equivalent to the uniform distribution over all ANFs.

$\longrightarrow$ each monomial appears with probability $\frac{1}{2}$.

Uniform distribution over all permutations:
open problem.

- all coordinates of a permutation of $\mathbf{F}_2^n$ have degree at most $(n-1)$.
- almost all permutations of $\mathbf{F}_2^n$ have degree $(n-1)$ [Wells 69], [Das 02], [Konyagin-Pappalardi 02]

Higher-order differential attacks

Higher-order differentials [Lai 94]

Differential (derivative) of f w.r.t. $\alpha \in \mathbb{F}_2^n$:

$$D_\alpha f : x \mapsto f(x \oplus \alpha) \oplus f(x)$$

Higher-order differential.

$$D_{\alpha_1} D_{\alpha_2} \ldots D_{\alpha_d} f(x) = \bigoplus_{v \in \langle \alpha_1, \ldots, \alpha_d \rangle} f(x \oplus v) := D_{\langle \alpha_1, \ldots, \alpha_d \rangle} f(x)$$

$$\deg D_V f \leq \deg f - \dim V$$

Then, for any subspace V with $\dim V > \deg f$

$$\forall a \in \mathbb{F}_2^n, \; D_V f(a) = \bigoplus_{v \in V} f(a \oplus v) = 0$$

Higher-order differential attacks [Knudsen 94]

for all keys, f_k has a low degree in x

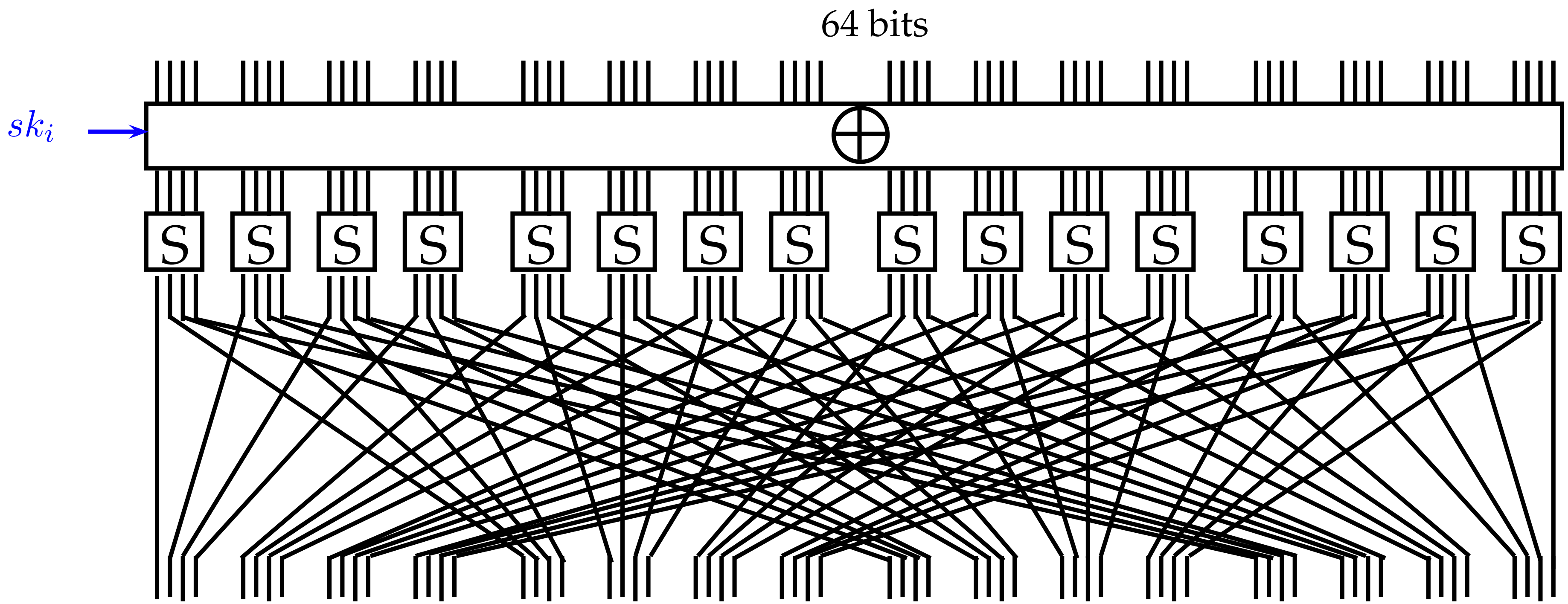
For any affine subspace $a + V$ with $\dim V > \deg_x f$

$$\bigoplus_{x \in a+V} f_k(x) = 0$$

Distinguisher with data complexity proportionnal to $2^{\deg_x f}$

Degree of an iterated block cipher

PRESENT [Bogdanov et al. 07]



31 rounds (+ a key addition)

Algebraic degree of PRESENT with a fixed key

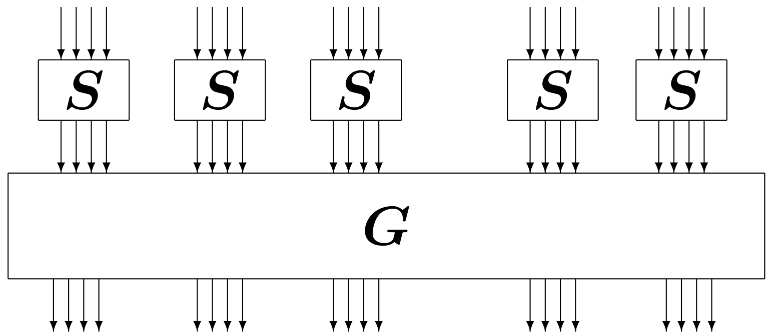
$\deg S = 3$

After r rounds,

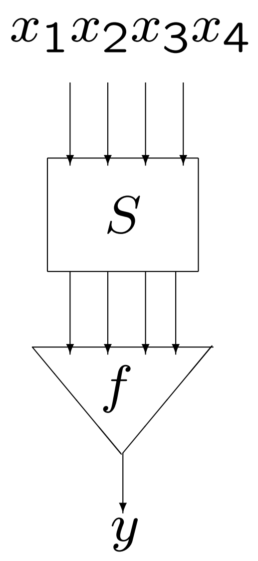
$$\deg_x E^r \leq 3^r$$

r	1	2	3	4	5	6	7
deg	3	9	27	63	63	63	63

Using the particular form of the Sbox layer



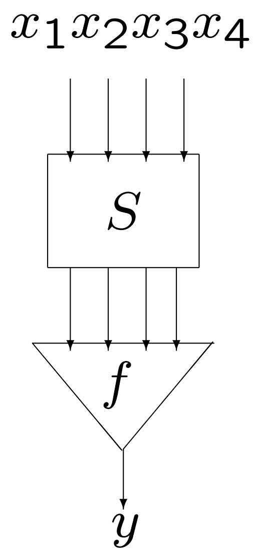
Exercise: find the maximal degree of the following 4-bit function



with $f(y_1, y_2, y_3, y_4) = y_1y_2$.

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	f	e	b	c	6	d	7	8	0	3	9	a	4	2	1	5
$S_1(x)$	1	0	1	0	0	1	1	0	0	1	1	0	0	0	1	1
$S_2(x)$	1	1	1	0	1	0	1	0	0	1	0	1	0	1	0	0
$S_3(x)$	1	1	0	1	1	1	1	0	0	0	0	0	1	0	0	1
$S_4(x)$	1	1	1	1	0	1	0	1	0	0	1	1	0	0	0	0

Exercise: find the maximal degree of the following 4-bit function

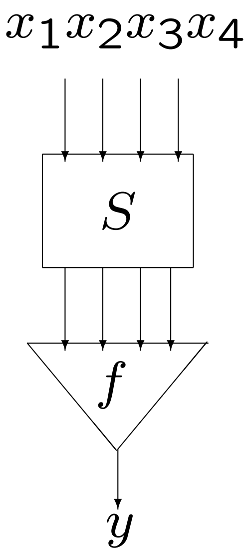


with $f(y_1, y_2, y_3, y_4) = y_1 y_2$.

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	f	e	b	c	6	d	7	8	0	3	9	a	4	2	1	5
$S_1(x)$	1	0	1	0	0	1	1	0	0	1	1	0	0	0	1	1
$S_2(x)$	1	1	1	0	1	0	1	0	0	1	0	1	0	1	0	0
$S_3(x)$	1	1	0	1	1	1	1	0	0	0	0	0	1	0	0	1
$S_4(x)$	1	1	1	1	0	1	0	1	0	0	1	1	0	0	0	0

$$\deg S_1 S_2 \leq 3$$

Exercise: find the maximal degree of the following 4-bit function



with $f(y_1, y_2, y_3, y_4) = y_1y_2y_3$.

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S(x)$	f	e	b	c	6	d	7	8	0	3	9	a	4	2	1	5
$S_1(x)$	1	0	1	0	0	1	1	0	0	1	1	0	0	0	1	1
$S_2(x)$	1	1	1	0	1	0	1	0	0	1	0	1	0	1	0	0
$S_3(x)$	1	1	0	1	1	1	1	0	0	0	0	0	1	0	0	1
$S_4(x)$	1	1	1	1	0	1	0	1	0	0	1	1	0	0	0	0

$$\deg S_1S_2S_3 \leq 3$$

Product of some coordinates of a permutation

δ_k = maximal degree of the product of k coordinates of S

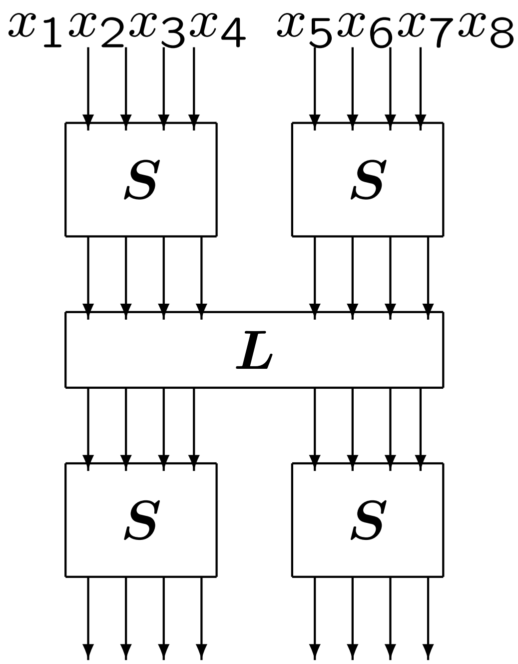
In our example:

k	1	2	3	4
δ_k	3	3	3	4

Proposition. If S is a permutation of $\mathbf{F}_2^n$,

$\delta_k = n$ if and only if $k = n$

Exercise: find the maximal degree of the following 2-round cipher

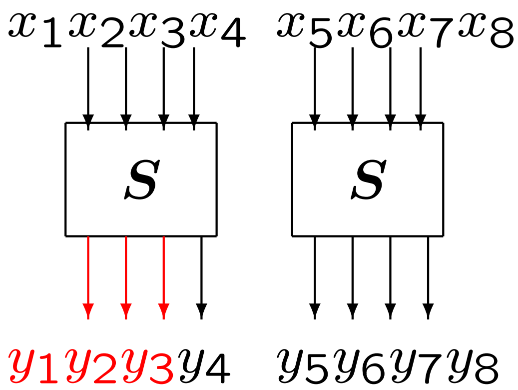


where the 4-bit Sbox S is such that the maximal degree δ_k of the product of k of its coordinates is given by

k	1	2	3	4
δ_k	3	3	3	4

Exercise: find the maximal degree of the following 2-round cipher

f is a function of degree ≤ 3 of the output of the first Sbox layer



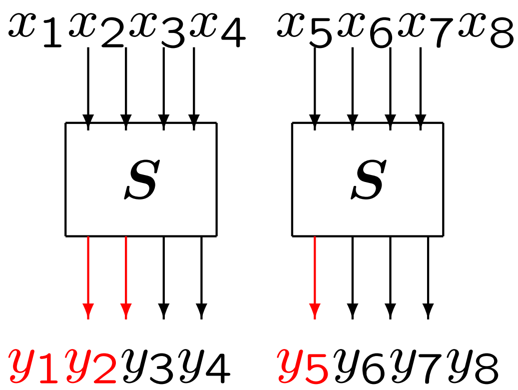
First case:

$$f(x_1, \ldots, x_8) = y_1 y_2 y_3 = S_1(x_1, \ldots, x_4) S_2(x_1, \ldots, x_4) S_3(x_1, \ldots, x_4)$$

$$\Rightarrow \deg f \leq 3$$

Exercise: find the maximal degree of the following 2-round cipher

f is a function of degree ≤ 3 of the output of the first Sbox layer



Second case:

$$f(x_1, \ldots, x_8) = y_1 y_2 y_5 = S_1(x_1, \ldots, x_4) S_2(x_1, \ldots, x_4) S_1(x_5, \ldots, x_8)$$

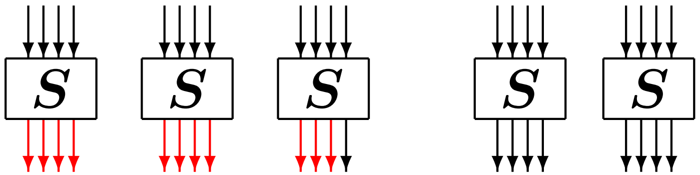
$$\Rightarrow \deg f \leq 3 + 3 = 6$$

Degree of the product f of d output coordinates

A fundamental parameter:

$\boldsymbol{\delta_k}$ = maximal degree of the product of $\boldsymbol{k}$ coordinates of $\boldsymbol{S}$

Example: $\boldsymbol{d = 11}$



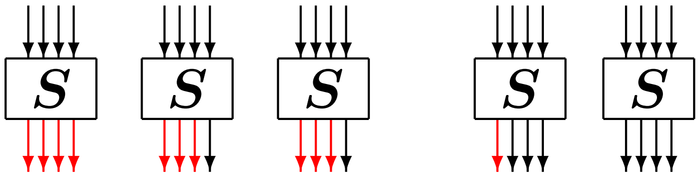
$$\deg \boldsymbol{f} \leq \boldsymbol{2\delta_4 + \delta_3}$$

Degree of the product f of d output coordinates

A fundamental parameter:

δ_k = maximal degree of the product of k coordinates of S

Example: $d = 11$



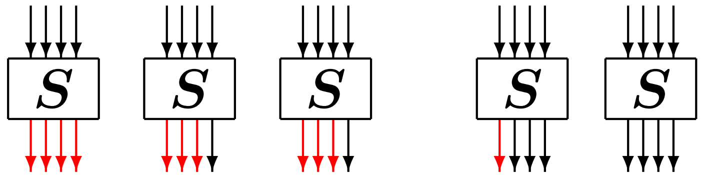
$$\deg f \leq \delta_4 + 2\delta_3 + \delta_1$$

Degree of the product f of d output coordinates

A fundamental parameter:

δ_k = maximal degree of the product of k coordinates of S

Example: $d = 11$



$$\deg f \leq \max_{(x_1,\ldots,x_4)} (x_1\delta_1 + \ldots + x_4\delta_4)$$

with $x_1 + 2x_2 + 3x_3 + 4x_4 = d$.

A new bound [Boura, C., De Cannière 09]

Theorem. Let $F = (S, \ldots, S)$ where S is a permutation of $\mathbf{F}_2^m$. Then,

$$\deg(G \circ F) \leq n - \frac{n - \deg G}{\gamma(S)}$$

where

$$\gamma(S) = \max_{1 \leq k \leq m-1} \frac{m-k}{m - \delta_k(S)} .$$

Most notably,

$$\gamma(S) \leq m - 1 \text{ with equality iff } \deg S = m - 1.$$

Our example

$$\gamma(S) = \max_{1\le k<4} \frac{4-k}{4-\delta_k(S)}$$

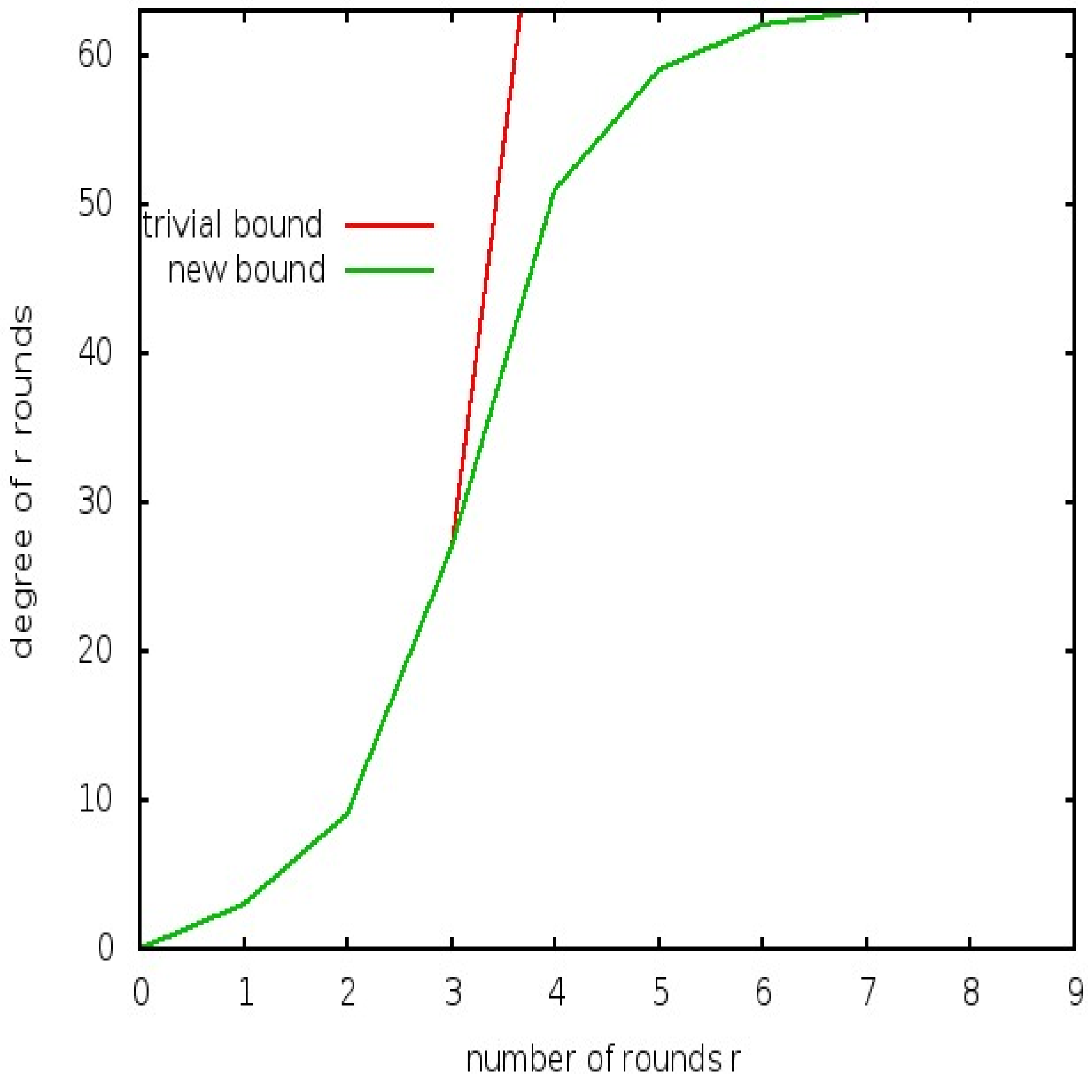
k	1	2	3	4
$\delta_k(S)$	3	3	3	4

$$\gamma(S) \le \max\left(\frac{3}{1},\ \frac{2}{1},\ \frac{2}{1}\right) = 3$$

We deduce

$$\deg(G \circ F) \le n - \frac{n - \deg G}{3}$$

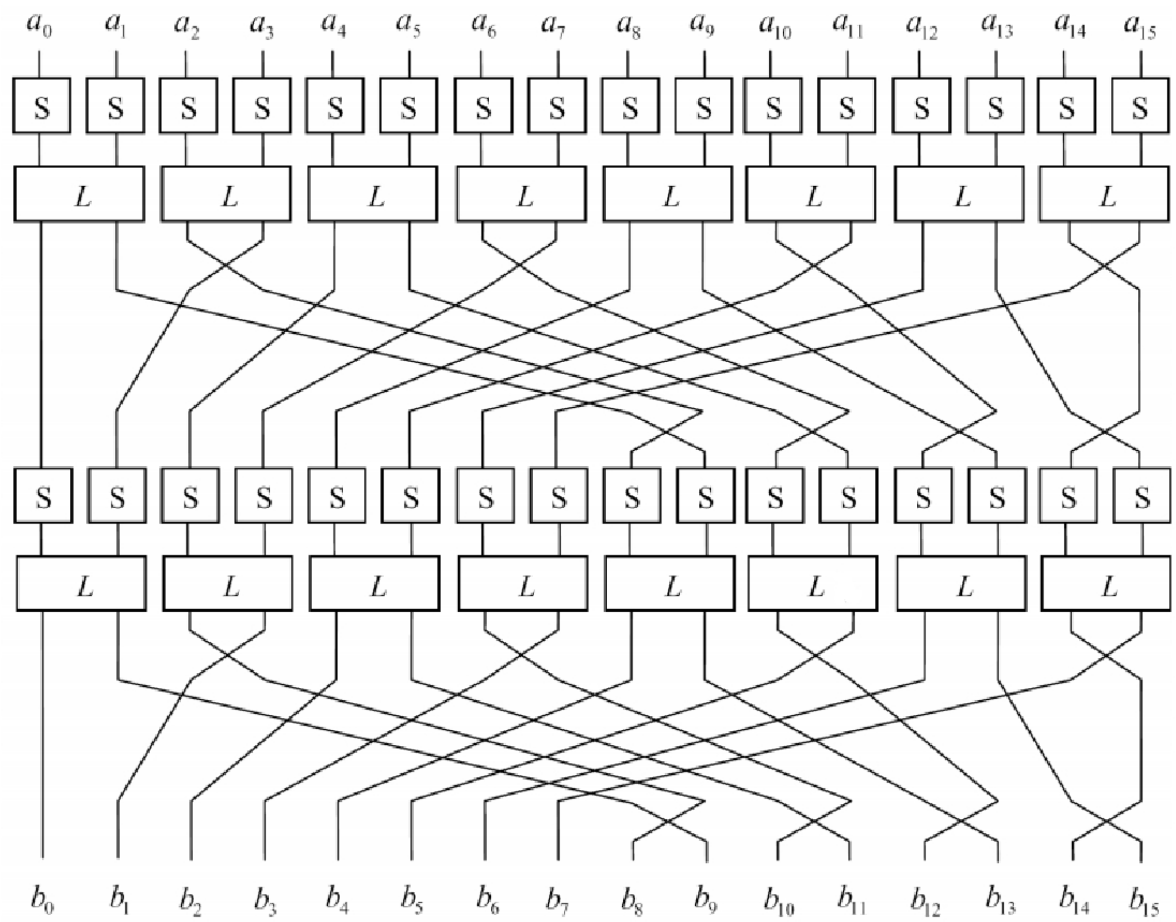
Our example


60
50
40
30
20
10
0
trivial bound
new bound
degree of r rounds
0
1
2
3
4
5
6
7
8
9
number of rounds r

How does the degree increase with the number of rounds?

Inspired by JH [Wu 2007]

block size = 64 bits with a 4-bit Sbox



How does the degree increase with the number of rounds?

SuperBox view:

r rounds without the last linear layer = each output depends on 2^{r+1} input bits.

For $r = 3$:

$n = 16$, $\gamma(S) = 3$, $\deg G \leq 6$

$$\Rightarrow \deg(G \circ F) \leq 16 - \frac{16-6}{3} = 12.66$$

How does the degree increase with the number of rounds?

SuperBox view:

r rounds without the last linear layer = each output depends on 2^{r+1} input bits.

For $r = 3$:

$n = 16,\ \gamma(S) = 3,\ \deg G \le 6$

$$\Rightarrow \deg(G \circ F) \le 16 - \frac{16-6}{3} = 12.66$$

# rounds	block size	bound
2	8	6
3	16	12
4	32	25
5	64	51
6	64	59
7	64	62
8	64	63

Key recovery on 6 rounds

Let G_k be the first 5 rounds of the cipher

$$\deg G_k \leq 51$$

$$\Rightarrow \bigoplus_{x \in a+V} G_k(x) = 0 \text{ for any } (a+V) \text{ with } \dim V = 52$$

Key recovery on 6 rounds

Input: 2^{52} pc pairs (m_i, c_i) with m_i in an affine subspace of dim 52

For recovering the j-th key byte k_j:

For each possible value for k_j
    For i from 0 to $(2^{52} - 1)$
        $y_{i,j} \leftarrow \mathcal{R}^{-1}(c_{i,j} \oplus k_j)$
    If $\bigoplus_i y_{i,j} = 0$
        Return k_j

$$D = 2^{52} \text{ and } T = 8 \times 2^{60}$$

Are there better choices for V?

For $V = a + \langle e_0, \dots, e_{27} \rangle$,

$V =$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$

- invariant under the key addition and the first Sbox layer
- Let H_k = first linear layer + rounds 2 to 5.

  Since $\deg H_k \leq 25$ and $\dim V = 28$,

$$\bigoplus_{x \in a+V} H_k(x) = \bigoplus_{x \in b+V} G_k(x) = D_V G_k(b) = 0$$

$$\Rightarrow D = 2^{28} \text{ and } T = 8 \times 2^{36}$$

What does it mean?

$$G_k(x) = \sum_{u \in \mathbf{F}_2^n} c_u x^u \text{ with } c_u = \bigoplus_{x \preceq u} G_k(x)$$

For $V = \langle e_0, \ldots, e_{27} \rangle = \{x : x \preceq 0 \ldots 0 \underbrace{1 \ldots 1}_{28}\}$

$$\Rightarrow \bigoplus_{x \in a+V} G_k(x) = 0 \text{ for all } a \in \mathbf{F}_2^n$$

means that the ANF of G_k does not contain any monomial multiple of $x_0 \ldots x_{27}$

⇒Cube distinguisher with data complexity $2^{wt(u)}$[Aumasson, Dinur, Meier, Shamir 09]

Many refinements

- zero-sum distinguishers on SHA-3 [Boura, C., De Cannière 11]
- interpolation attack on Low-MC [Dinur, Liu, Meier, Wang 15]
- attack on the full Misty-1 using the division property [Todo 15]