Exact maximum expected differential and linear probability for 2-round Kuznyechik

Vitaly Kiryukhin

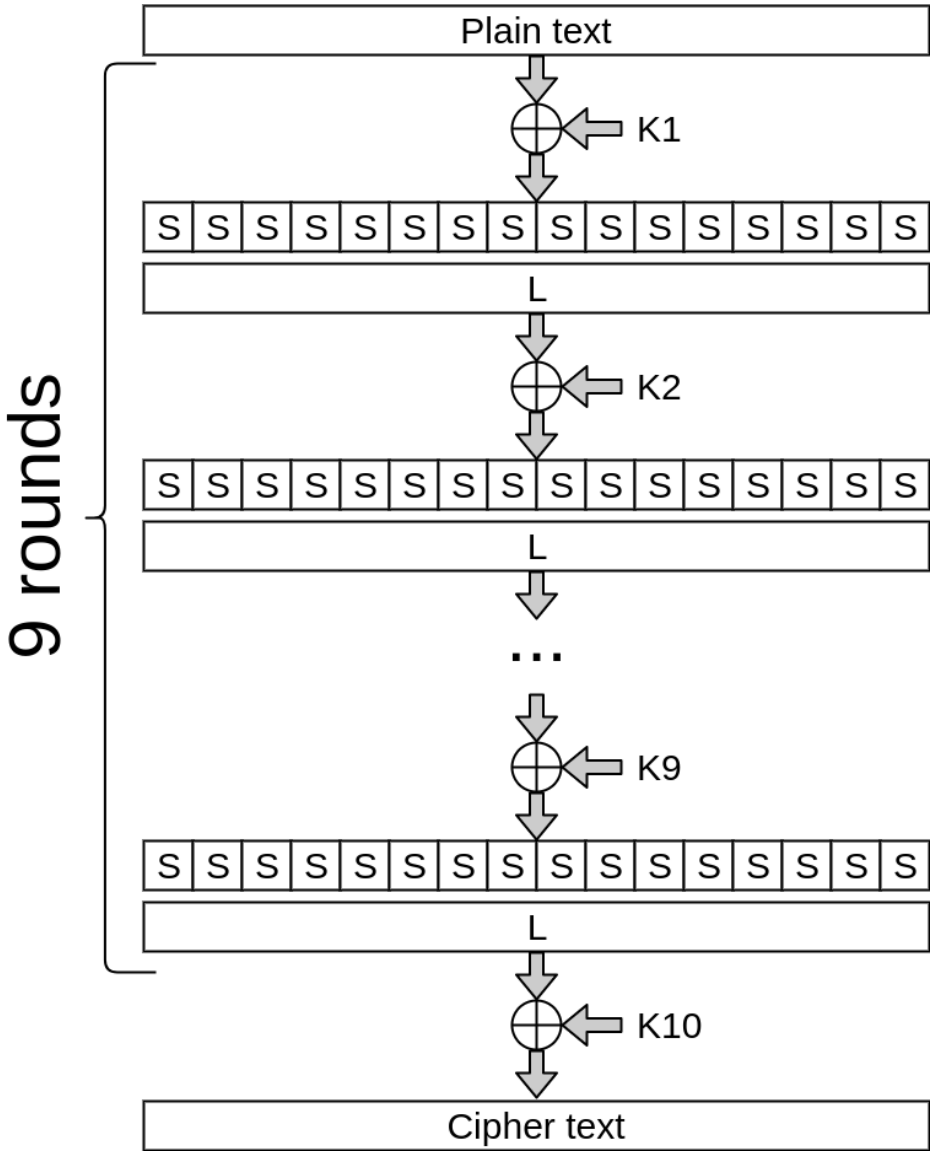
JSC «InfoTeCS»

CTCrypt'18

May 29, 2018

vitaly.kiryukhin@infotecs.ru

GOST 34.12-2015 – «Kuznyechik»



Kuznyechik is an LSX block cipher

Block size – 128 bit ($n = 16$ byte)

Key size – 256 bit

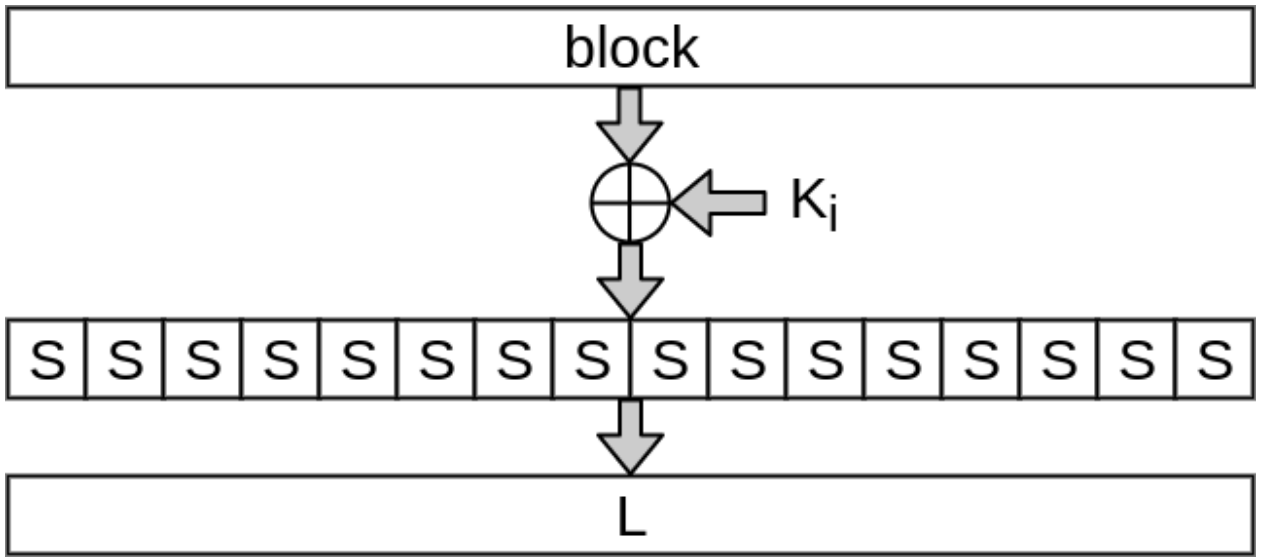
It contains 9 full rounds

Round transformations

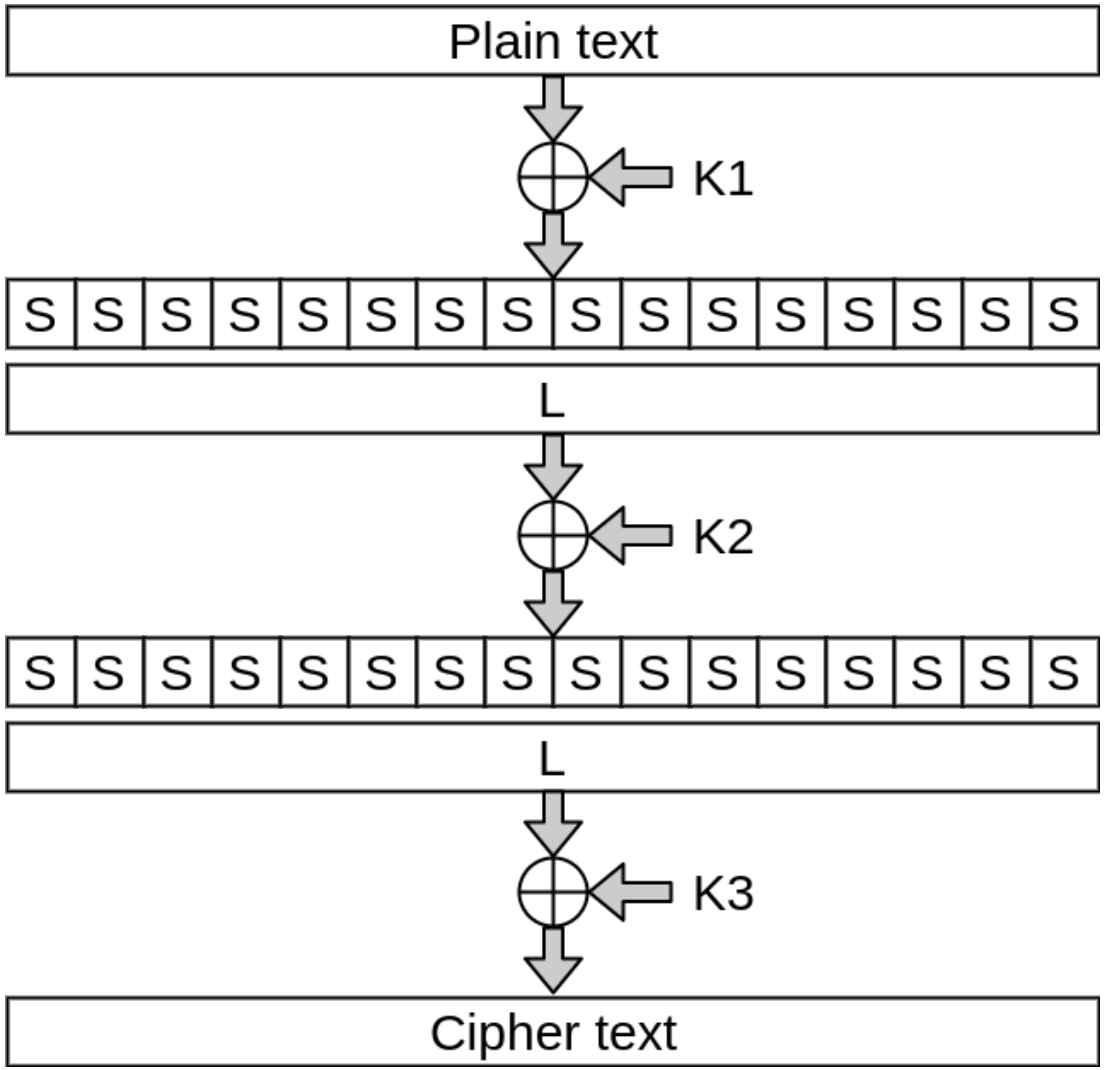
X – modulo 2 addition of an input block with an iterative key

S – parallel application of a fixed bijective byte substitution

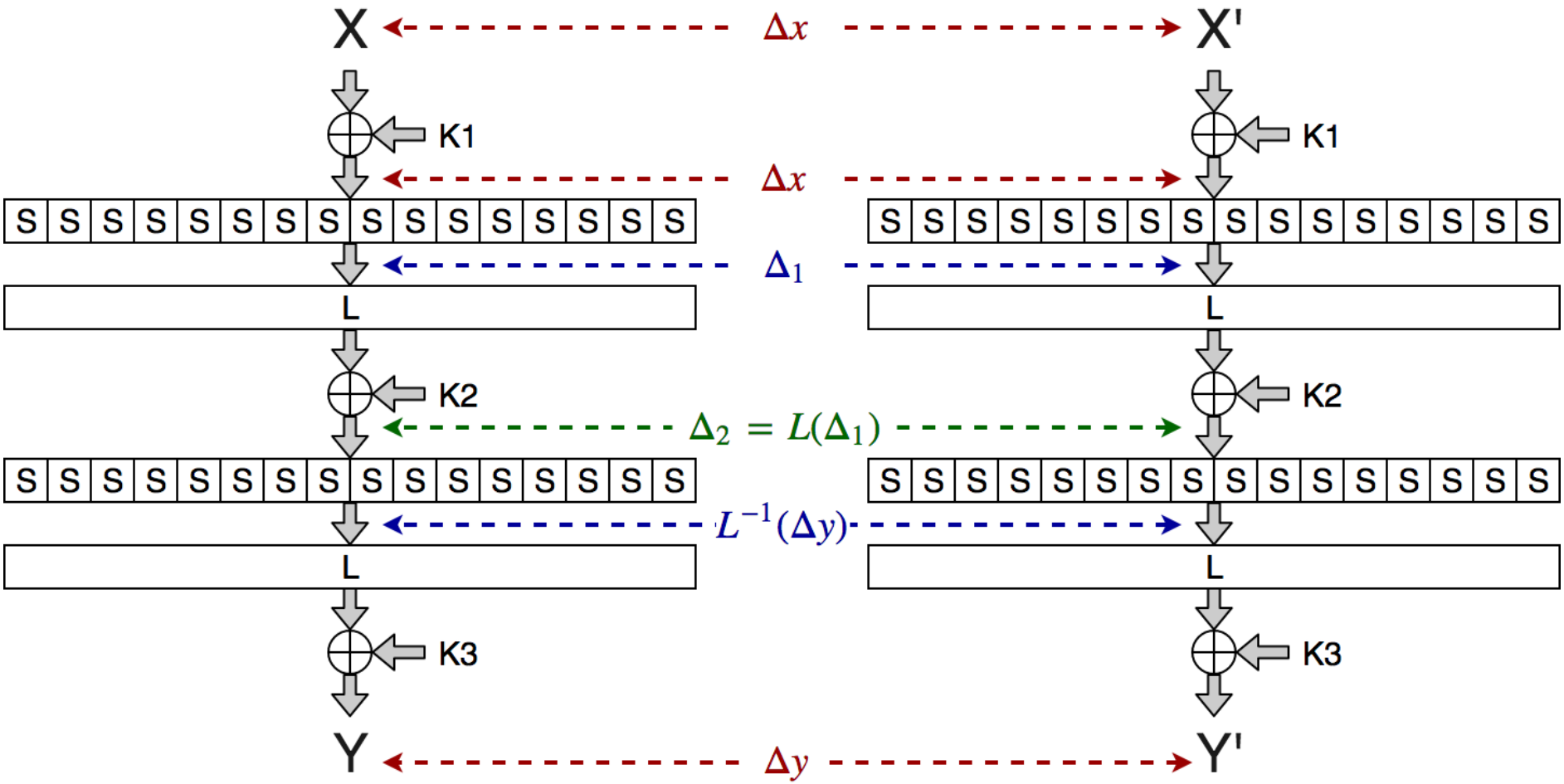
L – linear transformation – MDS(32, 16, 17), optimal diffusion operation, branch number (minimal code distance) $\mathcal{B} = 17$



2-round Kuznyechik



Differential trail in 2-round Kuznyechik

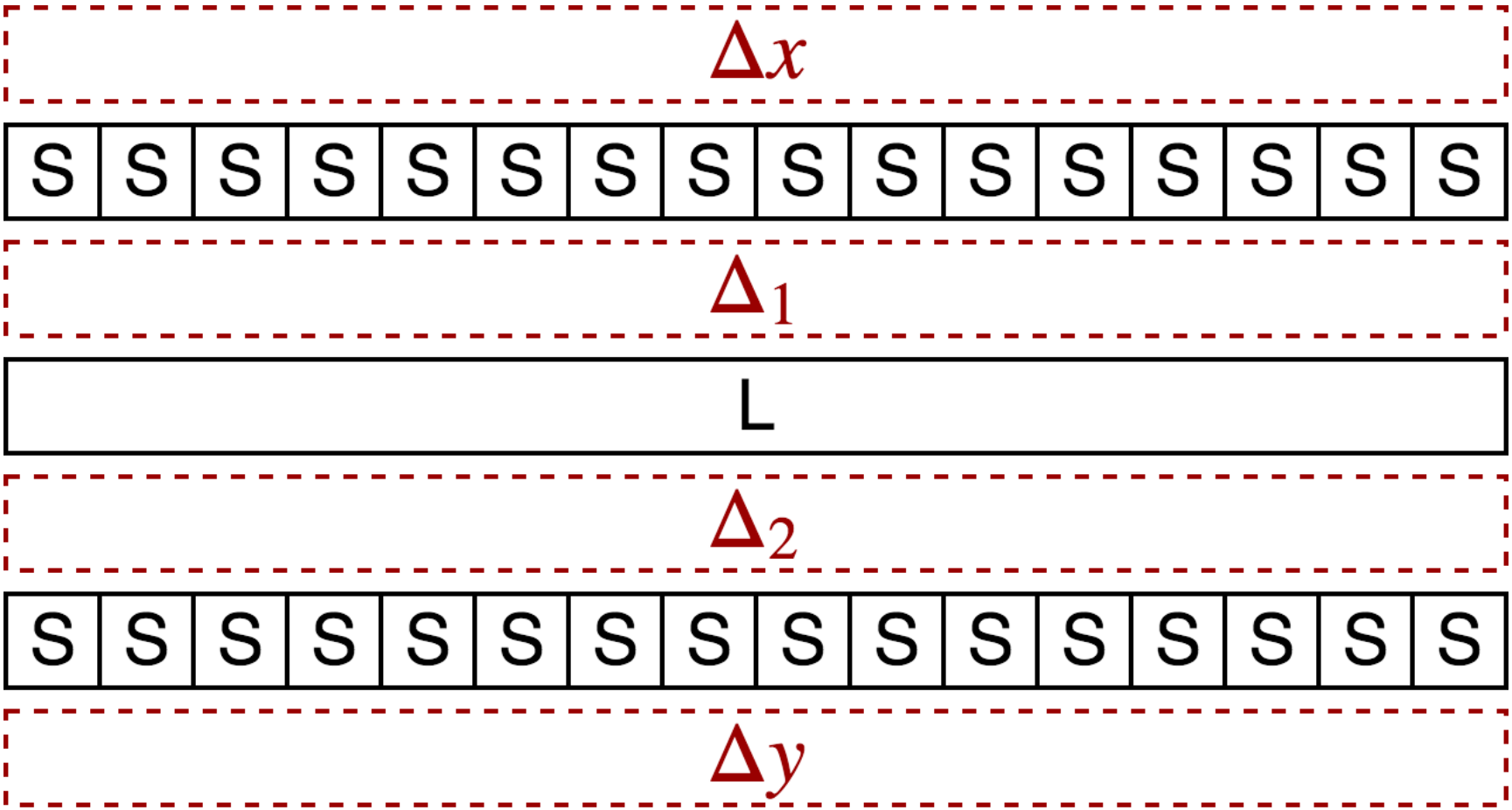


The linear transformation on the second round can be omitted without loss of generality.

Differential trail

$$\Omega = \Delta x \xrightarrow{S} \Delta_1 \xrightarrow{L} \Delta_2 \xrightarrow{S} \Delta y$$

$$\Pr(\Omega) = \prod_{i=1}^{n} \Pr(\Delta x[i] \to \Delta_1[i]) \cdot \prod_{i=1}^{n} \Pr(\Delta_2[i] \to \Delta y[i])$$

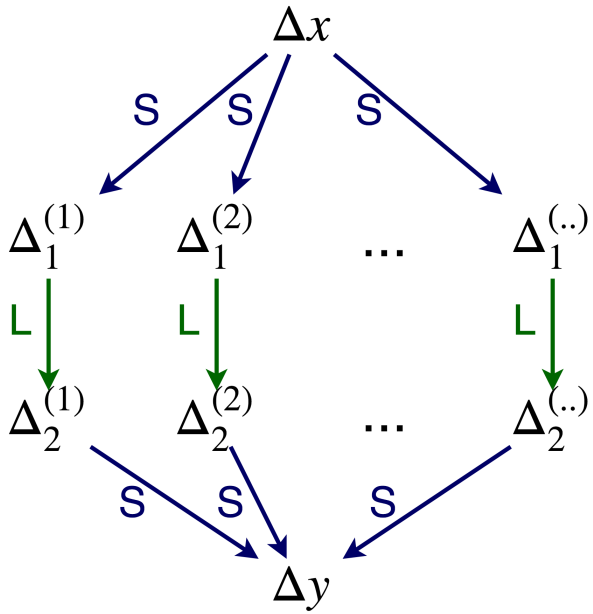


Differential

$$DIFF(\Delta x, \Delta y) = \{\Omega : \Omega = \Delta x \to \ldots \to \Delta y\}$$

$$\Pr\left(DIFF(\Delta x, \Delta y)\right) = \sum_{\Omega \in DIFF(\Delta x, \Delta y)} \Pr(\Omega)$$

$$MEDP = \max_{DIFF(\Delta x, \Delta y) \setminus (0,0)} \Pr\left(DIFF(\Delta x, \Delta y)\right)$$



Our target

1) The best differential trail

$$\max_{\Omega \setminus 0} \Pr(\Delta x \xrightarrow{S} \Delta_1 \xrightarrow{L} \Delta_2 \xrightarrow{S} \Delta y)$$

2) The best differential

$$MEDP = \max_{DIFF(\Delta x, \Delta y) \setminus (0,0)} \Pr\left(DIFF(\Delta x, \Delta y)\right)$$

Estimate of differential trail probability

Theorem

Θ – *minimum number of active S-boxes*

p_{max} – *maximum local differential probability*

$\Omega = \Delta x \to \ldots \to \Delta y$ – *any differential trail*

$$\Pr(\Omega) \leq p_{max}^{\Theta}$$

2-round Kuznyechik

$\Theta = \mathcal{B} = 17$ – property of MDS code

$p_{max} = \left(\frac{8}{256}\right)$ – maximum local differential probability

$\Omega = \Delta x \xrightarrow{S} \Delta_1 \xrightarrow{L} \Delta_2 \xrightarrow{S} \Delta y$

$$\Pr(\Omega) \leq p_{max}^{\Theta} = \left(\frac{8}{256}\right)^{17} = 2^{-85}$$

The best differential trail

$\Omega = \Delta x \xrightarrow{S} \Delta_1 \xrightarrow{L} \Delta_2 \xrightarrow{S} \Delta y$ is 2-round Kuznyechik trail.

$w(\Omega) = w(\Delta_1, \Delta_2)$ is the number of non-zero bytes in Δ_1 and Δ_2 (the number of active S-boxes in Ω)

For Kuznyechik: $\mathcal{B} = 17 \leq w(\Omega) \leq 32$

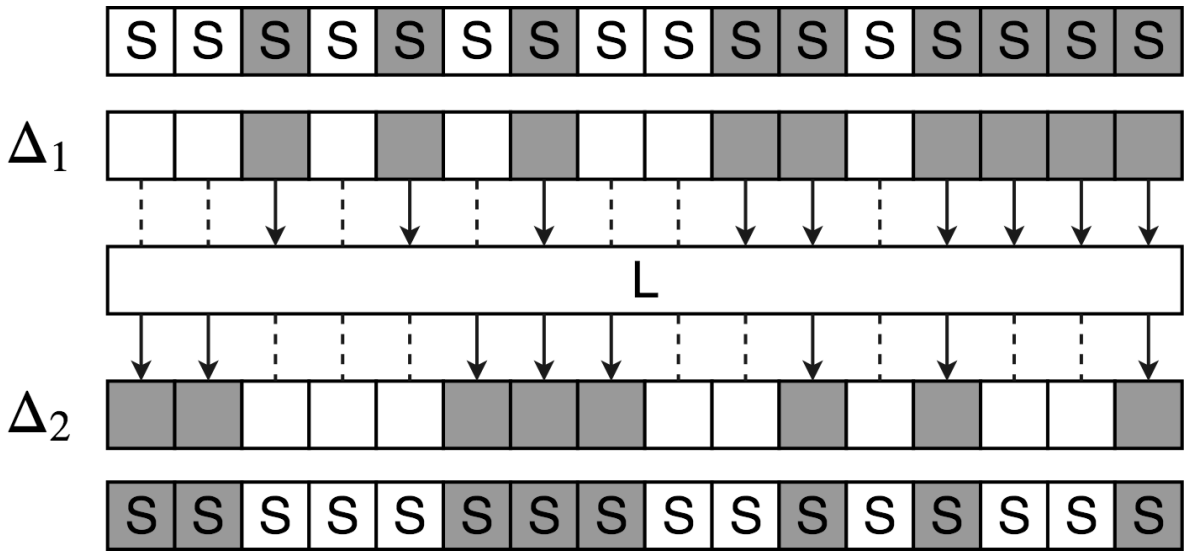
Idea to find the best trail:

1. we will finding all $(\Delta_1, \Delta_2) : w(\Delta_1, \Delta_2) = 17$
2. for each (Δ_1, Δ_2) there is an optimal pair Δx, Δy
3. among the constructed trails we will find the trail with the maximum probability
4. we will show that any trail $\Omega : \ w(\Omega) > 17$ has less probability than the best trail

Algorithm for finding codewords with the smallest byte weight

How to find all $(\Delta_1, \Delta_2) : w(\Delta_1, \Delta_2) = 17$?

1) Fix locations of active S-boxes of first and second layers



Algorithm for finding codewords with the smallest byte weight

2) Let's present the linear transformation as a system of equations $\Delta_1\mathbb{L} = \Delta_2$ over $GF(2^8)$.

$\Delta_1 = (x_1, x_2, x_3, \ldots, x_{16})$, $\Delta_2 = (y_1, y_2, y_3, \ldots, y_{16})$

3) Select and solve the subsystem $\mathbb{S}$ in $\Delta_1\mathbb{L} = \Delta_2$

Let $w(\Delta_1) = t$, $w(\Delta_2) = r$, $t + r = 17$.

$\mathbb{S}$ consists of t variables («columns» with non-zero x_j) and $n - r = t - 1$ equations («rows» with zero y_j)

After solving the system we have a set of solutions $\Delta_1^{(i)}\mathbb{L} = \Delta_2^{(i)}$, $i = \overline{1, 255}$

Algorithm for finding codewords with the smallest byte weight

Number of different equation systems is

$$\sum_{(t,r):t+r=n+1} \binom{n}{t}\binom{n}{r} = \binom{2n}{n+1}$$

Each system has $|GF(2^8)| - 1 = 255$ solutions.

Number of the (Δ_1, Δ_2) : $w(\Delta_1, \Delta_2) = 17$ is

$$255 \cdot \binom{2n}{n+1} = 255 \cdot \binom{32}{17} \approx 2^{37}$$

In other words, we found all codewords with the smallest byte weight in MDS code.

The best trail

The result of a search among code words (Δ_1, Δ_2) : $w(\Delta_1, \Delta_2) = 17$ is the best trail:

$$\max_{(\Delta_1,\Delta_2):\ w(\Delta_1,\Delta_2)=17} \Pr(\Delta x \to \Delta_1 \to \Delta_2 \to \Delta y) = \left(\frac{8}{256}\right)^{13} \left(\frac{6}{256}\right)^{4} = 2^{-86.66\ldots}$$

For any $w > 17$ holds:

$$P(\Delta x \to \Delta_1 \to \Delta_2 \to \Delta y) \leq \left(\frac{8}{256}\right)^{w} \leq \left(\frac{8}{256}\right)^{18} = 2^{-90} < 2^{-86.66\ldots}$$

Hence we find the best trail. It contains 17 active S-boxes.

The best trail

Two different pairs (Δ_1, Δ_2) were found that generate the best trails.

It is shown that there are no other such.

One of the best trails:

00950000000008200adad1b00ff00007b	Δx
0 8 0 0 0 0 8 0 8 8 8 0 6 0 0 8	$P(\Delta x \to \Delta_1) \cdot 256$
001900000002d00b8b8950072000028	Δ_1
2a00000d2337f74d0082a80000009d1b	Δ_2
8 0 0 8 8 8 8 6 0 8 6 0 0 0 6 8	$P(\Delta_2 \to \Delta y) \cdot 256$
030000e08bec5a55002da90000005a95	Δy

Algorithm for finding the best differential

1. We will find for the best differential among differentials consisting of trails $\Omega: \ w(\Omega) = \mathcal{B} = 17$
2. Thereafter we analytically show that any differential $\{\Omega: \ w(\Omega) > \mathcal{B}\}$ is worse than the one found in the first step.

Algorithm for finding the best differential

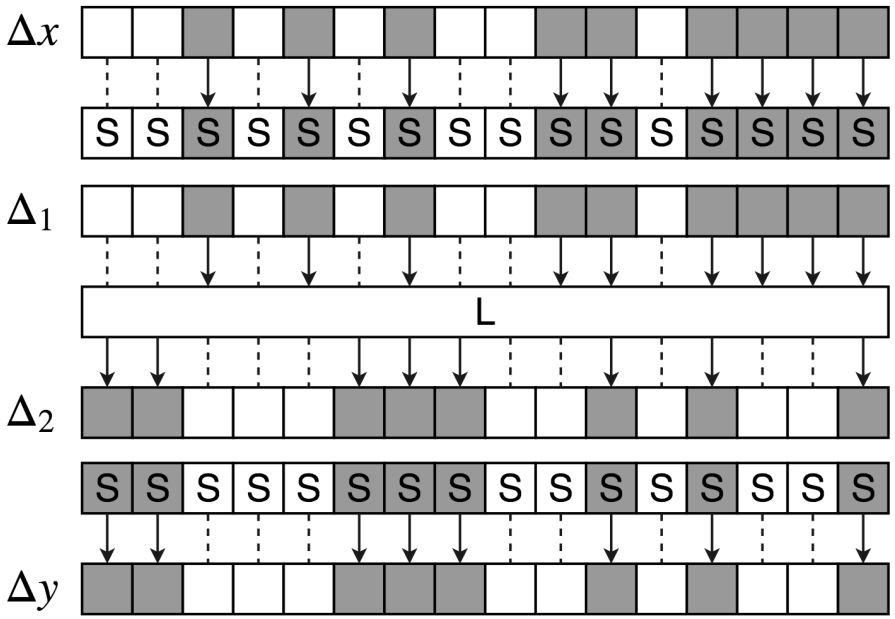
We will use the following considerations:

- The probability of the best *trail* is the lower bound for the probability of the best *differential.* This probability is a threshold value.
- Any $DIFF(\Delta x, \Delta y) = \{\Omega : \ w(\Omega) = \mathcal{B} = 17\}$ contains no more than 255 trails

Algorithm for finding the best differential

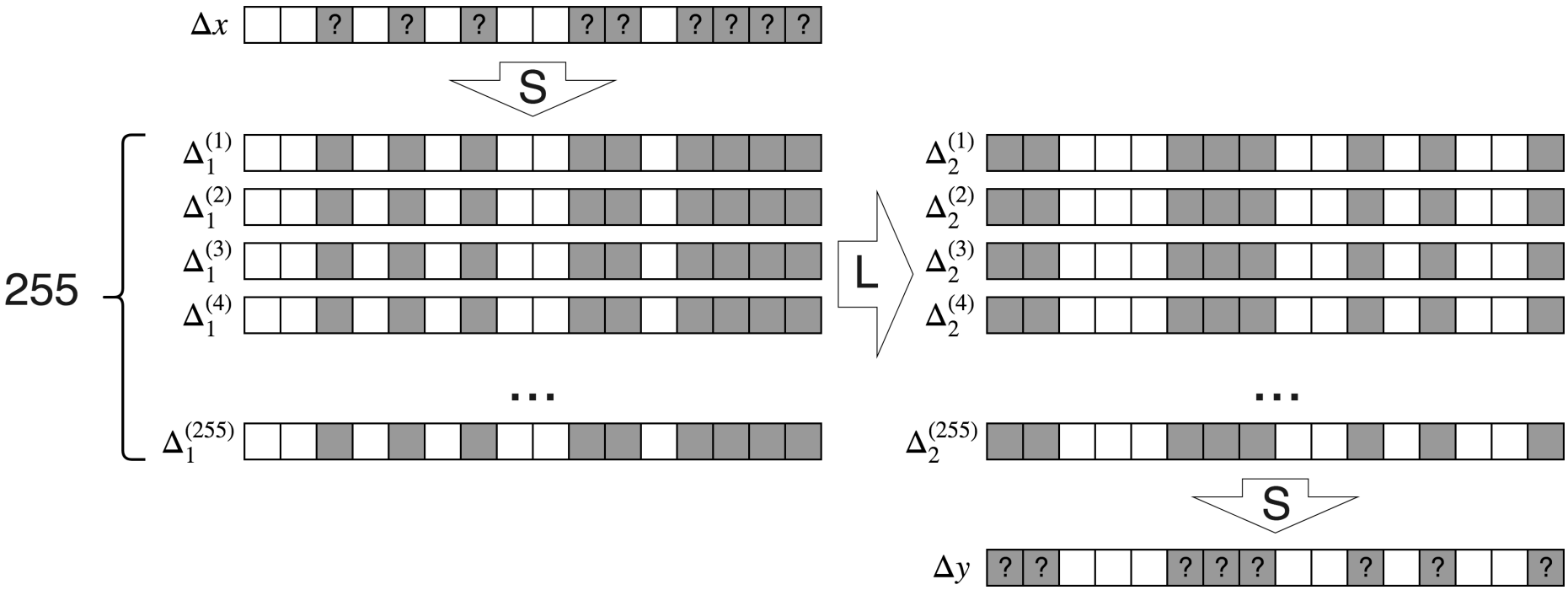
Any $DIFF(\Delta x, \Delta y) = \{\Omega : \ w(\Omega) = \mathcal{B} = 17\}$ contains no more than 255 trails.

Let we have differential $(\Delta x, \Delta y)$. It has some set of active S-boxes. Hence any trail $\Delta x \to \Delta_1 \to \Delta_2 \to \Delta y$ in the differential has the same active S-boxes. Therefore, all (Δ_1, Δ_2) of these trail are solutions of one subsystem $\mathbb{S}$ of equations. The number of solutions is 255.



Algorithm for finding the best differential

Solutions $\left(\Delta_1^{(1)}, \Delta_2^{(1)}\right)$, ..., $\left(\Delta_1^{(255)}, \Delta_2^{(255)}\right)$ of $\mathbb{S}$ can be represented as follows



Algorithm for finding the best differential

We have to choose the values $(\Delta x, \Delta y)$ so as to maximize the differential.

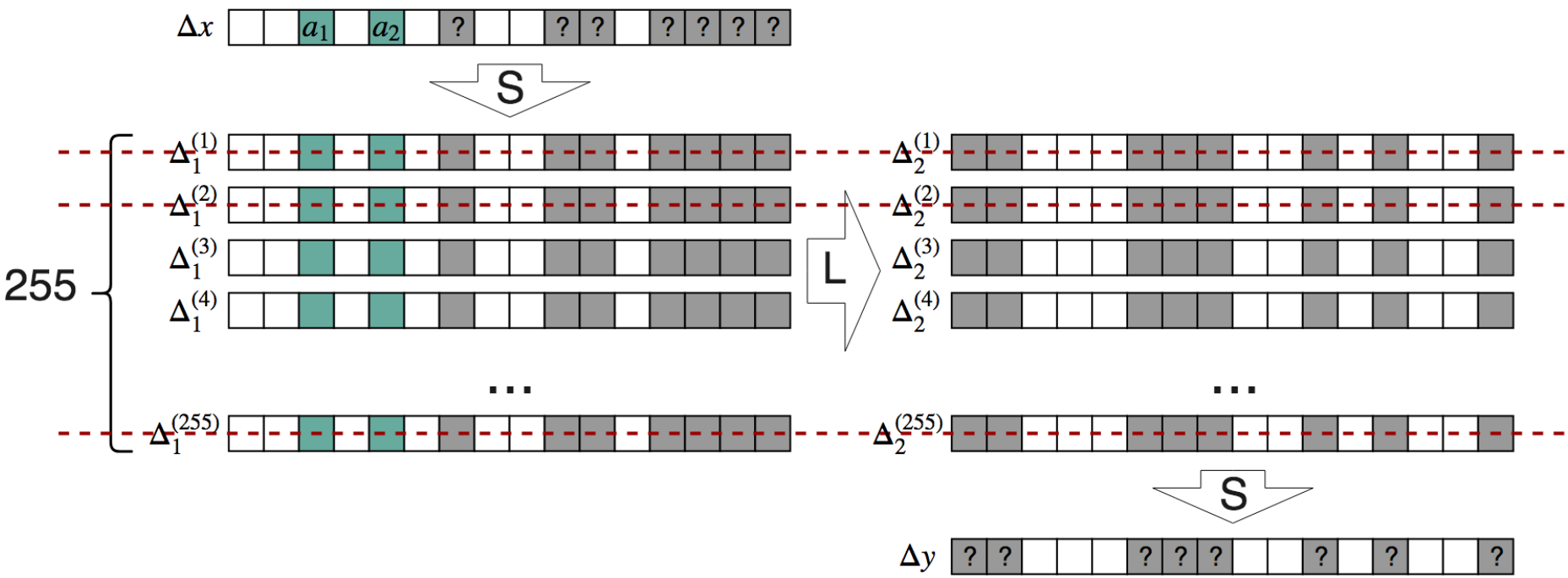
Optimal differential trail $\Omega^{(i)} = \Delta x^{(i)} \xrightarrow{S} \Delta_1^{(i)} \xrightarrow{L} \Delta_2^{(i)} \xrightarrow{S} \Delta y^{(i)}$ exists for each solution $\left(\Delta_1^{(i)}, \Delta_2^{(i)}\right)$.

$\sum_{i=1}^{255} \Pr(\Omega^{(i)})$ – is upper estimate for differential $(\Delta x, \Delta y)$.

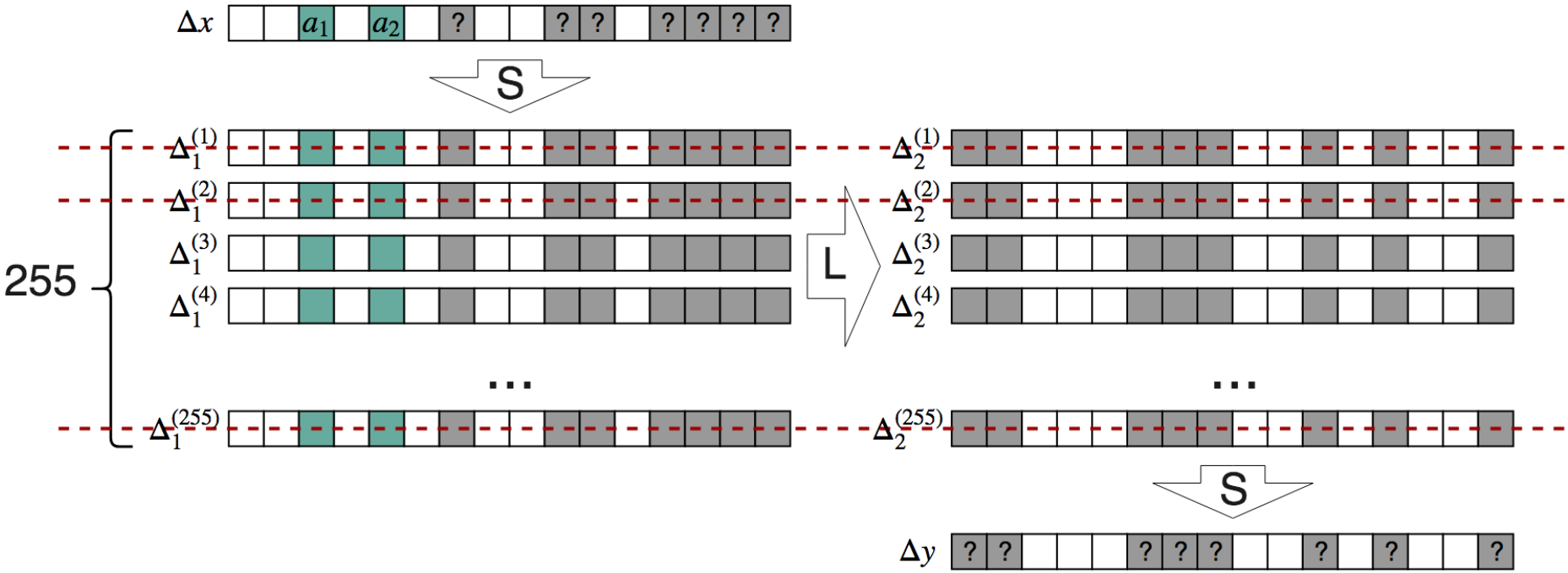
If this estimate is less than threshold value then we refuse to construct a differential on this set of solutions.

Algorithm for finding the best differential

We will *recursively* iterate over nonzero bytes in $(\Delta x, \Delta y)$. The algorithm uses the «pruning» of the branches of the search tree due to the constructed upper bounds.



Algorithm for finding the best differential



- All items in one column are different (the property of MDS code)
- Fixing one byte does not increase the number of trails in differential and the estimate of its probability.
- Number of trails in such differential is no more than 128 for any bijective S-box.
- In practice, fixing a single byte reduces the number of trails by about half.

Algorithm for finding the best differential

Thus, the algorithm for finding the best differential consists of:

- consideration of all sets of solutions $\Delta_1^{(i)}\mathbb{L} = \Delta_2^{(i)}$
- construction of differentials for each of these sets

The best differential

Thus, we have shown that for 2-round Kuznyechik the best differential $DIFF(\Delta x, \Delta y) = \{\Omega : \ w(\Omega) = \mathcal{B} = 17\}$ contains only one differential trail and its probability is

$$\left(\frac{8}{256}\right)^{13}\left(\frac{6}{256}\right)^{4} = 2^{-86.66...}$$

Estimate of differential with $\mathcal{B}+1$ active S-boxes

Theorem

Let DIFF$(\Delta x, \Delta y)$ is the differential in 2-round Kuznyechik. Let $\forall \Omega \in$ DIFF$(\Delta x, \Delta y)$, $w(\Omega) > \mathcal{B}$. Then

$$\Pr\left(DIFF(\Delta x, \Delta y)\right) \leq 2^{-87.469\ldots}$$

This theorem was proved analytically.

Corollary

We know that the best differential DIFF$(\Delta x, \Delta y) = \{\Omega : \ w(\Omega) = \mathcal{B}\}$ has probability $2^{-86.66\ldots}$. Hence for 2-round Kuznyechik

$$MEDP = 2^{-86.66\ldots}$$

Linear properties of 2-round Kuznyechik

There is a certain duality between differential and linear cryptanalysis. It allows us to apply the algorithms described above to calculate linear characteristics.

The best linear characteristic has probability

$$\left(\frac{56}{256}\right)^{2\cdot 8}\left(\frac{52}{256}\right)^{2\cdot 7}\left(\frac{48}{256}\right)^{2\cdot 2} = 2^{-76.936\ldots}$$

The best linear hull contains 48 linear characteristics and has probability

$$\mathit{MELP} = 2^{-76.936\ldots} + 2^{-134.601\ldots}$$

About full-round Kuznyechik

For any LSX cipher, the N-round MEDP (MELP) is the upper bound for $(N+t)$-round MEDP (MELP) $\forall t \geq 0$.

Therefore, the 2-round MEDP (MELP) of Kuznyechik is the upper bound for any larger number of rounds.

Codewords with small binary weight

Let $\mathbb{G} = \mathbb{E}|\mathbb{L}$ is a linear binary code, codeword length – 256 bits, infoword length – 128 bits.

It is shown that in $\mathbb{G}$ there are no codewords of binary weight 17, ... ,20.

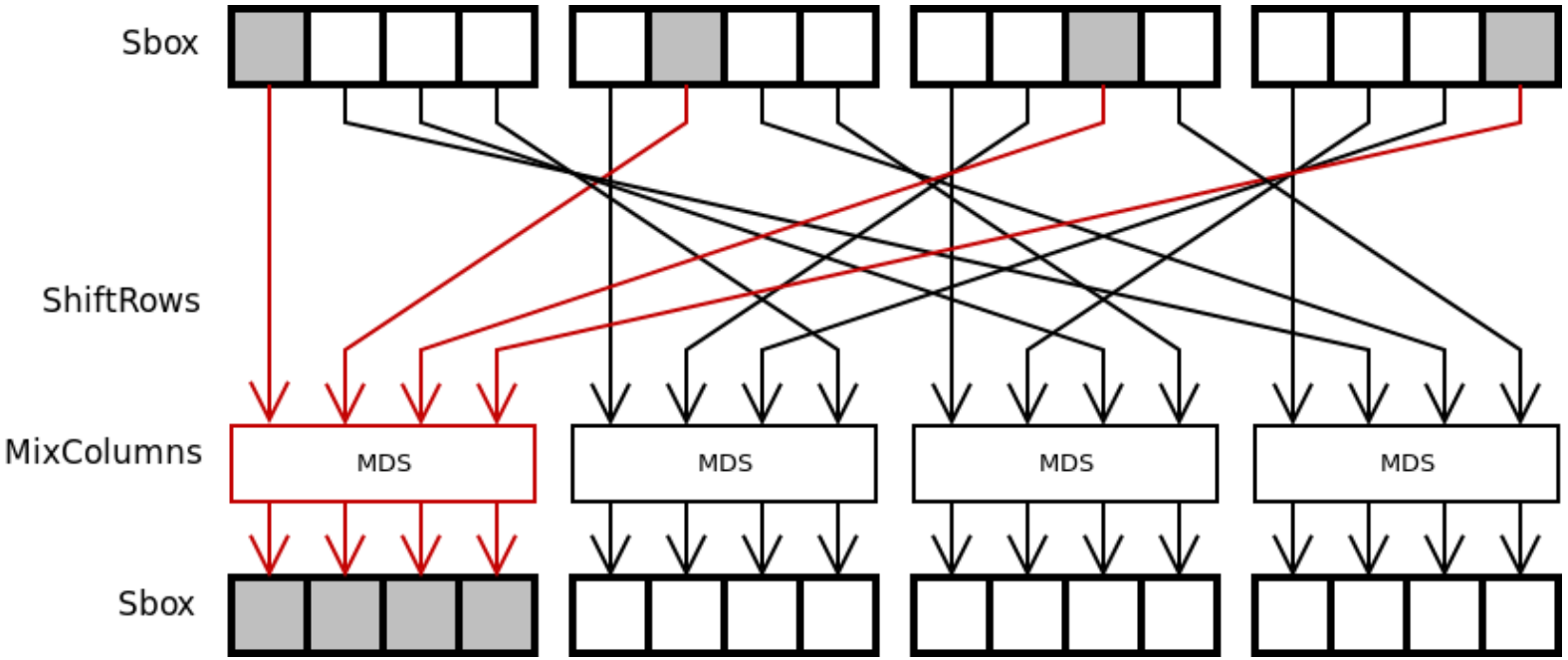
Two codewords with binary weight equal to 29 are found.

One of them:

0 2 0 2 2 0 0 0 0 2 1 0 1 2 0 1	w
009000a00300000000090100010900004	x
150400090100010900000000003a00090	$y = x\mathbb{L}$
3 1 0 2 1 0 1 2 0 0 0 0 2 2 0 2	w

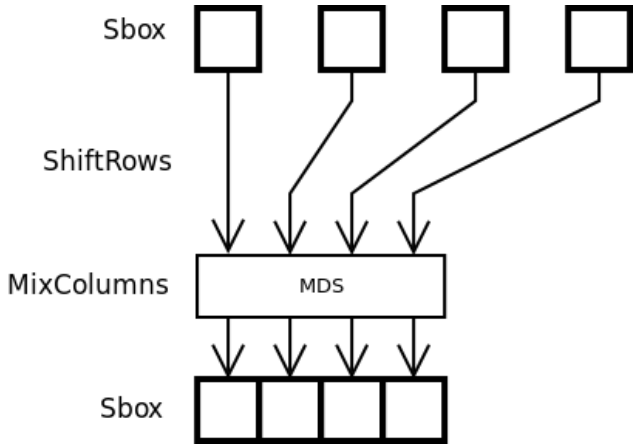
The comparison with AES

2-round AES:



The comparison with AES

The problem is easier due to the smaller block size (32 bit) and «analytical» form of S-box.



For 2-round AES: $\mathrm{MEDP} = 2^{-28.272\ldots}$, $\mathrm{MELP} = 2^{-27.287\ldots}$

Liam Keliher and Jiayuan Sui. *Exact Maximum Expected Differential and Linear Probability for 2-Round Advanced Encryption Standard (AES)*, Cryptology ePrint archive, Report 2005/321, 2005, `https://eprint.iacr.org/2005/321`

Conclusion

We presented algorithms:

- for finding codewords with the small byte weight in MDS-codes
- for finding all the best differential trails (linear characteristics) and differentials (linear hulls) in 2-round Kuznyechik

It was shown that in 2-round Kuznyechik:

- the best differential contains one differential trail
- the best linear hull contains 48 linear characteristics.

$MEDP = 2^{-86.66\ldots}$ and $MELP = 2^{-76.936\ldots}$ was proved.

Thank you for attention!

Questions?