Using Multiple Differentials...

On the LLR and χ^2 Statistical Tests in Differential Context

Céline Blondeau

Aalto University

Luxembourg, January 2013

joint work with Benoît Gérard and Kaisa Nyberg

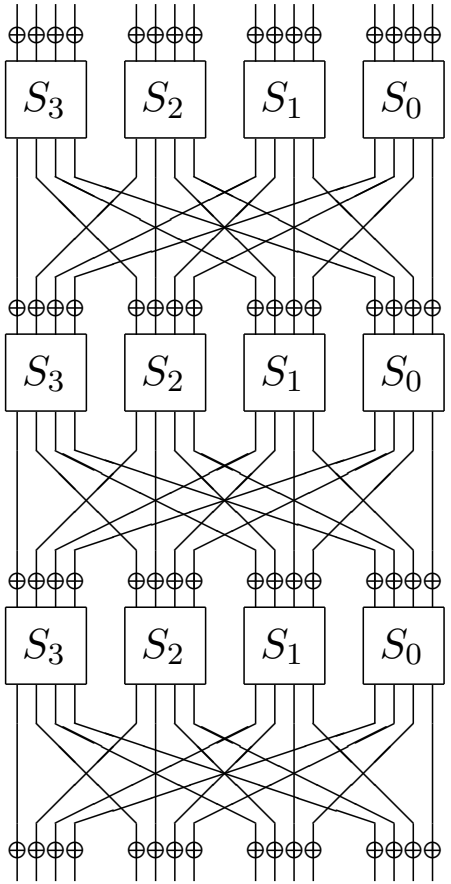
Outline

Outline

Block ciphers

$$E_K : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$$

- K: Master key
- F: Round function
- K_i: Round key



SMALLPRESENT-[4]

Statistical Attacks

Statistical attacks:

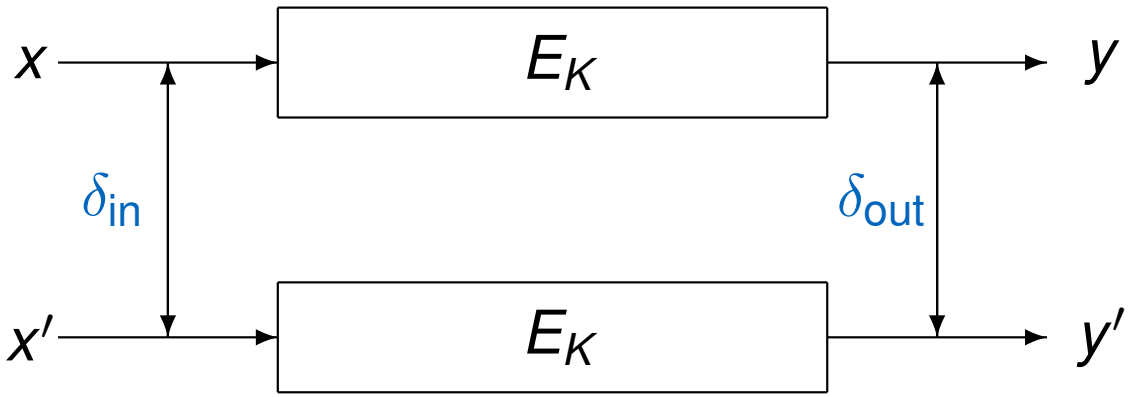
- Take advantage of a non-uniform behavior of the cipher
- Two families: Linear and Differential cryptanalysis

Improvement of differential cryptanalysis

- Differential cryptanalysis [Biham Shamir 91]
- Truncated differential cryptanalysis [Knudsen 95]
- Impossible differential cryptanalysis [Biham Biryukov Shamir 99]
- Higher order differential cryptanalysis [Lai 94] [Knudsen 95]
- Bulk Multiple differential cryptanalysis [Blondeau Gérard 11]

Differential Cryptanalysis

Given an input difference between two plaintexts, some output differences occur more often than others.



Differential: pair of input and output difference $(\delta_{\text{in}}, \delta_{\text{out}})$

Differential probability: $p = P_{X,K}[\, E_K(x) \oplus E_k(x \oplus \delta_{\text{in}}) = \delta_{\text{out}} \,]$

Uniform probability: $\theta = 2^{-m}$

Using Multiple differentials...

- Truncated differential cryptanalysis
- Impossible differential cryptanalysis
- Higher order differential cryptanalysis
- Bulk differential cryptanalysis

Using Multiple differentials...

- Truncated differential cryptanalysis
- Impossible differential cryptanalysis
- Higher order differential cryptanalysis
- Bulk differential cryptanalysis

One non-uniform probability is used for comparison with uniform probability

Bulk differential cryptanalysis [FSE 2011]

- Set of differences $(\delta_{\text{in}}^{(v)}, \delta_{\text{out}}^{(v)})$, with probabilities p_v.
- $p = \frac{1}{\Delta_{in}} \sum_v p_v$ expected probability.
- $\theta = \frac{1}{\Delta_{in}} \sum_v \frac{1}{2^m}$ uniform probability.

Bulk differential cryptanalysis [FSE 2011]

- Set of differences $({\delta_{\text{in}}}^{(v)}, {\delta_{\text{out}}}^{(v)})$, with probabilities p_v.
- $p = \frac{1}{\Delta_{in}} \sum_v p_v$ expected probability.
- $\theta = \frac{1}{\Delta_{in}} \sum_v \frac{1}{2^m}$ uniform probability.

Frequencies are summed up.

Bulk differential cryptanalysis [FSE 2011]

- Set of differences $(\delta_{\text{in}}^{(v)}, \delta_{\text{out}}^{(v)})$, with probabilities p_v.
- $p = \frac{1}{\Delta_{in}} \sum_v p_v$ expected probability.
- $\theta = \frac{1}{\Delta_{in}} \sum_v \frac{1}{2^m}$ uniform probability.

Frequencies are summed up.

How to use the probability of each differential individually?

Related Work

Linear Cryptanalysis [Matsui 93]:

- Multiple linear cryptanalysis [Baignères Junod Vaudenay 04]
- Multidimensional linear cryptanalysis [Hermelin Cho Nyberg 08]

Both use LLR and/or χ^2 statistical tests.

Differential Cryptanalysis:

Recently :

- How to apply LLR and/or χ^2 statistical tests?
- How to partition the output differences?

Multiple Differential Cryptanalysis

- Fixed input difference δ_{in} (To simplify the analysis)
- Vector of "differences": $V = [\delta_{\text{out}}^{(v)}]$ after r rounds,
- $p = [p_v]_{v \in V}$ vector of expected probabilities.
- $\theta = [\theta_v]_{v \in V}$ vector of uniform probabilities.
- $q^k = [q_v^k]_{v \in V}$ vector of observed probabilities for the key k.

Recent Work

[Albrecht Leander 12]

LLR statistical test

Application to SMALLPRESENT-[4] ($m = 16$) and KATAN-32 ($m = 32$)

[Blondeau Gérard Nyberg 12]

LLR and χ^2 statistical tests.

What to do when $m > 32$?

$\Rightarrow$ Introduction of partitioning functions

Outline

Partitioning Functions

We analyze two “orthogonal” cases

- Unbalanced partitioning
  - Take a subset of simple differences
- Balanced partitioning
  - Group the differences in order to be able to use information of the whole output space.

Partitioning Functions

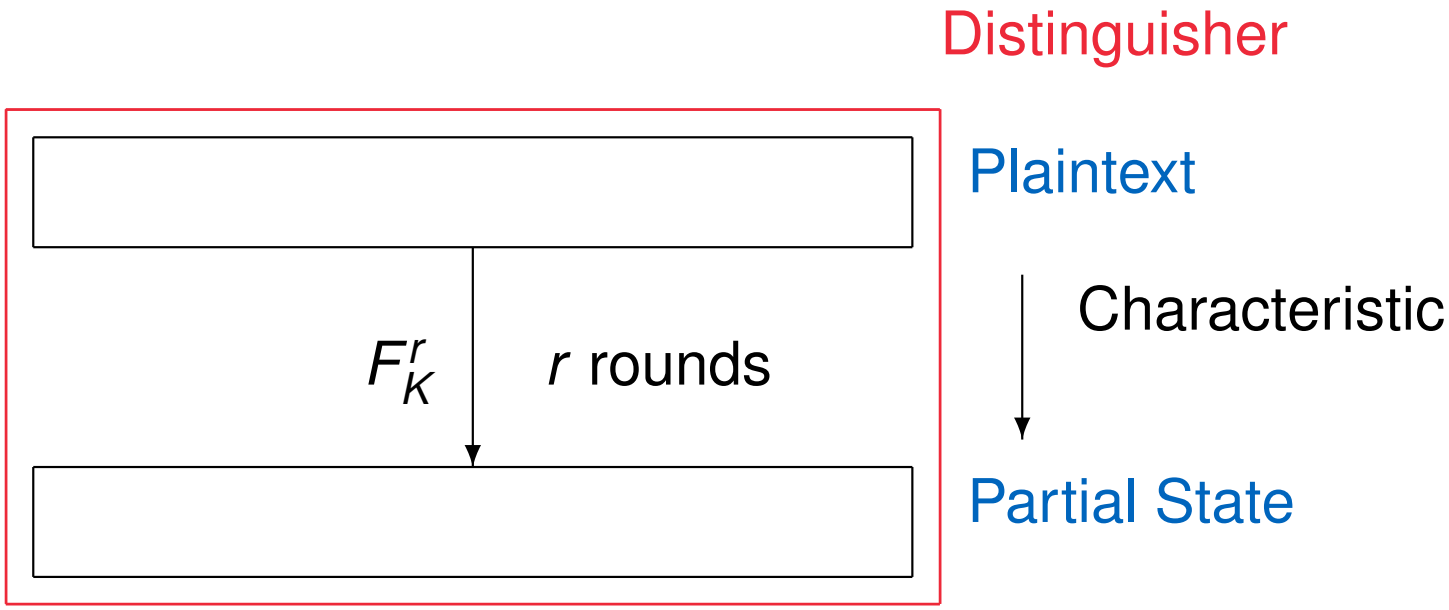
We analyze two “orthogonal” cases

- Unbalanced partitioning
  - Take a subset of simple differences
- Balanced partitioning
  - Group the differences in order to be able to use information of the whole output space.

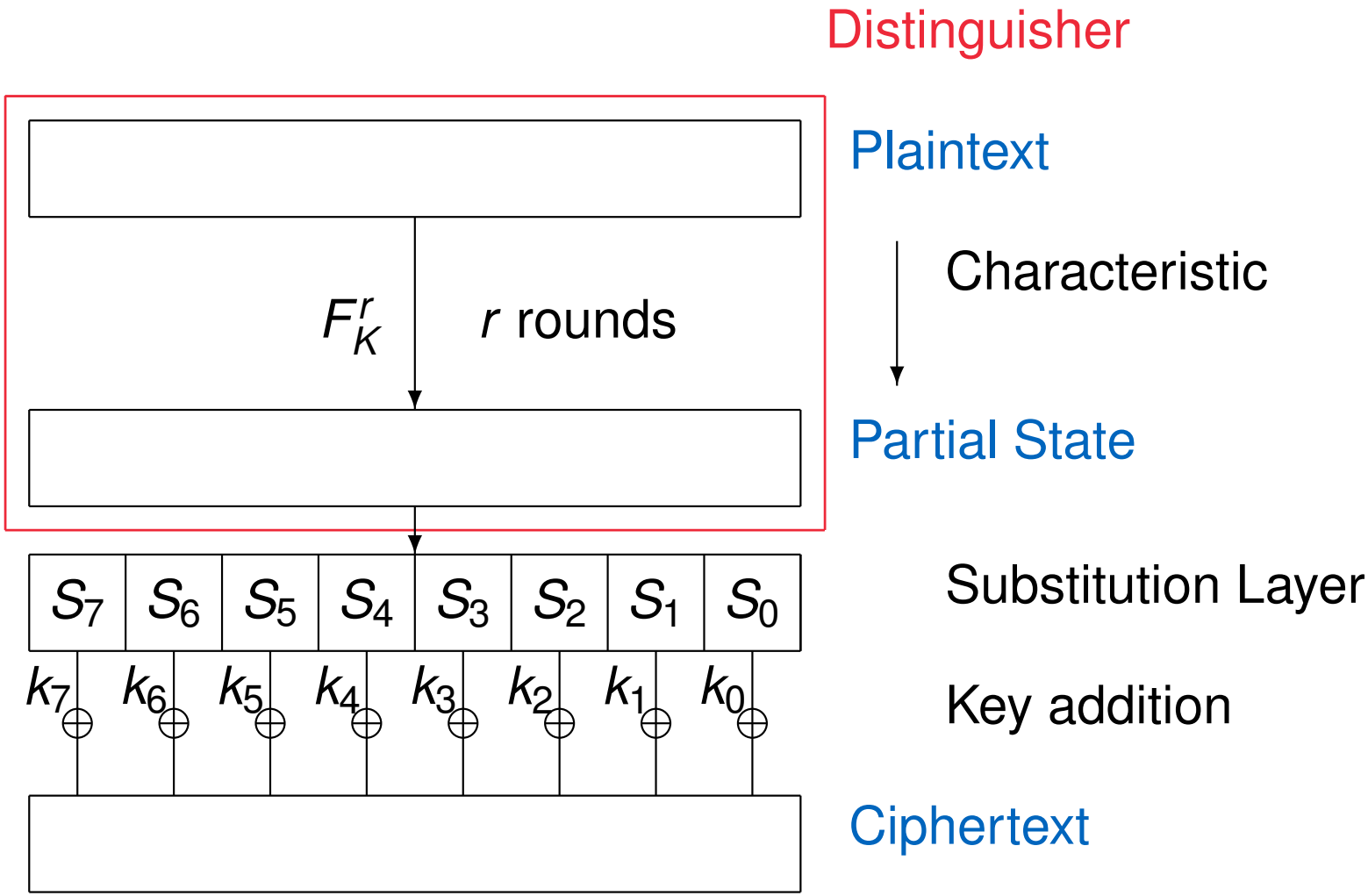
Aim:

- Compare Time, Memory and Data complexity of the different methods.

Last Round Attack



Last Round Attack



Unbalanced Partitioning

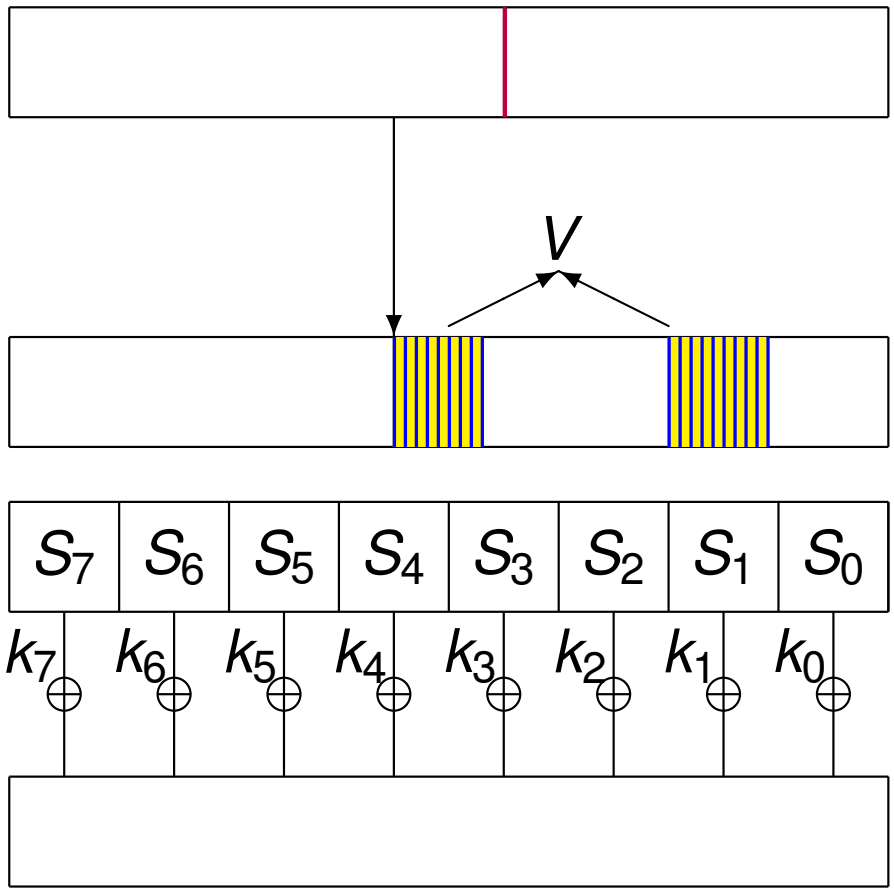
Idea: Subset of simple differences

- Output differences $(\delta_{out}^{(v)})_{1 \leq v \leq A}$,
- Counter for each of these differentials q_v^k.
- As $\sum_{i=1}^{A} q_v^k \neq 1$
- We have a "trash" counter q_0^k which gathers all other output differences.

Last Round Attack: We increment the counter q_v^k

if the difference $\delta_{out}^{(v)}$ is obtained after partial deciphering.

Unbalanced Partitioning: Last Round Attack

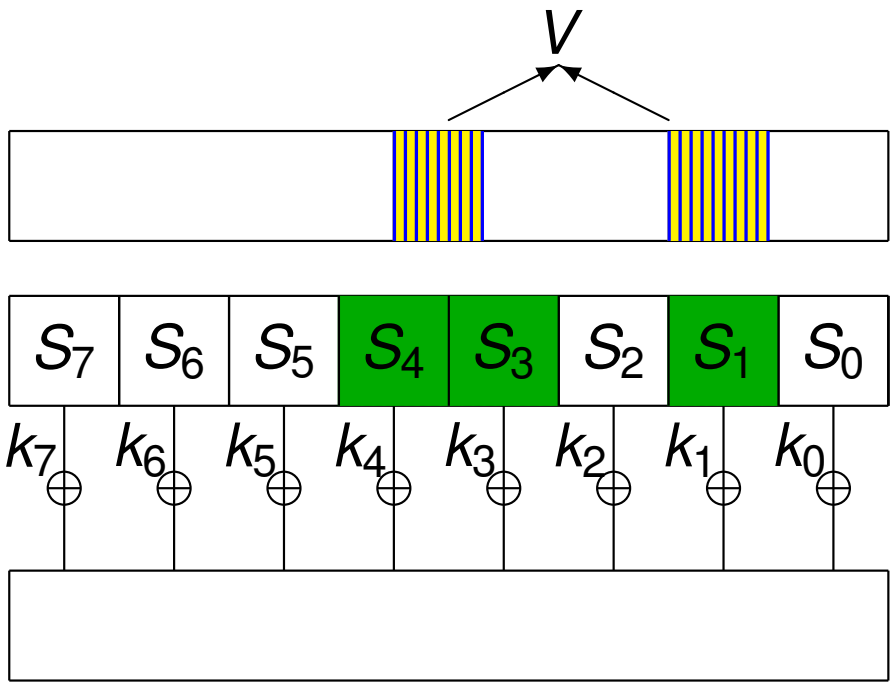


δ_{in}

$V = [\delta_{\text{out}}^{(v)}]_v$

Substitution Layer

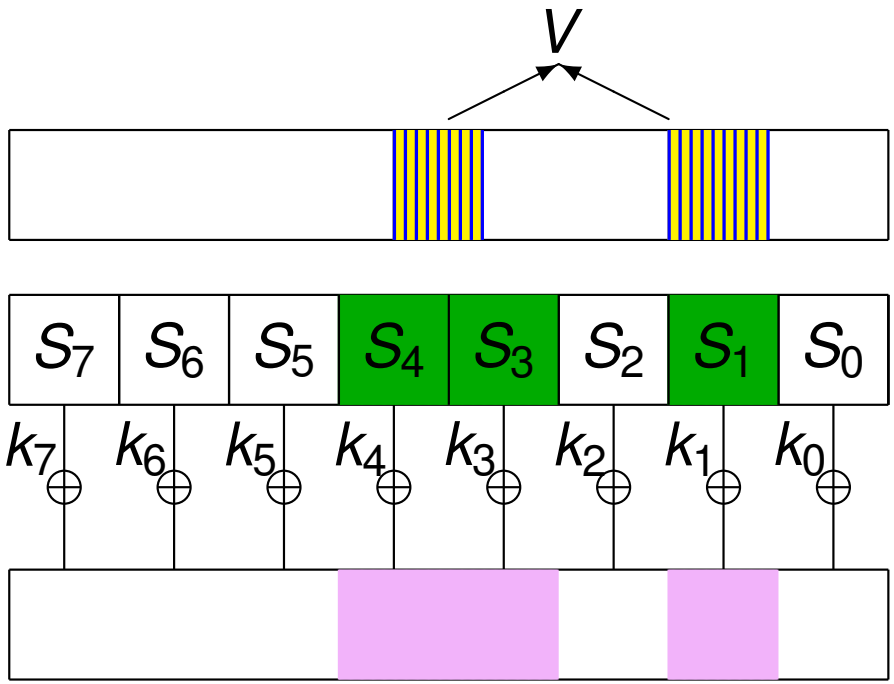
Unbalanced Partitioning: Last Round Attack



$V = [\delta_{\text{out}}^{(v)}]_v$

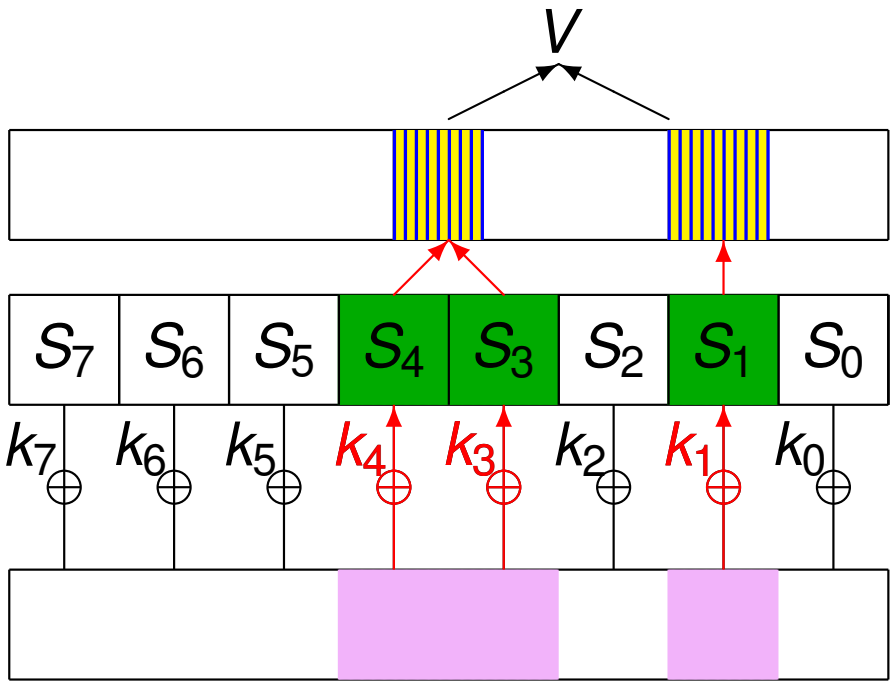
Active Sboxes

Unbalanced Partitioning: Last Round Attack



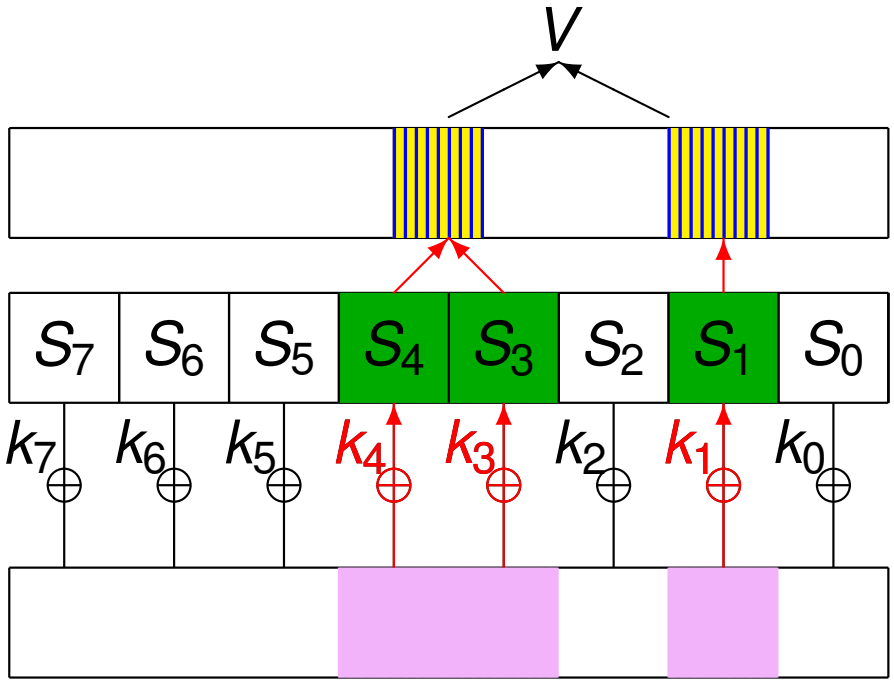
Sieving process
Discard some ciphertext pairs

Unbalanced Partitioning: Last Round Attack



For all key candidates, partially decipher

Unbalanced Partitioning: Last Round Attack



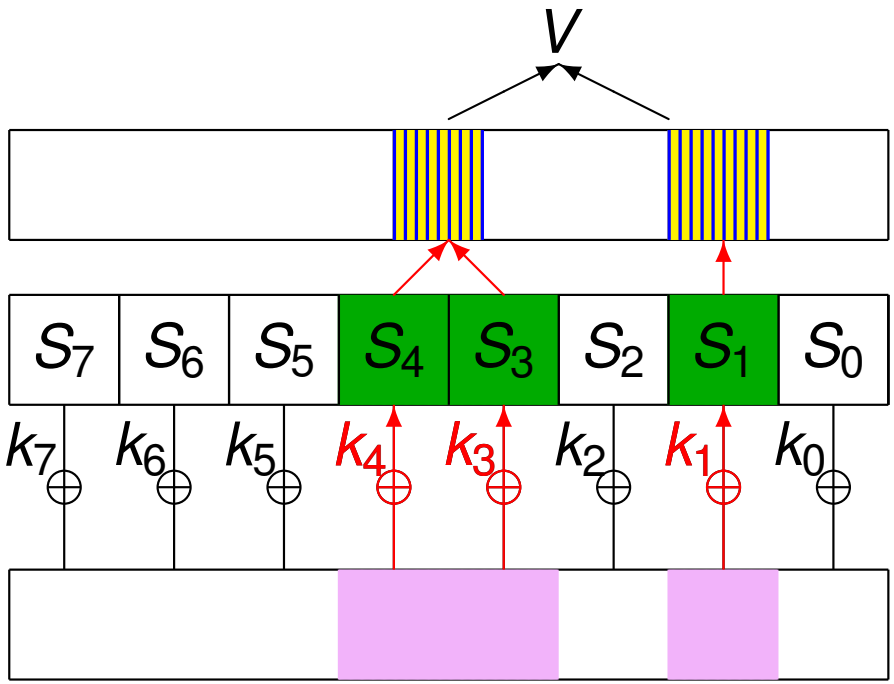
If $\delta = \delta_{\text{out}}^{(v)}$

Increment q_v^k

Otherwise

Increment q_0^k

Unbalanced Partitioning: Last Round Attack



If $\delta = \delta_{\text{out}}^{(v)}$
Increment q_v^k

Otherwise
Increment q_0^k

Analyse the vectors q^k for each key

Unbalanced Partionning: Remarks

Corresponding known/former attacks:

- Differential cryptanalysis.

Advantage:

- A sieving process $\Rightarrow$ “smaller” time complexity

Disadvantage:

- Subset of output space $\Rightarrow$ Not all information
- Small probabilities $\Rightarrow$ Non-tightness of the information

Balanced Partitioning

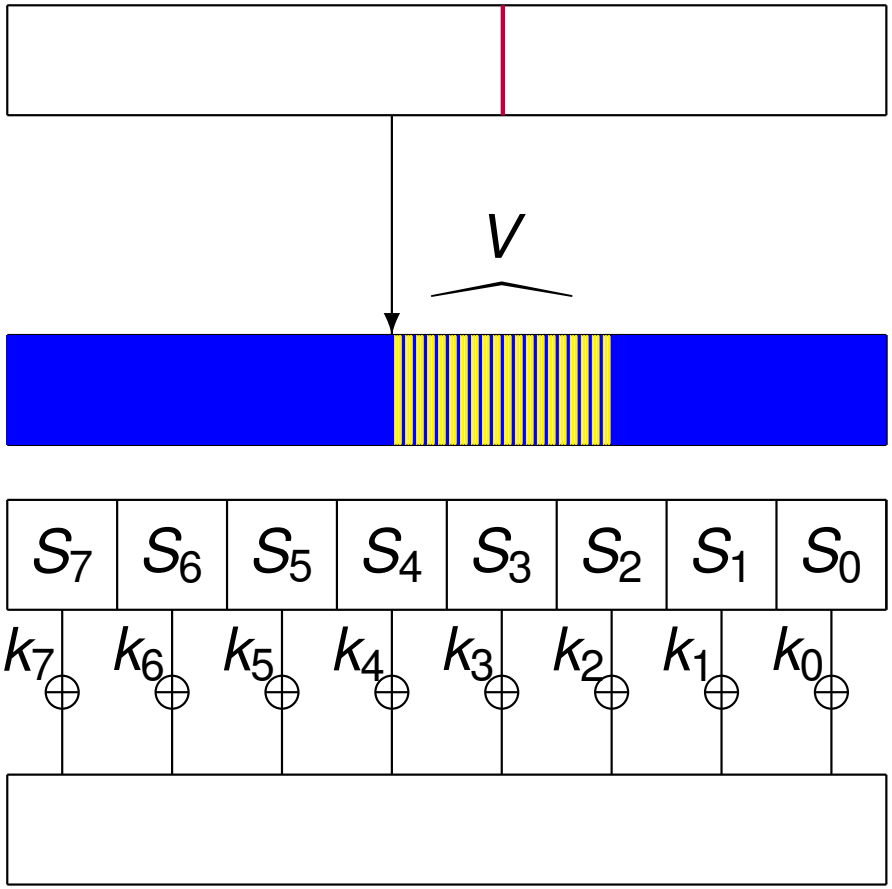
Idea: Using information from all differences by grouping them.

Let $V = [\delta_{\text{out}}^{(v)}]_v$ a subspace of $\mathbb{F}_2^m$

A group of differences $\Delta_{\text{out}}^{(v)} = \delta_{\text{out}}^{(v)} \oplus \bar{V}$ $\qquad (\bar{V} \oplus V = \mathbb{F}_2^m)$

A counter q_v^k for each group of differences.

Balanced Partitioning: Last Round Attack

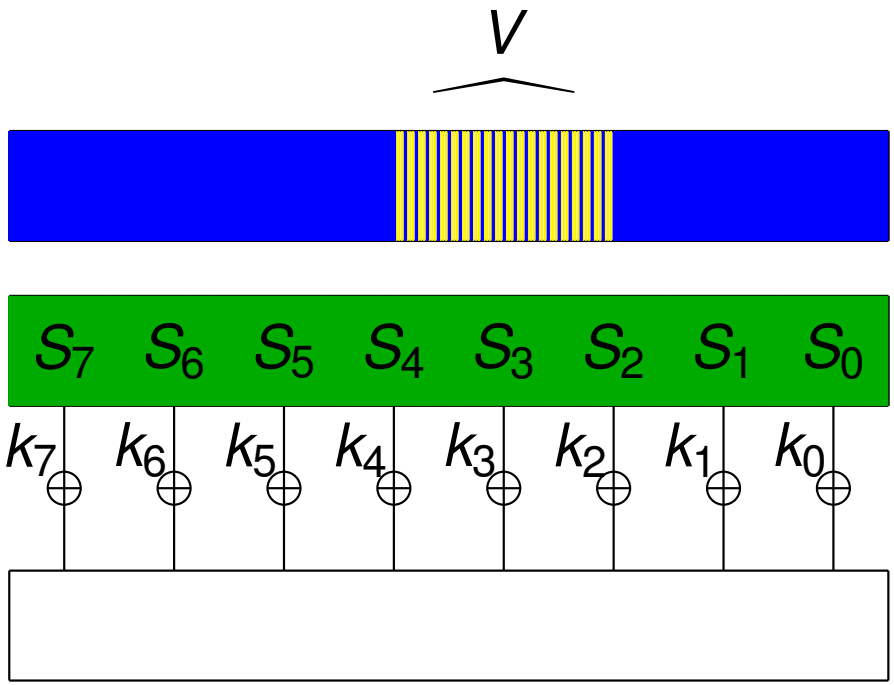


δ_{in}

$\Delta_{\text{out}} = \delta_{\text{out}} \oplus \bar{V}$

Substitution Layer

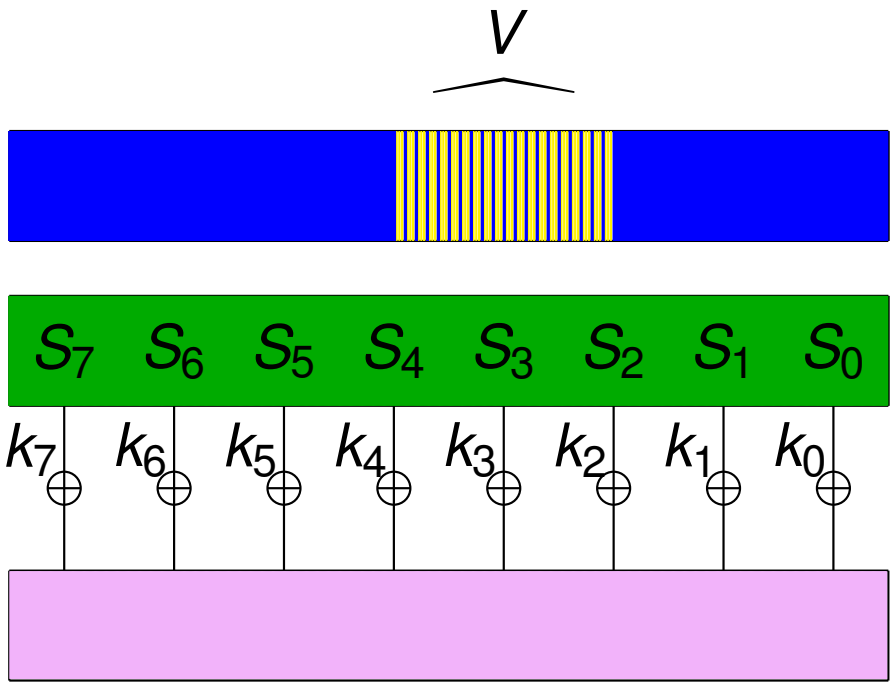
Balanced Partitioning: Last Round Attack



$$\Delta_{\text{out}} = \delta_{\text{out}} \oplus \bar{V}$$

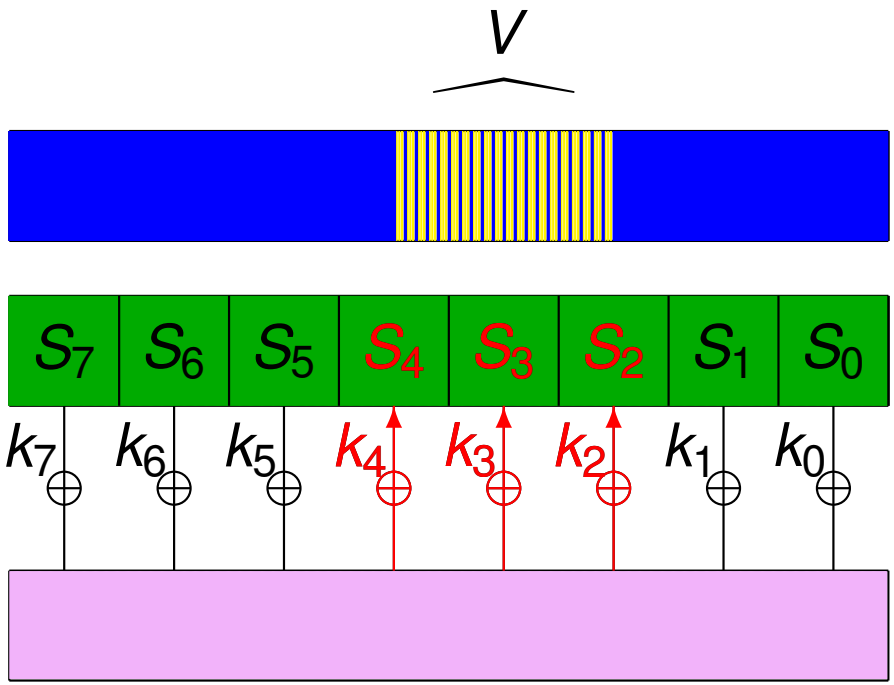
Active Sboxes

Balanced Partitioning: Last Round Attack



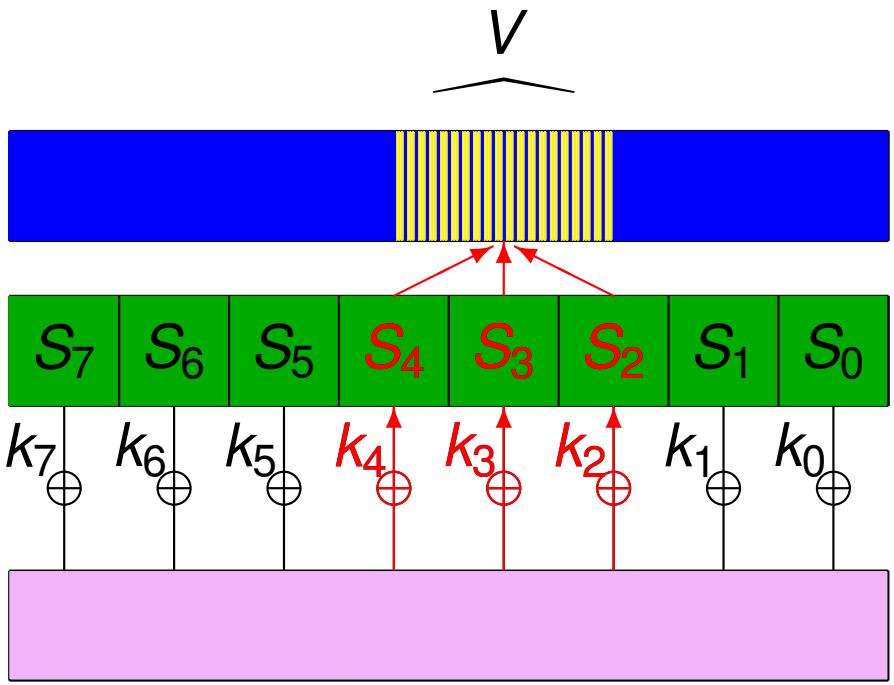
No sieving process
Partially decipher for all pairs

Balanced Partitioning: Last Round Attack



For all key candidates, partially decipher

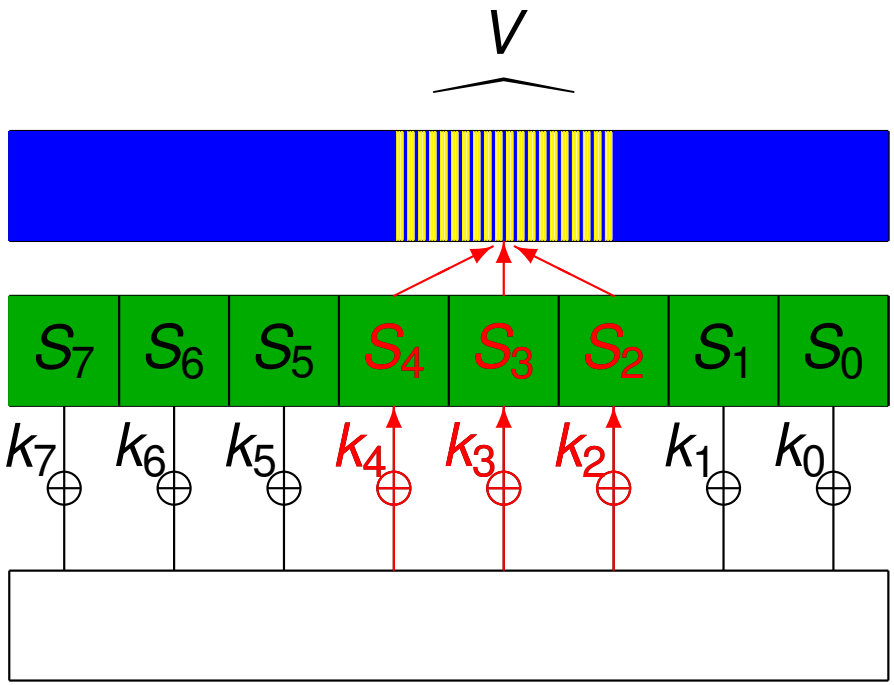
Balanced Partitioning: Last Round Attack



If $\delta \in \delta_{\text{out}}^{(v)} \oplus \bar{V}$

Increment q_v^k

Balanced Partitioning: Last Round Attack



If $\delta \in \delta_{\text{out}}^{(v)} \oplus \bar{V}$

Increment q_v^k

Analyse the vectors q^k for each key

Balanced Partitioning: Remarks

Corresponding known/former attacks:

- Truncated Differential cryptanalysis.

Advantage:

- Whole output space $\Rightarrow$ More information
- Bigger Probabilities $\Rightarrow$ Tightness of the information

Disadvantage:

- No sieving process $\Rightarrow$ Larger time complexity

Outline

Statistical Tests

Probability distribution vectors

- Expected: $p = [p_v]_{v \in V}$
- Uniform: $\theta = [\theta_v]_{v \in V}$
- Observed: q^k (for a given key candidate k)

`LLR` test: requires the knowledge of the theoretical probability p.

$$S_k = \mathtt{LLR}_k(q^k, p, \theta) \stackrel{\text{def}}{=} N_s \sum_{v \in V} q_v^k \log\left(\frac{p_v}{\theta_v}\right).$$

χ^2 test: Does not require the knowledge of p for the attack

$$S_k = \chi_k^2(q^k, \theta) = N_s \sum_{v \in V} \frac{(q_v^k - \theta_v)^2}{\theta_v}.$$

Complexities

Let $S(k)$ be the statistic obtained for a key candidate k.

$$S_k = \mathtt{LLR}_k(q^k, p, \theta) \text{ or } = \chi^2_k(q^k, \theta)$$

Then,

$$S_k \sim \begin{cases} \mathcal{N}(\mu_R, \sigma_R^2) & \text{if } k = K_r, \\ \mathcal{N}(\mu_W, \sigma_W^2) & \text{otherwise.} \end{cases}$$

[Selçuk 07]:

- Estimates of the value of $\mu_R, \mu_W, \sigma_R, \sigma_w$ for both $\mathtt{LLR}$ and χ^2 statistical tests.
- Estimates of the Data Complexity

Asymptotic Complexity when $P_S = 0.5$

a: Advantage

$\Phi_{0,1}$: cumulative function of standard normal distribution

LLR test:

$$N \approx \frac{Var_p(\log(\frac{p}{\theta}))}{\left[E_p(\log(\frac{p}{\theta})) - E_\theta(\log(\frac{p}{\theta}))\right]^2} \Phi_{0,1}^{-2}(1 - 2^{-a})$$

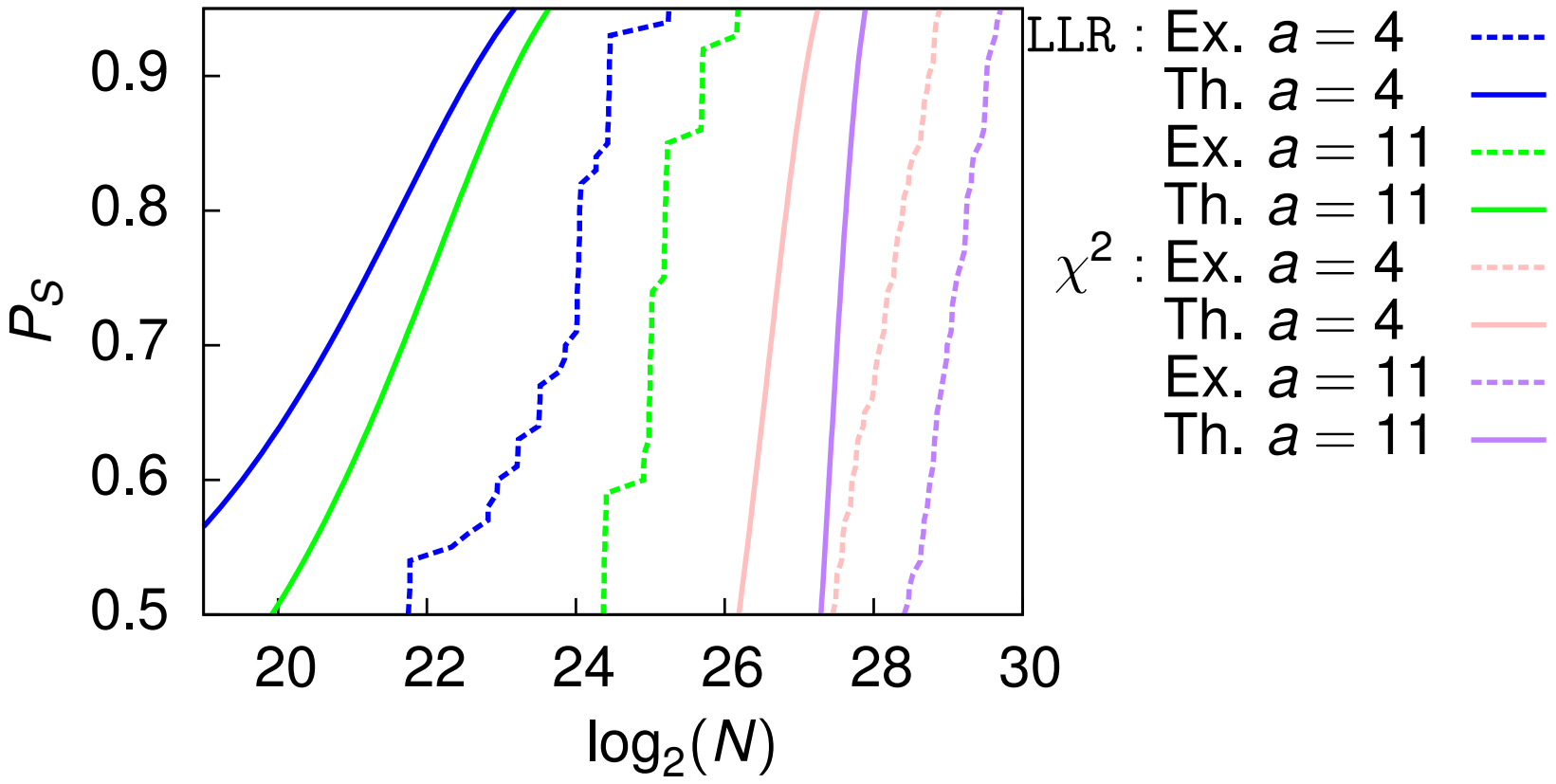
χ^2 test:

$$N \approx \frac{\sqrt{2|V|}}{C(p)} \Phi_{0,1}^{-1}(1 - 2^{-a})$$

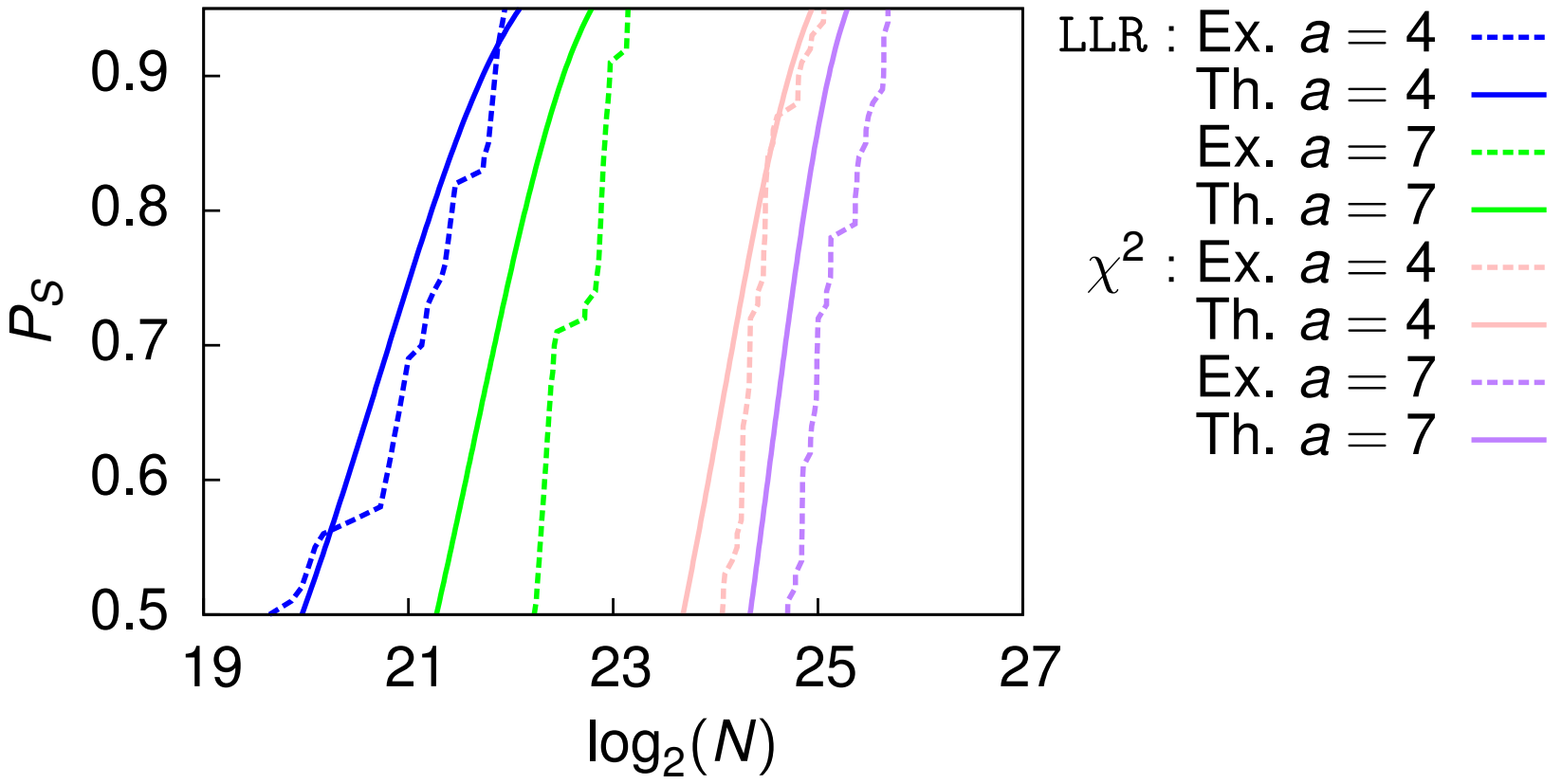
where $C(p) = \sum_{v \in V} \frac{(p_v - \theta_v)^2}{\theta_v}$

Outline

Unbalanced Partitioning: Simple differences



Balanced Partitioning: Group of output differences



Conclusions

Balanced or Unbalanced partitioning ?

- Time Complexity: unbalanced $\Rightarrow$ faster attack.
- Data Complexity: depends of the cipher.

LLR or χ^2?

- If we have a good estimate of the expected probabilities $\Rightarrow$ LLR provides better Data and Memory complexities
- Otherwise LLR is not effective