Techniques for Estimating the Data Complexity of Statistical Cryptanalyses: a Brief Overview.

Benoît Gérard

CryptoGroup - Université catholique de Louvain - Belgium

COSIC seminar - 10/02/2012

Encryption primitives

Asymmetric.

- ✓ Security proofs.
- x Lack of efficiency in many scenarios.

Symmetric.

- ✓ Efficient.
- x No security proofs*.

* *Except decorrelation theory ⇒ secure ciphers are not efficient.*

Introduction

Symmetric primitives are useful (and used)

$\Downarrow$

security analysis required

Designers provide arguments on the resistance against such and such attacks among which statistical attacks.

Outline

Statistical cryptanalysis

Basic estimates

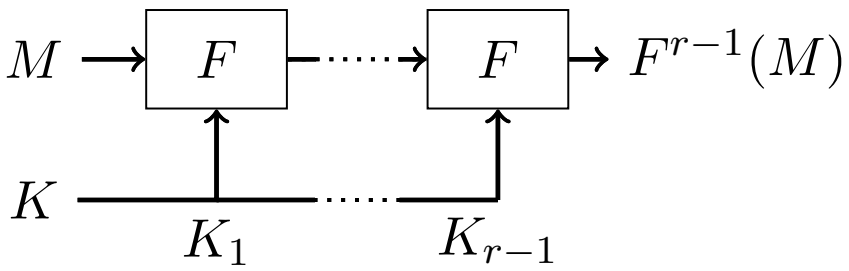
Toward general estimates

Advanced topics

Outline

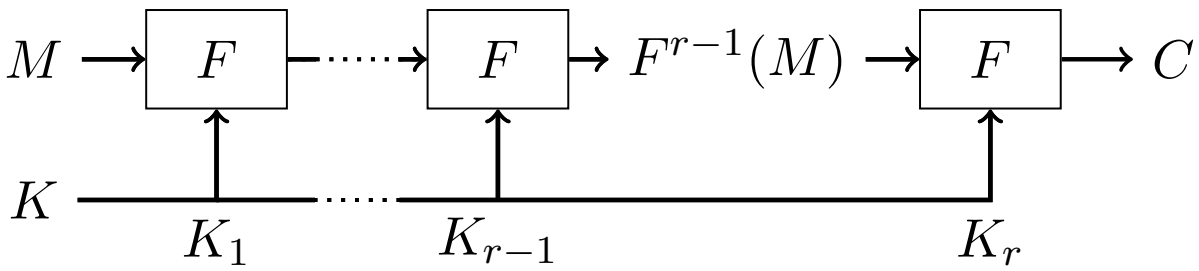
Last-round attacks

- Find a unexpected behavior ($r-1$ rounds).



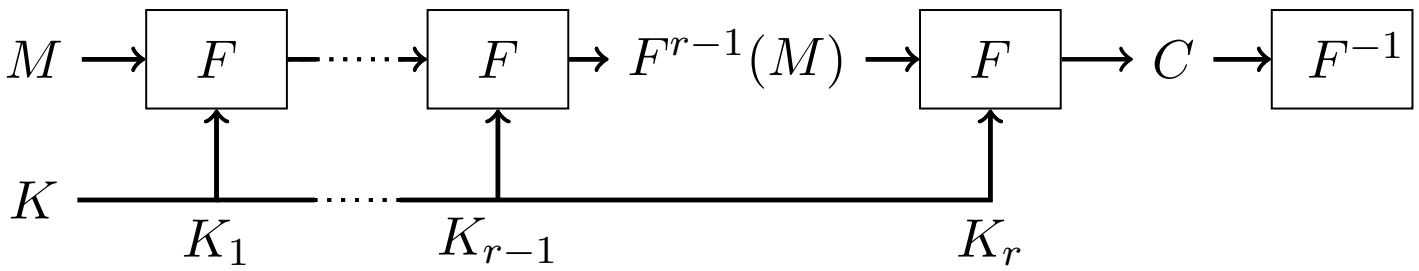
Last-round attacks

- Find a unexpected behavior ($r-1$ rounds).
- Collect r-round samples.



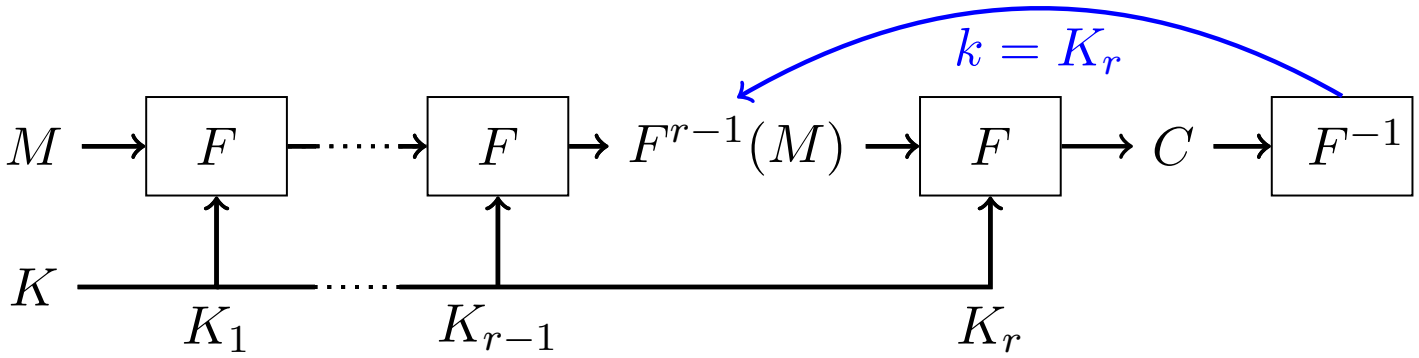
Last-round attacks

- Find a unexpected behavior ($r-1$ rounds).
- Collect r-round samples.
- Partially decrypt samples by one round.



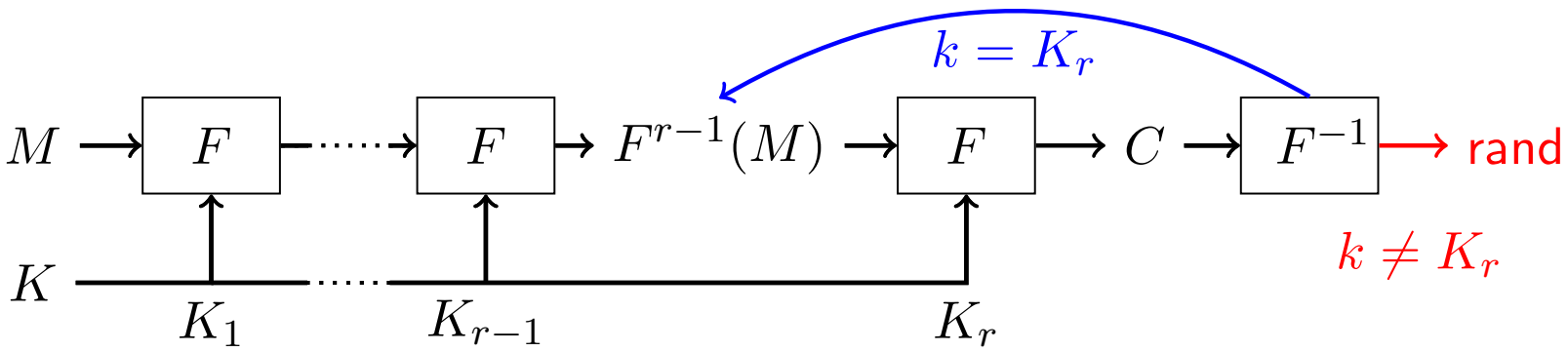
Last-round attacks

- Find a unexpected behavior ($r-1$ rounds).
- Collect r-round samples.
- Partially decrypt samples by one round.



Last-round attacks

- Find a unexpected behavior ($r-1$ rounds).
- Collect r-round samples.
- Partially decrypt samples by one round.



Wrong Key Randomization Hypothesis

Statistical cryptanalysis

Consists in 3 steps:

1. **Distillation.** Extracting relevant statistic from samples.
2. **Analysis.** Processing statistics to rank subkeys.
3. **Search.** Test subkeys until the correct one is found.

Example 1: linear cryptanalysis

- Unexpected behavior:

$$\Pr_{M,K}\left[\langle \pi, M\rangle \oplus \langle \gamma, F_K^{r-1}(M)\rangle = 0\right] = \frac{1}{2} + \varepsilon.$$

- WKRH (notion of random):

Bias obtained for wrong keys is zero.

$$p = \frac{1}{2} \quad \text{and} \quad p_* = \frac{1}{2} + \varepsilon.$$

- Statistic used for the attack:

$$\#\Big\{(m,c) | \langle \pi, m\rangle \oplus \langle \gamma, F_k^{-1}(c)\rangle = 0\Big\}.$$

Example 2: differential cryptanalysis

- Unexpected behavior:

$$\Pr_{M,K}\left[F_K^{r-1}(M) \oplus F_K^{r-1}(M \oplus \delta_1) = \delta_2\right] = p_*.$$

- WKRH (notion of random):

  Differences are uniformly distributed over $\mathbb{F}_2^s \backslash \{0\}$.

$$p = \frac{1}{2^s - 1} \approx 2^{-s} \quad \text{and} \quad p_* > 2^{-s}.$$

- Statistic used for the attack:

$$\#\Big\{(m_1, m_2, c_1, c_2) | m_2 \oplus m_1 = \delta_1 \text{ and } F_k^{-1}(c_1) \oplus F_k^{-1}(c_2) = \delta_2\Big\}.$$

Data complexity

Statistical attacks are based on detecting an unexpected behavior.

How many samples are required to detect this behavior?

The answer depends on the available computational power in the offline phase.

Outline

Linear cryptanalysis I

- [Matsui, EUROCRYPT 1993]: *Heavy formula based on Gaussian law.*

Idea:

The statistics considered follow a binomial distribution that can be approximated by a Gaussian law.

Linear cryptanalysis I

- [Matsui, EUROCRYPT 1993]: *Heavy formula based on Gaussian law.*
- [Junod, SAC 2001]: *Gaussian law + Beta distribution.*

Idea:

Using order statistics simplifies Matsui's formula.

Linear cryptanalysis I

- [Matsui, EUROCRYPT 1993]: *Heavy formula based on Gaussian law.*
- [Junod, SAC 2001]: *Gaussian law + Beta distribution.*
- [Selçuk , JoC 2008]: *Compact formula (Gaussian law).*

Idea:

Order statistics distributions tend toward the Gaussian one $\rightarrow$ more simplifications.

Linear cryptanalysis I

- [Matsui, EUROCRYPT 1993]: *Heavy formula based on Gaussian law.*
- [Junod, SAC 2001]: *Gaussian law + Beta distribution.*
- [Selçuk , JoC 2008]: *Compact formula (Gaussian law).*

Folklore:

$$N = O\left(\varepsilon^{-2}\right).$$

Differential cryptanalysis I

- [Biham and Shamir, JoC 1991]: *rule of thumb.*

Idea:

If $p_* \gg p$ then observing one good pair is enough $\Rightarrow$ roughly $1/p_*$ samples is ok.

Differential cryptanalysis I

- [Biham and Shamir, JoC 1991]: *rule of thumb.*
- [Gilbert, PhD thesis 1997]: *Poisson distribution.*

Idea:

The statistics considered follow a binomial distribution that can be approximated by a Poisson distribution if $p_* > 4\,p$.

Differential cryptanalysis I

- [Biham and Shamir, JoC 1991]: *rule of thumb.*
- [Gilbert, PhD thesis 1997]: *Poisson distribution.*
- [Selçuk , JoC 2008]: *Compact formula (Gaussian law).*

Idea:

The statistics considered follow a binomial distribution that can be approximated by a Gaussian law . . . or not.

Differential cryptanalysis I

- [Biham and Shamir, JoC 1991]: *rule of thumb.*
- [Gilbert, PhD thesis 1997]: *Poisson distribution.*
- [Selçuk , JoC 2008]: *Compact formula (Gaussian law).*

Folklore:

$$N = O\left(p_*^{-1}\right).$$

Question:

Single approach for all statistical attacks?

Outline

Kullback-Leibler divergence

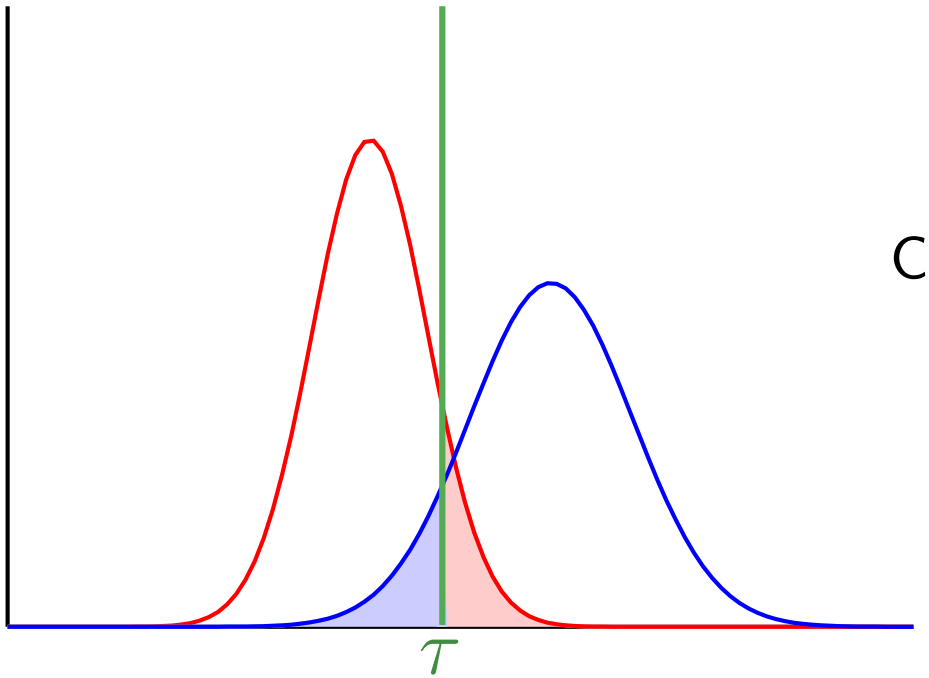
For two probability functions p and q:

$$D(p||q) = \sum_x p(x) \ln \left(\frac{p(x)}{q(x)} \right).$$

- Also known as *relative entropy*.
- "Distance" between two distributions.

Binary hypothesis testing

$$\alpha = \Pr\left[\Sigma_{K_r} < \tau\right] \quad , \quad \beta = \Pr\left[\Sigma_k \geq \tau\right].$$



Cryptanalysis settings:

- $1 - \alpha$ non negligible;
- β the smallest possible.

$$P_*(x) = \Pr\left[\Sigma_{K_r} = x\right] \quad , \quad P(x) = \Pr\left[\Sigma_k = x\right].$$

Theorem (asymptotic)

If $N = \dfrac{d}{2D(P_*||P)}$,

$$P_e = \alpha + \beta \approx \Phi(-\sqrt{d}/2).$$

- Asymptotic: convergence of the log likelihood ratio (LLR).
- Not really suited for cryptanalysis: $\beta \ll \alpha$.

- attacks where P_* and P are binomial distributions.
- error probabilities are of the form $Q(N)2^{-L(N)}$.

Last slide: $L(N) = N \cdot D(P_*||P)$.

Estimate for N (α fixed to 0.5 and small enough β)

$$N \approx -\frac{\log(2\sqrt{\pi}\beta)}{D(P_*||P)}.$$

$+$ formula to compute α for N and β fixed.

Outline

Facing reality

Carefully designed ciphers $\Rightarrow$ pushing to the limits.

- Using quasi-indistinguishable phenomenons.
- Using more than one statistical phenomenon.

Pushing to the limits $\Rightarrow$ new problems appear.

- The distribution P_* depends on the key.
- Wrong Key Randomization Hypothesis is not well stated.

Multiple differential cryptanalysis

Combination technique.
Scores obtained for differentials are added.

Different probabilities $\Rightarrow$ P_* and P not binomial anymore.

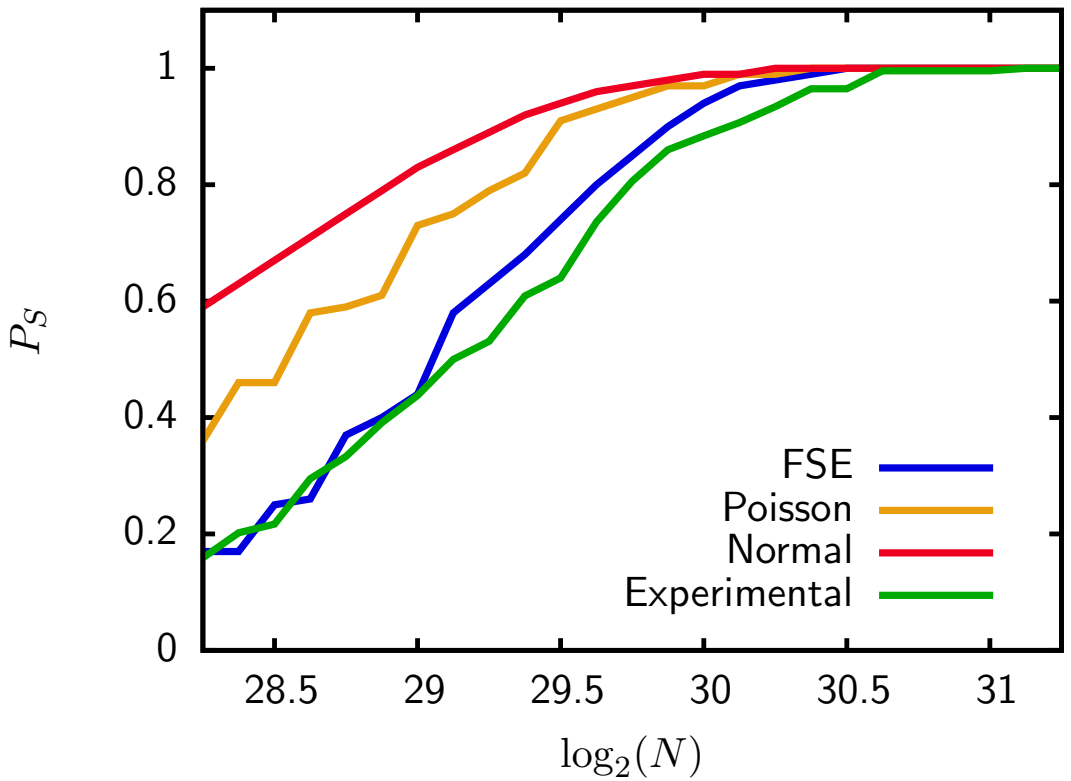
[Blondeau and Gérard, FSE 2011]

Under a restrictive assumption on the differentials used:

- Estimate for the (tail of) sum of binomial distributions.
- Formula for the success probability as a function of N.

Multiple differential cryptanalysis

Fixed number of tested keys.



Using many linear approximations

Enough papers for a single 2-hours talk.

I will only consider:

- [Biryukov *et al.*, CRYPTO 2004].
  any set of approximations, assuming statistical independence.
- [Hermelin *et al.*, 2008-2010].
  structured sets of approximations, takes profit of statistical dependence.

Linear cryptanalysis: multiple vs multidimensional

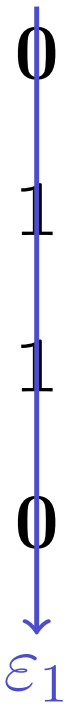
	(π_1, γ_1)	(π_2, γ_2)	(π_3, γ_3)
(m_1, c_1)	**0**	**1**	**0**
(m_2, c_2)	**1**	**0**	**1**
(m_3, c_3)	**1**	**0**	**1**
(m_4, c_4)	**0**	**0**	**1**

Linear cryptanalysis: multiple vs multidimensional

multiple	(π_1, γ_1)	(π_2, γ_2)	(π_3, γ_3)
(m_1, c_1)	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{0}$
(m_2, c_2)	$\mathbf{1}$	$\mathbf{0}$	$\mathbf{1}$
(m_3, c_3)	$\mathbf{1}$	$\mathbf{0}$	$\mathbf{1}$
(m_4, c_4)	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{1}$

Linear cryptanalysis: multiple vs multidimensional

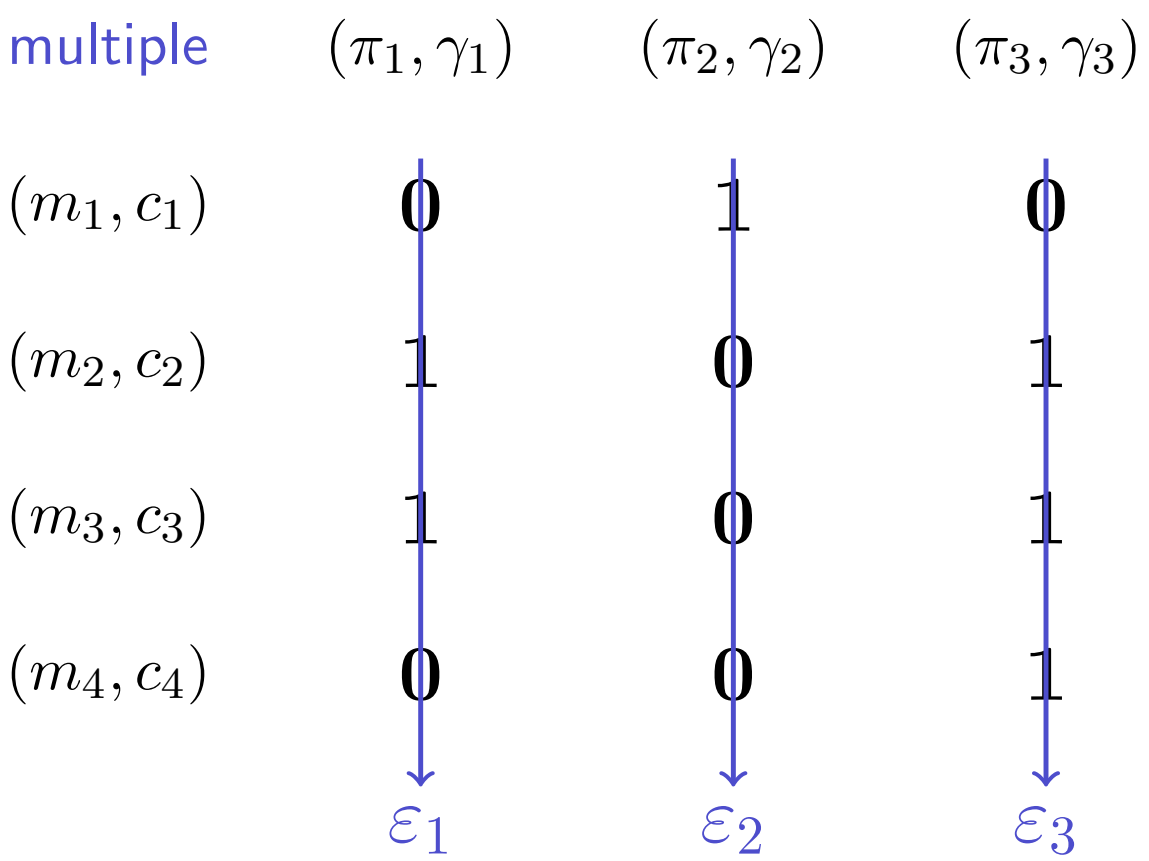
multiple	(π_1, γ_1)	(π_2, γ_2)	(π_3, γ_3)
(m_1, c_1)	**0**	**1**	**0**
(m_2, c_2)	1	**0**	**1**
(m_3, c_3)	1	**0**	**1**
(m_4, c_4)	**0**	**0**	**1**



Linear cryptanalysis: multiple vs multidimensional

multiple	(π_1, γ_1)	(π_2, γ_2)	(π_3, γ_3)
(m_1, c_1)	**0**	1	**0**
(m_2, c_2)	1	**0**	**1**
(m_3, c_3)	1	**0**	**1**
(m_4, c_4)	**0**	**0**	**1**
	ε_1	ε_2	

Linear cryptanalysis: multiple vs multidimensional



Linear cryptanalysis: multiple vs multidimensional

multidim.	(π_1, γ_1)	(π_2, γ_2)	(π_3, γ_3)
(m_1, c_1)	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{0}$
(m_2, c_2)	$\mathbf{1}$	$\mathbf{0}$	$\mathbf{1}$
(m_3, c_3)	$\mathbf{1}$	$\mathbf{0}$	$\mathbf{1}$
(m_4, c_4)	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{1}$

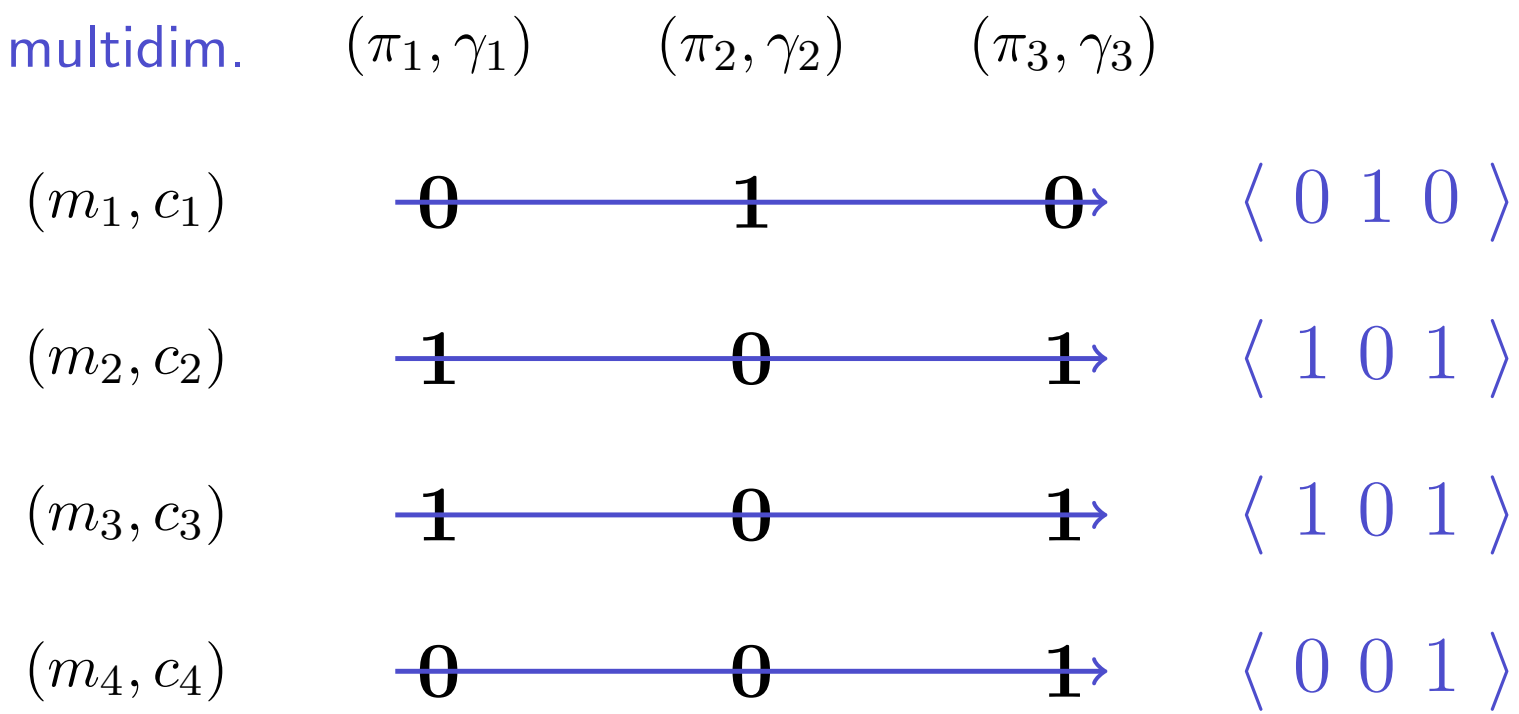
Linear cryptanalysis: multiple vs multidimensional

multidim.	(π_1, γ_1)	(π_2, γ_2)	(π_3, γ_3)	
(m_1, c_1)	**0**	**1**	**0**	$\langle\, 0\ 1\ 0 \,\rangle$
(m_2, c_2)	**1**	**0**	**1**	
(m_3, c_3)	**1**	**0**	**1**	
(m_4, c_4)	**0**	**0**	**1**	

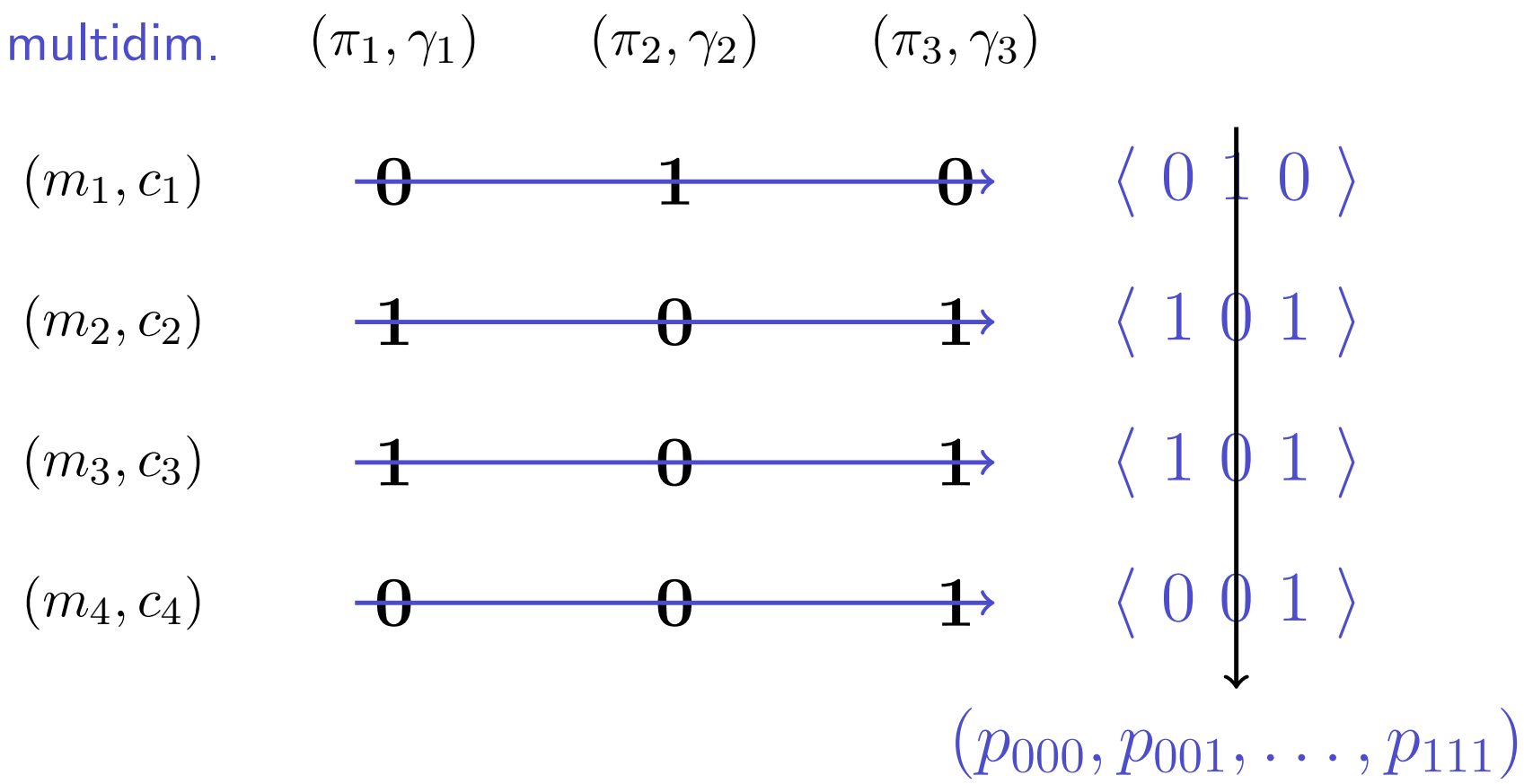
Linear cryptanalysis: multiple vs multidimensional

multidim.	(π_1, γ_1)	(π_2, γ_2)	(π_3, γ_3)	
(m_1, c_1)	**0**	**1**	**0**	$\langle\, 0\ 1\ 0 \,\rangle$
(m_2, c_2)	**1**	**0**	**1**	$\langle\, 1\ 0\ 1 \,\rangle$
(m_3, c_3)	**1**	**0**	**1**	
(m_4, c_4)	**0**	**0**	**1**	

Linear cryptanalysis: multiple vs multidimensional



Linear cryptanalysis: multiple vs multidimensional



Multiple linear cryptanalysis

- Framework based on the likelihood of subkeys.
- Extracting vector of biases from samples.
- Likelihoods determined by the Euclidean distance to a theoretical vector.
- Gain: metric based on the mean rank of the key.

Remarks in the literature:

- underlying the framework: Jensen inequality is an equality.
- considering mean rank of the correct key is pessimistic.

Multidimensional linear cryptanalysis

- Boolean functions theory approach.
- Extracting small vectors of parities from samples.
- Scores obtained by comparing empirical distributions of these vectors to theoretical ones.
  - LLRs asymptotically normally distributed.
  - χ^2 distribution quickly tends toward Gaussian one.

Counterparts:

- restriction on approximations used.
- theoretical estimates are not so good.

An entropy based approach

Main idea

Estimating mutual information between key and extracted statistics.

Data complexity for the correct key to be top-ranked among $2^{n_{key}}$ candidates:

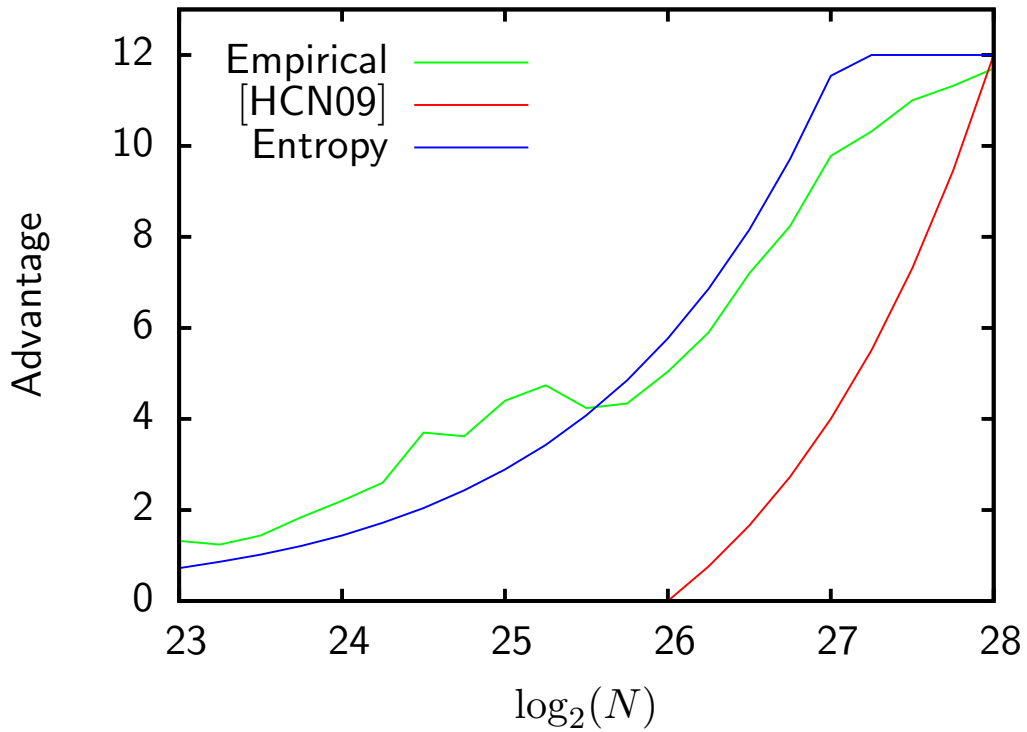
x Gain $\rightarrow N \approx \frac{n_{key}+1}{\sum_j \varepsilon_j^2}$.

✓ Entropy $\rightarrow N \approx \frac{n_{key}}{2\sum_j \varepsilon_j^2}$.

An entropy based approach

Main idea

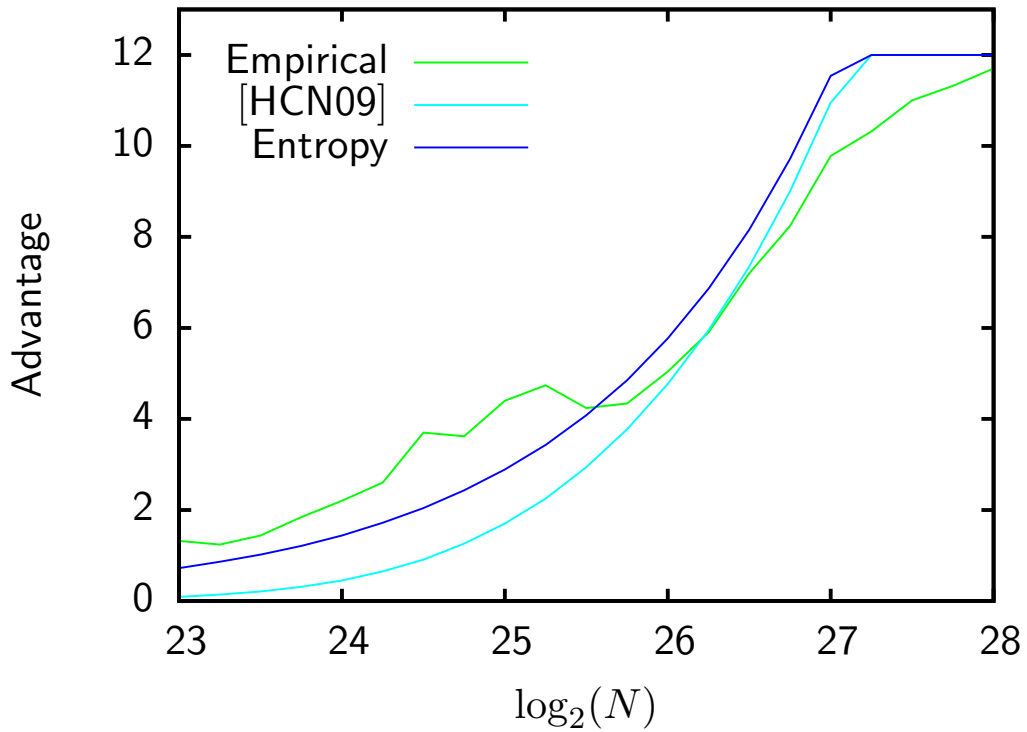
Estimating mutual information between key and extracted statistics.



An entropy based approach

Main idea

Estimating mutual information between key and extracted statistics.



Key dependence

Differential cryptanalysis.

[Daemen and Rijmen, JMC 2007]

$$p_* \sim \mathcal{P}oisson(\cdot)$$

Linear cryptanalysis.

[Daemen and Rijmen, JMC 2007]

$$p_* \sim \mathcal{N}orm(\cdot, \cdot)$$

[Röck and Nyberg, WCC 2010]
Exploiting the explicit dependence on the key.

WKRH formulation

WKRH: using a wrong value for k we obtain a random (vectorial) Boolean function.

Based on [Daemen and Rijmen, JMC 2007],

- most of wrong keys will lead to non-zero biases;
- some wrong keys will result in differential probabilities larger than $2^{-(s-1)}$.

Elmar presented its joint work with Andrey during the last PhD Day. They obtained new estimates for linear cryptanalysis

Andrey and Vincent proposed attacks using linear approximations having bias 0.