Side channel attacks based on linear approximations

Thomas Roche, Cédric Tavernier

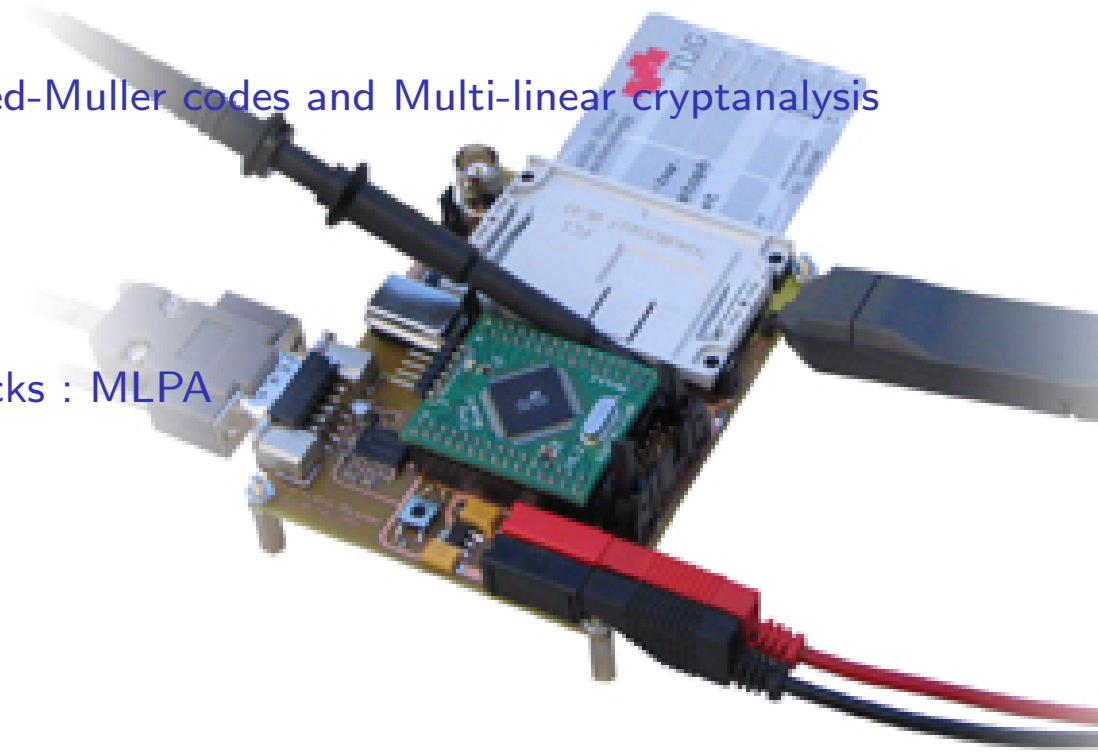
Laboratoire LIG, Grenoble, France.

Communications and Systems, Le Plessis Robinson, France.

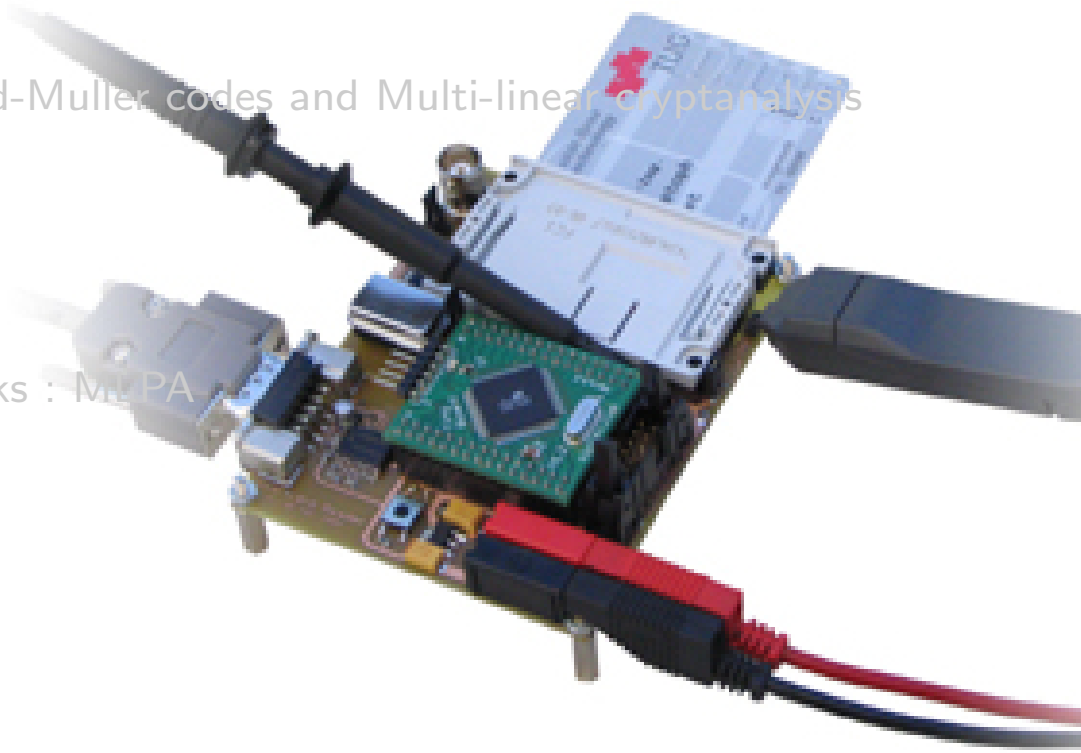
First International Alternative Workshop on Aggressive Computing and Security

ESIEA - Laval, 23-25 Oct. 2009

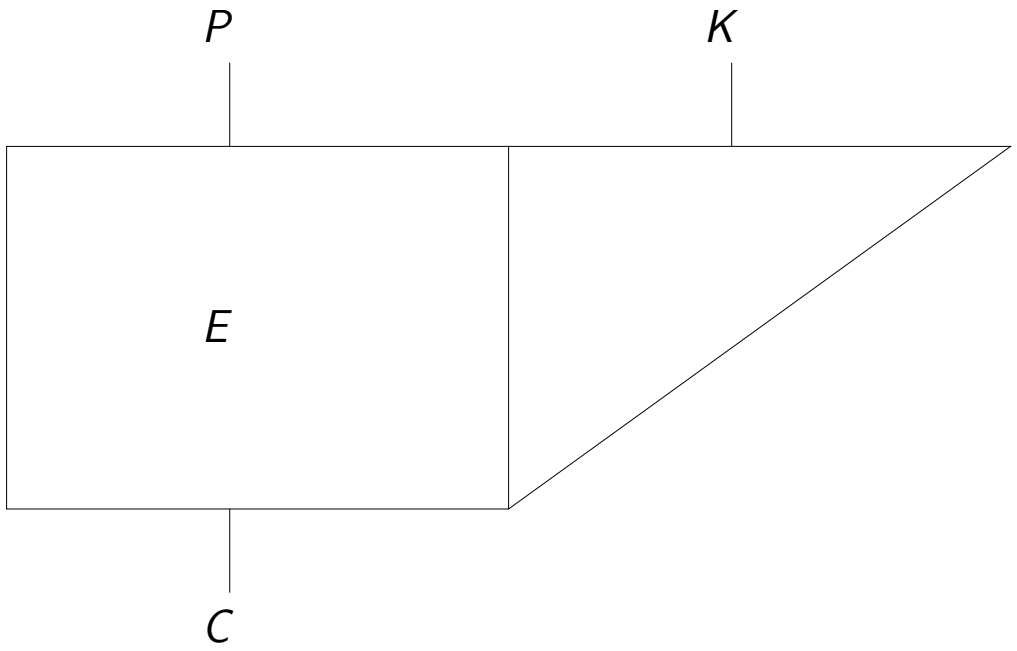
Outline



Outline

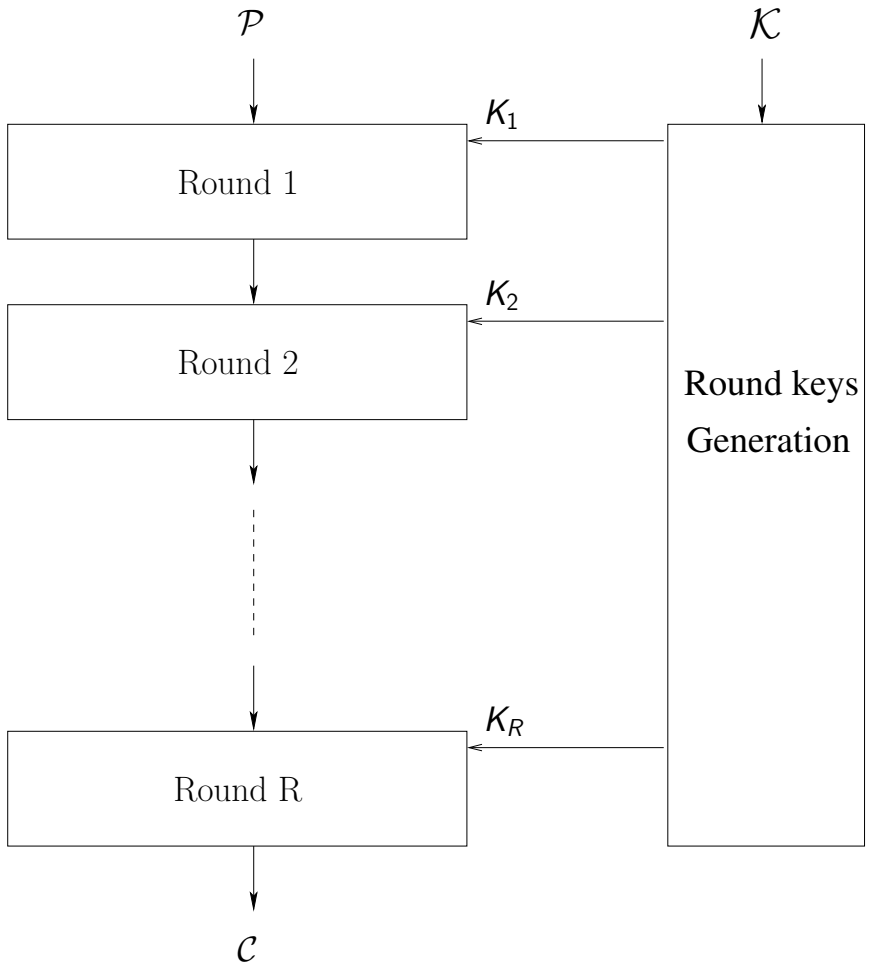


Block cipher

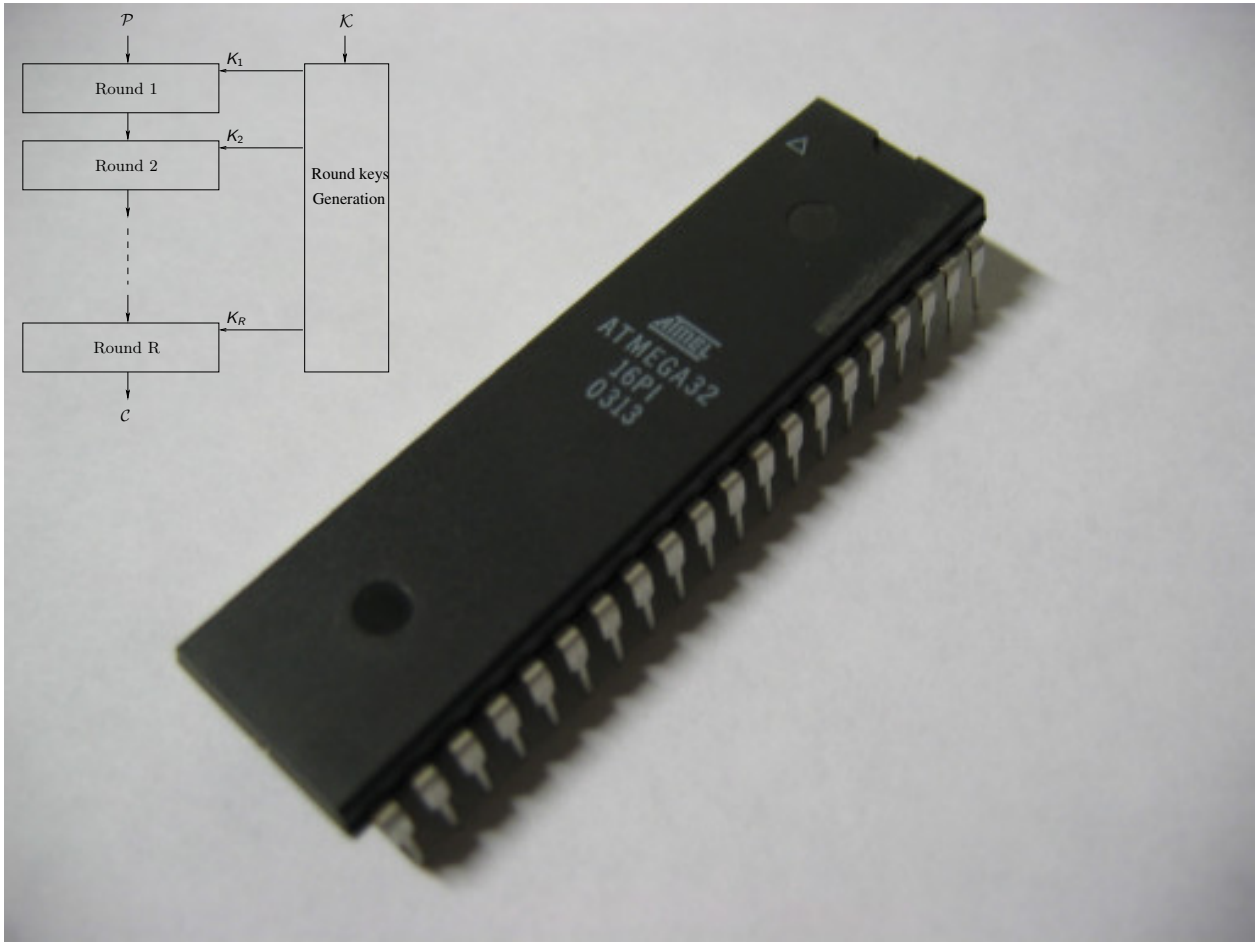


Iterated block cipher : A Mathematical object

- DES (1974)
  - $\mathcal{P}$: 64 bits
  - $\mathcal{K}$: 56 bits
  - best attack $\sim 2^{56}$ (brute force)
- AES (2000)
  - $\mathcal{P}$: 128 bits
  - $\mathcal{K}$: 128/192/256 bits
  - best attack $\sim 2^{128}/2^{192}/2^{256}$ (brute force)

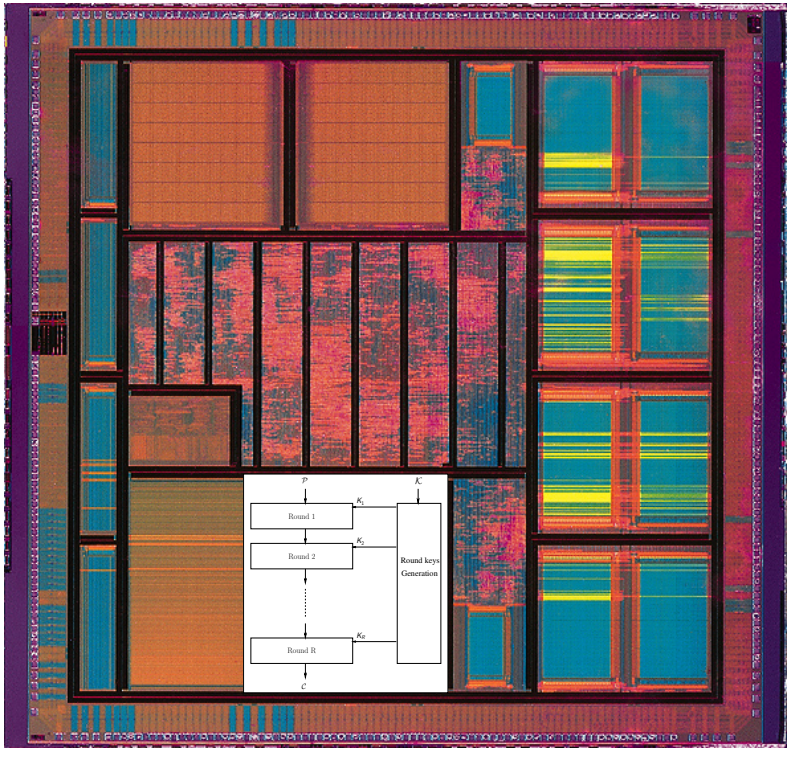


Iterated block cipher : Software Implementation

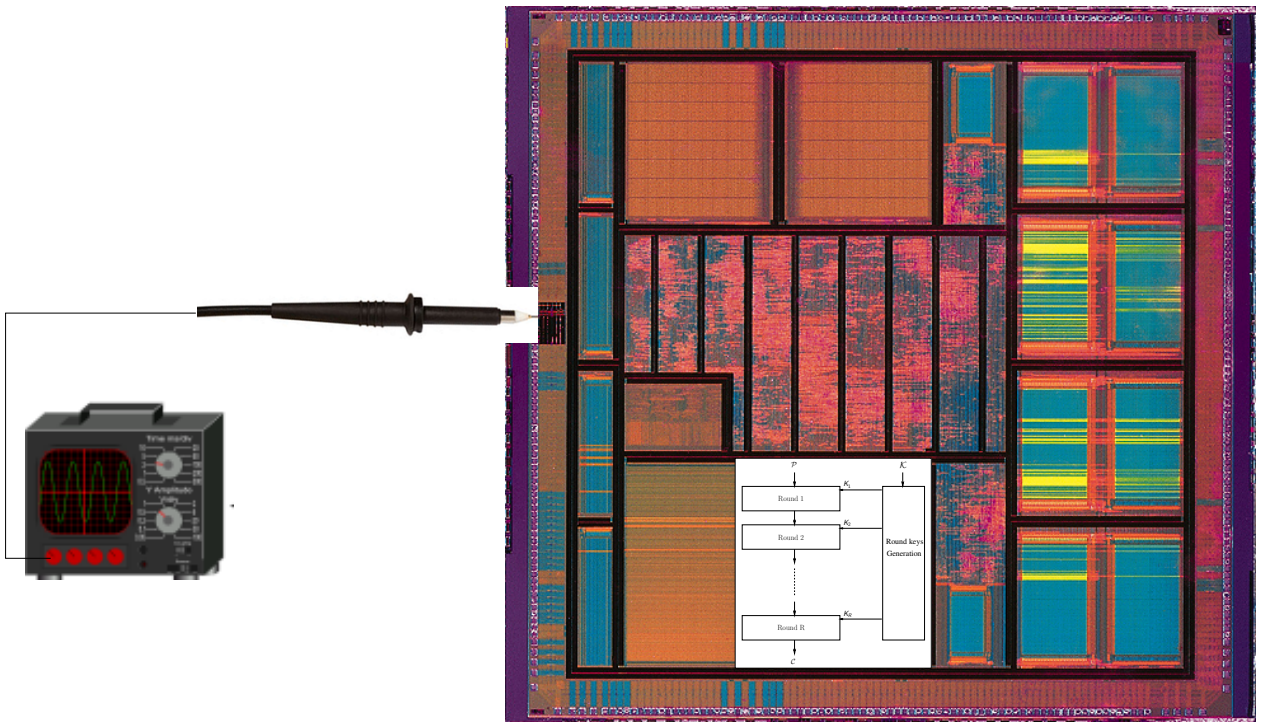


Iterated block cipher : Hardware Implementation

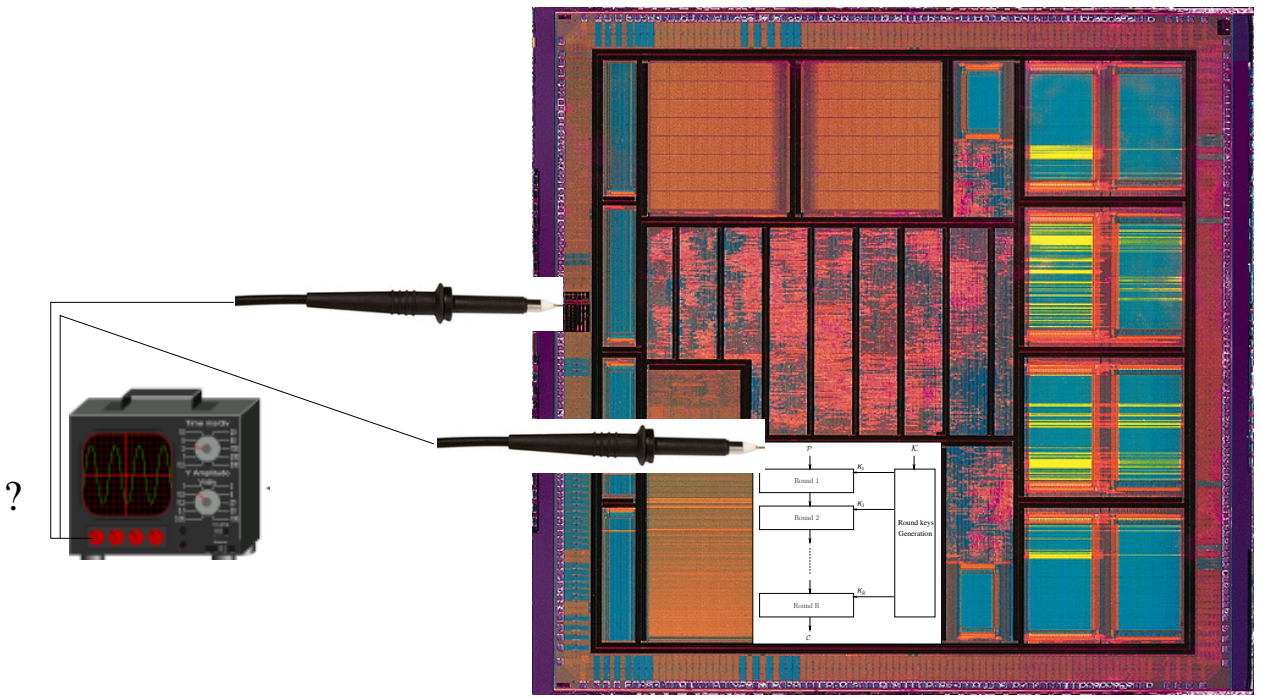
- Hardware : ASIC, FPGA
  - asynchronous vs synchronous implementation
  - iterative vs unrolled implementation



Side channel attack using consumption leakage



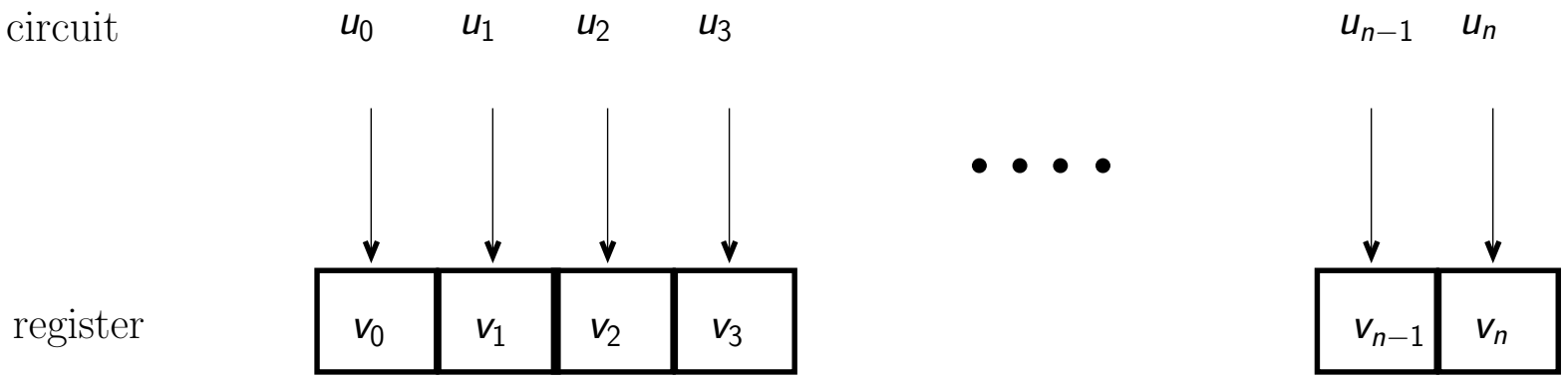
Side channel attack using consumption leakage



Side channel attack using consumption leakage

Consumption Model

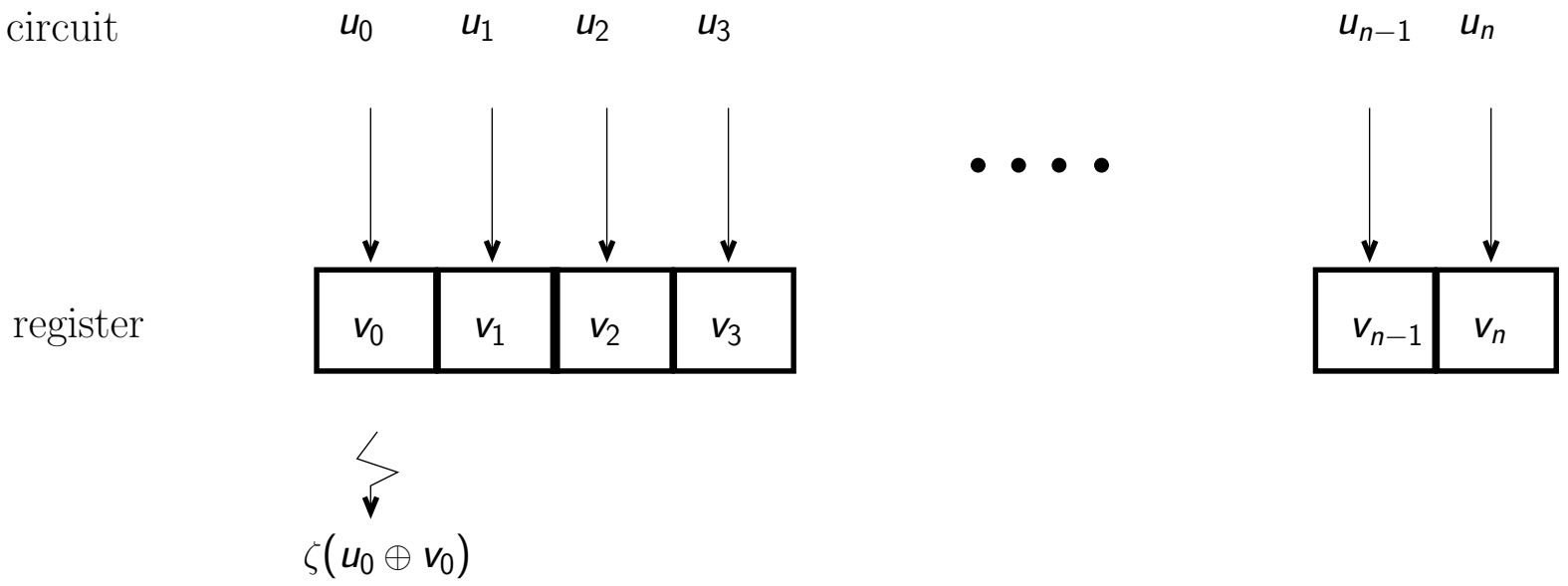
A CMOS transistor consumption is high during a bit-flip otherwise it is negligible.



Side channel attack using consumption leakage

Consumption Model

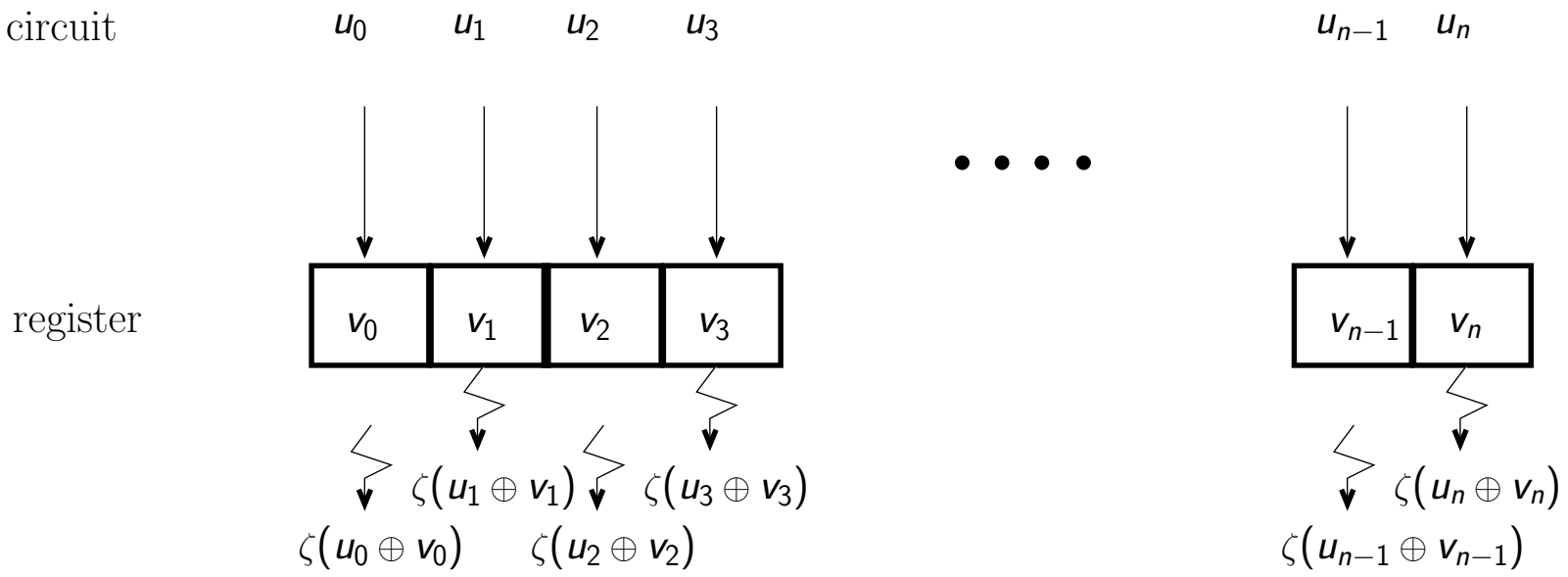
A CMOS transistor consumption is high during a bit-flip otherwise it is negligible.



Side channel attack using consumption leakage

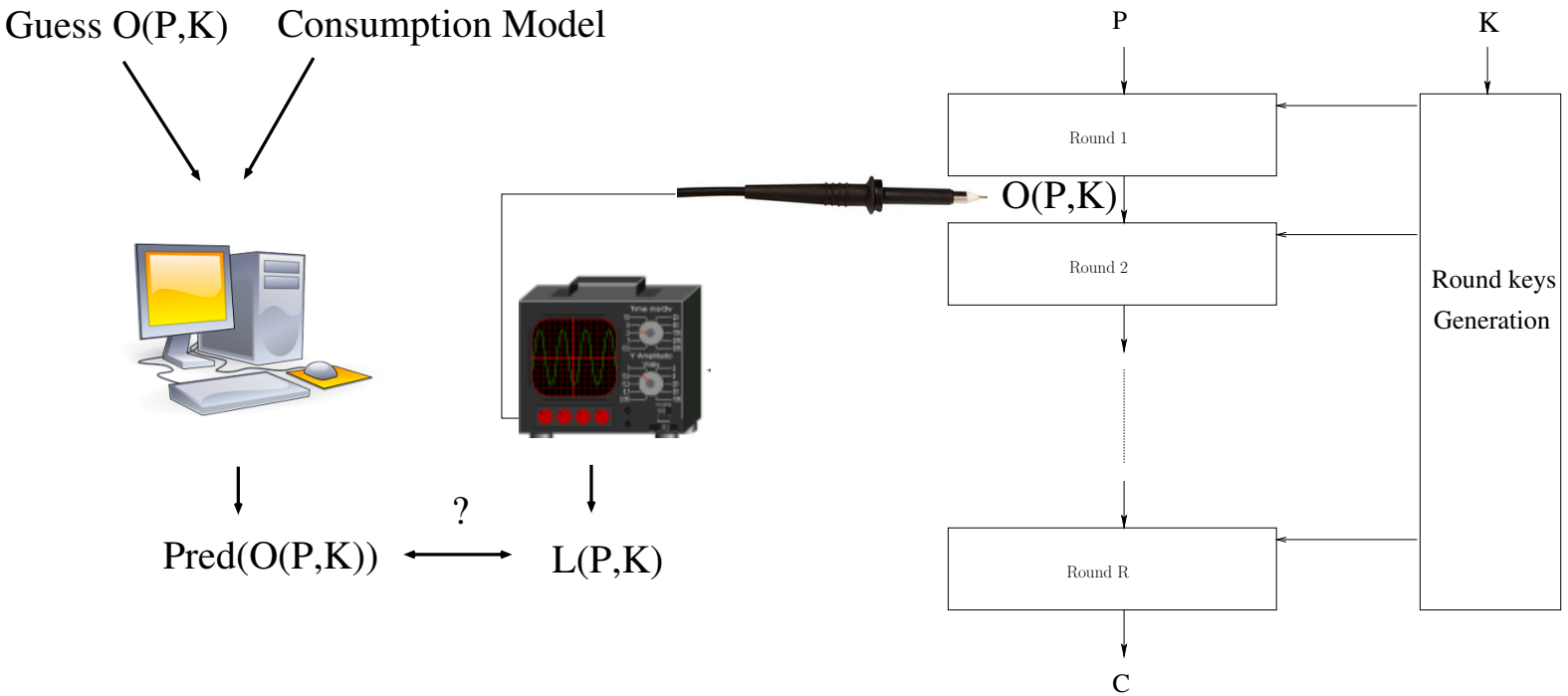
Consumption Model

A CMOS transistor consumption is high during a bit-flip otherwise it is negligible.

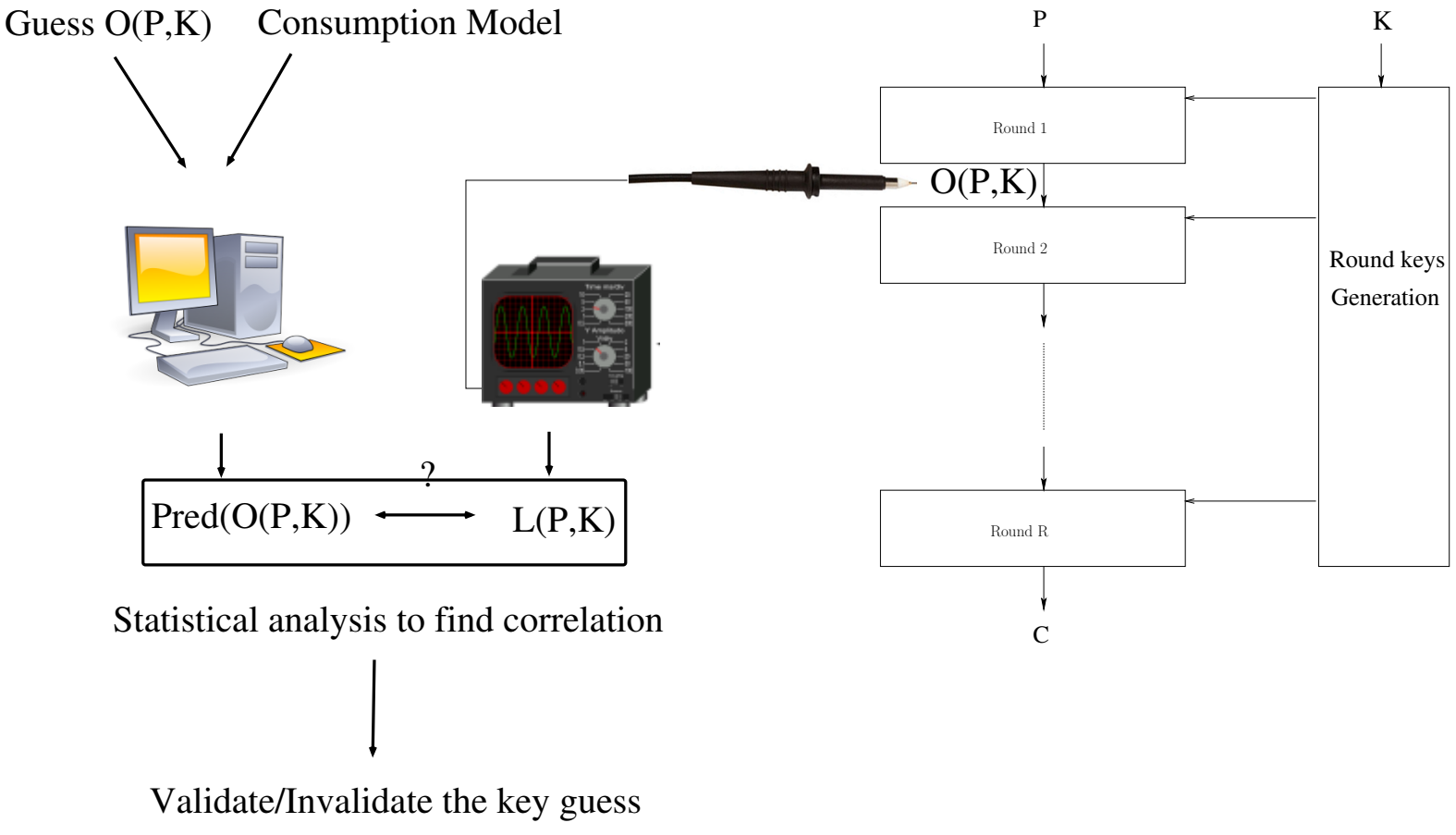


$$\text{Consumption} \sim \zeta \times HW(u \oplus v)$$

Classical Power Analysis Attacks (DPA [Kocher 98])



Classical Power Analysis Attacks (DPA [Kocher 98])



Classical Power Analysis Attacks (DPA [Kocher 98])

- DES (1974)
  - $\mathcal{P}$: 64 bits
  - $\mathcal{K}$: 56 bits
  - best attack $\sim 2^{56}$ (brute force)
  - Power Analysis Attack : few hundreds known plaintexts
- AES (2000)
  - $\mathcal{P}$: 128 bits
  - $\mathcal{K}$: 128/192/256 bits
  - best attack $\sim 2^{128}/2^{192}/2^{256}$ (brute force)
  - Power Analysis Attack : few hundreds known plaintexts

Countermeasures

The implementation is safe if one can shut down all information leakages :

Suppress synchronization elements. (buses and registers)

and/or Randomize the data processed.
(Random Masks [Goub 99, Char 99, Akka 01, Lv 05, Riva 08])

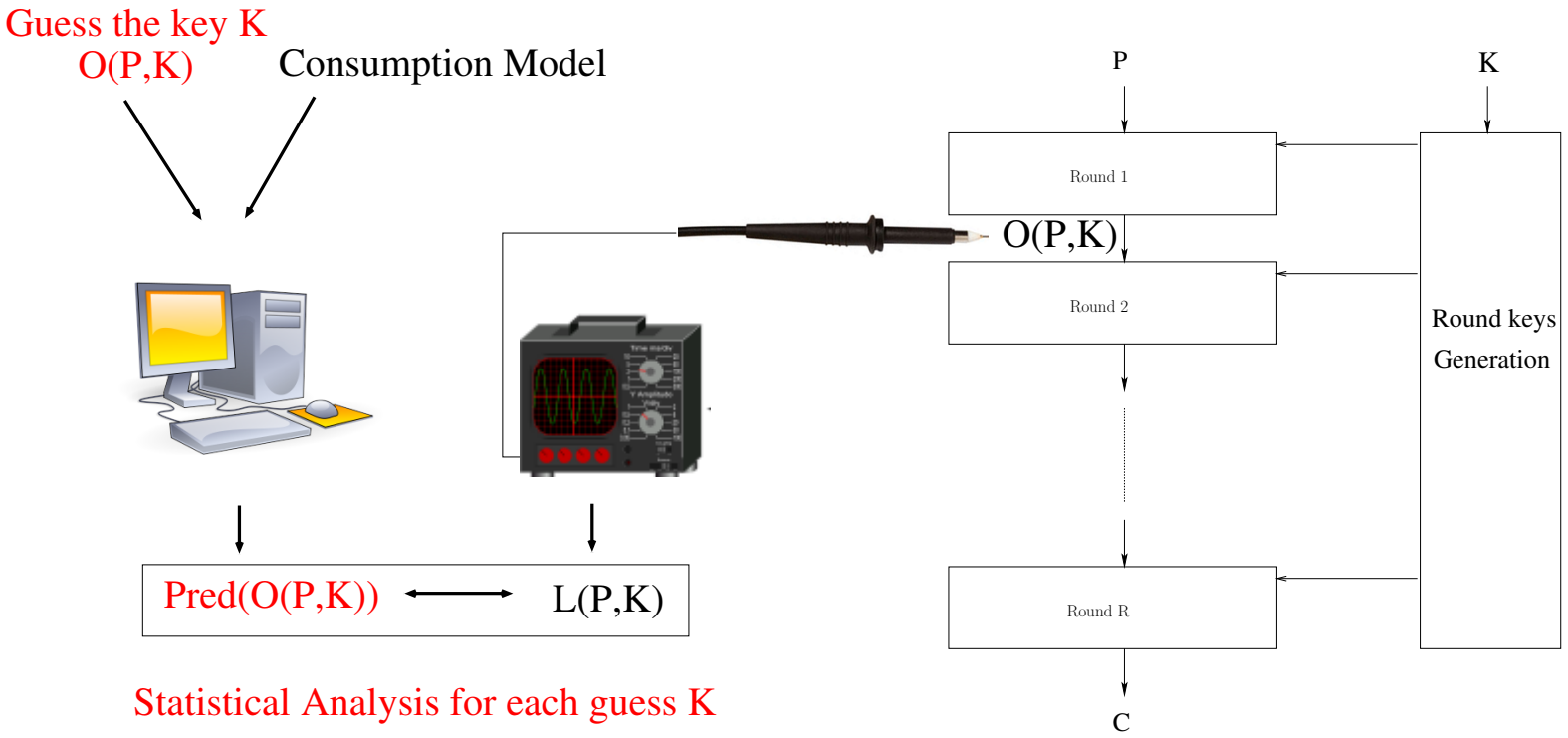
and/or Add random useless computations.

and/or balanced dynamic dual-rail gates designs.

and/or ...

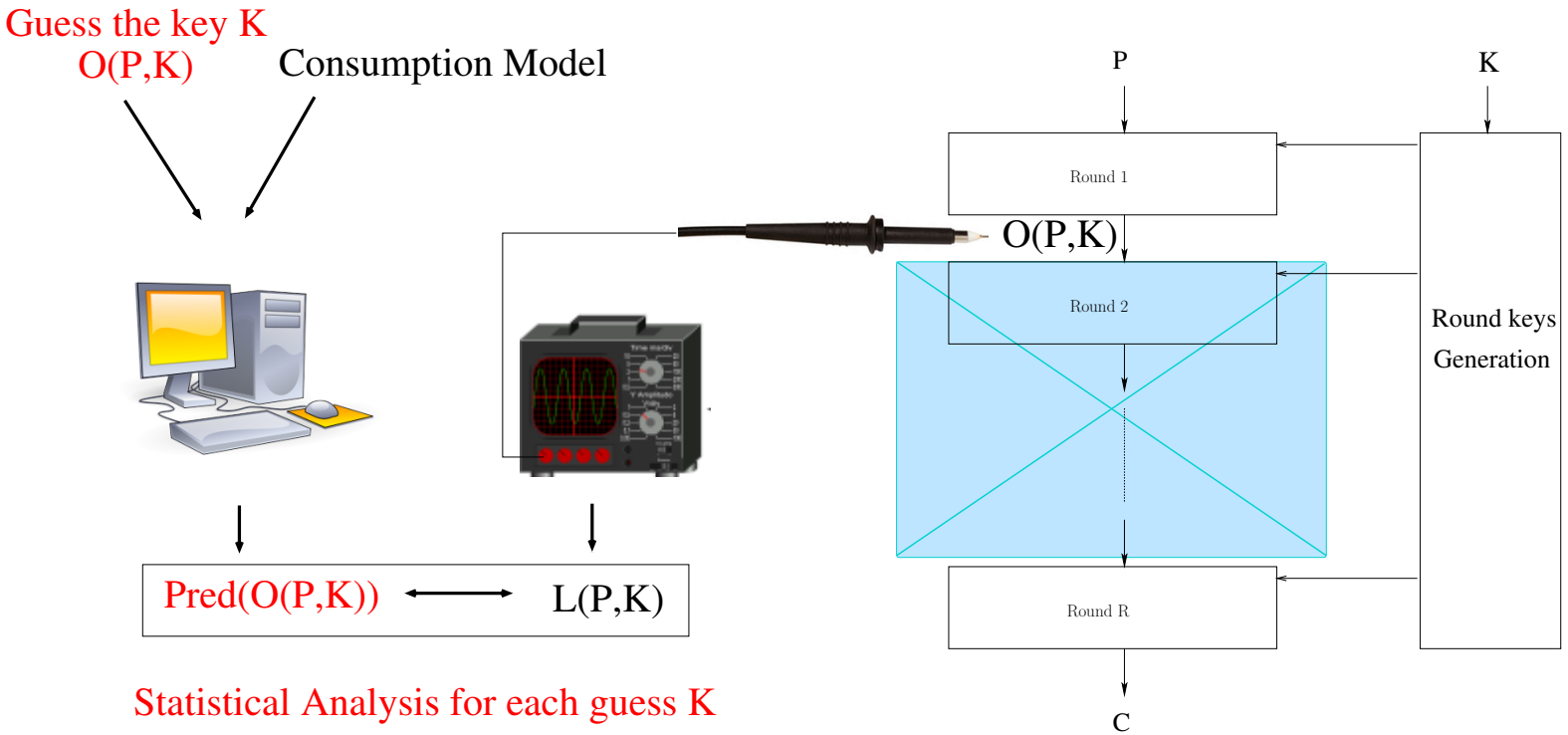
The sound countermeasures are very expensive !

Limitations of classical Power Analysis Attacks



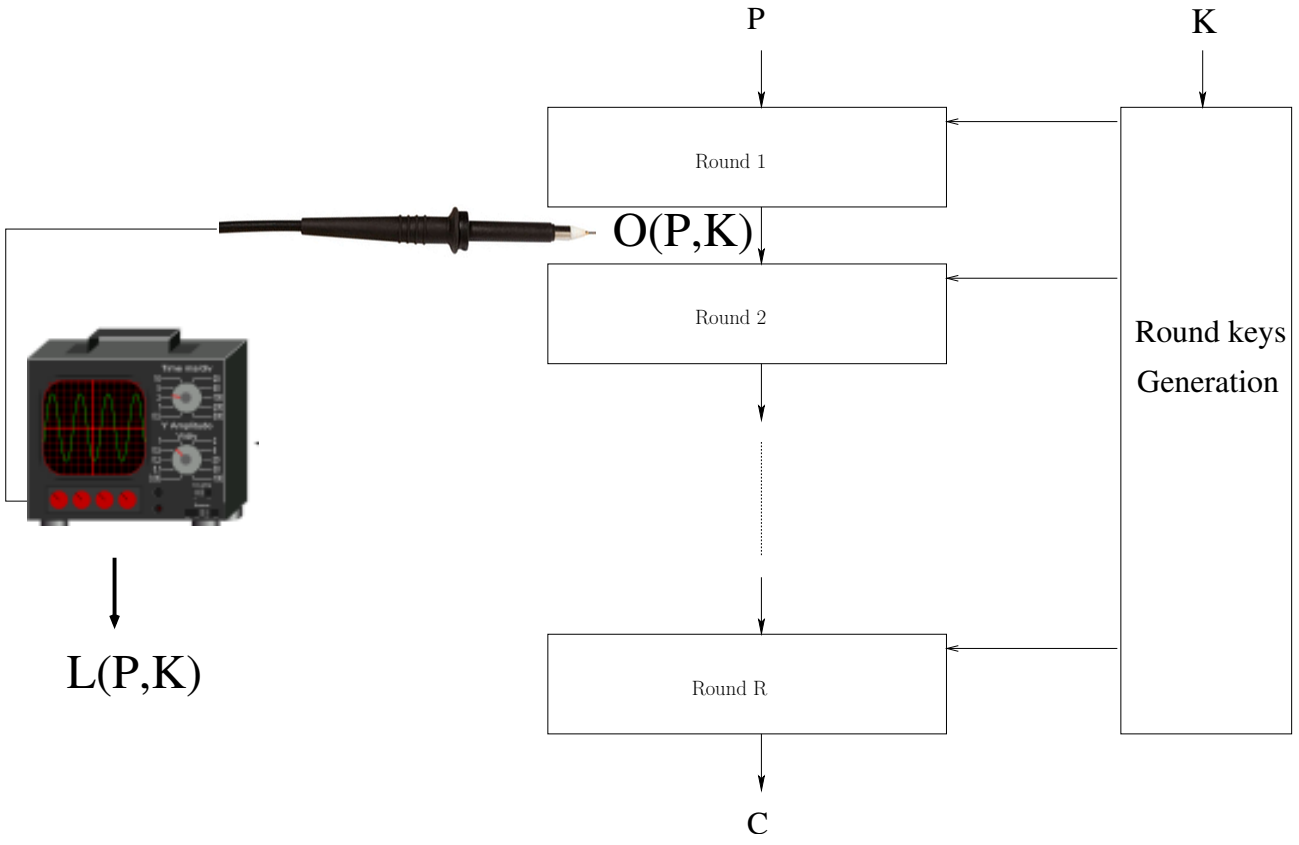
Restriction to intermediate values $O(P, K)$ dependent to less than 32 key-bits.
i.e. first and last round of the cipher.

Limitations of classical Power Analysis Attacks

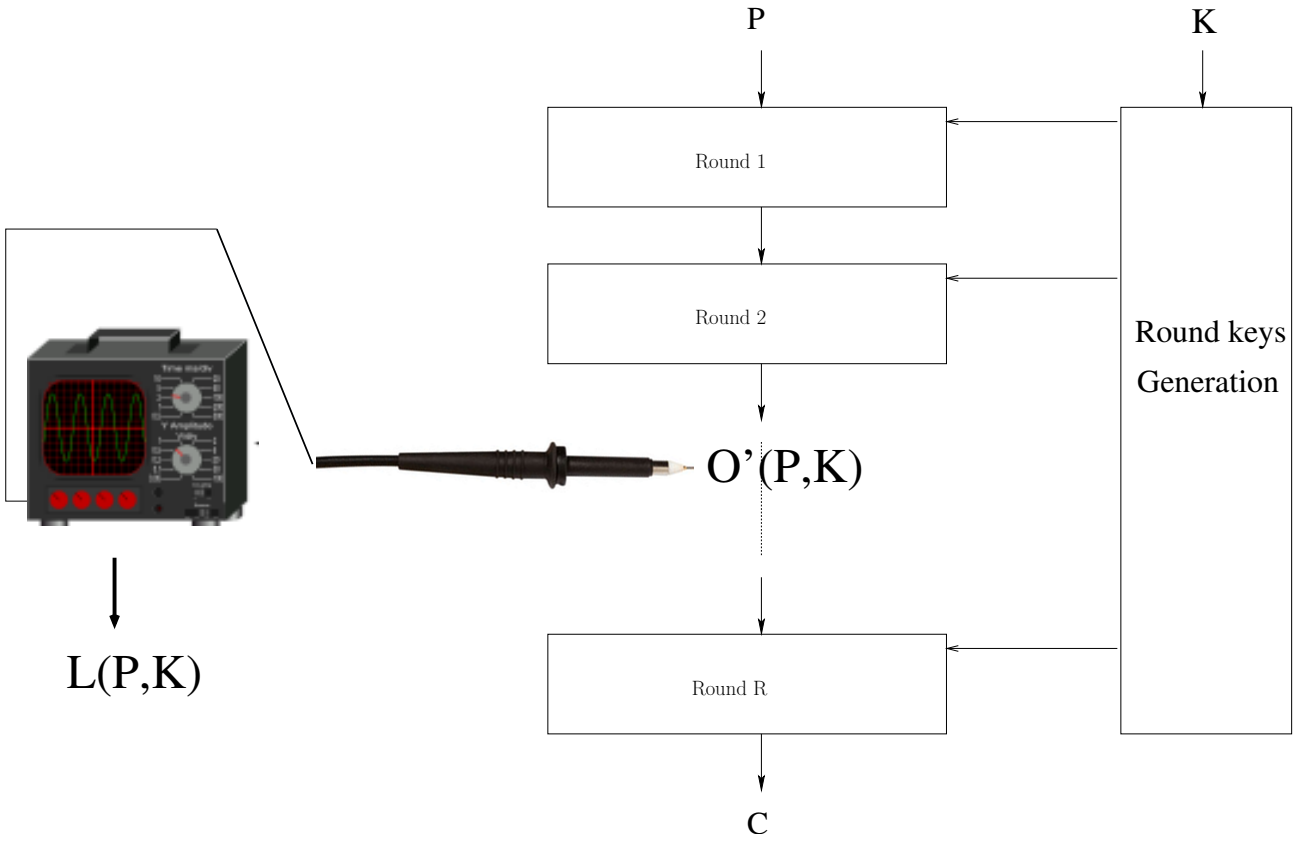


Restriction to intermediate values $O(P, K)$ dependent to less than 32 key-bits.
i.e. first and last round of the cipher.

We want to use deeper leakage points



We want to use deeper leakage points



Glued Blocks

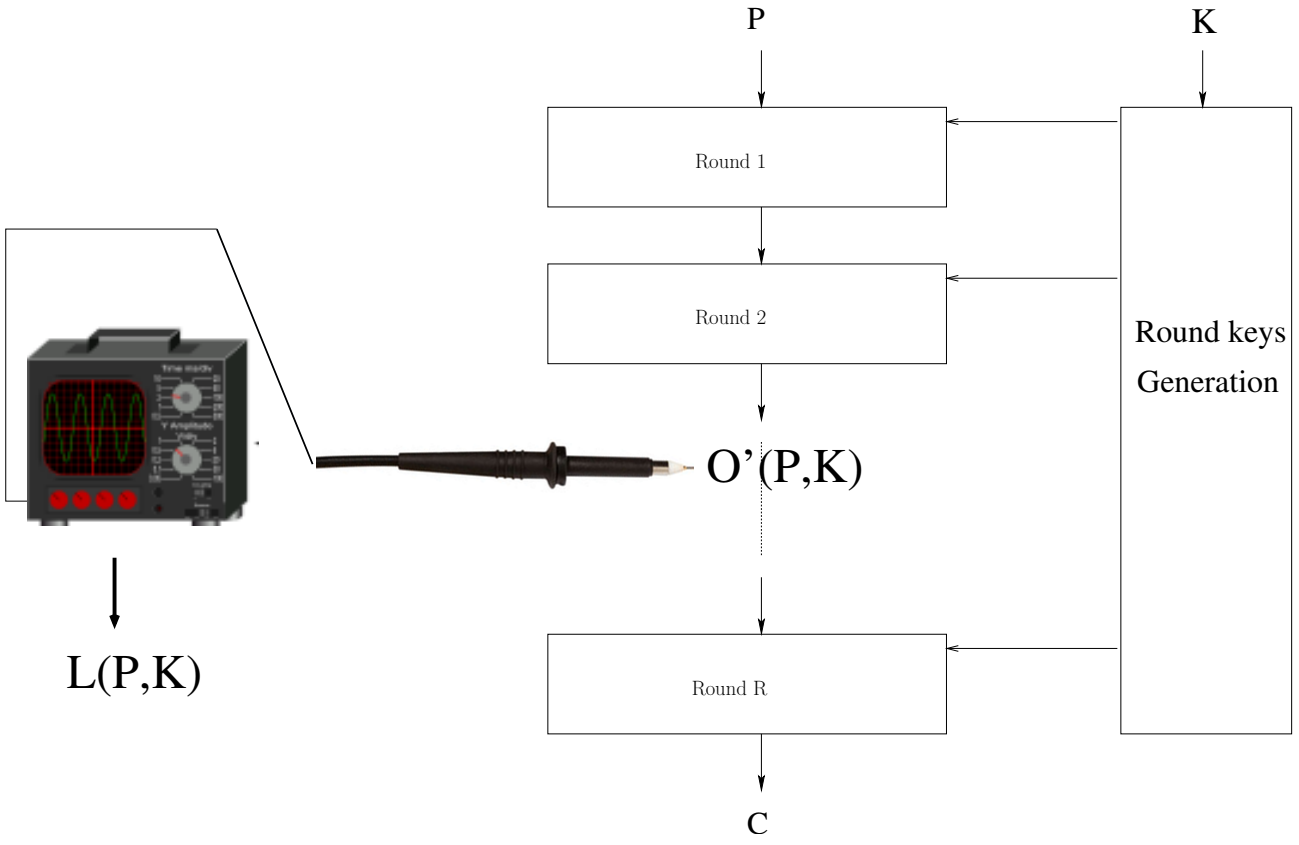
light countermeasures [Akkar Goubin 03]

Concentrate countermeasures on the firsts and lasts rounds.
i.e. no information leak during these critical rounds

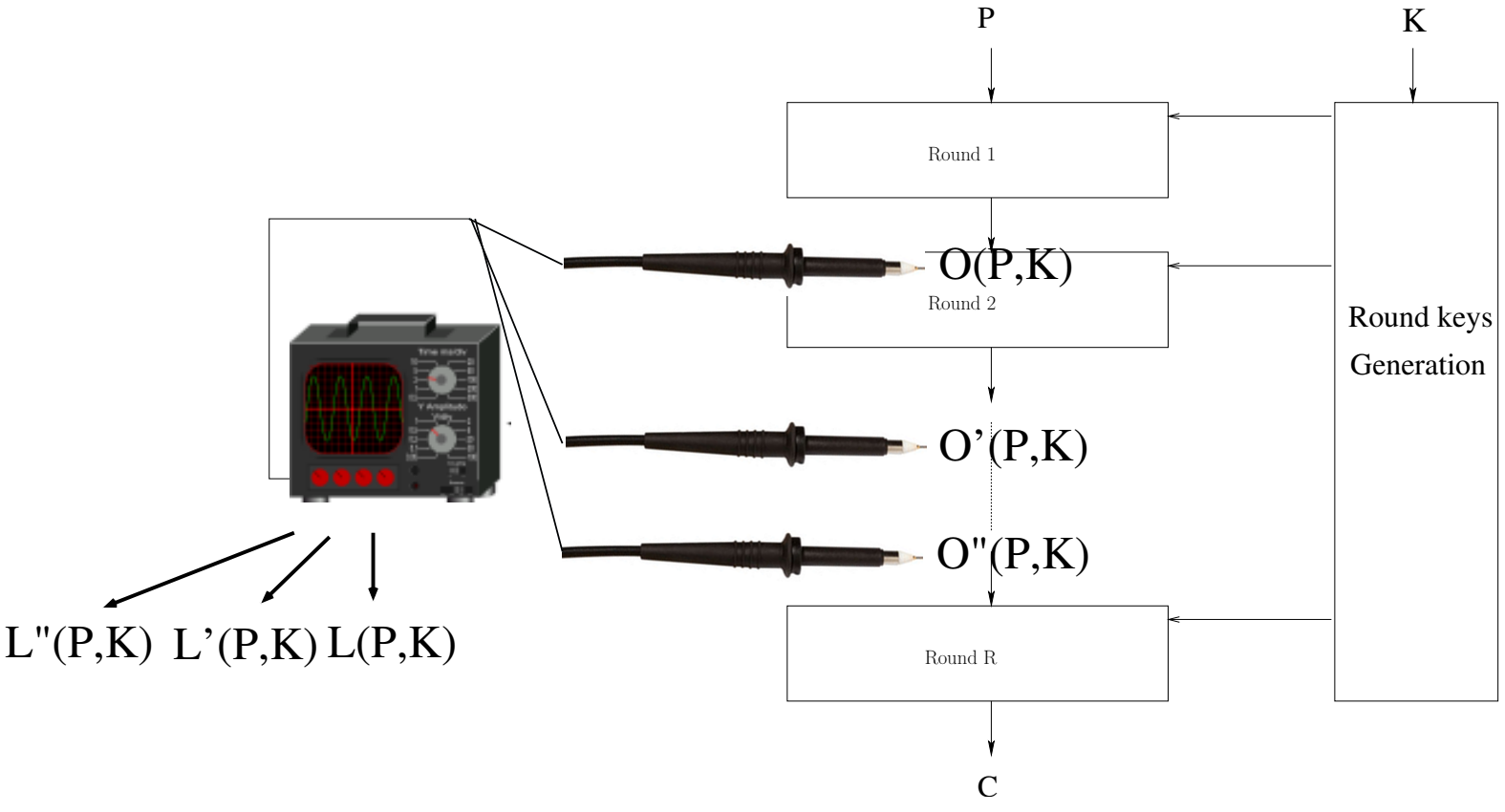
$\Rightarrow$ No observable intermediate value is dependent to less than 32 key bits.

Sufficient countermeasure against classical power analysis attacks : DPA, CPA, MIA, etc ...

Simultaneous multiple side channel measurements



Simultaneous multiple side channel measurements



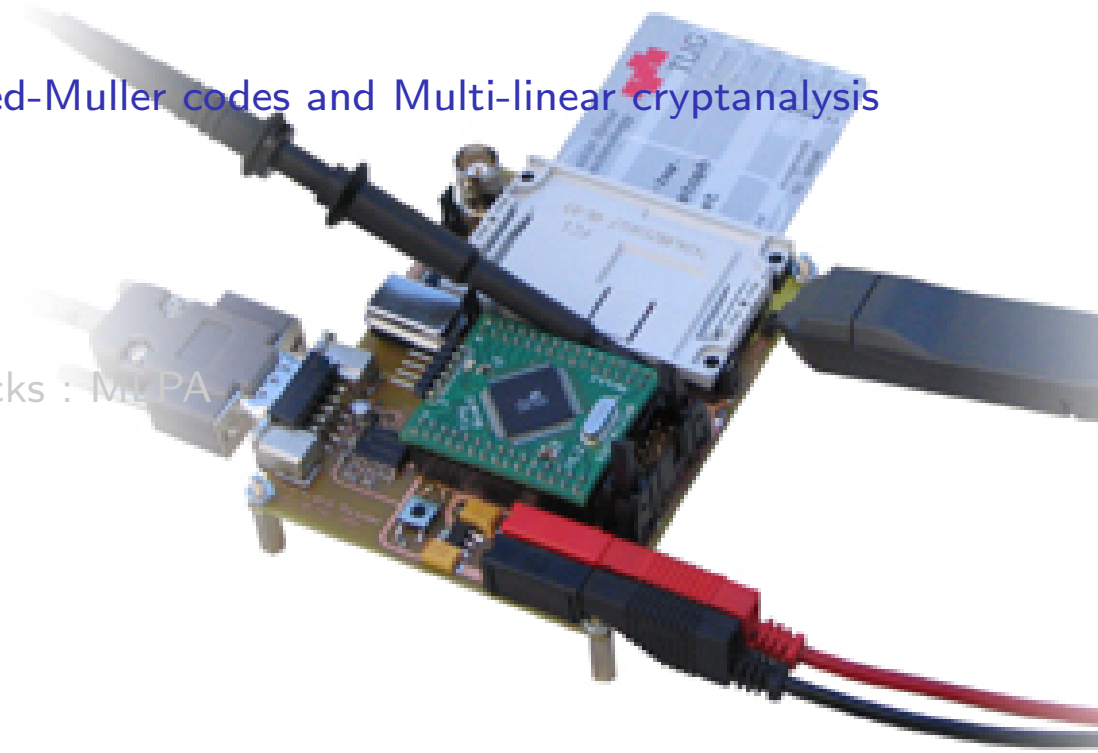
Related work

- Differential-based side channel [Carl 05, Hand 06]
- Collision-based side channel
  [Schr 03, Schr 04, Ledi 04, Bogd 07, Bogd 08b, Bogd 08a]
- Algebraic side channel [Rena 09]

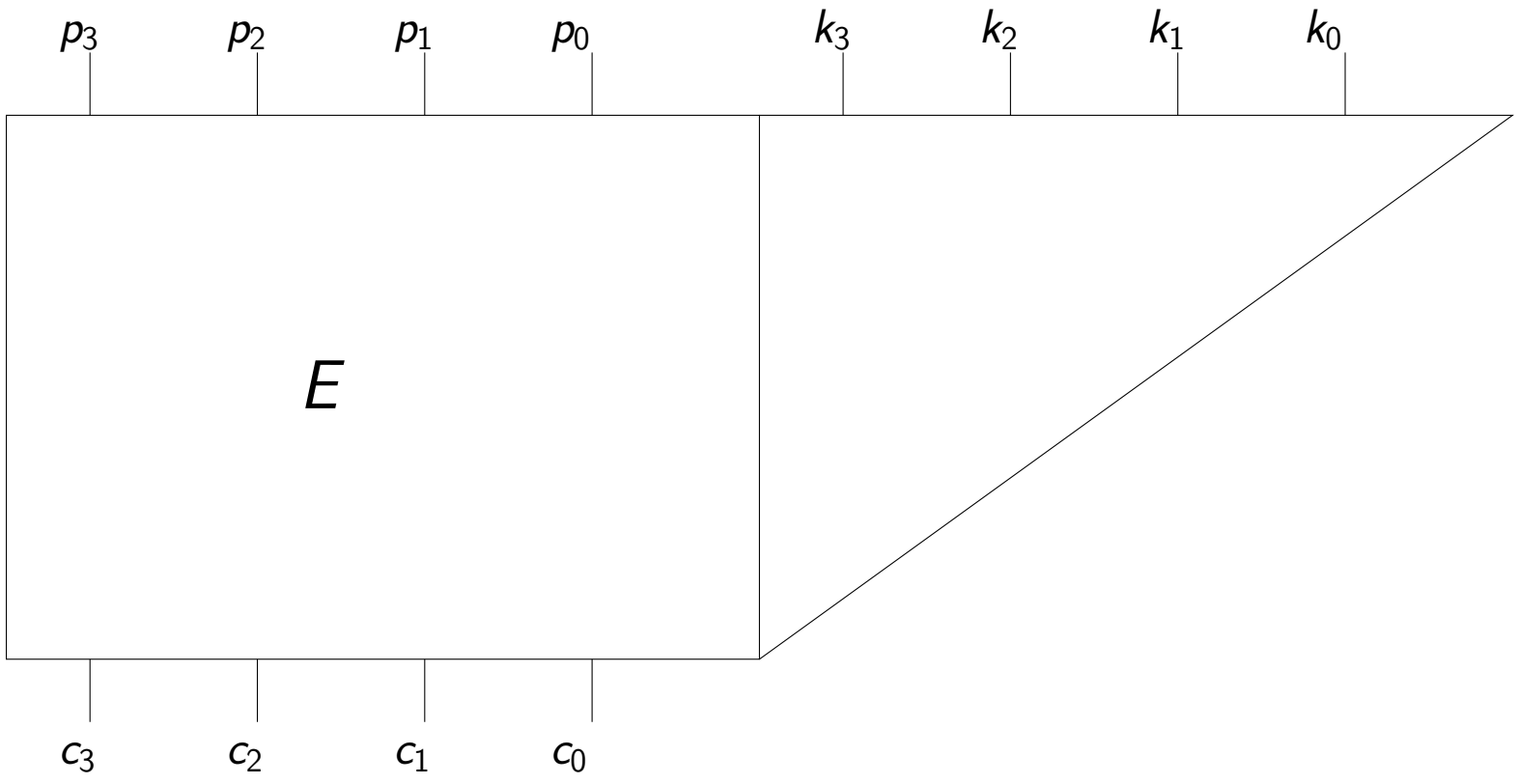
Limitation

These method usually need to have very precise measures (template-like attacks).
Hypothesis : The attacker must be able to precisely evaluate $O(P, K)$ from $L(P, K) \Rightarrow$ Access to a training board.

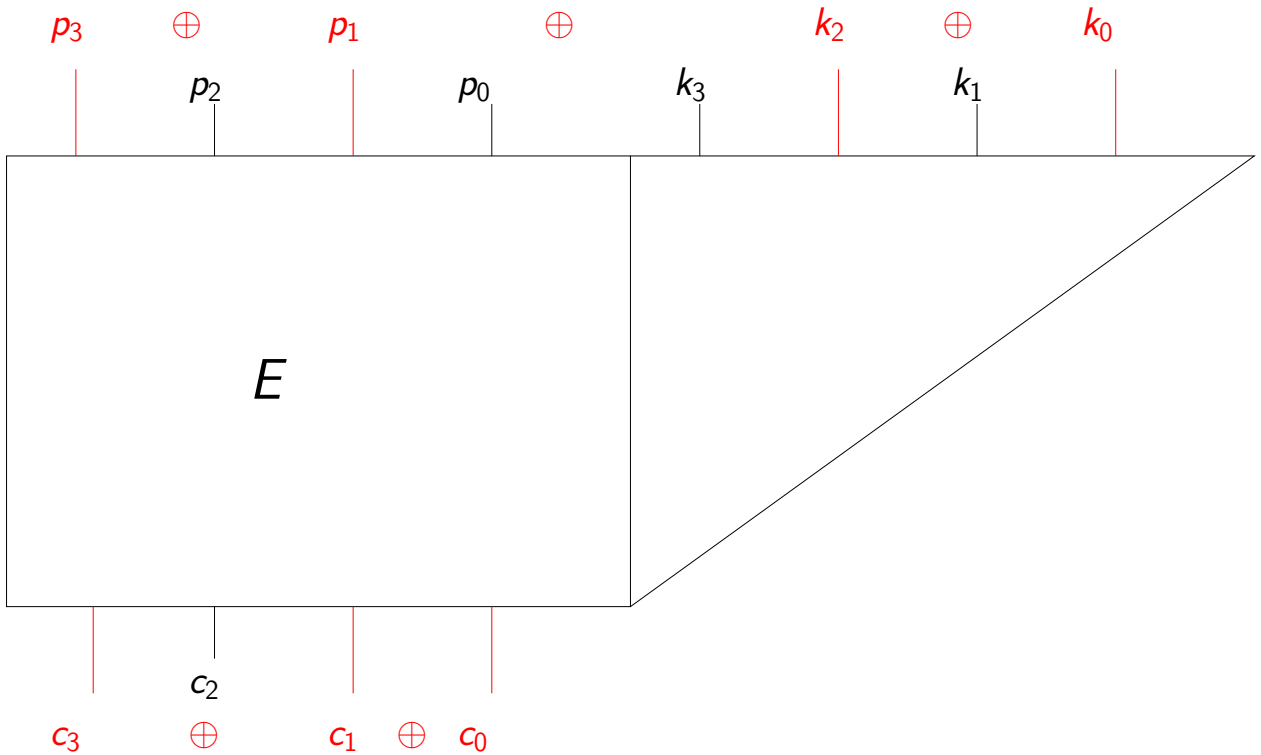
Outline



Symmetric cipher (4-bits plaintexts, 4-bits key)



Linear approximations



linear approximation

$p_1 \oplus p_3 \oplus k_0 \oplus k_2 = c_0 \oplus c_1 \oplus c_3$
hold with probability $p = 1/2 + \epsilon$.

Multi-linear cryptanalysis

$$\left\{\begin{array}{lcl|l} k_0 \oplus k_2 & = & p_1 \oplus p_3 \oplus c_0 \oplus c_1 \oplus c_3 & p = 1/2 + \epsilon_1 \\ k_0 \oplus k_1 \oplus k_2 & = & p_0 \oplus p_2 \oplus c_2 \oplus c_3 & p = 1/2 + \epsilon_2 \\ k_1 \oplus k_3 & = & p_2 \oplus p_3 \oplus c_1 \oplus c_3 & p = 1/2 + \epsilon_3 \\ k_1 \oplus k_2 \oplus k_3 & = & p_0 \oplus p_1 \oplus p_2 \oplus p_3 \oplus c_2 \oplus c_3 & p = 1/2 + \epsilon_4 \end{array}\right.$$

Complexity of the attack [Biry 04]

Given n linear approximations $\langle \alpha_i, P \rangle \oplus \langle \mu_i, K \rangle = \langle \beta_i, E(P,K) \rangle$

$$\#Plaintexts = O(\frac{1}{\sum_i (\epsilon_i^2)})$$

Multi-linear cryptanalysis

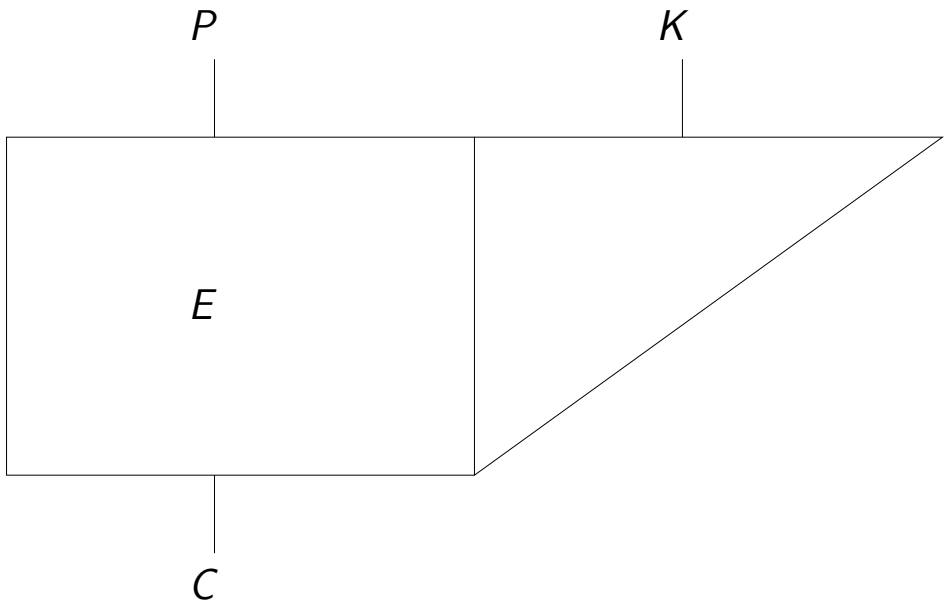
$$\left\{\begin{array}{lcl|l} k_0 \oplus k_2 & = & p_1 \oplus p_3 \oplus c_0 \oplus c_1 \oplus c_3 & p = 1/2 + \epsilon_1 \\ k_0 \oplus k_1 \oplus k_2 & = & p_0 \oplus p_2 \oplus c_2 \oplus c_3 & p = 1/2 + \epsilon_2 \\ k_1 \oplus k_3 & = & p_2 \oplus p_3 \oplus c_1 \oplus c_3 & p = 1/2 + \epsilon_3 \\ k_1 \oplus k_2 \oplus k_3 & = & p_0 \oplus p_1 \oplus p_2 \oplus p_3 \oplus c_2 \oplus c_3 & p = 1/2 + \epsilon_4 \end{array}\right.$$

Complexity of the attack [Biry 04]

Given n linear approximations $\langle \alpha_i, P \rangle \oplus \langle \mu_i, K \rangle = \langle \beta_i, E(P, K) \rangle$

$$\#Plaintexts = O(\frac{1}{\sum_i (\epsilon_i^2)})$$

Multivariate degree 1 polynomial reconstruction

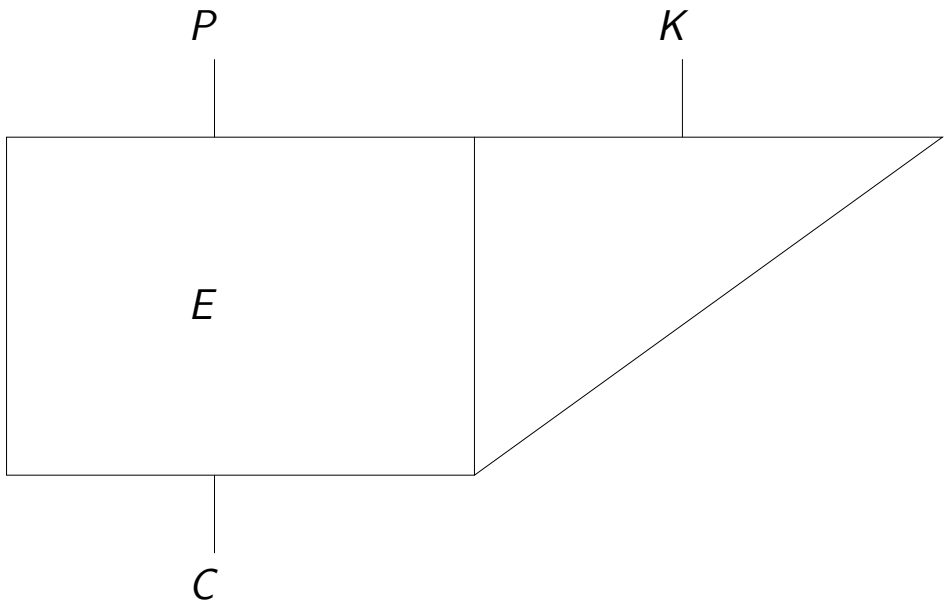


$$\langle \alpha, P \rangle \oplus \langle \mu, K \rangle = \langle \beta, E(P, K) \rangle$$

An interpretation of the problem

By fixing β, we fall in a problem list decoding of the first order Reed-Muller code, we have to decode the noisy codeword $\langle \beta, E(P, K) \rangle$.

Multivariate degree 1 polynomial reconstruction



$$\langle \alpha, P \rangle \oplus \langle \mu, K \rangle = \langle \beta, E(P, K) \rangle$$

An interpretation of the problem

By fixing β, we fall in a problem list decoding of the first order Reed-Muller code, we have to decode the noisy codeword $\langle \beta, E(P, K) \rangle$.

Reed-Muller code properties

Definition of $RM(1, m)$

- $RM(1, m) = \{f \in GF(2)^{(1)}[x_1, x_2, \cdots, x_m]\}$;
- Usual representation : $(f(0), f(1), \cdots, f(2^m - 1))$;
- Boolean representation : $f = f_1 x_1 \oplus f_2 x_2 \oplus \cdots \oplus f_m x_m$
- code of length $n = 2^m$ and minimal distance $d = n/2$.

Classical Problem

Given a Boolean function g, we want to construct the list $\{f \in RM(1, m) \mid d_H(f, g) \leq n(1/2 - \epsilon)\}$, which is equivalent to $L_g(\epsilon) = \{f \in RM(1, m) \mid I^{(g)}(f) = \sum_{x \in GF(2)^m} (-1)^{f(x) \oplus g(x)} \geq 2\epsilon n\}$.

Johnson Bound

In fact $\|L_g(\epsilon)\| \leq \frac{1}{4\epsilon^2}$

List Decoding Algorithms

A simple idea

$2\epsilon n \leq |l^{(g)}(f)| \leq \sum_{s \in GF(2)^{m-i}} |\sum_{r \in GF(2)^i} (-1)^{g(r,s) \oplus f^{(i)}(r)}|$ where

$f^{(i)} = f_1 x_1 \oplus \cdots \oplus f_i x_i$.

Screening process : we suggest f_i and we check if the inequality is satisfied.

$\Rightarrow L_g^{(i)}(\epsilon) = \{f \in RM(1,i) \mid \sum_s |\sum_{r \in GF(2)^i} (-1)^{g(r,s) \oplus f(r)}| \geq 2\epsilon n\}$.

Complexity

Worst case complexity

The complexity of this algorithm is in $\mathcal{O}(n \log_2^2(\epsilon))$ [I Du 07].

The complexity of the prob. version is in $\mathcal{O}(m^2/\epsilon^6)$ [Kaba 04].

The size of the result can be of size $m/2\epsilon^2$, thus optimal complexity could be in $\mathcal{O}(m/\epsilon^2)$.

Optimal complexity

In fact Goldreich and Levin algorithm : $\mathcal{O}(m/\epsilon^4)$.

I. Dumer, G. Kabatiansky and C. Tavernier [Four 09] :

$$\mathcal{O}(m/\epsilon^2)$$

Conclusion

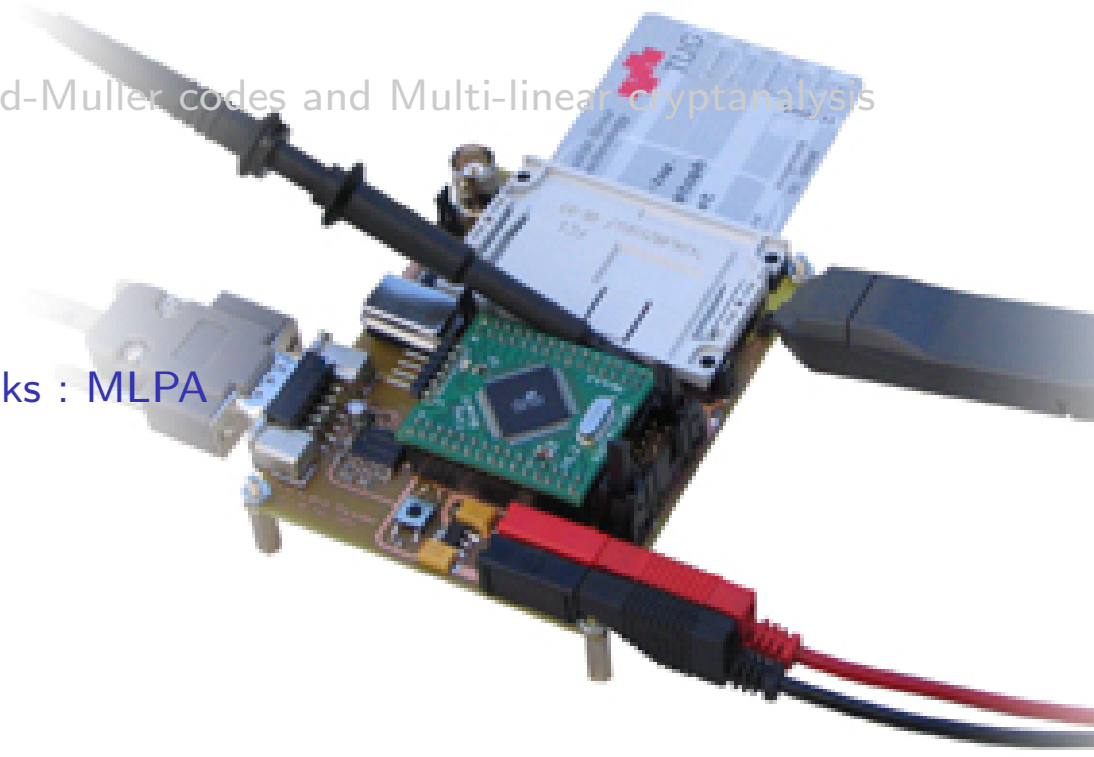
List Decoding in $RM(1, m)$ vs Piling-Up Lemma

Using a list decoding algorithm in $RM(1, m)$,

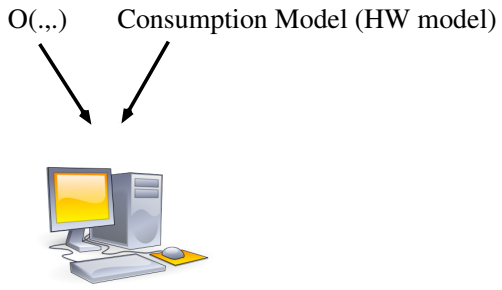
- we do not have to rely on the S-Boxes independence assumption.
- we do not have to know the structure of the target function (seen as a black box).

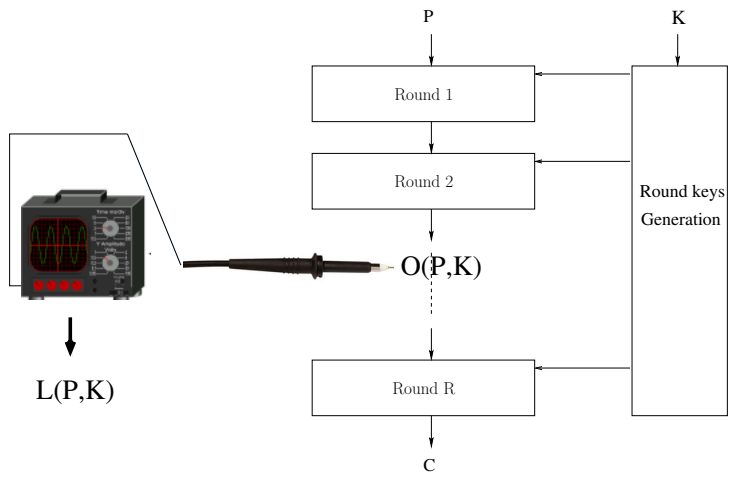
However, the piling-up lemma allows to find approximations with very small biases.

Plan

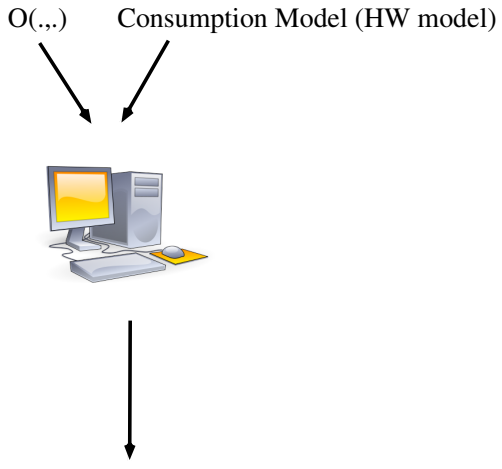


Side channel measurements

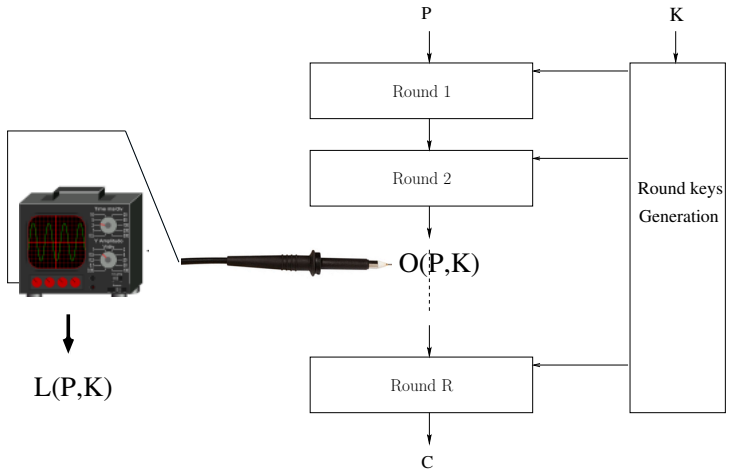




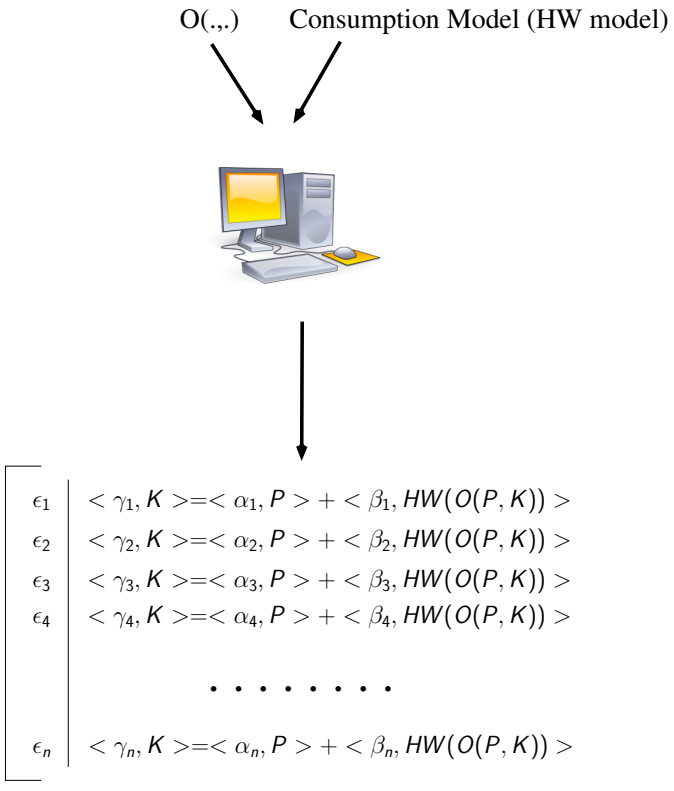
Side channel measurements

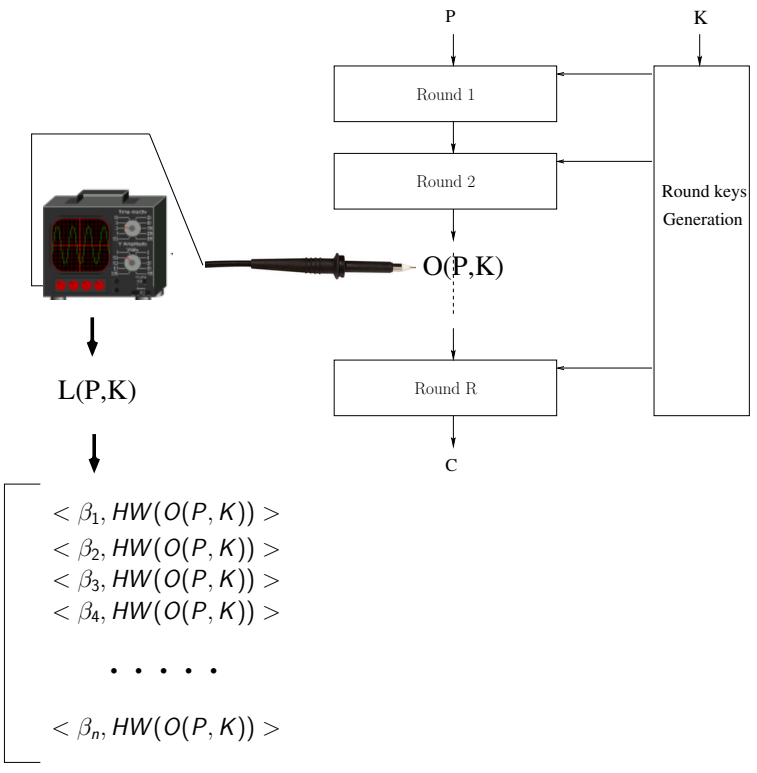


$$
\begin{array}{|l|l|}
\hline
\epsilon_1 & <\gamma_1, K>=<\alpha_1, P>+<\beta_1, HW(O(P,K))> \\
\epsilon_2 & <\gamma_2, K>=<\alpha_2, P>+<\beta_2, HW(O(P,K))> \\
\epsilon_3 & <\gamma_3, K>=<\alpha_3, P>+<\beta_3, HW(O(P,K))> \\
\epsilon_4 & <\gamma_4, K>=<\alpha_4, P>+<\beta_4, HW(O(P,K))> \\
 & \cdots\cdots\cdots \\
\epsilon_n & <\gamma_n, K>=<\alpha_n, P>+<\beta_n, HW(O(P,K))> \\
\hline
\end{array}
$$

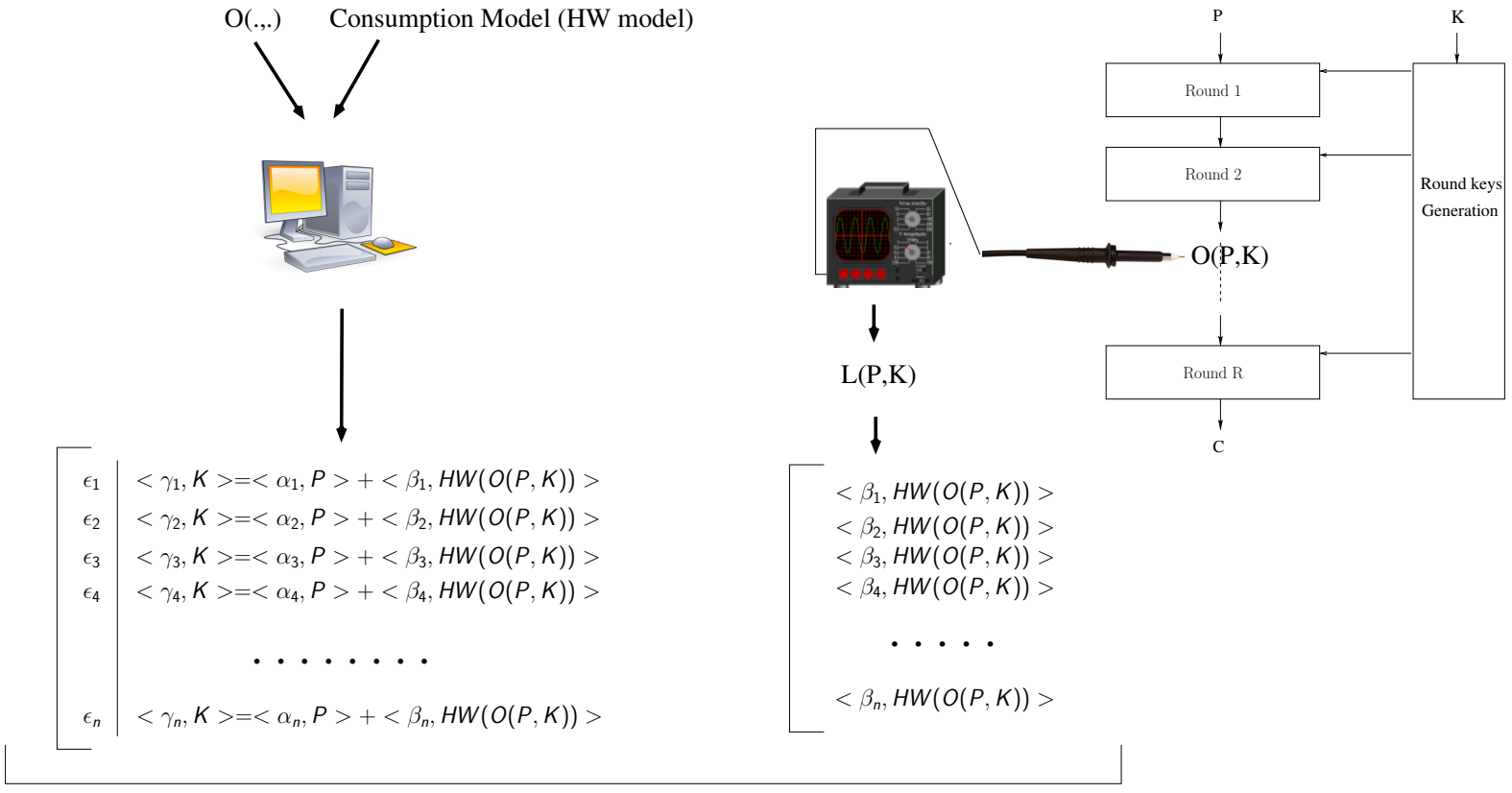


Side channel measurements





Side channel measurements



Get a subpart of K by solving the linear system

Hamming weight approximation : DES implementation

From traces "secmatv1_2006_04_0809"
http://www.dpacontest.org/

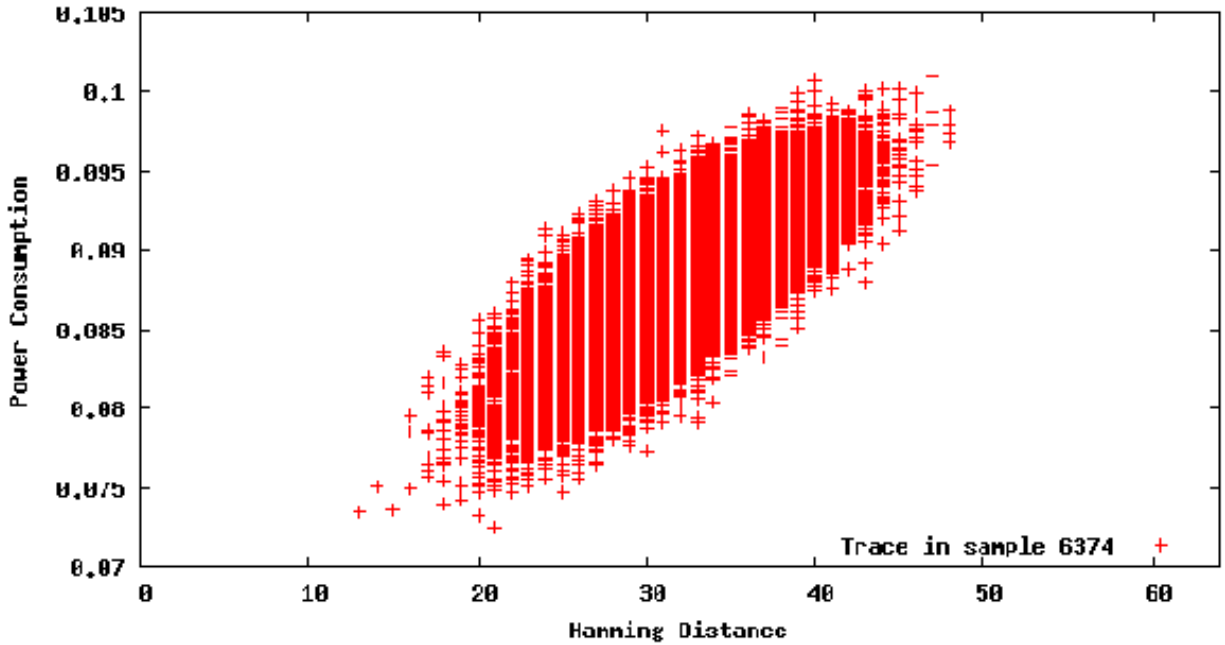


FIG.: Repartition of power consumption with respect to the Hamming Distance model

Hamming weight approximation : DES implementation

From traces "secmatv1_2006_04_0809"
http://www.dpacontest.org/

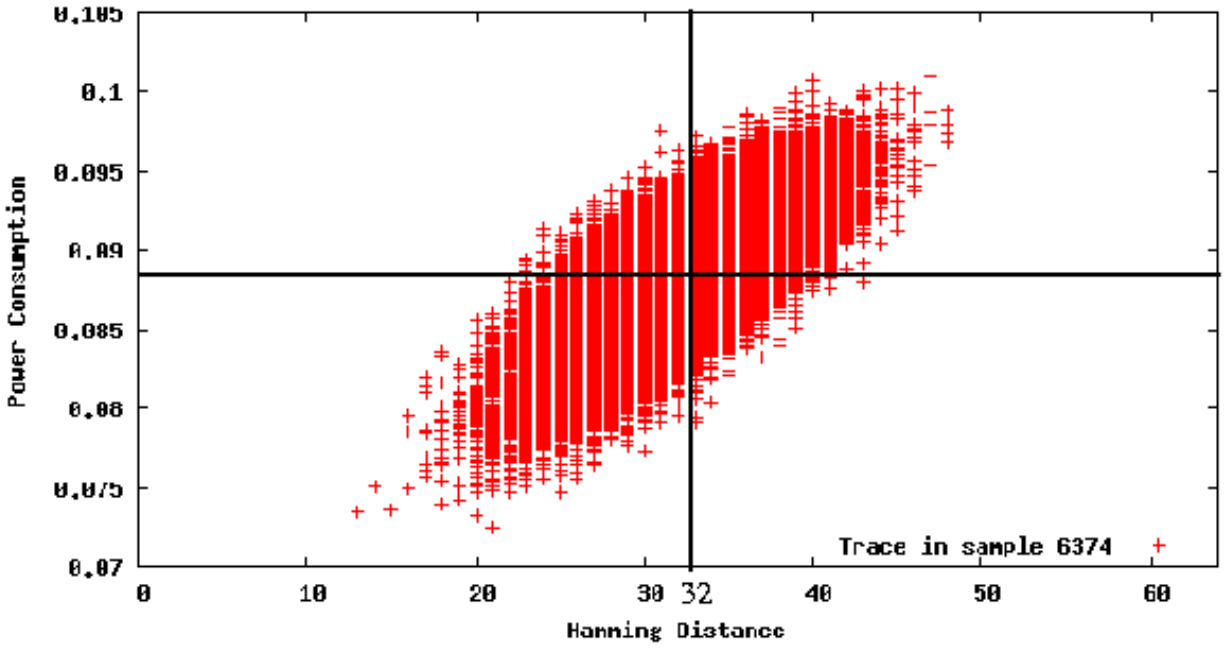


FIG.: Repartition of power consumption with respect to the Hamming Distance model

Attack algorithm and first results on DPA-contest traces

1. Offline static computation : Find many and good approximations of the intermediate data Hamming weight (for every output mask).
2. Online attack : multi-linear cryptanalysis assuming "Leaked information = Hamming distance".

Cipher	rounds	# linear equ.	# key bits	# traces
DES	1	533	40	500
DES	2	1452	27	500
DES	15	488	41	500
DES	16	446	37	500

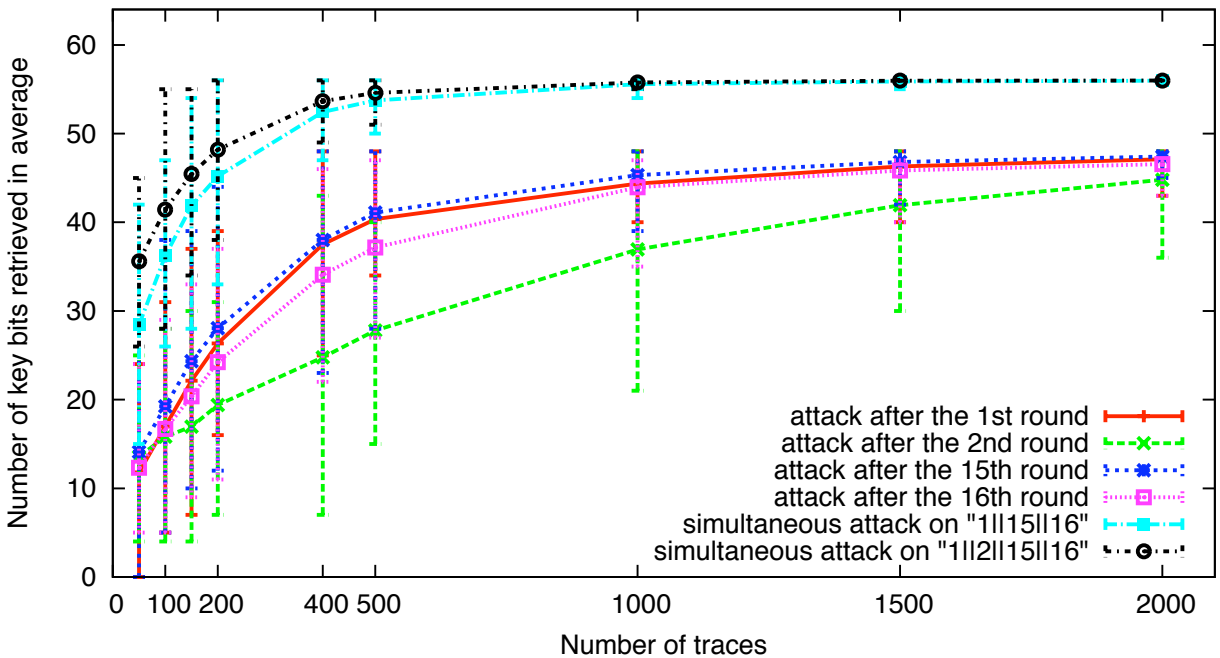
Approximation examples

Output Mask β (in binary) : 100000
Taken on the hamming weight of the bit flips after two rounds.

Bias	Equations (plain part)	Equations (key part)
0.0215	1 + P[5, 26, 27, 31, 45, 53, 61]+	K[6, 7, 29, 38, 52]
0.0134	0 + P[28, 29, 31, 37, 45, 53]+	K[6, 7, 29, 61]
0.0156	1 + P[5, 28, 29, 31, 37, 45]+	K[6, 29, 38, 61]
0.0189	1 + P[5, 28, 29, 31, 37, 53]+	K[7, 29, 38, 61]
0.0163	0 + P[5, 8, 9, 37, 45, 53, 61]+	K[6, 7, 38, 52, 61]
0.0223	0 + P[5, 14, 15, 31, 37, 45, 61]+	K[6, 29, 38, 52, 61]
0.0182	0 + P[5, 28, 29, 31, 37, 53, 61]+	K[7, 29, 38, 52, 61]
0.0157	1 + P[5, 26, 27, 31, 37, 53, 61]+	K[7, 29, 38, 52, 61]
0.0191	0 + P[5, 26, 27, 31, 37, 45, 53, 61]+	K[6, 7, 29, 38, 52, 61]

Simultaneous attack results on DPA-contest traces

Cipher	rounds	# linear equ.	# key bits	# traces
DES	1||15||16	1467	36	100
DES	1||15||16	1467	53	500
DES	1||2||15||16	2919	41	100
DES	1||2||15||16	2919	54	500



Training device attack

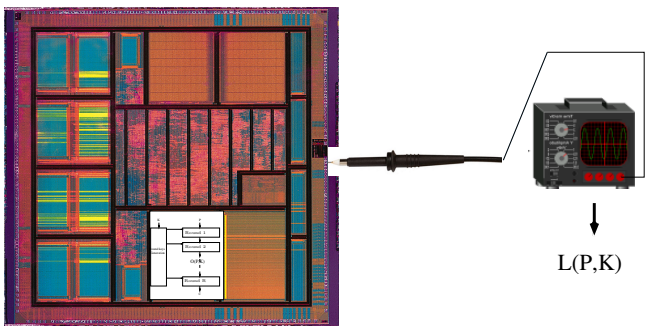
Remark

The List-decoding algorithm operates on the target boolean function as a black-box.

Getting the linear approximations from a twin board
i.e. Chosen plaintexts and keys

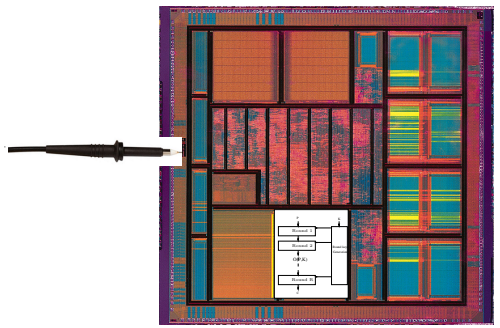
- Approximations directly linked to the leaked information.
  much more accurate.
- No need to choose a power consumption model.
- No need to know the target block cipher.

A template-like attack



Training Board

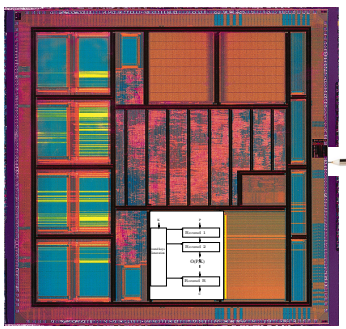
Chosen plaintext
Chosen Key

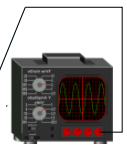


Target Board

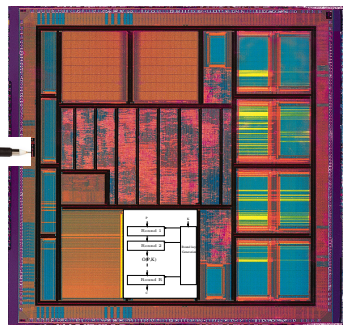
Known plaintext
Secret Key

A template-like attack





L(P,K) Consumption Model (HW model)



Training Board

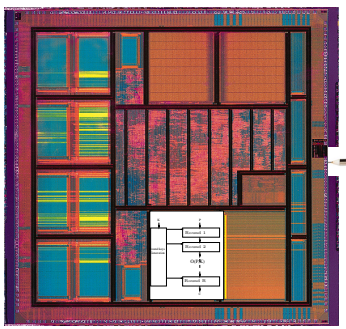
Chosen plaintext
Chosen Key

Target Board

Known plaintext
Secret Key

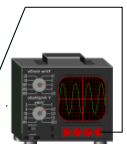
$$
\begin{array}{|c|l|}
\hline
\epsilon_1 & <\gamma_1, K>=<\alpha_1, P>+<\beta_1, L(P,K)> \\
\epsilon_2 & <\gamma_2, K>=<\alpha_2, P>+<\beta_2, L(P,K)> \\
\epsilon_3 & <\gamma_3, K>=<\alpha_3, P>+<\beta_3, L(P,K)> \\
\epsilon_4 & <\gamma_4, K>=<\alpha_4, P>+<\beta_4, L(P,K)> \\
 & \cdots\cdots\cdots \\
\epsilon_n & <\gamma_n, K>=<\alpha_n, P>+<\beta_n, L(P,K)> \\
\hline
\end{array}
$$

A template-like attack



Training Board

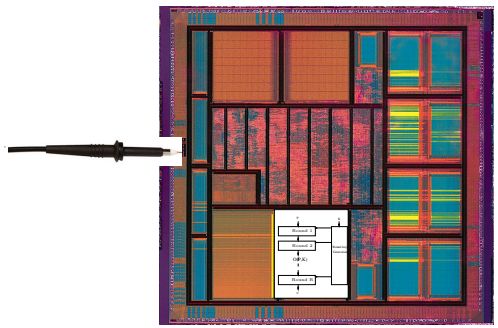
Chosen plaintext
Chosen Key



L(P,K)

Consumption Model (HW model)

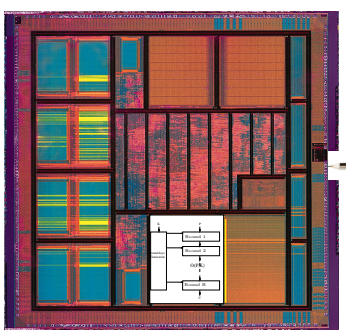
$$\begin{array}{|c|l|}
\epsilon_1 & <\gamma_1, K>=<\alpha_1, P>+<\beta_1, L(P,K)> \\
\epsilon_2 & <\gamma_2, K>=<\alpha_2, P>+<\beta_2, L(P,K)> \\
\epsilon_3 & <\gamma_3, K>=<\alpha_3, P>+<\beta_3, L(P,K)> \\
\epsilon_4 & <\gamma_4, K>=<\alpha_4, P>+<\beta_4, L(P,K)> \\
 & \cdots\cdots\cdots \\
\epsilon_n & <\gamma_n, K>=<\alpha_n, P>+<\beta_n, L(P,K)>
\end{array}$$



Target Board

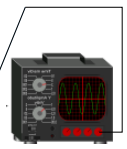
Known plaintext
Secret Key

A template-like attack



Training Board

Chosen plaintext
Chosen Key

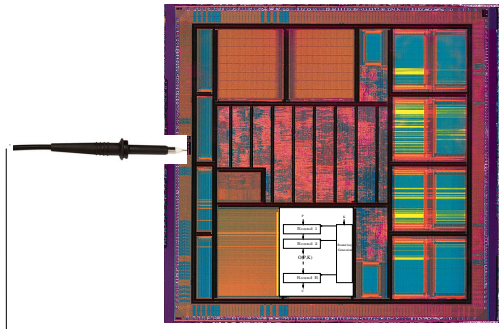


L(P,K)

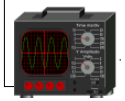
Consumption Model (HW model)

$$\begin{array}{|c|l|}
\epsilon_1 & <\gamma_1, K>=<\alpha_1, P>+<\beta_1, L(P,K)> \\
\epsilon_2 & <\gamma_2, K>=<\alpha_2, P>+<\beta_2, L(P,K)> \\
\epsilon_3 & <\gamma_3, K>=<\alpha_3, P>+<\beta_3, L(P,K)> \\
\epsilon_4 & <\gamma_4, K>=<\alpha_4, P>+<\beta_4, L(P,K)> \\
 & \cdots\cdots\cdots \\
\epsilon_n & <\gamma_n, K>=<\alpha_n, P>+<\beta_n, L(P,K)>
\end{array}$$

Get a subpart of K by solving the linear system



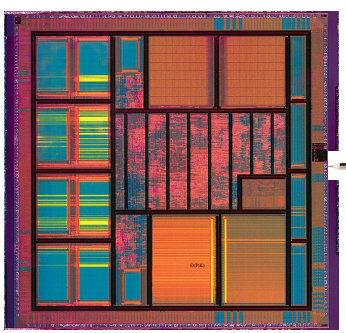
Target Board

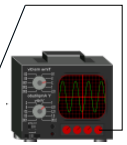


Known plaintext
Secret Key

L(P,K)

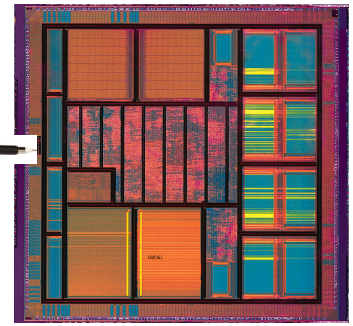
A template-like attack





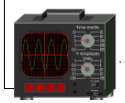
L(P,K)

Consumption Model (HW model)



Target Board

Known plaintext
Secret Key



Training Board

Chosen plaintext
Chosen Key

$$\begin{array}{l|l} \epsilon_1 & <\gamma_1, K>=<\alpha_1, P>+<\beta_1, L(P,K)> \\ \epsilon_2 & <\gamma_2, K>=<\alpha_2, P>+<\beta_2, L(P,K)> \\ \epsilon_3 & <\gamma_3, K>=<\alpha_3, P>+<\beta_3, L(P,K)> \\ \epsilon_4 & <\gamma_4, K>=<\alpha_4, P>+<\beta_4, L(P,K)> \\ & \cdots\cdots\cdots \\ \epsilon_n & <\gamma_n, K>=<\alpha_n, P>+<\beta_n, L(P,K)> \end{array}$$

L(P,K)

Get a subpart of K by solving the linear system

Next Steps on MLPA :

- MLPA on a consumption model refinement.
- HO-MLPA.
- MLPA on masked implementation.
- Other block ciphers.
- Better or More linear approximations.
- MLPA template attack.
- Unknown block cipher attack.
- *Higher degree approximations.*

The end.

references I


M.-L. Akkar and C. Giraud.
"An Implementation of DES and AES, Secure against Some Attacks".
In : Çetin Kaya Koç, D. Naccache, and C. Paar, Eds., *CHES*, pp. 309–318, Springer, 2001.


A. Biryukov, C. D. Cannière, and M. Quisquater.
"On Multiple Linear Approximations".
In : M. K. Franklin, Ed., *CRYPTO*, pp. 1–22, Springer, 2004.


A. Bogdanov.
"Improved Side-Channel Collision Attacks on AES".
In : C. M. Adams, A. Miri, and M. J. Wiener, Eds., *Selected Areas in Cryptography*, pp. 84–95, Springer, 2007.

references II


A. Bogdanov.
"Multiple-Differential Side-Channel Collision Attacks on AES".

In : E. Oswald and P. Rohatgi, Eds., *CHES*, pp. 30–44, Springer, 2008.


A. Bogdanov, I. Kizhvatov, and A. Pyshkin.
"Algebraic Methods in Side-Channel Collision Attacks and Practical Collision Detection".
In : D. R. Chowdhury, V. Rijmen, and A. Das, Eds., *INDOCRYPT*, pp. 251–265, Springer, 2008.

references III

V. Carlier, H. Chabanne, E. Dottax, and H. Pelletier.
"Generalizing Square Attack using Side-Channels of an AES Implementation on an FPGA".
In : T. Rissa, S. J. E. Wilton, and P. H. W. Leong, Eds., *FPL*, pp. 433–437, IEEE, 2005.

S. Chari, C. S. Jutla, J. R. Rao, and P. Rohatgi.
"Towards Sound Approaches to Counteract Power-Analysis Attacks".
In : M. J. Wiener, Ed., *CRYPTO*, pp. 398–412, Springer, 1999.

R. Fourquet, P. Loidreau, and C. Tavernier.
"Finding good linear approximations of block ciphers and its application to cryptanalysis of reduced round DES".
In : *Workshop on Coding Theory and Cryptography, Ullensvang, Norvège*, 2009.

references IV

L. Goubin and J. Patarin.
"DES and Differential Power Analysis (The "Duplication" Method)".
In : Çetin Kaya Koç and C. Paar, Eds., *CHES*, pp. 158–172, Springer, 1999.

H. Handschuh and B. Preneel.
"Blind Differential Cryptanalysis for Enhanced Power Attacks".
In : E. Biham and A. M. Youssef, Eds., *Selected Areas in Cryptography*, pp. 163–173, Springer, 2006.

G. K. I. Dumer and C. Tavernier.
"List Decoding of the First Order Binary Reed Muller Codes".
Problems of Information Transmission, Vol. 43, No. 3, pp. 225–232, 2007.

references V


M. Joye and J.-J. Quisquater, Eds.
Cryptographic Hardware and Embedded Systems - CHES 2004 : 6th International Workshop Cambridge, MA, USA, August 11-13, 2004. Proceedings, Springer, 2004.


G. Kabatiansky and C. Tavernier.
"List decoding with Reed Muller codes of order one".
In : *nine International Workshop On Algebraic and Combinatorial Coding Theory*, pp. 230–236, 2004.


H. Ledig, F. Muller, and F. Valette.
"Enhancing Collision Attacks".
In : M. Joye and J.-J. Quisquater, Eds., *CHES*, pp. 176–190, Springer, 2004.

references VI

J. Lv and Y. Han.
"Enhanced DES Implementation Secure Against High-Order Differential Power Analysis in Smartcards".
In : C. Boyd and J. M. G. Nieto, Eds., *ACISP*, pp. 195–206, Springer, 2005.

M. Renauld and F.-X. Standaert.
"Algebraic Side-Channel Attacks".
Cryptology ePrint Archive, Report 2009/279, 2009.
http://eprint.iacr.org/.

M. Rivain, E. Dottax, and E. Prouff.
"Block Ciphers Implementations Provably Secure Against Second Order Side Channel Analysis".
In : K. Nyberg, Ed., *FSE*, pp. 127–143, Springer, 2008.

references VII

K. Schramm, T. J. Wollinger, and C. Paar.
"A New Class of Collision Attacks and Its Application to DES".
In : T. Johansson, Ed., *FSE*, pp. 206–222, Springer, 2003.

K. Schramm, G. Leander, P. Felke, and C. Paar.
"A Collision-Attack on AES : Combining Side Channel- and Differential-Attack".
In : M. Joye and J.-J. Quisquater, Eds., *CHES*, pp. 163–175, Springer, 2004.