Improved Higher-Order Side-Channel Attacks with FPGA Experiments

Eric Peeters*, François-Xavier Standaert,
Nicolas Donckers, Jean-Jacques Quisquater

UCL Crypto Group

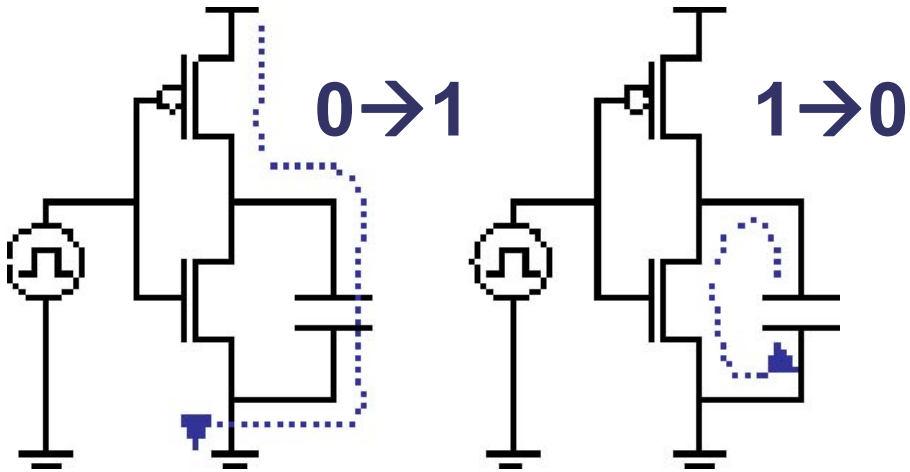
CHES 2005

Outline

- Power Analysis: CMOS models
- Block Cipher
- Boolean Masking
- High-Order Power Attack
  - With perfect measurements
  - With real measurements
- Comparison
- Conclusions

Power Consumption Model

- Most IC's ➔ CMOS



$$O \cong H[d \oplus d']$$

perfect

real

Different Contexts

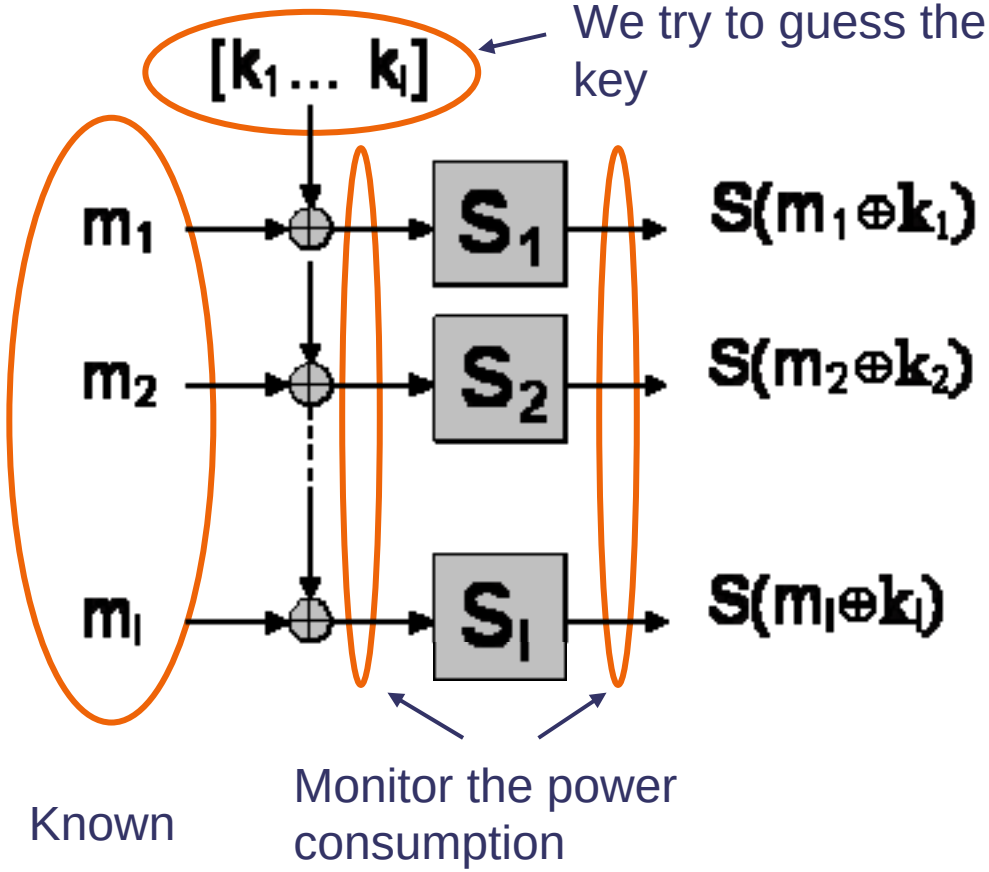
- **Hamming Weight** model

$$O = H[d]$$

- **Hamming Distance** model: depends on d and d'

$$O = H[d \oplus d']$$

First Round of a Block Cipher



In our analysis all buses are 4-bit wide

(N_b=4)

Targeting FPGA

- Pipelined Implementation
  - Structure fed with a new message every clock cycle
- Hamming Distance model
  - Consumption before the S-box

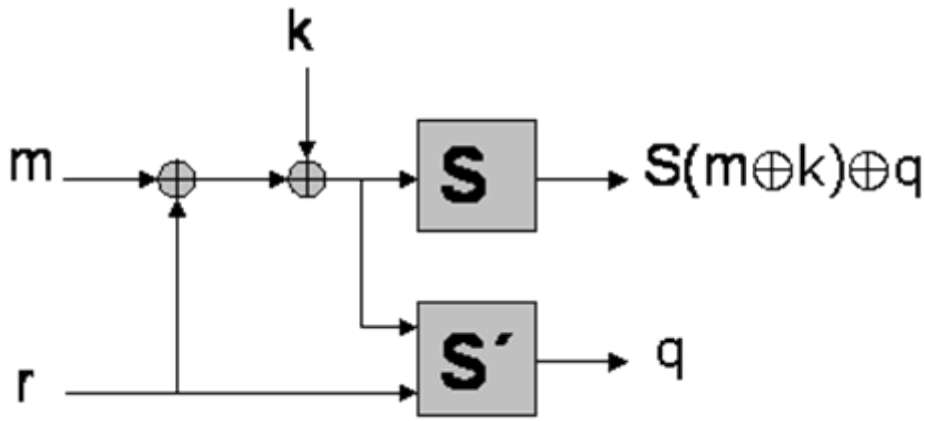
$$O = H[m \oplus k \oplus m' \oplus k]$$

$$O = H[m \oplus m']$$

  - Consumption after the S-box

$$O = H[S(m \oplus k) \oplus S(m' \oplus k)]$$

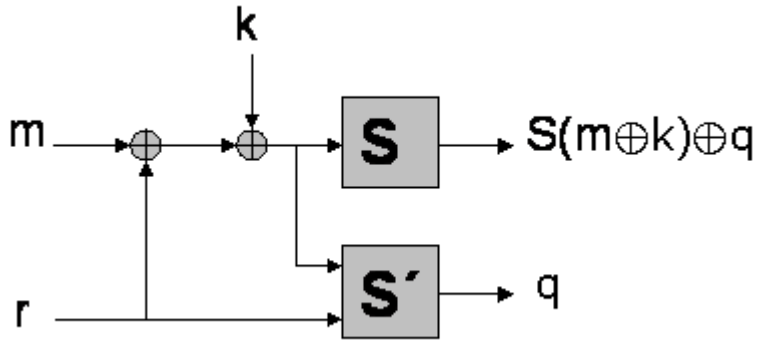
Boolean Masking



[Goubin et al. CHES '99]

- Focus on 2nd order Boolean masking
- Extendable to higher-order Boolean masking

HO Attack with Perfect Measurements



$$O = H[(S(m \oplus k) \oplus q) \oplus (S(m' \oplus k) \oplus q')] + H[q \oplus q']$$

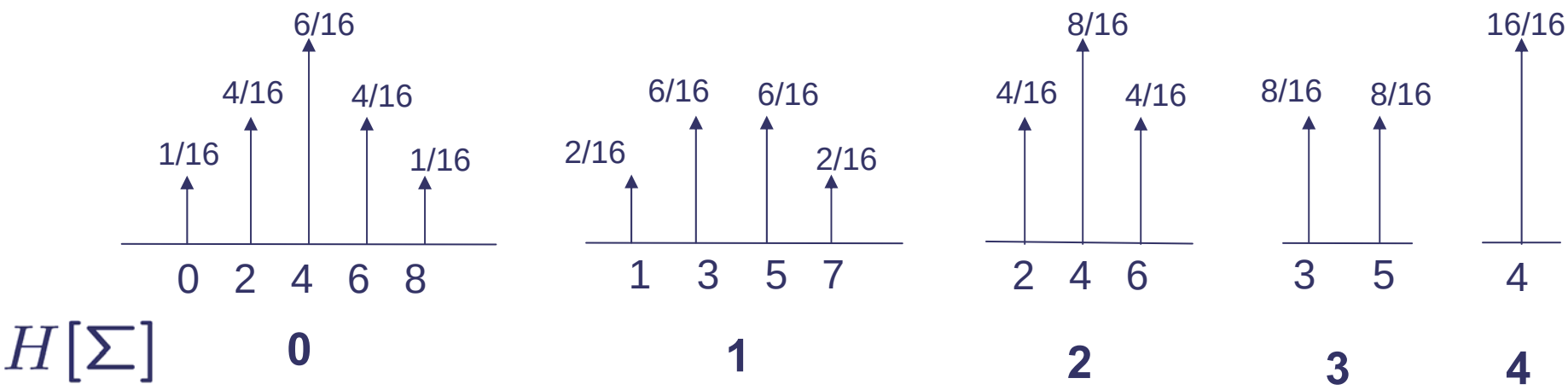
- Let us define the **machine state**

$$\Sigma = S(m \oplus k) \oplus S(m' \oplus k)$$

- And the random state $R = [q \oplus q']$

$$O(\Sigma, R) = H[\Sigma \oplus R] + H[R]$$

- **Machine state** Σ directly related to a Probability Density Functions (PDFs)



PDFs $P[O|\Sigma = \sigma_i]$ with $N_b = 4$

Maximum Likelihood Approach

$$\begin{aligned}
\Sigma^*(k_0) &= \{\sigma_1(k_0), \sigma_2(k_0), ..., \sigma_n(k_0)\};\\
\Sigma^*(k_1) &= \{\sigma_1(k_1), \sigma_2(k_1), ..., \sigma_n(k_1)\};\\
\Sigma^*(k_2) &= \{\sigma_1(k_2), \sigma_2(k_2), ..., \sigma_n(k_2)\};\\
&\vdots
\end{aligned}$$

$$\begin{aligned}
P[O^*|\Sigma^*(k_0)] &= P[O = o_1|\Sigma = \sigma_1(k_0)] \times P[O = o_2|\Sigma = \sigma_2(k_0)] \times ...\\
P[O^*|\Sigma^*(k_1)] &= P[O = o_1|\Sigma = \sigma_1(k_1)] \times P[O = o_2|\Sigma = \sigma_2(k_1)] \times ...\\
P[O^*|\Sigma^*(k_2)] &= P[O = o_1|\Sigma = \sigma_1(k_2)] \times P[O = o_2|\Sigma = \sigma_2(k_2)] \times ...\\
&\vdots
\end{aligned}$$

Decision in favor of k

$$k = \underset{\forall\ k_j}{argmax}\, P[O^*|\Sigma^*(k_j)]$$

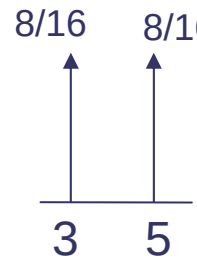
Example

- Correct key guess: ‘0001’
- Sequence of observation **O*={5,3,6,5,…}**

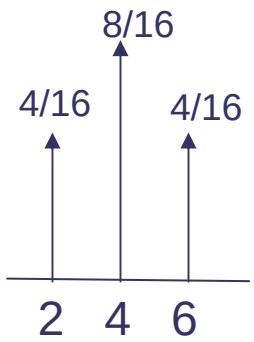
- M*={5,12,10,14,0,…}
- R*={3,11,6,12,9,…}
- 4-bit S-box: S=[9,2,15,8,3,14,5,11,13,4,10,6,0,7,1,12]

First Guess k_0='0000'?

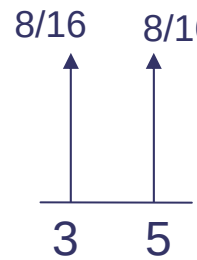
O*={5,3,6,5,…}

$H[\Sigma_1(k_0)] = 3$ → 

→ $P[O_1 = 5|\Sigma_1(k_0)] = \frac{8}{16}$

$H[\Sigma_2(k_0)] = 2$ → 

→ $P[O_2 = 3|\Sigma_2(k_0)] = 0$

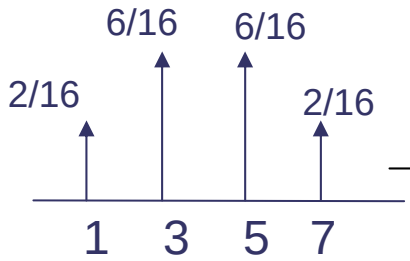
$H[\Sigma_3(k_0)] = 3$ → 

→ $P[O_3 = 6|\Sigma_3(k_0)] = 0$

Correct Guess k_1='0001'

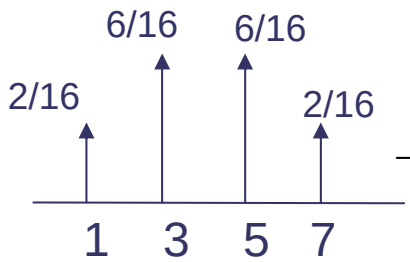
O*={5,3,6,5,…}

$$H[\Sigma_1(k_1)] = 1 \longrightarrow$$



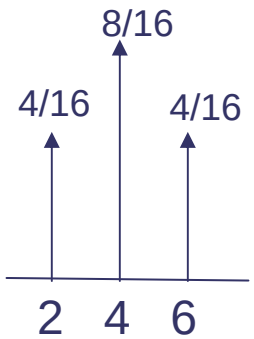
$$\longrightarrow P[O_1 = 5|\Sigma_1(k_1)] = \frac{6}{16}$$

$$H[\Sigma_2(k_1)] = 1 \longrightarrow$$



$$\longrightarrow P[O_2 = 3|\Sigma_2(k_1)] = \frac{6}{16}$$

$$H[\Sigma_3(k_1)] = 2 \longrightarrow$$



$$\longrightarrow P[O_3 = 6|\Sigma_3(k_1)] = \frac{4}{16}$$

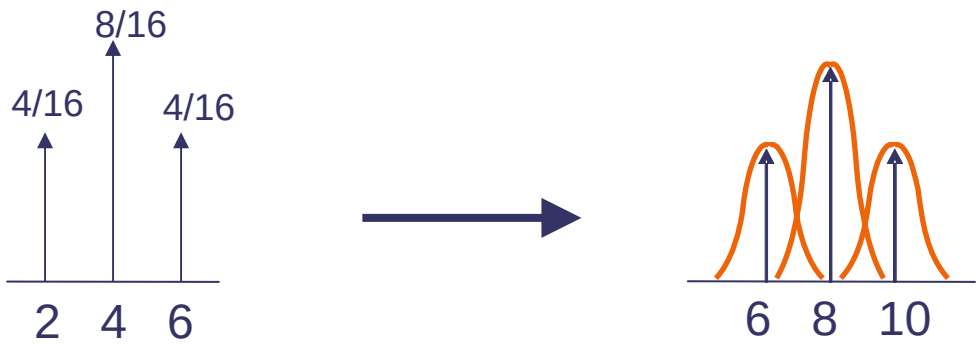
Compute the Likelihood for Each Key Guess

$$
\begin{aligned}
P[O^*|\Sigma^*(k_0)] &= P[O_1 = 5|\Sigma_1(k_0)] \times P[O_2 = 3|\Sigma_2(k_0)] \times \ldots \\
&= \frac{8}{16} \times 0 \times \ldots = 0
\end{aligned}
$$

$$
\begin{aligned}
P[O^*|\Sigma^*(k_1)] &= P[O_1 = 5|\Sigma_1(k_1)] \times P[O_2 = 3|\Sigma_2(k_1)] \times \ldots \\
&= \frac{6}{16} \times \frac{6}{16} \times \ldots
\end{aligned}
$$

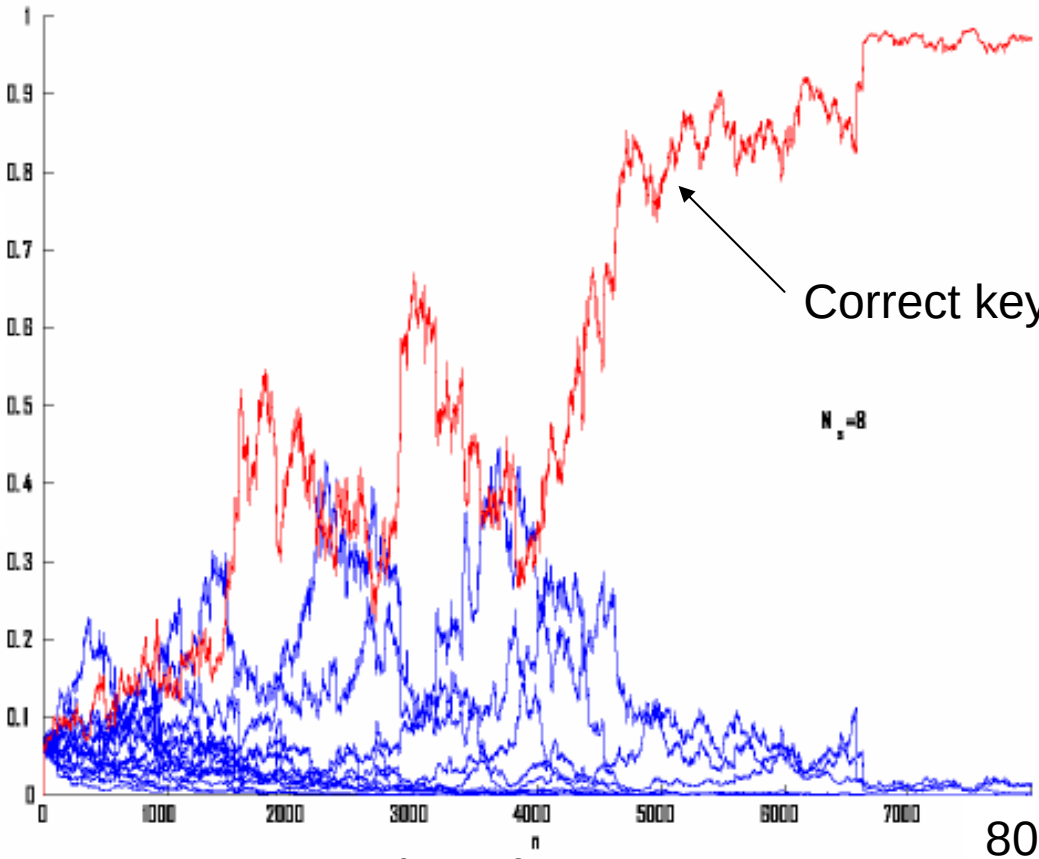
HO Attacks with Algorithmic Noise

- Due to the bit transitions of non targeted S-boxes
- The noise is assumed to be Gaussian

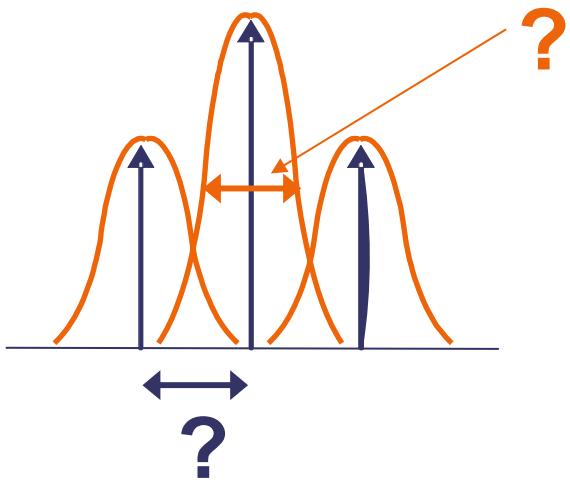


Simulated Attack With 8 S-boxes in Parallel: N_S=8

$P[O^*|\Sigma^*(k_i)]$



HO Attacks with Real Measurements

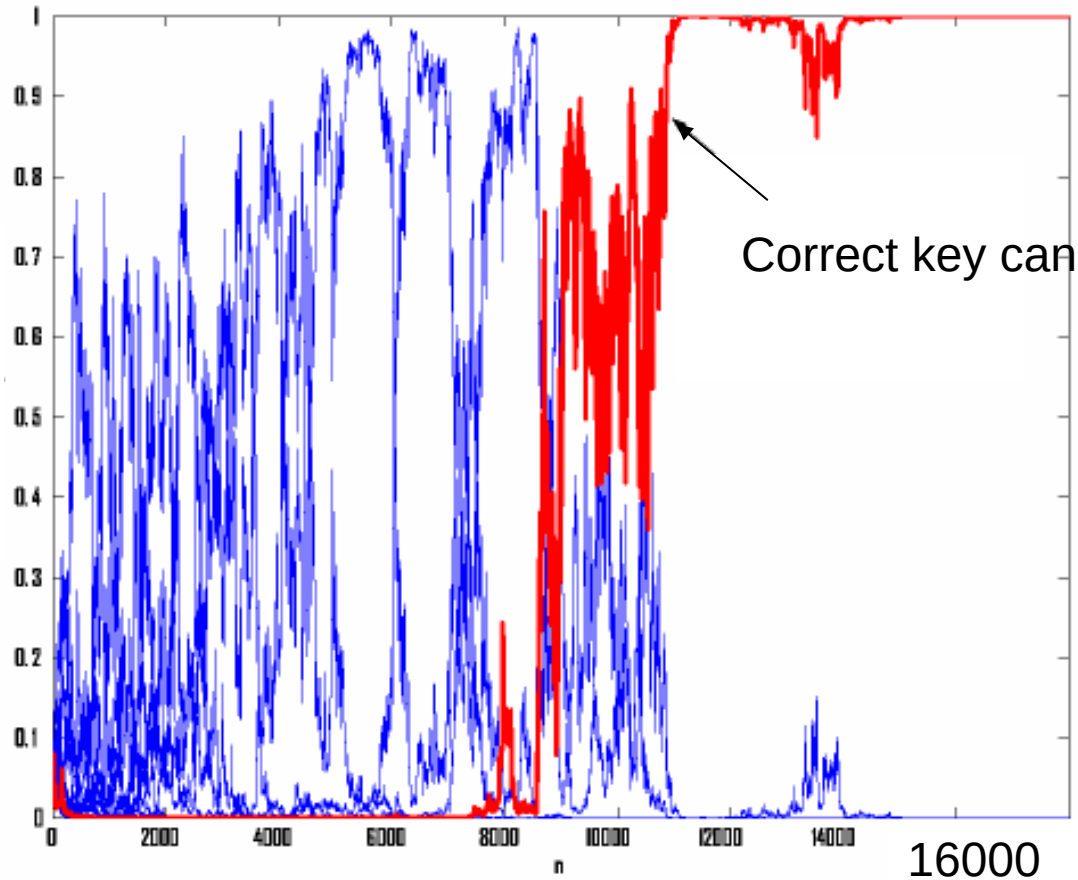


- PDFs? ➔ Require **similar chip** to evaluate the mean and variance of each Gaussians

  ➔ Or **machine learning** methods to extract the mixture of Gaussians' characteristics

FPGA Results

$P[O^*|\Sigma^*(k_i)]$



Correct key candidate

16000

Number of measurements

Comparison with Previous Work

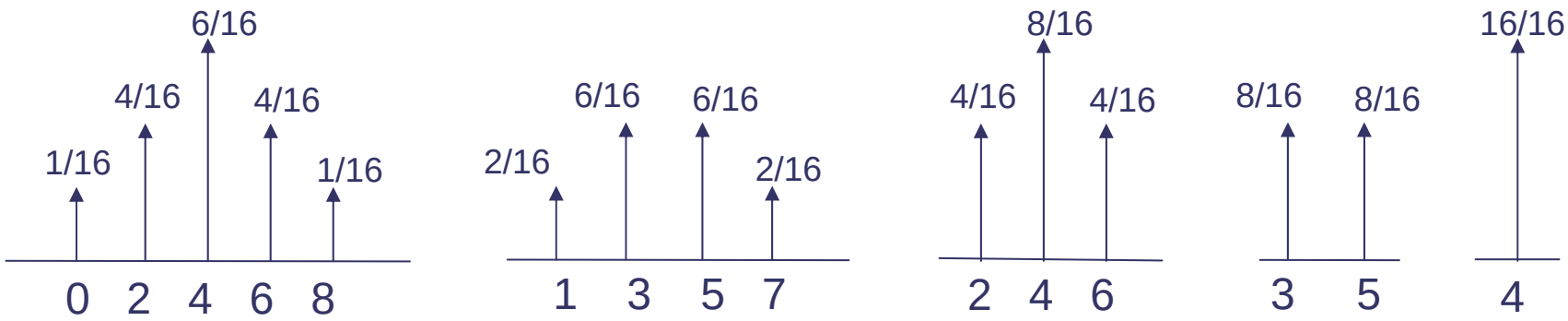
- Second Order Power Attack on FPGA

Correlation	>>	Maximum likelihood
~130,000		~12,000
N_S=1		N_S=8
[Standaert et al. ITCC '05]		**Ours**

≈ Multiple bits Waddle's attack

[Waddle et al. CHES '04]

Remark



- $E[x^2]=E[x]^2+V[x]$ **[Waddle et al. CHES '04]**

➔ Squaring the trace yields key dependency since the variances are key dependent

Conclusions

- A key guess defines a succession of **machine states** related to different **PDFs**

  ➔ We can compute the likelihood of the different key candidates from a sequence of observations
- Maximum likelihood is very efficient in this context
- FPGAs considered to be a challenging devices for SCA: noise
- Applicable to other devices: simulation with a 8-bit microprocessor → ~ 50 measurements

Questions?