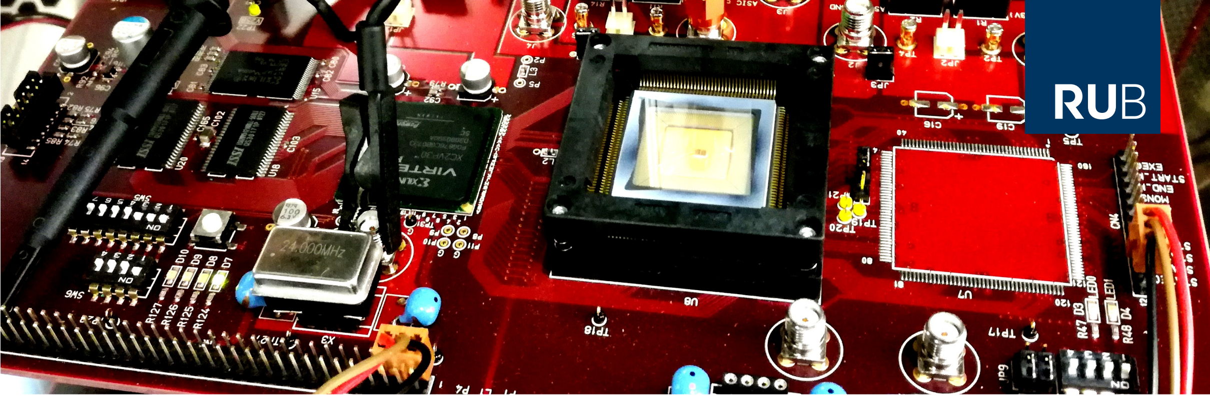


RUB

RUHR-UNIVERSITÄT BOCHUM

On the Easiness of Turning Higher-Order Leakages into First-Order

Thorben Moos and Amir Moradi
Horst-Görtz Institute for IT Security
Ruhr-Universität Bochum

14th April, 2017

DFG Deutsche Forschungsgemeinschaft

Outline

Leakage Assumption: Noisy Hamming Weight Model

Masked and Unmasked Leakage

RUB

Unmasked Implementation

$$l(x) = HW(x) + \mathcal{N}(\mu, \delta^2)$$

$$x \in \{0,1\}^4,\ \mu = 0,\ \delta = 2$$

First-Order Boolean Masked Implementation

$$l(x_m) + l(m) = HW(x_m) + HW(m) + \mathcal{N}(\mu, \delta^2)$$

$$x \in \{0,1\}^4,\ m \leftarrow \{0,1\}^4,\ x_m = x \oplus m,\ \mu = 0,\ \delta = 2$$

Unmasked Implementation

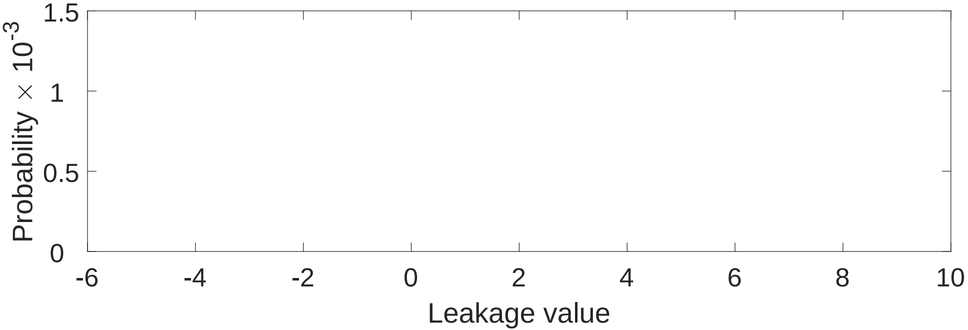
Introduction

$$x = 0000_2$$

$$l(x) = HW(0000_2) + \mathcal{N}(0, 2^2)$$

$$x = 1111_2$$

$$l(x) = HW(1111_2) + \mathcal{N}(0, 2^2)$$

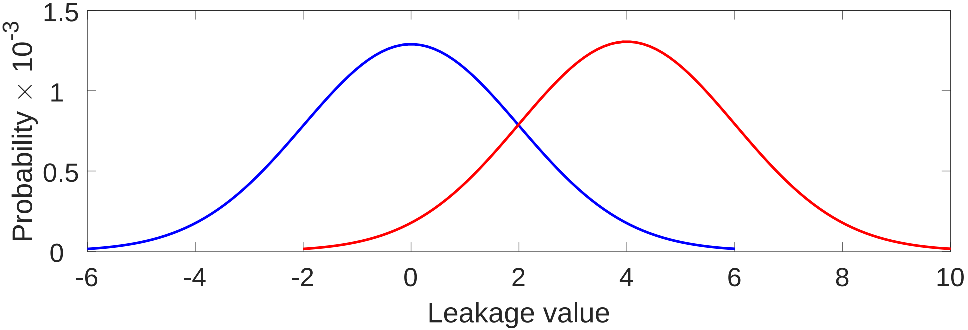


Unmasked Implementation

Masked and Unmasked Leakage

$$x = 0000_2$$
$$l(x) = HW(0000_2) + \mathcal{N}(0, 2^2)$$
$$l(x) = 0 + \mathcal{N}(0, 2^2)$$

$$x = 1111_2$$
$$l(x) = HW(1111_2) + \mathcal{N}(0, 2^2)$$
$$l(x) = 4 + \mathcal{N}(0, 2^2)$$

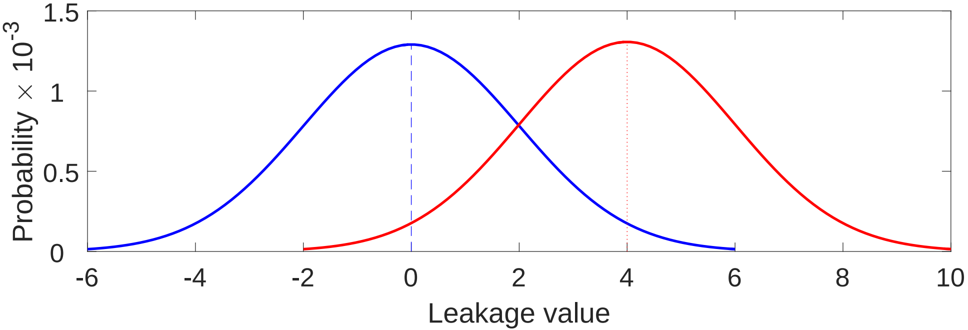


Unmasked Implementation

Masked and Unmasked Leakage

$$x = 0000_2$$
$$l(x) = HW(0000_2) + \mathcal{N}(0, 2^2)$$
$$l(x) = 0 + \mathcal{N}(0, 2^2)$$
$$E(l(x)) = 0$$

$$x = 1111_2$$
$$l(x) = HW(1111_2) + \mathcal{N}(0, 2^2)$$
$$l(x) = 4 + \mathcal{N}(0, 2^2)$$
$$E(l(x)) = 4$$



First-Order Boolean Masked Implementation

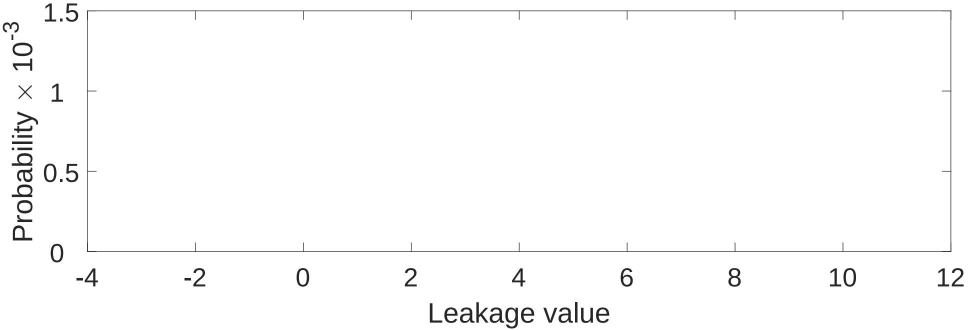
Masked and Unmasked Leakage

$x = 0000_2$

$l(x_m) + l(m) = HW(0000_2 \oplus m) + \ldots$

$x = 1111_2$

$l(x_m) + l(m) = HW(1111_2 \oplus m) + \ldots$

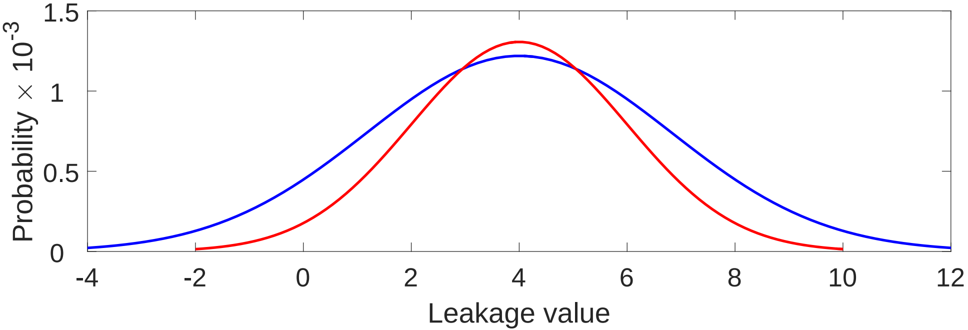


First-Order Boolean Masked Implementation

Masked and Unmasked Leakage

$$x = 0000_2$$
$$l(x_m) + l(m) = HW(0000_2 \oplus m) + \ldots$$
$$l(x_m) + l(m) = 2 \cdot HW(m) + \mathcal{N}(0, 2^2)$$

$$x = 1111_2$$
$$l(x_m) + l(m) = HW(1111_2 \oplus m) + \ldots$$
$$l(x_m) + l(m) = 4 + \mathcal{N}(0, 2^2)$$



First-Order Boolean Masked Implementation

Masked and Unmasked Leakage

$$x = 0000_2$$

$$l(x_m) + l(m) = HW(0000_2 \oplus m) + \ldots$$

$$l(x_m) + l(m) = 2 \cdot HW(m) + \mathcal{N}(0, 2^2)$$

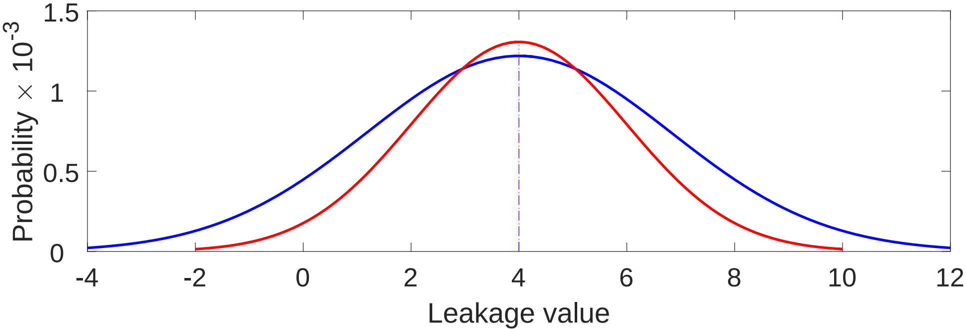
$$E(l(x_m) + l(m)) = 4$$

$$x = 1111_2$$

$$l(x_m) + l(m) = HW(1111_2 \oplus m) + \ldots$$

$$l(x_m) + l(m) = 4 + \mathcal{N}(0, 2^2)$$

$$E(l(x_m) + l(m)) = 4$$

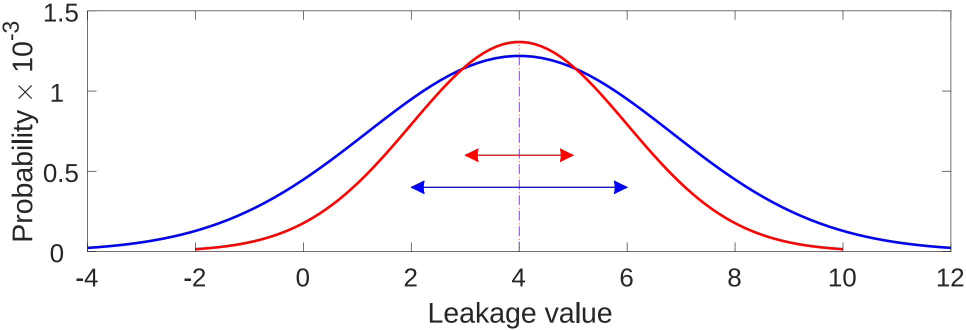


Higher-Order Statistical Moments

Masked and Unmasked Leakage

Usually assumed adversarial strategy:

Estimating second-order centered moments (= variances) to distinguish distributions



Higher-Order Statistical Moments

Masked and Unmasked Leakage

Usually assumed adversarial strategy:

Estimating second-order centered moments (= variances) to distinguish distributions

BUT: There are some limitations

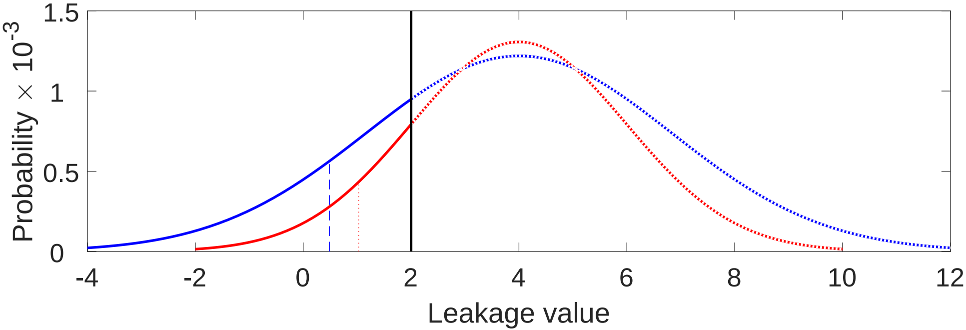
- Complexity increases exponentially with the order to be estimated
- Estimation is very sensitive to the noise level

Any Simple Alternatives?

Novel Approach

Our observation:

First-order moments (= means) can be used to distinguish **slices** of the distributions



Any Simple Alternatives?

Novel Approach

Our observation:

First-order moments (= means) can be used to distinguish **slices** of the distributions

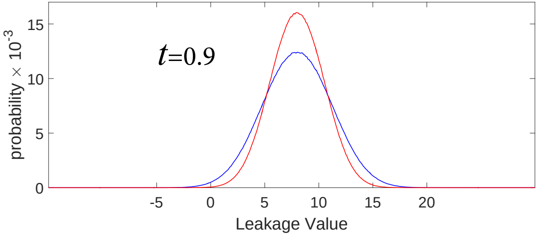
Can this be useful or advantageous in practice?

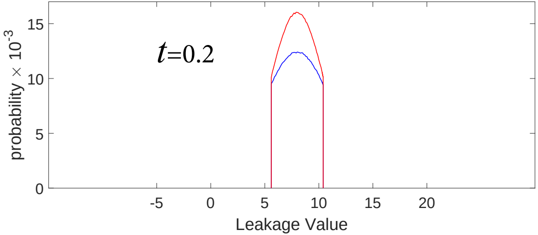
1. How to choose the slices/thresholds?
2. Does the concept apply to higher-order masking as well?
3. Is it able to outperform higher-order distinguishers (for specific settings)?
4. Is it suitable for real-world measurements (i.e. not perfectly gaussian noise)?

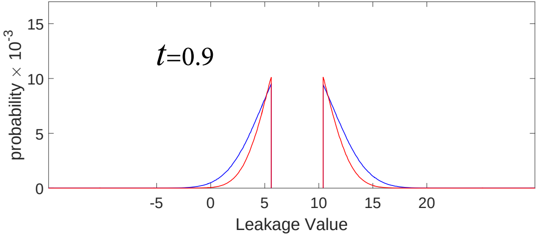
t Statistics: First-Order Masking – Unsuitable Slices

Distinguishability

- 1 million simulations
- two different $\mathrm{x} \in \{0, 1\}^8$
- random/uniform $\mathrm{m} \leftarrow \{0, 1\}^8$
- $\mu = 0, \ \delta = 2$

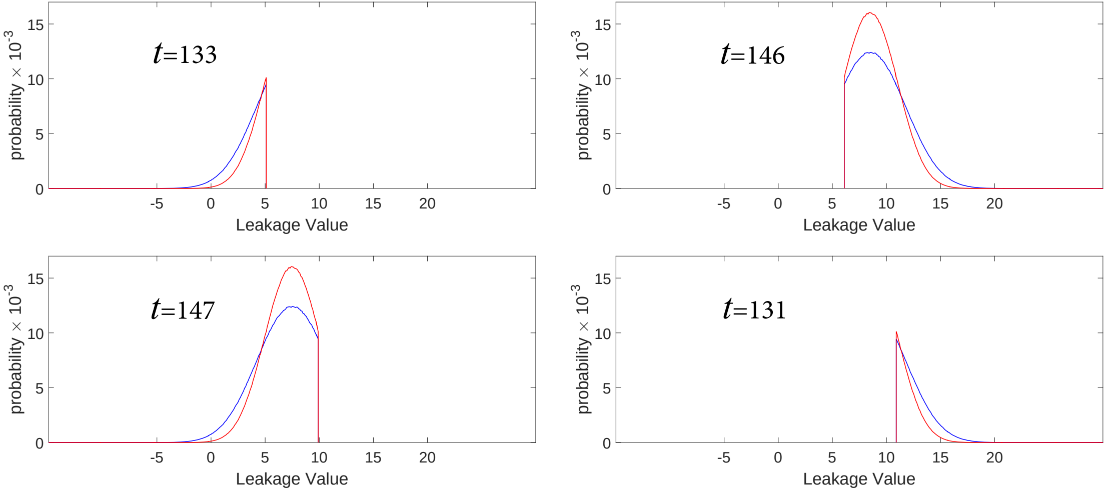






t Statistics: First-Order Masking – Suitable Slices

Distinguishability

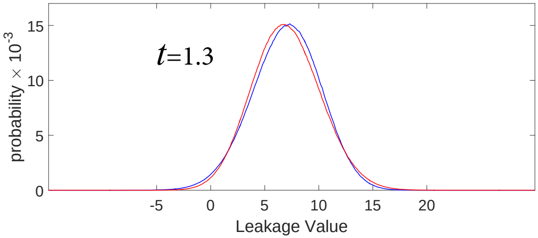


t Statistics: Second-Order Masking – Unsuitable Slices

Distinguishability

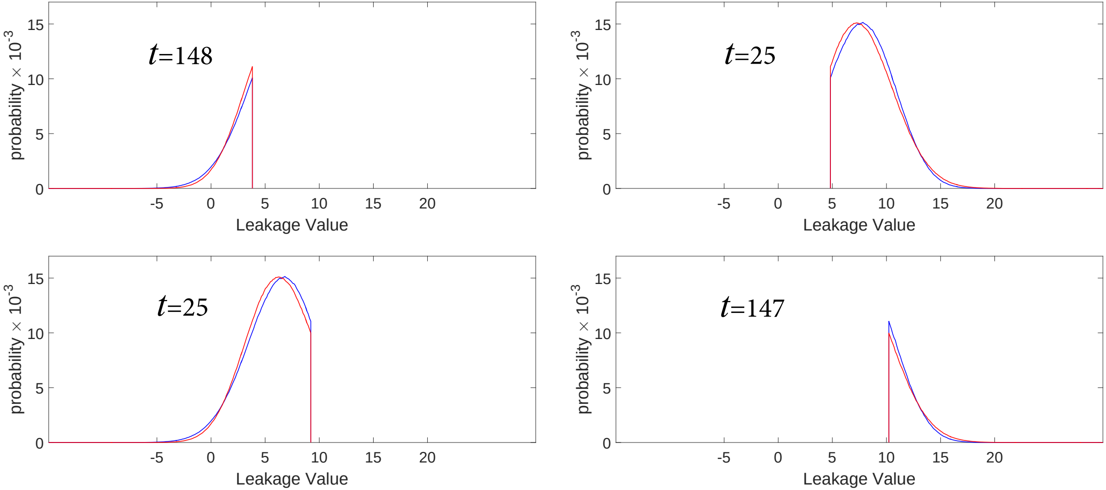
Note: Second-order masked leakage distributions are usually distinguished by their third-order statistical moment (= skewness)

- 1 million simulations
- two different $\mathrm{x} \in \{0,1\}^8$
- random/uniform $\mathrm{m} \leftarrow \{0,1\}^8$
- $\mu = 0,\ \delta = 2$



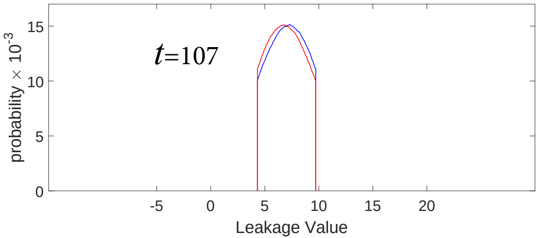
t Statistics: Second-Order Masking – Suitable Slices

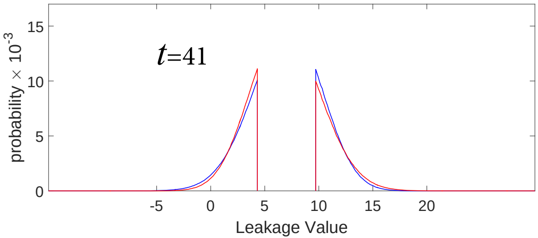
Distinguishability



t Statistics: Second-Order Masking – Suitable Slices

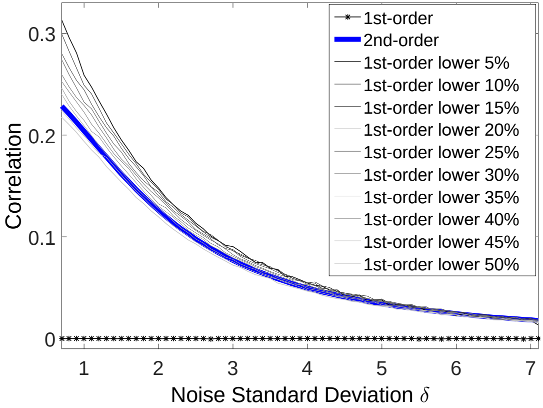
Distinguishability

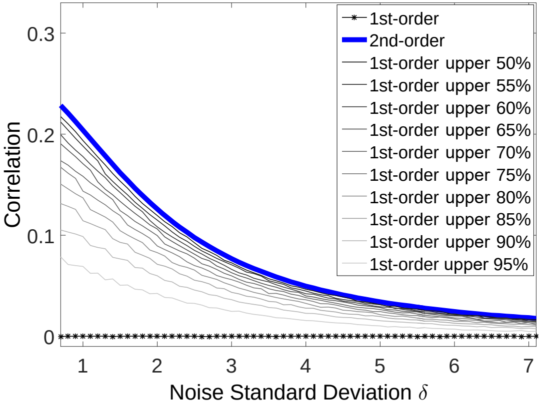




Different Slices – First-Order Masking

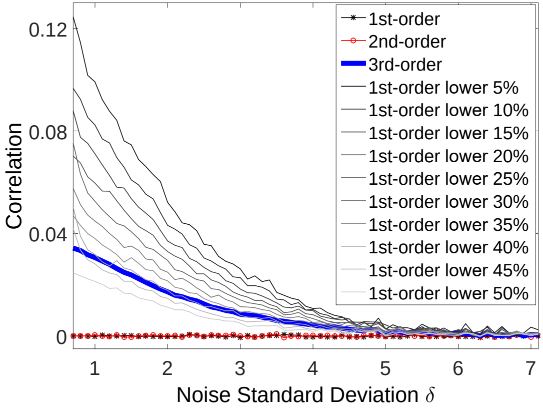
Correlation Comparison

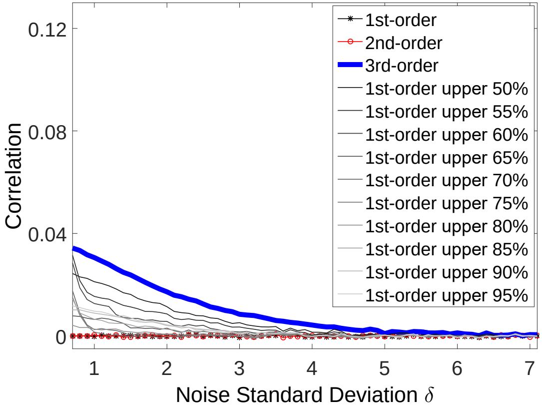




Different Slices – Second-Order Masking

Correlation Comparison



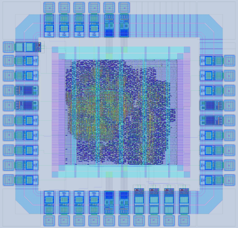


PRESENT-80 Threshold Implementation Chip

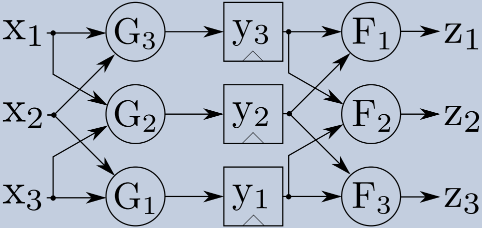
Target

RUB

150 nm ASIC Prototype with nibble-serial PRESENT-80 Threshold Implementation Core



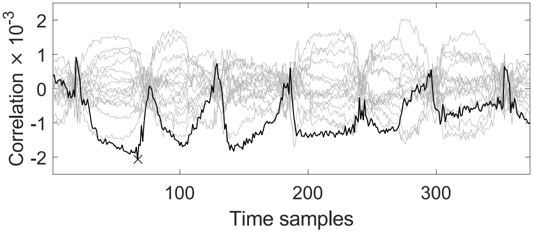
(a) Layered view of 150nm ASIC

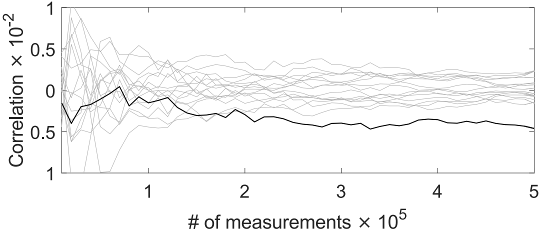


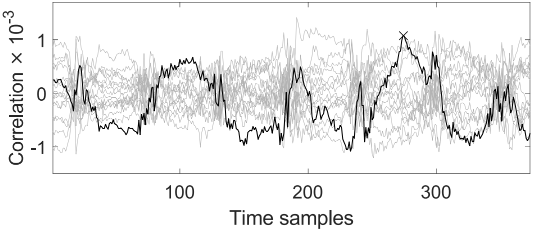
(b) Threshold implementation of the 4-bit PRESENT-80 S-Box

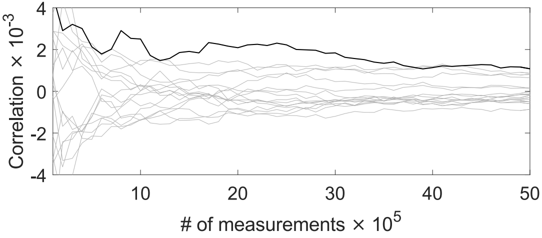
Conventional Second- and Third-Order CPA

Results



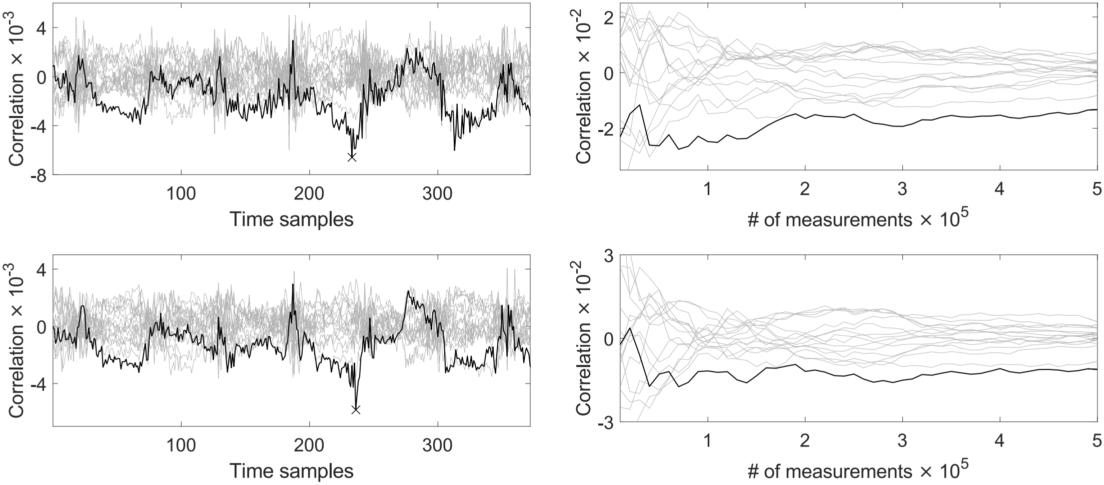






First-Order CPA on Upper 20% and Upper 30% Slices

Results



Quantitative Comparison

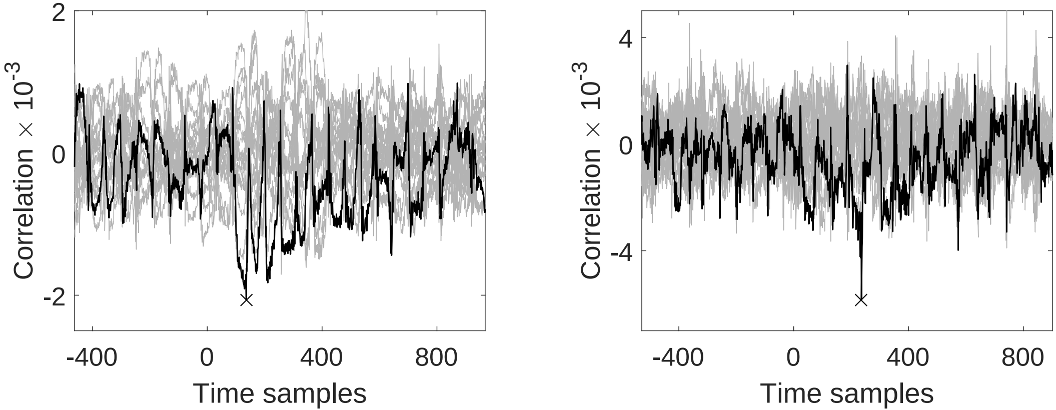
Results

Up to 4 Times Less Traces Required

Stat. Order	Slice	MTD
1st	100 %	–
2nd	100 %	200,000
3rd	100 %	>5,000,000
1st	Upper 15 %	700,000
1st	Upper 20 %	50,000
1st	Upper 25 %	70,000
1st	Upper 30 %	70,000
1st	Upper 35 %	90,000
1st	Upper 40 %	800,000

Visual Comparison

Results



Conclusion and Future Work

Conclusion

RUB

Conclusion

- Masked leakage distributions can be attacked by first-order distinguishers
- No estimation of higher-order moments required
- Might be able to relax sensitivity of higher-order evaluations to the noise level
- Case study shows that it can succeed with fewer measurements

Future Work

- More quantitative case study – Implementations with Masking + Hiding
- Combine attacks on different slices (Useful for leakage detection?)

Thank you for your attention.

Any questions?