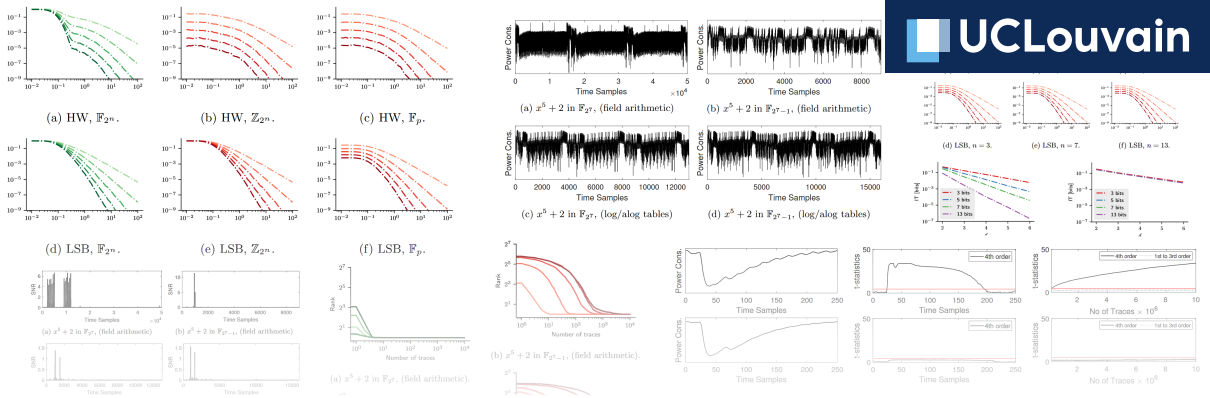


UCLouvain Institute of Information and Communication Technologies, Electronics and Applied Mathematics (ICTEAM)

Effective and Efficient Masking with Low Noise using Small-Mersenne-Prime Ciphers

Loïc Masure, Pierrick Méaux, Thorben Moos, François-Xavier Standaert

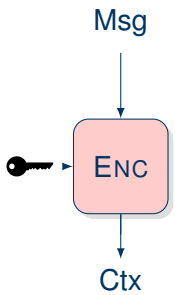
Crypto Group, ICTEAM Institute, UCLouvain, Louvain-la-Neuve, Belgium.

Luxembourg University, SnT, Luxembourg.

April 24th, 2023

European Research Council

Context : Side-Channel Analysis (SCA)

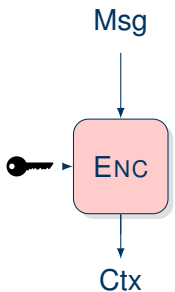


: N bits

- Black-box cryptanalysis: 2^N

Context : Side-Channel Analysis (SCA)

"Cryptographic algorithms don't run on paper,

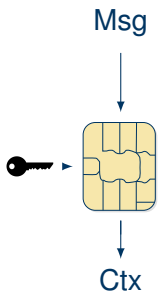


: N bits

- Black-box cryptanalysis: 2^N

Context : Side-Channel Analysis (SCA)

"Cryptographic algorithms don't run on paper, they run on physical devices"

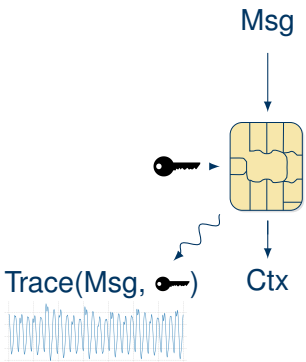


: N bits

- Black-box cryptanalysis: 2^N

Context : Side-Channel Analysis (SCA)

“Cryptographic algorithms don’t run on paper, they run on physical devices”



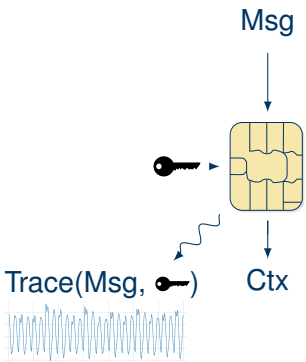
: N bits

- Black-box cryptanalysis: 2^N

Trace : power, EM, acoustics, runtime, . . .

Context : Side-Channel Analysis (SCA)

"Cryptographic algorithms don't run on paper, they run on physical devices"



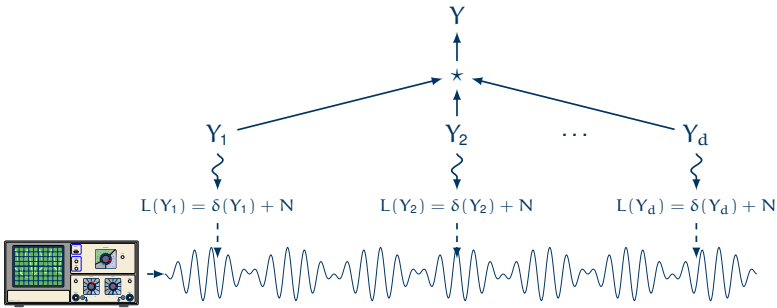
: N bits

- Black-box cryptanalysis: 2^N
- Side-Channel Analysis: $2^n \cdot \frac{N}{n}, n \ll N$

Trace : power, EM, acoustics, runtime, ...

Masking

Masking, aka *MPC on silicon*: linear secret sharing over a finite field $(\mathbb{F}, \star, \cdot)$



The Effect of Masking

Simulation, for $\mathbb{F}_{2^n}$: $L(Y_i) = \mathcal{hw}(Y_i) + \mathcal{N}(0; \sigma^2)$, $\mathcal{hw}$ = Hamming weight

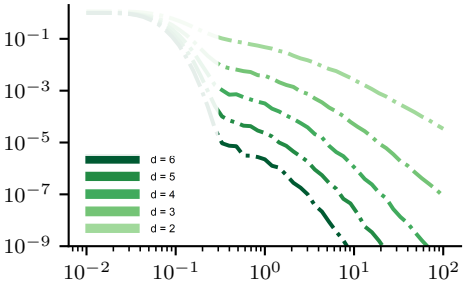


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

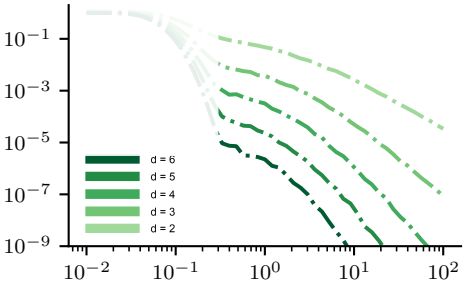
Observation: "Masking amplifies noise"
Constant gap between each curve (log scale)

$\Longleftrightarrow$

exponential security w.r.t. #shares d

The Effect of Masking

Simulation, for $\mathbb{F}_{2^n}$: $L(Y_i) = hw(Y_i) + \mathcal{N}(0; \sigma^2)$, $hw =$ Hamming weight



Explanation: "Masking amplifies noise"

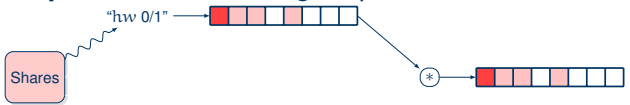
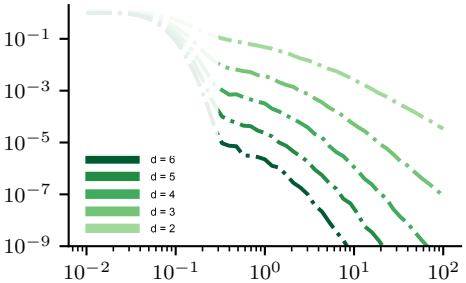


Figure $MI(Y; Trace)$ vs. σ^2, $2 \leqslant d \leqslant 6$

The Effect of Masking

Simulation, for $\mathbb{F}_{2^n}$: $L(Y_i) = hw(Y_i) + \mathcal{N}(0; \sigma^2)$, $hw =$ Hamming weight



Explanation: "Masking amplifies noise"

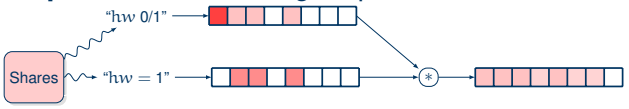
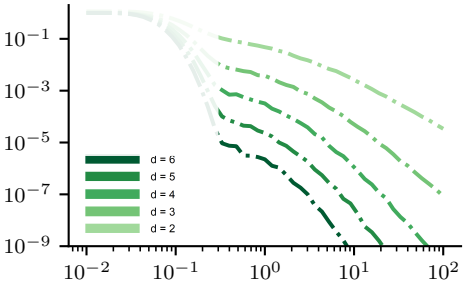


Figure $MI(Y; Trace)$ vs. σ^2, $2 \leqslant d \leqslant 6$

The Effect of Masking

Simulation, for $\mathbb{F}_{2^n}$: $L(Y_i) = hw(Y_i) + \mathcal{N}(0; \sigma^2)$, $hw =$ Hamming weight



Explanation: "Masking amplifies noise"

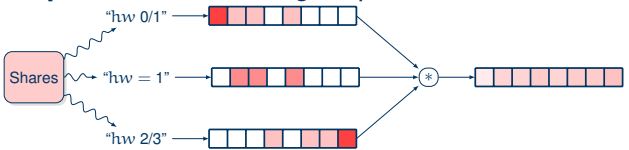


Figure $MI(Y; Trace)$ vs. σ^2, $2 \leqslant d \leqslant 6$

Masking in a Low-Noise Setting

Does masking always work in a low-noise setting ?

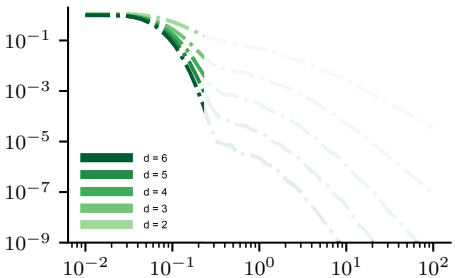


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant \mathrm{d} \leqslant 6$

Observation:

Secret always leaks $>$ 1 bit, regardless of d !

Masking in a Low-Noise Setting

Does masking always work in a low-noise setting ?

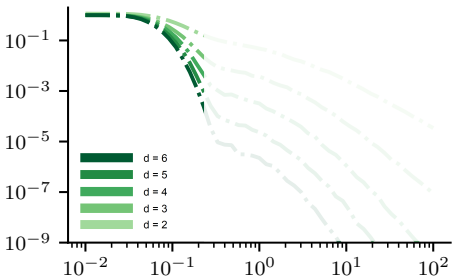


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

Explanation:

$$hw(Y_1 \oplus \ldots \oplus Y_d) = \sum_i hw(Y_i) - 2 \cdot (\ldots)$$

Masking in a Low-Noise Setting

Does masking always work in a low-noise setting ?

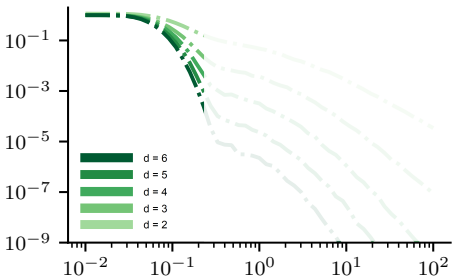


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

Explanation:

$hw(Y_1 \oplus \ldots \oplus Y_d) = \sum_i hw(Y_i) - 2 \cdot (\ldots)$

Parity of $hw(Y)$ stable by $\oplus$: **subgroup of** $\mathbb{F}_{2^n}$

Masking in a Low-Noise Setting

Does masking always work in a low-noise setting ?

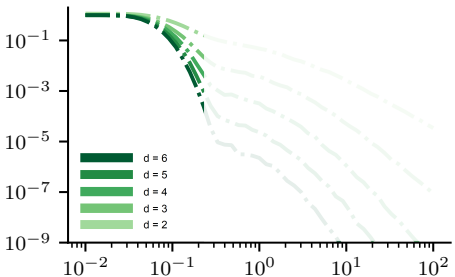
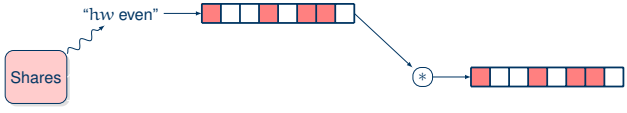


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

Explanation:

$hw(Y_1 \oplus \ldots \oplus Y_d) = \sum_i hw(Y_i) - 2 \cdot (\ldots)$

Parity of $hw(Y)$ stable by $\oplus$: **subgroup of** $\mathbb{F}_{2^n}$



Masking in a Low-Noise Setting

Does masking always work in a low-noise setting ?

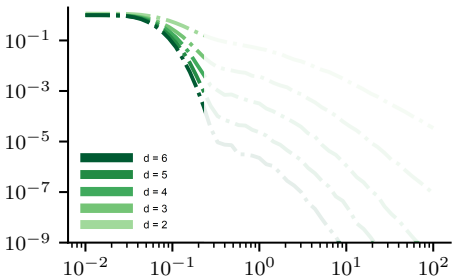
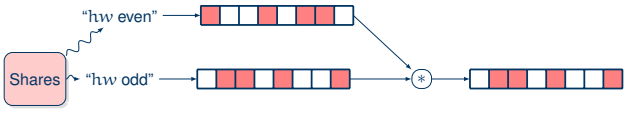


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

Explanation:

$\mathrm{hw}(Y_1 \oplus \ldots \oplus Y_d) = \sum_i \mathrm{hw}(Y_i) - 2 \cdot (\ldots)$

Parity of $\mathrm{hw}(Y)$ stable by $\oplus$: **subgroup of** $\mathbb{F}_{2^n}$



Masking in a Low-Noise Setting

Does masking always work in a low-noise setting ?

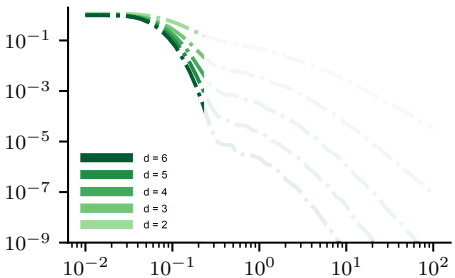
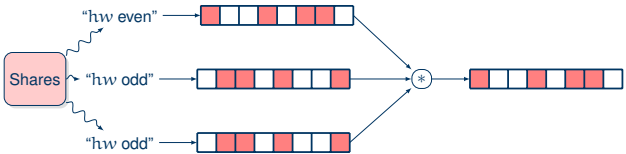


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

Explanation:

$\mathrm{hw}(Y_1 \oplus \ldots \oplus Y_d) = \sum_i \mathrm{hw}(Y_i) - 2 \cdot (\ldots)$

Parity of $\mathrm{hw}(Y)$ stable by $\oplus$: **subgroup of** $\mathbb{F}_{2^n}$



Masking in a Low-Noise Setting

Does masking always work in a low-noise setting ?

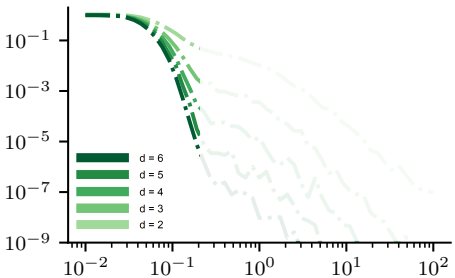
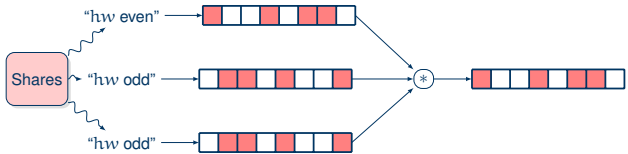


Figure $\mathrm{MI}(Y; \mathrm{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

Explanation:

$\mathrm{hw}(Y_1 \oplus \ldots \oplus Y_d) = \sum_i \mathrm{hw}(Y_i) - 2 \cdot (\ldots)$

Parity of $\mathrm{hw}(Y)$ stable by $\oplus$: **subgroup of** $\mathbb{F}_{2^n}$



Corollary: parallelism is no cure either

Two Solutions for Sound Masking

What conditions the distributions  of each share must fit?

Theorem[1]

Soundness $\iff$ support not contained in any non-trivial subgroup (or coset) of $\mathbb{F}$

Two solutions:

- Leak < 1 bit per share:[2]
  - Support of PMF always larger than any coset
  - Work with any $\mathbb{F}$ (usually chosen to fit the cipher) ✓
  - **Leakage-dependent: not always verified** ✗
- Choose $\mathbb{F}$ without any non-trivial subgroup $\implies$ **work over $\mathbb{F}_p$, with p prime**:
  - May not fit every block cipher ✗
  - No assumption on the leakage ✓

[1] Stromberg 1960; Dziembowski, Faust, and Skórski 2016.

[2] Béguinot et al. 2023.

Two Solutions for Sound Masking

What conditions the distributions  of each share must fit?

Theorem[1]

Soundness $\iff$ support not contained in any non-trivial subgroup (or coset) of $\mathbb{F}$

Two solutions:

- Leak $<$ 1 bit per share:[2]
  - Support of PMF always larger than any coset
  - Work with any $\mathbb{F}$ (usually chosen to fit the cipher) ✓
  - **Leakage-dependent: not always verified ✗**
- Choose $\mathbb{F}$ without any non-trivial subgroup $\implies$ **work over $\mathbb{F}_p$, with p prime**:
  - May not fit every block cipher ✗
  - No assumption on the leakage ✓

[1] Stromberg 1960; Dziembowski, Faust, and Skórski 2016.

[2] Béguinot et al. 2023.

Two Solutions for Sound Masking

What conditions the distributions  of each share must fit?

Theorem[1]

Soundness $\iff$ support not contained in any non-trivial subgroup (or coset) of $\mathbb{F}$

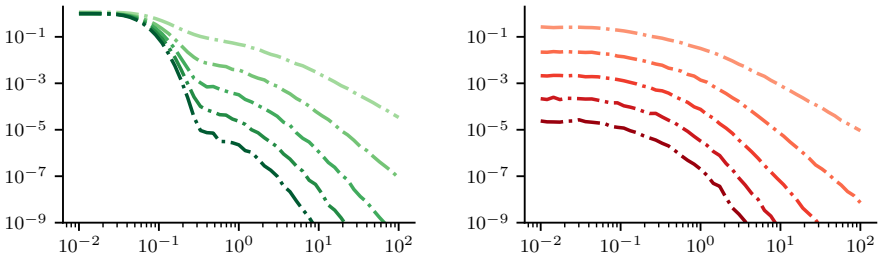
Two solutions:

- Leak $<$ 1 bit per share:[2]
  - Support of PMF always larger than any coset
  - Work with any $\mathbb{F}$ (usually chosen to fit the cipher) ✓
  - **Leakage-dependent: not always verified** ✗
- Choose $\mathbb{F}$ without any non-trivial subgroup $\implies$ **work over $\mathbb{F}_p$, with p prime**:
  - May not fit every block cipher ✗
  - No assumption on the leakage ✓

[1] Stromberg 1960; Dziembowski, Faust, and Skórski 2016.

[2] Béguinot et al. 2023.

Second Solution: work over $\mathbb{F}_p$



(a) HW, $\mathbb{F}_{2^n}$ (b) HW, $\mathbb{F}_p$

Figure $\mathrm{MI}(Y; \mathtt{Trace})$ vs. σ^2, $2 \leqslant d \leqslant 6$

- Not only masking over $\mathbb{F}_p$ safer in low-noise settings ...
- The advantage is also kept in high-noise settings !

How to leverage?

Q: How can we make use of masking in $\mathbb{F}_p$ to effectively and efficiently protect crypto implementations?

A: Ideally, we need algorithms that work in implementation-friendly prime fields, such as small-Mersenne-prime fields ($\mathbb{F}_{2^n-1}$), and use only simple field arithmetic $(+,\ -,\ \cdot)$

AES-prime: An AES-like toy cipher adapted for prime-field masking

- Based on arithmetic addition/multiplication modulo a prime, applied to 4×4 state
- Small Mersenne prime, i.e., $p = 2^7 - 1$, for efficient reduction (and constant mult.)
- Sbox is based on a small power map in $\mathbb{F}_p$ (bijection without fixed point)
- MixColumns is a 4×4 MDS matrix over $\mathbb{F}_p$
- Security claim: Attack complexity $\geqslant 2^{7 \cdot 16}$ with 14 cipher rounds

$$S(x) = x^5 + 2 \bmod p \qquad M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 16 \\ 1 & 4 & 16 & 2 \\ 1 & 16 & 2 & 4 \end{bmatrix}$$

Complex in Software? Not really!

Field Addition in $\mathbb{F}_{2^n-1}$ in C/C++ and ARM Assembly ($c = a + b \bmod p$)

```
c = a+b;                                    ADD r0,r0,r1
                                            UBFX r1,r0,#0,#n
c = (c & p) + (c >> n);                     ADD r0,r1,r0,ASR #n
```

Field Multiplication in $\mathbb{F}_{2^n-1}$ in C/C++ and ARM Assembly ($c = a \cdot b \bmod p$)

```
c = a*b;                                    MUL r0,r1,r0
                                            UBFX r1,r0,#0,#n
c = (c & p) + (c >> n);                     ADD r0,r1,r0,ASR #n
                                            UBFX r1,r0,#0,#n
c = (c & p) + (c >> n);                     ADD r0,r1,r0,ASR #n
```

- Only works for sufficiently small integers ($<$ 16 bit for multiplication operands on ARM Cortex-M3)
- If $c < p$ is strictly needed for the addition result, then $c \stackrel{?}{=} p$ needs to be checked after reduction

Complex in Hardware? Not really!

Field Addition in $\mathbb{F}_{2^n-1}$ in VHDL ($c = a + b \bmod p$)

```
ab   <= ('0' & a) + ('0' & b);
c    <= ab(n-1 downto 0) + ('0' & ab(n));
```

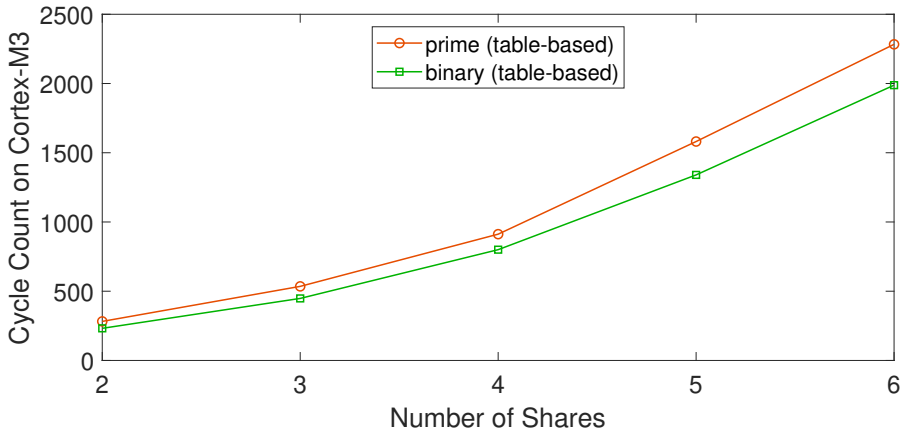
Field Multiplication in $\mathbb{F}_{2^n-1}$ in VHDL ($c = a \cdot b \bmod p$)

```
ab   <= a * b;
ab_r <= ('0' & ab(n-1 downto 0)) + ('0' & ab(2*n-1 downto n));
c    <= ab_r(n-1 downto 0) + ('0' & ab_r(n));
```

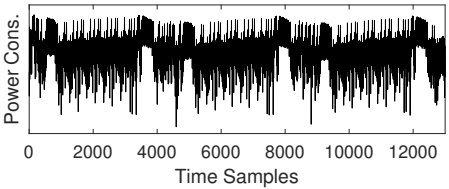
- Works with `NUMERIC_STD` package as well as the proprietary `STD_LOGIC_ARITH` & `STD_LOGIC_UNSIGNED` packages
- If $c < p$ is strictly needed for the addition result, then $c \stackrel{?}{=} p$ needs to be checked after reduction

Software Case Study: Masked S-box

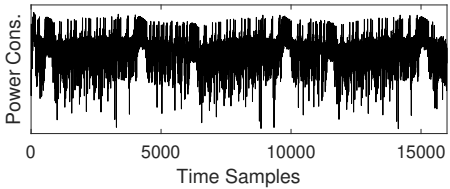
Naive implementation of masked $x^5 + 2$ using 3 consecutive ISW multiplications:



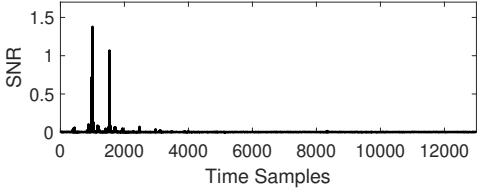
Masked $x^5 + 2$ (naive) in Software, Log/Alog tables



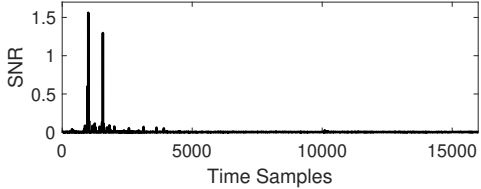
(a) Cortex-M3 sample trace, field $\mathbb{F}_{2^7}$.



(b) Cortex-M3 sample trace, field $\mathbb{F}_{2^7-1}$.

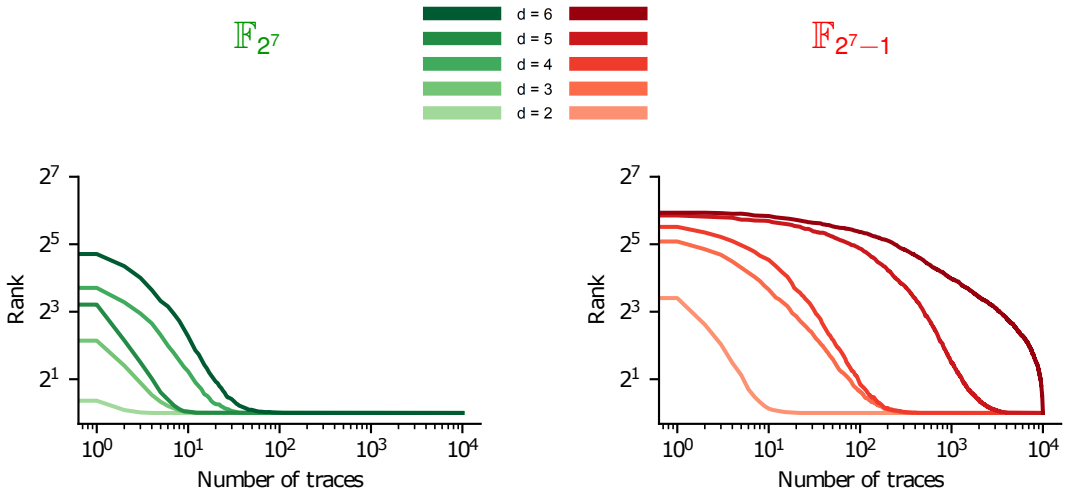


(c) SNR of input share 0, field $\mathbb{F}_{2^7}$.

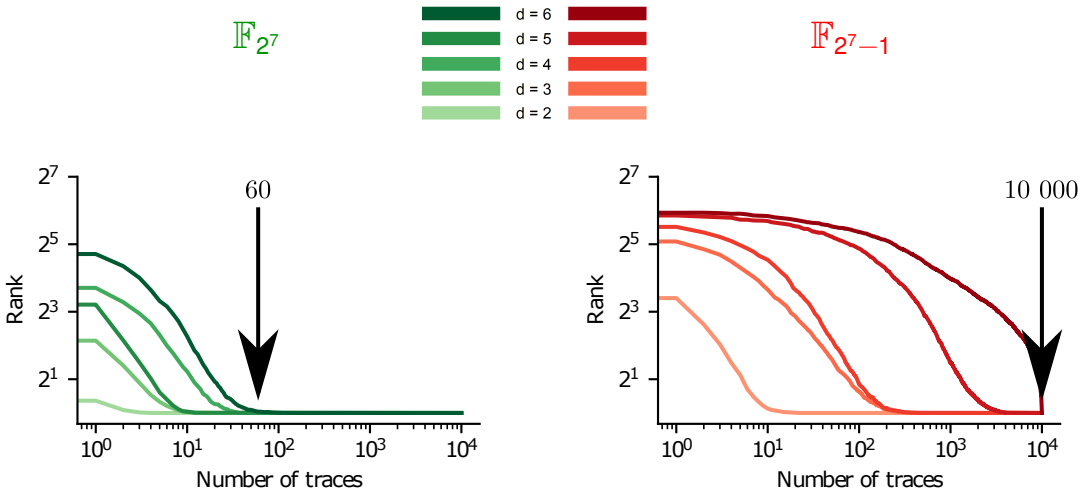


(d) SNR of input share 0, field $\mathbb{F}_{2^7-1}$.

Software, Horizontal SASCA Attack for 2-6 Shares



Software, Horizontal SASCA Attack for 2-6 Shares



Follow-Up Work at TCHES

- A follow-up of this work is already published at TCHES Volume 2023 Issue 2
- "Prime-Field Masking in Hardware and its Soundness against Low-Noise SCA Attacks"
- More hardware focused, introduces new arbitrary-order PINI gadgets for secure **squaring** in prime fields, also optimized constructions for masking the `AES-prime` S-box
- Evaluation of performance vs. physical security tradeoff between AES vs. `AES-prime`
- `https://doi.org/10.46586/tches.v2023.i2.482-518`
- `https://github.com/uclcrypto/prime_field_masking_hardware`

Conclusion

- Additive masking in small and implementation-friendly prime fields seems promising for efficient physically secure cryptography
- We can mask securely without the need to guarantee a notable amount of noise
- We demonstrated: Security advantages over Boolean masking can reach **multiple orders of magnitude** against low-noise attacks in practical experiments
- Advantages against SCA attacks with high noise, as well as fault attacks are also expected
- **New dedicated ciphers for efficient masking in prime fields are needed to explore the interest of this design space**

References I

Béguinot, J. et al. (2023). “Removing the Field Size Loss from Duc et al.’s Conjectured Bound for Masked Encodings”. In: Constructive Side-Channel Analysis and Secure Design - 14th International Workshop, CC Ed. by E. B. Kavun and M. Pehl. Vol. 13979. Lecture Notes in Computer Science. Springer, pp. 86–104. DOI: 10.1007/978-3-031-29497-6_5. URL: https://doi.org/10.1007/978-3-031-29497-6_5.

Dziembowski, S., S. Faust, and M. Skórski (2016). “Optimal Amplification of Noisy Leakages”. In: TCC (A2). Vol. 9563. LNCS. Springer, pp. 291–318.

Stromberg, K. (1960). “Probabilities on a Compact Group”. In: Transactions of the American Mathematical Society 94.2, pp. 295–309. ISSN: 00029947. URL: http://www.jstor.org/stable/1993313.