Affine Masking against Higher-Order Side Channel Analysis

Guillaume Fumaroli[1] Ange Martinelli[1] Emmanuel Prouff[2] Matthieu Rivain[3]

THALES

Oberthur Technologies

CRYPTOEXPERTS

Dec 3rd, 2010

Outline

Outline

Side Channel Analysis (SCA)

- **Physical leakage** depends on **intermediate variables**
- **Sensitive variable** depends on both the known input or output and on a guessable part of the secret key
- **SCA** exploits the physical leakage on a sensitive variable for key recovery
  - DPA, CPA *[BrierClavierOlivier04]*
  - MIA *[GierlichBatinaTuylsPreneel08]*
  - Template Attacks *[ChariRaoRohatgi02]*

Higher-Order SCA (HO-SCA)

- **d^{th}-order SCA (dO-SCA)** exploits the leakages at d different times
- Circumvented by **d^{th}-order masking (dO-masking)**
  - Every **sensitive variable** Z manipulated as $d+1$ **shares** $(R_i)_{0 \leq i \leq d}$
  - Any d-tuple of the shares R_i **independent of** Z
  - Ensure **completeness:** $Z = R_0 \star \cdots \star R_d$ for some group operation $\star$
  - Shares R_i **processed separately**
- **Parameter d:** impacts both security & efficiency
  - **+:** complexity of dO-SCA grows exponentially with d due to intrinsic leakage noise *[ChariJutlaRaoRohatgi99]*
  - **−:** efficiency of dO-masking quickly decreases with d

Higher-Order SCA (HO-SCA)

- dO-masking provides perfect dO-SCA resistance
  but remains vulnerable to $(d+1)$O-SCA
- Practical HO-SCA resistance depends on masking operation $\star$
- Block cipher masking mainly rely on Boolean masking ($\star = \oplus$)
- Unfortunately when $\star = \oplus$:
  - Leakages can be efficiently combined
    to reveal information on underlying unmasked data
  - Satisfying HO-SCA resistance obtained only for large d:
    prohibitive implementation costs in some settings
- **Alternative approach:** design an efficient scheme that
  - may not provide **perfect** dO-SCA resistance for some d
  - but provides good **practical** dO-SCA resistance for all d

Outline

Affine Masking of the AES

Description

- Every **sensitive variable** Z manipulated as

$$G(Z) = R_0 + R_1 \times Z$$

  - R_0: additive mask
  - R_1: multiplicative mask

- $Z \in \mathrm{GF}(2^n)$, $R_0 \in \mathrm{GF}(2^n)$, $R_1 \in \mathrm{GF}(2^n)^\star$
- **Related work:**
  - *[AkkarGiraud01]*: $R_0 = 0$
  - *[VonWillich01]*: $R_0 \in \mathrm{GF}(2)^n, R_1 \in \mathrm{GL}_n(\mathrm{GF}(2))$

Affine Masking of the AES

- Masking the AES amounts to mask every round transformations
  - **AddRoundKey, (Inv)ShiftRows, (Inv)MixColumns**
    - compatible with affine masking
  - (Inv)SubBytes

SubBytes

$$\underbrace{s'}_{\text{output}} = S\,[\,\underbrace{s}_{\text{input}}\,]$$

Affine Masking of the AES

- Masking the AES amounts to mask every round transformations
  - AddRoundKey, (Inv)ShiftRows, (Inv)MixColumns
    - compatible with affine masking
  - **(Inv)SubBytes**

SubBytes

$$\underset{s'}{\text{output}} = S\,[\,\underset{s}{\text{input}}\,]$$

Affine Masking of the AES

- Masking the AES amounts to mask every round transformations
  - AddRoundKey, (Inv)ShiftRows, (Inv)MixColumns
    - compatible with affine masking
  - **(Inv)SubBytes**

SubBytes

$$\underbrace{G(s')}_{\text{output}} = \tilde{S}[\underbrace{G(s)}_{\text{input}}]$$

where $\tilde{S}, G$ verify $\forall x, \tilde{S}[G(x)] = G(S[x])$

Affine Masking of the AES

- Masking the AES amounts to mask every round transformations
  - AddRoundKey, (Inv)ShiftRows, (Inv)MixColumns
    - compatible with affine masking
  - **(Inv)SubBytes**

SubBytes

$$\underset{\text{output}}{G(s')} \quad = \quad \underset{\text{input}}{\tilde{S}[G(s)]}$$

where $\tilde{S}, G$ verify $\forall x, \tilde{S}[G(x)] = G(S[x])$

- **Ref.:** $G, \tilde{S}$: pre-computation, G^{-1} : on-the-fly
  - For all i, $G[i] \leftarrow R_1 \cdot i \oplus R_0$
  - For all i, $\tilde{S}[G[i]] \leftarrow G[S[i]]$
- **Variation 1:** $G, \tilde{S}, G^{-1}$: pre-computation
- **Variation 2:** $\tilde{S}$: pre-computation, G, G^{-1}: on-the-fly

Comparison of AES implementations

Method	Reference	Cycles	RAM (bytes)	ROM (bytes)
Unprotected Implementation				
No Masking	Na.	2×10^3	32	1150
Provably Secure 1O-SCA Resistant Implementation				
1O Bool. Mask.	*[Mess00]*	9×10^3	$256 + 35$	1744
Aff. Mask. (ref.)	This paper	29×10^3	$512 + 37$	2857
Aff. Mask. (Var. 1)	This paper	28×10^3	$768 + 36$	2985
Aff. Mask. (Var. 2)	This paper	38×10^3	$256 + 37$	3252
Provably Secure 2O-SCA Resistant Implementation				
2O Bool. Mask.	*[SP06]*	594×10^3	$512 + 90$	2336
2O Bool. Mask.	*[RDP08]*	672×10^3	$256 + 86$	2215

Table: Comparison of several 8-bit 8051 AES implementations

Comparison of AES implementations

Method	Reference	Cycles	RAM (bytes)	ROM (bytes)
Unprotected Implementation				
No Masking	Na.	2×10^3	32	1150
Provably Secure 1O-SCA Resistant Implementation				
1O Bool. Mask.	*[Mess00]*	9×10^3	$256 + 35$	1744
Aff. Mask. (ref.)	This paper	29×10^3	$512 + 37$	2857
Aff. Mask. (Var. 1)	This paper	28×10^3	$768 + 36$	2985
Aff. Mask. (Var. 2)	This paper	38×10^3	$256 + 37$	3252
Provably Secure 2O-SCA Resistant Implementation				
2O Bool. Mask.	*[SP06]*	594×10^3	$512 + 90$	2336
2O Bool. Mask.	*[RDP08]*	672×10^3	$256 + 86$	2215

Table: Comparison of several 8-bit 8051 AES implementations

Outline

Leakage Model

- Each sensitive variable Z_i manipulated as

$$U_i = R_1 Z_i + R_0$$

- Intermediate variable U_i associated to a **leakage** L_i *s.t.*

$$L_i = \varphi(U_i) + B_i$$

with

  - φ: **deterministic leakage function**
  - B_i: **independent additive noise**

Definition (d-th order leakage (dO-leakage))

Tuple $(L_{i_1}, \ldots, L_{i_d})$ *s.t.* $(U_{i_1}, \ldots, U_{i_d})$ jointly depend on some sensitive variable

- No 1O-leakage because of the additive mask

2O-leakage of Affine Masking

Lemma

The pairs (U_1, U_2) *and* $(U_1', U_2') = (G(Z), R_0)$ *are identically distributed.*

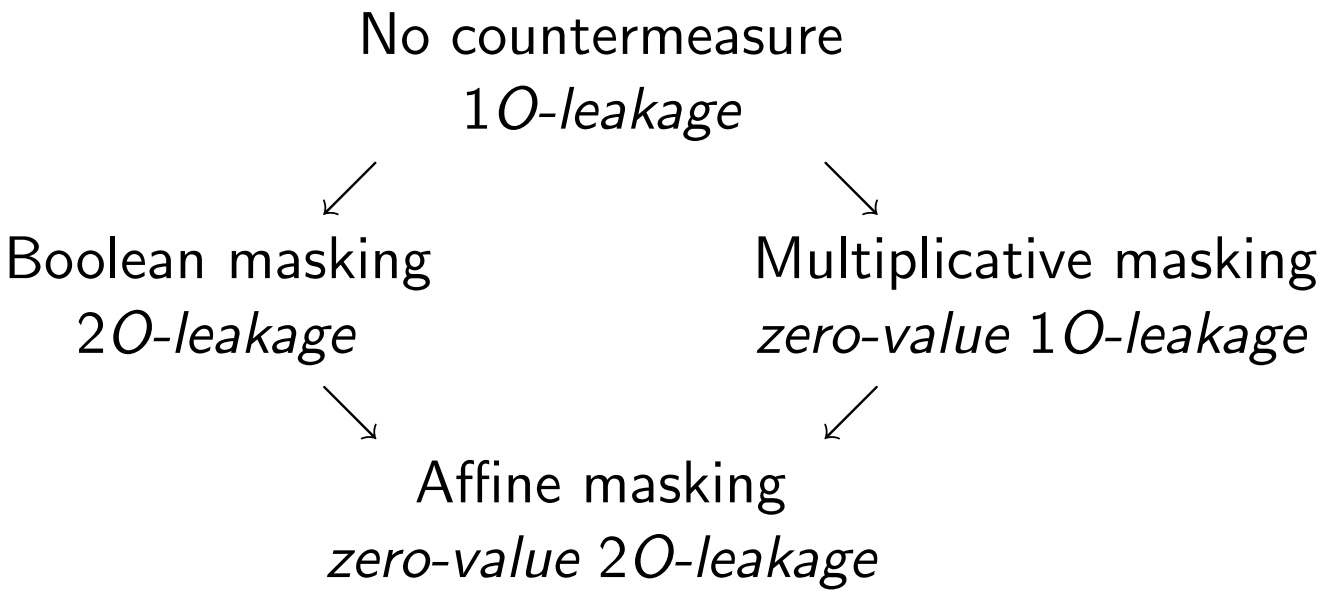
- **Proof:** just consider $Z = Z_1 + Z_2$
- Focus on 2O-leakages associated with $(U_1, U_2) = (G(Z), R_0)$ only

Lemma

The random pair $\big((L_1, L_2)|Z = z\big)$ *is identically distributed for every* $z \in \mathrm{GF}(2^n)^*$ *and the random pair* $\big((L_1, L_2)|Z = 0\big)$ *has a distinct distribution.*

- 2O-leakage only leaks information about whether $Z = 0$ or $Z \neq 0$

2O-leakage of Affine Masking



3O-leakage of Affine Masking

- 2O-leakage (L_1, L_2) associated to the pair $(U_1, U_2) = (R_1 Z + R_0, R_0)$ only leaks information about whether $Z = 0$ or $Z \neq 0$
- A natural extension is to consider the 3O-leakage (L_1, L_2, L_3) **with L_3 associated to $U_3 = R_1$**
- Because (U_1, U_2, U_3) **reveals the full value of Z** as

$$Z = U_3^{-1} \cdot (U_1 + U_2)$$

- However in practive (L_1, L_2, L_3) **does not reveal** (U_1, U_2, U_3)
- In practice, 3O-leakage provides less information than 2O-leakage
- Extracting information on Z is more difficult with affine masking than with Boolean masking

Information Theoretic Evaluation

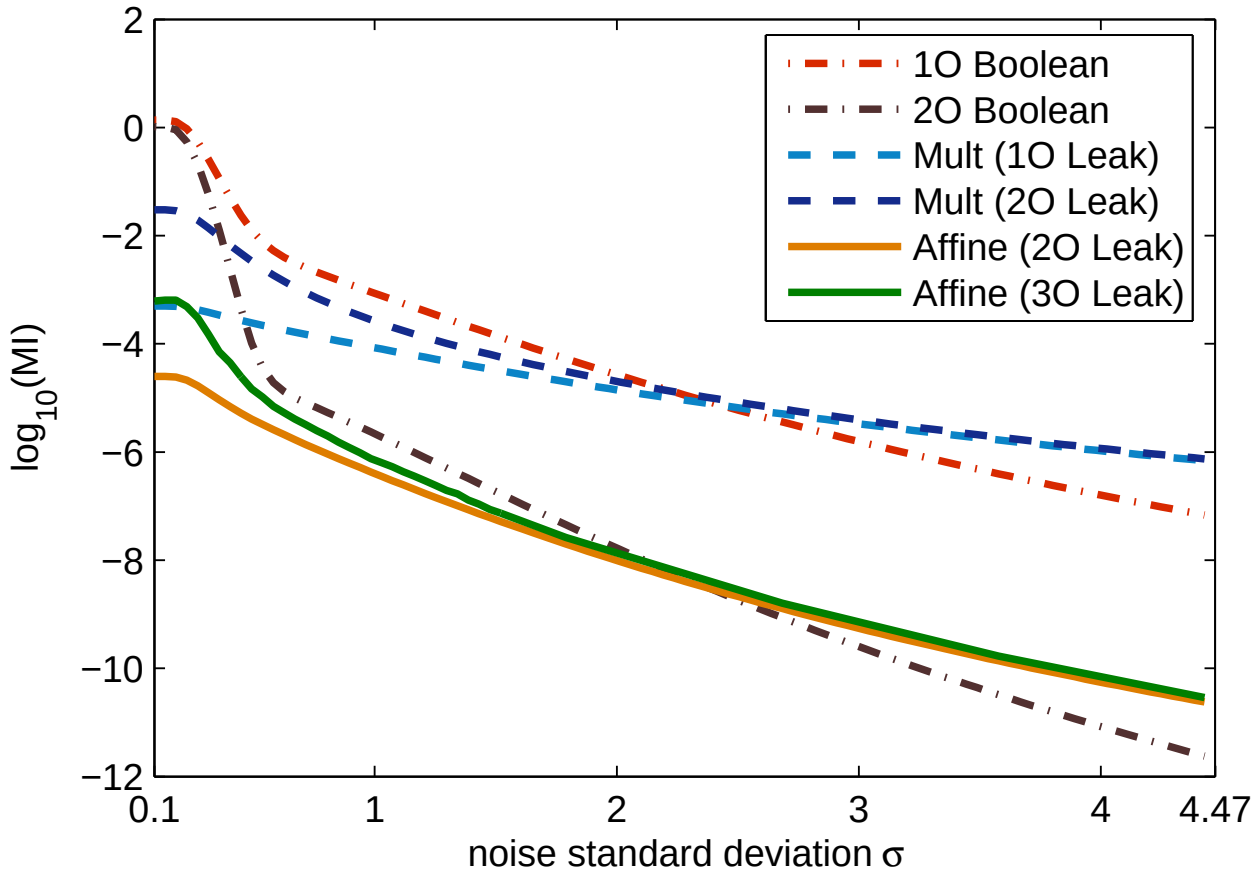


Figure: MI between leakage and sensitive variable *w.r.t.* noise deviation

Outline

Attack simulations - Unprofiled

Attack \ SNR	$+\infty$	1	1/2	1/5	1/10
Unprofiled Attacks against Boolean Masking					
2O-DPA on 1O Bool. Mask.	150	500	1500	6000	20 000
2O-MIA on 1O Bool. Mask.	100	5000	15 000	50 000	160 000
3O-DPA on 2O Bool. Mask.	1500	9000	35 000	280 000	$> 10^6$
3O-MIA on 2O Bool. Mask.	160	160 000	650 000	$> 10^6$	$> 10^6$
Unprofiled Attacks against Multiplicative Masking					
1O-DPA on Mult. Mask.	900	1500	2500	4000	7500
1O-MIA on Mult. Mask.	700	2500	3500	5500	15000
2O-DPA on Mult. Mask.	2500	7500	20 000	60 000	220 000
2O-MIA on Mult. Mask.	4000	35 000	55 000	100 000	200 000
Unprofiled Attacks against Affine Masking					
2O-DPA on Affine Mask.	6500	20 000	45 000	170 000	650 000
2O-MIA on Affine Mask.	5500	100 000	600 000	$> 10^6$	$> 10^6$
3O-DPA on Affine Mask.	$> 10^6$	$> 10^6$	$> 10^6$	$> 10^6$	$> 10^6$
3O-MIA on Affine Mask.	100 000	$> 10^6$	$> 10^6$	$> 10^6$	$> 10^6$

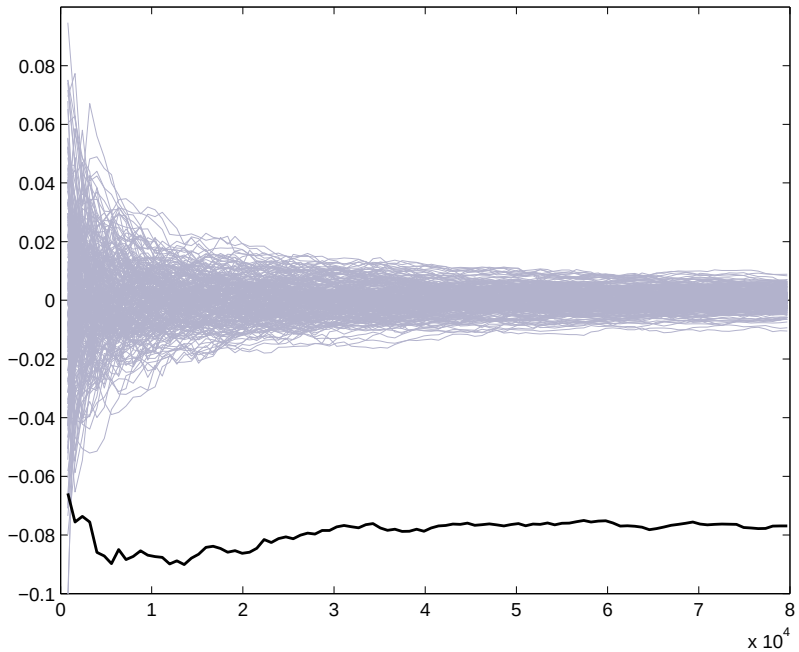
Table: Number of leakage measurements for a 90% success rate.

Attack simulations - Profiled

Attack \ SNR	$+\infty$	1	1/2	1/5	1/10
Profiled Attacks					
2O-TA on Bool. Mask.	20	500	1200	7000	20 000
3O-TA on 2O Bool. Mask.	20	8000	35 000	300 000	$> 10^6$
1O-TA on Mult. Mask.	500	1300	1900	4000	7000
2O-TA on Mult. Mask.	60	900	1400	4000	8000
2O-TA on Aff. Mask.	1300	15 000	45 000	200 000	$> 10^6$
3O-TA on Aff. Mask.	260	15 000	35 000	200 000	10^6

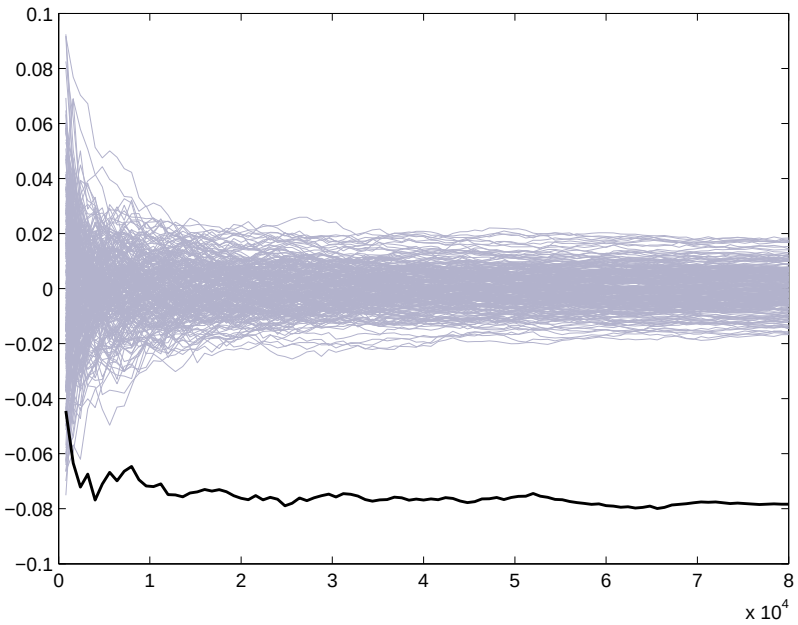
Table: Number of leakage measurements for a 90% success rate.

Practical Attacks - Multiplicative Masking

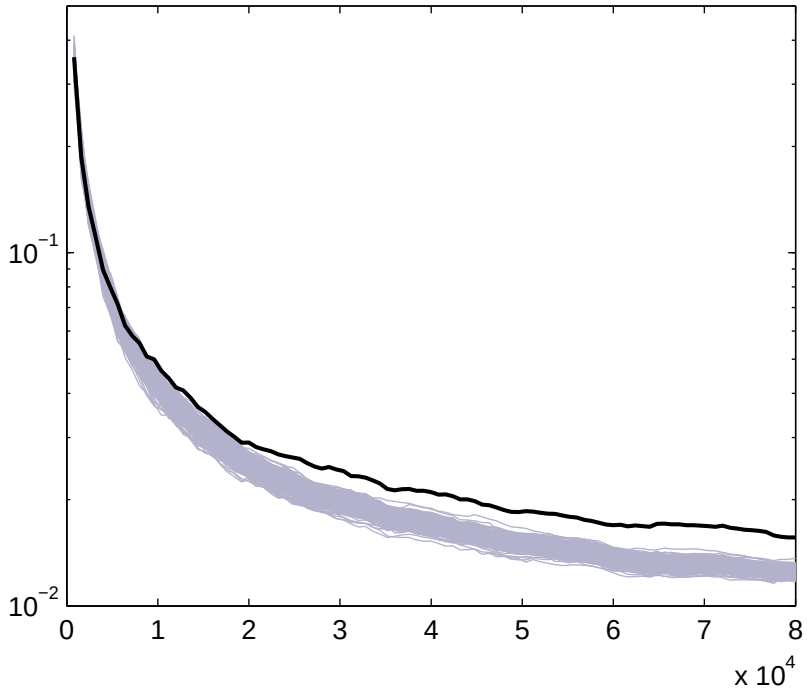


1O-DPA

Practical Attacks - Boolean Masking

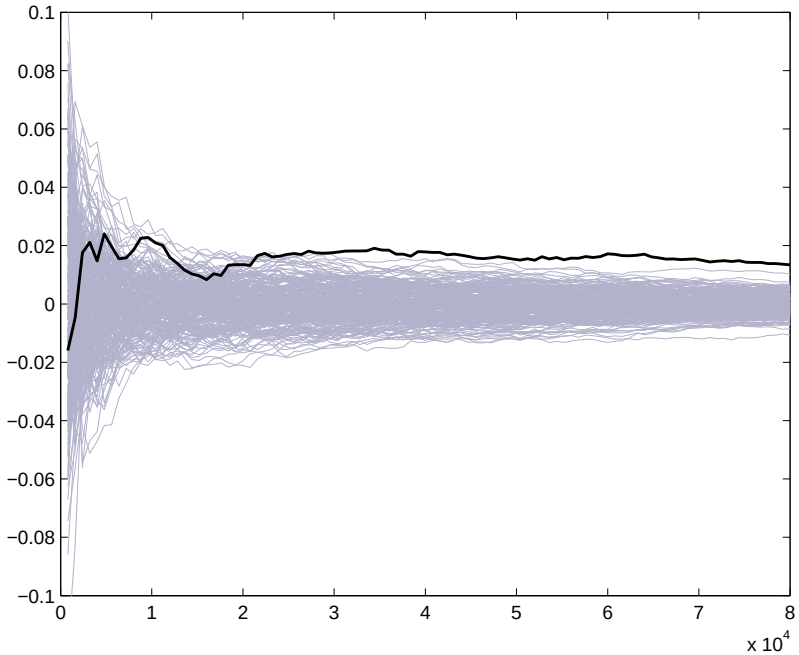


2O-DPA



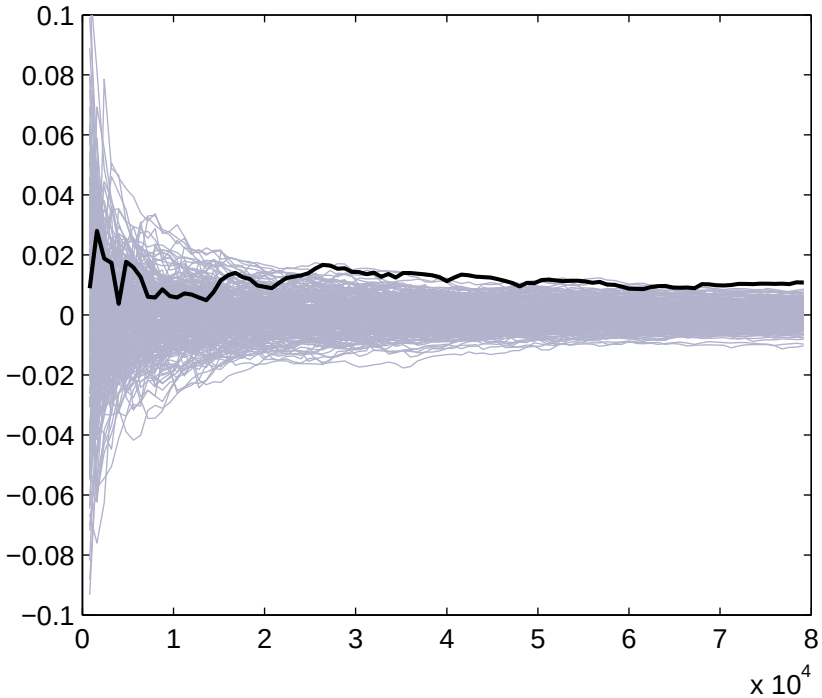
2O-MIA

Practical Attacks - Boolean Masking

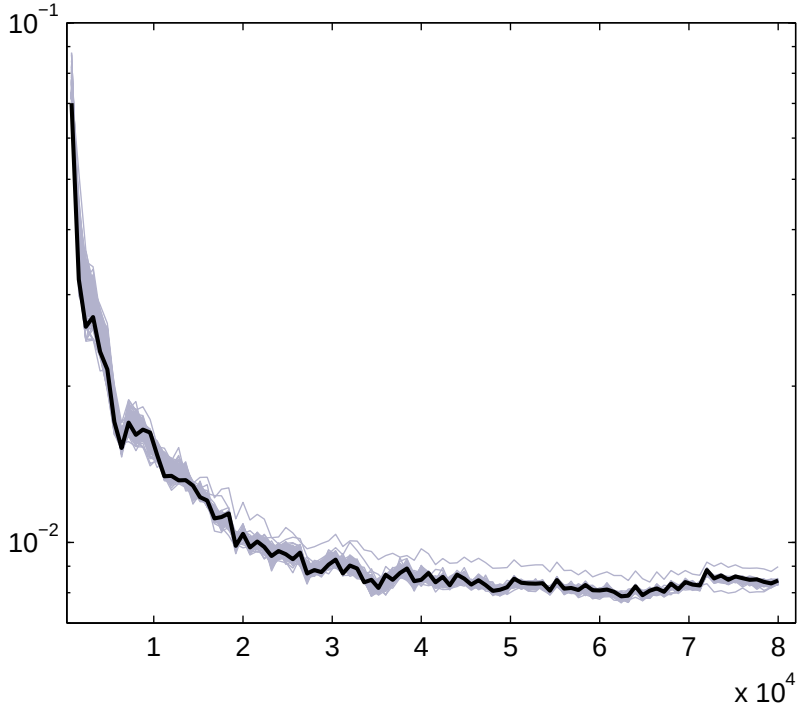


3O-DPA

Practical Attacks - Affine Masking



DPA



MIA

Outline

Conclusion

- Introduce affine masking over GF(2^n) as a possible alternative to the commonly used Boolean masking
- Provide an in-depth theoretical and practical analysis
- Prove that affine masking achieves a very good performance-security trade-off compared to existing countermeasures

Details can be found in the ePrint report 2010/523
(SAC 2010 extended version)

Thank you for your attention

Questions?