Optimal Codes for Inner Product Masking

Wei Cheng, Sylvain Guilley, Claude Carlet, Jean-Luc Danger and Alexander Schaub
wei.cheng@telecom-paristech.fr
Jun 24, 2019

Outlines

Why IPM?

Higher concrete security level (security order at bit-level)

Backgrounds

Masking is the most popular countermeasure to protect cryptographic implementations against side-channel analysis.

For Boolean masking, also named Perfect masking [CG18] with n shares in $\mathbb{K} = \mathbb{F}_{2^k}$ can be expressed in a coding format:

$$\mathrm{Z} = (Z_1, \ldots, Z_n) = \left(X + \sum_{i=2}^{n} M_i, M_2, M_3, \ldots, M_n \right) = X\mathbf{G} + M\mathbf{H}, \quad (1)$$

where $\mathbf{G}$ and $\mathbf{H}$ are generating matrix of C and D, respectively.

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \end{pmatrix} \in \mathbb{K}^{1\times n}$$

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathbb{K}^{(n-1)\times n}$$

Why IPM?

Inner Product Masking (IPM)

IPM was proposed by Balasch *et al.* [BFGV12, BFG15, BFG$^+$17], where random masks are involved by using Inner Product operation.

Let $\mathrm{X} \in \mathbb{F}_{2^k}$ denotes a field elements, L= $(L_1, L_2, \ldots, L_n)$ with $L_i \in \mathbb{F}_{2^k} \backslash \{0\}$ denotes a vector with n elements. The secret is $X = \langle L, Z \rangle = \sum_i^n L_i Z_i$.
Then IPM, also can be expressed in a coding format:

$$\mathrm{Z} = \left(X + \sum_{i=2}^{n} L_i M_i, M_2, M_3, \ldots, M_n \right) = X\mathbf{G} + M\mathbf{H} \tag{2}$$

where $\mathbf{G}$ and $\mathbf{H}$ are generating matrix of C and D, respectively, as follows.

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \end{pmatrix} \in \mathbb{K}^{1\times n}$$

$$\mathbf{H} = \begin{pmatrix} L_2 & 1 & 0 & \cdots & 0 \\ L_3 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_n & 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathbb{K}^{(n-1)\times n}$$

Defining parameters of codes

Definition 1 (Weight Enumerator Polynomial)

For a linear code D of parameters $[n, l, d_D]$,

$$W_D(X, Y) = \sum_{i=0}^{n} B_i X^{n-i} Y^i \tag{3}$$

where $B_i = |\{d \in D | w_H(d) = i\}|$ and w_H is the Hamming weight function.

Example 2

e.g., for linear code [8,4,4], we have $W_D(X, Y) = X^8 + 14X^4Y^4 + Y^8$, also denoted as: [<0,1>,<4,14>,<8,1>]. Thus, we have $B_0 = 1,\ B_4 = 14,\ B_8 = 1$.

Definition 3 (Dual Code)

The dual code of D, denoted as $D^\perp$, is: $D^\perp = \{x \mid \forall d \in D,\ \langle x, d\rangle = 0\}$.

Recall that $\mathrm{Z} = X\mathbf{G} + M\mathbf{H}$, where $\mathbf{G}$ and $\mathbf{H}$ are generating matrices of code C and D, respectively. Thus the generating matrix of dual code $D^\perp$ is

$$\mathbf{H}^\perp = (1, L_2, L_3, \ldots, L_n). \tag{4}$$

Why IPM?

Higher concrete security level (security order at bit-level)

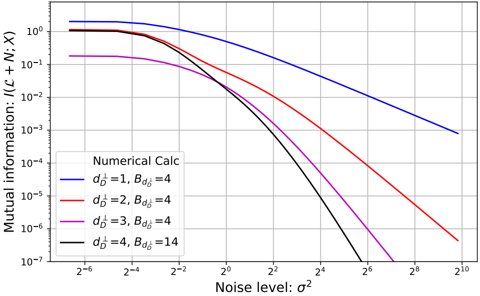


Figure 1: Mutual information I$[\mathcal{L}+N;X]$ between leakages ($\mathcal{L}=w_H(Z)$) and X in IPM.

From Fig. 1, obviously,

- Boolean masking's security level is lower than IPM (Note that if $L_2 = X^0$ $(=1)$, the IPM is degraded to Boolean masking)
- IPM's security depends on the choices of L_i
- The security level is related to $d_D^\perp$ as in [PGS$^+$17, BFG$^+$17, CG18]

Why IPM?

Higher concrete security level (security order at bit-level)

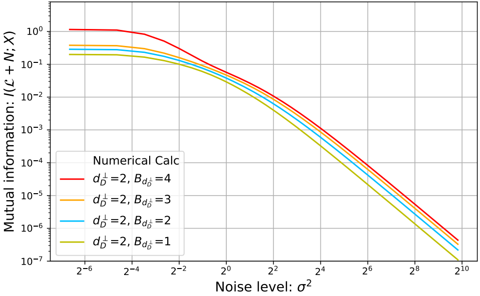


Figure 2: Mutual information I$[\mathcal{L}+N;X]$, using codes with the same $d_D^\perp$.

But, even with the same $d_D^\perp$, we can see that:

- IPM with different codes have different security level
- $d_D^\perp$ is not enough as a leakage metric
- **Question: how to concretely characterize the security of IPM?**

Related Works

The state-of-the-art

Security order

Two kinds of security order d_w and d_b under probing model are:

- **Word-level (k-bit) security order d_w**: leakages of word-level computation or data
- **Bit-level security order d_b**: in practice, each bit of sensitive variable can be investigated independently

In order to analyze the security order of IPM at bit-level, we introduce:

Sub-field representation

By using sub-field representation, we decompose $\mathbb{F}_{2^k}$ into $\mathbb{F}_2^k$ as

$$SubfieldRepresentation : (1, L_2, \ldots, L_n)_{2^k} \to (I_k, \mathbb{L}_2, \ldots, \mathbb{L}_n)_2 \quad (5)$$

So by *sub-field representation*, a $(1 \times n)$ vector $(1, L_2, \ldots, L_n)$ at word-level is converted to $(k \times nk)$ matrix $(I_k, \mathbb{L}_2, \ldots, \mathbb{L}_n)$ at bit-level.

The state-of-the-art

Table 1: Summaries of security analysis on IPM and DSM.

	Security order	Code parameters	Metrics	Comments
Balasch *et al.* [BFG15]	d_w	–	MI	MI varies for different L vector
Wang *et al.* [WSY+16]	d_b	$d_D^\perp$	MI	O_{min} $(= d_D^\perp)$ was used (the lowest key-dependent statistical moment)
Poussier *et al.* [PGS+17]	$d_w,\ d_b$	$d_D^\perp$	MI	
Balasch *et al.* [BFG+17]	d_w, d_b	–	MI	d_{bound} $(\approx d_b)$ is in bound moment model
Claude *et al.* [CG18]	$d_w,\ d_b$	$d_D^\perp$	MI, SR	SR of *optimal attack* [BGHR14]
This work	$d_w,\ d_b$	$d_D^\perp,\ \boldsymbol{B_{d_D^\perp}}$	MI, SR, **SNR**	An unified framework to analyze all IPM codes by closed-form expression

- Here $d_w,\ d_b$ are word- and bit-level security order, respectively, and $d_w = n - 1$.
- Bit-level security order d_b equals to $d_D^\perp - 1$ in [PGS+17], [CG18] and in this work.

Outlines

Concrete security level of IPM

SNR as a metric

SNR is a commonly used in side-channel analysis as a leakage metric. Let

$$\mathcal{L} = P(Z) + N$$

denotes the leakages where N denotes the independent noise, we have

$$Var(E(P(Z) + N|X)) = Var(E(P(Z)|X))$$

and then define *SNR* as:

$$SNR = \frac{Var(E(\mathcal{L}|X))}{Var(N)} = \frac{Var(E(P(Z)|X))}{\sigma^2}. \tag{6}$$

Let $\widehat{P}(z)$ be the Fourier transform of $P(z)$ defined as:

Definition 4 (Fourier Transformation)

The Fourier transformation of a pseudo-Boolean function $P : \mathbb{F}_2^n \to \mathbb{R}$ is denoted by $\widehat{P} : \mathbb{F}_2^n \to \mathbb{R}$, and defined as: $\widehat{P}(z) = \sum_y P(y)(-1)^{y \cdot z}$.

Concrete security level of IPM

Therefore, we have following theorem:

Theorem 5 (*SNR* of IPM)

For IPM scheme with $Z = X\mathbf{G} + M\mathbf{H}$, the SNR between secret X and leakages is

$$SNR = \frac{2^{-2n}}{\sigma^2} \sum_{x \in D^\perp \setminus \{0\}} \left(\widehat{P}(x)\right)^2. \tag{7}$$

Theorem 6 (Security order of IPM)

If $d^\circ P < d_D^\perp$, the attack fails with SNR$= 0$, thus the security order of IPM scheme in bounded moment model is $d = d_D^\perp - 1$.

Therefore the security order is the minimum value of $d^\circ P$ such that $SNR \neq 0$, where SNR is quantitative metric to quantify the leakages.

Concrete security level of IPM

Hamming weight leakage model

We use $P(z) = w_H(z)^d$ as higher order leakage model. Clearly, the degree of P is $d^\circ P = d$. Thus we have following theorem for *SNR*.

Theorem 7 (*SNR* of IPM)

For SNR of the Hamming weight leakages with respect to secret variable X which protected by IPM, we have

$$SNR = \begin{cases} 0 & \text{if } d^\circ P < d_D^\perp \\ \frac{1}{\sigma^2} B_{d_D^\perp} \left(\frac{d_D^\perp !}{2^{d_D^\perp}} \right)^2 & \text{if } d^\circ P = d_D^\perp \end{cases} \tag{8}$$

Surprisingly, the *SNR* of IPM is quantitatively connected to $d_D^\perp$ and $B_{d_D^\perp}$, which is determined by selecting $L = (L_1, L_2, \ldots, L_n)$.

Mutual information as a metric

In the presence of noise $N \sim \mathcal{N}(0, \sigma^2)$, the mutual information between the noisy leakage $\mathcal{L} + N$ and X can be developed using a Taylor's expansion [CDG+14]:

$$\mathsf{I}[\mathcal{L}+N; X] \approx \frac{1}{\ln 2} \sum_{d=0}^{+\infty} \frac{1}{2\,d!} \sum_{x \in \mathbb{F}_2^k} \mathbb{P}(X=x) \frac{\left(k_d(\mathcal{L} \mid X=x) - k_d(\mathcal{L})\right)^2}{\left(Var(\mathcal{L}) + \sigma^2\right)^d}$$

$$= \frac{1}{\ln 2} \sum_{d=0}^{+\infty} \frac{1}{2\,d!} \frac{Var(k_d(\mathcal{L} \mid X))}{\left(Var(\mathcal{L}) + \sigma^2\right)^d} \;, \qquad (9)$$

where k_d are order d cumulants [Car03].

The term $Var(E(k_d(\mathcal{L} \mid X)))$ is null for $d < d_D^{\perp}$, and equals $Var(\mu_d(\mathcal{L} \mid X)) = Var(E(\mathcal{L}^{d_D^{\perp}} \mid X))$ for $d = d_D^{\perp}$. Thus, under Hamming weight leakage model, the mutual information can be developed at first order in $1/\sigma^{2d_D^{\perp}}$ as:

$$\mathsf{I}[\mathcal{L}+N; X] = \frac{d_D^{\perp}! B_{d_D^{\perp}}}{2 \ln 2 \cdot 2^{2d_D^{\perp}}} \times \frac{1}{\sigma^{2d_D^{\perp}}} + \mathcal{O}\left(\frac{1}{\sigma^{2(d_D^{\perp}+1)}}\right) \quad \text{when } \sigma \longrightarrow +\infty \qquad (10)$$

Mutual information as a metric

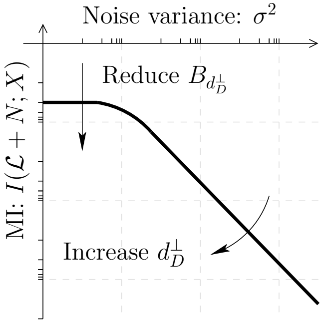


Figure 3: Two concomitant objectives to reduce the mutual information.

From Fig. 3, we can see that:

- the slope in the log-log representation of the *MI* versus the noise standard deviation is all the more steep as $d_D^\perp$ is high, and
- the vertical offset is adjusted by $B_{d_D^\perp}$; the smaller it is the smaller the *MI*.

Choosing optimal codes for IPM

Using $d_D^\perp$ and $B_{d_D^\perp}$ as a unified evaluation framework

A unified evaluation framework for IPM

For IPM with $Z = \left(X + \sum_{i=2}^{n} L_i M_i, M_2, M_3, \ldots, M_n\right) = X\mathbf{G} + M\mathbf{H}$, its concrete security level can be characterized by two defining parameters $d_D^\perp$ and $B_{d_D^\perp}$, where code D is generated by $\mathbf{H}$.

Example 8

For $n = 2$ with $L_2 \in \mathbb{F}_{2^4}$, by subfield representation:

- $d_D^\perp = 2$ for $L_2 \in \{X^i\}$ for $i \in \{0, 1, 2, 3, 12, 13, 14\}$
- $d_D^\perp = 3$ for $L_2 \in \{X^i\}$ for $i \in \{4, 5, 6, 7, 8, 9, 10, 11\}$

In particular, for $d_D^\perp = 2$, we have:

- IPM with $L_2 = X^0$: [<0,1>,<2,4>,<4,6>,<6,4>,<8,1>]
- IPM with $L_2 = X^1,\ X^{14}$: [<0,1>,<2,3>,<3,2>,<4,3>,<5,4>,<6,1>,<7,2>]
- IPM with $L_2 = X^2,\ X^{13}$: [<0,1>,<2,2>,<3,3>,<4,3>,<5,4>,<6,2>,<7,1>]
- IPM with $L_2 = X^3,\ X^{12}$: [<0,1>,<2,1>,<3,4>,<4,3>,<5,4>,<6,3>]

Choosing optimal codes for IPM

Using $B_{d_D^\perp}$ and $d_D^\perp$ as a unified evaluation framework

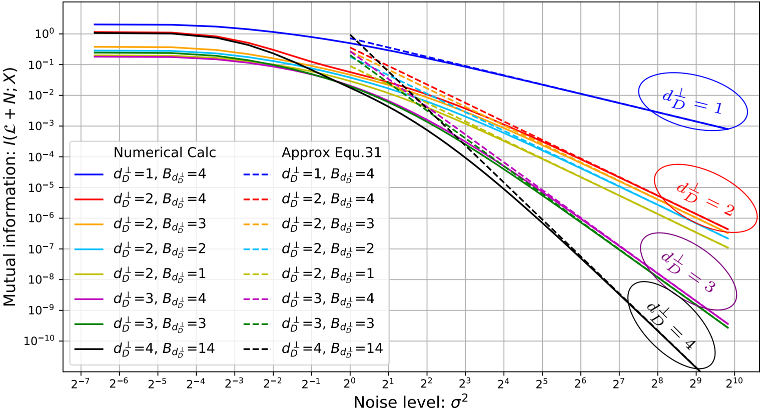


Figure 4: Numerical calculation and approximation of mutual information $\mathrm{I}[\mathcal{L}+N;X]$ between leakages and X in IPM.

Choosing optimal codes for IPM

Using $B_{d_D^\perp}$ and $d_D^\perp$ as a unified evaluation framework

By this unified evaluation framework, it is easy to select optimal codes for IPM, which with the highest side-channel resistance.

Algorithm 1: Optimal Code Selection

Result: Optimized $d_D^\perp$ and $B_{d_D^\perp}$

```
Maximize d_D^⊥;
if mean{B_i < n/2} then
    goto 1;
else
    Minimize B_{d_D^⊥};
return d_D^⊥ and B_{d_D^⊥}
```

Outlines

Success rate as an attack metric

Practical security evaluation

Optimal Attack [BGHR14]

For each attack, the targeted variable is:

$$\mathbf{z} = (w_H(t_q + k + m_2L_2 + \cdots + m_nL_n), w_H(m_2), w_H(m_3), \ldots, w_H(m_n))$$

for n-dimensional attack (e.g., *Attack_2D*), and

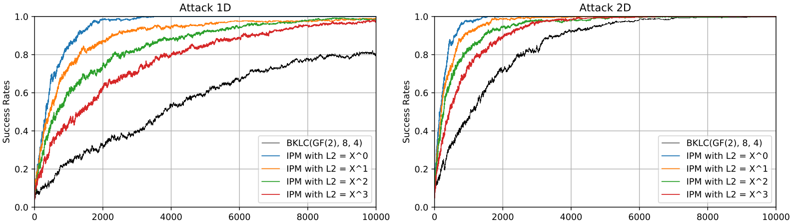
$$\begin{aligned} z &= w_H(t_q + k + m_2L_2 + \cdots + m_nL_n) + w_H(m_2) + w_H(m_3) + \cdots + w_H(m_n) \\ &= z_1 + z_2 + \cdots + z_n \end{aligned}$$

for 1-dimensional attack (e.g., *Attack_1D*).

The success rate is the metric for evaluating attacks on different codes (refer to Appendix for attacks).

What about the codes with the same $d_D^\perp$?

- Seting-up: $n = 2$, $k = 4$, $L_2 \in \{X^0, \ldots, X^3\}$, $T = 10,000$, $\sigma = 1.50$



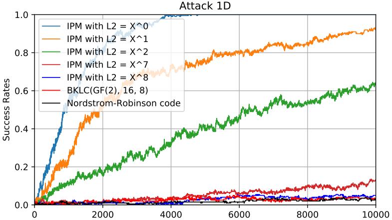
For $d_b = 2$, we have

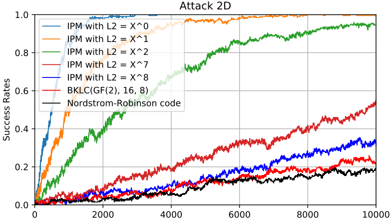
- IPM with $L_2 = X^0$: [<0,1>,<2,4>,<4,6>,<6,4>,<8,1>]
- IPM with $L_2 = X^1$: [<0,1>,<2,3>,<3,2>,<4,3>,<5,4>,<6,1>,<7,2>]
- IPM with $L_2 = X^2$: [<0,1>,<2,2>,<3,3>,<4,3>,<5,4>,<6,2>,<7,1>]
- **IPM with $L_2 = X^3$: [<0,1>,<2,1>,<3,4>,<4,3>,<5,4>,<6,3>]**
- *BKLC(GF(2), 8, 4)*: [<0,1>,<4,14>,<8,1>] $\longrightarrow$Not IPM codes

What about the codes with the same $d_D^\perp$?

Codes with the same $d_D^\perp$ while different $B_{d_D^\perp}$

- Seting-up: $n = 2$, k=8, $L_2 \in \{X^0, \ldots, X^7\}$, $T = 10,000$, $\sigma = 1.50$





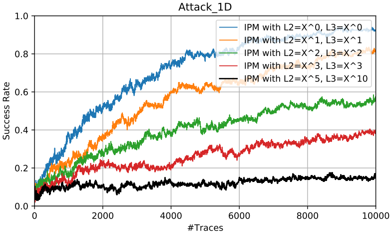
Concerning $d_D^\perp$, we have

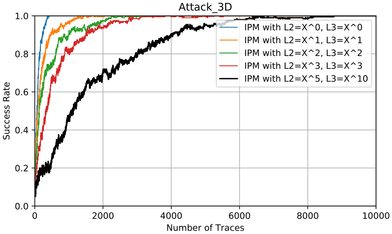
- IPM with $L_2 = X^0$: [<0,1>,<2,8>,<4,28>,<6,56>,<8,70>,...,<16,1>]
- IPM with $L_2 = X^1$: [<0,1>,<2,7>,<4,21>,<5,8>,<6,35>,...,<14,1>]
- **IPM with $L_2 = X^7$: [<0,1>,<2,1>,<4,1>,<5,23>,<6,36>,...,<14,2>]**
- IPM with $L_2 = X^8$: [<0,1>,<4,3>,<5,25>,<6,34>,<7,36>,...,<14,2>]
- *BKLC(GF(2), 16, 8)*: [<0,1>,<5,24>,<6,44>,<7,40>,<8,45>,...,<12,10>]
- *Nordstrom-Robinson code*: (16, 256, 6)

What about the codes with the same $d_D^\perp$?

Codes with the same $d_D^\perp$ while different $B_{d_D^\perp}$

- Seting-up: n=3, k=4, $L_2, L_3 \in \{X^0, \ldots, X^3\}$, $T = 10,000$, $\sigma = 1.50$





Concerning $d_D^\perp$, we have

- IPM with $L_2 = X^0, L_3 = X^0$: [<0,1>,<3,4>,<6,6>,<9,4>,<12,1]
- IPM with $L_2 = X^1, L_3 = X^1$: [<0,1>,<3,3>,<4,1>,<5,1>,...,<11,1>]
- IPM with $L_2 = X^2, L_3 = X^2$: [<0,1>,<3,2>,<4,1>,<5,3>,...,<11,1>]
- **IPM with $L_2 = X^3, L_3 = X^3$: [<0,1>,<3,1>,<4,1>,<5,4>,...,<10,1>]**
- IPM with $L_2 = X^5, L_3 = X^{10}$: [<0,1>,<6,12>,<8,3>]≡*BKLC*(GF(2),12,4)

Summary of Results

Table 2: Optimizing IPM in several scenarios

#Shares	$\mathbb{F}_{2^k}$	Word-level (IPM)	Bit-level (BKLC)	Δ	Comments
$n=2$	$k=4$	$max\{d_D^\perp\}=3$ $mean\{B_i\}=4$ $min\{B_d\}=4$	[8, 4, 4]: $d_D^\perp=4$ $mean\{B_i\}=4$ $B_d=14$	-1	[WSY+16, CG18]
	$k=8$	$max\{d_D^\perp\}=4$ $mean\{B_i\}=8$ $min\{B_d\}=3$	[16, 8, 5]: $d_D^\perp=5$ $mean\{B_i\}=4$ $B_d=24$	-1	[PGS+17], Try one NR non-linear code (16, 256, 6)
$n=3$	$k=4$	$max\{d_D^\perp\}=6$ $mean\{B_i\}=6$ $min\{B_d\}=12$	[12, 4, 6]: $d_D^\perp=6$ $mean\{B_i\}=6$ $B_d=12$	0	**New**, the best IPM code is equivalent to BKLC code
	$k=8$	$max\{d_D^\perp\}=8$ $mean\{B_i\}=12$ $min\{B_d\}=7$	[24, 8, 8]: $d_D^\perp=8$ $mean\{B_i\}=10$ $B_d=130$	0	[PGS+17], but the BKLC code can't be used

Outlines

Conclusions

With the concepts from coding theory, we propose a unified framework to analyze and optimize the concrete security level of IPM scheme.

- Two leakage metric *SNR* and *MI* to quantitatively characterize the the SCA resistance of IPM
- By adding $B_{d^{\perp}_{D}}$, we propose a unified framework to systemically evaluate all codes for IPM
- By using attack metric *SR*, we validate the effective of our unified framework
- Propose a simple method to choose optimal codes for IPM, also with examples:
  - with n=2 shares: 4-bit and 8-bit variables
  - with n=3 shares: 4-bit and 8-bit variables
- IPM is not optimal compared to *BKLC* codes, especially for $n = 2$ with $k = 4$ and $k = 8$ bits

Question?

References I

Josep Balasch, Sebastian Faust, and Benedikt Gierlichs.

Inner product masking revisited.

In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I*, volume 9056 of *Lecture Notes in Computer Science*, pages 486–510. Springer, 2015.

Josep Balasch, Sebastian Faust, Benedikt Gierlichs, Clara Paglialonga, and François-Xavier Standaert.

Consolidating Inner Product Masking.

In Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology - ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I*, volume 10624 of *Lecture Notes in Computer Science*, pages 724–754. Springer, 2017.

Josep Balasch, Sebastian Faust, Benedikt Gierlichs, and Ingrid Verbauwhede.

Theory and Practice of a Leakage Resilient Masking Scheme.

In Xiaoyun Wang and Kazue Sako, editors, *Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings*, volume 7658 of *Lecture Notes in Computer Science*, pages 758–775. Springer, 2012.

References II

Nicolas Bruneau, Sylvain Guilley, Annelie Heuser, and Olivier Rioul.

Masks Will Fall Off – Higher-Order Optimal Distinguishers.

In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology – ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014, Proceedings, Part II*, volume 8874 of *Lecture Notes in Computer Science*, pages 344–365. Springer, 2014.

Jean-François Cardoso.

Dependence, Correlation and Gaussianity in Independent Component Analysis.

Journal of Machine Learning Research, 4:1177–1203, 2003.

Claude Carlet, Jean-Luc Danger, Sylvain Guilley, Houssem Maghrebi, and Emmanuel Prouff.

Achieving side-channel high-order correlation immunity with leakage squeezing.

J. Cryptographic Engineering, 4(2):107–121, 2014.

Claude Carlet and Sylvain Guilley.

Statistical properties of side-channel and fault injection attacks using coding theory.

Cryptography and Communications, 10(5):909–933, 2018.

Romain Poussier, Qian Guo, François-Xavier Standaert, Claude Carlet, and Sylvain Guilley.

Connecting and Improving Direct Sum Masking and Inner Product Masking.

In Thomas Eisenbarth and Yannick Teglia, editors, *Smart Card Research and Advanced Applications - 16th International Conference, CARDIS 2017, Lugano, Switzerland, November 13-15, 2017, Revised Selected Papers*, volume 10728 of *Lecture Notes in Computer Science*, pages 123–141. Springer, 2017.

References III

Weijia Wang, François-Xavier Standaert, Yu Yu, Sihang Pu, Junrong Liu, Zheng Guo, and Dawu Gu.
Inner Product Masking for Bitslice Ciphers and Security Order Amplification for Linear Leakages.
In Kerstin Lemke-Rust and Michael Tunstall, editors, *Smart Card Research and Advanced Applications - 15th International Conference, CARDIS 2016, Cannes, France, November 7-9, 2016, Revised Selected Papers*, volume 10146 of *Lecture Notes in Computer Science*, pages 174–191. Springer, 2016.

Appendix I. IPM codes with $n = 2$ in $\mathbb{F}_{2^4}$

Table 3: IPM for $n = 2$ and $k = 4$

L_2	Weight Enumeration Polynomial	$I(x, k)$
X^0	[<0, 1>, <2, 4>, <4, 6>, <6, 4>, <8, 1>]	1.151963
X^1	[<0, 1>, <2, 3>, <3, 2>, <4, 3>, <5, 4>, <6, 1>, <7, 2>]	0.380288
X^2	[<0, 1>, <2, 2>, <3, 3>, <4, 3>, <5, 4>, <6, 2>, <7, 1>]	0.287149
X^3	[<0, 1>, <2, 1>, <3, 4>, <4, 3>, <5, 4>, <6, 3>]	0.199569
X^4	[<0, 1>, <3, 4>, <4, 5>, <5, 4>, <6, 2>]	0.181675
X^5	[<0, 1>, <3, 3>, <4, 7>, <5, 4>, <7, 1>]	0.246318
X^6	[<0, 1>, <3, 4>, <4, 5>, <5, 4>, <6, 2>]	0.181675
X^7	[<0, 1>, <3, 4>, <4, 5>, <5, 4>, <6, 2>]	0.181675
X^8	[<0, 1>, <3, 4>, <4, 5>, <5, 4>, <6, 2>]	0.181675
X^9	[<0, 1>, <3, 4>, <4, 5>, <5, 4>, <6, 2>]	0.181675
X^{10}	[<0, 1>, <3, 3>, <4, 7>, <5, 4>, <7, 1>]	0.246318
X^{11}	[<0, 1>, <3, 4>, <4, 5>, <5, 4>, <6, 2>]	0.181675
X^{12}	[<0, 1>, <2, 1>, <3, 4>, <4, 3>, <5, 4>, <6, 3>]	0.199569
X^{13}	[<0, 1>, <2, 2>, <3, 3>, <4, 3>, <5, 4>, <6, 2>, <7, 1>]	0.287149
X^{14}	[<0, 1>, <2, 3>, <3, 2>, <4, 3>, <5, 4>, <6, 1>, <7, 2>]	0.380288

Appendix II. Two optimal attacks

For two attacks *Attack_1D* and *Attack_2D*, we refer to *Optimal Attack* [BGHR14] as:

- The monovariate attack measures the sum of leakages for each trace q $(1 \leq q \leq Q)$, hence the optimal attack guesses the correct key k^* as:

$$\hat{k^*} = \arg\max_{k \in \mathbb{F}_2^k} \sum_{q=1}^{Q} \log \sum_{m_2 \in \mathbb{F}_2^k} exp\backslash - \frac{1}{4\sigma^2}\left\{\left(l_q^{(1)} + l_q^{(2)} - w_H(t_q \oplus k \oplus F[l_2][m_2], m_2)\right)^2\right\} \tag{11}$$

- The bivariate attack measures each of two shares l_q^1 and l_q^2 independently, the optimal attack guesses the correct key k^* as:

$$\hat{k^*} = \arg\max_{k \in \mathbb{F}_2^k} \sum_{q=1}^{Q} \log \sum_{m_2 \in \mathbb{F}_2^k} exp\backslash - \frac{1}{2\sigma^2}\left\{\left(l_q^{(1)} - w_H(t_q \oplus k \oplus F[l_2][m_2])\right)^2 + \left(l_q^{(2)} - w_H(m_2)\right)^2\right\} \tag{12}$$