Optimality and Efficiency of Higher-Order Rank-Based Side-Channel Attacks against Inner Product Masking Schemes

Haitao Zhan, Wei Cheng✉, *Member, IEEE*, Yuejun Liu, Jingdian Ming, Sylvain Guilley, *Senior Member, IEEE*, Yongbin Zhou

***Abstract*—Higher-Order Spearman Coefficient Analysis (HO-SPA) is of particular interest in non-profiling Side-Channel Attacks (SCA) due to its effectiveness against monotonic nonlinear leakage. However, the optimal form, effective order combinations, and performance of HO-SPA (or its variants) remain underexplored. This work systematically revisits HO-SPA in the presence of Boolean Masking (BM) and Inner Product Masking (IPM). Specifically, we propose a new variant of HO-SPA, called HO-ASPA (Higher-Order Adjusted Spearman Coefficient Analysis), and derive the optimal form of HO-ASPA for two-share BM and IPM across a wide range of order combinations under the Hamming weight (HW) leakage model. We first derive the optimal form of HO-ASPA by extending the Spearman Coefficient Analysis to higher-order moments, while preserving the centered-product combination of the share leakages. We then establish an upper bound on the maximal correlation induced by the rank transformation and characterize the relationship between the signal-to-noise ratio and the optimal correlation coefficient in the rank-domain, and extend these analyses to IPM. To enable a fair comparison, we also derive the optimal forms of Higher-Order Correlation Power Analysis (HO-CPA) and HO-ASPA for both BM and IPM under the Least-Significant-Bit (LSB) leakage model. Our analysis highlights a key distinction: unlike HO-CPA, where higher orders often outperform the lowest feasible order, HO-ASPA shows no significant performance gap between higher-order and lower-order attacks. Moreover, under the HW leakage model, our HO-ASPA not only aligns with existing results and succeeds in counteracting the security order amplification, but also significantly outperforms Kruskal-Wallis Analysis (KWA)-based attacks at various order combinations. We further validate the effectiveness of the proposed HO-ASPA through simulated and real-world experimental evaluations.**

***Index Terms*—Side-Channel Attacks, Non-profiling Distinguishers, Higher-Order Correlation Analysis, Inner Product Masking.**

I. Introduction

SIDE-CHANNEL Attacks (SCA) exploit information leakage (e.g., power consumption [1], electromagnetic radiation [2], etc.) generated by the physical implementation of cryptographic algorithms to infer secret data being processed. In essence, the leakage is correlated with the internal operations of the device, thereby posing a serious threat to the practical security of cryptographic systems [3]. Since Kocher's pioneering work [1], [4], SCA has attracted considerable attention from both academia and industry due to its strong practical effectiveness.

Depending on the attacker's capabilities, SCA can be broadly classified into profiling attacks, which require a controllable device replica for leakage characterization, and non-profiling attacks, such as classical DPA [1], in which the attacker observes the device's behavior without a profiling phase. The latter model covers most real-world embedded-security scenarios (e.g., smart cards, IoT devices) and is the focus of practical security evaluation, as well as the scope of this work.

A. Masking as an Effective Countermeasure

To defend against SCA, masking has become a widely adopted countermeasure due to its provable security [5], [6]. The core idea is to randomize the dependence between a sensitive intermediate variable X and its leakage by splitting X into n shares [7], thereby breaking the direct link between the sensitive variable and its physical leakage. Taking Boolean Masking (BM) as an example [3], a sensitive variable X is randomly split into n shares $S_1, \ldots, S_n$ such that $X = S_1 \oplus S_2 \oplus \cdots \oplus S_n$. Typically, the first $n-1$ shares (called masks) are chosen randomly, while the last share (the masked variable) is computed to ensure that their XOR equals X. This is referred to as $(n-1)$th-order masking, where $(n-1)$ denotes the security order. Under the corresponding security model, an adversary cannot learn information about X by accessing at most $n-1$ shares simultaneously. When $(n-1)$th-order masking is used, an adversary may attempt a higher-order SCA by combining the leakage signals produced during the processing of the n shares $S_1, \ldots, S_n$, thereby recovering information about X.

To enhance the side-channel resilience, Inner Product Masking (IPM) [8], [9] has been proposed as a promising scheme that provides higher-order security while keeping the number of shares n small. In particular, BM is a special case of IPM. In fact, studies have demonstrated that IPM possesses the Security Order Amplification (SOA) property [9]–[12]: by selecting appropriate IPM parameters, an IPM implementation with n shares can resist attacks of orders much greater than n. In this work, we say that an attack is "breaking SOA" if it achieves reliable key recovery at an effective attack order lower than the minimum order predicted by the SOA security

This work is supported in part by the National Natural Science Foundation of China (No.U2336205, No.62502207, No.62302226, No.62302224), Yunnan Provincial Major Science and Technology Special Plan Projects (202502AD080008) and Yunnan Provincial New R&D Institution Cultivation Project (202404BQ040148). This work has also received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No.101092766 (Allegro Project).

H. Zhan, W. Cheng, Y. Liu, J. Ming, and Y. Zhou are with the School of Cyber Science and Engineering, Nanjing University of Science and Technology, Nanjing, China (email: {zhanhaitao, wei.cheng, liuyuejun, mingjingdian, zhouyongbin}@njust.edu.cn).

S. Guilley is with Secure-IC S.A.S., 75014 Paris, France (e-mail: sylvain.guilley@secure-ic.com). W. Cheng and S. Guilley are also with LTCI, Télécom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France.

analysis, under the specific leakage and attacker models considered. This wording does not imply a universal break of IPM under all possible leakage configurations. Additionally, various masking techniques such as secret sharing schemes [13] and direct sum masking [14] have been proposed for side-channel protection. More generally, code-based masking [15] provides a unified framework for these masking schemes.

B. Non-Profiling Attacks against Masked Implementations

To evaluate the effectiveness of masking schemes, various higher-order SCA methods have been proposed. Among them, non-profiling attacks have gained widespread attention due to their practicality. Correlation Power Analysis (CPA) [16] is a widely adopted non-profiling method whose effectiveness has been extensively validated on various datasets. In [17], second-order CPA was introduced to assess the security of BM. It was shown that, under the Hamming weight (HW) leakage model and with (independent) Gaussian noise, combining the two leakages via the so-called centered-product yields the optimal attack efficiency. Subsequently, the approach was extended to higher-order attacks against IPM in [18]. More recently, [19] derived the optimal form of Higher-Order CPA (HO-CPA) under the HW leakage model for both BM and IPM. Notably, the findings in [18], [19] indicate that HO-CPA is effective only when the attack order exceeds the security order $n-1$.

However, CPA can only capture linear relationships. Consequently, several studies have investigated distinguishers that exploit nonlinear dependencies [12], [20], [21]. In the context of masking, Yan et al. [21] proposed a non-profiling distinguisher based on Kruskal–Wallis analysis (KWA), which can exploit nonlinear information. Their experiments demonstrated that KWA outperformed Mutual Information Analysis [22] and Kolmogorov–Smirnov test-based distinguishers [23] against BM. Subsequently, Wu et al. [12] extended KWA to IPM and established its equivalence with the Adjusted Second-Order Spearman Coefficient Analysis (A-2O-SPA). Their experiments in [12] showed that rank-based distinguishers such as KWA and A-2O-SPA can break the SOA of IPM. Specifically, for IPM with public parameters 14 and 23, which have security orders of 2 and 3, respectively, the results reported in [18], [19] indicate that HO-CPA requires third- or fourth-order attacks to succeed. In contrast, rank-based distinguishers, such as KWA and Spearman Coefficient Analysis (SPA), can succeed using only second-order attacks. However, these attacks remain relatively inefficient, often requiring hundreds of thousands of traces for successful key recovery.

In summary, Cheng et al. [19] derived the optimal form of HO-CPA and evaluated its effectiveness against both BM and IPM. Wu et al. [12] systematically analyzed KWA and SPA, showing that KWA is a special case of SPA. However, their analysis was limited to the second order, leaving the optimal forms of higher-order KWA and Higher-Order SPA (HO-SPA) for BM and IPM unexplored. Specifically, it remains unclear which form and order of HO-SPA or its variant is the most effective, and whether they can outperform HO-CPA against BM and IPM.

C. Our Contributions

In this work, we investigate the optimal form of Higher-Order Adjusted Spearman Coefficient Analysis (HO-ASPA) by leveraging higher-order statistical moments of side-channel leakage, and further compare the effectiveness of HO-ASPA, KWA, and HO-CPA under various leakage models.

First, we derive the optimal HO-ASPA for two-share BM across different order combinations under the HW leakage model. This extends the applicability of the second-order SPA for two-share BM in two directions. First, we propose a higher-order-moment-based HO-ASPA for two-share BM at higher attack orders. Second, we identify certain ineffective HO-ASPA configurations whose attack orders can substantially exceed the security order $n-1$. Furthermore, we analyze the signal-to-noise ratio (SNR) and characterize its relationship with the optimal correlation coefficient in the rank-domain. This SNR–correlation relationship also applies to IPM. In addition, we derived the optimal forms of HO-CPA and HO-ASPA under the Least-Significant-Bit (LSB) leakage model.

Second, we derive the optimal form of two-share HO-ASPA against IPM under the HW leakage model. We show that, unlike HO-CPA, for which higher-order attacks can be more effective than attacks at the lowest feasible order, HO-ASPA exhibits no significant performance difference across attack orders. This finding is consistent with the optimal rank-domain hypothetical prediction function derived in this work. Our derived HO-ASPA is also consistent with the findings in [12] and can likewise break the SOA. Moreover, under the HW leakage model, HO-ASPA significantly outperforms KWA in attack performance. Additionally, similar to BM, we derive the optimal forms of HO-ASPA under the LSB leakage model for IPM and conduct experimental comparisons. We also perform simulation-based experimental validation on a three-share IPM to further demonstrate the applicability of our HO-ASPA beyond the two-share setting.

Finally, we systematically evaluate the proposed HO-ASPA methods against the state-of-the-art distinguishers, especially HO-CPA [19] and A-2O-SPA [12]. Specifically, when targeting the two-share IPM, we exhaustively evaluate all order combinations $(o_1, o_2) \in [1, 7]^2$ under the HW leakage model. By introducing the optimal rank-domain hypothetical prediction function for HO-ASPA, the Guessing Entropy (GE) [24] reaches 0.01 with $30,500$ traces under the HW leakage model at $\sigma^2 = 0.01$, whereas A-2O-SPA [12] has $GE = 1.10$ after $4,500,000$ traces, indicating a trace number advantage of over $147\times$ while retaining the ability to break SOA. Furthermore, we identify alternative HO-CPA order combinations that achieve better key-recovery performance than that reported in [19]. For IPM, these combinations reduce the number of traces required by factors of $3\times$ and $5\times$ at noise levels of $\sigma^2 = 0.01$ and 1.00, respectively.

In summary, by deriving the optimal form of HO-ASPA, this work connects and extends the study of optimal HO-CPA combinations in [19] and the second-order rank-based analysis in [12]. Compared with [12], our approach retains the ability to break SOA while substantially improving attack efficiency. Overall, these findings enrich the range of non-profiling SCA

methods and provide a broader basis for evaluating the practical security of cryptographic implementations.

Remark 1. *Throughout this work, the term "optimality" of the proposed distinguisher and the term "optimal" in expressions such as "optimal hypothetical prediction function" and "optimal rank-domain hypothetical prediction function" refer to the prediction function that maximizes the (Pearson or Spearman) correlation coefficient between the leakage and the hypothetical leakage under the considered leakage assumptions, including the centered-product combination, the specified leakage model (HW or LSB), and the rank transformation. For brevity and readability, we do not explicitly restate these assumptions throughout the remainder of the paper.*

II. Preliminary

We first introduce the notation used throughout this work. Then, we present the rank-based distinguishers in this work, a formal presentation of IPM, and the associated definitions.

A. Notations

Let $\mathbb{F}_q$ denote a finite field, where $q = 2^\ell$ and $\mathbb{R}$ is the real number field. The parameter ℓ corresponds to the bit-width of the data path in cryptographic algorithms (typically taking a value of 4 or 8). Random variables are denoted by uppercase letters (e.g., X), while their specific realizations are denoted by lowercase letters (e.g., x).

Let $w_H(X)$ denote the HW of X, i.e., the number of ones in its binary representation. Let $LSB(X)$ denote the LSB of X, and let "$\oplus$" denote the XOR operation. Let $\widehat{X} = R(X)$ denote the rank after rank transformation with ties averaged. The expectation, variance, and covariance of random variables are denoted by $\mathbb{E}[X]$, $\mathbb{V}[X]$, and $\mathrm{Cov}[X, Y]$, respectively, while $\sigma[X]$ represents the standard deviation of X. During an attack, the attacker assumes a leakage model f and computes hypothetical leakage values $f(X_k)$, where X_k is an intermediate value related to a key guess k, and $\tilde{k}$ is the guessed key according to the distinguisher score.

B. Rank-Based Analysis and Evaluation Metrics

SCA exploits the statistical dependency between physically observable leakage (e.g., power consumption) and the hypothetical leakage values of internal states, which are dependent on the secret key. Statistical distinguishers are then employed to identify the correct key hypothesis. This section introduces two non-profiling, rank-based distinguishers: KWA and HO-SPA. Below, we first introduce the leakage models adopted in this work.

A sensitive variable X is split into n shares, denoted by $(S_1, S_2, \ldots, S_n)$, such that $X = S_1 \oplus S_2 \oplus \cdots \oplus S_n$. The physical leakage associated with each share S_i is denoted by L_i and is modeled as:

$$L_i = \mathcal{L}(S_i) + N_i, \quad i = 1, \ldots, n, \tag{1}$$

where $\mathcal{L}(\cdot)$ is the deterministic leakage function (e.g., w_H), and N_i represents independent and identically distributed (i.i.d.) additive white Gaussian noise with $N_i \sim \mathcal{N}(0, \sigma^2)$.

In higher-order SCA, the attacker combines multiple leakage signals $L_1, \ldots, L_n$ into a combined leakage C. The focus of this work is on the centered-product combination, which has been shown to be the most effective combination method [17]. We take the combined leakage parameterized by the order vector $\vec{o} = (o_1, \ldots, o_n)$ as in [19]:

$$C = (L_1 - \mathbb{E}[L_1])^{o_1} \cdot \cdots \cdot (L_n - \mathbb{E}[L_n])^{o_n}, \tag{2}$$

where $o_1, \ldots, o_n \geq 0$ specify the statistical order applied to each share's leakage.

The core of an attack is the distinguisher D, which uses statistical tools to differentiate the correct key guess from incorrect ones. Given a set of hypothetical leakage values $f(X_k)$ related to a key guess k and the combined leakage C, the distinguisher outputs a distinguishing score d_k. The guessed key $\tilde{k}$ is obtained by maximizing this score:

$$\tilde{k} = \underset{k \in \mathcal{K}}{\operatorname{argmax}}\, d_k = \underset{k \in \mathcal{K}}{\operatorname{argmax}}\, \mathrm{D}(f(X_k), C), \tag{3}$$

where $\mathcal{K}$ is the key guess space.

Since the distinguishers HO-SPA and KWA require a rank transformation over the entire sample space, let Ω be the sample set:

$$\Omega \subset \mathbb{R}. \tag{4}$$

When Ω is taken as the set of all possible values of the combined leakage $C(x, y)$ over $(x, y) \in \mathbb{F}_q \times \mathbb{F}_q$, its size is $|\Omega| = 2^{2\ell}$ (for $\ell = 8$, $|\Omega| = 65,536$). The rank transformation is then defined as follows.

Definition 1 (Rank Transformation $R(\cdot)$ and Ranked Leakage $\widehat{C}$)**.** *For any $c \in \mathrm{supp}(C)$, define the population counts*

$$\begin{aligned} r_<(c) &\triangleq \big|\{\omega \in \Omega : \omega < c\}\big|, \\ r_=(c) &\triangleq \big|\{\omega \in \Omega : \omega = c\}\big|. \end{aligned} \tag{5}$$

The mid-rank of C is then defined as

$$R(c) \triangleq r_<(c) + \frac{r_=(c) + 1}{2}. \tag{6}$$

The ranked leakage is $\widehat{C} = R(C)$, and when $r_=(c)$ is even, $R(c)$ may be a half-integer but still lies within $[1, |\Omega|]$.

Next, we introduce the rank-based distinguisher KWA [21], which is a non-parametric higher-order distinguisher. Its core idea is to test whether, under each key hypothesis k, the distributions of the samples of the combined leakage C, grouped according to the hypothetical leakage values $f(X_k)$, exhibit significant differences.

Definition 2 (Kruskal-Wallis Analysis, KWA)**.** *Given a grouping function $g : \mathcal{X} \to \mathcal{G}$ (e.g., $g(X_k) = f(X_k)$), which maps the key-dependent intermediate value to a finite set $\mathcal{G}$, the KWA distinguishing score is defined as:*

$$d_k = (m-1)\frac{\sum_{i \in \mathcal{G}} m_i \left(\mathbb{E}[\widehat{C}_i | g(X_k) = i] - \mathbb{E}[\widehat{C}]\right)^2}{\sum_{i \in \mathcal{G}} \sum_{j=1}^{m_i} \left(\widehat{C}_{ij} - \mathbb{E}[\widehat{C}]\right)^2}, \tag{7}$$

where m is the total number of traces, and m_i is the number of traces in group i (with $\sum_i m_i = m$). $\widehat{C}_{ij}$ is the rank of the

j-th trace in group i, $\mathbb{E}[\widehat{C}_i]$ is the average rank of group i, and $\mathbb{E}[\widehat{C}] = (m+1)/2$ is the overall average rank.

The rationale of KWA is that, when the key guess is incorrect ($k \neq k^*$), the grouping is usually close to random, and the group average ranks $\mathbb{E}[\widehat{C}_i]$ will converge to the overall average rank $\mathbb{E}[\widehat{C}]$, resulting in a d_k close to 0. When $k = k^*$, the correct grouping maximizes the inter-group differences, yielding a significantly higher distinguishing score.

HO-SPA [12] is another non-parametric distinguisher. It applies the traditional Pearson correlation attack to the ranked data, enabling it to capture monotonic nonlinear leakages.

Definition 3 (Higher-Order Spearman Coefficient Analysis, HO-SPA)**.** *HO-SPA obtains the distinguishing score by calculating the Spearman rank correlation coefficient between the hypothetical leakage $f(X_k)$ and the combined leakage C:*

$$d_k = \left|\widehat{\rho}(f(X_k), C)\right|, \tag{8}$$

where $\widehat{\rho}(X,Y) = \rho(\widehat{X},\widehat{Y})$ denotes the Pearson correlation coefficient (as defined in [17]) after converting variables X and Y to their ranks (denoted $\widehat{X}$ and $\widehat{Y}$).

C. Inner Product Masking

To counteract SCA, masking techniques are widely employed. The core idea is to randomly split a sensitive intermediate value X into multiple shares, ensuring that the leakage of any single share or a small number of shares does not directly reveal the secret. This work focuses on two important masking schemes: BM and IPM. In particular, BM is a special case of IPM, and both can be described within a unified code-based masking framework [15].

IPM is a masking scheme with higher algebraic complexity [8], [9]. It utilizes the inner product operation over finite fields to construct shares, thereby enabling the SOA property. In an n-share IPM scheme, a sensitive variable $X \in \mathbb{F}_q$ is encoded as follows:

$$\vec{S} = (S_1, \ldots, S_n) = \left(X \oplus \sum_{i=2}^{n} \alpha_i \cdot Y_i,\ Y_2,\ \ldots,\ Y_n \right), \tag{9}$$

where $Y_2, \ldots, Y_n \in \mathbb{F}_q$ are uniformly random masks, and $\vec{\alpha} = (\alpha_2, \ldots, \alpha_n)$ is a public IPM parameter with $\alpha_i \in \mathbb{F}_q \setminus \{0\}$.

For concreteness, Fig. 1 illustrates the two-share case ($n = 2$) of the IPM encoding and the corresponding side-channel leakage observation, which is the focus of our subsequent analysis. This example can be easily extended to more shares with corresponding leakages.

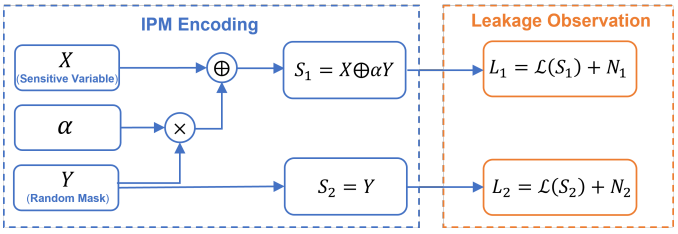


Fig. 1: Two-share IPM sharing, and the corresponding side-channel leakage observation.

From a coding-theoretic perspective [9], the encoding process of IPM can be represented as:

$$\vec{S} = X \cdot \mathbf{G} + \vec{Y} \cdot \mathbf{H}, \tag{10}$$

where $\mathbf{G}$ and $\mathbf{H}$ are the generator matrices of the linear codes $\mathcal{C}$ and $\mathcal{D}$, respectively, and $\vec{Y} = (Y_2, \ldots, Y_n)$ is the mask vector. The security of IPM is intrinsically linked to the dual distance $d_{\mathcal{D}}^{\perp}$ of $\mathcal{D}$. Specifically, its bit-level probing security order is given by $t_b = d_{\mathcal{D}}^{\perp} - 1$, implying that an attacker must combine the leakage from at least $t_b + 1$ shares to successfully recover the secret key.

The key advantage of IPM lies in the fact that, by carefully selecting the IPM parameter $\vec{\alpha}$ (i.e., defining the linear code $\mathcal{D}$), it is possible to achieve $t_b > t_w$, where $t_w = n - 1$ is the conventional word-level probing security order. This property, known as SOA, enables IPM to resist higher-order attacks.

With Eq. (10), one can easily obtain BM from IPM by taking the following specific parameters: $\alpha_2 = \cdots = \alpha_n = 1$. Within the coding framework, the dual code $\mathcal{D}^{\perp}$ corresponding to BM has a specific structure. Its generator matrix, after binary expansion, typically takes the form $(I_\ell, I_\ell, \ldots, I_\ell)$, where I_ℓ is the $\ell \times \ell$ identity matrix. This results in the bit-level probing security order t_b of BM being equal to its word-level security order t_w, i.e., $t_b = t_w = n - 1$. Consequently, BM cannot achieve security order amplification.

In summary, IPM provides a unified masking framework, with BM being its simplest instance. The subsequent analysis of HO-SPA will be conducted within this IPM framework.

III. HO-ASPA Against Boolean Masking

This section first introduces an original framework of the HO-SPA distinguisher. Based on this framework, we then propose the optimal rank-domain hypothetical prediction function for HO-ASPA under the HW leakage model. Next, we derive the optimal rank-domain hypothetical prediction function for both HO-CPA and HO-ASPA under the LSB leakage model. Finally, we present the SNR analysis.

A. HO-ASPA against BM under the HW Leakage Model

The original HO-SPA framework can be implemented by applying a rank transformation to the parameters of HO-CPA and then performing the HO-CPA computation. It is crucial to note, however, that while HO-CPA measures linear correlation, HO-SPA can capture the nonlinear correlations. This ability to extract nonlinear leakage information can be highly advantageous in certain scenarios.

Lemma 1 (Original HO-SPA Computation)**.** *Following [17], the correlation coefficient $\widehat{\rho}$ in HO-SPA is computed as:*

$$\begin{aligned} \widehat{\rho}(f(X), C) &= \rho(\widehat{f(X)}, \widehat{C}) \\ &= \rho(\widehat{f(X)}, \mathbb{E}[\widehat{C}|X]) \cdot \rho(\mathbb{E}[\widehat{C}|X], \widehat{C}). \end{aligned} \tag{11}$$

This observation leads to a practical method for HO-SPA: simply apply a rank transformation to the hypothetical prediction function and the combined leakage, and then perform HO-CPA. Indeed, this approach is reasonably effective and constitutes the basis of the original HO-SPA framework.

While Lemma 1 provides a straightforward method for computing HO-SPA, it is generally hard to find the form of f such that $\widehat{f}$ is optimal in Eq. (11). Referring back to Definition 3, the attack efficiency can be quantified by $d_k = |\rho|$ under ideal conditions. As in Eq. (11), the term $\rho(\mathbb{E}[\widehat{C}|X], \widehat{C})$ is fixed for a given set of traces. Therefore, maximizing $\widehat{\rho}(f(X), C)$ is equivalent to maximizing $\rho(\widehat{f(X)}, \mathbb{E}[\widehat{C}|X])$. This requires $\widehat{f(X)}$ to equal $\mathbb{E}[\widehat{C}|X]$. However, directly inverting the rank mapping to find an f satisfying $\widehat{f(X)} = \mathbb{E}[\widehat{C}|X]$ is impractical. We therefore adopt a simpler construction: instead of searching for f and then rank-transforming it, we directly construct the rank-domain hypothetical prediction function $f(X)$, which already resides in the rank-value domain $[1, |\Omega|]$. The correlation coefficient is then computed as $\rho(f(X), \widehat{C})$ without further rank-transforming but directly using f. We refer to this construction as HO-ASPA and define it as follows.

Definition 4 (Higher-Order Adjusted Spearman Coefficient Analysis, HO-ASPA)**.** *HO-ASPA obtains the distinguishing score by calculating the Pearson correlation coefficient between the hypothetical leakage $f(X_k)$ and the rank-transformed leakage $\widehat{C}$:*

$$d_k = \left|\rho(f(X_k), \widehat{C})\right|. \tag{12}$$

The optimality is defined at the statistical level by maximizing Pearson correlation after ranking (on the complete sample space Ω). The resulting f_{opt} is unique up to affine transformations $a \cdot f_{opt} + b$ with $a \neq 0$, which leave the Pearson correlation unchanged. We adopt the canonical choice $a = 1$, $b = 0$.

Accordingly, the optimal rank-domain hypothetical prediction function for HO-ASPA is [1]

$$f_{opt} = \mathbb{E}[\widehat{C}|X]. \tag{13}$$

Under the above condition, $\rho(f_{opt}(X), \widehat{C}) = \rho(\mathbb{E}[\widehat{C}|X], \widehat{C})$, and the ideal correlation coefficient ρ^s_{opt} is given by [2]:

$$\rho^s_{opt} = \frac{\sigma[\mathbb{E}[\widehat{C}|X]]}{\sigma[\widehat{C}]}. \tag{14}$$

The complete attack flow among the different distinguishers and their relationships is shown in Fig. 2, including HO-CPA, the original HO-SPA, and our HO-ASPA. In addition, KWA is a special case of HO-ASPA. It does not look for the function $f(x)$ or $\widehat{f(x)}$, but only relies on the last factor $\rho(\mathbb{E}[\widehat{C}|X], \widehat{C})$ after the original HO-SPA expansion (Eq. (11)) to obtain the distinguishing score (see [12] for the proof).

Considering the two-share BM under the HW leakage model, we shall ignore the independent noise when studying the higher-order moment effects. The centered-product combination in Eq. (2) can be expressed as:

$$C = \left(w_H(X \oplus Y) - \frac{\ell}{2}\right)^{o_1} \cdot \left(w_H(Y) - \frac{\ell}{2}\right)^{o_2}, \tag{15}$$

[1]Following the core idea of [12], we directly adopt $f_{opt} = \mathbb{E}[\widehat{C}|X]$ as the prediction function for subsequent analysis.

[2]A detailed proof can be found in [17], [19].

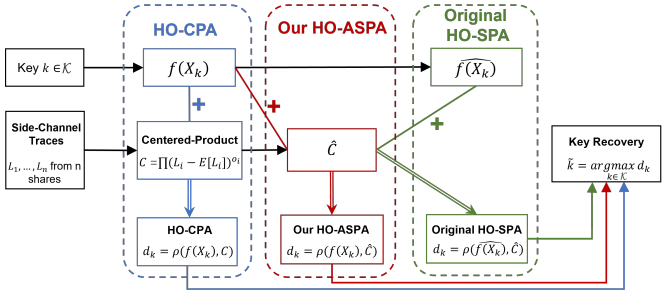


Fig. 2: Comparison of the attack workflow in HO-CPA, original HO-SPA, and our HO-ASPA.

where $X \in \{0,1\}^\ell$ is the sensitive intermediate variable and $Y \in \{0,1\}^\ell$ is a random mask. Since Y is uniformly distributed, the expectation of its leakage satisfies $\mathbb{E}[L_1] = \mathbb{E}[L_2] = \frac{\ell}{2}$. Then, based on Eq. (15), we present the optimal rank-domain hypothetical prediction function for HO-ASPA:

Theorem 1 (Optimal Rank-Domain Hypothetical Prediction Function for HO-ASPA under the HW Leakage Model for BM)**.** *Let $o_1 = 1$ and let o_2 be odd (so $o_1 + o_2$ is even). For any x (denoting $h = w_H(x)$) we have:*

$$\mathbb{E}[\widehat{C}|X = x] = G_\ell(h), \tag{16}$$

where $G_\ell(h)$ is given by the following explicit finite sum:

$$G_\ell(h) = \sum_{t=0}^{\ell} \sum_{j} \frac{\binom{h}{j}\binom{\ell-h}{t-j}}{2^\ell} R\Big(\big(h + t - 2j - \tfrac{\ell}{2}\big)\big(t - \tfrac{\ell}{2}\big)^{o_2}\Big), \tag{17}$$

and the summation index j is an integer that runs over

$$j \in [\max(0, t - (\ell - h)),\ \min(h, t)]. \tag{18}$$

Proof. Given the symmetry between the two shares in BM, we only prove the case where $o_1 = 1$ and $o_2 \geq 1$. Fix $X = x$ and let $h = w_H(x)$. Let $t = w_H(Y)$, $j = w_H(x \wedge Y)$. Then

$$w_H(x \oplus Y) = h + t - 2j.$$

Moreover, for fixed h and t, j follows a hypergeometric distribution $\Pr(j|t,h) = \binom{h}{j}\binom{\ell-h}{t-j}/\binom{\ell}{t}$. Since Y is uniform over $\{0,1\}^\ell$, we have $\Pr(t) = \binom{\ell}{t}/2^\ell$, hence $\Pr(t, j|h) = \binom{h}{j}\binom{\ell-h}{t-j}/2^\ell$. Taking the total expectation of $R(C(x, Y))$ over (t, j) yields Eq. (17). □

Particularly, for $\ell = 8$, exhaustive enumeration (see Table IV in Appendix A) shows that for all tested values of o_2, the function $G_8(h)$ closely approximates an affine function of h with a constant β (residual sum of squares below 10^{-3}), i.e.,

$$\mathbb{E}[\widehat{C}|X = x] = G_8(h) \approx \beta(w_H(x) - 4) + \frac{|\Omega| + 1}{2}. \tag{19}$$

Theorem 1 addresses the case where $o_1 + o_2$ is even, and the affine approximation in Eq. (19) is supported by numerical validation (see Table IV in Appendix A): for $\ell = 8$ and $o_1 = 1$, the residual sum of squares (RSS) between the function $G(h)$ and an affine function of h remains below 10^{-3} across a series of odd values of o_2. Next, we proceed to investigate the scenario where $o_1 + o_2$ is odd. We then have:

Corollary 1 (Conditional Expectation for Odd-Order Combinations in HO-ASPA)**.** *Under the same setting, if the total order $o_1 + o_2$ is odd, then for any x:*

$$\mathbb{E}[\widehat{C}|X = x] = \frac{|\Omega| + 1}{2}. \tag{20}$$

Proof. As in Theorem 1, we only prove the case where $o_1 = 1$ and $o_2 \geq 1$. When o_2 is even, for a fixed x, consider a pair $(y, \bar{y})$ with $\bar{y} = y \oplus 1^\ell$. Then $C(x, \bar{y}) = -C(x, y)$. By the symmetry of the overall distribution, $R(-c) = |\Omega| + 1 - R(c)$. Hence each pair $(y, \bar{y})$ satisfies $R(C(x, y)) + R(C(x, \bar{y})) = |\Omega| + 1$. Averaging over y gives $\mathbb{E}[\widehat{C}|X = x] = (|\Omega| + 1)/2$. □

The above results hold strictly under the rank definition. Corollary 1 shows that when the total order is odd, the conditional expected rank is constant, making HO-ASPA attacks ineffective. When the total order is even, the conditional expected rank depends on the HW of the sensitive value and exhibits an approximately affine relation to it, providing a theoretical incentive for a successful attack.

B. HO-ASPA and HO-CPA against BM under the LSB Leakage Model

In the previous subsection, we discussed the optimal rank-domain hypothetical prediction function for HO-ASPA under the HW leakage model. Here we further extend our analysis to the LSB leakage model. We first present the optimal hypothetical prediction function $\mathbb{E}[C|X]$ for HO-CPA under this model (the optimal hypothetical prediction function for the HW leakage model has been presented in [17]). From Eq. (2), we obtain the following combined leakage and the corresponding theorem for the LSB leakage model.

$$C = (LSB(X \oplus Y) - \frac{1}{2})^{o_1} \cdot (LSB(Y) - \frac{1}{2})^{o_2}. \tag{21}$$

Theorem 2 (Optimal Hypothetical Prediction Function for HO-CPA under the LSB Leakage Model for BM)**.** *Let $X, Y \in \{0,1\}^\ell$ be independent and uniformly distributed random variables and $p_1 = o_1 \bmod 2$, $p_2 = o_2 \bmod 2$. Let $o = o_1 + o_2$; the conditional expectation $\mathbb{E}[C|X]$ satisfies:*

$$\mathbb{E}[C|X] = \begin{cases} \frac{1}{2^o} & \text{if } p_1 = 0,\ p_2 = 0 \\ 0 & \text{if } p_1 \neq p_2 \\ \frac{1}{2^o}(1 - 2LSB(X)) & \text{if } p_1 = 1,\ p_2 = 1. \end{cases} \tag{22}$$

Specifically, the conditional expectation $\mathbb{E}[C|X]$ carries information about the sensitive intermediate value X only when both o_1 and o_2 are odd. The proof of Theorem 2 is provided in Appendix B.

We now proceed to derive the optimal rank-domain hypothetical prediction function for HO-ASPA under the LSB leakage model.

Theorem 3 (Optimal Rank-Domain Hypothetical Prediction Function for HO-ASPA under the LSB Leakage Model for BM)**.** *Let o_1 and o_2 be odd. Following the same setup as Theorem 2, we obtain:*

$$f_{opt} = \mathbb{E}[\widehat{C}|X] = |\Omega|\big(\frac{3}{4} - \frac{1}{2}LSB(X)\big) + \frac{1}{2}. \tag{23}$$

Proof. Let $a = LSB(X)$, $b = LSB(Y)$ and $o = o_1 + o_2$. Because o_1 and o_2 are odd, C simplifies to

$$C = \left(\frac{1}{2}\right)^o \times (-1)^{2-(a \oplus b)-b}.$$

Thus, conditioned on X, the random variable C takes only two distinct values:

$$c_0 = \left(\frac{1}{2}\right)^o \text{ (if } a = 0), \quad c_1 = -\left(\frac{1}{2}\right)^o \text{ (if } a = 1).$$

Because X is uniformly distributed, the two values appear with equal frequency in the overall population. Let $R(C)$ denote the rank after rank transformation. For $c_1 < c_0$, we have:

$$\widehat{c}_1 = R(c_1) = \frac{|\Omega|}{4} + \frac{1}{2}\ ,\ \widehat{c}_0 = R(c_0) = \frac{3|\Omega|}{4} + \frac{1}{2}\ .$$

Therefore, substituting $R(C)$ into the conditional expectation yields Eq. (23). □

Theorem 3 addresses the case where both o_1 and o_2 are odd. For the other parity combinations, the results are analogous to Theorem 2, and their proofs follow the same lines. When both o_1 and o_2 are even, C is a deterministic constant; consequently $\mathbb{E}[\widehat{C}|X]$ carries no information about X. When the parities of o_1 and o_2 differ, the conditional distribution of C is independent of X, and $\mathbb{E}[\widehat{C}|X]$ is again a constant. Hence, the exploitable leakage in the combined leakage C exists only when both o_1 and o_2 are odd.

C. SNR Analysis of HO-ASPA against BM

Based on the optimal form of HO-ASPA, we further analyze its theoretical properties. We first characterize the information-theoretic upper bound introduced by the rank transformation, and then establish the relationship between the SNR and the optimal correlation in the rank-transformed leakage domain.

Theorem 4 (Upper Bound on Maximal Correlation of Rank Transformation)**.** *Let C be the original combined leakage and $\widehat{C} = R(C)$ be its rank-transformed version. Define the Rényi maximal correlations [25] as:*

$$\begin{aligned} \rho_C^* &= \sup_{f_1,\, g_1} \rho\big(f_1(X),\, g_1(C)\big), \\ \rho_{\widehat{C}}^* &= \sup_{f_2,\, g_2} \rho\big(f_2(X),\, g_2(\widehat{C})\big), \end{aligned} \tag{24}$$

where the supremum is taken over all measurable functions f and g with finite variance. Then one has:

$$\rho_{\widehat{C}}^* \leq \rho_C^*. \tag{25}$$

Proof. Since $\widehat{C} = R(C)$ is a deterministic function of C, the σ-algebra generated by $\widehat{C}$ is contained in that generated by C:

$$\sigma(\widehat{C}) \subseteq \sigma(C). \tag{26}$$

For any measurable functions f and g with finite variance, the composition $g(\widehat{C})$ is $\sigma(C)$-measurable, so there exists a measurable function k such that $g(\widehat{C}) = k(C)$.

Hence, the function class induced by $\widehat{C}$ satisfies:

$$\big\{g_2(\widehat{C}) | g_2 \text{ measurable}\big\} \subseteq \big\{g_1(C) | g_1 \text{ measurable}\big\}. \tag{27}$$

Taking the supremum over these classes gives:

$$\begin{aligned}\rho^*_{\widehat{C}} &= \sup_{f_2,\, g_2} \rho\big(f_2(X),\, g_2(\widehat{C})\big) \\ &\le \sup_{f_1,\, g_1} \rho\big(f_1(X),\, g_1(C)\big) = \rho^*_C,\end{aligned} \tag{28}$$

which is a direct consequence of the data processing inequality for maximal correlation. □

Theorem 4 establishes an information-theoretic upper bound: the maximal achievable correlation from the rank-transformed leakage $\widehat{C}$ cannot exceed that from the original leakage C. However, this result characterizes only the theoretical upper bound under the maximal correlation criterion and does not directly determine the finite-sample performance of practical distinguishers. In practical attacks, the observed performance may further depend on factors such as the specific distinguisher, the leakage function, and the noise level of the target device.

To further connect the information-theoretic limitation with the traditional SNR metric used in SCA, we recall the relationship between SNR and the maximum achievable Pearson correlation coefficient. For a leakage variable C, the maximum achievable Pearson correlation coefficient is obtained by using the optimal prediction function $\mathbb{E}[C|X]$. Following [19, Eq. (11)], the SNR is defined as:

$$\mathrm{SNR}_C = \frac{\mathrm{Var}[\mathbb{E}[C|X]]}{\mathrm{Var}[C] - \mathrm{Var}[\mathbb{E}[C|X]]}. \tag{29}$$

Under the optimal hypothetical prediction function $\mathbb{E}[C|X]$, the maximum achievable Pearson correlation coefficient from [19] is:

$$\rho^p_{opt} = \frac{\sigma[\mathbb{E}[C|X]]}{\sigma[C]} = \frac{1}{\sqrt{1 + \frac{1}{\mathrm{SNR}_C}}}. \tag{30}$$

Equivalently, the SNR can be expressed as:

$$\mathrm{SNR}_C = \frac{(\rho^p_{opt})^2}{1 - (\rho^p_{opt})^2}. \tag{31}$$

When we perform a rank transformation, i.e., apply $\widehat{C} = R(C)$, the SNR of $\widehat{C}$ is given by:

$$\mathrm{SNR}_{\widehat{C}} = \frac{\mathrm{Var}[\mathbb{E}[\widehat{C}|X]]}{\mathrm{Var}[\widehat{C}] - \mathrm{Var}[\mathbb{E}[\widehat{C}|X]]}. \tag{32}$$

We have the following lemma.

Lemma 2 (SNR and Optimal Spearman Correlation Relationship)**.** *Let $\widehat{C} = R(C)$ be the rank-transformed leakage. The maximum achievable correlation coefficient obtained with the optimal prediction function $\mathbb{E}[\widehat{C}|X]$ satisfies:*

$$\rho^s_{opt} = \frac{\sigma[\mathbb{E}[\widehat{C}|X]]}{\sigma[\widehat{C}]} = \frac{1}{\sqrt{1 + \frac{1}{\mathrm{SNR}_{\widehat{C}}}}}. \tag{33}$$

Proof. The result directly follows by combining Eq. (14) with the definition of $\mathrm{SNR}_{\widehat{C}}$ in Eq. (32). □

Lemma 2 shows that the SNR-optimal correlation coefficient relationship can be directly established for the rank-transformed leakage $\widehat{C}$. However, this result only characterizes the relationship between SNR and the optimal correlation within each corresponding leakage domain, and does not directly provide a comparison between the optimal correlation coefficient of HO-ASPA and HO-CPA.

IV. HO-ASPA against Inner Product Masking

In this section, we begin by establishing the leakage model for IPM, and then derive the optimal rank-domain hypothetical prediction function for HO-ASPA under the HW leakage model. We also derive the optimal forms for both HO-CPA and HO-ASPA under the LSB leakage model. Similar to the BM case, the analyses and optimal distinguisher forms presented in this section are not limited to two-share masking. The three-share IPM experiment presented later in this work further illustrates the applicability of our HO-ASPA beyond the two-share setting.

A. Leakage Model

We first analyze the leakage model for IPM. It is important to note that the nonlinear relationship between the two-share vectors in IPM critically depends on the IPM parameter α. To thoroughly investigate and formally describe this nonlinear interaction in IPM at the bit level, we can represent the field multiplication operation (i.e., the effect of α) by an $\ell \times \ell$ linear transformation matrix, as in [19].

Let X denote the sensitive intermediate value, let α be the public IPM parameter, and let Y denote the random mask. According to the IPM scheme in Eq. (9), the leaked shares are given by $\vec{S} = (S_1, S_2)$, where $S_1 = X \oplus \alpha Y,\ S_2 = Y$. For $\alpha \in \mathbb{F}_q$, the multiplication map

$$T_\alpha : Y \mapsto \alpha Y, \tag{34}$$

is an $\mathbb{F}_2$-linear transformation, and it can be implemented with an $\ell \times \ell$ binary matrix $M_\alpha \in \mathbb{F}_2^{\ell \times \ell}$. The entries of M_α are:

$$(M_\alpha)_{i,j} = \mathrm{bit}_i\left(\alpha \cdot (1 \ll j)\right), \tag{35}$$

where $\mathrm{bit}_i(\cdot)$ extracts the i-th bit and $\ll$ denotes bit shifting. This matrix fully characterizes the leakage term:

$w_H(\alpha Y) = w_H(M_\alpha Y)$ or $LSB(\alpha Y) = LSB(M_\alpha Y)$.

The matrix representation M_α established above not only provides an explicit mathematical characterization of the nonlinear interaction in two-share IPM, but also serves as the basis for our subsequent analysis of higher-order leakages. In the following, we will leverage this formulation to systematically derive the optimal rank-domain hypothetical prediction function for HO-ASPA and quantitatively evaluate its distinguishing capability through theoretical SNR analysis.

B. HO-ASPA against IPM under the HW Leakage Model

We follow the IPM leakage setting described in [19]. In particular, the centered-product leakage C admits a character-sum expansion, and the conditional expectation $\mathbb{E}[C|X]$ can be expressed in closed form as shown in [19, Eq. (35)]. We do not reproduce the derivation here because our focus is on the rank-transformed leakage $\widehat{C} = R(C)$.

Let $X, Y \in \mathbb{F}_q$ be independent and uniformly distributed. Given a public parameter $\alpha \in \mathbb{F}_q$, consider the centered-product combination as in Eq. (2). The combined leakage C is given by:

$$C = \Big(w_H(X \oplus \alpha Y) - \frac{\ell}{2}\Big)^{o_1} \cdot \Big(w_H(Y) - \frac{\ell}{2}\Big)^{o_2}. \tag{36}$$

Theorem 5 gives the optimal rank-domain hypothetical prediction function for HO-ASPA against IPM. It parallels Theorem 1 for BM: while the BM optimal function reduces to a function of $w_H(x)$, the IPM counterpart depends on the IPM parameter α through the combined leakage $C(x, Y)$ and is computed by averaging after rank mapping $R(\cdot)$ over the random mask Y.

Theorem 5 (Optimal Rank-Domain Hypothetical Prediction Function for HO-ASPA under the HW Leakage Model for IPM)**.** *For any fixed $x \in \mathbb{F}_q$ under the setting described above, we have:*

$$\begin{aligned}\mathbb{E}[\widehat{C}|X = x] &= \sum_{c \in \mathrm{supp}(C)} R(c) \Pr\big(C = c|X = x\big) \\ &= \mathbb{E}_Y\big[R\big(C(x, Y)\big)\big].\end{aligned} \tag{37}$$

Proof. By definition, $\widehat{C} = R(C)$. The discrete expansion of the conditional expectation gives

$$\mathbb{E}[\widehat{C}|X = x] = \sum_c R(c)\ \Pr\big(C = c|X = x\big).$$

Under the condition $X = x$, the leakage satisfies $C \equiv C(x, Y)$. Therefore, one has

$$\sum_c R(c)\ \Pr\big(C = c|X = x\big) = \mathbb{E}_Y\big[R\big(C(x, Y)\big)\big]. \qquad \square$$

To apply Theorem 5, an attacker can follow four steps for an IPM instance with parameter α:

1) Enumerate all $y \in \mathbb{F}_q$;
2) Compute $C(x, y)$ via Eq. (36);
3) Apply the rank mapping $R(\cdot)$ by Definition 1;
4) Take the average over Y to obtain $f_{opt}(x)$.

Corollary 2. *If C is independent of X, i.e. $\Pr(C = c|X = x) = \Pr(C = c)$ for all c, x, then for any $X = x$,*

$$\mathbb{E}[\widehat{C}|X = x] = \frac{|\Omega| + 1}{2}. \tag{38}$$

Proof. Independence implies that the conditional distribution $C|X = x$ coincides with the distribution of C. Hence

$$\mathbb{E}[\widehat{C}|X = x] = \mathbb{E}[R(C)].$$

Now $R(C)$ is the mid-rank computed over the whole set Ω (see Eq. (4)): the average of the mid-ranks over all $|\Omega|$ samples is always $(|\Omega| + 1)/2$. Note that using mid-ranks for ties does not change the overall average. $\square$

Remark 2. *For $\ell = 8$, the size of the whole set $|\Omega| = 65,536$ is finite. Consequently, both the rank mapping $R(\cdot)$ and*

$$f_{opt}(x) \triangleq \mathbb{E}[\widehat{C}|X = x] = \mathbb{E}_Y\big[R\big(C(x, Y)\big)\big], \tag{39}$$

can be exactly computed by a full enumeration over all pairs $(x, y) \in \mathbb{F}_q \times \mathbb{F}_q$, without any Monte-Carlo error.

C. HO-ASPA against IPM under the LSB Leakage Model and SNR Analysis

In Section III-B, we investigated the optimal forms for BM in both HO-CPA and HO-ASPA under the LSB leakage model, which depends solely on the LSB of the sensitive intermediate value. However, in IPM, due to the presence of the public parameter α, neither HO-CPA nor HO-ASPA can capture leakage information.

We therefore extend the LSB leakage to multiple bits to examine the role of α more precisely. The combined leakage C considered in this analysis is constructed from $LSB_\zeta(X \oplus \alpha Y)$ and $LSB_\zeta(Y)$ using the same centered-product combination defined in Eq. (2), where $LSB_\zeta(\cdot)$ denotes the lower ζ bits.

Since multiplication by α is an $\mathbb{F}_2$-linear map, we split Y into its lower ζ bits Y_L and the remaining $\ell - \zeta$ bits Y_H, yielding $u = LSB_\zeta(\alpha Y) = A_\alpha Y_L \oplus B_\alpha Y_H$, where $A_\alpha \in \mathbb{F}_2^{\zeta \times \zeta}$ and $B_\alpha \in \mathbb{F}_2^{\zeta \times (\ell - \zeta)}$ are the sub-matrices of M_α (refer to Eq. (35)), and B_α controls how much of Y_H is injected into the observed bits. Let $r = \mathrm{rank}(B_\alpha)$. For fixed Y_L, $B_\alpha Y_H$ can attain exactly 2^r distinct values as Y_H varies.

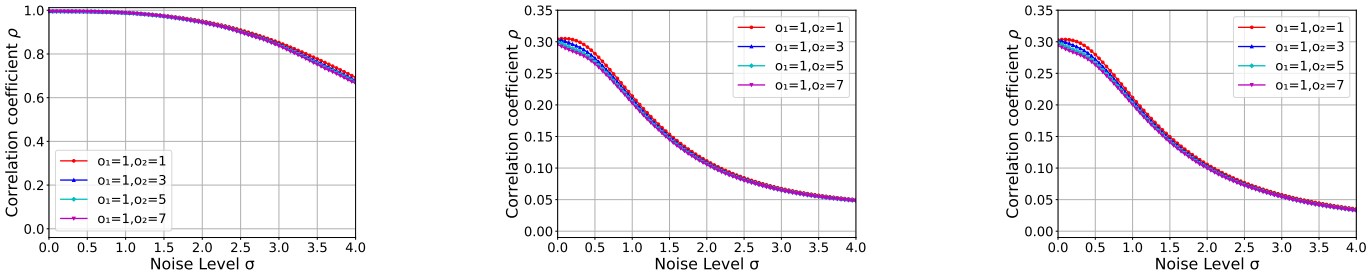
Lemma 3 (Full-Rank LSB Uniformity)**.** *If, for the selected bit width ζ, $\mathrm{rank}(B_\alpha) = \zeta$ (i.e., B_α has full row rank), then for any fixed Y_L, the random vector $u = LSB_\zeta(\alpha Y)$ is uniformly distributed over $\mathbb{F}_2^\zeta$, i.e., $\Pr(u|Y_L) = 2^{-\zeta}$.*

Proof. Since the masks Y are uniformly distributed, for fixed Y_L, the linear map $Y_H \mapsto B_\alpha Y_H$ is surjective onto $\mathbb{F}_2^\zeta$ (as B_α has full row rank). Hence each value of $B_\alpha Y_H$ occurs exactly $2^{\ell - 2\zeta}$ times, so it is uniform. Adding the fixed vector $A_\alpha Y_L$ does not affect uniformity, thus u is uniform. $\square$

By Lemma 3, when the rank condition holds, we have $\Pr(u|Y_L) = 2^{-\zeta}$, so u is independent of Y_L. The first share's leakage involves $LSB_\zeta(X \oplus \alpha Y) = LSB_\zeta(X) \oplus u$. Moreover, (u, Y_L) is a function of Y, and Y is independent of X (the random masks), so $(u, Y_L) \perp\!\!\!\perp X$. From $(u, Y_L) \perp\!\!\!\perp X$ we have $\Pr(u, Y_L, X) = \Pr(u, Y_L)\Pr(X)$, and from $Y \perp\!\!\!\perp X$ we have $\Pr(Y_L, X) = \Pr(Y_L)\Pr(X)$. Hence $\Pr(u|Y_L, X) = \Pr(u, Y_L, X)/\Pr(Y_L, X) = \Pr(u|Y_L) = 2^{-\zeta}$, and u is independent of (Y_L, X). Hence the joint distribution of $(LSB_\zeta(X \oplus \alpha Y), LSB_\zeta(Y))$ is identical for all $X = x$. Since C is a deterministic function of this pair, it also satisfies $C \perp\!\!\!\perp X$. The same holds for $\widehat{C} = R(C)$. Therefore, $\mathbb{E}[C|X]$ and $\mathbb{E}[\widehat{C}|X]$ are constant under the leakage model on the observed bits, with the centered-product construction and uniformly distributed independent masks. As a result, neither HO-CPA nor HO-ASPA can extract information about the sensitive variable from the LSB leakage under IPM.

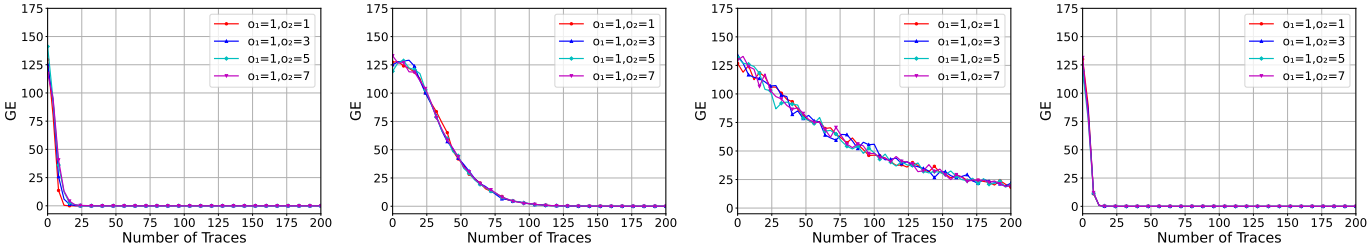
The condition $r < \zeta$ requires an α that fails to fully mix Y_H into the observed bits. The canonical example is $\alpha = 1$ (BM), where $B_\alpha = 0$ and $r = 0$, so the LSB attack succeeds. For the specific non-degenerate IPM parameter $\alpha = 23$ considered in our analysis, the matrix B_α has full row rank, leading to $r = \zeta$ in this case and the independence conclusion above.

We now consider the SNR of the rank-transformed leakage $\widehat{C}$ under IPM. As in the BM case, the SNR-correlation relationship established in Eq. (32) and Eq. (33) remains applicable, with $\widehat{C} = R(C)$ for IPM leakage. Therefore, the



(a) $\rho(f_{opt}, E[\widehat{C}|X])$ (b) $\rho(E[\widehat{C}|X], \widehat{C})$ (c) $\rho(f_{opt}, \widehat{C})$

Fig. 3: Evaluation of correlation coefficients on the two-share BM under HW leakage as noise levels increase.



(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 4: GE results of various non-profiling distinguishers on the two-share BM under LSB leakage at a noise level of $\sigma^2 = 0.01$.

SNR of $\widehat{C}$ is given by Eq. (32), and the corresponding optimal rank-domain correlation ρ^s_{opt} satisfies Eq. (33).

V. Experimental Results

Sections III and IV presented a theoretical analysis of BM and IPM in the noiseless setting. However, how different attack parameters actually behave in a realistic noisy environment remains an open question that requires empirical validation. Therefore, this section verifies our theoretical findings through systematic experimental evaluation.

As noted in [12], the effectiveness of KWA depends critically on the grouping function $g(X)$. In our experiments, we evaluate two grouping strategies: 1) $g(X) = X \bmod 128$, which corresponds to the grouping strategy employed in [12]; and 2) $g(X) = w_H(X)$, where we retain the same distinguisher but partition the traces according to the HW leakage function w_H. Correspondingly, we implement four distinct attack configurations:

- KWA$_{128}$: KWA with modulo-128 grouping;
- KWA$_{HW}$: KWA grouped by w_H;
- HO-CPA$_{opt}$: HO-CPA employs the optimal prediction function derived in [19];
- HO-ASPA$_{opt}$: HO-ASPA using the derived optimal rank-domain hypothetical prediction function f_{opt}.

To quantify and compare the attack efficiency across these configurations, we use GE [24] as the performance metric for different distinguishers. Note that we take the empirical means over 100 attacks to estimate GE when comparing different attacks. In each simulated run, the secret key remains fixed, while the plaintexts, masks, and Gaussian noise are randomly generated. For real-world datasets, traces are randomly selected. The reported GE is therefore the average over these 100 runs. All baselines (HO-CPA$_{opt}$, KWA$_{128}$, KWA$_{HW}$, HO-ASPA$_{opt}$) are evaluated on identical traces under identical preprocessing, ensuring a fair comparison.

A. Stability of the Results

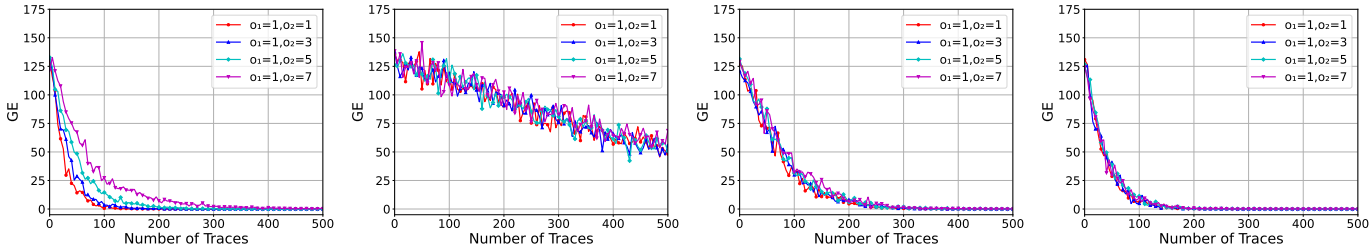
To assess the stability of the experimental results, we computed the standard deviation of the GE over 100 repeated attacks. We present two representative examples: for successful attacks, e.g., when $GE \leq 0.10$ as shown in Fig. 5a, the maximum standard deviation among the four curves at the points where the average GE reaches 0.10 is 0.0449, obtained for $(o_1, o_2) = (1, 5)$ at around 360 traces. For unsuccessful attacks, as shown in Fig. 5b, the maximum standard deviation among the four curves at 500 traces is 5.5143, obtained for $(o_1, o_2) = (1, 1)$. Therefore, as demonstrated by the representative cases examined, the standard deviations confirm the run-to-run stability of the GE-based evaluation under the tested conditions, which remains consistent across other cases.

B. Simulated Experimental Setting

In the theoretical analysis, we ignore noise to investigate the effects of different order combinations. To align with real-world scenarios, we incorporate Gaussian noise $N_i \sim \mathcal{N}(0, \sigma^2)$ into the leakage model in the experiments. To illustrate the effectiveness of HO-ASPA, we mainly conducted experiments on two-share BM and IPM. Moreover, we conducted experiments on a three-share IPM to show the applicability of our attack beyond the two-share setting. The sensitive variable is defined as $X = \texttt{Sbox}(P \oplus K)$, corresponding to the `Sbox` output in the AES implementation. The leakage of each share is modeled as $L_i = \mathcal{L}(S_i) + N_i$.

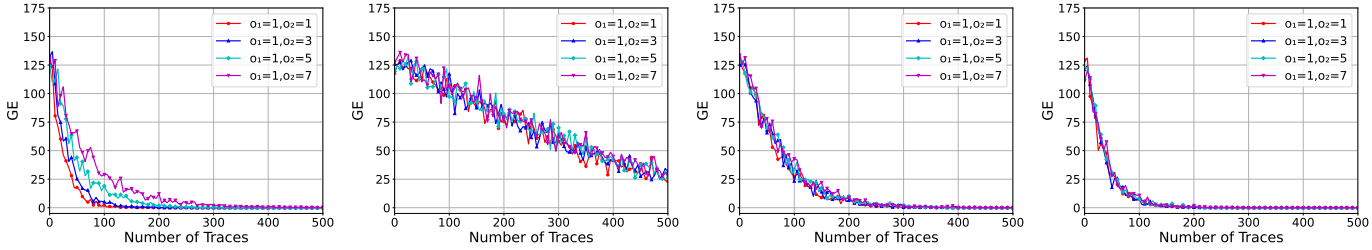
Regarding the leakage function $\mathcal{L}$, we considered three types of leakage models in our simulations:

- HW where $\mathcal{L}(S) = w_H(S) = \sum_{i=1}^{\ell} S[i]$, with $S[i]$ denoting the i-th bit of share S;



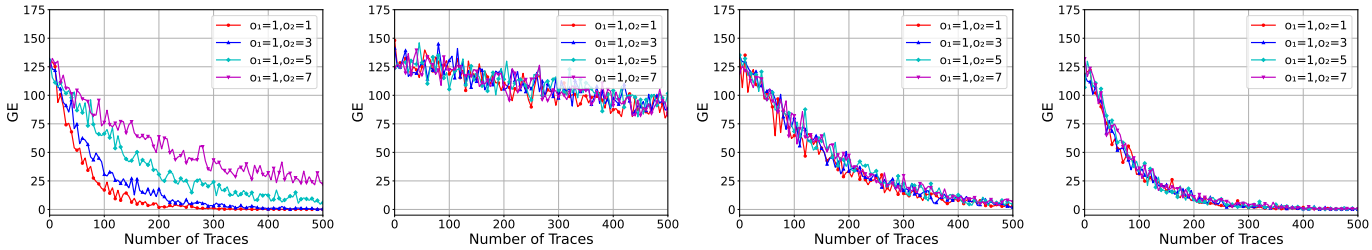
(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 5: GE results of various non-profiling distinguishers on the two-share BM under HW leakage at a noise level of $\sigma^2 = 0.01$.



(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 6: GE results of various non-profiling distinguishers on the two-share BM under RW leakage at a noise level of $\sigma^2 = 0.01$.



(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 7: GE results of various non-profiling distinguishers on the two-share BM under HW leakage at a noise level of $\sigma^2 = 1.0$.

- Randomly weighted (RW [3]) bits where $\mathcal{L}(S) = w_R(S) = \sum_{i=1}^{\ell} w[i]S[i]$ with each weight $w[i] \in [-1, +1]$;
- LSB where $\mathcal{L}(S) = LSB(S) = S\&1$ with $S\&1$ denoting the bitwise AND operation that extracts the LSB of S.

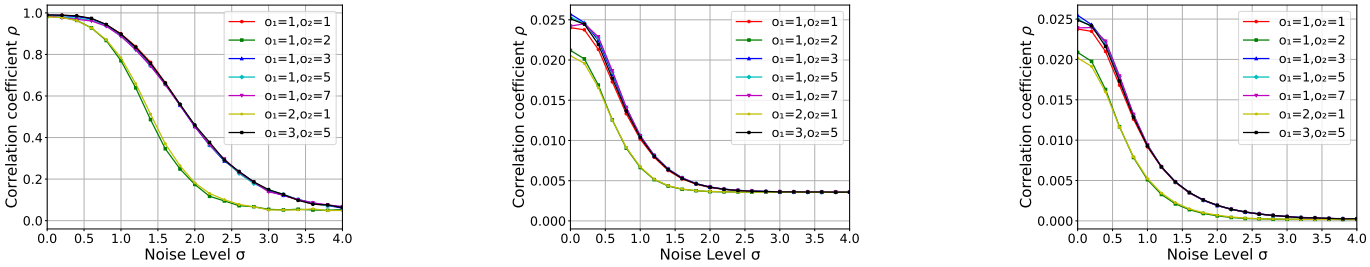
C. Various Non-Profiling Distinguishers against BM

In Theorem 1, we derived the optimal rank-domain hypothetical prediction function for HO-ASPA as $f_{opt} = \mathbb{E}[\widehat{C}|X]$. Therefore, we first show the variation of correlation coefficients along with noise levels in Fig. 3, where the product of the first two corresponds to the last one. As the noise level increases, all correlation coefficients decrease accordingly. In particular, when the noise level approaches zero, the correlation coefficient in Fig. 3a is close to 1, which verifies the optimal rank-domain hypothetical prediction function f_{opt} derived in Theorem 1. It is worth emphasizing that in Fig. 3, the correlation coefficients of each effective order combination show little difference. By Eq. (14), we can deduce that although higher-order combinations introduce substantial noise, the rank-based distinguisher HO-ASPA ensures that $\widehat{C}$ remains similar. This explains why the correlation coefficients for various order combinations exhibit minor variation.

Next, we examine the GE results of all distinguishers shown in Fig. 4 to Fig. 7. Under the LSB leakage model (see Fig. 4), KWA$_{128}$ is more effective than KWA$_{HW}$, which is opposite to the observation under the HW leakage model. Further observations of different order combinations for the HO-CPA$_{opt}$ distinguisher across all figures reveal that the most effective order combination in every leakage model is the lowest effective order, i.e., $(o_1, o_2) = (1, 1)$. This is primarily because higher-order combinations amplify noise. However, for rank-based distinguishers – KWA$_{128}$, KWA$_{HW}$, and HO-ASPA$_{opt}$ – there is no significant difference in attack performance among all effective order combinations. The reason for this is the same as in our analysis of Fig. 3. Comparing KWA$_{128}$ and KWA$_{HW}$ across the figures, it is evident that KWA$_{HW}$ outperforms KWA$_{128}$ under both HW and RW leakages. The fundamental reason lies in the fact that in BM, the leakage information in these two leakage models is related to the HW. It should be noted that KWA itself does not rely on HW; however, our experimental results demonstrate that a carefully designed grouping function can significantly enhance its distinguishing capability.

In all four figures (Fig. 4 to Fig. 7), HO-CPA$_{opt}$ and HO-ASPA$_{opt}$ achieve the best overall performance, and their results are very close. Recalling Theorem 1, when $o_1 = 1$ the optimal hypothetical prediction function for HO-CPA$_{opt}$ is $f_{opt} = w_H$, whose only difference from HO-ASPA$_{opt}$ lies in whether the traces are rank-transformed or not. However, it should be noted that HO-CPA$_{opt}$ achieves the strongest

[3]The simulated weights are $\vec{w}$ = [-0.25, 0.90, 0.46, 0.20, -0.69, -0.69, -0.88, 0.73], ordered from the least to the most significant bits.



(a) $\rho(f_{opt}, E[\widehat{C}|X])$ (b) $\rho(E[\widehat{C}|X], \widehat{C})$ (c) $\rho(f_{opt}, \widehat{C})$

Fig. 8: Evaluation of correlation coefficients on the two-share IPM under HW leakage as noise levels increase.

attack effectiveness with its most effective order combination. In summary, the takeaway is that both HO-CPA and HO-ASPA distinguishers demonstrate significant effectiveness against BM.

D. Non-Profiling Distinguishers against IPM

Following the same experimental setup used for the two-share BM case, we conducted experiments on IPM with public parameter $\alpha = 23$ (denoted IPM23 in the sequel). As shown in [9], this particular IPM instance offers a security order of $t_b = 3$. This implies that combined attacks of at least fourth order are required to be effective against IPM23.

Similar to BM, we first present the evolution of its correlation coefficients in Fig. 8. Compared with Fig. 3 for BM, the correlation coefficients in Fig. 8 clearly demonstrate the strength of IPM. Furthermore, the results in Fig. 8a when the noise level is close to 0 verify the optimal form of the rank-domain hypothetical prediction function proposed in our Theorem 5. Second, as we demonstrated in Section IV-C, under the LSB leakage model, when the public parameter α leads to $t_b > 1$, the combined leakages C and $\widehat{C}$ are independent of the sensitive variable, thus GE remains around 128 for all four distinguishers (we omit the results for brevity).

Comparing all distinguishers in Fig. 9 to Fig. 11, the most efficient distinguisher is HO-CPA$_{opt}$. Compared with KWA$_{128}$ and HO-ASPA$_{opt}$, we find that HO-CPA fails to recover the key when the combination order is lower than the security order t_b. However, Fig. 9b and Fig. 9d demonstrate that both KWA$_{128}$ and HO-ASPA$_{opt}$ are capable of breaking the SOA. Moreover, compared with KWA$_{128}$, HO-ASPA$_{opt}$ is a more powerful distinguisher.

TABLE I: GE of four distinguishers under optimal order combinations (with $(o_1, o_2) \in [1, 7]^2$) using 10,000 traces, under the HW leakage model with a noise level of $\sigma^2 = 0.01$.

Distinguishers	IPM1	IPM2	IPM3	IPM14	IPM23
HO-CPA$_{opt}$	(1,1): 0.0	(1,1): 0.0	**(1,4): 0.0**	**(1,7): 0.0**	**(7,1): 0.1**
KWA$_{128}$	(1,1): 0.0	(1,1): 0.0	(1,6): 85.5	(1,4): 107.0	(1,4): 109.0
KWA$_{HW}$	(1,1): 0.0	(1,1): 0.0	(4,1): 3.0	(3,7): 108.7	(2,7): 117.1
HO-ASPA$_{opt}$	(1,1): 0.0	(1,1): 0.0	(4,1): 0.1	(7,3): 10.2	(1,3): 20.3

Finally, in Tables I and II, we provide the GE values for the optimal order combinations of each distinguisher, where orders o_1 and o_2 vary from 1 to 7. From Tables I and II, we observe that HO-CPA and HO-ASPA are the most effective distinguishers. Furthermore, the above simulation experiments show that, compared with HO-CPA, HO-ASPA can break SOA, and significantly outperforms KWA.

TABLE II: GE of four distinguishers under optimal order combinations (with $(o_1, o_2) \in [1, 7]^2$) using 100,000 traces, under the HW leakage model with a noise level of $\sigma^2 = 1.00$.

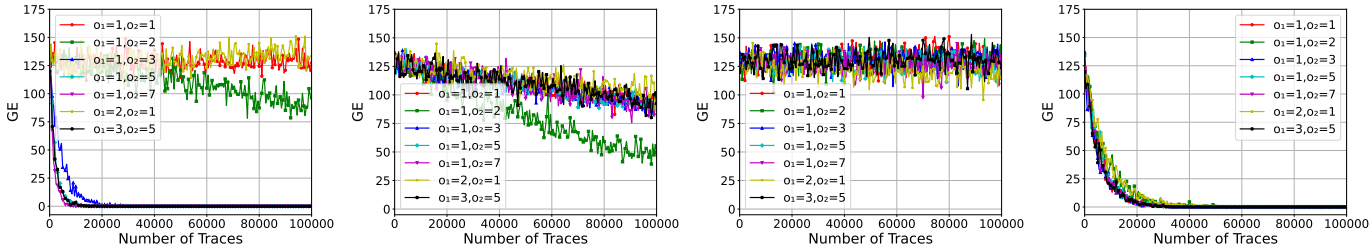
Distinguishers	IPM1	IPM2	IPM3	IPM14	IPM23
HO-CPA$_{opt}$	(1,1): 0.0	(1,1): 0.0	**(1,4): 0.0**	**(1,4): 0.0**	**(5,1): 0.0**
KWA$_{128}$	(1,1): 0.0	(1,1): 0.0	(1,6): 12.8	(1,4): 92.3	(3,6): 121.7
KWA$_{HW}$	(1,1): 0.0	(1,1): 0.0	(1,6): 0.2	(5,1): 78.7	(1,4): 111.5
HO-ASPA$_{opt}$	(1,1): 0.0	(1,1): 0.0	**(1,6): 0.0**	(7,1): 2.1	(1,5): 7.5

E. Extending to a Three-share IPM

To demonstrate the applicability of our analysis beyond the two-share setting, we extend our simulated experiments to three shares. Specifically, the parameter n in the IPM encoding in Eq. (9) is set to 3, and the centered-product leakage combination is applied using Eq. (2). We present experimental results on a three-share IPM in Fig. 12, where the IPM parameters are set to $\alpha_2 = 1$ and $\alpha_3 = 23$, corresponding to a security order of $t_b = 4$. We fix the first-share order to $o_1 = 1$ and vary the orders of the second and third shares, i.e., (o_2, o_3). Thus, each two-dimensional order label (o_2, o_3) in Fig. 12 corresponds to the complete order combination $(o_1, o_2, o_3) = (1, o_2, o_3)$, with total attack order $o = 1 + o_2 + o_3$. It can be observed that KWA distinguishers do not successfully recover the key regardless of whether modulo-128 grouping or HW grouping is used, while HO-CPA and HO-ASPA still retain distinguishing capability. Among them, the best distinguishing performance is still achieved by HO-CPA, which is consistent with Theorem 4. However, compared with HO-CPA, HO-ASPA still retains the advantage of breaking SOA. Specifically, under the parameter settings of $(o_1, o_2, o_3) = (1, 1, 1)$ and $(o_1, o_2, o_3) = (1, 1, 2)$, HO-ASPA can effectively reduce GE while maintaining a total order that does not exceed the security order.

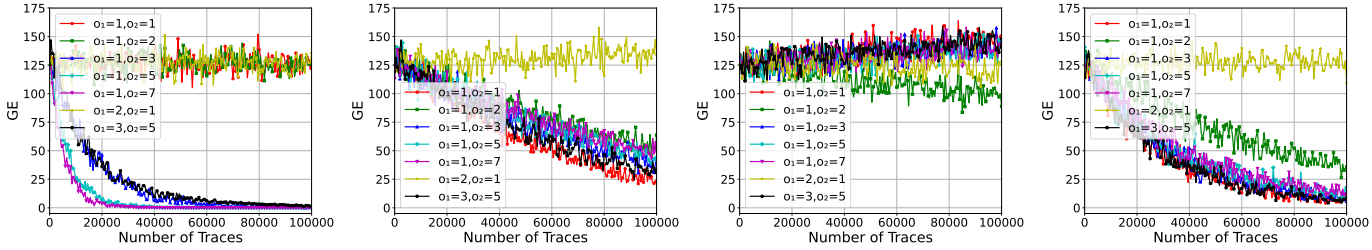
F. Real-World Attack Experiments

In this section, we validate our statistical analysis of non-profiling side-channel distinguishers against BM and IPM through real-world measurements. The first dataset is ASCAD, which contains traces from a masked AES implementation using two-share BM on an ATmega8515 MCU. Our measurements are acquired using a ChipWhisperer-Lite platform with an STM32F405 target at a sampling rate of 29.5 MS/s,



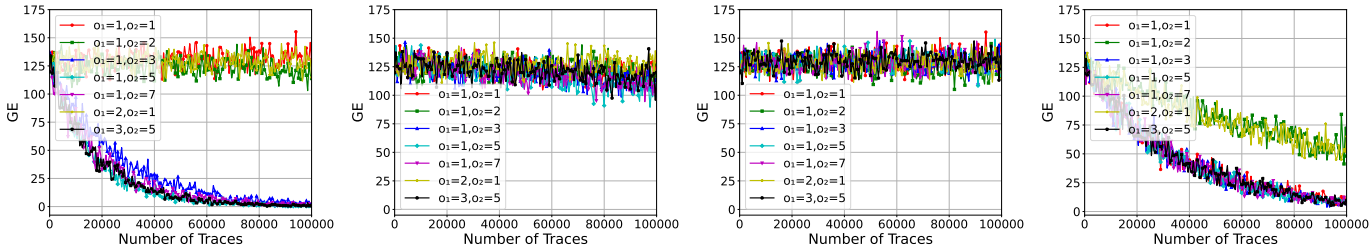
(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 9: GE results of various non-profiling distinguishers on the two-share IPM under HW leakage at a noise level of $\sigma^2 = 0.01$.



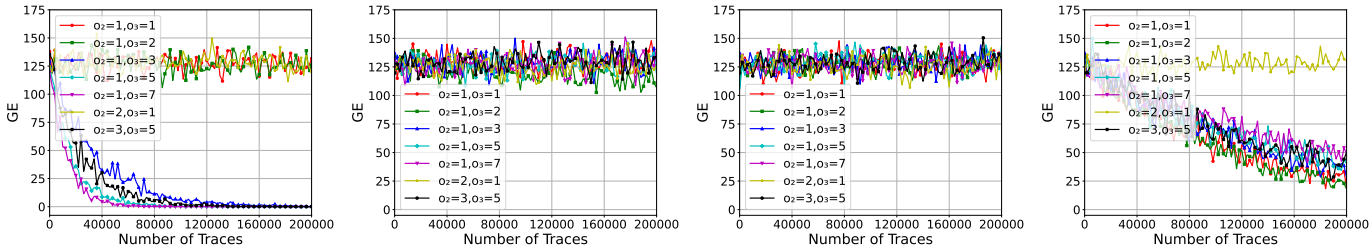
(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 10: GE results of various non-profiling distinguishers on the two-share IPM under RW leakage at a noise level of $\sigma^2 = 0.01$.



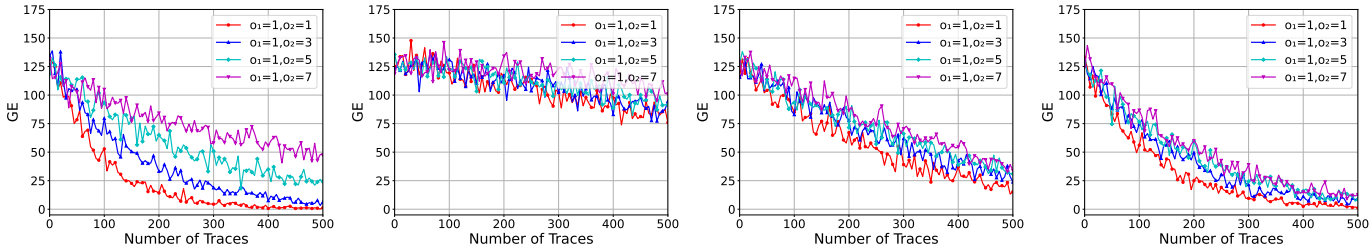
(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 11: GE results of various non-profiling distinguishers on the two-share IPM under HW leakage at a noise level of $\sigma^2 = 1.0$.



(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

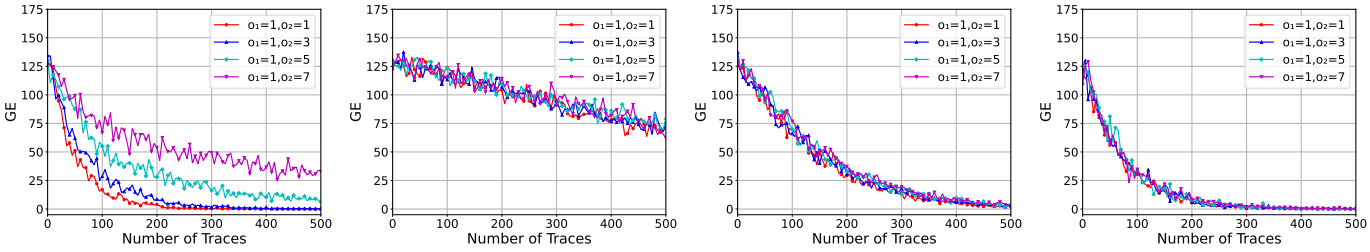
Fig. 12: GE results of various non-profiling distinguishers against three-share IPM under the HW leakage model, with o_1 fixed at 1 and a noise level of $\sigma^2 = 0.01$.



(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

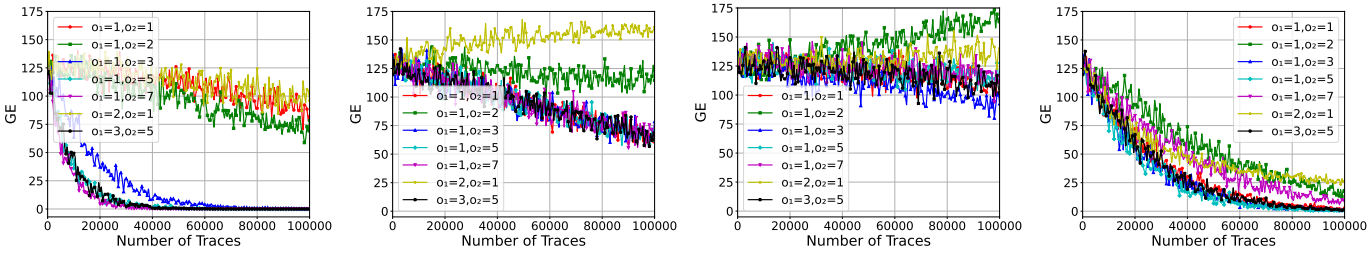
Fig. 13: GE results of various non-profiling distinguishers on the two-share BM of ASCAD on real-world measurements.

with $200,000$ power traces collected per configuration. We implement two-share IPM ($\alpha = 23$) and BM ($\alpha = 1$) for the AES S-box, following the publicly available higher-order IPM implementation of Balasch et al. [26]. In the preprocessing stage, we select the three points for each share that have the highest Pearson correlation coefficients with the HW of that



(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 14: GE results of various non-profiling distinguishers on the two-share BM of our database on real-world measurements.



(a) HO-CPA$_{opt}$ (b) KWA$_{128}$ (c) KWA$_{HW}$ (d) HO-ASPA$_{opt}$

Fig. 15: GE results of various non-profiling distinguishers on the two-share IPM of our database on real-world measurements.

TABLE III: Comparison with existing works by focusing on the BM and IPM, with $\alpha = 23$ for IPM, under the HW leakage model. Note that the metric for evaluating different distinguishers is the number of traces to achieve $GE \leq 0.10$, and the two noise levels are $\sigma^2 = 0.01$ (max #traces $= 500,000$) and $\sigma^2 = 1.00$ (max #traces $= 1,000,000$), where *fail* indicates unsuccessful attack with the max #traces. The best attack orders $(o_1, o_2) \in [1, 7]^2$ are also added for higher-order cases.

Related Works	**Experimental Type**	**Main Method**	**Break SOA**	**BM** (#Traces)	**IPM23** (#Traces)
Prouff et al. [17] IEEE TC 2009	Simulation	2O-CPA	No	$\sigma^2 = 0.01$: **170** $\sigma^2 = 1.00$: **390**	/
Ming et al. [18] IEEE TIFS 2022	Simulation + Real Device	HO-CPA$_{HW}$	No	/	$\sigma^2 = 0.01$: 21,100 (1,6) $\sigma^2 = 1.00$: 263,900 (1,6)
Yan et al. [21] ICISC 2023	Simulation + Real Device	2O-KWA$_{16}$	/	$\sigma^2 = 0.01 : 1,040$ $\sigma^2 = 1.00 : 2,380$	/
Cheng et al. [19] IEEE TC 2024	Simulation + Real Device	HO-CPA$_{opt}$	No	$\sigma^2 = 0.01$: **170 (1,1)** $\sigma^2 = 1.00$: **390 (1,1)**	$\sigma^2 = 0.01 : 13,900$ (1,5) $\sigma^2 = 1.00 : 98,900$ (1,5)
Wu et al. [12] IEEE TIFS 2025	Simulation + Real Device	2O-SPA	**Yes**	$\sigma^2 = 0.01 : 280$ $\sigma^2 = 1.00 : 560$	$\sigma^2 = 0.01$: Fail, $GE = 157.33$ $\sigma^2 = 1.00$: Fail, $GE = 127.68$
		A-2O-SPA$_{HW}$	**Yes**	$\sigma^2 = 0.01 : 240$ $\sigma^2 = 1.00 : 540$	$\sigma^2 = 0.01$: Fail, $GE = 157.94$ $\sigma^2 = 1.00$: Fail, $GE = 120.65$
		2O-KWA$_{128}$/A-2O-SPA$_{opt}$	**Yes**	$\sigma^2 = 0.01 : 2,390$ $\sigma^2 = 1.00 : 4,980$	$\sigma^2 = 0.01$: Fail, $GE = 39.37$ $\sigma^2 = 1.00$: Fail, $GE = 61.43$
This work	Simulation + Real Device	HO-CPA$_{opt}$	No	$\sigma^2 = 0.01$: **170 (1,1)** $\sigma^2 = 1.00$: **390 (1,1)**	$\sigma^2 = 0.01$: **7,500 (7,1)** $\sigma^2 = 1.00$: **52,600 (5,1)**
		KWA$_{128}$	**Yes**	$\sigma^2 = 0.01 : 2,390$ (1,1) $\sigma^2 = 1.00 : 4,980$ (1,1)	$\sigma^2 = 0.01 : 453,000$ (1,6) $\sigma^2 = 1.00$: Fail, $GE = 54.81$ (1,6)
		KWA$_{HW}$	**Yes**	$\sigma^2 = 0.01 : 370$ (1,1) $\sigma^2 = 1.00 : 820$ (1,1)	$\sigma^2 = 0.01$: Fail, $GE = 72.46$ (2,1) $\sigma^2 = 1.00$: Fail, $GE = 72.21$ (1,4)
		HO-ASPA$_{opt}$	**Yes**	$\sigma^2 = 0.01 : 220$ (1,1) $\sigma^2 = 1.00 : 500$ (1,1)	$\sigma^2 = 0.01 : 30,500$ (1,3) $\sigma^2 = 1.00 : 211,900$ (1,5)

share. Then, the leakage samples of the selected points are combined using a centered product to form a higher-order combined leakage C.

The results for BM and IPM are shown in Figs. 13, 14 and 15. In particular, Fig. 13 and Fig. 14 clearly demonstrate that for BM the most effective order combination among all distinguishers is still the lowest effective order, and the best-performing distinguisher remains HO-CPA$_{opt}$. All other conclusions align with our simulation results.

The IPM results are presented in Fig. 15. Similar to the simulation experiments, the poorest distinguisher is still KWA$_{HW}$, while the best distinguisher is HO-CPA$_{opt}$ (considering only the optimal order combination). As in the simulations, HO-CPA$_{opt}$ fails when the order is below the security order. In contrast, both HO-ASPA$_{opt}$ and KWA$_{128}$ retain the ability to break the SOA, and this capability is significantly stronger for HO-ASPA$_{opt}$ than for KWA$_{128}$.

In summary, we have verified our theoretical derivation of the HO-ASPA distinguisher for two-share BM and IPM through simulations and real-world measurements. The optimal rank-domain hypothetical prediction function derived in this work, when used as the hypothetical prediction function of HO-ASPA, demonstrates strong distinguishing power against both BM and IPM.

G. Comparative Analysis of Higher-Order Attacks

In the context of higher-order non-profiling SCA, numerous attack methods targeting BM and IPM have been proposed

in existing works. Therefore, we systematically compare the higher-order attacks in the simulated context by considering different order combinations (o_1, o_2) for each scheme. The comparative results are detailed in Table III.

As shown in Table III, the first observation is that the optimal order for all higher-order distinguishers when attacking two-share BM is $(1, 1)$, and second-order CPA (2O-CPA) maintains the highest efficiency against BM. The second observation is that HO-CPA remains the most efficient distinguisher at its optimal order. For IPM, the HO-ASPA$_{opt}$ scheme proposed in this work not only successfully breaks the SOA of IPM but also significantly outperforms other distinguishers (variants of KWA and SPA) capable of breaking SOA. Compared to other methods capable of breaking SOA, such as KWA and A-2O-SPA, HO-ASPA demonstrates a trace advantage of more than $147\times$ at $\sigma^2 = 0.01$ under the HW leakage model, with $(o_1, o_2) \in [1, 7]^2$ exhaustively searched. Specifically, HO-ASPA$_{opt}$ reaches $GE = 0.01$ with 30,500 traces, whereas KWA$_{128}$/A-2O-SPA$_{opt}$ still has $GE = 1.10$ after $4,500,000$ traces.

The results in Table III clarify the practical capability boundaries of different attack methods and confirm the necessity of systematically optimizing the attack order for accurately evaluating the resilience of masking in practice.

VI. Discussion and Conclusions

This work has systematically investigated the effectiveness of HO-ASPA in higher-order settings, with particular focus on BM and IPM. The theoretical analysis establishes that rank-based distinguishers are subject to an information-theoretic upper bound: the maximal correlation of the rank-transformed leakage cannot exceed that of the original leakage (Theorem 4). A link between HO-CPA and HO-ASPA could be found, e.g., leveraging hypotheses made in [27]. It is worth noting that these theoretical comparisons do not directly reflect the performance of attacks on finite samples observed in practice. Under the HW leakage model, for BM, the optimal rank-domain hypothetical prediction function is numerically observed to be approximately affine for the evaluated $\ell = 8$ setting, whereas for IPM the optimal rank-domain hypothetical prediction function exhibits a more complex structure determined by the underlying IPM parameters.

Extensive experimental evaluations confirm several key findings. In BM scenarios, the attack efficiency of different distinguishers peaks at the lowest effective order combination. For IPM, however, the most effective attack orders vary significantly across different distinguishers. In particular, KWA$_{HW}$ performs well against BM but performs poorly against IPM. Finally, compared with KWA$_{128}$, HO-ASPA not only retains the ability to break SOA but also significantly outperforms it.

These results provide important insights for both attack optimization and countermeasure design. Attackers can leverage the optimal hypothetical prediction function to enhance distinguisher efficiency, while defenders gain a clearer understanding of how noise injection and algebraic masking techniques can provide effective protection even against sophisticated rank-based attacks. Future work may extend this analysis to more complex leakage models and explore the application of these findings to emerging cryptographic implementations.

Acknowledgements

The authors sincerely thank the reviewers and the (associate) editors for their constructive and insightful comments, which have substantially strengthened the contribution of this work.

Appendix A
Numerical Verification of Theorem 1

TABLE IV: Normalized conditional expectation of rank on $G(h)$ for different values of o_2 and the HW value h, with $\ell = 8$ and $o_1 = 1$.

h	$o_2 = 1$	$o_2 = 3$	$o_2 = 5$	$o_2 = 7$
0	0.761750	0.758706	0.758599	0.758599
1	0.685850	0.687479	0.687465	0.687465
2	0.623005	0.620940	0.620846	0.620846
3	0.559898	0.559751	0.559768	0.559768
4	0.500000	0.500000	0.500000	0.500000
5	0.440102	0.440249	0.440232	0.440232
6	0.376995	0.379060	0.379154	0.379154
7	0.314150	0.312521	0.312535	0.312535
8	0.238250	0.241294	0.241401	0.241401

Appendix B
Proof of Theorem 2

Proof. Let $a = LSB(X)$, $b = LSB(Y)$, and $c = LSB(X \oplus Y)$. Then $c = a \oplus b$.

Let $p_i = o_i \bmod 2$ ($p_i = 0$ for even, 1 for odd). Then

$$C = \left(\tfrac{1}{2}\right)^{o_1+o_2}(2c-1)^{p_1}(2b-1)^{p_2}.$$

We compute $\mathbb{E}[C|X]$ in three cases:

1) If $p_1 = p_2 = 0$ (both o_1, o_2 even): $C = (\frac{1}{2})^{o_1+o_2}$ is constant, hence $\mathbb{E}[C|X] = (\frac{1}{2})^{o_1+o_2}$.

2) If $p_1 \neq p_2$ (one even, one odd): C contains the factor $(2c-1)$ or $(2b-1)$, which has mean zero and is independent of the other factor, so $\mathbb{E}[C|X] = 0$.

3) If $p_1 = p_2 = 1$ (both o_1, o_2 are odd): Here $C = (\frac{1}{2})^{o_1+o_2}(2c-1)(2b-1)$. Notice that

$$(2c-1)(2b-1) = \big[2(a \oplus b) - 1\big](2b-1) = 1 - 2a,$$

since the expression equals $1 - 2a$ for both $b = 0$ and $b = 1$. Therefore, $\mathbb{E}[C|X] = (\frac{1}{2})^{o_1+o_2}(1-2a)$. □

Appendix C
Pseudocode of the Optimal HO-ASPA Attack

References

[1] P. C. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," in *Advances in Cryptology - CRYPTO '99*, ser. LNCS, M. J. Wiener, Ed., vol. 1666. Springer, 1999, pp. 388–397. [Online]. Available: https://doi.org/10.1007/3-540-48405-1_25

[2] D. Agrawal, B. Archambeault, J. R. Rao, and P. Rohatgi, "The EM side-channel(s)," in *Cryptographic Hardware and Embedded Systems - CHES 2002*, ser. LNCS, B. S. K. Jr., Ç. K. Koç, and C. Paar, Eds., vol. 2523. Springer, 2002, pp. 29–45. [Online]. Available: https://doi.org/10.1007/3-540-36400-5_4

Algorithm 1 HO-ASPA$_{opt}$ Attack

Input: m traces $\{(L_1^{(j)}, \ldots, L_n^{(j)})\}_{j=1}^{m}$; order vector $\vec{o} = (o_1, \ldots, o_n)$; sample space $\Omega \subset \mathbb{R}$; key space $\mathcal{K}$
Output: Recovered key $\tilde{k}$

// Centered-product combination
for $i = 1$ **to** n **do**
  $\bar{L}_i \leftarrow \frac{1}{m}\sum_{j=1}^{m} L_i^{(j)}$ // sample mean of i-th share
end for
$C \leftarrow \prod_{i=1}^{n}(L_i - \bar{L}_i)^{o_i}$ // Eq. (2)
// Rank transformation
$\widehat{C} \leftarrow R(C)$
// Key recovery
for each $k \in \mathcal{K}$ **do**
  $f_{opt}(X_k) \leftarrow \mathbb{E}[\widehat{C}|X_k]$ // Theorem 1 / Theorem 5
  $d_k \leftarrow |\rho(f_{opt}(X_k), \widehat{C})|$ // Eq. (12)
end for
$\tilde{k} \leftarrow \underset{k \in \mathcal{K}}{\operatorname{argmax}}\, d_k$ // Eq. (3)
return $\tilde{k}$

[3] S. Chari, C. S. Jutla, J. R. Rao, and P. Rohatgi, "Towards sound approaches to counteract power-analysis attacks," in *Advances in Cryptology - CRYPTO '99*, ser. LNCS, M. J. Wiener, Ed., vol. 1666. Springer, 1999, pp. 398–412. [Online]. Available: https://doi.org/10.1007/3-540-48405-1_26

[4] P. C. Kocher, "Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems," in *Advances in Cryptology - CRYPTO '96*, ser. LNCS, N. Koblitz, Ed., vol. 1109. Springer, 1996, pp. 104–113. [Online]. Available: https://doi.org/10.1007/3-540-68697-5_9

[5] Y. Ishai, A. Sahai, and D. A. Wagner, "Private circuits: Securing hardware against probing attacks," in *Advances in Cryptology - CRYPTO 2003*, ser. LNCS, D. Boneh, Ed., vol. 2729. Springer, 2003, pp. 463–481. [Online]. Available: https://doi.org/10.1007/978-3-540-45146-4_27

[6] S. Mangard, E. Oswald, and T. Popp, *Power analysis attacks - revealing the secrets of smart cards*. Springer, 2007.

[7] L. Goubin and J. Patarin, "DES and differential power analysis (the "duplication" method)," in *Cryptographic Hardware and Embedded Systems*, ser. LNCS, Ç. K. Koç and C. Paar, Eds., vol. 1717. Springer, 1999, pp. 158–172. [Online]. Available: https://doi.org/10.1007/3-540-48059-5_15

[8] J. Balasch, S. Faust, and B. Gierlichs, "Inner product masking revisited," in *Advances in Cryptology - EUROCRYPT 2015*, ser. LNCS, E. Oswald and M. Fischlin, Eds., vol. 9056. Springer, 2015, pp. 486–510. [Online]. Available: https://doi.org/10.1007/978-3-662-46800-5_19

[9] W. Cheng, S. Guilley, C. Carlet, S. Mesnager, and J. Danger, "Optimizing inner product masking scheme by a coding theory approach," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 220–235, 2021. [Online]. Available: https://doi.org/10.1109/TIFS.2020.3009609

[10] W. Wang, F. Standaert, Y. Yu, S. Pu, J. Liu, Z. Guo, and D. Gu, "Inner product masking for bitslice ciphers and security order amplification for linear leakages," in *CARDIS 2016*, ser. LNCS, K. Lemke-Rust and M. Tunstall, Eds., vol. 10146. Springer, 2016, pp. 174–191. [Online]. Available: https://doi.org/10.1007/978-3-319-54669-8_11

[11] R. Poussier, Q. Guo, F. Standaert, C. Carlet, and S. Guilley, "Connecting and improving direct sum masking and inner product masking," in *CARDIS 2017*, ser. LNCS, T. Eisenbarth and Y. Teglia, Eds., vol. 10728. Springer, 2017, pp. 123–141. [Online]. Available: https://doi.org/10.1007/978-3-319-75208-2_8

[12] Q. Wu, W. Cheng, F. Zhang, and S. Guilley, "Statistical analysis of non-profiling higher-order distinguishers against inner product masking," *IEEE Trans. Inf. Forensics Secur.*, vol. 20, pp. 4008–4023, 2025. [Online]. Available: https://doi.org/10.1109/TIFS.2025.3558601

[13] L. Goubin and A. Martinelli, "Protecting AES with shamir's secret sharing scheme," in *Cryptographic Hardware and Embedded Systems - CHES 2011*, ser. LNCS, B. Preneel and T. Takagi, Eds., vol. 6917. Springer, 2011, pp. 79–94. [Online]. Available: https://doi.org/10.1007/978-3-642-23951-9_6

[14] J. Bringer, C. Carlet, H. Chabanne, S. Guilley, and H. Maghrebi, "Orthogonal direct sum masking - A smartcard friendly computation paradigm in a code, with builtin protection against side-channel and fault attacks," in *WISTP 2014*, ser. LNCS, D. Naccache and D. Sauveron, Eds., vol. 8501. Springer, 2014, pp. 40–56. [Online]. Available: https://doi.org/10.1007/978-3-662-43826-8_4

[15] W. Cheng, S. Guilley, C. Carlet, J. Danger, and S. Mesnager, "Information leakages in code-based masking: A unified quantification approach," *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, vol. 2021, no. 3, pp. 465–495, 2021. [Online]. Available: https://doi.org/10.46586/tches.v2021.i3.465-495

[16] E. Brier, C. Clavier, and F. Olivier, "Correlation power analysis with a leakage model," in *Cryptographic Hardware and Embedded Systems - CHES 2004*, ser. LNCS, M. Joye and J. Quisquater, Eds., vol. 3156. Springer, 2004, pp. 16–29. [Online]. Available: https://doi.org/10.1007/978-3-540-28632-5_2

[17] E. Prouff, M. Rivain, and R. Bevan, "Statistical analysis of second order differential power analysis," *IEEE Trans. Computers*, vol. 58, no. 6, pp. 799–811, 2009. [Online]. Available: https://doi.org/10.1109/TC.2009.15

[18] J. Ming, Y. Zhou, W. Cheng, and H. Li, "Optimizing higher-order correlation analysis against inner product masking scheme," *IEEE Trans. Inf. Forensics Secur.*, vol. 17, pp. 3555–3568, 2022. [Online]. Available: https://doi.org/10.1109/TIFS.2022.3209890

[19] W. Cheng, J. Ming, S. Guilley, and J. Danger, "Statistical higher-order correlation attacks against code-based masking," *IEEE Trans. Computers*, vol. 73, no. 10, pp. 2364–2377, 2024. [Online]. Available: https://doi.org/10.1109/TC.2024.3424208

[20] Y. Souissi, N. Debande, S. Mekki, S. Guilley, A. Maalaoui, and J. Danger, "On the optimality of correlation power attack on embedded cryptographic systems," in *WISTP 2012*, ser. LNCS, I. G. Askoxylakis, H. C. Pöhls, and J. Posegga, Eds., vol. 7322. Springer, 2012, pp. 169–178. [Online]. Available: https://doi.org/10.1007/978-3-642-30955-7_15

[21] Y. Yan, E. Oswald, and A. Roy, "Not Optimal but Efficient: A Distinguisher Based on the Kruskal-Wallis Test," in *ICISC 2023*, ser. LNCS, H. Seo and S. Kim, Eds., vol. 14561. Springer, 2023, pp. 240–258. [Online]. Available: https://doi.org/10.1007/978-981-97-1235-9_13

[22] B. Gierlichs, L. Batina, P. Tuyls, and B. Preneel, "Mutual information analysis," in *Cryptographic Hardware and Embedded Systems - CHES 2008*, ser. LNCS, E. Oswald and P. Rohatgi, Eds., vol. 5154. Springer, 2008, pp. 426–442. [Online]. Available: https://doi.org/10.1007/978-3-540-85053-3_27

[23] C. Whitnall, E. Oswald, and L. Mather, "An Exploration of the Kolmogorov-Smirnov Test as a Competitor to Mutual Information Analysis," in *Smart Card Research and Advanced Applications - 10th IFIP WG 8.8/11.2 International Conference, CARDIS 2011*, ser. LNCS, E. Prouff, Ed., vol. 7079. Springer, 2011, pp. 234–251. [Online]. Available: https://doi.org/10.1007/978-3-642-27257-8_15

[24] F. Standaert, T. Malkin, and M. Yung, "A unified framework for the analysis of side-channel key recovery attacks," in *Advances in Cryptology - EUROCRYPT 2009*, ser. LNCS, A. Joux, Ed., vol. 5479. Springer, 2009, pp. 443–461. [Online]. Available: https://doi.org/10.1007/978-3-642-01001-9_26

[25] A. Rényi, "On measures of dependence," *Acta mathematica hungarica*, vol. 10, no. 3-4, pp. 441–451, 1959.

[26] J. Balasch, S. Faust, B. Gierlichs, C. Paglialonga, and F. Standaert, "Consolidating inner product masking," *IACR Cryptol. ePrint Arch.*, vol. 2017, p. 1047, 2017. [Online]. Available: http://eprint.iacr.org/2017/1047

[27] N. Bruneau, S. Guilley, A. Heuser, and O. Rioul, "Masks Will Fall Off - Higher-Order Optimal Distinguishers," in *Advances in Cryptology - ASIACRYPT 2014, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014, Proceedings, Part II*, ser. Lecture Notes in Computer Science, P. Sarkar and T. Iwata, Eds., vol. 8874. Springer, 2014, pp. 344–365. [Online]. Available: https://doi.org/10.1007/978-3-662-45608-8_19