Beyond Side-Channels: Evaluating Inner Product Masking Against SIFA

Qianmei Wu[1], Sayandeep Saha[2], Wei Cheng[3,4], Fan Zhang[1], and Shivam Bhasin[5,6]

[1] School of Cyber Science and Technology, Zhejiang University, Hangzhou 310007, China qianmei;fanzhang@zju.edu.cn
[2] Department of Computer Science and Engineering, Indian Institute of Technology Bombay sayandeepsaha@cse.iitb.ac.in
[3] LTCI, Institut Polytechnique de Paris, T´el´ecom Paris, 91120 Palaiseau, France
[4] Secure-IC S.A.S., 75014 Paris, France wei.cheng@telecom-paris.fr
[5] Temasek Laboratories, Nanyang Technological University, Singapore
[6] National integrated Centre For Evaluation, Nanyang Technological University, Singapore sbhasin@ntu.edu.sg

Abstract. Statistical Ineffective Fault Attack (SIFA) presents a critical threat to cryptographic implementations by circumventing conventional detection-based countermeasures effective against traditional fault attacks. Particularly, SIFA operates via two mechanisms: SIFA-1 exploits fault effectiveness dependency on target values, while SIFA-2 leverages conditional propagation of faulted values based on sensitive intermediates. Recent studies suggest that, masking, mainly a side-channel protection, also exhibits promising resistance to SIFA-1, such as prime masking. In this paper, we systematically evaluate the resilience of Inner Product Masking (IPM) against SIFA, which has been established in prior works as a powerful side-channel-resistant alternative to Boolean masking. Specifically, with regard to SIFA-1, our theoretical analysis demonstrates that Inner Product (IP) encoding provides stronger SIFA-1 resistance than both Boolean and prime masking under generic multi-bit fault models using various fault types. More interestingly, an equivalence between Side-channel and SIFA-1 security has been theoretically established for IP encoding, indicating that optimal IP encoding exists that simultaneously provides the highest side-channel resistance and maximizes the complexity of effective SIFA-1 attacks. For SIFA-2, our analysis reveals that IPM's protection remains fundamentally bounded by the computational field size, consistent with previous results in this regard, e.g., for prime field masking. However, some vulnerabilities to *persistent faults* are anticipated for the most recently proposed IPM multiplication gadget. Given the compatibility with existing ciphers and demonstrated superior resistance against SIFA-1, IPM emerges as a more promising fault-resistant encoding technique compared to prime masking.

Keywords: Statistical Ineffective Fault Attack · Inner Product Masking · Side-channel Resistance.

1 Introduction

Side-Channel Attack (SCA) [24, 27] and Fault Injection Attack (FIA) [8, 7] constitute critical threats to physical cryptographic implementation. Side-channel adversaries exploit observable execution characteristics, including power consumption, electromagnetic emissions or timing variations, to extract sensitive information leaked during computation. To mitigate SCAs, *masking* (regarded as mainstream countermeasure) utilizes a randomized function f_{enc} (termed as *encoding function*), to map the sensitive variable Z to a set of random shares S_i with $1 \leq i \leq l$, such that $f_{\mathsf{enc}} : Z \rightarrow \{S_1, ..., S_l\}$, thus de-correlating Z from intermediate values S_i during execution and enhancing side-channel security. On the other hand, by actively manipulating either processed data or execution flow, fault-injection adversaries can subvert security mechanisms or force the disclosure of sensitive information, including various fault analysis techniques such as Differential Fault Analysis (DFA) [1], Fault Sensitivity Analysis [25], Differential Fault Intensity Analysis [22], etc. To defeat FIA, algorithmic protections primarily rely on computational redundancy (e.g., in space, time or information) to detect or correct maliciously injected errors [3].

Statistical Ineffective Fault Attack (SIFA) [18, 17] and Fault Template Attack (FTA) [34] have recently emerged as a great concern due to their inherent ability to exploit correct ciphertexts (or only correctness information) and thereby bypass most of the conventional redundancy-based and detection-based defenses [4]. SIFA manifests in two distinct variants [35, 29]: SIFA-1 exploits dependencies between fault effectiveness and target values (e.g., LSB stuck-at-0 faults affecting only odd numbers), while SIFA-2 (and FTA) leverages conditional propagation of faulted values based on a sensitive intermediate through non-linear operations (e.g., multiplication of a faulted value by 0 prevents fault propagation). Several mitigation strategies have been proposed so far to prevent SIFA. There predominantly exist two classes of countermeasures: 1) Redundant masked encoding [35, 14], in which the data-dependency of fault ineffectiveness is mitigated by leveraging masking (against SIFA-1) and per-share error-correction (against SIFA-2), ensuring protection up to a certain number of faulted bits; and 2) Modified masking, where fault propagation to the output is ensured for each fault, thereby making each fault effective [13, 16]. While initial versions of these countermeasures were found vulnerable to combined SCA and FIA [33, 36], more recently they have been enhanced to prevent such combined attacks [19, 6, 37].

The cost of preventing all these attacks is, however, substantial. Especially for combined attacks, the countermeasures utilizing Boolean masking (a cheap masking in terms of overheads) rely on Redundant masked encoding, requiring several registers and error-correction blocks within a multiplication gadget. Other proposals rely on polynomial masking. Most importantly, all of the proposed countermeasures provide security with respect to the *threshold fault model* [19, 20], which assumes that the adversary can corrupt up to k gates/wires in an implementation. While motivated by the threshold probing model [23] (which became highly successful for designing and proving the security of masking schemes), the threshold fault model does not reflect the complete reality,

in which a fault may also corrupt multiple bits with possibly weaker control over the bit locations being corrupted. Recent research has demonstrated that the multi-bit fault model represents a realistic threat scenario [4, 32, 37] in practice. The theoretical models have also been extended recently with the *random fault model* [15], and its generalized variant [5]. However, few provably secure constructions currently exist in such (multi-bit) theoretical models.

Although SCA and FIA represent seemingly orthogonal attack vectors, the "positive" role of masking in preventing SIFA, especially for the SIFA-1, is well-established [35, 36, 13]. The crux here is that even if one share remains uncorrupted and uniformly random, the entire state (even if faulty) remains uniformly random by Piling-up Lemma [28]. While this argument only holds for the threshold fault model (a fixed number of bits lower than the order of the Boolean masking can be faulted), recently Moos et al. [29] demonstrated that masking can provide something more. In particular, additive masking in prime fields can enhance the attack complexity exponentially with the share count for SIFA-1 faults, even if (a subset of) bits from all shares are faulted. The security argument holds even if such multi-bit faults are deterministic, that is, they corrupt the same set of bits with 100% repeatability. Such a fault model, while practically hard to achieve, provides an upper bound on the power of the adversary. For SIFA-2 faults, the security enhancement is limited (for multi-bit corruption in a single share), but still proportional to the field size. The impact of [29] is that it provides a seemingly simple way for preventing SIFA only with an error-check at the ciphertext (called *simple detection countermeasure* in this work), without requiring costly error-correction/detection inside the cipher computation, and almost for free, as prime-field masking shows extra resilience against SCA at low-noise scenarios. The fault model is precise multi-bit faults, which is even stronger than what most practical multi-bit adversaries can achieve in reality.

Following this line of research, we demonstrate in this paper that Inner Product Masking (IPM) [2] with *simple detection countermeasure* against traditional fault attacks, emerges as a new promising and more powerful countermeasure for both SCA and SIFA-1 security. Regarding side-channel security, IPM is a powerful alternative to Boolean Masking (BM), featuring three key advantages: *Security Order Amplification* (SOA) [41, 30], quantifiable security level [11] and inherent resistance to transitional leakage [2]. Theoretical analyses [11, 10] and empirical evaluations [42] collectively demonstrate that IPM achieves a higher bit-probing security order than Boolean masking under linear leakages, while utilizing identical share counts. While recent work [43] demonstrates that SOA can be partially compromised by rank-based distinguishers, extensive attack simulations in [43] confirm that IPM still provides substantially stronger side-channel resistance than Boolean masking (reaching up to 1000× improvement) under both linear and non-linear leakages against diverse higher-order distinguishers. Interestingly, in this work we show that IPM maintains significantly stronger resistance than BM against SIFA-1, mirroring its established SCA protection, thereby providing strong *dual* resistance against both SCA and faults. In ad-

dition, unlike prime field masking, which requires designing new ciphers on the prime field, IPM can be applied directly for traditional ciphers like AES.

1.1 Our Contribution

In this paper, we systematically evaluate the resilience of IPM against SIFA. First, we theoretically establish the resistance capability of Inner Product (IP) encoding to defeat SIFA-1, demonstrating its effectiveness as a fault-resistant encoding technique, which holds with *simple detection countermeasure*. In particular, we employ a multi-bit fault model, incorporating four extensively analyzed fault injection types: *stuck-at*, *toggle/bit-flip*, *constant-byte* and *random-byte* faults. Our key observations are as follows:

- Under *stuck-at* faults, we establish a theoretical framework to analyze the resilience of IP encoding against SIFA-1. Our analysis reveals that its security hinges on the interplay between the parameterized encoding and the fault model employed. Specifically, we develop two *matrix-rank parameters* (derived from the matrix constructed through both encoding parameters and fault models), $\mathcal{R}_{\overline{\mathcal{F}}}$ and $\mathcal{R}_{\mathcal{F}}$, to jointly characterize the security properties of IP encoding. First, one key rank parameter $\mathcal{R}_{\overline{\mathcal{F}}}$ determines whether a given IPM instance can theoretically withstand specific fault models. Second, in cases where vulnerability exists, both parameters $\mathcal{R}_{\overline{\mathcal{F}}}$ and $\mathcal{R}_{\mathcal{F}}$ help to qualitatively indicate and quantitatively assess the attack complexity.
- Under *toggle/bit-flip* and *random-byte* faults, we theoretically prove that SIFA-1 is incapable of defeating IP encoding.
- Under *constant-byte* faults, we demonstrate that successful SIFA-1 requires simultaneous corruption of all IPM shares, representing the full control of each bit during fault injections with high precision and thus imposing severe practical constraints on adversaries.

Similar to prime masking, IPM also provides limited protection against SIFA-2 attacks. In particular, corrupting only one share is sufficient for an attack. However, the probability of obtaining an ineffective fault is inversely proportional to the field size if the fault is injected at the input of a multiplication gadget. In case the fault is induced inside a gadget, we have demonstrated (based on the analysis of the most recent IPM gadget from CHES 2025 [21]) that the security margin remains the same, i.e., proportional to the field size. In other words, faulting inside this gadget does not result in more favorable cases for the adversary. However, we inherently assume protection against *persistent faults* [44], which can be relevant for IPM, at least for the existing implementations, highlighting constructing robust SIFA-2 IPM gadgets as an important future work.

Interestingly, our analysis reveals a fundamental connection between SIFA-1 resilience and SCA resistance of IP encoding under stuck-at faults, bridging two previously considered orthogonal attack classes. In particular, we formally establish that the coding-theoretic parameters characterizing side-channel resistance [11] in IP encoding can be repurposed to indicate its SIFA-1 resilience,

which is quantified by the minimum bit-wise faults required for effective attack and the amount of valid fault models at this minimum cost. Thus, there exist IP encoding that concurrently provide both optimal SCA protection and maximum resilience against a broad and extremely challenging class of SIFA-1 fault models. Given its compatibility with existing ciphers and demonstrated superior resistance against SIFA-1, IPM emerges as a more promising fault-resistant encoding technique compared to prime masking [29].

2 Preliminary

2.1 Notations

We denote the finite field of order q as $\mathbb{F}_q$ and limit that $q = 2^n$ in this paper. The field addition operation can be represented by "$\oplus$", while the symbol "$\cdot$" represents field multiplication over $\mathbb{F}_{2^n}$ or the inner product of two vectors over $\mathbb{F}_{2^n}^l$ according to the context. Vectors are denoted as $\vec{X}$ over $\mathbb{F}_{2^n}^l$ with $X[i]$ the i-th element and $|\vec{X}|$ is the size of vector. Let calligraphic $\mathcal{F}$ denote sets and $|\mathcal{F}|$ denote the cardinality of the set $\mathcal{F}$. In particular, we define $X[\mathcal{F}]$ (resp., $X[\overline{\mathcal{F}}]$) as the ordered sub-vector composed of components $\{X[i]|i \in \mathcal{F}\}$ (resp., $\{X[i]|i \notin \mathcal{F}\}$) of $\vec{X}$, where the elements appear in their original index order. For a natural number n we denote with $[n]$ the set of integers from 1 to n, both included. We use bold capital letters $\mathbf{A}$ to represent matrices over $\mathbb{F}_{2^n}^{r\times c}$, constructed with r rows and c columns. $\mathbf{A}[*, j]$ (resp., $\mathbf{A}[i, *]$) denotes the j-th column (resp., i-th row) of $\mathbf{A}$, and $\mathbf{A}[*, \mathcal{F}]$ (resp., $\mathbf{A}[*, \overline{\mathcal{F}}]$) with $\mathcal{F} \subseteq [c]$ denotes the matrix made up of the j-th column of $\mathbf{A}$ for $j \in \mathcal{F}$ (resp., $j \notin \mathcal{F}$), where the columns appear in their original index order. The symbol $\mathbf{A}^T$ denotes the transpose of $\mathbf{A}$. We use $\mathsf{Rank}(\mathbf{A})$ or $\mathcal{R}(\mathbf{A})$ to denote the rank of $\mathbf{A}$. For two matrices $\mathbf{A}$ and $\mathbf{B}$, we denote their product as $\mathbf{A} \times \mathbf{B}$ (or in short $\mathbf{AB}$), and $[\mathbf{A}, \mathbf{B}]$ is the concatenation of columns of $\mathbf{A}$ and $\mathbf{B}$. Unless otherwise noted, we assume the vectors are column vectors in this paper, and the row vectors are denoted as $\vec{X}^T$. In addition, the probability of the event $[X = x]$ is denoted by $Pr[X = x]$. Note that all indices in this paper start from 1 instead of 0.

An $[n, k, d]_q$-linear code $\mathcal{C}$ over $\mathbb{F}_q$ is a k-dimensional subspace of $\mathbb{F}_q^n$ with minimal (Hamming) distance d such that $k = \dim(\mathcal{C})$. The Hamming distance d_H between two vectors (codewords) of equal length is the number of positions at which the corresponding symbols differ. The minimum distance of a linear code $\mathcal{C}$ is defined as $d_{\mathcal{C}} = \min_{\vec{c_1},\vec{c_2}\in\mathcal{C}} d_H(\vec{c_1}, \vec{c_2})$, which equals the minimum weight of its nonzero vector (codeword). The (Euclidean) dual or orthogonal code $\mathcal{C}^\perp$ of linear code $\mathcal{C}$ is the set of vectors which are orthogonal to all vectors (codewords) of $\mathcal{C}$, such that $\mathcal{C}^\perp := \{\vec{u}|\vec{u} \cdot \vec{c} = 0, \forall \vec{c} \in \mathcal{C}\}$. Let $L : \mathcal{B} \to \mathcal{Z}$ be a linear transformation, represented by the matrix $\mathbf{L} \in \mathbb{F}^{r\times c}$ such that $L(\vec{b}) = \mathbf{L} \times \vec{b}$ for $\vec{b} \in \mathcal{B}$, where $\dim(\mathcal{B}) = c$ and $\dim(\mathcal{Z}) = r$. The *kernel* and *image* of L are denoted by $\ker(L)$ and $\mathrm{im}(L)$, respectively, with $\dim(\mathrm{im}(L)) = \mathsf{Rank}(\mathbf{L})$.

Since each $X \in \mathbb{F}_{2^n}$ can be represented by a polynomial $\sum_{i=0}^{n-1} a_i x^i$, where $\{1, x, \cdots, x^{n-1}\}$ form the basis of $\mathbb{F}_2^n$, we denote by $X \in \mathbb{F}_{2^n}$ and $[X]_2 \in \mathbb{F}_2^n$

the sensitive variable at n-bit word-level and bit-level, respectively. Since this paper involves transformation between $\mathbb{F}_2$ and $\mathbb{F}_{2^n}$, we formally present two key definitions as follows, which align with those given in [11].

Definition 1 (Sub-field Representation). *Let $z \in \mathbb{F}_{2^n}$, then the sub-field representation of z is $[z]_2 \in \mathbb{F}_2^n$. Let $\vec{Z} \in \mathbb{F}_{2^n}^l$, then the sub-field representation of $\vec{Z}$ is column vector over $\mathbb{F}_2^{nl}$ such that $[\vec{Z}]_2 = [[Z_1]_2, [Z_2]_2, \cdots, [Z_l]_2]$. We use the symbol $[\vec{Z}]_2[i]$ to denote the i-th bit (component) of $[\vec{Z}]_2$.*

For instance, considering $\vec{Z} \in \mathbb{F}_{2^8}^2$ and $\vec{Z} = \{1, 234\}$, its sub-field representation is $[\vec{Z}]_2 = [[1]_2, [234]_2] = [1,0,0,0,0,0,0,0,0,1,0,1,0,1,1,1]$.

Definition 2 (Code Expansion). *Let $L \in \mathbb{F}_{2^n}$, we can construct an $n \times n$ matrix $\mathbf{L}$ over $\mathbb{F}_2^{n \times n}$, where the j-th column of $\mathbf{L}$ corresponds to the coefficients of the polynomial $L \times x^{j-1}$. Hence for a vector $\vec{L} \in \mathbb{F}_{2^n}^l$, we can construct a matrix of size $n \times nl$ over $\mathbb{F}_2$.*

$$CodeExpansion : \vec{L} = [L[1], L[2], \cdots, L[l]]_{2^n} \rightarrow \mathbf{L} = [\mathbf{L_1}, \mathbf{L_2}, \cdots, \mathbf{L_l}]_2. \quad (1)$$

For instance, if $\vec{L} \in \mathbb{F}_{2^8}^2$ under the irreducible polynomial $x^8 + x^4 + x^3 + x + 1$ (the one used in AES) and let $\vec{L} = \{1, 23\}$, its code expansion is represented as in the left-most matrix. Assume an index set $\mathcal{F} = \{1, 3, 5, 8, 11, 12, 15\}$, then $\mathbf{L}[*, \mathcal{F}]$ and $\mathbf{L}[*, \overline{\mathcal{F}}]$ are shown in the middle and the right-most matrix, respectively.

$$\begin{bmatrix} 1&0&0&0&0&0&0&0&1&0&0&0&1&0&1&1 \\ 0&1&0&0&0&0&0&0&1&1&0&0&1&1&1&0 \\ 0&0&1&0&0&0&0&0&1&1&1&0&0&1&1&1 \\ 0&0&0&1&0&0&0&0&0&1&1&1&1&0&0&0 \\ 0&0&0&0&1&0&0&0&1&0&1&1&0&1&1&1 \\ 0&0&0&0&0&1&0&0&0&1&0&1&1&0&1&1 \\ 0&0&0&0&0&0&1&0&0&0&1&0&1&1&0&1 \\ 0&0&0&0&0&0&0&1&0&0&0&1&0&1&1&0 \end{bmatrix} \begin{bmatrix} 1&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1 \\ 0&1&0&0&1&0&1 \\ 0&0&0&0&0&1&0 \\ 0&0&1&0&1&1&1 \\ 0&0&0&0&0&1&1 \\ 0&0&0&0&1&0&0 \\ 0&0&0&1&0&1&1 \end{bmatrix} \begin{bmatrix} 0&0&0&0&1&0&1&0&1 \\ 1&0&0&0&1&1&1&1&0 \\ 0&0&0&0&1&1&0&1&1 \\ 0&1&0&0&0&1&1&0&0 \\ 0&0&0&0&1&0&0&1&1 \\ 0&0&1&0&0&1&1&0&1 \\ 0&0&0&1&0&0&1&1&1 \\ 0&0&0&0&0&0&0&1&0 \end{bmatrix} \quad (2)$$

At last, the dual distance and the kissing number of a linear code, formally defined in [11, 42], are given as follows.

Definition 3 (Dual Distance [26] and Kissing Number [38]). [7] *Considering a linear code $\mathcal{C}$, its dual distance $d_{\mathcal{C}}^{\perp}$ is the minimum Hamming weight $w_H(u)$ of nonzero $u \in \mathbb{F}^n$, such that $\sum_{c \in \mathcal{C}} (-1)^{c \cdot u} \neq 0$. The dual distance $d_{\mathcal{C}}^{\perp}$ of a linear code $\mathcal{C}$ coincides with the minimum distance $d_{\mathcal{C}^{\perp}}$ of the dual code $\mathcal{C}^{\perp}$. Accordingly, the kissing number B_d is the number of codewords in $\mathcal{C}$ at minimum distance d to the zero codeword: $B_d = |\{x \in \mathcal{C} \,|\, w_H(x) = d\}|$.*

[7] The kissing number is a geometric notion originally defined for lattices and sphere packings. Recently, it is adapted for linear codes over finite fields [38].

2.2 Inner Product Masking

The IPM encoding function $f_{\mathsf{enc}} : Z \to \vec{B}$ is defined via the inner product of two vectors, ensuring $Z = \vec{L} \cdot \vec{B}$, where $\vec{L}$ is a public and fixed vector, referred to as *IPM Parameters* in this paper. Note that all the components $L[i]$ of the vector $\vec{L}$ are non-zero elements over $\mathbb{F}_{2^n} \backslash \{0\}$. The shares vector $\vec{B}$ is structured as:

$$\vec{B} = \left[B[1] = Z \oplus \bigoplus_{i=2}^{l} B[i] \cdot L[i],\, B[2],\, \cdots,\, B[l] \right], \tag{3}$$

with $B[i]$ (for $i \in [2, l]$) randomly sampled from $\mathbb{F}_{2^n}$. For efficient implementation, the first element of $\vec{L}$ is usually assigned as $L[1] = 1$. In particular, Boolean Masking (BM) is a special case of IPM, where all the IPM parameters are equal to 1 such that $L[i] = 1$ for $i \in [l]$. Throughout this paper, we denote the IPM parameters vector by $\vec{L}$ and the shares vector by $\vec{B}$.

As established in [30, 11], the encoding function of IPM can alternatively be implemented through a coding-theoretic formulation:

$$\vec{B} = Z \times \mathbf{G} \oplus [B[2], \cdots, B[l]] \times \mathbf{H}, \tag{4}$$

where $\mathbf{G}$ and $\mathbf{H}$ are generator matrices of two linear codes $\mathcal{C}$ and $\mathcal{D}$, respectively. Specifically, two generator matrices over $\mathbb{F}_{2^n}$ and $\mathbb{F}_2$ are as follows:

$$\begin{aligned}
&\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{F}_{2^n}^{1 \times l} \qquad \mathbf{G} = \begin{bmatrix} \mathbf{I} & \mathbf{O} & \mathbf{O} & \cdots & \mathbf{O} \end{bmatrix} \in \mathbb{F}_2^{n \times nl} \\
&\mathbf{H} = \begin{bmatrix} L[2] & 1 & 0 & \cdots & 0 \\ L[3] & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L[l] & 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{F}_{2^n}^{(l-1) \times l}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{L_2}^T & \mathbf{I} & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{L_3}^T & \mathbf{O} & \mathbf{I} & \cdots & \mathbf{O} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{L_l}^T & \mathbf{O} & \mathbf{O} & \cdots & \mathbf{I} \end{bmatrix} \in \mathbb{F}_2^{(l-1)n \times nl},
\end{aligned} \tag{5}$$

where $\mathbf{I}$ and $\mathbf{O}$ denotes the identify and zero matrix, respectively. The coding form allows to quantify the concrete side-channel security of IPM by coding theory. Following [11], the side-channel resistance of IPM is characterized by two coding-theoretic parameters of linear code $\mathcal{D}$ over $\mathbb{F}_2$: the dual distance $d_{\mathcal{D}}^{\perp}$ and the kissing number $B_{d_{\mathcal{D}}^{\perp}}$ (of $\mathcal{D}$'s dual code). Precisely, $d_{\mathcal{D}}^{\perp}$ indicates the bit-probing security order [30] and a smaller $B_{d_{\mathcal{D}}^{\perp}}$ implies a higher security level at the same $d_{\mathcal{D}}^{\perp}$.

For the linear code $\mathcal{D}$ with its generator matrix as $\mathbf{H}$, the generator matrix of its dual code $\mathcal{D}^{\perp}$ can be constructed by the code expansion of IPM parameters vector $\vec{L}$ such that $\mathbf{L} = [\mathbf{L_1}, \mathbf{L_2}, \mathbf{L_3}, \cdots, \mathbf{L_l}]_2$ (recall $\mathbf{L_1} = \mathbf{I}$) since $\mathbf{H} \times \mathbf{L}^T = \mathbf{O}$ holds. Directly, we have the following corollary.

Corollary 1. *Let the IPM parameters vector $\vec{L}$ over $\mathbb{F}_{2^n}$ with $|\vec{L}| = l$ has its code expansion as $\mathbf{L} \in \mathbb{F}_2^{n \times nl}$ and assume a $[nl, n, d_{\mathcal{L}}]$-linear code $\mathcal{L}$ over $\mathbb{F}_2$ is generated by $\mathbf{L}$, we have $d_{\mathcal{L}} = d_{\mathcal{D}}^{\perp}$ and $B_{\mathcal{L}} = B_{d_{\mathcal{D}}^{\perp}}$.*

Proof. For the linear code $\mathcal{D}$, we have $d_{\mathcal{D}}^{\perp} = d_{\mathcal{D}^{\perp}}$ (recall Definition 3). Since $\mathcal{L}$ generated by $\mathbf{L}$ is the dual code of $\mathcal{D}$, directly we have $d_{\mathcal{L}} = d_{\mathcal{D}^{\perp}} = d_{\mathcal{D}}^{\perp}$ and thus $B_{\mathcal{L}} = B_{d_{\mathcal{D}^{\perp}}} = B_{d_{\mathcal{D}}^{\perp}}$. This completes the proof.

2.3 SIFA and FTA

Statistical Ineffective Fault Attacks (SIFA) and Fault Template Attacks (FTA) were proposed in CHES 2018 [18] and Eurocrypt 2020 [34], respectively, as effective techniques to overcome detection-based fault countermeasures. Both of them exploit statistical bias in a cipher's internal state distribution that emerge from fault injections when conditioned on fault ineffectiveness. Interestingly, faults exhibit value-dependent effectiveness, altering the internal state for certain sensitive values while leaving others unaffected. For example, consider a stuck-at-0 fault injected in the Least Significant Bit (LSB) of a word. This fault only affects the computation when the word's numerical value is odd, since the LSB of even numbers has already been 0 in binary representation. This value-dependent effectiveness creates the statistical bias that SIFA exploits. When an adversary exclusively observes fault-free computational results (where all faulty executions are filtered by detection mechanisms), the remaining internal states at the fault location introduce statistically deviation from uniform distribution (since no odd values are collected). Then, key recovery can be performed by testing candidate keys through backward computation to the fault location and verifying distribution uniformity, since only the correct key guess will produce the expected statistical bias at the fault site. According to the classification introduced in [35, 29], such attack instance is categorized as SIFA-1. This methodology can be extended to exploit fault *propagation* [34] through nonlinear operations. For instance, when a fault is injected into one input of an AND gate or field multiplication, it propagates to the output only if the other input is nonzero. This value-dependent fault propagation can also be used to obtain a distinguisher for key recovery. Attacks leveraging this mechanism are classified as SIFA-2 [35, 29]. While FTA combines SIFA-1 and SIFA-2 with template-building capabilities, enabling attacks on intermediate cipher rounds even without requiring plaintext/ciphertext access. Note that as shown in [29], the same mitigation strategies work for countering SIFA and FTA.

3 Inner Product Encoding vs. SIFA-1

In this section, we examine the resistance capability of IP encoding against SIFA-1 under a generic and realistic multi-bit adversary model, encompassing *stuck-at*, *toggle/bit-flip*, *constant-byte* and *random-byte* faults.

3.1 Threat Model

Given the focus on a generalized analysis of IP encoding resilience against SIFA-1, we adopt a generic yet strong multi-bit threat model. In particular, we assume that $\vec{C} = \mathsf{Alg}(\vec{P}, \vec{k})$, where $\vec{C}, \vec{P}, \vec{k}$ denotes unmasked outputs, inputs and sensitive keys, respectively. Let Alg denote an IP masked cryptographic circuit containing both linear and non-linear operations, equipped with *simple detection countermeasure* against fault attacks. By repeatedly injecting faults into the shares of a component $P[i]$ of inputs, the adversary aims to recover (part of) the secret key $\vec{k}$. The faulty inputs and the corresponding faulty outputs are denoted

as $\vec{P}^{'}$ and $\vec{C}^{'}$, respectively. The detection mechanism performs a post-decoding comparison between $\vec{C}$ and $\vec{C}^{'}$, producing outputs only if $\vec{C} = \vec{C}^{'}$.

Let $Z = P[i] = \vec{L} \cdot \vec{B}$ where $Z \in \mathbb{F}_{2^n}$ and $\vec{L}, \vec{B} \in \mathbb{F}_{2^n}^l$ (note that the elements of $\vec{L}$ are nonzero) be the vulnerable variable. We assume arbitrary set of bit-wise faults are injected directly into the bit-wise shares $[\vec{B}]_2$ simultaneously with perfect precision. Bit-wise fault captures practical adversarial behavior, manipulating binary signals and individual bits, which has been extensively considered in the prior works [36, 29, 33, 34, 17, 5]. While this fault model, which assumes perfectly precise, synchronous and repeated multi-bit injections across all shares, is challenging to realize in practice, it establishes an upper bound on the adversarial capabilities, thereby defining a theoretical lower bound on the resistance of IP encoding. Let $\mathcal{F} \subseteq \{1, \cdots, nl\}$ denote the set of fault locations with $|\mathcal{F}| \leq nl$. The faulty shares are represented by $\vec{B}^{'}$ with its sub-field representation as $[\vec{B}^{'}]_2$. To reiterate, we assume that all effective faults trigger the *simple detection countermeasure*, resulting in secure abortion rather than propagation of corrupted outputs. Consequently, adversaries can only obtain $\vec{C}^{'} = \mathsf{Alg}(\vec{P}^{'}, \vec{k})$ with ineffective faults $Z^{'} = Z$ to launch attacks.

Since SIFA-1 employs biased distribution of $Z^{'}$ with ineffective faults and its dependency on the value of Z to retrieve the correct key $\vec{k}$, we first introduce the following fundamental corollary.

Corollary 2 (Conditions for Effective SIFA-1). *Let $Z \in \mathbb{F}_{2^n}$ be vulnerable variable and $Z^{'} \in \mathbb{F}_{2^n}$ its faulty counterpart. SIFA-1 can be effective if satisfying:*

- $0 < Pr[Z = Z^{'}] < 1$,
- *$Z^{'}$ follows a non-uniform distribution over $\mathbb{F}_{2^n}$ under the ineffective faults when satisfying $[Z = Z^{'}]$.*

Proof. The probability $Pr[Z = Z^{'}]$ indicates the possibility of ineffective faults occurring, thereby implying that adversaries may potentially obtain the corresponding outputs, hence we have $Pr[\vec{C} = \vec{C}^{'}] = Pr[Z = Z^{'}]$. If $Pr[\vec{C} = \vec{C}^{'}] = Pr[Z = Z^{'}] = 0$, then no ineffective faults occur, preventing adversaries from obtaining valid outputs and resulting in attack failure; if $Pr[\vec{C} = \vec{C}^{'}] = Pr[Z = Z^{'}] = 1$, then Alg is fault-free, implying that the ineffectiveness depends on Z for none of the cases. Whereas $0 < Pr[\vec{C} = \vec{C}^{'}] = Pr[Z = Z^{'}] < 1$ indicates that adversaries can obtain outputs at a certain possibility and the ineffectiveness relies on the values of Z. However, even when $0 < Pr[C = C^{'}] < 1$ holds, the uniform distribution of $Z^{'}$ over $\mathbb{F}_{2^n}$ with ineffective faults when satisfying $[Z = Z^{'}]$ prevents distinguishing the true key $\vec{k}$ among all key candidates, resulting in SIFA-1 failure. This completes the proof.

To further clarify the case where $Pr[Z = Z^{'}] = 1$, we utilize a two-share Boolean masking as a simple example. Assume that bit-flip faults are directly injected to the LSB of Boolean shares $\vec{B} = [B[1], B[2]]$ such that $B[1]^{'} = B[1] \oplus 1$, $B[2]^{'} = B[2] \oplus 1$. Hence, we have $B[1] \oplus B[2] = B[1]^{'} \oplus B[2]^{'}$ for all the cases of $\vec{B}$, resulting in $Pr[Z = Z^{'}] = 1$ and thereby attack failure.

SIFA-1 Attack Complexity. Under a valid SIFA-1 model, which satisfies the conditions established in Corollary 2, the *attack complexity* of SIFA-1 informally depends on three factors:

- the probability $Pr[C = C^{'}]$ of obtaining valid outputs;
- the *uniformity degree* of the distribution of $Z^{'}$ with ineffective faults;
- the distinguisher's capability (e.g., Squared Euclidean Imbalance, SEI) to identify the correct key among candidate keys.

The first factor quantifies the number of ineffective faults the adversary obtains per fixed number of fault attempts, while the latter two factors determine how many correct outputs must be collected to accumulate sufficient distributional information for a successful SIFA attack. As this work focuses on IPM's resistance to SIFA-1, we exclude the investigation of distinguisher capabilities and instead consider the remaining two factors. The term *uniformity degree* characterizes the *distance* (e.g., SEI) of the distribution of $Z^{'}$ under ineffective faults from uniform distribution with same support. It directly impacts attack efficiency: the higher the degree, the greater the attack complexity becomes. For instance, considering an extreme case where $Z^{'}$ follows a uniform distribution, such attacks even cannot be effective, as established in Corollary 2.

A Toy Circuit for Simulated Attack. In this section, to assess the resilience of IP encoding against attacks, we employ a minimal toy circuit, enabling a streamlined analysis free from dependencies on concrete primitives. Specifically, we adopt the minimum working example devised in [29] for simulated attacks, as illustrated in Figure 1. Particularly, $P[0], P[1]$ are plaintext words, $k[0], k[1]$ are key words, and $C[0], C[1]$ are ciphertext words, each of size $n = 8$ bits. M is a linear operation, implemented as multiplication by a matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Meanwhile, S represents a non-linear operation, specifically an AES S-box under irreducible polynomial $x^8 + x^4 + x^3 + x + 1$, as also assumed in [29]. We consider SIFA-1 that repeatedly injects faults in the encoding of $P[0]$ (thus $P[0]$ is the vulnerable variable Z) in order to obtain a distinguisher for $\vec{k} = [k[0], k[1]]$ from the collection of valid outputs $\vec{C} = [C[0], C[1]]$ with ineffective faults. Note that we omit the redundant detection circuitry here, operating under the assumption that the ciphertext $\vec{C}$ will be obtained by adversaries exclusively when $\vec{C} = \vec{C}'$ holds.

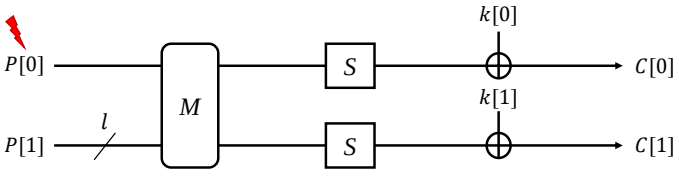
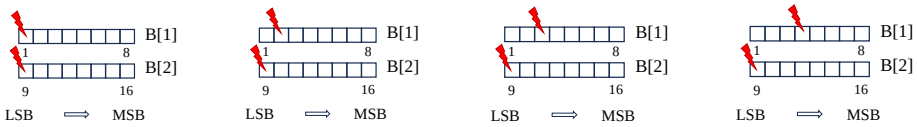


Fig. 1: Toy cipher circuit for SIFA-1 simulation, with fault location denoted by a red lightning symbol.

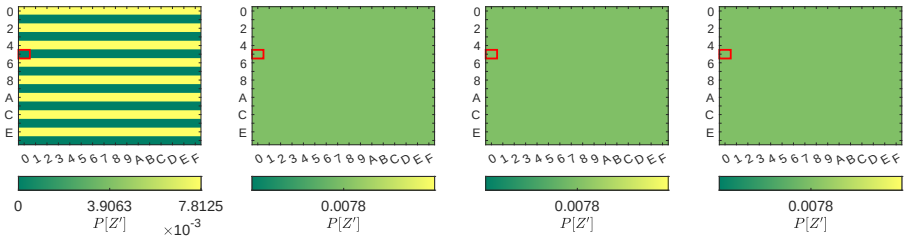
3.2 Initial Observation and Example

To obtain an initial intuition for the behavior of IP encoding under SIFA-1, we first examine a concrete fault scenario where the adversary injects stuck-at-0 faults into the LSB of every share $B[i]$ for $i \in [l]$, as also assumed in [29]. Simply we consider $l = 2$ and thus we have $\mathcal{F} = \{1, 9\}$ as shown in Figure 2a. For two-share IPM, 4 different IPM parameters are considered, including $L[2] \in \{1, 247, 245, 90\}$. Recall that the IPM instance with $L[2] = 1$ is exactly Boolean masking. Figure 3 shows the resulting probability distributions of $Z^{'} \in \mathbb{F}_{2^8}$ when the fault is ineffective, where the value x (resp., y) on X-axis (resp., Y-axis) represents the nibble value of the 4 most (resp., least) significant bits, and the color intensity corresponds to the probability $Pr[Z^{'} = 16x + y]$. As shown in Figure 3, the distribution under BM is biased (as also observed in [29, Figure 1]), whereas the distributions corresponding to other selected IPM parameters remain uniform. Moreover, we observe that this uniform distribution occurs not only in the selected instances but extends to all other two-share IPM instances where $L[2] \in \mathbb{F}_{2^8} \setminus \{0, 1\}$.



(a) $\mathcal{F} = \{1, 9\}$ (b) $\mathcal{F} = \{2, 9\}$ (c) $\mathcal{F} = \{3, 9\}$ (d) $\mathcal{F} = \{4, 9\}$

Fig. 2: Schematic diagram illustrating bit-wise fault injection under the specified fault model $\mathcal{F}$, with fault location denoted by a red lightning symbol.



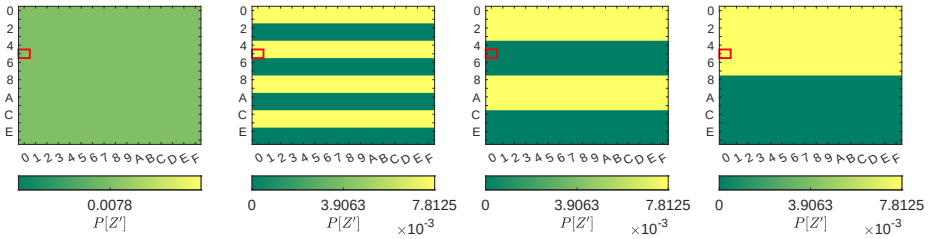
(a) $\vec{L} = [1, 1]$ (b) $\vec{L} = [1, 247]$ (c) $\vec{L} = [1, 245]$ (d) $\vec{L} = [1, 90]$

Fig. 3: Probability distribution of $Z^{'}$ over $\mathbb{F}_{2^8}$ when given knowledge of outputs obtained for ineffective faults after biasing LSB of each share with stuck-at-0 faults. The red box highlights $Pr[Z^{'} = \texttt{0x05}]$ as an example. As is indicated, $Pr[Z^{'} = \texttt{0x05}] = 0$ for BM and $Pr[Z^{'} = \texttt{0x05}] = 0.0078$ for other cases.

This observation indicates that two-share IPM (excluding BM case) thwarts SIFA-1 under the fault model of stuck-at-0 injections biasing LSB of each share (while this LSB model is valid to compromise two-share prime masking [29]). Naturally a critical question emerges: does IP encoding maintain its resilience against SIFA-1 when the fault model is changed, for instance, when varying fault locations or increasing the number of corrupted bits? Here we first present illus-

trative examples to provide preliminary insights, followed by a comprehensive theoretical analysis in subsequent sections that delivers systematic answers.

Now we consider distinct fault models for different IPM instances. Specifically, for the cases where $L[2] \in \{1, 247\}$, we employ $\mathcal{F} = \{2, 9\}$ which corresponds to injecting stuck-at-0 faults at the second least significant bit of $B[1]$ and the LSB of $B[2]$, as shown in Figure 2b; when $L[2] = 245$ and $L[2] = 90$, we utilize $\mathcal{F} = \{3, 9\}$ and $\mathcal{F} = \{4, 9\}$, as depicted in Figure 2c and 2d, respectively, applying similar fault injection principles. Figure 4 presents the probability distribution of $Z^{'} \in \mathbb{F}_{2^8}$ with ineffective faults under various IPM parameters and different fault models. As evidenced in Figure 4, under our specifically constructed fault models, all selected IPM instances except BM exhibit biased distributions. Through enumerate testing, we find that the specifically constructed fault model selectively biases only the target IPM parameters vector. For instance, the fault model $\mathcal{F} = \{2, 9\}$ proves effective exclusively for $\vec{L} = [1, 247]$, while Figure 4a confirms that BM maintains a uniform distribution under this fault model.



(a) $\vec{L} = [1, 1]$ (b) $\vec{L} = [1, 247]$ (c) $\vec{L} = [1, 245]$ (d) $\vec{L} = [1, 90]$

Fig. 4: Probability distribution of $Z^{'}$ over $\mathbb{F}_{2^8}$ with ineffective faults under different fault models $\mathcal{F}$ using stuck-at-0 faults. The red box highlights $Pr[Z^{'} = \texttt{0x05}]$ as an example. Specifically, $Pr[Z^{'} = \texttt{0x05}] = 0.0078$ for BM, for $\vec{L} = [1, 247]$ and $\vec{L} = [1, 90]$, we have $Pr[Z^{'} = \texttt{0x05}] = 7.8125 \times 10^{-3}$, while for $\vec{L} = [1, 245]$, there is $Pr[Z^{'} = \texttt{0x05}] = 0$.

Combining with Figure 3 and 4, our initial findings indicate that the effectiveness of IPM's defeating capability against SIFA-1 under stuck-at faults is strongly dependent on both its IPM parameters and the specific fault model employed. Although our preliminary experimental results show that corrupting two specific bits can successfully bias the distribution of $Z^{'}$ (thus compromise IPM) under controlled conditions, we will formally demonstrate in the following sections that this (simplest) case does not generalize. Moreover, even for those several selected IPM instances, the adversaries must achieve two-bit fault coverage with precise location (as shown in Figure 2) to effectively attack IPM.

3.3 Stuck-at Faults

As observed in Section 3.2, the resistance of IP encoding against SIFA-1 under stuck-at-0 faults exhibits strong dependency on both IPM parameters and the fault model $\mathcal{F}$. This observation naturally leads to two fundamental questions:

1) how do variations in fault models $\mathcal{F}$ and IPM parameters collectively determine the mitigation capability of IP encoding against SIFA-1 under stuck-at faults; 2) does the optimal IPM parameters exist to maximize protection against SIFA-1, and if so, how can it be systematically determined? In this subsection, we address both questions by developing a theoretical framework to characterize IPM resistance relative to both IPM parameters and fault models $\mathcal{F}$, while simultaneously identifying the optimal parameters.

Theoretical Framework to Characterize IPM Resistance. Recall that $Z = \vec{L} \cdot \vec{B}$ is the vulnerable variable over $\mathbb{F}_{2^n}$, where $\vec{L}$ is the IPM parameters vector and $\vec{B}$ (with $\vec{B}'$ its faulty counterpart) represents the shares vector. Let $\vec{e} = \vec{B} \oplus \vec{B}'$, effective SIFA-1 requires the condition $0 < Pr[Z = Z'] < 1$ such that $0 < Pr[\vec{L} \cdot \vec{e} = 0] < 1$ according to Corollary 2. Under both stuck-at-0 and stuck-at-1 faults, the addition vector $\vec{e}$ satisfies $[\vec{e}]_2[i] = 0$ if $i \notin \mathcal{F}$ and otherwise $[\vec{e}]_2[i]$ uniformly distributed over $\mathbb{F}_2$. Consequently, the probability satisfies $0 < Pr[\vec{L} \cdot \vec{e} = 0] < 1$. This holds since:

- when $\vec{e} = \vec{0}$, we trivially have $\vec{L} \cdot \vec{e} = 0$, yielding $Pr[\vec{L} \cdot \vec{e} = 0] \neq 0$;
- there exists cases where only one index $i \in \mathcal{F}$ satisfies $[\vec{e}]_2[i] = 1$, while all other element $[\vec{e}]_2[j] = 0$ if $j \neq i$. Under this condition, we have $\vec{L} \cdot \vec{e} = L[\lfloor i/n \rfloor + 1] \cdot e[\lfloor i/n \rfloor + 1] \neq 0$ (recall that $L[k] > 0$ for all $k \in [l]$), where $\lfloor \cdot \rfloor$ denotes the floor function (rounding down to the nearest integer). Hence, there exists cases when $\vec{L} \cdot \vec{e} > 0$, yielding $Pr[\vec{L} \cdot \vec{e} = 0] \neq 1$.

The probability $Pr[\vec{L} \cdot \vec{e} = 0]$ critically determines the adversary's ability to obtain valid outputs $\vec{C}'$, thereby relating to the attack complexity, which will be quantified in the subsequent analysis (please refer to Theorem 4). We now continue to rigorously examine the necessary conditions for SIFA-1 effectiveness established in Corollary 2. Given that $0 < Pr[\vec{L} \cdot \vec{e} = 0] < 1$, the resistance of IPM depends decisively on the distribution of faulty variable Z' with ineffective faults. We now characterize its distribution through the following theorem.

Theorem 1. *Let $Z = \vec{L} \cdot \vec{B}$ over $\mathbb{F}_{2^n}$ with $\vec{B} \in \mathbb{F}_{2^n}^l$ be the vulnerable variable and $\mathcal{F} \subseteq [nl]$ denote the stuck-at fault locations, where bit-wise faults are directly injected into the shares $[\vec{B}]_2$. Then, if $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = n$ holds, the faulty variable Z' with ineffective faults (satisfying $[Z = Z']$) is uniformly distributed over $\mathbb{F}_{2^n}$, where $\mathbf{L}$ is the code expansion of $\vec{L}$.*

Please refer to Appendix A for detailed proof. Based on Corollary 2 and Theorem 1, we establish the theoretical relationship between the fault model $\mathcal{F}$ and IPM parameters $\vec{L}$ in determining the resistance of IP encoding against SIFA-1 under stuck-at faults. Specifically, the rank of matrix $\mathbf{L}[*, \overline{\mathcal{F}}]$, deriving from both fault model $\mathcal{F}$ and IPM parameters $\vec{L}$, determines whether the corresponding IPM instance can resist SIFA-1. If $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = n$ holds, then such a IPM instance is capable of defeating SIFA-1 under this fault model; otherwise

it cannot. This allows us to explain the distinct distributions observed in Section 3.2. Table 1 records the rank of $\mathbf{L}[*, \overline{\mathcal{F}}]$ corresponding to the specific IPM parameters $\vec{L}$ with $|\vec{L}| = 2$ and fault model $\mathcal{F}$. It is revealed in Table 1 that all cases exhibiting non-uniform distributions in Section 3.2 correspond to matrices $\mathbf{L}[*, \overline{\mathcal{F}}]$ with its rank less than n, whereas uniform distributions emerge when $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = n$, where $n = 8$.

Table 1: Rank of $\mathbf{L}[*, \overline{\mathcal{F}}]$ under various fault models and IPM parameters

$\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$	$\mathcal{F} = \{1, 9\}$	$\mathcal{F} = \{2, 9\}$	$\mathcal{F} = \{3, 9\}$	$\mathcal{F} = \{4, 9\}$
$\vec{L} = [1, 1]$	7	8	8	8
$\vec{L} = [1, 247]$	8	7	8	8
$\vec{L} = [1, 245]$	8	8	7	8
$\vec{L} = [1, 90]$	8	8	8	7
other $\vec{L}$	8	8	8	8

According to Theorem 1, the following lemmas naturally derive from an adversarial perspective.

Lemma 1. *Let $Z = \vec{L} \cdot \vec{B}$ over $\mathbb{F}_{2^n}$ with $\vec{B} \in \mathbb{F}_{2^n}^l$ be the vulnerable variable. A successful SIFA-1 with effective fault model $\mathcal{F}$ necessitates the corruption of at least one bit fault in each share $B[i]$ for $i \in [l]$.*

Proof. The code expansion $\mathbf{L_i}$ for each component $L[i]$ of $\vec{L}$ is full-rank such that $\mathsf{Rank}(\mathbf{L_i}) = n$ across $i \in [l]$. If not all the shares are corrupted, we assume that $B[i]$ is the fault-free share, then we have $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = \mathsf{Rank}(\mathbf{L_i}) = n$. This completes the proof.

Lemma 2. *Let $\mathcal{F} \subseteq [nl]$ be the set of bit-wise stuck-at fault locations and $\mathbf{L} \in \mathbb{F}_2^{n \times nl}$ the code expansion of IPM parameters $\vec{L}$, which is a generator matrix of a $[nl, n, d_{\mathcal{L}}]$-linear code. For any vector (codeword) $\vec{u} \in \mathcal{L}$, define its non-zero positions $\mathcal{I}_u := \{i | u[i] = 1\}$. Then, the fault model $\mathcal{F}$ is valid (enabling $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$) if and only if there exists a nonzero codeword $\vec{u} \in \mathcal{L} \setminus \{\vec{0}\}$ whose non-zero positions are entirely contained in $\mathcal{F}$ such that $\mathcal{F} \cap \mathcal{I}_u = \mathcal{I}_u$.*

Please refer to Appendix B for detailed proof. Lemma 2 further derives the characteristic of valid SIFA-1 fault models $\mathcal{F}$ that promisingly compromise specific IPM instances based on Theorem 1. Its security implications will be analyzed in the following sub-section. While Lemma 1 provides the theoretical foundation for prior empirical observation [29, 35] that successful SIFA-1 model requires stuck-at faults across all shares of Boolean masking. According to Lemma 1, the lower bound for successful SIFA-1 is formally established as l bit-wise faults (same as share count), while the upper bound for IPM resistance is $(l-1)n$ bit-wise faults. This establishes the security transition phase between IP encoding and SIFA-1: IPM resist SIFA-1 below l bit-wise faults but become vulnerable beyond $(l-1)n$ bit-wise faults. Subsequently, we present Theorem 2 to demonstrate the tightness of the bounds for security transition phase.

Theorem 2 (Tightness of Security Transition Bounds). *Let $Z = \vec{L} \cdot \vec{B}$ over $\mathbb{F}_{2^n}$ with $\vec{B} \in \mathbb{F}_{2^n}^l$ be the vulnerable variable and $\mathcal{F} \subseteq [nl]$ denote the stuck-at fault locations, where bit-wise faults are directly injected into the shares $[\vec{B}]_2$. During the security transition phase that $l \leq |\mathcal{F}| \leq (l-1)n$, we have*

C1: Under any given fault model $\mathcal{F}$, there exists IPM parameters $\vec{L}$ with $|\vec{L}| = l$ and $L[1] = 1$ enabling $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = n$;

C2: Under any given IPM parameters vector $\vec{L}$ with $|\vec{L}| = l$, there exists at least one fault model $\mathcal{F}$ enabling $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$.

Please refer to Appendix C for detailed proof. As a simple empirical verification for $C1$ of Theorem 2, we enumerate all possible fault models $\mathcal{F}$ for the case where $l = 2$ and $n = 8$. For each candidate $\mathcal{F}$ satisfying $|\mathcal{F}| \leq (l-1)n$, we confirm the existence of an IPM parameters vector $\vec{L}$ with $|\vec{L}| = 2$ and $L[1] = 1$ enabling $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = n$. While for the verification of $C2$, an empirical proof will be provided in the subsequent analysis.

As established in Theorem 1 and 2, during security transition phase $[l, (l-1)n]$, an effective SIFA-1 attack requires a fault model $\mathcal{F}$ enabling $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$ to compromise a specific IPM instance with $\vec{L}$. This implies that a SIFA-1 adversary must achieve both $|\mathcal{F}|$-bit fault injection coverage and precise bit-location control (e.g., Figure 2d) to compromise IPM. We define the term *attack setup complexity* to capture the adversary's burden in constructing an effective attack environment, encompassing both the required number of bit-wise corruptions and the precision of bit-location control needed for effective SIFA-1 fault models. This term specifically distinguishes the concept from that of *attack complexity*, which is *relevant* only when an attack can be effective. As analyzed in Section 3.1, the *attack complexity* of SIFA-1 relates to the *uniformity degree* of faulty variable's distribution with ineffective faults and the probability of obtaining valid outputs by adversaries. Subsequently, we present Theorem 3 and 4, deriving two *matrix-rank parameters*, to characterize the SIFA-1 attack complexity under valid fault models.

Theorem 3 (Uniformity Degree). *Let $Z' = \vec{L} \cdot \vec{B}'$ be the vulnerable variable under a stuck-at fault model $\mathcal{F} \subseteq [nl]$ that effectively compromises IPM with the parameters vector $\vec{L}$. Let $\triangle_u(Z')$ denotes the uniformity degree of the distribution of Z' with ineffective faults, then we have*

$$\triangle_u(Z') \propto \mathcal{R}(\mathbf{L}[*, \overline{\mathcal{F}}]). \tag{6}$$

Please refer to Appendix D for detailed proof. We characterize the *uniformity degree* qualitatively through $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ rather than quantitatively, as a precise metric would require explicit modeling of the distinguisher's capabilities, that is beyond the scope of this work (left as an interesting future work). Note that the extreme case [8] occurs when $|\mathcal{F}| = nl$ (all share bits are corrupted), for

[8] We exclude the other boundary case where $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = n$ since this indicates an ineffective attack, making the analysis of *attack complexity* inapplicable.

which we define $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 0$. Now we quantify the probability of obtaining valid outputs by adversaries through Theorem 4 below.

Theorem 4 (Probability of Obtaining Valid Outputs). *Assume that $\mathcal{F} \subseteq [nl]$ is an effective fault model under stuck-at faults for compromising IPM with $\vec{L}$. Then, The probability of obtaining valid outputs can be quantified as:*

$$Pr[C = C^{'}] = \frac{1}{2^{\mathcal{R}(\mathbf{L}[*,\mathcal{F}])}}. \tag{7}$$

Please refer to Appendix E for detailed proof. Since the matrix $\mathbf{L}[*, \mathcal{F}]$ has size $n \times |\mathcal{F}|$, its maximum possible rank is the smaller of n and $|\mathcal{F}|$. Hence the probability of obtaining valid outputs is lower bounded by $\frac{1}{2^n}$. As established in Theorem 3 and 4, the attack complexity of SIFA-1 is jointly influenced by two rank parameters $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ and $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$, which characterize the distribution uniformity and valid output probability, respectively. However, as the fault set size $|\mathcal{F}|$ increases, $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ decreases monotonically, reducing the uniformity degree and thus lowering attack complexity; while $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ increases correspondingly, decreasing the probability of valid outputs and thus increasing attack difficulty. Consequently, these two rank parameters exhibit countervailing effects on *attack complexity*. However, we should underline that increasing the size of fault model $\mathcal{F}$ itself imposes significant practical challenges for adversaries. For simplicity, we define $\mathcal{R}_{\overline{\mathcal{F}}} := \mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ and $\mathcal{R}_{\mathcal{F}} := \mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ throughout this paper.

How to Select Optimal IPM Parameters. Now we aim to identify the *optimal* IPM parameters during the security transition phase $[l, (n-1)l]$ (all IPM instances exhibit identical resistance effectiveness beyond this phase).

Let us consider a practical attack scenario. Usually, cryptography has been already protected by IPM with fixed and specific $\vec{L}$, which is public to adversaries. Hence, the adversary can tailor attack strategy accordingly. In practice, a valid fault model $\mathcal{F}$ forces adversaries to simultaneously corrupt (multiple) bits across all shares (recall Lemma 1). As the number of precise faulty bits increases, the *attack setup complexity* significantly increases as such complex faults are possible with use of advanced injection equipments like electromagnetic glitches, lasers, etc., which incur high equipment cost and specialized attacker's expertise. Given these practical constraints, we reasonably assume that SIFA-1 adversaries prioritize minimizing the *attack setup complexity* before optimizing the *attack complexity*, that is, minimizing fault injections $|\mathcal{F}|$ while achieving $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$. Therefore, *optimal* IPM parameters should first maximize the *attack setup complexity* of SIFA-1 by

- maximize the minimum fault injection cost $|\mathcal{F}|$ to achieve $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$;
- minimize the number of valid fault models $\mathcal{F}$ at this minimal setup cost.

Note that fewer valid models $\mathcal{F}$ at minimum bit-injection cost implies stricter fault precision requirements and thus increasing *attack setup complexity*. In the

following, we propose Theorem 5 to quantify the minimum amount of faulty bits required to satisfy $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$ and the number of effective fault models at this minimal bit-injection cost, given specific IPM parameters vector $\vec{L}$.

Theorem 5. *Let $\mathcal{F} \subseteq [nl]$ be the set of bit-wise fault locations and $\mathbf{L} \in \mathbb{F}_2^{n \times nl}$ the code expansion of IPM parameters $\vec{L}$, which is a generator matrix of an $[nl, n, d_{\mathcal{L}}]$-linear code. Then, the minimum size $|\mathcal{F}|$ to enable $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$ is the distance $d_{\mathcal{L}}$ of linear code $\mathcal{L}$; and the number of effective fault models $\mathcal{F}$ at minimum size $|\mathcal{F}| = d_{\mathcal{L}}$ equals the kissing number $B_{d_{\mathcal{L}}}$ of linear code $\mathcal{L}$.*

Proof. Directly from Lemma 2, the minimum size of $\mathcal{F}$ is

$$|\mathcal{F}|_{\min} := \min(|\mathcal{F}|) = \min_{\vec{u} \in \mathcal{L} \setminus \{\vec{0}\}} (|\mathcal{I}_u|) = d_{\mathcal{L}}, \tag{8}$$

where $\mathcal{I}_u$ denote non-zero positions of the vector $\vec{u} \in \mathcal{L} \setminus \{\vec{0}\}$, such that $\mathcal{I}_u := \{i | u[i] = 1\}$. Accordingly by Definition 3 and Lemma 2, the number of valid fault models $\mathcal{F}$ at size $|\mathcal{F}| = d_{\mathcal{L}}$ is the kissing number $B_{d_{\mathcal{L}}}$. This completes the proof.

Interestingly, Theorem 5 and Corollary 1 establish a theoretical connection between side-channel and stuck-at-fault SIFA-1 resistance of IP encoding:

- the *distance* $d_{\mathcal{L}}$ (equivalently $d_{\mathcal{D}^\perp}$) determines both the minimum number of bits that must be probed and corrupted, to launch an effective higher-order side-channel (excluding rank-based [43]) attack and SIFA-1 attack under stuck-at faults, respectively;
- the *kiss number* $B_{d_{\mathcal{L}}}$ (equivalently $B_{d_{\mathcal{D}^\perp}}$) demonstrates the redundancy rate for an effective attack (including both SCA and stuck-at-fault SIFA-1) at minimal bit cost (whether through probing or corrupting).

In other words, the coding-theoretic parameters $d_{\mathcal{L}}$ and $B_{d_{\mathcal{L}}}$ serve as joint metrics both for the side-channel mitigation efficacy and SIFA-1 resistance under stuck-at faults. Furthermore, it suggests that a set of IPM parameters can be identified accordingly which simultaneously optimize side-channel resistance and maximize the *attack setup complexity* of stuck-at-fault SIFA-1, providing dual protection against both passive and active adversaries.

After constructing valid fault models $\mathcal{F}$ at lowest *attack setup complexity*, an adversary should minimize both $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ and $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ to optimally reduce *attack complexity*. However, there is an inherent trade-off between these rank parameters: increasing $|\mathcal{F}|$ monotonically decreases $\mathcal{R}_{\overline{\mathcal{F}}}$ (reducing attack complexity) while increasing $\mathcal{R}_{\mathcal{F}}$ (amplifying complexity). Notably, we observe that $\mathcal{R}_{\mathcal{F}}$ exhibits greater sensitivity to increasing $|\mathcal{F}|$ than $\mathcal{R}_{\overline{\mathcal{F}}}$, owing to the full-rank structure of each $n \times n$ block $\mathbf{L_i}$ in matrix $\mathbf{L}$. For instance, considering the simplest l-share Boolean case: reducing $\mathcal{R}_{\overline{\mathcal{F}}}$ from r to $r-1$ requires faulting at least l bits (e.g., the same i-th bit of all shares), whereas increasing $\mathcal{R}_{\mathcal{F}}$ from r to $r+1$ requires only a single additional bit to $\mathcal{F}$ (when $|\mathcal{F}|$ is minimized). In addition, we empirically test the combined effect of these two rank parameters (under conditions of equal sensitivity) on attack complexity through experiments (see Appendix F). It is shown that, with decreasing $\mathcal{R}_{\overline{\mathcal{F}}}$ and increasing $\mathcal{R}_{\mathcal{F}}$, the

net effect shows an overall increase in attack complexity, implying that $\mathcal{R}_{\mathcal{F}}$ emerges as the dominant parameter governing attack complexity, outweighing the influence of $\mathcal{R}_{\overline{\mathcal{F}}}$[9]. Building on those observations (differential sensitivity and distinct weight effects), SIFA-1 adversaries would prefer to focus more on keeping $\mathcal{R}_{\mathcal{F}}$ as low as possible. This assumption has another more important advantage: it naturally aligns with minimizing fault set size $|\mathcal{F}|$. Recall that reducing $\mathcal{R}_{\overline{\mathcal{F}}}$ requires increasing $|\mathcal{F}|$. While lowering $\mathcal{R}_{\mathcal{F}}$ moves in tandem with reducing $|\mathcal{F}|$, both objectives push toward smaller fault injections.

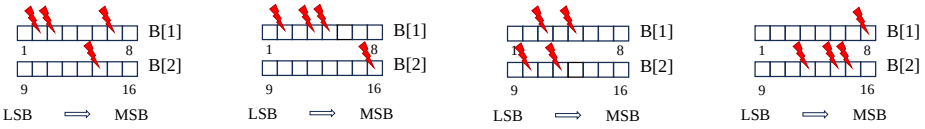
Therefore, the *optimal* IPM parameters are defined through a dual optimization criterion with priority: first maximizing *attack setup complexity* by maximizing the minimum fault injection cost $|\mathcal{F}|$ required for $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$ while minimizing viable fault models at this minimum cost; second at this lowest *attack setup complexity*, maximizing $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ to amplify *attack complexity*. We use the symbol $\mathcal{R}_{\min}$ to denote the minimum $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ related to the specific IPM parameters $\vec{L}$ and the valid fault models with $|\mathcal{F}| = d_{\mathcal{L}}$. Thus, *optimal* IPM parameters should maximize $\mathcal{R}_{\min}$. As established in Theorem 5, the *attack setup complexity* of specific IPM parameters can be quantified through two coding-theoretical parameters $d_{\mathcal{L}}$ and $B_{d_{\mathcal{L}}}$. While regarding *attack complexity*, Lemma 2 identifies the valid fault models $\mathcal{F} = \mathcal{I}_{u_d}$ with minimal size $|\mathcal{F}| = d_{\mathcal{L}}$, where $\mathcal{I}_{u_d}$ represents the non-zero positions of non-zero vector $\vec{u} \in \mathcal{L}$ with minimal Hamming weight $d_{\mathcal{L}}$, hence $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ and thus $\mathcal{R}_{\min}$ can be derived directly with no need to exhaustive search.

Let consider two-share IPM over $\mathbb{F}_{2^s}$ as a concrete example. We report the resistance ability (indicated by coding-theoretic parameters and $\mathcal{R}_{\min}$) of various two-share IPM parameters in Table 4 of Appendix G. This table is generated through exhaustive search rather than derived directly from our coding-theoretic parameters. As expected, the candidates determined by $d_{\mathcal{L}}$ and $B_{d_{\mathcal{L}}}$ perfectly match those identified by the same $d_{\mathcal{D}}^{\perp}$ and $B_{d_{\mathcal{D}}^{\perp}}$ in Table II of [12]. Furthermore, we perform an identical exhaustive search for three-share IPM and confirm that the results align with those reported in [9] as well. These results provide empirical validation for both Theorem 5 and Corollary 1. In addition, our exhaustive-search experiments confirm that all two-share (resp., three-share) IPM instances are vulnerable to valid fault models with size satisfying $|\mathcal{F}| \leq (l-1)n = 8$ (resp., $|\mathcal{F}| \leq (l-1)n = 16$), providing empirical validation for $C2$ of Theorem 2.

Verified by both theoretical parameters and empirical enumerating, the optimal two-share IPM parameters have 12 candidates with $d_{\mathcal{L}} = 4$ and $\mathcal{R}_{\min} = 4$ (shown in Table 4 in Appendix G). Hence, to compromise optimal two-share IPM instances, an adversary must corrupt 4 bits simultaneously and repeatedly across two shares, with attack complexity increasing to $\frac{1}{2^4}$ (with the maximum *uniformity degree* $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 7 = n - 1$ that is confirmed). While as shown in [29], two-share prime masking only requires 2 bit-injection cost with a valid output probability of roughly $\frac{1}{2^2}$, demonstrating the superior resistance of IP encoding against stuck-at-fault SIFA-1. Moreover, only $B_{d_{\mathcal{L}}} = 4$ combinations of

[9] The influence of $\mathsf{Rank}([*, \overline{\mathcal{F}}])$ on attack complexity depends also on the capabilities of distinguishing algorithms, we limit this conclusion within our attack setup.

bit-fault locations are effective for two-share optimal IPM instances at the minimum bit-injection cost. Considering a specific IPM instance with $\vec{L} = [1, 23]$ as an example, the effective fault models contain $\mathcal{F}_1 = \{2, 3, 7, 14\}$, $\mathcal{F}_2 = \{2, 4, 5, 16\}$, $\mathcal{F}_3 = \{3, 5, 10, 12\}$, $\mathcal{F}_4 = \{8, 12, 14, 15\}$, with the fault-injection schematic diagram illustrated in Figure 5. To the best of our knowledge, such precise multi-bit faults across shares have not been reported in public literature related to fault injection techniques. We should underline that only those 4 fault models (out of $C_{16}^4 = 1820$ possible combinations) depicted in Figure 5 are effective against IPM instance with $\vec{L} = [1, 23]$ at 4-bit injection cost. Our exhaustive simulations confirm that attacks with fewer than 4-bit corruptions invariably fail. While, increasing the number of corrupted bits beyond 4 will expand the set of effective fault models. However, this comes at the cost of increased *attack setup complexity* and elevated risk of higher attack complexity.



(a) $\{2, 3, 7, 14\}$ (b) $\{2, 4, 5, 16\}$ (c) $\{3, 5, 10, 12\}$ (d) $\{8, 12, 14, 15\}$

Fig. 5: Schematic diagram illustrating bit-wise corruption under effective fault models $\mathcal{F}$ for $\vec{L} = [1, 23]$, with fault location denoted by a red lightning symbol.

Regarding three-share IPM instances over $\mathbb{F}_{2^8}$, both exhaustive search experiments and theoretical coding parameters identify the following optimal IPM parameter sets as in [12]: $\vec{L} = [1, 240, 218]$, $\vec{L} = [1, 218, 240]$, $\vec{L} = [1, 91, 204]$, $\vec{L} = [1, 204, 91]$, $\vec{L} = [1, 27, 196]$, $\vec{L} = [1, 196, 27]$, requiring a minimum $d_{\mathcal{L}} = 8$ bits for effective SIFA-1 attacks, with $B_{d_{\mathcal{L}}} = 6$ effective fault models achievable at this minimum bit-injection cost. Furthermore, they maintain a low probability of obtaining valid outputs at $\frac{1}{2^{\mathcal{R}_{\min}}} = \frac{1}{2^7}$. While for Boolean and prime masking, 3 bit-wise injections suffice for effective attacks with valid outputs probability given by $\frac{1}{2^3}$. We should highlight that for IPM, both of the minimum number of bit-wise faults $|\mathcal{F}|$ required for effective SIFA-1 and the attack complexity (the *uniformity degree* $\mathcal{R}_{\overline{\mathcal{F}}} = 7$ maintain in both two-share and three-share optimal cases) exhibit a non-linear (super-linear) growth with increasing share count (e.g., from 4 bits for 2 shares to 8 bits for 3 shares; from 2^{-4} to 2^{-7}), significantly increasing both *attack setup complexity* and *attack complexity* at the cost of merely adding one share. While for both Boolean and prime (with small share counts) masking, the *attack setup complexity* grows linearly with the number of shares. This demonstrates the strong protection against SIFA-1 offered by IP encoding. While valid multi-bit (across all shares) fault models already present substantial challenges for SIFA-1 adversaries (Figure 5), it is crucial to note that this complexity assumes perfect 100% bit-wise fault injection precision per execution. In practical scenarios, lower precision rates would further increase the number of fault attempts required for successful SIFA-1 attacks. Last but not the

least, those there-share (resp., two-share) optimal parameter sets also provide optimal SCA protection, achieving bit-probing security of $7 = d_{\mathcal{D}^\perp} - 1 = d_{\mathcal{L}} - 1$ (resp., 3) with kissing number as $B_{d^\perp_{\mathcal{D}}} = B_{d_{\mathcal{L}}} = 6$ (resp., $B_{d_{\mathcal{L}}} = 4$).

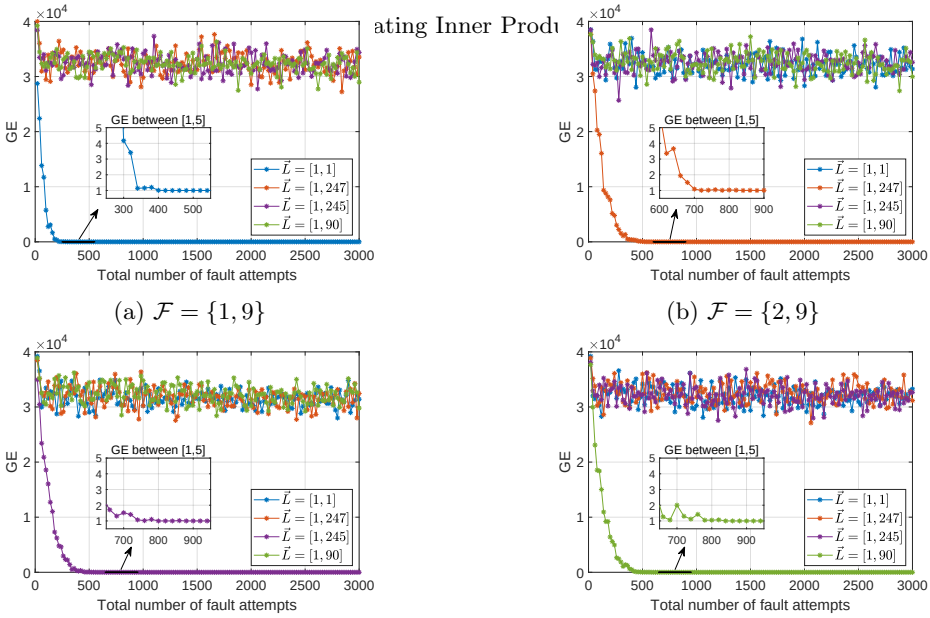
Simulated Attacks. In this sub-section, we experimentally validate our developed theoretical framework through comprehensive SIFA-1 simulations under stuck-at faults. Limited by paper size, please refer to Appendix H for simulations of optimal two-share IPM parameters.

As demonstrated in prior analysis, stuck-at-0 and stuck-at-1 faults exhibit identical attack capabilities. Hence, our simulations focus exclusively on stuck-at-0 faults. We evaluate the resistance of IP encoding using a minimal toy circuit (depicted in Figure 1). Without loss of generality, our implementation employs a two-share IPM scheme, leveraging the provably secure gadgets from [2]. The attack operates by performing multiple fault injection attempts on all shares of $P[0]$ for randomly chosen input data, while collecting outputs from ineffective fault injections (when $[\vec{C} = \vec{C}']$). Although not explicitly depicted in Figure 1, the shared representations of $C[0]$ and $C[1]$ are decoded (unmasked) before being released as outputs. For each valid output, the adversary performs reverse computation through the circuit to the injection point by enumerating key hypotheses for $k[0]$ and $k[1]$, then tests the distribution uniformity for each key candidate by calculating Squared Euclidean Imbalance (SEI) [18], where the correct key is identified as the hypothesis yielding the maximum SEI value among all possible candidates. In our simulated evaluation, Guessing Entropy (GE) [39] is employed to quantify and demonstrate the attack complexity.

Table 2: Rank of both $\mathbf{L}[*, \overline{\mathcal{F}}]$ and $\mathbf{L}[*, \mathcal{F}]$ under specific IPM parameters $\vec{L}$ and fault model $\mathcal{F}$. $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 8$ indicates attack failure (Theorem 1), rendering attack complexity (based on $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$) unnecessary, thereby marked "-".

$\mathcal{F}$	$\vec{L} = [1, 1]$		$\vec{L} = [1, 247]$		$\vec{L} = [1, 245]$		$\vec{L} = [1, 90]$	
	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$
$\mathcal{F} = \{1, 9\}$	7	1	8	-	8	-	8	-
$\mathcal{F} = \{2, 9\}$	8	-	7	2	8	-	8	-
$\mathcal{F} = \{3, 9\}$	8	-	8	-	7	2	8	-
$\mathcal{F} = \{4, 9\}$	8	-	8	-	8	-	7	2

We begin with the attack validation of the initial examples presented in Section 3.2. Table 2 records the involving four IPM instances including $L[2] = \{1, 247, 245, 90\}$, four fault models $\mathcal{F} = \{i, 9\}$ for $i \in [4]$ (Figure 2), and the corresponding matrix-rank parameters. Recall that in Section 3.2, we initially observed that (shown in Figure 3 and 4), the fault model $\mathcal{F} = \{i, 9\}$ can exclusively bias the distribution of Z' (with ineffective faults) protected by IPM parameters $\vec{L} = [1, \alpha[i]]$, respectively, where $\vec{\alpha} = [1, 247, 245, 90]$ and $i \in [4]$. Figure 6 illustrates the GE results of SIFA-1 targeting those four IPM instances under those four fault models. As shown in Figure 6a, 6b, 6c and 6d, the attack model $\mathcal{F} = \{i, 9\}$ only successfully compromises the IPM instance with $\vec{L} = [1, \alpha[i]]$ for $i \in [4]$, respectively, where $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 7 < n$ in each case. However,



(a) $\mathcal{F} = \{1, 9\}$ (b) $\mathcal{F} = \{2, 9\}$

(c) $\mathcal{F} = \{3, 9\}$ (d) $\mathcal{F} = \{4, 9\}$

Fig. 6: GE results of various IPM parameters under different fault models, with inset plots highlighting the convergence behavior as GE approaches 1.

for other combinations of fault models $\mathcal{F}$ and IPM parameters $\vec{L}$, we observe $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 8 = n$, which leads to unsuccessful attacks. These experimental results provide strong empirical validation for Corollary 2 and Theorem 1.

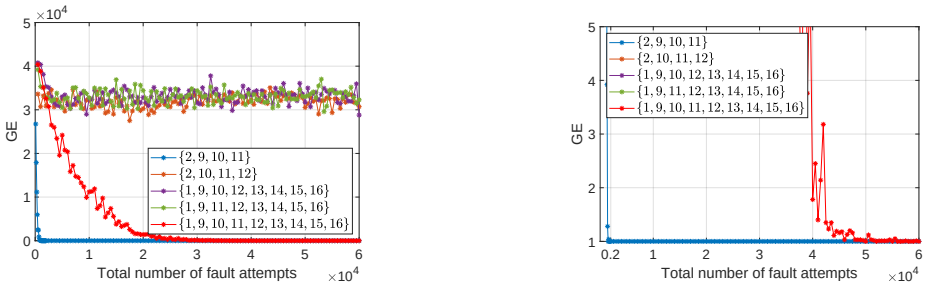
Moreover, a closer examination of the inset plots in Figure 6 reveals distinct attack complexity (expressed by amount of fault attempts required for decreasing GE to 1). For $\mathcal{F} = \{1, 9\}$ attacking Boolean instance (with $\vec{L} = [1, 1]$), approximately 400 fault attempts suffice. While attacks using $\mathcal{F} = \{1, 247\}$, $\mathcal{F} = \{1, 245\}$ and $\mathcal{F} = \{1, 90\}$ against the respective IPM instances $\vec{L} = [1, 247]$, $\vec{L} = [1, 245]$ and $\vec{L} = [1, 90]$ require roughly 800 fault attempts, which is twice the complexity of the Boolean case. This difference is predicted by the indicator $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ (note that all instances share identical *uniformity degree* evidenced by $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 7$ in Table 2). As Table 2 shows, $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}]) = 1$ for Boolean case under $\mathcal{F} = \{1, 9\}$, corresponding to a theoretical attack complexity of $\frac{1}{2}$. For other effective fault models targeting the respective IPM instances, $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}]) = 2$ yields an attack complexity of $\frac{1}{4}$, precisely half that of the Boolean case, providing empirical validation for Theorem 4.

To further validate the indicating capability of two matrix-rank parameters for attack effectiveness and complexity of SIFA-1 under stuck-at faults, we evaluate the IPM instance $\vec{L} = [1, 247]$ under additional fault models. Table 3 reports the rank of both $\mathbf{L}[*, \overline{\mathcal{F}}]$ and $\mathbf{L}[*, \mathcal{F}]$ for IPM instance with $\vec{L} = [1, 247]$ under various fault models. Figure 7 presents the corresponding GE results attacking this IPM instance. The experimental results reconfirm that successful SIFA-1 compromising IPM instance occurs exclusively when the rank condition

$\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$ is satisfied. Specifically, the fault model $\mathcal{F} = \{2, 9, 10, 11\}$ subsumes the effective 2-bit fault model $\mathcal{F} = \{2, 9\}$ (recall Table 2). This inclusion preserves the critical rank condition $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 7 < n$, thereby maintaining the capability to successfully compromise $\vec{L} = [1, 247]$. However, since $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}]) = 3$ for this extended fault model, the attack complexity increases to $\frac{1}{2^3}$, two times higher than the $\frac{1}{2^2}$ complexity of $\mathcal{F} = \{2, 9\}$. The experimental result aligns with the theoretical prediction, as evidenced in Figure 7b, where successful attack requires approximately 1600 fault attempts, matching the expected complexity derived from $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ (note that both fault models share identical *uniformity degree* indicated by $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 7$). For other alternative 4-bit fault models such as $\mathcal{F} = \{2, 10, 11, 12\}$, despite increasing bit-wise fault injections (4-bit vs. 2-bit for $\mathcal{F} = \{2, 9\}$), the failure to satisfy $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$ prevents any GE reduction. Figure 7 empirically confirms that $\mathcal{F} = \{2, 10, 11, 12\}$ results in ineffective attacks.

Table 3: Rank of both $\mathbf{L}[*, \overline{\mathcal{F}}]$ and $\mathbf{L}[*, \mathcal{F}]$ under $\vec{L} = [1, 247]$ and various fault models $\mathcal{F}$. $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 8$ indicates attack failure (Theorem 1), rendering attack complexity (based on $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$) unnecessary, thereby marked "-".

$\mathcal{F}$	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$
$\mathcal{F} = \{2, 9, 10, 11\}$	7	3
$\mathcal{F} = \{2, 10, 11, 12\}$	8	-
$\mathcal{F} = \{1, 9, 10, 12, 13, 14, 15, 16\}$	8	-
$\mathcal{F} = \{1, 9, 11, 12, 13, 14, 15, 15\}$	8	-
$\mathcal{F} = \{1, 9, 10, 11, 12, 13, 14, 15, 16\}$	7	8



(a) GE between $[1, 50000]$. (b) GE between $[1, 5]$.

Fig. 7: GE results of IPM instances with $\vec{L} = [1, 247]$ under different fault models.

Notably, even when scaling fault injection from the security transition's lower bound $l = 2$ to its upper bound $(l - 1)n = 8$, any fault model failing to satisfy $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$ remains ineffective, as shown in Figure 7, where both maximal 8-bit (during the security transition phase) fault models $\mathcal{F} = \{1, 9, 10, 12, 13, 14, 15, 16\}$ and $\mathcal{F} = \{1, 9, 10, 11, 13, 14, 15, 16\}$ fail com-

pletely, due to $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}]) = 8 = n$. This demonstrates that increased fault injection magnitude alone during security transition phase cannot compensate for unfavorable rank conditions. However, if the amount of fault injection exceeds the security transition upper bound $(l-1)n = 8$, SIFA-1 attacks are invariably successful, as validated in Figure 7, where the 9-bit fault model $\mathcal{F} = \{1, 9, 10, 11, 12, 13, 14, 15, 16\}$ achieves full key recovery. This fault model corresponds to corrupting the first bit of $B[1]$ and all bits of $B[2]$. Although the 9-bit model successfully compromises the IPM instance, Figure 7b shows it requires over $50,000$ fault attempts to reduce GE to 1. This aligns well with the theoretical complexity predicted by $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}]) = 8$ (with the *uniformity degree* remaining $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 7$), yielding attack complexity of $\frac{1}{2^8}$, 64 times higher than the optimal $\mathcal{F} = \{2, 9\}$ with complexity as $\frac{1}{2^2}$ (there is $800 \times 64 = 51200$).

This experiment further reveals the distinct sensitivity between $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ and $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ to increasing fault model size: when $|\mathcal{F}|$ increases from 2 to 4 (resp., 9), $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ remains constant while $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ increases from 2 to 3 (resp., 8). These findings substantiate our selection of $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$ rather than $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ as the primary determinant of attack complexity when identifying optimal IPM parameters. From an adversary perspective, $\mathcal{F} = \{2, 9\}$ proves more cost-efficient than other evaluated fault models since others require 4-bit and 9-bit (versus 2-bit) precise fault injections and higher attack complexity with 2^{-3} and 2^{-8} (versus 2^{-2}). Consequently, SIFA-1 adversaries with knowledge of the IPM parameters $\vec{L} = [1, 247]$ would rationally prefer $\mathcal{F} = \{2, 9\}$ as the better (optimal) attack strategy. This rationale motivates our selection of $d_{\mathcal{L}}$ (minimum fault model size) and $\mathcal{R}_{\min}$ as key metrics for determining optimal IPM parameters. Please refer to Appendix H for the simulated attacks on optimal two-share IPM instances.

3.4 Toggle/Bit-Flip Faults

Following the investigation of stuck-at faults, we now examine the resistance of IP encoding against SIFA-1 under toggle/bit-flip faults, another widely utilized fault type. Under the threat model assumptions established in Section 3.1, the fault model $\mathcal{F}$ induces bit-flips in the specified positions and thus the bits in $[\vec{B}]_2[\mathcal{F}]$ are flipped, thereby generating the faulty $Z^{'}$. Let us first analyze the probability $Pr[C = C^{'}] = Pr[Z = Z^{'}]$ for valid outputs with ineffective faults. For the equality $Z = Z^{'}$ to hold, the condition $\vec{L} \cdot \vec{B} = \vec{L} \cdot \vec{B}^{'}$ should be satisfied, which implies $\vec{L}(\vec{B} \oplus \vec{B}^{'}) = 0$. Assuming $\vec{e} = \vec{B} \oplus \vec{B}^{'}$, we have $[\vec{e}]_2[i] = 1$ if $i \in \mathcal{F}$ and $[\vec{e}]_2[i] = 0$ otherwise, with this property being independent of the value of $\vec{B}$. At this stage, the fault only yields two distinct cases:

- If $\vec{L} \cdot \vec{e} = 0$, then $Pr[Z = Z^{'}] = 1$ holds deterministically. Note that the value of $\vec{e}$ depends exclusively on the fault model, precisely the specific faults locations $\mathcal{F}$. According to Corollary 2, the attack necessarily fails;
- If $\vec{L} \cdot \vec{e} \neq 0$, then $Pr[Z = Z^{'}] = 0$ holds deterministically, preventing adversaries from obtaining outputs as $Pr[C = C^{'}] = 0$ and resulting in attack failure, as also established in the proof of Corollary 2.

In conclusion, IP encoding fundamentally thwarts SIFA-1 under toggle/bit-flip faults. This finding aligns with the established result in [29] claiming the ineffectiveness of toggle/bit-flip faults attacking BM.

3.5 Constant-Byte Faults

This subsection assesses IP encoding's resilience to SIFA-1 when subjected to constant-byte fault injections that force an arbitrary subset of the shares vector $\vec{B}$ to constant values per share (in that subset) per execution. Under this scenario, the fault model should accordingly be adjusted slightly: let $\mathcal{F} \subseteq \{1, \cdots, l\}$ denote the set of byte-wise fault locations. Note that Corollary 2 remains valid under this fault model. We assume that $\vec{r} \in \mathbb{F}_{2^n}^l$ is the faulty vector satisfying $r[i] = 0$ if $i \notin \mathcal{F}$ and otherwise $r[i]$ denotes the constant value that is injected at the i-th share. Under this scenario, the resulting faulty shares $B^{'}[i]$ are fixed for the entire attack for $i \in \mathcal{F}$. Hence, the addition vector $\vec{e} = \vec{B} \oplus \vec{B}^{'}$ satisfies $e[i] = 0$ if $i \notin \mathcal{F}$ and otherwise $e[i]$ uniformly distributed over $\mathbb{F}_{2^n}$. Consequently, the probability of obtaining ineffective faults is given by

$$Pr[Z = Z^{'}] = Pr[\vec{L} \cdot \vec{e} = 0] = \frac{(2^n)^{|\mathcal{F}|-1}}{(2^n)^{|\mathcal{F}|}} = \frac{1}{2^n}.$$

Now consider the distribution of $Z^{'} = \vec{L} \cdot \vec{B}^{'}$ with ineffective faults, where $\vec{B}^{'}$ satisfies $B^{'}[i] = r[i]$ if $i \in \mathcal{F}$ and otherwise $B^{'}[i]$ uniformly distributed over $\mathbb{F}_{2^n}$, resulting in that

$$Z^{'} = \vec{L} \cdot \vec{B}^{'} = \left(\bigoplus_{i \in \mathcal{F}} L[i] \cdot r[i] \right) \oplus \left(\bigoplus_{i \notin \mathcal{F}} L[i] \cdot B^{'}[i] \right) = a \oplus \left(\bigoplus_{i \notin \mathcal{F}} L[i] \cdot B^{'}[i] \right), \tag{9}$$

where a is a constant value over $\mathbb{F}_{2^n}$. According to Equation 9, we have that

$$\begin{cases} Z^{'} = a, \text{ if } |\mathcal{F}| = l; \\ Z^{'} = a \oplus \left(\bigoplus_{i \notin \mathcal{F}} L[i] \cdot B^{'}[i] \right), \text{ if } |\mathcal{F}| < l. \end{cases} \tag{10}$$

For the latter condition such that $|\mathcal{F}| < l$, the vulnerable variable $Z^{'}$ with ineffective faults follows a uniform distribution, and thus results in SIFA-1 failure, as established in Corollary 2. While for the former case where $|\mathcal{F}| = l$, the distribution of $Z^{'}$ with ineffective faults satisfies $Pr[Z^{'} = a] = 1$ and $Pr[Z^{'} \neq a] = 0$. Hence the distribution of $Z^{'}$ with ineffective faults is biased and can be exploited to attack IPM. Please refer to Appendix I for simulated attacks under constant-byte model. In conclusion, under constant-byte faults where constant values are injected per execution, SIFA-1 can successfully compromise IPM only when all shares $B[i]$ are faulted (implying $|\mathcal{F}| = l$). This scenario is equivalent to constructing $|\mathcal{F}| = nl$ under stuck-at faults. It suggests that the SIFA-1 adversary must have a full control of all-bit fault injection across all shares with high precision (lower precision will definitely increase attack complexity) for successful attack. However, this requirement significantly constraints the practical feasibility of such attacks against IPM-protected implementations.

3.6 Random-Byte Faults

In this sub-section, we evaluate the resistance ability of IP encoding against SIFA-1 under random-byte faults, which inject random values into an arbitrary subset of the shares vector $\vec{B}$, representing a particularly weak adversarial fault model requiring minimal injection precision. Similar to constant-byte faults, the fault model should be adjusted slightly: let $\mathcal{F} \subseteq \{1, \cdots, l\}$ denote the set of byte-wise fault locations. Also Corollary 2 remains valid under this fault model. After fault injection, the resulting faulty shares $B^{'}[i]$ are randomly (uniformly) distributed for the entire attack for $i \in \mathcal{F}$. Hence, the addition vector $\vec{e} = \vec{B} \oplus \vec{B}^{'}$ satisfies $e[i] = 0$ if $i \notin \mathcal{F}$ and otherwise $e[i]$ uniformly distributed over $\mathbb{F}_{2^n}$. Consequently, the probability of obtaining ineffective faults for each execution is given by $P(\vec{L} \cdot \vec{e} = 0) = \frac{1}{2^n}$, resulting in $P(\vec{L} \cdot \vec{e} = 0) = \frac{1}{2^{nT}}$ for the entire attack, where T denotes the number of executions. Further considering the distribution of $Z^{'} = \vec{L} \cdot \vec{B}^{'}$ with ineffective faults, each component $B^{'}[i]$ is uniformly distributed over $\mathbb{F}_{2^n}$, regardless of whether index i belongs to $\mathcal{F}$ or not. Consequently, the resulting output $Z^{'}$ also follows a uniform distribution, making it ineffective for SIFA-1 attacking IPM, according to Corollary 2. In conclusion, IP encoding fundamentally resists SIFA-1 under random-byte faults.

3.7 Summary: What we have Achieved So Far

This section theoretically establishes IP encoding's inherent resistance to SIFA-1, rooted in the algebraic independence of its IPM parameter structure. Under *stuck-at* faults, we observe that l-share IPM requires $\geq l$-bit fault injections (while for both Boolean and prime masking, l bit-injections suffice) across all shares for effective SIFA-1, with only a few specific bit-location combinations proving effective at the minimal fault-injection cost. For example, optimal two-share (resp., three-share) IPM instances need minimum 4 (resp., 8) bit-wise injection for effective attack, and only 4 (resp., 6) fault models out of $C_{16}^4 = 1820$ (resp., $C_{24}^6 = 134596$) possible combinations are valid at the minimum bit-injection cost. This creates non-trivial implementation barriers for SIFA-1 adversaries. In addition, the minimum required bit corruption grows super-linearly with increasing share count. Another interesting finding is that given a specific IPM instance, the number of bits requiring corruption for successful SIFA-1 attacks exactly matches the probing requirements for effective SCA. Hence, those optimal IPM instances which maximize the *attack setup complexity* simultaneously provide highest side-channel resistance. While under both *toggle/bit-flip* and *random-byte* faults, IP encoding inherently resist SIFA-1. Regarding *constant-byte* faults, SIFA-1 can be effective only when the adversary has full control of all bits across all shares with high precision and thus can change all shares simultaneously to constant values, which is a very strong adversary model and becomes practically infeasible.

4 Inner Product Masking vs. SIFA-2

In this section, we evaluate the resistance of IPM against SIFA-2. Unlike SIFA-1 that relies on non-uniformity of intermediate fault-affected variables, SIFA-2

exploits the conditional *propagation* of faults through non-linear, non-bijective operations (dependent on the unmasked sensitive inputs to these operations). This approach is fundamentally capable of successfully compromising IPM (and other masking schemes as well) if no specialized protection is added, as practical implementations necessarily incorporate non-linear operations. As also claimed in [29], SIFA-2 can be performed in two ways:

- **Gadget-independent SIFA-2**. This type corrupts the input of a non-bijective, non-linear (masked) operation such as an AND gate or, more generally, a field multiplication gadget of arbitrary size.
- **Gadget-dependent SIFA-2**. This type operates by injecting faults inside a masked gadget computing non-linear operation (e.g., multiplication).

In the following analysis, we will quantify the attack complexity of gadget-independent SIFA-2 attacking IPM and give a analysis on a state-of-the-art multiplication gadget proposed in CHES 2025 paper [21]. Given that fault detection requires at least temporarily effective injections (distinguishing only between propagated and non-propagated faults), we will proceed under the assumption that the fault model can induce effective single-bit fault on unmasked inputs, that is extensively assumed in previous literature works [29, 33, 17].

4.1 Gadget-independent SIFA-2

Among IP-masked gadgets, only the multiplication gadget exhibits both non-linearity and non-bijectivity. Therefore, our analysis focuses exclusively on the multiplication gadget in this sub-section. Under this scenario, gadget-independent SIFA-2 exploits the fact that a fault in one of the inputs of a multiplication gadget (of any size) will fail to propagate to the output when the other (unmasked) input equals 0. Consequently, the observation of any valid outputs (resulting from ineffective faults) directly leaks information about the input value being 0. However, the probability of an input being zero is field-size dependent ($\frac{1}{2^n}$ for $\mathbb{F}_{2^n}$, i.e., $\frac{1}{2}$ for an AND gate). This demonstrates an inverse relationship between field size and vulnerability, with smaller fields being more susceptible to such attacks. Consequently, word-level masking offers inherent security advantages over bit-level approaches. In conclusion, IPM demonstrates equivalent resistance to gadget-independent SIFA-2 attacks as BM, with both schemes exhibiting field-size-dependent security characteristics, with attack complexity lower-bounded by $\frac{1}{\text{field size}}$. Please refer to Appendix J for simulated attacks under gadget-independent SIFA-2.

4.2 Gadget-dependent SIFA-2

Regarding gadget-dependent SIFA-2 scenarios, the resistance of IPM demonstrates significant variation depending on specific gadget realizations, instead of the specific IPM parameters (as in SIFA-1 models). As analyzed in [29], a fault injected in a gadget's internal operation can become exploitable through two mechanisms: 1) the fault propagates to all output shares, rendering the combinations statistically biased dependent on unmasked values; 2) if not propagating

to all output shares, the propagation of the fault may be conditioned on all the input shares, thereby making the correctness of gadget conditioned on the unmasked input (this typically happens when a faulted variable is multiplied by all the shares of an input variable).

According to above-mentioned two principles, we analyze the state-of-the-art multiplication gadget proposed in [21]. The gadget operates in three stages: 1) generating constant matrices $\mathbf{S}$ (the inner product of zero) and $\mathbf{V}$ (field-size dependent matrix) in Algorithm 1 of [21]; 2) cross products of input shares in Algorithm 2 and 3 of [21]; 3) share compressing using IPM parameters vector $\vec{L}$ in Algorithm 3 of [21]. Considering the cross products, corrupting any single bit of internal variables (bit-wise or byte-wise) leaks at most one share per input variable and propagates to just one output share. While regarding stage 1 and 3, there are three possible types of variables that can be corrupted:

- internal variables $z_{i,j}$ and z_j (related to the output shares) that depend on single input/output shares and propagate faults to at most one output share;
- randomness variables that has no effect on the correctness of unmasked states and thus such a fault cannot leak information by ineffectiveness;
- some constant variables such as $\mathbf{V}$ and $\vec{L}$. Corrupting those constant variables could transform the IP encoding into another "valid" IP encoding, might introducing some new vulnerabilities. However, this scenario involves persistent/permanent fault models, which falls outside the scope of our current investigation focusing exclusively on transitional fault models.

Therefore, faulting the internals of such an IPM multiplication gadget provides no additional advantage over input faulting adversaries (gadget-independent SIFA-2 scenarios). However, there exist some potential vulnerabilities against persistent/permanent faults, requiring additional protection mechanisms for these constant variable. This deeper investigation is left as a future work.

5 Discussion

5.1 Strong Dual Resistance of IP encoding to SCA and SIFA-1

This work theoretically establishes an equivalence between the side-channel and SIFA-1 (under stuck-at faults) resistance in IP encoding through the derivation of two indicating coding-theoretic parameters $d_{\mathcal{L}} = d_{\mathcal{D}^\perp}$ and $B_{d_{\mathcal{L}}} = B_{d_{\mathcal{D}^\perp}}$. Consequently, optimal IPM instances exist offering highest side-channel resistance simultaneously maximizing the *attack setup complexity* required for SIFA-1. Specifically, l-share IPM shares identical $l-1$ probing security order [23] with l-share Boolean masking. However, some IPM instances feature higher bit-probing security order [30, 11] (indicated by $d_{\mathcal{L}}-1 > l-1$) than that of BM (with $l-1$ bit-probing security order), thus requiring substantially more physical measurements (reaching up to 1000× [43] under two-share case) for successful SCA. Interestingly, those IPM instances also require at least $d_{\mathcal{L}} > l$ bit-wise corruption for effective SIFA-1 under stuck-at faults, while BM only needs l faulty bits, thereby significantly amplifying the *attack setup complexity* of SIFA-1. Taking

two-share IPM as an example which is reported to be $1.5\times$ [2, 42] costly than BM in software implementation, IPM requires 4 bit corruption with only 4 effective fault models (e.g., Figure 5), while BM needs 2 faulty bit with 8 effective fault location combinations. Furthermore, with increase in share count, IPM exhibits a super-linear growth in required minimal bit-wise fault injections (e.g., from 4 to 8 if increasing from 2 to 3 shares), while BM shows a linear growth (e.g., from 2 to 3 if increasing from 2 to 3 shares). Therefore, while incurring higher implementation cost (potentially reducible in hardware) compared to BM, IP encoding delivers substantially stronger protection against both side-channel analysis and SIFA-1 attacks, making it a more promising practical countermeasure.

5.2 Comparison between IPM and Prime Masking [29] on SIFA-1

Under toggle/bit-flip fault conditions, IPM fundamentally prevents SIFA-1 attacks, while prime masking remains vulnerable to such exploitation as shown in [29]. For random-byte and constant-byte faults, we omit discussion as these scenarios were not considered in prior work [29]. Regarding stuck-at fault models, IPM demonstrates higher *attack setup complexity* than prime masking, requiring both a greater number of bit corruptions (4 vs. 2 for two-share; 8 vs. 3 for three-share) and higher-precision control over fault location injection (e.g., Figure 5), which imposes more substantial practical challenges for adversaries. In terms of *attack complexity*, IPM achieves a lower probability of yielding valid outputs compared to prime masking under effective fault models and small-share conditions. Recall the probability of IP encoding is quantified by $2^{-\mathcal{R}(\mathbf{L}[*,\mathcal{F}])}$, while for prime masking, the probability is $Pr[C = C^{'}] = \left(\frac{p+1}{2p}\right)^{l} > \frac{1}{2^{l}}$, where p denotes the prime field size. Therefore, the probability of obtaining valid outputs exhibits fundamentally different exponential scaling properties between masking schemes: for prime masking it grows exponentially with the number of shares, while for IPM the complexity grows exponentially with the rank of $\mathbf{L}[*, \mathcal{F}]$, demonstrating that IPM's security depends on algebraic independence rather than just share count. However, when the number of shares in prime masking exceeds the word size of IPM (i.e., $l > n$), prime masking demonstrates superior *attack complexity*; while for cases where $l \leq n$, IPM is shown to be superior under small-share conditions. Remarkably, IPM offers the crucial advantage of direct applicability to existing cipher implementations, whereas prime masking necessitates the design of new cipher architectures.

6 Conclusion

In this paper, we systematically evaluate the resistance capability of IPM against SIFA, including both IP encoding against SIFA-1 and IP masked gadgets against SIFA-2. Our theoretical analysis establishes that IP encoding inherently resists SIFA-1 attacks under four widely-studied fault models, including *stuck-at*, *toggle/bit-flip*, *constant-byte* and *random-byte* faults. Specifically under *stuck-at* faults, a theoretical framework is developed to characterize SIFA-1 security of IP

encoding, deriving two matrix-rank parameters to determine both the effectiveness and complexity of SIFA-1 attack, and thus the optimal IPM parameters can be identified to maximize SIFA-1 defense. Interestingly on this basis, we build a theoretical equivalence of side-channel protection and SIFA-1 resistance in IP encoding. We show by concrete optimal IPM instances that can provide strong (optimal) dual protection against both SCA and SIFA-1 under stuck-at faults.

Regarding the resistance of IP masked gadgets against SIFA-2, we consider both gadget-independent and gadget-dependent scenarios. For the former case, we reveal that the attack complexity is lower bounded by $\frac{1}{\text{field size}}$. Considering the latter, we give a SIFA-2 analysis of the state-of-the-art masked multiplication gadget and identifies some vulnerabilities to persistent/permanent fault models. We highlight the critical need to develop SIFA-2-resistant IPM gadgets as an open research challenge, which would extend the native and strong SIFA-1 protection already inherent in IP encoding.

References

1. Ali, S., Mukhopadhyay, D., Tunstall, M.: Differential fault analysis of AES: towards reaching its limits. J. Cryptogr. Eng. **3**(2), 73–97 (2013). https://doi.org/10.1007/S13389-012-0046-Y, https://doi.org/10.1007/s13389-012-0046-y
2. Balasch, J., Faust, S., Gierlichs, B., Paglialonga, C., Standaert, F.: Consolidating inner product masking. In: Takagi, T., Peyrin, T. (eds.) 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. Lecture Notes in Computer Science, vol. 10624, pp. 724–754. Springer (2017). https://doi.org/10.1007/978-3-319-70694-8_25, https://doi.org/10.1007/978-3-319-70694-8_25
3. Bar-El, H., Choukri, H., Naccache, D., Tunstall, M., Whelan, C.: The sorcerer's apprentice guide to fault attacks. Proc. IEEE **94**(2), 370–382 (2006). https://doi.org/10.1109/JPROC.2005.862424, https://doi.org/10.1109/JPROC.2005.862424
4. Bartkewitz, T., Bettendorf, S., Moos, T., Moradi, A., Schellenberg, F.: Beware of insufficient redundancy an experimental evaluation of code-based FI countermeasures. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2022**(3), 438–462 (2022). https://doi.org/10.46586/TCHES.V2022.I3.438-462, https://doi.org/10.46586/tches.v2022.i3.438-462
5. Belaïd, S., Feldtkeller, J., Güneysu, T., Guinet, A., Richter-Brockmann, J., Rivain, M., Sasdrich, P., Taleb, A.R.: Formal definition and verification for combined random fault and random probing security. In: Chung, K., Sasaki, Y. (eds.) Advances in Cryptology - ASIACRYPT 2024 - 30th International Conference on the Theory and Application of Cryptology and Information Security, Kolkata, India, December 9-13, 2024, Proceedings, Part VII. Lecture Notes in Computer Science, vol. 15490, pp. 167–200. Springer (2024). https://doi.org/10.1007/978-981-96-0941-3_6, https://doi.org/10.1007/978-981-96-0941-3_6
6. Berndt, S., Eisenbarth, T., Faust, S., Gourjon, M., Orlt, M., Seker, O.: Combined fault and leakage resilience: Composability, constructions and compiler. In: Handschuh, H., Lysyanskaya, A. (eds.) Advances in Cryptology - CRYPTO 2023 - 43rd

Annual International Cryptology Conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings, Part III. Lecture Notes in Computer Science, vol. 14083, pp. 377–409. Springer (2023)

7. Biham, E., Shamir, A.: Differential fault analysis of secret key cryptosystems. In: Jr., B.S.K. (ed.) Advances in Cryptology - CRYPTO '97, 17th Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 1997, Proceedings. Lecture Notes in Computer Science, vol. 1294, pp. 513–525. Springer (1997). https://doi.org/10.1007/BFB0052259, https://doi.org/10.1007/BFb0052259
8. Boneh, D., DeMillo, R.A., Lipton, R.J.: On the importance of checking cryptographic protocols for faults (extended abstract). In: Fumy, W. (ed.) Advances in Cryptology - EUROCRYPT '97, International Conference on the Theory and Application of Cryptographic Techniques, Konstanz, Germany, May 11-15, 1997, Proceeding. Lecture Notes in Computer Science, vol. 1233, pp. 37–51. Springer (1997). https://doi.org/10.1007/3-540-69053-0_4, https://doi.org/10.1007/3-540-69053-0_4
9. Cheng, W.: linear code candidates for three-share ipm. https://github.com/Qomo-CHENG/Optimal_code_in_CBM_aes/blob/main/outputs/log_ipm_n3_aes.log (2022)
10. Cheng, W., Guilley, S., Carlet, C., Danger, J., Mesnager, S.: Information leakages in code-based masking: A unified quantification approach. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2021**(3), 465–495 (2021). https://doi.org/10.46586/TCHES.V2021.I3.465-495, https://doi.org/10.46586/tches.v2021.i3.465-495
11. Cheng, W., Guilley, S., Carlet, C., Mesnager, S., Danger, J.: Optimizing inner product masking scheme by a coding theory approach. IEEE Trans. Inf. Forensics Secur. **16**, 220–235 (2021). https://doi.org/10.1109/TIFS.2020.3009609, https://doi.org/10.1109/TIFS.2020.3009609
12. Cheng, W., Guilley, S., Danger, J.: Information leakage in code-based masking: A systematic evaluation by higher-order attacks. IEEE Trans. Inf. Forensics Secur. **17**, 1624–1638 (2022). https://doi.org/10.1109/TIFS.2022.3167914, https://doi.org/10.1109/TIFS.2022.3167914
13. Daemen, J., Dobraunig, C., Eichlseder, M., Groß, H., Mendel, F., Primas, R.: Protecting against statistical ineffective fault attacks. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2020**(3), 508–543 (2020). https://doi.org/10.13154/TCHES.V2020.I3.508-543, https://doi.org/10.13154/tches.v2020.i3.508-543
14. Dhooghe, S., Nikova, S.: My gadget just cares for me - how NINA can prove security against combined attacks. In: Jarecki, S. (ed.) Topics in Cryptology - CT-RSA 2020 - The Cryptographers' Track at the RSA Conference 2020, San Francisco, CA, USA, February 24-28, 2020, Proceedings. Lecture Notes in Computer Science, vol. 12006, pp. 35–55. Springer (2020)
15. Dhooghe, S., Nikova, S.: The random fault model. In: Carlet, C., Mandal, K., Rijmen, V. (eds.) Selected Areas in Cryptography - SAC 2023 - 30th International Conference, Fredericton, Canada, August 14-18, 2023, Revised Selected Papers. Lecture Notes in Computer Science, vol. 14201, pp. 191–212. Springer (2023)
16. Dhooghe, S., Ovchinnikov, A., Toprakhisar, D.: Stati: Protecting against fault attacks using stable threshold implementations. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2024**(1), 229–263 (2024)

17. Dobraunig, C., Eichlseder, M., Groß, H., Mangard, S., Mendel, F., Primas, R.: Statistical ineffective fault attacks on masked AES with fault countermeasures. In: Peyrin, T., Galbraith, S.D. (eds.) Advances in Cryptology - ASIACRYPT 2018 - 24th International Conference on the Theory and Application of Cryptology and Information Security, Brisbane, QLD, Australia, December 2-6, 2018, Proceedings, Part II. Lecture Notes in Computer Science, vol. 11273, pp. 315–342. Springer (2018). https://doi.org/10.1007/978-3-030-03329-3_11, https://doi.org/10.1007/978-3-030-03329-3_11
18. Dobraunig, C., Eichlseder, M., Korak, T., Mangard, S., Mendel, F., Primas, R.: SIFA: exploiting ineffective fault inductions on symmetric cryptography. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2018**(3), 547–572 (2018). https://doi.org/10.13154/TCHES.V2018.I3.547-572, https://doi.org/10.13154/tches.v2018.i3.547-572
19. Feldtkeller, J., Güneysu, T., Moos, T., Richter-Brockmann, J., Saha, S., Sasdrich, P., Standaert, F.: Combined private circuits - combined security refurbished. In: Meng, W., Jensen, C.D., Cremers, C., Kirda, E. (eds.) Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security, CCS 2023, Copenhagen, Denmark, November 26-30, 2023. pp. 990–1004. ACM (2023)
20. Feldtkeller, J., Richter-Brockmann, J., Sasdrich, P., Güneysu, T.: CINI MINIS: domain isolation for fault and combined security. In: Yin, H., Stavrou, A., Cremers, C., Shi, E. (eds.) Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022. pp. 1023–1036. ACM (2022). https://doi.org/10.1145/3548606.3560614, https://doi.org/10.1145/3548606.3560614
21. Gaspoz, J., Dhooghe, S.: Bit t-sni secure multiplication gadget for inner product masking. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2025**(1), 104–127 (2025). https://doi.org/10.46586/TCHES.V2025.I1.104-127, https://doi.org/10.46586/tches.v2025.i1.104-127
22. Ghalaty, N.F., Yuce, B., Taha, M.M.I., Schaumont, P.: Differential fault intensity analysis. In: Tria, A., Choi, D. (eds.) 2014 Workshop on Fault Diagnosis and Tolerance in Cryptography, FDTC 2014, Busan, South Korea, September 23, 2014. pp. 49–58. IEEE Computer Society (2014). https://doi.org/10.1109/FDTC.2014.15, https://doi.org/10.1109/FDTC.2014.15
23. Ishai, Y., Sahai, A., Wagner, D.A.: Private circuits: Securing hardware against probing attacks. In: Boneh, D. (ed.) Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings. Lecture Notes in Computer Science, vol. 2729, pp. 463–481. Springer (2003). https://doi.org/10.1007/978-3-540-45146-4_27, https://doi.org/10.1007/978-3-540-45146-4_27
24. Kocher, P.C.: Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems. In: Koblitz, N. (ed.) Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings. Lecture Notes in Computer Science, vol. 1109, pp. 104–113. Springer (1996). https://doi.org/10.1007/3-540-68697-5_9, https://doi.org/10.1007/3-540-68697-5_9
25. Li, Y., Sakiyama, K., Gomisawa, S., Fukunaga, T., Takahashi, J., Ohta, K.: Fault sensitivity analysis. In: Mangard, S., Standaert, F. (eds.) Cryptographic Hardware and Embedded Systems, CHES 2010, 12th International Workshop, Santa Barbara, CA, USA, August 17-20, 2010. Proceedings. Lecture Notes in Computer Science,

vol. 6225, pp. 320–334. Springer (2010). https://doi.org/10.1007/978-3-642-15031-9_22, https://doi.org/10.1007/978-3-642-15031-9_22
26. MacWilliams, F.J., Sloane, N.J.A.: The theory of error-correcting codes, vol. 16. Elsevier (1977)
27. Mangard, S., Oswald, E., Popp, T.: Power analysis attacks - revealing the secrets of smart cards. Springer (2007)
28. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings. Lecture Notes in Computer Science, vol. 765, pp. 386–397. Springer (1993). https://doi.org/10.1007/3-540-48285-7_33, https://doi.org/10.1007/3-540-48285-7_33
29. Moos, T., Saha, S., Standaert, F.: Prime masking vs. faults - exponential security amplification against selected classes of attacks. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2024**(4), 690–736 (2024). https://doi.org/10.46586/TCHES.V2024.I4.690-736, https://doi.org/10.46586/tches.v2024.i4.690-736
30. Poussier, R., Guo, Q., Standaert, F., Carlet, C., Guilley, S.: Connecting and improving direct sum masking and inner product masking. In: Eisenbarth, T., Teglia, Y. (eds.) Smart Card Research and Advanced Applications - 16th International Conference, CARDIS 2017, Lugano, Switzerland, November 13-15, 2017, Revised Selected Papers. Lecture Notes in Computer Science, vol. 10728, pp. 123–141. Springer (2017). https://doi.org/10.1007/978-3-319-75208-2_8, https://doi.org/10.1007/978-3-319-75208-2_8
31. Rivain, M., Prouff, E.: Provably secure higher-order masking of AES. In: Mangard, S., Standaert, F. (eds.) Cryptographic Hardware and Embedded Systems, CHES 2010, 12th International Workshop, Santa Barbara, CA, USA, August 17-20, 2010. Proceedings. Lecture Notes in Computer Science, vol. 6225, pp. 413–427. Springer (2010). https://doi.org/10.1007/978-3-642-15031-9_28, https://doi.org/10.1007/978-3-642-15031-9_28
32. Sadhukhan, R., Saha, S., Mukhopadhyay, D.: Antisifa-cad: A framework to thwart sifa at the layout level. In: 2022 IEEE/ACM International Conference On Computer Aided Design (ICCAD). pp. 1–9 (2022)
33. Saha, S., Bag, A., Jap, D., Mukhopadhyay, D., Bhasin, S.: Divided we stand, united we fall: Security analysis of some SCA+SIFA countermeasures against sca-enhanced fault template attacks. In: Tibouchi, M., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2021 - 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6-10, 2021, Proceedings, Part II. Lecture Notes in Computer Science, vol. 13091, pp. 62–94. Springer (2021). https://doi.org/10.1007/978-3-030-92075-3_3, https://doi.org/10.1007/978-3-030-92075-3_3
34. Saha, S., Bag, A., Roy, D.B., Patranabis, S., Mukhopadhyay, D.: Fault template attacks on block ciphers exploiting fault propagation. In: Canteaut, A., Ishai, Y. (eds.) Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part I. Lecture Notes in Computer Science, vol. 12105, pp. 612–643. Springer (2020). https://doi.org/10.1007/978-3-030-45721-1_22, https://doi.org/10.1007/978-3-030-45721-1_22
35. Saha, S., Jap, D., Roy, D.B., Chakraborty, A., Bhasin, S., Mukhopadhyay, D.: A framework to counter statistical ineffective fault analysis of block ci-

phers using domain transformation and error correction. IEEE Trans. Inf. Forensics Secur. **15**, 1905–1919 (2020). https://doi.org/10.1109/TIFS.2019.2952262, https://doi.org/10.1109/TIFS.2019.2952262

36. Saha, S., Ravi, P., Jap, D., Bhasin, S.: Non-profiled side-channel assisted fault attack: A case study on DOMREP. In: Design, Automation & Test in Europe Conference & Exhibition, DATE 2023, Antwerp, Belgium, April 17-19, 2023. pp. 1–6. IEEE (2023). https://doi.org/10.23919/DATE56975.2023.10137176, https://doi.org/10.23919/DATE56975.2023.10137176
37. Shahmirzadi, A.R., Moradi, A.: Clock glitch versus sifa. In: 2020 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT). pp. 1–6 (2020)
38. Solé, P., Liu, Y., Cheng, W., Guilley, S., Rioul, O.: Linear programming bounds on the kissing number of q-ary codes. In: IEEE Information Theory Workshop, ITW 2021, Kanazawa, Japan, October 17-21, 2021. pp. 1–5. IEEE (2021). https://doi.org/10.1109/ITW48936.2021.9611478, https://doi.org/10.1109/ITW48936.2021.9611478
39. Standaert, F., Malkin, T., Yung, M.: A unified framework for the analysis of side-channel key recovery attacks. In: Joux, A. (ed.) Advances in Cryptology - EUROCRYPT 2009, 28th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cologne, Germany, April 26-30, 2009. Proceedings. Lecture Notes in Computer Science, vol. 5479, pp. 443–461. Springer (2009). https://doi.org/10.1007/978-3-642-01001-9_26, https://doi.org/10.1007/978-3-642-01001-9_26
40. Veyrat-Charvillon, N., Gérard, B., Renauld, M., Standaert, F.: An optimal key enumeration algorithm and its application to side-channel attacks. In: Knudsen, L.R., Wu, H. (eds.) Selected Areas in Cryptography, 19th International Conference, SAC 2012, Windsor, ON, Canada, August 15-16, 2012, Revised Selected Papers. Lecture Notes in Computer Science, vol. 7707, pp. 390–406. Springer (2012). https://doi.org/10.1007/978-3-642-35999-6_25, https://doi.org/10.1007/978-3-642-35999-6_25
41. Wang, W., Standaert, F., Yu, Y., Pu, S., Liu, J., Guo, Z., Gu, D.: Inner product masking for bitslice ciphers and security order amplification for linear leakages. In: Lemke-Rust, K., Tunstall, M. (eds.) Smart Card Research and Advanced Applications - 15th International Conference, CARDIS 2016, Cannes, France, November 7-9, 2016, Revised Selected Papers. Lecture Notes in Computer Science, vol. 10146, pp. 174–191. Springer (2016). https://doi.org/10.1007/978-3-319-54669-8_11, https://doi.org/10.1007/978-3-319-54669-8_11
42. Wu, Q., Cheng, W., Guilley, S., Zhang, F., Fu, W.: On efficient and secure code-based masking: A pragmatic evaluation. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2022**(3), 192–222 (2022). https://doi.org/10.46586/TCHES.V2022.I3.192-222, https://doi.org/10.46586/tches.v2022.i3.192-222
43. Wu, Q., Cheng, W., Zhang, F., Guilley, S.: Statistical analysis of non-profiling higher-order distinguishers against inner product masking. IEEE Trans. Inf. Forensics Secur. **20**, 4008–4023 (2025). https://doi.org/10.1109/TIFS.2025.3558601, https://doi.org/10.1109/TIFS.2025.3558601
44. Zhang, F., Lou, X., Zhao, X., Bhasin, S., He, W., Ding, R., Qureshi, S., Ren, K.: Persistent fault analysis on block ciphers. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2018**(3), 150–172 (2018)

A Proof of Theorem 1

Proof. We begin with the proof of stuck-at-0 faults, and then demonstrate that stuck-at-1 faults exhibit equivalent attack characteristics. Recall $Z^{'} = \vec{L} \cdot \vec{B}^{'}$. By using the code expansion of $\vec{L}$ and the sub-field representation of $\vec{B}^{'}$, we can expand the inner product of vectors $\vec{L} \cdot \vec{B}^{'}$ over $\mathbb{F}_{2^n}^l$ into a matrix multiplication with size $n \times nl$ and size $nl \times 1$ over $\mathbb{F}_2$:

$$\vec{L} \cdot \vec{B}^{'} = [\mathbf{L_1}, \cdots, \mathbf{L_l}]_2 \times [\vec{B}^{'}]_2. \tag{11}$$

Under stuck-at-0 faults, the faulty shares $\vec{B}^{'}$ with ineffective faults satisfies $[\vec{B}^{'}]_2[i] = 0$ if $i \in \mathcal{F}$ and otherwise $[\vec{B}^{'}]_2[i]$ uniformly distributed over $\mathbb{F}_2$. By splitting the fault-free and fault-affected components (resp., columns) within $[\vec{B}^{'}]_2$ (resp., $\mathbf{L}$), we have

$$\begin{aligned}
[\mathbf{L_1}, \cdots, \mathbf{L_l}]_2 \times [\vec{B}^{'}]_2 &= \mathbf{L}[*, \mathcal{F}] \times [\vec{B}]_2[\mathcal{F}] \oplus \mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}] \\
&= \vec{0} \oplus \mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}] \\
&= \mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}].
\end{aligned} \tag{12}$$

Hence $\mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}]$ characterizes the distribution of $Z^{'}$ with ineffective faults. Let regard $[Z^{'}]_2 = \mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}]$ as a linear transformation $L_{\overline{\mathcal{F}}} : \mathcal{B} \rightarrow \mathcal{Z}^{'}$ from $\mathbb{F}_2^{nl-|\mathcal{F}|}$ to $\mathbb{F}_2^n$, represented by $\mathbf{L}[*, \overline{\mathcal{F}}]$. The following two conditions must simultaneously hold to ensure $Z^{'}$ is uniformly distributed over $\mathbb{F}_{2^n}$:

- $nl - |\mathcal{F}| \geq n$, namely $|\mathcal{F}| \leq (l-1)n$;
- The rank of matrix $\mathbf{L}[*, \overline{\mathcal{F}}]$ should be n such that $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = n$.

If $|\mathcal{F}| > (l-1)n$, then the rank of matrix $\mathbf{L}[*, \overline{\mathcal{F}}]$ is necessarily less than n (recall that the size of $\mathbf{L}[*, \overline{\mathcal{F}}]$ is $n \times (nl - |\mathcal{F}|)$ over $\mathbb{F}_2$), which violates the second condition. Therefore, we can relax the first condition and rely solely on the second condition to determine whether $Z^{'}$ is uniformly distributed. The proof for stuck-at-0 faults is now complete.

Extending the proof for stuck-at-1 faults, the faulty shares $\vec{B}^{'}$ satisfies $[\vec{B}^{'}]_2[i] = 1$ if $i \in \mathcal{F}$ and otherwise $[\vec{B}^{'}]_2[i]$ uniformly distributed over $\mathbb{F}_2$. Applying identical analytical methods to those used for stuck-at-0 faults yields:

$$\begin{aligned}
[\mathbf{L_1}, \cdots, \mathbf{L_l}]_2 \times [\vec{B}^{'}]_2 &= \mathbf{L}[*, \mathcal{F}] \times [\vec{B}]_2[\mathcal{F}] \oplus \mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}] \\
&= [a]_2 \oplus \mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}],
\end{aligned} \tag{13}$$

where a is a constant value over $\mathbb{F}_{2^n}$. Hence the distribution of $Z^{'}$ with ineffective faults under stuck-at-1 faults is also determined by the distribution of $\mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}]$, mirroring its behavior under stuck-at-0 faults. This completes the proof.

B Proof of Lemma 2

Proof. For each $\vec{u} \in \mathcal{L}\backslash\{\vec{0}\}$, there exists a row vector $\vec{c} \in \mathbb{F}_2^{1\times n}$ enabling $\vec{u} = \vec{c}\times\mathbf{L}$. We define its non-zero positions $\mathcal{I}_c := \{i|c[i] = 1\}$. Hence, we have

$$\vec{u} = \vec{c}\times\mathbf{L} = \bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,*] \oplus \bigoplus_{i\in[n]\backslash\mathcal{I}_c}\mathbf{L}[i,*] = \bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,*], \tag{14}$$

where $\left(\bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,*]\right)[j] = 1$ if $j \in \mathcal{I}_u$, and otherwise $\left(\bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,*]\right)[j] = 0$. If the fault model $\mathcal{F}$ satisfies $\mathcal{F}\cap\mathcal{I}_u = \mathcal{I}_u$ and we assume $\mathcal{F} = \mathcal{I}_u \cup \mathcal{O}$, there is

$$\vec{c}\times\mathbf{L}[*,\overline{\mathcal{F}}] = \bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,\overline{\mathcal{F}}] = \bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,\overline{\mathcal{I}_u\cup\mathcal{O}}] = \vec{0}. \tag{15}$$

Hence, the i-th row vectors of $\mathbf{L}[*,\overline{\mathcal{F}}]$ for $i \in \mathcal{I}_c$ are linear dependent. Considering the extreme case where $|\mathcal{I}_c| = 1$, it implies that there exists one all-zero row vector in $\mathbf{L}[*,\overline{\mathcal{F}}]$. Hence, the rank of matrix $\mathbf{L}[*,\overline{\mathcal{F}}]$ is strictly less than n such that $\mathsf{Rank}(\mathbf{L}[*,\overline{\mathcal{F}}]) < n$.

If for any $\vec{u} \in \mathcal{L}\backslash\{\vec{0}\}$, the fault model fails to contain all its non-zero positions such that $\mathcal{F}\cap\mathcal{I}_u \neq \mathcal{I}_u$, then there always exists a position $j \in \mathcal{I}_u \cap \overline{\mathcal{F}}$. Hence, for any $\vec{c} \in \mathbb{F}_2^n \setminus \{\vec{0}\}$, there exists a j enabling that

$$\vec{c}\times\mathbf{L}[*,\overline{\mathcal{F}}] = \bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,\overline{\mathcal{F}}] \neq \vec{0}, \tag{16}$$

since we have $(\bigoplus_{i\in\mathcal{I}_c}\mathbf{L}[i,\overline{\mathcal{F}}])[j] = 1$. Thus, the row vectors of $\mathbf{L}[*,\overline{\mathcal{F}}]$ are linear independent, yielding $\mathsf{Rank}(\mathbf{L}[*,\overline{\mathcal{F}}]) = n$. This completes the proof.

C Proof of Theorem 2

Proof. Assume for each share $B[i]$ for $i \in [l]$, the set of fault-free locations is denoted as $\mathcal{M}_i \subseteq [n]$, satisfying $\bigcup_{i=1}^{l}\{(i-1)n + x|x \in \mathcal{M}_i\} = \{1,\cdots,nl\}\setminus\mathcal{F}$ and $\sum_{i=1}^{l}|\mathcal{M}_i| = |\overline{\mathcal{F}}|$. Hence we have

$$\mathbf{L}[*,\overline{\mathcal{F}}] = [\mathbf{L_1}[*,\mathcal{M}_1],\mathbf{L_2}[*,\mathcal{M}_2],\cdots,\mathbf{L_l}[*,\mathcal{M}_l]]_2, \tag{17}$$

where the column vectors in each matrix $\mathbf{L_i}[*,\mathcal{M}_i]$ are linear independent, for $i \in [l]$. Let $\mathcal{V}_i$ for $i \in [l]$ be the linear subspace with dimension $|\mathcal{M}_i|$ of the vector space $\mathbb{F}_2^n$, and its generating matrix is $\mathbf{L_i}[*,\mathcal{M}_i]^T$.

Proof of $C1$. Let's begin with the worst-scenario attack where $|\mathcal{F}| = (l-1)n$, resulting in $\sum_{i=1}^{l}\dim(\mathcal{V}_i) = \sum_{i=1}^{l}|\mathcal{M}_i| = n$. Hence, we should prove that for any given $\mathcal{F}$ with its size satisfying $|\mathcal{F}| = (l-1)n$, there exits $\vec{L}$ enabling $\bigcap_{i=1}^{l}\mathcal{V}_i = \vec{0}$. Next, we employ proof by contradiction to demonstrate this proposition.

Given a specific fault model $\mathcal{F}$ with $|\mathcal{F}| = (l-1)n$, assume that for all $\vec{L} \in \mathbb{F}_{2^n}^l$, there exists i, j and one non-zero $\vec{v}$ enabling $\mathcal{V}_i \cap \mathcal{V}_j = \vec{v} \neq \vec{0}$. Equivalently, for all $\vec{L}$ with $|\vec{L}| = l$, there exists $\vec{a} \in \mathbb{F}_2^{|\mathcal{M}_i|} \setminus \{\vec{0}\}$, $\vec{b} \in \mathbb{F}_2^{|\mathcal{M}_j|} \setminus \{\vec{0}\}$, enabling that

$$\vec{a}^T \times \mathbf{L_i}[*, \mathcal{M}_i]^T = \vec{b}^T \times \mathbf{L_j}[*, \mathcal{M}_j]^T.$$

Converting $\mathbb{F}_2$ to $\mathbb{F}_{2^n}$, we have

$$\bigoplus_{k=1}^{|\mathcal{M}_i|} a[k] \cdot L[i] \cdot 2^{\mathcal{M}_i[k]-1} = \bigoplus_{y=1}^{|\mathcal{M}_j|} b[y] \cdot L[j] \cdot 2^{\mathcal{M}_j[y]-1}.$$

Thus for all $\vec{L}$ with $|\vec{L}| = l$, there exists $\vec{a} \in \mathbb{F}_2^{|\mathcal{M}_i|} \setminus \{\vec{0}\}$, $\vec{b} \in \mathbb{F}_2^{|\mathcal{M}_j|} \setminus \{\vec{0}\}$, enabling

$$L[i] = \left(\frac{\bigoplus_{y=1}^{|\mathcal{M}_j|} b[y] \cdot 2^{\mathcal{M}_j[y]-1}}{\bigoplus_{k=1}^{|\mathcal{M}_i|} a[k] \cdot 2^{\mathcal{M}_i[k]-1}} \right) \cdot L[j] = x \cdot L[j]. \tag{18}$$

Equation 18 implies that for all $\vec{L}$ with $|\vec{L}| = l$, $L[i]$ and $L[j]$ should satisfy the constraint defined by $L[i] = x \cdot L[j]$. Let $\mathcal{X}$ be the value space of x, hence the maximum size of $\mathcal{X}$ is $(2^{|\mathcal{M}_i|} - 1)(2^{|\mathcal{M}_j|} - 1) < 2^n - 1$. However, there exists $\vec{L}$ with $L[1] = 1$ enabling that $L[i] = \overline{x} \cdot L[j]$ with $\overline{x} \notin \mathcal{X}$, that violates Equation 18 and accordingly causes a contradiction. This completes the proof under worst scenario where $|\mathcal{F}| = (l-1)n$.

Now consider the generic case where $|\mathcal{F}| \leq (l-1)n$. Let $\mathcal{F}^*$ represents the worst ones with $|\mathcal{F}^*| = (l-1)n$, then we have for each $\mathcal{F}$ with $|\mathcal{F}| < (l-1)n$, there exist one $\mathcal{F}^*$ such that $\mathcal{F} \subset \mathcal{F}^*$. Hence, there at least exists $\vec{L}$ with $|\vec{L}| = l$ enabling $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}^*}]) = n$, allowing the corresponding $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = \mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}^*}]) = n$. This completes the proof.

Proof of $C2$. We should prove that given any IPM parameters vector $\vec{L}$ with $|\vec{L}| = l$, there exists $\mathcal{F}$ with $l \leq |\mathcal{F}| \leq (l-1)n$ enabling $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) < n$. We seek to identify at least one valid $\mathcal{F}$ that satisfies the required conditions.

Assume that for a fault model $\mathcal{F}$ with $|\mathcal{F}| = (l-1)n$, there exists i, j enabling $|\mathcal{M}_i| + |\mathcal{M}_j| = n$ and the other remaining $|\mathcal{M}_k| = 0$ for $k \in [l] \setminus \{i, j\}$. Specifically, we define that $|\mathcal{M}_i| = n-1$ and $|\mathcal{M}_j| = 1$. Hence for $\mathcal{M}_i \subset [n]$, it has n possible candidates $\mathcal{M}_i^k$ denoting $\mathcal{M}_i = [n] \setminus \{k\}$ for $k \in [n]$. Similarly for $\mathcal{M}_j$, it also has n possible candidates $\mathcal{M}_j^y$ denoting $\mathcal{M}_j = \{y\}$ for $y \in [n]$. Correspondingly, for each linear subspace $\mathcal{V}_i^k$ with generating matrix $\mathbf{L_i}[*, \mathcal{M}_i^k]^T$, its size satisfies $|\mathcal{V}_i^k| = 2^{n-1}$. Hence the size of $\bigcup_{k=1}^n \mathcal{V}_i^k$ satisfies

$$\left| \bigcup_{k=1}^{n} \mathcal{V}_i^k \right| = \sum_{\mathcal{S} \subseteq [n] \setminus \emptyset} (-1)^{|\mathcal{S}|+1} \left| \bigcap_{k \in \mathcal{S}} \mathcal{V}_i^k \right| = (-1) \times \sum_{i \in [n] \setminus \{0\}} C_n^i 2^{n-i} (-1)^i = 2^n - 1. \tag{19}$$

While the size of $\mathcal{V}_j^y$ is $|\mathcal{V}_j^y| = 1$ and thus $\bigcup_{y=1}^n \mathcal{V}_j^y$ satisfies $|\bigcup_{y=1}^n \mathcal{V}_j^y| = n$. This indicates that there exists $\vec{v} \neq \vec{0}$ that $(\bigcup_{k=1}^n \mathcal{V}_i^k) \cap (\bigcup_{y=1}^n \mathcal{V}_j^y) = \vec{v}$. Hence we can

always find y, k that enables $\mathcal{V}_i^k \cap \mathcal{V}_j^y = \vec{v}$. Thus, we can always construct a valid fault model $\mathcal{F}$ enabling that

$$\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = \mathsf{Rank}([\mathbf{L_i}[*, \mathcal{M}_i^k], \mathbf{L_j}[*, \mathcal{M}_j^y]]_2) < n.$$

This completes the proof.

D Proof of Theorem 3

Proof. As indicated by Equation 12, the distribution of $Z^{'}$ is characterized by $\mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{B}]_2[\overline{\mathcal{F}}]$, which is a linear transformation $L_{\overline{\mathcal{F}}} : \mathcal{B} \to \mathcal{Z}^{'}$ from $\mathbb{F}_2^{nl-|\mathcal{F}|}$ to $\mathbb{F}_2^n$. By the rank–nullity theorem, we have

$$\dim(\ker(L_{\overline{\mathcal{F}}})) = \dim(\mathcal{B}) - \mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]).$$

Thus, the distribution of $Z^{'}$ satisfies

$$Pr[Z^{'} = z^{'}] = \frac{2^{\dim(\ker(L_{\overline{\mathcal{F}}}))}}{2^{\dim(\mathcal{B})}} = \frac{1}{2^{\mathcal{R}(\mathbf{L}[*, \overline{\mathcal{F}}])}}, \text{ for } z^{'} \in \mathcal{Z}^{'}.$$

Note that $\mathrm{im}(L_{\overline{\mathcal{F}}}) = \mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$, hence $Z^{'}$ is uniformly distributed over $\mathcal{Z}^{'}$ with its size $|\mathcal{Z}^{'}| = 2^{\mathcal{R}(\mathbf{L}[*, \overline{\mathcal{F}}])}$. However for the fault-free counterpart Z over $\mathbb{F}_{2^n}$, the cardinality is $|\mathcal{Z}| = 2^n$. Thus for faulty $Z^{'}$, there are $2^n - 2^{\mathcal{R}(\mathbf{L}[*, \overline{\mathcal{F}}])}$ values unreachable, yielding:

$$\begin{cases} P(Z^{'} = z^{'}) = 2^{-\mathcal{R}(\mathbf{L}[*, \overline{\mathcal{F}}])}, \text{ if } z^{'} \in \mathcal{Z}^{'} \text{ with } |\mathcal{Z}^{'}| = \mathcal{R}(\mathbf{L}[*, \overline{\mathcal{F}}]), \\ P(Z^{'} = z) = 0, \text{ if } z \in \mathcal{Z} \setminus \mathcal{Z}^{'}. \end{cases} \tag{20}$$

Therefore, a reduction in $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}])$ directly corresponds to decreased uniformity in the distribution of $Z^{'}$. This completes the proof.

E Proof of Theorem 4

Proof. As established previously, the likelihood of obtaining valid outputs $\vec{C}^{'}$ is governed by the probability $Pr[Z = Z^{'}] = Pr[\vec{L} \cdot \vec{e} = 0]$. Recall that the addition vector $\vec{e} = \vec{B} \oplus \vec{B}^{'}$ satisfies $[\vec{e}]_2[i] = 0$ if $i \notin \mathcal{F}$ and otherwise $[\vec{e}]_2[i]$ uniformly distributed over $\mathbb{F}_2$ under stuck-at faults. Hence we have

$$\begin{aligned} \vec{L} \cdot \vec{e} = [\mathbf{L_1}, \cdots, \mathbf{L_l}]_2 \times [\vec{e}]_2 &= \mathbf{L}[*, \mathcal{F}] \times [\vec{e}]_2[\mathcal{F}] \oplus \mathbf{L}[*, \overline{\mathcal{F}}] \times [\vec{e}]_2[\overline{\mathcal{F}}] \\ &= \mathbf{L}[*, \mathcal{F}] \times [\vec{e}]_2[\mathcal{F}] \oplus \vec{0} \\ &= \mathbf{L}[*, \mathcal{F}] \times [\vec{e}]_2[\mathcal{F}]. \end{aligned} \tag{21}$$

If $\vec{L} \cdot \vec{e} = 0$, then $\mathbf{L}[*, \mathcal{F}] \times [\vec{e}]_2[\mathcal{F}] = \vec{0}$ should hold. Regarding $\mathbf{L}[*, \mathcal{F}]$ as a matrix representing the linear transformation $L_{\mathcal{F}} : \mathcal{E} \to \mathcal{V}$ from $\mathbb{F}_2^{|\mathcal{F}|}$ to $\mathbb{F}_2^n$, then

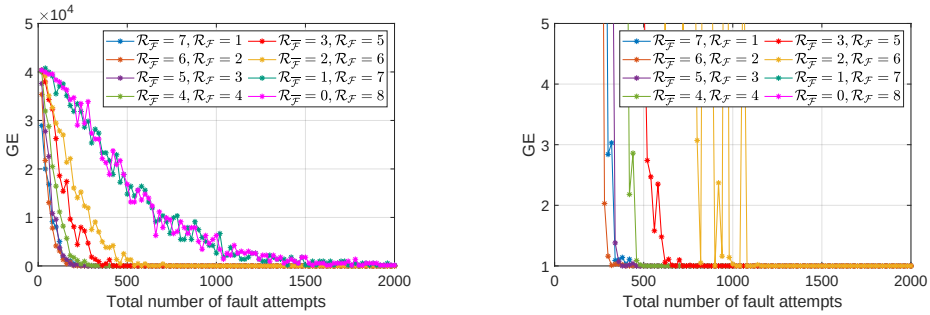
$\mathbf{L}[*,\mathcal{F}] \times [\vec{e}]_2[\mathcal{F}] = \vec{0}$ has $2^{\dim(\ker(L_\mathcal{F}))}$ solutions, by the rank-nullity theorem (that is $\dim(L_\mathcal{F}) = \dim(\ker(L_\mathcal{F})) + \mathsf{Rank}(\mathbf{L}[*,\mathcal{F}])$), we have

$$Pr[\vec{L}\cdot\vec{e}=0] = Pr[\mathbf{L}[*,\mathcal{F}]\times[\vec{e}]_2[\mathcal{F}]=\vec{0}] = \frac{2^{\dim(\ker(L_\mathcal{F}))}}{2^{|\mathcal{F}|}} = \frac{1}{2^{\mathcal{R}(\mathbf{L}[*,\mathcal{F}])}}. \quad (22)$$

This completes the proof.

F Attack Complexity Under Coupled Rank Parameters

We construct 8 fault models $\mathcal{F}$ to experimentally study how both rank parameters influence attack complexity under conditions of equal sensitivity. Specifically, for 2-share Boolean case, corrupting x-bit Least Significant Bits across all shares, we have $\mathsf{Rank}([*,\overline{\mathcal{F}}]) = n - x$ and $\mathsf{Rank}(\mathbf{L}[*,\mathcal{F}]) = x$. Figure 8 illustrates the Guessing Entropy (GE) results of SIFA-1 under stuck-at-0 faults when x increases from 1 to 8 (the attack details are elaborated in Section 3.3 of paper).



(a) GE between $[1, 50000]$. (b) GE between $[1, 5]$.

Fig. 8: GE results of 2-share Boolean instance over $\mathbb{F}_{2^8}$ with various $\mathcal{R}_{\overline{\mathcal{F}}} = \mathsf{Rank}(\mathbf{L}[*,\overline{\mathcal{F}}])$ and $\mathcal{R}_\mathcal{F} = \mathsf{Rank}(\mathbf{L}[*,\mathcal{F}])$.

G Resistance Ability of Two-share IPM Parameters

Table 4 records the resistance ability of various two-share IPM parameters, evaluated according to the coding-theoretic and rank parameters, where $\mathcal{R}_{\min}$ gives the minimum $\mathsf{Rank}(\mathbf{L}[*,\mathcal{F}])$. For each metric $d_\mathcal{L}$, the table includes both the candidate exhibiting largest and the smallest $B_{d_\mathcal{L}}$.

H Simulation of Optimal Two-share IPM Parameters

We examine the optimal two-share IPM instances identified in this paper. From the 12 candidates shown in Table 4, we select four optimal instances with

Table 4: Resistance of various two-share IPM parameters, summarized as in [12].

$d_{\mathcal{L}}$	$B_{d_{\mathcal{L}}}$	$\mathcal{R}_{\min}$	Candidates of $L[2]$
1	-	-	$\emptyset$
2	8	1	{1}
2	1	2	{16, 17, 34, 39, 60, 70, 90, 115, 116, 119, 120, 133, 140, 180, 182, 201, 207, 215, 230, 234, 247}
3	6	3	{3,83,101,137,158,166,202,246}
3	1	3	{14,15,19,20,40,44,48,49,52,56,61,67,69,75,76,80,94,97,99, 103,112,113,114,118,121,124,136,138,142,149,151,153,155, 161,169,172,173,183,185,186,187,188,189,193,199, 200,206,211,213,214,219,226,229,231,237,242,243}
4	17	3	{29, 64}
4	4	4	{23,46,51,54,81,92,95,102,108,162,165,184}

$L[2] = \{23, 54, 184, 102\}$, as documented in Table 5. Notably, the $\vec{L} = [1, 23]$ instance was previously evaluated in [42] and confirmed as the optimal two-share IPM against side-channel attacks, while the $\vec{L} = [1, 54]$ instance was originally proposed in [40]. Table 4 indicates that all optimal candidates share identical security characteristics: the minimal fault injection requirement is $d_{\mathcal{F}} = 4$ bits, with exactly $B_{d_{\mathcal{L}}} = 4$ distinct fault models achieving this minimum bit-injection cost. For each selected instance, we randomly choose two effective 4-bit fault models (reported in Table 5) and evaluate their resistance against SIFA-1 using all chosen fault models. Recall that two of the effective fault models for $\vec{L} = [1, 23]$ are illustrated in Figure 5 in the paper. The corresponding GE results are illustrated in Figure 9. As demonstrated in Figure 9, only the two identified effective fault models successfully compromise the respective optimal IPM instances, requiring approximately $3,500$ fault attempts to achieve complete key recovery. The observed attack complexity is approximately $8\times$ higher than the Boolean case (requiring approximately 400 attempts), which well matches the theoretical prediction derived from the rank conditions: with the same *uniformity degree* $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 7$, $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}]) = 4$ yields a success probability of $\frac{1}{2^4}$, one-eighth of the Boolean case's $\frac{1}{2}$ probability.

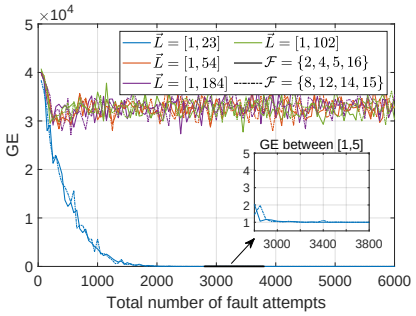
Note that our exhaustive simulations verify two critical security properties of these optimal two-share IPM instances. First, any fault model with fewer than 4 bits invariably fails to compromise the protection. Second, among all possible $C_{16}^4 = 1820$ combinations of 4-bit fault models, only 4 specific candidates prove effective. Furthermore, these four effective fault models impose stringent precision requirements for bit-level fault injection (as exemplified in Figure 5 for $\vec{L} = [1, 23]$), demanding exact targeting of specific bit positions.

I Simulated Attacks Under Constant-byte Faults

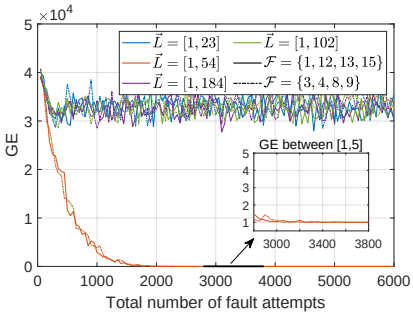
For the sake of a direct comparison with stuck-at faults, we choose the same two-share IPM instances over $\mathbb{F}_{2^8}$ evaluated in Section 3.3, containing $L[2] =$

Table 5: Rank of both $\mathbf{L}[*, \overline{\mathcal{F}}]$ and $\mathbf{L}[*, \mathcal{F}]$ under part of optimal two-share IPM parameters $\vec{L}$ and various fault models $\mathcal{F}$. $\mathsf{Rank}(\mathbf{L}[*, \overline{\mathcal{F}}]) = 8$ indicates attack failure (Theorem 1), rendering attack complexity (based on $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}])$) unnecessary, thereby marked "-".

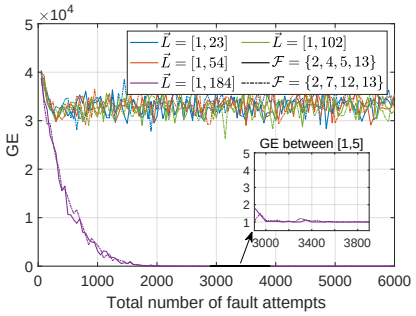
$\mathcal{F}$	$\vec{L} = [1, 23]$		$\vec{L} = [1, 54]$		$\vec{L} = [1, 184]$		$\vec{L} = [1, 102]$	
	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$	$\mathbf{L}[*, \overline{\mathcal{F}}]$	$\mathbf{L}[*, \mathcal{F}]$
$\mathcal{F} = \{2, 4, 5, 16\}$	7	4	8	-	8	-	8	-
$\mathcal{F} = \{8, 12, 14, 15\}$	7	4	8	-	8	-	8	-
$\mathcal{F} = \{1, 12, 13, 15\}$	8	-	7	4	8	-	8	-
$\mathcal{F} = \{3, 4, 8, 9\}$	8	-	7	4	8	-	8	-
$\mathcal{F} = \{2, 4, 5, 13\}$	8	-	8	-	7	4	8	-
$\mathcal{F} = \{2, 7, 12, 13\}$	8	-	8	-	7	4	8	-
$\mathcal{F} = \{3, 4, 6, 16\}$	8	-	8	-	8	-	7	4
$\mathcal{F} = \{5, 12, 13, 14\}$	8	-	8	-	8	-	7	4



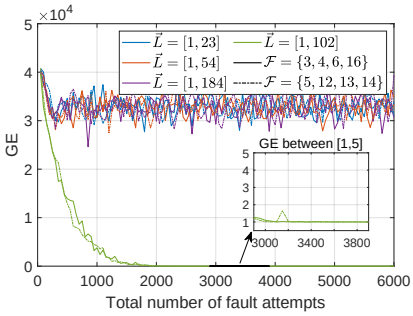
(a) GE results



(b) GE results



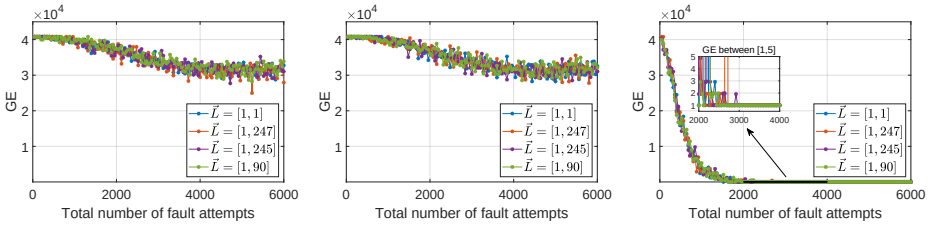
(c) GE results



(d) GE results

Fig. 9: GE results of optimal two-share IPM parameters under various fault models, with inset plots highlighting the convergence behavior as GE approaches 1.

$\{1, 247, 245, 90\}$. Given the two-share IPM scheme ($l = 2$), the constant-byte fault model yields exactly three possible fault models, including $\mathcal{F} = \{1\}$, $\mathcal{F} = \{2\}$ and $\mathcal{F} = \{1, 2\}$, with the last one corrupting all shares. Figure 10 presents the GE results of SIFA-1 attacking the selected IPM instances under all three possible fault models. The experimental results clearly demonstrate that successful attacks only occur when using the fault model that corrupts all shares ($\mathcal{F} = \{1, 2\}$), while partial corruption attempts ($\mathcal{F} = \{1\}$ and $\mathcal{F} = \{2\}$) exhibit ineffective attacks against IPM instances. Considering the attack complexity of effective fault model, it requires approximately $3,000$ fault attempts for full key recovery. Recall that its probability of obtaining valid outputs is $\frac{1}{2^8}$. While in the simulated attacks under stuck-at faults (shown in Figure 7), the same valid output probability $\frac{1}{2^8}$ corresponds to nearly $55,000$ fault attempts. The disparity in required fault attempts stems from difference in the *uniformity degree* of the distribution of vulnerable variables with ineffective faults. Compared to stuck-at faults (recall that $\mathsf{Rank}(\mathbf{L}[*, \mathcal{F}]) = 7$), this fault model yields a more non-uniform distribution of $Z^{'}$ with ineffective faults, consequently increasing its vulnerability to be compromised.



(a) $\mathcal{F} = \{1\}$ (b) $\mathcal{F} = \{2\}$ (c) $\mathcal{F} = \{1, 2\}$

Fig. 10: GE results of various IPM instances under constant-byte fault models, with inset plots highlighting the convergence behavior as GE approaches 1.

J Simulated Attacks of Gadget-independent SIFA-2

We conduct a simulated gadget-independent SIFA-2 attack targeting the non-bijective nature of IP-masked multiplication gadget in $\mathbb{F}_{2^n}$. Note that any fault injection attempt into the input will result in an ineffective fault with probability $\frac{1}{2^n}$ at the output. If utilizing the toy circuit shown in Figure 1 of the paper, the final multiplication yields a deterministic zero result. This enables direct key recovery as the adversary observes $C[1] = k[1] \oplus (0x63)$ (where 0x63 is the affine transformation of zero) and no statistical analysis is needed. Hence, we now consider the case where fault injection occurs one round earlier, as illustrated in Figure 11, featuring an AES S-box S realized through the efficient chain [31]. Specifically, S-box is decomposed into a carefully sequenced combination of finite field multiplications, squaring operations, and a final affine transformation.

The symbolic notation and operational representation maintain consistency with Figure 1. This scenario also aligns with the simulation presented in [29], where similar early-round fault injections were analyzed for SIFA-2 on prime masking.

We evaluate a two-share IPM-protected implementation over $\mathbb{F}_{2^8}$, following the construction of masked gadgets established in [2]. The attack works by injecting a single bit-flip fault into the LSB of share $B[1]$ from one input of multiplication gadget. Crucially, ineffective faults occur only when the other input equals 0, which happens with probability $\frac{1}{2^8}$. Figure 12 illustrates the GE results of SIFA-2 attacking IPM instances, including $L[2] = \{1, 247, 245, 90\}$, under such gadget-independent scenarios. As demonstrated in Figure 12, all IPM instances exhibit consistent vulnerability to SIFA-2, requiring less than $3,000$ fault attempts for full key recovery. This empirical result confirms the theoretical attack complexity of $\frac{1}{2^8}$, matching the behavior observed in the SIFA-1 model under random-byte faults (Figure 10). Note that the distribution of fault-free sensitive input variable Z_2 satisfies $Pr[Z_2 = 0] = 1$ and $Pr[Z_2 \neq 0] = 0$ under gadget-independent SIFA-2 scenarios, which shares identical *uniformity degree* with the effective SIFA-1 model under constant-byte faults.

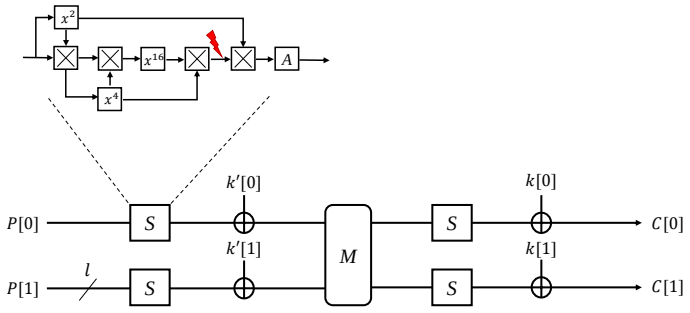
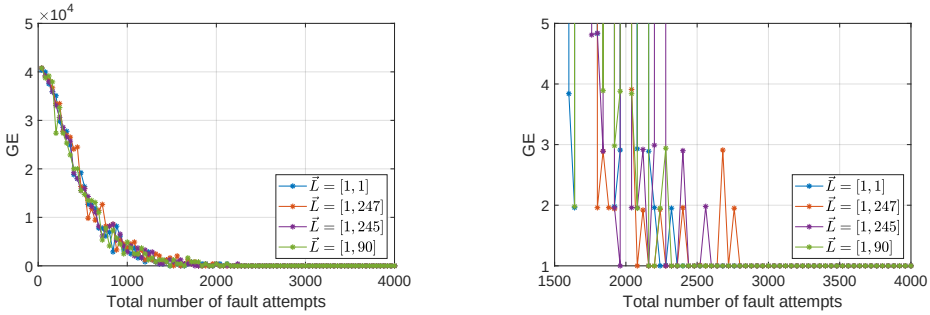


Fig. 11: Toy cipher circuit for SIFA-2 simulation, with fault location denoted by a red lightning symbol. The cross symbol and "A" represents the field multiplication and affine transformation, respectively.



(a) GE between $[1, 50000]$. (b) GE between $[1, 5]$.

Fig. 12: GE results of various IPM instances under SIFA-2 fault models in the gadget-independent scenarios.