Information Leakages in Code-based Masking: A Unified Quantification Approach

Wei Cheng[1], Sylvain Guilley[2,1], Claude Carlet[3],
Jean-Luc Danger[1,2] and Sihem Mesnager[4]

[1] LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France
{wei.cheng,jean-luc.danger}@telecom-paris.fr
[2] Secure-IC S.A.S., Tour Montparnasse (27th floor), Paris, France
sylvain.guilley@secure-ic.com
[3] LAGA, Department of Mathematics, University of Paris VIII, Paris, France and University of Bergen, Norway claude.carlet@gmail.com
[4] Department of Mathematics, University of Paris VIII, F-93526 Saint-Denis, University Sorbonne Paris Cité, LAGA, UMR 7539, CNRS, 93430 Villetaneuse and Telecom Paris, Polytechnic Institute of Paris, 91120 Palaiseau, France smesnager@univ-paris8.fr

Abstract.
This paper presents a unified approach to quantifying the information leakages in the most general code-based masking schemes. Specifically, by utilizing a uniform representation, we highlight first that all code-based masking schemes' side-channel resistance can be quantified by an all-in-one framework consisting of two easy-to-compute parameters (the dual distance and the number of conditioned codewords) from a coding-theoretic perspective. In particular, we use signal-to-noise ratio (SNR) and mutual information (MI) as two complementary metrics, where a closed-form expression of SNR and an approximation of MI are proposed by connecting both metrics to the two coding-theoretic parameters. Secondly, considering the connection between Reed-Solomon code and SSS (Shamir's Secret Sharing) scheme, the SSS-based masking is viewed as a particular case of generalized code-based masking. Hence as a straightforward application, we evaluate the impact of public points on the side-channel security of SSS-based masking schemes, namely the polynomial masking, and enhance the SSS-based masking by choosing optimal public points for it. Interestingly, we show that given a specific security order, more shares in SSS-based masking leak more information on secrets in an information-theoretic sense. Finally, our approach provides a systematic method for optimizing the side-channel resistance of every code-based masking. More precisely, this approach enables us to select optimal linear codes (parameters) for the generalized code-based masking by choosing appropriate codes according to the two coding-theoretic parameters. Summing up, we provide a best-practice guideline for the application of code-based masking to protect cryptographic implementations.

Keywords: Side-channel attacks · Leakage quantification · Signal to Noise Ratio (SNR) · Mutual Information (MI) · Inner Product Masking (IPM) · Shamir's Secret Sharing (SSS) · Generalized Code-based Masking (GCM) · Coding theory.

1 Introduction

Masking is one of the most well-studied countermeasures to protect cryptographic implementations against side-channel attacks due to the favorable provable security it provides. The core idea underlying any masking scheme is to split the sensitive (key-dependent) variables into several shares and perform independent computations on masked variables only. Indeed, the rationale is that, given a sufficient amount of noise, the attack complexity increases exponentially with the number of shares [CJRR99, PR13], while the implementation cost increases only quadratically (or only cubically in higher-order glitch-free implementations [GSF13]).

Two key ingredients of a masking scheme are the encoding for randomizing the sensitive variables, and the masked operations for manipulating the random shares. Regarding the latter, the secure masked operations can be constructed effectively [ISW03, RP10] for both bit- and word-oriented variables. Furthermore, thanks to the well-established concept of (Strong) Non-Inference (NI and SNI) introduced by Barthe et al. [BBD+16], the basic gadgets carrying out the elementary operations (e.g., addition, multiplication, etc.) can be composed to construct the whole implementation without losing the claimed security properties. Regarding the former, the encoding is a more fundamental ingredient in masking that provides the achievable upper bounds of side-channel security order with tunable public parameters. Indeed, firstly, the side-channel security order of the full implementation cannot exceed the security order of the corresponding encoding, and secondly, when implemented ideally, the security order of an implementation can be guaranteed by its encoding. However, evaluating the concrete side-channel resistance of the encoding in general cases remains an open problem since many different encodings in various masking schemes behave differently when fed with diverse parameters. Therefore, a unified quantification approach would formalize and compare the security of different encodings and find optimal parameters for a specific masking scheme.

1.1 State-of-the-Art

1.1.1 Unifying Masking Schemes by Generalization

Generalization is a promising approach to unify different masking schemes. In this trend, the code-based masking generalizes many existing schemes, including Boolean masking, Inner Product masking (IPM)[1] [BFG15, BFG+17], Leakage Squeezing (LS) [CDG+14, CG18] and Direct Sum masking (DSM) [BCC+14, PGS+17]. To the best of our knowledge, the generalized code-based masking (GCM) [WMCS20] is the most generic scheme in this respect. In particular, polynomial masking [GM11, PR11] is also a special case of GCM, which is built upon Shamir's secret sharing (SSS) scheme [Sha79].

Let $X \in \mathbb{F}_{2^\ell}^k$ and $Y \in \mathbb{F}_{2^\ell}^t$ be respectively the sensitive variable and t random masks. Then the encoded variable in GCM writes:

$$Z = X\mathbf{G} + Y\mathbf{H} \in \mathbb{F}_{2^\ell}^n,$$

given that $k + t \leq n$, where $\mathbf{G}$ and $\mathbf{H}$ are generator matrices of two codes $\mathcal{C}$ and $\mathcal{D}$, respectively. For the sake of simplicity, we take $k = 1$, but essentially, the GCM can use packed secret sharing techniques [GSF13, WMCS20] to improve the performance by parallelism. However, the side-channel security evaluation of encoding is similar to any k, since each of the k sensitive variables is encoded similarly. The overview of connections between these masking schemes is shown in Fig. 1, where the four intersecting areas are:

- Intersection I: as pointed out in [CS21], Boolean masking can be considered as a special case of polynomial masking for small enough parameters ($n \leq 6$ or equivalently $t \leq 5$).
- Intersection II: in [BFG15], the authors claimed that the polynomial masking is a special case of IPM. However, this generalization does not indicate the exact connections between SSS-scheme and RS codes. Indeed, if we take the polynomial evaluations in encoding into consideration, the generalization from SSS-based masking to IPM is valid only when $n = 2$ and $t = 1$.

[1] We consider the improved IPM [BFG15] rather than the original one [BFGV12], since firstly, there exist some first-order information leakages in the latter [PRR14], and secondly the performance of the latter is much lower than the former, which makes it impractical.

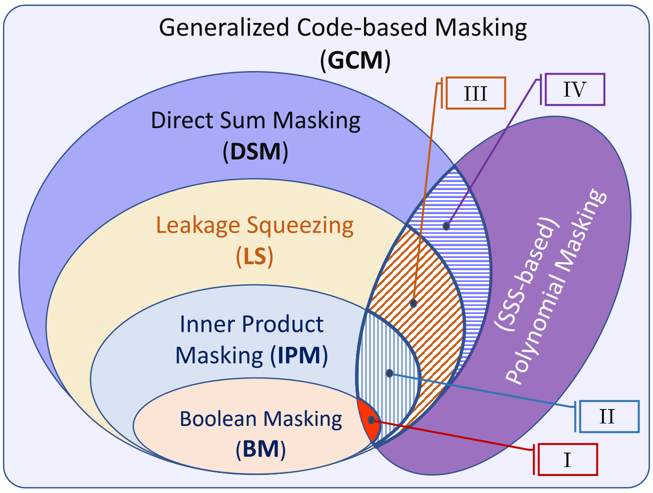

Generalized Code-based Masking (GCM)
Direct Sum Masking (DSM)
Leakage Squeezing (LS)
Inner Product Masking (IPM)
Boolean Masking (BM)
(SSS-based) Polynomial Masking
I
II
III
IV

Figure 1: Overview of code-based masking schemes. In particular, all intersections I, II, III, and IV mean that $n = t + 1$ in SSS-based masking, where the two codes $\mathcal{C}$ and $\mathcal{D}$ are complementary.

- Intersections III and IV: in SSS-based masking, if $n = t + 1$, the codes $\mathcal{C}$ and $\mathcal{D}$ are complementary, therefore they can be viewed as DSM (or LS) scheme. Otherwise, if $n > t + 1$, the corresponding masking schemes are out of DSM's scope. On the other side, the linear codes for DSM may not be converted into SSS-based schemes since the codes in SSS are endowed with a specific algebraic structure.

The most significant benefit of utilizing code-based masking is the higher security order than the simple Boolean masking given the same number of shares. Taking 2-share IPM over $\mathbb{F}_{2^8}$ [BFG$^+$17, CGC$^+$21] as an instance, when appropriate public parameters are chosen, the side-channel security order can be maximized to 3 under the bit-probing model [PGS$^+$17], which is higher than 1 in Boolean masking. Moreover, the security orders are enlarged to 7 vs. 2 (IPM vs. Boolean one) in 3-share scenarios [CGC$^+$21, Tab. 2].

Currently, the side-channel security order of GCM has been connected to the dual distance of $\mathcal{D}$ [PGS$^+$17, CG18], which is denoted as $d_{\mathcal{D}}^{\perp}$. As a special case, the security order t in IPM and DSM is equal to $d_{\mathcal{D}}^{\perp} - 1$ since the two codes $\mathcal{C}$ and $\mathcal{D}$ are complementary. However, as pointed out in [CGC$^+$21], the dual distance of $\mathcal{D}$ is not sufficient to characterize the concrete side-channel resistance of IPM, hence a new framework with a new parameter (more precisely $B_{d_{\mathcal{D}}^{\perp}}$, which counts the number of codewords of Hamming weight equal to $d_{\mathcal{D}}^{\perp}$ in $\mathcal{D}^{\perp}$) is proposed to model IPM's concrete security level more accurately. Nevertheless, this framework is not applicable to GCM since $\mathcal{C}$ and $\mathcal{D}$ may not be complementary anymore.

1.1.2 Public Points in SSS and Polynomial Masking

To construct a t-th order secure polynomial masking, a polynomial of degree t is firstly selected: $f_X(\mathsf{X}) = X + \sum_{i=1}^{t} u_i \mathsf{X}^i$, where the secret X is then associated as the constant term in $f_X(\mathsf{X})$. Secondly, $f_X(\mathsf{X})$ is evaluated in n distinct points α_i for $1 \leq i \leq n$, which are called "public points" in the scheme. As a result, the secret X is encoded by using the private parameters u_i (which are random masks viewed in the context of masking).

As observed in [CMP18], the public points in SSS play a significant role in the side-channel resistance of SSS-based masking schemes. In fact, this problem of public points is inherent in the SSS scheme and can be dated back to Massey [Mas93] who claimed that SSS scheme "*can be attacked with the well-developed tools of algebraic coding theory*". The

SSS-based masking provides a practical example whereby changing the public points in polynomial masking, the concrete security level can be significantly different.

However, to the best of our knowledge, there are neither qualitative principles for selecting good or even optimal public points in SSS-based masking nor a quantitative approach to evaluate the role of public points played in the side-channel resistance of SSS-based masking. In this paper, we propose solutions to the two problems by utilizing a coding-theoretic quantitative approach.

1.1.3 Independence Assumption behind Masking Schemes

The independence assumption is an indispensable condition behind the security proofs when extending from the probing model to the bounded moment model or noisy leakage models [BDF+17, DDF14]. For instance, if this independence condition is violated due to physical defaults (e.g., couplings through the ground or parasitic capacitances, glitches, etc.), the side-channel security order will decrease accordingly [DFS15]. However, this independence condition is essentially related to inter-share leakages from different shares in masking and treats each share as a whole.

Moreover, the independence issue also happens in intra-share cases where the leakages of different bits in the same share leak jointly. This kind of leakage is often called non-linear leakages and comes, e.g., from registers or memory units of real devices. In fact, both intra-share and inter-share independence issues can happen simultaneously. Taking AES implemented on ARM Cortex-M4 as an example, where the registers are 32-bit, and each share is in $\mathbb{F}_{2^8}$, four shares can be manipulated at the same time. Consequently, the register will leak jointly, including intra-share and inter-share leakages. To the best of our knowledge, the intra-share independence issue has not yet been studied thoroughly in the sense of security order reduction. We will show that essentially, the intra-share independence is the condition for higher security orders under the bounded moment model [BDF+17].

1.2 Our Contributions

In view of the above state-of-the-art, our contributions are threefold as follows.

A Unified Leakage Quantification Approach for GCM. We derive a closed-form expression for SNR to quantify the information leakages in GCM for any leakage functions. In particular, we present a simplified expression for the Hamming weight leakage model. In fact, this new result generalizes the framework proposed in [CGC+21] for IPM. Furthermore, we use mutual information (MI) to quantify the information leakages of GCM in an information-theoretic sense. Both SNR and MI are connected to two properties (namely the dual distance and the number of conditioned codewords) of the linear codes used in GCM. Relying on a theoretical analysis of SNR and MI, we propose a unified approach to quantify information leakage in GCM. Then we show how to select optimal codes for GCM by optimizing the two properties. The experimental results confirm that the MI can be minimized by utilizing optimal codes, which indicates the improved concrete security level of the corresponding masking scheme.

Optimal Public Points for SSS-based Polynomial Masking. As an application of our unified approach, we characterize the side-channel resistance of polynomial masking from a coding-theoretic point of view. The first outcome is a more accurate characterization of information leakage and the second outcome is a straightforward method to choose optimal linear codes (parameters) for SSS-based masking. For the first time, we quantify the impact of combining different public points in SSS-based masking in the context of side-channel analysis and show that more shares leak more information (given a specific t). In particular, our coding-theoretic approach can exactly depict the observations made in [CMP18]. Using MI, we present the quantitative results of information leakages in SSS-based masking, which again validate our unified approach. For the first time, we exhibit several optimal tuples of public points (the linear codes in a coding-theoretic perspective) for SSS-based masking in the sense of side-channel resistance.

Revisited Independence Condition in Masking Schemes. Independence condition requires that the information leakages from different variables are statistically independent. In the context of masking, it exists in two cases: inter-share and intra-share. Specifically, the former means that the leakages of different shares are independent, which is well-studied in literature [BDF$^+$17]. The latter deals with the leakages from one share, in which different bits in this share may leak independently or not. To capture both of them, we introduce the leakage function P, where its numerical degree indicates both cases' independence conditions. For instance, the commonly assumed Hamming weight leakage model has a numerical degree equal to 1, a perfect independent case. Moreover, we show how the degree of P affects the side-channel security order of a masking scheme.

We underline that all mathematical derivations presented in this paper have been verified formally with `Magma` computational algebra system [Uni]. The open sources of this paper are available on `Github` [CG20].

1.3 Difference between this work and [CGC$^+$21]

In this work, we study GCM by using a similar coding-theoretic approach as in [CGC$^+$21]. However, two key differences make this work significantly different from [CGC$^+$21].

Firstly, GCM generalizes IPM by allowing $\mathcal{C}$, and $\mathcal{D}$ to be non-complementary, which allows deriving security metrics in a more general manner. In [CGC$^+$21], the authors prove that the side-channel security of IPM only depends on the code $\mathcal{D}$. While in this work, for the first time, we show that the side-channel security depends on both $\mathcal{C}$ and $\mathcal{D}$. In particular, the quantitative findings enable us to put forward optimal GCM encodings which are new upon [CS21]: given the same parameters n and t (the number of shares and security order), we decrease the information leakage in GCM to the lesser possible extent.

Secondly, GCM allows for protections in much more general contexts. Namely, GCM can be used to withstand glitches [PR11] and to detect errors against fault injection attacks on top of preventing side-channel attacks. Therefore, our work has broader implications for the protection of realistic platforms. In a nutshell, GCM opens a new path to derive unified countermeasures against both fault injection and side-channel attacks.

2 Preliminaries

2.1 Encoding in Code-based Masking

Let n, k be positive integers and $\mathbb{K} = \mathbb{F}_{2^\ell}$ be a finite field. Let $\mathcal{C}$ be an $[n, k]_q$ linear code parameter with generator matrix $\mathbf{G}$ defined over $\mathbb{F}_q$ (here we use $q = 2^\ell$). Let the irreducible polynomial be $g(\alpha) = \alpha^8 + \alpha^4 + \alpha^3 + \alpha^2 + 1$ to generate the field $\mathbb{K} = \mathbb{F}_{2^8}$. Recall that for an $(n,\, t)$-SSS scheme, the secret X is split into n shares, and the sharing is t-privacy, where any $t + 1$ shares can be used to recover the secret but not for less than t shares. Note that the $(n,\, t)$-SSS scheme is also connected to the Reed-Solomon (RS) code with parameters $[n, t + 1]$.

Let $X \in \mathbb{K}^k$, $Y \in \mathbb{K}^t$ and $Z \in \mathbb{K}^n$ be the sensitive variable, the random masks, and the shared variable; we use Eqn. 1 as the uniform representation of encoding in GCM which is used throughout the paper:

$$Z = X\mathbf{G} + Y\mathbf{H} \in \mathbb{F}_{2^\ell}^n, \tag{1}$$

where $k + t \leq n$, $\mathbf{G}$ and $\mathbf{H}$ are two generator matrices of the two codes $\mathcal{C}$ and $\mathcal{D}$ with $\mathcal{C} \cap \mathcal{D} = \{0\}$.

In this paper, we focus on GCM, which is the most general case of code-based maskings[2]. By using the uniform representation as Eqn. 1, we revisit the encodings of code-based masking schemes as in Tab. 1.

Table 1: Encodings in IPM, LS, DSM, SSS-based masking and GCM, revisited.

	IPM [BFG15, BFG+17]	LS [3] [CDG+14]	DSM [BCC+14, PGS+17]	SSS-based masking [GM11, PR11]	**GCM** [WMCS20]
Conditions on $\mathcal{C}$ and $\mathcal{D}$	$\mathcal{C}\cap\mathcal{D}=\{0\}$, $\mathcal{C}+\mathcal{D}=\mathbb{K}^n$	$\mathcal{C}\cap\mathcal{D}=\{0\}$, $\mathcal{C}+\mathcal{D}=\mathbb{K}^n$	$\mathcal{C}\cap\mathcal{D}=\{0\}$, $\mathcal{C}+\mathcal{D}=\mathbb{K}^n$	$\mathcal{C}\cap\mathcal{D}=\{0\}$	$\mathcal{C}\cap\mathcal{D}=\{0\}$
$\mathbf{G}\in\mathbb{K}^{k\times n}$	$\begin{pmatrix}1 & 0 & 0 & \cdots & 0\end{pmatrix}$	$\mathbf{G}\in\mathbb{K}^{k\times n}$	$\mathbf{G}\in\mathbb{K}^{k\times n}$	$\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}$	$\mathbf{G}\in\mathbb{K}^{k\times n}$
$\mathbf{H}\in\mathbb{K}^{t\times n}$	$\begin{pmatrix}\alpha_1 & 1 & 0 & \cdots & 0\\ \alpha_2 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \alpha_t & 0 & 0 & \cdots & 1\end{pmatrix}$	$\mathbf{H}\in\mathbb{K}^{t\times n}$	$\mathbf{H}\in\mathbb{K}^{t\times n}$	$\begin{pmatrix}\alpha_1^1 & \alpha_2^1 & \cdots & \alpha_n^1\\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_1^t & \alpha_2^t & \cdots & \alpha_n^t\end{pmatrix}$	$\mathbf{H}\in\mathbb{K}^{t\times n}$
Security parameters: n, k, t	$k=1$, $n=t+1$	$n=k+t$. $\mathbf{G}$, $\mathbf{H}$ can be any matrices	$n=k+t$. $\mathbf{G}$, $\mathbf{H}$ can be any matrices	$n\geq k+t$ and $f_X(\mathsf{X})$. In glitch-free case, $n\geq 2t+1$ [PR11]	$n\geq k+t$. $\mathbf{G}$, $\mathbf{H}$ can be any matrices

2.2 Linear Codes

We recall several known definitions and properties on linear codes, which hold respectively when the base field is $\mathbb{K}=\mathbb{F}_2$ or $\mathbb{K}=\mathbb{F}_{2^\ell}$. Given a linear code $\mathcal{C}$ with parameters $[n,k,d_\mathcal{C}]$ where $d_\mathcal{C}$ is the minimum distance, its weight enumerator is defined as follows.

Definition 1 (Weight Enumerator [MS77, §5.2])**.** The weight enumerator of a linear code specifies the number of codewords $\mathcal{C}$ of each possible Hamming weight in $\mathcal{C}$. Specifically, we have

$$W_\mathcal{C}(\mathsf{X},\mathsf{Y})=\textstyle\sum_{i=0}^{n} B_i\mathsf{X}^{n-i}\mathsf{Y}^i \tag{2}$$

where $B_i=|\{c\in\mathcal{C}|w_H(c)=i\}|$ and $w_H(\cdot)$ denotes the Hamming weight function.

For instance, given a linear code $\mathcal{C}$ we have $B_0=1$, $B_1=\cdots=B_{d_\mathcal{C}-1}=0$, $B_{d_\mathcal{C}}>0$.

Note that two linear codes are said to be equivalent if one can be obtained from the other by a series of operations of the following two types: 1) an arbitrary permutation of the coordinate positions and 2) in any coordinate position, multiplication by any nonzero scalar. Straightforwardly, equivalent linear codes have the same weight enumerator.

Definition 2 (Dual Code [MS77, §1.8])**.** The dual code of $\mathcal{C}$ is the linear code $\mathcal{C}^\perp=\{u\in\mathbb{K}^n|\forall c\in\mathcal{C},\ c\cdot u=0\}$.

Definition 3 (Dual Distance [MS77])**.** The dual distance $d_\mathcal{C}^\perp$ of a linear code $\mathcal{C}$ is the minimum Hamming weight $w_H(u)$ of nonzero $u\in\mathbb{K}^n$, such that $\sum_{c\in\mathcal{C}}(-1)^{c\cdot u}\neq 0$.

Corollary 1. *For a linear code $\mathcal{C}$, we have $d_\mathcal{C}^\perp=d_{\mathcal{C}^\perp}$.*

According to [MP13, Theorem 5.1.18], there exists a self-dual basis of $\mathbb{F}_{q^\ell}$ over $\mathbb{F}_q$ if and only if either q is even or both q and ℓ are odd. We call this a sub-field representation.

Definition 4 (Sub-field representation [MS77, §7.7])**.** Let $x\in\mathbb{F}_{2^\ell}$, the sub-field representation of x is $[x]_2\in\mathbb{F}_2^\ell$.

[2] As a special case of IPM, a Boolean masking can be obtained by taking $\alpha_i=1$ for $1\leq i\leq t$ in Tab. 1.

[3] LS consists of the application of an arbitrary bijection on the shares. Although it has only been studied on vectors of bits (on $\mathbb{F}_2$), it can be trivially extended to vectors on $\mathbb{F}_{2^\ell}$. When the bijections are linear, LS is thus equivalent to DSM.

Definition 5 (Code Expansion [MS77, §7.7])**.** By using sub-field representation, the elements in $\mathbb{F}_{2^\ell}$ are decomposed over $\mathbb{F}_2$. Consider a generating matrix of a linear code of size $k \times n$ in $\mathbb{F}_{2^\ell}$. It becomes a generating matrix of size $k\ell \times n\ell$ in $\mathbb{F}_2$. Any linear codes of parameters $[n, k]_{2^\ell}$ contain $(2^\ell)^k = 2^{k\ell}$ codewords, hence is turned into a $[n\ell, k\ell]_2$ linear code in $\mathbb{F}_2$. The latter code is called the expansion code of the former.

Summing up, the two definitions build a direct link between word- and bit-level representation. This allows to connect the word (or register)-level probing and the bit-level probing security models, depending on the granularity of the attacker spying tool.

2.3 Properties of Complementary Space Vectors

In this subsection, we derive relevant properties of complementary space vectors that will be needed to derive our results. Let $\mathcal{E}$ a space vector of $\mathbb{K}^n$. The indicator of $\mathcal{E}$ is the application

$$x \in \mathbb{K}^n \mapsto \mathbb{1}_{\mathcal{E}}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 1. *Let $\mathcal{C}$ and $\mathcal{D}$ be two space vectors in $\mathbb{K}^n$ built from independent bases, meaning that $\mathcal{C} \cap \mathcal{D} = \{0\}$. Then $\mathcal{C}^\perp \cap \mathcal{D}^\perp = (\mathcal{C} \oplus \mathcal{D})^\perp$.*

Proof. First of all, we notice that $(\mathcal{C} \oplus \mathcal{D})^\perp \subseteq \mathcal{C}^\perp$. Indeed, a vector orthogonal to all vectors of $\mathcal{C} \oplus \mathcal{D}$ is in particular orthogonal to all vectors of $\mathcal{C} + 0 = \mathcal{C}$. In a symmetric way, we have that $(\mathcal{C} \oplus \mathcal{D})^\perp \subseteq \mathcal{D}^\perp$. Therefore, $(\mathcal{C} \oplus \mathcal{D})^\perp \subseteq \mathcal{C}^\perp \cap \mathcal{D}^\perp$.

Let us now prove the converse inclusion. Let $x \in \mathcal{C}^\perp \cap \mathcal{D}^\perp$. For any vector y in $\mathcal{C} \oplus \mathcal{D}$, there exists a unique pair $(c, d) \in \mathcal{C} \times \mathcal{D}$ (owing to the complementarity of space vectors $\mathcal{C}$ and $\mathcal{D}$), such that $y = c + d$. Now, $x \cdot y = x \cdot (c + d) = x \cdot c + x \cdot d = 0 + 0 = 0$. Indeed, $x \cdot c = 0$ because $x \in \mathcal{C}^\perp$ and $x \cdot d = 0$ because $x \in \mathcal{D}^\perp$. Therefore, we also have $\mathcal{C}^\perp \cap \mathcal{D}^\perp \subseteq (\mathcal{C} \oplus \mathcal{D})^\perp$. □

Lemma 2. *Let $\mathcal{C}$ and $\mathcal{D}$ two complementary space vectors, namely: $\mathcal{C} \cap \mathcal{D} = \{0\}$, and $\mathcal{C} \oplus \mathcal{D} = \mathbb{K}^n$. Then we have: $\mathcal{C}^\perp \cap \mathcal{D}^\perp = \{0\}$.*

Proof. By application of Lemma 1, we have that $\mathcal{C}^\perp \cap \mathcal{D}^\perp = (\mathcal{C} \oplus \mathcal{D})^\perp = (\mathbb{K}^n)^\perp$. Now, as $\mathbb{K}^n$ is the universe code, we have $(\mathbb{K}^n)^\perp = \{0\}$. □

In the remainder of this paper, we consider two cases:

- In GCM as a general case: $\mathcal{C} \cap \mathcal{D} = \{0\}$, and $\mathcal{C} \oplus \mathcal{D} \subseteq \mathbb{K}^n$. The redundant case $n > t + 1$ corresponds to the strict condition: $\mathcal{C} \oplus \mathcal{D} \subsetneq \mathbb{K}^n$ and then $\{0\} \subsetneq \mathcal{C}^\perp \cap \mathcal{D}^\perp$.
- In IPM or DSM as special cases: $\mathcal{C} \cap \mathcal{D} = \{0\}$, and $\mathcal{C} \oplus \mathcal{D} = \mathbb{K}^n$. This is the case of [CGC+21], where we have $\mathcal{C}^\perp \cap \mathcal{D}^\perp = \{0\}$ as shown in Lemma 2.

2.4 Basic Properties of Pseudo-Boolean Functions

Leakage functions turn a bitvector into a real value, which the attacker measures. Those functions are pseudo-Boolean functions $P : \mathbb{K}^{n\ell} \mapsto \mathbb{R}$, where $\mathbb{K} = \mathbb{F}_2$.

It is well-known that a pseudo-Boolean function P can be *uniquely* expressed in a *monomial basis* [CG99] called *Numerical Normal Form* (NNF) [NS94]:

$$P(Z) = \sum_{I \in \{0,1\}^{n\ell}} \beta_I Z^I,$$

where $Z^I = \prod_{i \in \{1,\ldots,n\ell\} \text{ s.t. } I_i = 1} Z_i$, and $\beta_I \in \mathbb{R}$. For instance, $Z^{(000\cdots0)_2} = 1$, $Z^{(100\cdots0)_2} = Z_1$ and $Z^{(110\cdots0)_2} = Z_1 Z_2$. We recall some basics of P as follows.

Definition 6 (Numerical Degree [CG99])**.** The numerical degree of a pseudo-Boolean function P denoted by $\deg(P)$ equals: $\deg(P) := d = \max\{w_H(I) | \beta_I \neq 0\}$.

Definition 7 (Fourier Transform)**.** The Fourier transform of a pseudo-Boolean function $P : \mathbb{K}^{n\ell} \mapsto \mathbb{R}$ is denoted by $\widehat{P} : \mathbb{K}^{n\ell} \mapsto \mathbb{R}$, and is defined as: $\widehat{P}(z) = \sum_{y \in \mathbb{K}^{n\ell}} P(y)(-1)^{y \cdot z}$.

Recall from [CG99, Car10] that, $\widehat{P}(z) = (-1)^{w_H(z)} \sum_{I \subseteq \{1,\dots,n\ell\}; supp(z) \subseteq I} 2^{n\ell - |I|} \beta_I$ where $\beta_I = 2^{-n\ell}(-2)^{|I|} \sum_{z \in \mathbb{F}_2^{n\ell}; I \subseteq supp(z)} \widehat{P}(z)$.

2.5 Connecting SSS Scheme to the RS code

We recall the $(n,\, t)$-SSS scheme by mainly referring to [CMP18, CRZ13]. Let $X \in \mathbb{K}$ again be the secret and can be split into n shares such that no tuple of shares with cardinality lower than t depends on X. The SSS scheme consists in selecting a random polynomial $f_X(\mathsf{X}) \doteq X + \sum_{i=1}^{t} u_i \mathsf{X}^i$ of degree t where u_i with $1 \leq i \leq t$ are t random coefficients (masks) in $\mathbb{K}$. The secret X is the constant term: $X = f_X(0)$. Then a (n, t)-sharing $(Z_1, Z_2, \dots, Z_n)$ of X is defined by evaluating the polynomial $f_X(\mathsf{X})$ in n distinct public non-zero points $\alpha_1, \alpha_2, \dots, \alpha_n$ in $\mathbb{K}$ such that $Z_i = f_X(\alpha_i)$. The recovery of X from its sharing consists in two steps: $f_X(\mathsf{X})$ is first recovered by using the Lagrange interpolation and second, $f_X(\mathsf{X})$ is evaluated in 0. Since in an $(n,\, t)$-SSS, any tuple of shares with cardinality greater than t can be used to recover X, we denote by U the selected shares ($|U| \geq t+1$), which is called the interpolation set.

Next, we recall the Reed-Solomon codes.

Definition 8 (Reed-Solomon Code [CMP18])**.** The Reed-Solomon code $\mathrm{RS}(\mathcal{S}, t+1) \subset \mathbb{K}^n$ of dimension $t+1$ over a finite field $\mathbb{K}$ and with evaluation subset $\mathcal{S} = \{\alpha_0, \alpha_1, \dots, \alpha_n\}$ of $\mathbb{K}$ is the subspace:

$$RS(\mathcal{S}, t+1) = \{(f(\alpha_0), f(\alpha_1), \dots, f(\alpha_n)); f(\mathsf{X}) \in \mathbb{K}[\mathsf{X}] \text{ and } \deg(f) \leq t\} \ .$$

Given the degree of $f(\mathsf{X})$ is t, then $t+1$ evaluations of it can be used to recover $f(\mathsf{X})$ itself and the codewords. In terms of RS codes, the sharing of X with SSS scheme is an encoding with a RS code RS $(\{\alpha_1, \dots, \alpha_n\}, t+1)$:

$$Z = (Z_1, Z_2, \dots, Z_n) = (X,\, Y) \begin{pmatrix} \mathbf{G} \\ \mathbf{H} \end{pmatrix} = X\mathbf{G} + Y\mathbf{H}, \tag{3}$$

where $\begin{pmatrix} \mathbf{G} \\ \mathbf{H} \end{pmatrix}$ is the generator matrix $(\alpha_i^j)_{i \in [1;\, n],\, j \in [0;\, t]}$. More precisely, $\mathbf{G}$ is an 1-by-n matrix equal to $(1, 1, \dots, 1)$ and $\mathbf{H}$ is a Vandermonde matrix. By denoting $\mathbf{G}_i$ and $\mathbf{H}_i$ the i-th column of $\mathbf{G}$ and $\mathbf{H}$ respectively, we have: $Z_i = f_X(\alpha_i) = X + \sum_{j=1}^{t} Y_j \alpha_i^j = X\mathbf{G}_i + (Y_1, \dots, Y_t)\ \mathbf{H}_i$.

Accordingly, the reconstruction of X from $Z = (Z_1, Z_2, \dots, Z_n)$ is done by taking Z_i to obtain an interpolation set U such that $|U| \geq t+1$. We also call this scheme the redundant sharing when $n > t+1$ since at least $t+1$ shares can recover X. We will show in Sec. 5 that more redundancies in sharing of SSS-based masking leak more information on X.

3 Quantifying Information Leakages in GCM

In this section, we use SNR as a leakage metric to evaluate the information leakages in GCM. In particular, SNR quantifies the key-dependent leakage at certain degrees. SNR is thus attractive in that if SNR at a given degree d is null, then one can conclude that the scheme is secure at order d.

3.1 Uniform Representation of Leakage Function

As the first step, we formalize the information leakages from a device. In this respect, we rely on the clarification on serial and parallel implementations proposed in [BDF+17].

Before formalization, we give an example to provide some intuition for the uniform leakage function P. Let $Z = (Z_1, Z_2, \dots, Z_n)$ denote the encoded intermediate with n shares and X be the secret. By ignoring the noise, we assume the leakage of each share is $\mathcal{L}_i = Z_i$ under the identity leakage model and $\mathcal{L} = \sum_i \mathcal{L}_i$ is the total leakage. To launch a successful attack, an adversary needs to find the (smallest) key-depend statics, namely raising d such that $\mathbb{E}\left[\mathcal{L}^d|X\right] \neq \mathbb{E}\left[\mathcal{L}^d\right]$, but $\mathbb{E}\left[\mathcal{L}^i|X\right] = \mathbb{E}\left[\mathcal{L}^i\right]$ for all $i < d$. Equivalently, an adversary needs the smallest d such that $\mathbb{V}\left[\mathbb{E}\left[\mathcal{L}^d|X\right]\right] \neq 0$, which measures the informative part in $\mathcal{L}$.

Formally, let $P = \varphi_P \circ \phi_P$ denote the leakage function, where ϕ_P is the leakage model for each share, and φ_P is the combination function that assembles the leakages from selected shares. In this paper, we call ϕ_P and φ_P the intra-share and inter-share leakage model, respectively. For instance, in serial implementations, the leakage of each share is: $\mathcal{L}_i = \phi_P(Z_i) + N_i$, then the exploitable leakages can be combined by φ_P. For instance, taking the Hamming weight model and centered product as leakage model and combination function, respectively, then $\mathcal{L}_i = \phi_P(Z_i) + N_i = w_H(Z_i) + N_i$ and $\mathcal{L} = \prod_{c=1}^{d}(\mathcal{L}_c - \mathbb{E}\left[\mathcal{L}_c\right]) = P(Z) + N_{total}$ where the latter combines leakages of d shares by the normalized product. Consequently, the highest order of key-dependent leakages is captured by P with numerical degree d.

Therefore, we use the following representation of P as a pseudo-Boolean function:

$$P(Z) = \sum_{I \in \{0,1\}^{n\ell}} \beta_I Z^I, \tag{4}$$

where $Z^I = \prod_{i \in \{1,\dots,n\ell\} \text{ s.t. } I_i = 1} Z_i$, and $\beta_I \in \mathbb{R}$ and $\deg(P) = \max\{w_H(I) \,|\, \beta_I \neq 0\}$.

Two Probing Models. For the purpose of a finer-grain analysis, we clarify the two kinds of probing model (see also [DGH+18, §2.2]) and corresponding security orders as follows:

- **Bit-probing model**: each probe only gets one bit at a time where each bit leaks independently or jointly. Correspondingly, ϕ_P is defined at bit-level and φ_P at certain degrees are used to combine the bit-level leakages. The security order under the bit-probing model is denoted by t_b.
- **Word-probing model**: each probe gets an ℓ-bit word at a time, where an ℓ-bit variable leaks as a whole. As a result, the degree of ϕ_P implies how many numbers of bits leaked jointly, in which the intra-share independence condition plays a role in security order reduction, as shown above. Similarly, the security order is then denoted by t_w.

When connected to coding-theoretic properties, the security orders t_b and t_w are related to the dual distance of the code $\mathcal{D}$ used in GCM over $\mathbb{F}_2$ and $\mathbb{F}_{2^\ell}$, respectively [PGS+17, CGC+21]. More precisely[4], we have $t_w = d_{\mathcal{D}}^{\perp} - 1$ and $t_b = d_{\mathcal{D}_2}^{\perp} - 1$ where $\mathcal{D}_2$ is the sub-field representation of $\mathcal{D}$. In the sequel, we call t the security order for the sake of simplicity, t_b and t_w should be unambiguous from the context (e.g., variables in $\mathbb{F}_2$ or $\mathbb{F}_{2^\ell}$).

3.2 SNR-based Information Leakage Quantification

Let $P(Z)$ be a leakage function as in Eqn 4 and let N denote the independent noise with zero mean and variance $\mathbb{V}[N] = \sigma_{total}^2 \propto \sigma^{2d}$ ($\propto$ means proportional to σ^{2d}) [CGC+21].

[4] In [WMCS20], a special case is presented with $t > d_{\mathcal{D}}^{\perp} - 1$. However, we always have $t = d_{\mathcal{D}}^{\perp} - 1$ if the optimal codes are used in GCM. Especially, the equality holds for all RS codes in SSS-based masking.

Then, the leakage is:

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

We have $\mathbb{V}\left[\mathbb{E}\left[P(Z) + N|X\right]\right] = \mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]$, where $Z = X\mathbf{G} + Y\mathbf{H} \in \mathbb{K}^n = \mathbb{F}_{2^\ell}^n$ is the encoding in GCM (Equ. 1). The SNR of leakages is defined as:

$$\mathrm{SNR} = \frac{\mathbb{V}\left[\mathbb{E}\left[\mathcal{L}|X\right]\right]}{\mathbb{V}\left[N\right]} = \frac{\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]}{\sigma_{total}^2}. \tag{5}$$

Therefore, we propose the following theorem to quantify the leakages in the GCM scheme by SNR.

Theorem 1. *Let a device be protected by the GCM scheme as $Z = X\mathbf{G} + Y\mathbf{H}$. Assume the leakages of the device can be represented in the form: $\mathcal{L} = P(Z) + N$. Then the SNR of the exploitable leakages is:*

$$SNR = \frac{\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]}{\sigma_{total}^2} = \frac{1}{2^{2n\ell} \cdot \sigma_{total}^2} \left(\sum_{x,\, y \in \mathcal{D}^\perp \setminus \mathcal{C}^\perp;\ x+y \in \mathcal{C}^\perp} \widehat{P}(x)\widehat{P}(y) \right), \tag{6}$$

where $\sigma_{total}^2 \propto \sigma^{2d}$ is the total noise and $\widehat{P}(\cdot)$ is the Fourier transform *of $P(\cdot)$*

The demonstration of Theorem 1 involves computing $\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]$, which can be derived by the following Lemma 3. In order to have the paper read fluently, its proof is relegated in Appendix A.1 which also proves Theorem 1.

Lemma 3. *Let a pseudo-Boolean function $P(Z)$ denote the leakage function, and taking the same notations as above, we have*

$$\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right] = \frac{1}{2^{2n\ell}} \sum_{x,\, y \in \mathcal{D}^\perp \setminus \mathcal{C}^\perp;\, x+y \in \mathcal{C}^\perp} \widehat{P}(x)\widehat{P}(y). \tag{7}$$

Remark 1. Note that Lemma 3 encompasses the core result in [CGC+21]. Indeed, as a special case, if $n = t + 1$ in SSS-based masking, the two codes $\mathcal{C}$ and $\mathcal{D}$ are complementary, as well as $\mathcal{C}^\perp$ and $\mathcal{D}^\perp$. Since by Lemma 2, we have $\mathcal{C}^\perp \cap \mathcal{D}^\perp = \{0\}$ and the only possible solution in Eqn. 7 is $x = y \neq 0$. Therefore, $\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]$ can be simplified into:

$$\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right] = \frac{1}{2^{2n\ell}} \sum_{x \in \mathcal{D}^\perp \setminus \{0\}} \widehat{P}(x)^2, \tag{8}$$

which is exactly the same result as in [CGC+21].

As a nutshell, the information leakages from GCM can be quantified by Theorem 1 under the generic leakage model characterized by P, which evaluates the SNR of the leakages. As a direct result, we have the following proposition, which connects the code property $d_\mathcal{D}^\perp$ and the security order in GCM.

Proposition 1. *The GCM is secure at the order $t = d_\mathcal{D}^\perp - 1$ under the bounded moment model and the probing model if* $\deg(P) < d_\mathcal{D}^\perp$.

Proof. Given a pseudo-Boolean function P, one has $\widehat{P}(z) = 0$ for all $z \in \mathbb{K}^n$ such that $w_H(z) > \deg(P)$ [CG99]. As a result, SNR will be zero since $\deg(P) < d_\mathcal{D}^\perp$ and all codewords of $\mathcal{D}^\perp \setminus \mathcal{C}^\perp$ as in Eqn. 6 have Hamming weight no less than $d_\mathcal{D}^\perp$. □

Consequently, the attacks on GCM fail if $\deg(P) < d_\mathcal{D}^\perp$. Conversely, for an attack to succeed, one must have $\deg(P) \geq d_\mathcal{D}^\perp$. This is, however, only a necessary condition, but not a sufficient one. Indeed, it is possible that attacks in the setting $\deg(P) \geq d_\mathcal{D}^\perp$ fail. This is illustrated in the next remark.

Remark 2. The security order can be even higher than $d^{\perp}_{\mathcal{D}}-1$ when there is no $x,\, y \in \mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp}$ such that $x+y \in \mathcal{C}^{\perp}$ which have weight $d^{\perp}_{\mathcal{D}}$. Indeed, in Eqn. 6, the sum will be empty if the degree of P is equal to $\deg(P) = d^{\perp}_{\mathcal{D}}$. Thus the SNR is equal to zero, and the security order increases accordingly. A specific example can be found in [WMCS20, Example 1] (shown in Appendix B.5), in which $d^{\perp}_{\mathcal{D}}$ equals 2 and the security order equals 2 as well.

4 Quantifying Hamming Weight Leakages

One realistic leakage model is the so-called "Hamming weight" leakage: each bit is leaking in a similar amount, though independently from others. It has been demonstrated to be practical in many works, such as [BCO04]. In this case, the attacker can measure a quantity $P(Z) = w_H(X\mathbf{G} + Y\mathbf{H})$. However, $\mathbb{E}\left[P(Z)|X\right] = \mathbb{E}\left[P(Z)\right]$ if the masking is perfect. But there exists a $d > 1$ such that for some x, $\mathbb{E}\left[P(Z)^d|X=x\right] \neq \mathbb{E}\left[P(Z)^d\right]$.

4.1 Simplifications

We use $P(z) = w_H(z)^d$ as the informative part in a leakage model, which captures the higher-order leakages where the numerical degree $\deg(P)$ equals d. Moreover, we have:

$$P(z) = \sum_{J_1+\cdots+J_{n\ell}=d} \binom{d}{J_1,\ldots,J_{n\ell}} \prod_{i=1}^{n\ell} z_i^{J_i} = \sum_{\substack{J\in\mathbb{N}^{n\ell},\text{ s.t. } w_H(J)<d;\\ \sum_{i=1}^{n\ell} J_i=d}} \binom{d}{J} z^J + d! \sum_{\substack{I\in\{0,1\}^{n\ell};\\ w_H(I)=d}} z^I \tag{9}$$

where $\mathbb{N} = \{0,1,\ldots\}$ is the set of integers. The multinomial coefficient $\binom{d}{J_1,\ldots,J_{n\ell}}$ is defined as $\frac{d!}{J_1!\cdots J_{n\ell}!}$ (recall that $J = (J_1,\ldots,J_{n\ell}) \in \mathbb{N}^{n\ell}$ with $\sum_{i=1}^{n\ell} J_i = d$). This coefficient equals $d!$ as long as for all i $(1 \leq i \leq n\ell)$, $J_i = 0$ or 1. Now, the terms in $P(z)$ are categorized into two cases:

- z^J where $J \in \mathbb{N}^{n\ell}$, $w_H(J) < d$, which consists in products of $< d$ bits of z, as $z^J = \prod_{i\in\{1,\ldots,n\ell\}\ s.t.\ J_i>0} z_i$,
- z^I where $I \in \{0,1\}^{n\ell}$, $w_H(I) = d$ which consists in products of d bits of z, as $z^I = \prod_{i\in\{1,\ldots,n\ell\}\ s.t.\ I_i=1} z_i$.

Indeed, let $i \in \{1,\ldots,n\ell\}$, then $z_i^{J_i} = 1$ if $J_i = 0$, and $z_i^{J_i} = z_i$ if $J_i > 0$. The first terms z^J have numerical degree $\deg(z^J) < d$, hence can be discarded in the analysis (they contribute nothing to the SNR). Remaining terms of numerical degree d are: $\sum_{I\in\{0,1\}^{n\ell},\, w_H(I)=d} z^I$.

Relying on decomposition in Eqn. 9, we can simplify lemma 3 as follows.

Lemma 4. *Let a pseudo-Boolean function $P(Z) = w_H(Z)$ denote the leakage function, and taking the same notations as above, we have*

$$\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right] = B'_d \left(\frac{d!}{2^d}\right)^2. \tag{10}$$

where B'_d denotes the adjusted coefficient in weight enumerator which is defined in Def. 9.

Before diving into the proof of Lemma 4, we define the parameter $B'_{d^{\perp}_{\mathcal{D}}}$ which count the number of codewords under certain conditions in $\mathcal{C}^{\perp}$ and $\mathcal{D}^{\perp}$.

Definition 9 (Adjusted coefficient in weight enumerator)**.** Let $\mathcal{C}$ and $\mathcal{D}$ denote two linear codes. The adjusted coefficient B'_d is defined as:

$$B'_d = \left|\{(x,y) \in (\mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp})^2 \,|\, x+y \in \mathcal{C}^{\perp},\ w_H(x) = w_H(y) = d\}\right|. \tag{11}$$

To be more precise, we use subscript 2 (if necessary) to indicate the subfield representation of a linear code. For instance, $\mathcal{D}_2$ denotes the subfield representation of $\mathcal{D}$ over $\mathbb{F}_2$. Therefore, we have the following lemma for B'_d.

Lemma 5. *Recall that $B_{d^{\perp}_{\mathcal{D}_2}}$ is the coefficient in weight enumerator of $\mathcal{D}_2^{\perp}$ defined in Def. 1, then we have the following inequality in SSS-based masking:*

$$B'_{d^{\perp}_{\mathcal{D}_2}} \geq B_{d^{\perp}_{\mathcal{D}_2}} \ .$$

Proof. $B'_{d^{\perp}_{\mathcal{D}_2}}$ is the number of pairs of codewords (x, y) in $\mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp}$ which satisfy the two conditions: their sum is in $\mathcal{C}^{\perp}$ and their weights are equal to $d^{\perp}_{\mathcal{D}_2}$. Clearly, this number is greater or equal to the same number of pairs where in addition, x and y are chosen to be identical. In the latter case, the number of codewords is equal to:

$$\left|\{x \in \mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp} | w_H(x) = d^{\perp}_{\mathcal{D}_2}\}\right| \ , \tag{12}$$

because $x + y = 0$ does always belong to $\mathcal{C}^{\perp}$ and that x and y have the same Hamming weight since they are equal. Now, Eqn. 12 is the minimum nonzero coefficient in the weight enumerator of $\mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp}$, which is equal to B_d in SSS-based masking. □

Hereafter, we demonstrate Lemma 4 by utilizing Eqn. 9 to simplify Lemma 3.

Proof of Lemma 4. Let $\varphi_I(z) = z^I$ where $I \in \{0,1\}^{n\ell}$. Thus

$$z^I = \prod_{i\in I} z_i = \prod_{i\in I} \frac{(1-(-1)^{z_i})}{2} = \frac{1}{2^d}\prod_{i\in I}(1-(-1)^{z_i}). \tag{13}$$

Since all monomials with numerical degree smaller than d have SNR $= 0$, we only focus on monomials with numerical degree equal to d. Taking $\varphi_I(z) = \phi_I(z) + \frac{(-1)^d}{2^d}(-1)^{\sum_{i\in I} z_i}$ where $\phi_I(z)$ is linear combination of monomials with numerical degree smaller than d in $\varphi_I(z)$, then the *Fourier transform* of $\varphi_I(z)$ is:

$$\begin{aligned}\widehat{\varphi}_I(y) &= \widehat{\phi}_I(y) + \frac{(-1)^d}{2^d}\sum_z (-1)^{z\cdot I}(-1)^{z\cdot y} = \widehat{\phi}_I(y) + \frac{(-1)^d}{2^d}\sum_z (-1)^{z\cdot(I+y)} \\ &= \widehat{\phi}_I(y) + \frac{(-1)^d}{2^{d-n\ell}}\mathbb{1}_{\{I\}}(y).\end{aligned} \tag{14}$$

We have $\widehat{\phi}_I(y) = 0$ for y with $w_H(y) \geq d^{\perp}_{\mathcal{D}} = t+1 > d$, since given a pseudo-Boolean function P, one has $\widehat{P}(z) = 0$ for all $z \in \mathbb{K}^n$ with $w_H(z) > \deg(P)$ [BCC+14, Lemma 1]. As a result, by combining Eqn. 14 with Eqn. 24, we have the following equation:

$$\begin{aligned}\mathbb{V}[\mathbb{E}[P(Z)|X]] &= \frac{1}{2^{2n\ell}} \sum_{\substack{x,\, y\in\mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp} \\ x+y\in\mathcal{C}^{\perp}}} \widehat{P}(x)\widehat{P}(y) \\ &= \frac{1}{2^{2n\ell}} \sum_{\substack{x,\, y\in\mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp} \\ x+y\in\mathcal{C}^{\perp}}} \left(\sum_{I|w_H(I)=d} \frac{(-1)^d}{2^{d-n\ell}}\binom{d}{I}\mathbb{1}_{\{I\}}(x)\right)\left(\sum_{I|w_H(I)=d} \frac{(-1)^d}{2^{d-n\ell}}\binom{d}{I}\mathbb{1}_{\{I\}}(y)\right) \\ &= 2^{-2d} \sum_{\substack{x,\, y\in\mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp} \\ x+y\in\mathcal{C}^{\perp}}} \left(\sum_{I|w_H(I)=d} \binom{d}{I}\mathbb{1}_{\{I\}}(x)\right)\left(\sum_{I|w_H(I)=d} \binom{d}{I}\mathbb{1}_{\{I\}}(y)\right) \\ &= \left(\frac{d!}{2^d}\right)^2 \sum_{x,\, y\in\mathcal{D}^{\perp}\backslash\mathcal{C}^{\perp};\ x+y\in\mathcal{C}^{\perp}} 1 \ = B'_d\left(\frac{d!}{2^d}\right)^2 ,\end{aligned} \tag{15}$$

where B'_d is the adjusted coefficient in weight enumerator defined in Def. 9. □

4.2 Connecting SNR with Code Properties

Taking Lemma 4 as an input to Theorem 1, we have the following theorem for Hamming weight leakages in GCM.

Theorem 2. *Let a device be protected by the GCM scheme as $Z = X\mathbf{G} + Y\mathbf{H}$. Assume the device is leaking in Hamming weight model in the form: $\mathcal{L} = P(Z) + N$. Then the SNR of the exploitable leakages is:*

$$SNR = \frac{\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]}{\sigma^2_{total}} = \frac{B'_{d^\perp_{\mathcal{D}}}}{\sigma^2_{total}} \left(\frac{d^\perp_{\mathcal{D}}!}{2^{d^\perp_{\mathcal{D}}}}\right)^2 ,$$

where σ^2_{total} is the total noise such that $\sigma^2_{total} \propto \sigma^{2d}$.

Proof. Obviously, substituting the expression of $\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]$ in Theorem 1 by Lemma 4 gives the proof. □

The takeaway point is, the Hamming weight leakages, in which $\deg(P) = 1$, are quantified by Theorem 2, in which the two parameters that have an impact on SNR are the dual distance $d^\perp_{\mathcal{D}}$ and the coefficient $B'_{d^\perp_{\mathcal{D}}}$. Therefore, the two parameters also affect the concrete security level of GCM. As a straightforward application of Theorem 2, the side-channel resistance of GCM can be optimized by increasing $d^\perp_{\mathcal{D}}$ and/or decreasing $B'_{d^\perp_{\mathcal{D}}}$.

4.3 MI-based Information-Theoretic Leakage Quantification

We extend the leakage quantification approach by using another metric, namely MI, in an information-theoretic sense. Let the secret X be encoded as in Eqn. 1, and let the leakages be $\mathcal{L} = P(Z) + N$, then the MI between $\mathcal{L}$ and X is defined as $\mathsf{I}[\mathcal{L}; X] = \mathsf{H}[\mathcal{L}] - \mathsf{H}[\mathcal{L}|X]$ where:

- the total entropy is: $\mathsf{H}[\mathcal{L}] = -\int_l \mathbb{P}\left[l\right] \log_2 \mathbb{P}\left[l\right] \mathrm{d}l$,
- the conditional entropy $\mathsf{H}[\mathcal{L}|X]$ is: $\mathsf{H}[\mathcal{L}|X] = -\sum_{x\in\mathbb{F}_2^\ell} \mathbb{P}\left[x\right] \int_l \mathbb{P}\left[l|x\right] \log_2 \mathbb{P}\left[l|x\right] \mathrm{d}l$.

In multivariate cases, two entropies are computed on $\mathcal{L} = (\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_d)$ for a d-variate MI by a d-D integral on continuous variables. While in monovariate cases, two entropies are computed by 1-D integrals. Moreover, $\mathsf{I}[\mathcal{L}; X]$ can be expanded using a Taylor's expansion[5] [CDG+14]:

$$\mathsf{I}[\mathcal{L}; X] = \sum_{d=0}^{+\infty} \frac{1}{2\,d!\ln 2} \sum_{x\in\mathbb{F}_2^\ell} Pr(x) \frac{\left(k_d(P(Z)|x) - k_d(P(Z))\right)^2}{\left(\mathbb{V}\left[P(Z)\right] + \sigma^2\right)^d} \tag{16}$$

where k_d is the d-th order cumulant [Car03].

Assuming the device is leaking in the Hamming weight model, we have the following theorem for quantifying the information leakages in GCM.

Theorem 3. *Let a device be protected by the GCM scheme as $Z = X\mathbf{G} + Y\mathbf{H}$. Assume the leakages of the device can be represented in the form: $\mathcal{L} = P(Z) + N$. Then the MI between $\mathcal{L}$ and X is estimated as:*

$$\mathsf{I}[\mathcal{L}; X] = \begin{cases} 0\,, & \textit{if } \deg(P) < d^\perp_{\mathcal{D}} \\ \frac{d^\perp_{\mathcal{D}}!\,B'_{d^\perp_{\mathcal{D}}}}{2\ln 2\cdot 2^{2d^\perp_{\mathcal{D}}}} \times \frac{1}{\sigma^{2d^\perp_{\mathcal{D}}}} + \mathcal{O}\left(\frac{1}{\sigma^{2(d^\perp_{\mathcal{D}}+1)}}\right), & \textit{if } \deg(P) = d^\perp_{\mathcal{D}},\ \textit{when } \sigma \to +\infty \end{cases} \tag{17}$$

where σ is the standard deviation of noise in the leakage of each share.

[5] The normalization by $\ln 2$ allows the mutual information to be expressed in unit of bits.

Proof. Since for a d-CI (*Correlation Immune*) function [Car10, Def. 1], all moments of order $i \le d$ are centered, so are the cumulants. Therefore, the first nonzero cumulant $k_d(X)$ is $k_{d_\mathcal{D}^\perp}(X)$ and it equals $\mu_{d_\mathcal{D}^\perp}(X)$. As a consequence, the term $\mathbb{E}\left[\left(k_d(P(Z)|X) - k_d(P(Z))\right)^2\right]$ in Eqn. 16 is null for all $d < d_\mathcal{D}^\perp$ and it is equal to $\mathbb{E}\left[\left(\mu_d(P(Z)|X) - \mu_d(P(Z))\right)^2\right] = \mathbb{V}\left[\mu_{d_\mathcal{D}^\perp}(P(Z)|X)\right] = \mathbb{V}\left[\mathbb{E}\left[P(Z)^{d_\mathcal{D}^\perp}|X\right]\right]$ for $d = d_\mathcal{D}^\perp$.

Assume that the device leaks in Hamming weight model, then $P(Z)^{d_\mathcal{D}^\perp}$ has a degree equal to $d_\mathcal{D}^\perp$. Hence the MI is equal to:

$$\mathsf{I}[\mathcal{L};X] = \frac{1}{2\ln 2 \cdot d_\mathcal{D}^\perp!} \frac{\mathbb{V}\left[\mathbb{E}\left[P(Z)^{d_\mathcal{D}^\perp}|X\right]\right]}{\left(\mathbb{V}\left[P(Z)\right] + \sigma^2\right)^{d_\mathcal{D}^\perp}} + \mathcal{O}\left(\frac{1}{\left(\mathbb{V}\left[P(Z)\right] + \sigma^2\right)^{d_\mathcal{D}^\perp + 1}}\right) , \qquad (18)$$

when $\sigma \to +\infty$. Finally, Eqn. 18 can be further developed at the first order in $1/\sigma^{2d_\mathcal{D}^\perp}$ as follows after involving Eqn. 15:

$$\mathsf{I}[\mathcal{L};X] = \frac{d_\mathcal{D}^\perp! B'_{d_\mathcal{D}^\perp}}{2\ln 2 \cdot 2^{2d_\mathcal{D}^\perp}} \times \frac{1}{\sigma^{2d_\mathcal{D}^\perp}} + \mathcal{O}\left(\frac{1}{\sigma^{2(d_\mathcal{D}^\perp+1)}}\right) ,$$

when $\sigma \to +\infty$, which proves Theorem 3. □

A comparison of MIs by estimation and numerical calculation is shown in Fig. 2. More precisely, the estimated MIs are converging to numerical one when $\log_{10}\sigma^2 \approx 1.5$, which verifies Theorem 3 numerically.

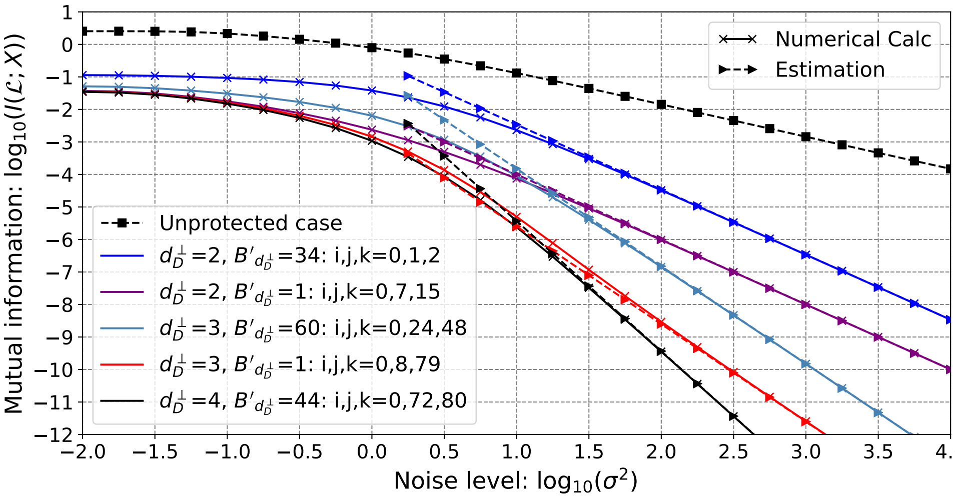


Figure 2: Numerical calculation and approximation of $\mathsf{I}[\mathcal{L};X]$ between leakage $\mathcal{L}$ and the sensitive variable X in $(3,1)$-SSS based masking. The three public points are $\alpha_1 = \alpha^i$, $\alpha_2 = \alpha^j$, $\alpha_3 = \alpha^k$.

Summing up, the information leakages of GCM under the Hamming weight model can be estimated by the two parameters $d_\mathcal{D}^\perp$ and $B'_{d_\mathcal{D}^\perp}$ in an information-theoretic sense. In the general case of leakage function P, the MI can be estimated similarly by applying different forms of P into Eqn. 18 to derive connections to coding properties correspondingly.

4.4 Optimal Codes for GCM

Thanks to Theorem 1, 2 and 3, we can compare the information leakages of GCM in a quantitative manner. More importantly, relying on the analytic characterization of information leakages, the three theorems enable us to choose optimal linear codes for GCM. Specifically, the codes with maximized $d^\perp_{\mathcal{D}}$ and minimized $B'_{d^\perp_{\mathcal{D}}}$ are the best candidates for GCM. Considering the SSS-based masking as a special case, the optimal public points can be determined straightforwardly by applying the two theorems.

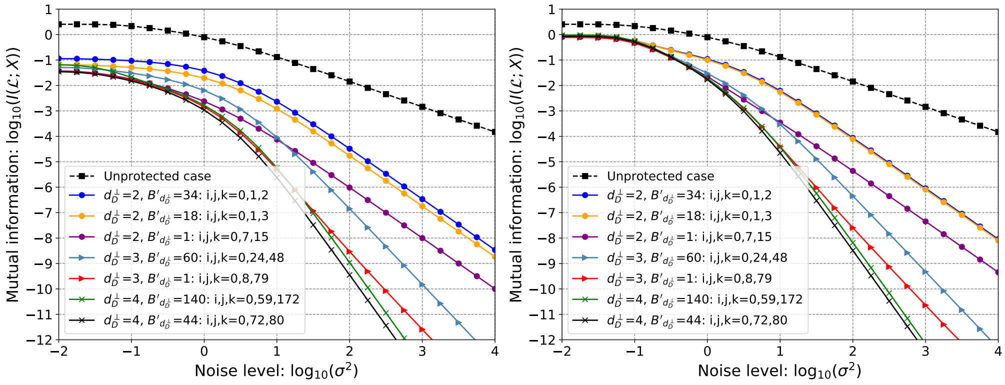
To thoroughly validate the optimal codes, we consider multivariate leakages. In particular, it is shown in [SVO+10] that comparing to sum, absolute difference, and normalized product, the joint distribution is the most efficient way to combine the multivariate leakages in side-channel analysis. In this paper, we consider both sum and joint distribution to exploit the multivariate leakages. A comparison of the two combination functions in an information-theoretic sense is presented in Appendix B.2.

We take $(3,1)$-SSS based masking as an example of GCM and specify it as follows. Let X be encoded into $Z = X\mathbf{G} + Y\mathbf{H}$ with $n = 3$ shares, the two generator matrices are:

$$\begin{aligned} \mathbf{G} \qquad\qquad\qquad\qquad &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} , \\ \mathbf{H} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix} &= \begin{pmatrix} 1 & \alpha^j & \alpha^k \end{pmatrix} . \end{aligned} \tag{19}$$

Considering the common "*Hamming weight* + *Gaussian noise*" model, the side-channel leakages are simulated as follows. Let $\mathcal{L} = (\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3)$ be 3-D leakages where $\mathcal{L}_i = \phi_P(Z_i) + N_i = w_H(Z_i) + N_i$ for $1 \leq i \leq 3$ and $N_i \sim \mathcal{N}(0, \sigma^2)$ is the Gaussian noise. To combine the 3-D leakages, other sum or joint distribution are applied wherein $\varphi_P(\mathcal{L}) = \sum_{i=0}^{3} \mathcal{L}_i$ is called 1-D leakages or $\varphi_P(\mathcal{L}) = (\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3)$ is called 3-D leakages, respectively.

The results are shown in Fig. 3(a) and 3(b) are 1-D MI and 3-D MI, respectively (more results over $\mathbb{F}_{2^4}$ are in Appendix B.1). The first observation is that the 3-D MI utilizing joint distribution exploits more key-dependent information existed in leakages, therefore the attack is more efficient when using the joint distribution of leakages [BGHR14]. Secondly, the numerical results in Fig. 3 are in accordance with the Theorem 2 and 3, where the two parameters $d^\perp_{\mathcal{D}}$ and $B'_{d^\perp_{\mathcal{D}}}$ in codes play a significant role in determining the side-channel resistance of GCM.



(a) 1-D MI on 8-bit case.

(b) 3-D MI on 8-bit case.

Figure 3: An information-theoretic evaluation of the leakages $\mathcal{L}$ and the sensitive variable X in $(3,1)$-SSS based masking. We choose seven codes with different values of $d^\perp_{\mathcal{D}}$ and/or $B'_{d^\perp_{\mathcal{D}}}$. The three public points are $\alpha_1 = \alpha^i$, $\alpha_2 = \alpha^j$, $\alpha_3 = \alpha^k$.

Thirdly, the strategy to choose the optimal codes for GCM is to maximize the dual distance $d^\perp_{\mathcal{D}}$ and/or to minimize the conditioned number of codewords $B'_{d^\perp_{\mathcal{D}}}$. Moreover, the concrete side-channel security level of GCM will be improved by optimizing either of the two parameters. Interestingly, when the noise levels are at certain intervals, the codes with smaller $d^\perp_{\mathcal{D}}$ (also with smaller $B'_{d^\perp_{\mathcal{D}}}$) may be better than that with larger $d^\perp_{\mathcal{D}}$. For instance, for the curves in purple (the fourth one) and in sky-blue (the fifth one) of

Fig. 3, the corresponding $d_{\mathcal{D}}^{\perp}$ are 2 and 3, respectively. When $\sigma^2 < 10$, the purple curve shows a better side-channel resistance than the sky-blue one.

5 Enhancing the SSS-based Polynomial Masking

In the context of masking, the random masks in SSS-based masking are u_i for $1 \leq i \leq t$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are n public points. Two main observations made in [CMP18] are:

- the choices of public points α_i can have an impact on side-channel resistance of the corresponding masking scheme, therefore, combining different $t+1$ tuples of Z_i, the efficiencies of corresponding template attacks are different,
- combining more than $t+1$ tuples of Z_i may improve the attack efficiency in the sense of the number of traces needed to recover the secret key.

Recall that the generator matrices in SSS-based masking (e.g., the RS code) from Tab. 1, **G** and **H** are the same as the generator matrices in DSM when $n = t+1$. In the context of masking, we only care about **G** and **H**, since the former is used to encode the secret X and the latter is for encoding the random masks (e.g., $u_1, \ldots, u_t$ in the case of SSS-based masking).

Note that **H** is a Vandermonde matrix, resulting in that the code $\mathcal{D}$ is a maximum distance separable (MDS) code, it is optimal at word-level. However, with different parameters α_i for $1 \leq i \leq n$, the codes have different impacts on side-channel resistance when they are adopted in masking schemes.

5.1 Further Clarifications

We further clarify the properties of the code $\mathcal{D}$ and its dual as follows. Let $\mathcal{D}$ be an RS code of parameters $[n,\, t,\, n-t+1]$ which is generated by **H** in Eqn. 3. Then its dual code $\mathcal{D}^{\perp}$ is also an RS code of parameters $[n,\, n-t,\, t+1]$ [MS77]. Recall the connections between the RS code and SSS scheme, $\mathcal{D}$ can be used to construct an $(n,\, t)$-SSS scheme.

Given that $n \geq t+1$, we assume that $t+1 \leq n' \leq n$, the code $\mathcal{D}'$ is constructed by selecting n' columns from the generator matrix **H** of $\mathcal{D}$ (or equivalently, remove $n - n'$ columns in **H**). Subsequently, the code $\mathcal{D}'$ has parameters $[n',\, t,\, n'-t+1]$. It is also an RS code and its dual code $\mathcal{D}'^{\perp}$ has parameters $[n',\, n'-t,\, t+1]$. Therefore, the dual distance of $\mathcal{D}'$ is equal to $\mathcal{D}$, namely $d_{\mathcal{D}'}^{\perp} = d_{\mathcal{D}}^{\perp} = t+1$. In summary, removing some coordinates ($n' \geq t+1$) in RS code does not decrease its dual distance (at word-level).

Remark 3. Note that for two arbitrary linear codes $\mathcal{D}$ and $\mathcal{D}'$ where the latter is generated from the former as above (by selecting some coordinates), we have the following lemma for their dual distances.

Lemma 6. $d_{\mathcal{D}}^{\perp} \leq d_{\mathcal{D}'}^{\perp}$.

Proof. Assume $u \in \mathcal{D}'^{\perp}$, by appending $n - n'$ zeros to u, then the new codeword $(u,\, 0_{n-n'})$ is also a codeword of $\mathcal{D}^{\perp}$. Therefore we have $d_{\mathcal{D}}^{\perp} \leq d_{\mathcal{D}'}^{\perp}$ [CGMÖ18]. □

Interestingly, Lemma 6 implies that given a fixed t, adding more shares in an (n, t)-SSS based masking cannot increase the security order of the corresponding masking scheme and can be more likely to lower the security order, especially under the bit-probing model.

5.2 Representing Linear Codes in Subfield $\mathbb{F}_2$

We take $\mathbb{F}_2$ as the subfield, then any codes over $\mathbb{F}_{2^\ell}$ can be expanded into subfields by code expansion Def. 5. We further investigate the properties of codes $\mathcal{D}$ and $\mathcal{D}'$.

Let $\mathcal{D}_2$ and $\mathcal{D}'_2$ denote the expanded codes of $\mathcal{D}$ and $\mathcal{D}'$ over $\mathbb{F}_2$, respectively. Since they are not MDS codes at the bit level, there is no straightforward method to compare the dual distances of $\mathcal{D}_2$ and $\mathcal{D}'_2$. However, by Lemma 6, it is obvious to have $d^{\perp}_{\mathcal{D}_2} \leq d^{\perp}_{\mathcal{D}'_2}$. This connection helps in SSS-based masking since, by increasing n, the dual distance at word-level keeps the same, but the dual distance at bit-level cannot be larger than in the case with $n' = t + 1$. Moreover, from the adversary's viewpoint, combining more than $t+1$ shares may be more efficient when attacking a specific SSS-based implementation.

From the quantitative results in Sec. 3, two parameters that have an impact on the side-channel resistance of GCM is the dual distance $d^{\perp}_{\mathcal{D}_2}$ and the coefficient $B'_{d^{\perp}_{\mathcal{D}_2}}$. Hereafter, we use the information-theoretic metric to show how the more redundant shares affect the concrete security level in SSS-based masking.

5.3 More Redundancy in Sharing Leaks More

We present an information-theoretic evaluation on $(3,1)$-SSS based polynomial masking. Taking $n = 3$ and $t = 1$, then the three public points $(\alpha_1, \alpha_2, \alpha_3)$ can be derived by setting $\alpha_1 = \alpha^i$, $\alpha_2 = \alpha^j$ and $\alpha_j = \alpha^k$, where i, j, k must be distinct integers. Due to the equivalence of the linear codes (Sec. 2.2), we can choose $i = 0$, $1 \leq j < k \leq 254$ and obtain 32131 candidates rather than $\binom{255}{3}$= 2731135 in total. Recall that the generator matrices $\mathbf{G}$ and $\mathbf{H}$ are as in Eqn. 19. Therefore, taking a random mask u_1, the X is encoded into:

$$Z = (Z_1,\, Z_2,\, Z_3) = X\mathbf{G} + u_1\mathbf{H} = (X + u_1\alpha_1,\, X + u_1\alpha_2,\, X + u_1\alpha_3)\ . \tag{20}$$

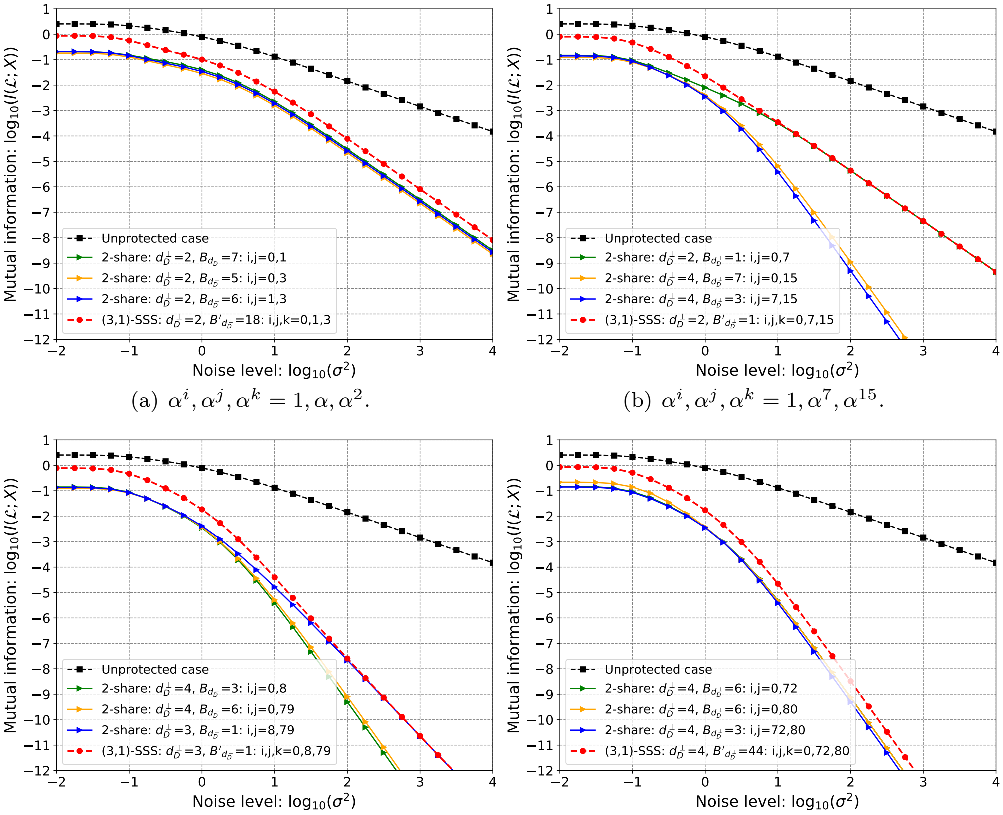
For all possible values of $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_{2^8}$, we study the dual distance $d^{\perp}_{\mathcal{D}}$ and the coefficient $B_{d^{\perp}_{\mathcal{D}}}$ at both word-level and bit-level. As expected, all codes have the same weight enumerator at word-level (they are all MDS codes and optimal at word-level). However, there are three possible values for $d^{\perp}_{\mathcal{D}}$ at bit-level, namely $d^{\perp}_{\mathcal{D}_2} \in \{2,\, 3,\, 4\}$. Hence, for each possible $d^{\perp}_{\mathcal{D}_2}$, we further study the possible values for the other parameter $B_{d^{\perp}_{\mathcal{D}_2}}$. In particular, for each case of $d^{\perp}_{\mathcal{D}_2}$, we show two or three codes with maximal and minimal values of $B_{d^{\perp}_{\mathcal{D}_2}}$. The specific properties of the codes are listed in Tab. 2 [6] and the MI between the leakages $\mathcal{L}$ and X are depicted in Fig. 3. The complete details of all linear codes for the $(3,1)$-SSS based masking are available in [CG20]. For the sake of brevity, we put more codes for $(3,1)$-SSS and $(5,2)$-SSS based masking in Appendix B.4.

Table 2: Exhibiting different codes in $(3,1)$-SSS scheme generated by Eqn. 20. Note that we take $\alpha_1 = \alpha^i = 1$, $\alpha_2 = \alpha^j$ and $\alpha_3 = \alpha^k$.

	$j=1$ $k=2$	$j=1$ $k=3$	$j=7$ $k=15$	$j=24$ $k=48$	$j=8$ $k=79$	$j=59$ $k=172$	$\boldsymbol{j=72}$ $\boldsymbol{k=80}$
Minimum distance $d_{\mathcal{D}}$	3	3	3	3	3	3	3
Dual distance (word) $d^{\perp}_{\mathcal{D}}$	2	2	2	2	2	2	2
Dual distance (bit) $d^{\perp}_{\mathcal{D}_2}$	2	2	2	3	3	**4**	**4**
Coefficient (bit) $B_{d^{\perp}_{\mathcal{D}_2}}$	20	18	**1**	22	**1**	76	**36**
Coefficient (bit) $B'_{d^{\perp}_{\mathcal{D}_2}}$	34	18	**1**	60	**1**	140	**44**

As shown in Tab. 2, for the first time, we exhibit an approach to find the optimal codes for SSS-based masking and present optimal codes for $(3,1)$-SSS based masking. Specifically, the code with $\alpha_1 = 1$, $\alpha_2 = \alpha^{72}$ and $\alpha_3 = \alpha^{80}$ (in the last column of Tab. 2) is one of the

[6] The data in Tab. 2 is formally verified by `Magma` [Uni]. Moreover, the scripts for calculating B'_d are also available on `Github` [CG20].



(a) $\alpha^i, \alpha^j, \alpha^k = 1, \alpha, \alpha^2$.

(b) $\alpha^i, \alpha^j, \alpha^k = 1, \alpha^7, \alpha^{15}$.

(c) $\alpha^i, \alpha^j, \alpha^k = 1, \alpha^8, \alpha^{79}$.

(d) $\alpha^i, \alpha^j, \alpha^k = 1, \alpha^{72}, \alpha^{80}$.

Figure 4: More shares leak more information, two study-cases on $(3,1)$-SSS based masking, where the three public points are: $\alpha_1 = \alpha^i$, $\alpha_2 = \alpha^j$, $\alpha_3 = \alpha^k$.

best candidates for $(3,1)$-SSS based masking. In addition, the generator matrices of all three optimal (nonequivalent) codes are shown in Appendix B.3. It is worth noting that the codes obtained by permuting the order of α_i for $1 \leq i \leq 3$ are equivalent, resulting in only three optimal codes for $(3,1)$-SSS based masking over $\mathbb{F}_{2^8}$.

Using the same settings of $(3,1)$-SSS based masking as in Sec. 4.4, the results of MI on the information leakages of 3-share and corresponding 2-share combinations are shown in Fig. 4. In each of four cases, the main takeaway point is that given a specific t in (n,t)-SSS based masking, all the more shares leak more key-dependent information. Specifically, we first highlight that the smallest security order determines the side-channel security of SSS-based masking among all $\binom{n}{t+1}$ combinations. In the context of coding theory, the dual distance of n-share SSS-based masking is determined by the minimum value of dual distances in truncated codes $\mathcal{D}'$. Two instances are in Fig. 4(b) and 4(c) where the minimum of dual distances are 2 and 3, respectively.

Secondly, when the codes in SSS and its truncated variants have the same dual distance, the parameter $B'_{d_{\mathcal{D}}^{\perp}}$ plays a role in side-channel resistance. More precisely, smaller $B'_{d_{\mathcal{D}}^{\perp}}$ brings improved concrete security for GCM. Two instances are shown in Fig. 4(a) and 4(d) where the dual distances of $\mathcal{D}$ are 2 and 4, respectively. Interestingly, a recent work [CS21] provides empirical comparisons on some instances of $(2,1)$-SSS and $(3,1)$-SSS based masking, which confirms our information-theoretic evaluation.

In summary, the information-theoretic evaluations in Fig. 4 confirms that more redundancy in sharing of GCM would leak more information. Besides, one way to find optimal codes for GCM is to build up from (sub-)optimal choices of the codes with less shares.

6 Discussions

6.1 Revisiting the Independence Condition

Failing to ensure the independence of the shares can ruin a masking scheme by revealing a lower order of key-dependent leakages than the designed security order. For instance, the unintentional physical coupling [BDF$^+$17] in the hardware device can combine leakages from different shares, hence degrade the concrete security level of a masked implementation. In this section, we investigate the intra-share independence issue and show the theoretical condition of higher-order security of code-based masking, especially in GCM as it is the most general case.

Another reason why the independence condition might be broken is the existence of glitches. Let us reason on a canonical example, namely that of the exclusive-or (XOR) gate. Let Z_1 and Z_2 be two single-bit shares, which enter an XOR gate. Recall that the leakage function is $P = \varphi_P \circ \phi_P$ as introduced in Sec. 3. Taking $\phi_P = 1$, then the leakage function is the pseudo-Boolean function φ_P, which lives in $\mathbb{F}_2 \times \mathbb{F}_2 \to \mathbb{R}$. It is equal to:

$$\varphi_P(Z_1, Z_2) = Z_1 \times Z_2 + (1 - Z_1) \times (1 - Z_2) = 2Z_1 \times Z_2 - Z_1 - Z_2 + 1.$$

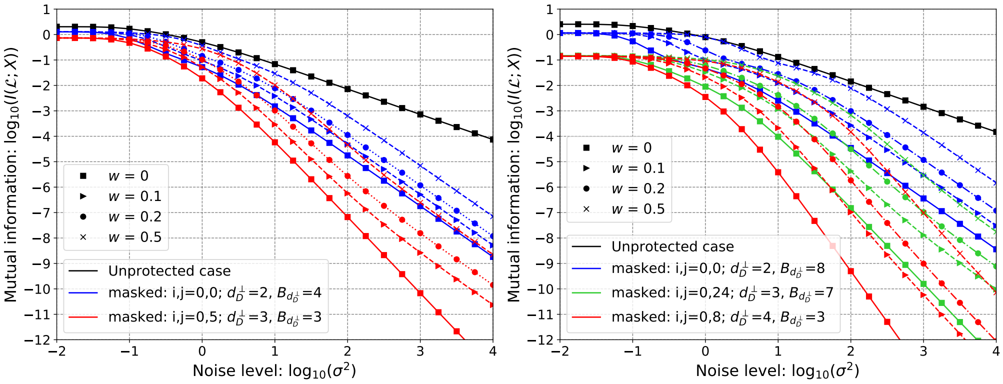
This function can glitch because of the term $Z_1 \times Z_2$. Indeed, if Z_1 changes, then the leading term still depends on Z_2 (derivative). Therefore, glitches are dreadful since they consist in combinations from within the chip, even before the measurement noise arrives.

An Information-Theoretic Evaluation of Intra-Share Independence. We consider the Hamming weight as leakage model in a perfect independent case and take the weighted square of Hamming weight as second-order (non-linear) leakages as follows:

$$\phi_P(Z_i) = \sum_{j=1}^{\ell} Z_{i,j} + w \sum_{j \neq k}^{\ell} Z_{i,j} Z_{i,k} = w_H(Z_i) + w \sum_{j \neq k}^{\ell} Z_{i,j} Z_{i,k} \tag{21}$$

where Z_i is an ℓ-bit share and w is the weight of second-order leakages. As a consequence, $P(Z) = \phi_P(Z)$ will be the same as Hamming weight model with $\deg(P) = 1$ if $w = 0$. Otherwise, there exists a different amount of second-order leakages indicated by w where the degree of P equals 2. The MI results on four candidates of w are shown in Fig. 5 for 4-bit and 8-bit variables, respectively. It is worthwhile to note that in 2-share settings with $n = 2$ and $t = 1$, the SSS-based masking can be transformed into IPM by changing the way of involving public parameters α_i for $1 \leq i \leq n$. Essentially, the two schemes are different because of the structure of $\mathbf{G}$ and $\mathbf{H}$ as in Tab. 1, but are comparable from a side-channel perspective.

The first observation from Fig. 5 is that MI increases along with the increasing amount of second-order leakages. More importantly, in the presence of second-order leakages, the security order under the bit-probing model [PGS$^+$17] (indicated by the slope of MI curves when the noise level is high) decreases by one since the degree of ϕ_P is 2. Similarly, the security order will reduce by two when the degree of ϕ_P equals 3 in the red curves of Fig. 5(b). However, the lowest security order under the bit-probing model is bounded by the Boolean masking under the word-probing model. More precisely, increasing the degree of ϕ_P only affects the intra-share independence and therefore decreases the security order under the bit-probing model, while the degree of φ_P (e.g., induced by couplings) affects the security order under the word-probing model.



(a) $\ell = 4$ for 4-bit case.

(b) $\ell = 8$ for 8-bit case.

Figure 5: The intra-share independence issue: the existence of higher-order leakages decreases the security of the corresponding masking scheme (two public parameters are $\alpha_1 = \alpha^i$, $\alpha_2 = \alpha^j$ as in Tab. 1). Note that the blue curves are for the Boolean masking.

6.2 Related Works

6.2.1 Differences with [CGC⁺21] in Detail

As summarized in Sec. 1.3, this work tackles GCM, which is a more general masking scheme than the one studied in [CGC$^+$21]. In fact, we utilize the same notion of the numerical degree and a similar coding-theoretic approach as in [CGC$^+$21], and also the same leakage assessment metrics like SNR and MI. However, generalizing [CGC$^+$21] to this work is not trivial at all, we show hereafter the technical differences from [CGC$^+$21].

We first highlight the different constructions of the generator matrices $\mathbf{G}$ and $\mathbf{H}$ in Tab. 1 for the codes $\mathcal{C}$ and $\mathcal{D}$, respectively. Indeed, $\mathcal{C}$ and $\mathcal{D}$ are not complementary in GCM, while they are complementary in IPM. In this respect, we show that Eqn. 7 is simplified as Eqn. 8 when $\mathcal{C}$ and $\mathcal{D}$ are complementary, thus we recover the main results in [CGC$^+$21] (see Remark 1). As a special case, the framework proposed in [CGC$^+$21] is applicable when $\mathcal{C}$ and $\mathcal{D}$ are complementary, e.g., when $n = t + 1$ in SSS-based masking.

Moreover, we prove that GCM requires introducing a more general parameter B'_d (see Def. 9), which is a novel parameter for linear codes. Particularly, in [CGC$^+$21] the parameter B_d only depends on $\mathcal{D}$. While B'_d depends on both $\mathcal{C}$ and $\mathcal{D}$, which indicates the importance of selecting appropriate candidates for both of them in practice. We also provide efficient magma scripts to evaluate this quantity [CG20].

Finally, we insist that the generalization in this work is a significant improvement that works for all GCMs. Since firstly, we show in Remark 2 that the security order can be greater than the dual distance minus one in GCM, which cannot be explained by the framework in [CGC$^+$21], but can be explained perfectly by this work in a quantitative manner. Secondly, the redundancies in GCM allow detecting faults (e.g., for glitch-free designs [PR11]), which is currently an active research topic. We leave open the question on the construction of coding-theoretic countermeasures against both side-channel and fault injection attacks for future investigation.

6.2.2 Connections with [CS21]

The SSS-based masking is also the topic of a recent work [CS21], in which Costes et al. showed that the Boolean masking is a special case of SSS-based masking when $n \leq 6$. More interestingly, their simulation-based multivariate attacks [BGHR14] confirm our mathematical derivations, in particular, the information-theoretic evaluation in Fig. 4.

More generally, this work provides a unified framework for quantifying information leakage of all GCM instances. As a straightforward application, Theorems 2 and 3 in this paper enable us to explain the empirical observations in practical attacks. For instance, the three codes for $(3, 1)$-SSS in Fig. 3 of [CS21] correspond to different $d_{\mathcal{D}}^{\perp}$ and/or $B_{d_{\mathcal{D}}^{\perp}}$. However, we stress that the three codes for $(2, 1)$-SSS in the same figure are not equivalent to each other but have the same $d_{\mathcal{D}}^{\perp}$ equal to 4 and closely distributed $B_{d_{\mathcal{D}}^{\perp}} \in \{11, 8, 8\}$. Moreover, this work presents a systematic way to select optimal codes for SSS-based masking and GCM, which is out of the scope of [CS21].

6.3 Efficient Implementations of GCM

In this paper, we optimize security without touching the performances of GCM (there is no tradeoff between security and performance). Our coding-theoretic approach shows that both SNR and MI security metrics concur that dual distance and adjusted coefficient in weight enumerator are the two drivers for security improvements. Essentially, we stick to the definition of GCM (recall the rightmost column in Tab. 1), and propose an effective way to tune the underlying codes.

In terms of performances, they are the same (with respect to memory and speed) as the generic GCM. A more detailed study could consist in attempting to represent the generator matrices $\mathbf{G}$ and $\mathbf{H}$ as compactly as possible (with as many zeros and ones in coefficients as possible, or with a specific structure, say "cyclic" for instance). Besides, Wang et al. [WMCS20] showed a complementary way to improve the overall performance of GCM implementations by an amortization technique. Both approaches would ease an efficient implementation of GCM, leaving an open problem for future study.

7 Conclusions and Perspectives

This paper presented a unified approach to quantifying the information leakages of code-based masking in the most general case, namely GCM, which already encompasses many state-of-the-art masking schemes. Firstly, by a uniform representation of encodings in GCM, we proposed a quantitative approach to evaluate the concrete security level of GCM. The signal-to-noise ratio and mutual information are used as two complementary metrics to quantify the lowest degree of key-dependent leakages. By this unified approach, we were able to quantify the impact of different codes in GCM and optimize it by choosing optimal codes for it. Next, we evaluated the impact of public points in Shamir's Secret Sharing in the context of masking. Thanks to the unified analytic approach, we showed the impact of public points in side-channel security orders of the corresponding masking. More importantly, we provided a roadmap to optimal linear codes for designers to optimize the SSS-based masking (also GCM) soundly. Lastly, we revisited the independence condition behind the masking scheme and showed that the intra-share dependence could ruin higher-order security under the bounded moment model. In particular, we showed how the higher-order intra-share leakages affect the side-channel security orders precisely.

However, the construction of optimal codes for a large number of shares is still an open problem. We launched an exhaustive study on $(3, 1)$-SSS based masking and presented some results on (5,2)-SSS in [CG20]. But the exhaustive enumeration would be computationally infeasible when n gets larger (e.g., $n > 8$) in SSS-based masking or, more generally, in GCM. A heuristic solution is to construct new (sub-) optimal codes by concatenating two optimal or sub-optimal codes, following a gradient descent idea. Alternatively, constructing the (sub-)optimal codes by an algebraic approach under certain constraints is a promising solution. We will explore both solutions for GCM in the future.

Acknowledgments

The authors sincerely thank the anonymous reviewers for their valuable comments, which significantly improved the quality of the paper. They also thank François-Xavier Standaert and Weijia Wang for their helpful advice regarding the computation of mutual information in multivariate cases.

This work has been partly financed by the BRAINE ("Big data pRocessing and Artificial Intelligence at the Network Edge") H2020 ECSEL European Project, under Grant Agreement N° 876967. Besides, the French FUI (AAP-22) program CSAFE+ also funded part of this work, related to secure implementation of cryptographic algorithm ready to be protected against perturbation attacks. Sylvain Guilley was supported in part by the National Natural Science Foundation of China (N° 61632020). Wei Cheng was supported by the scholarship from China Scholarship Council (CSC) under the grant N° 201809110144.

References

[BBD$^+$16] Gilles Barthe, Sonia Belaïd, François Dupressoir, Pierre-Alain Fouque, Benjamin Grégoire, Pierre-Yves Strub, and Rébecca Zucchini. Strong non-interference and type-directed higher-order masking. In Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, and Shai Halevi, editors, *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, pages 116–129. ACM, 2016.

[BCC$^+$14] Julien Bringer, Claude Carlet, Hervé Chabanne, Sylvain Guilley, and Houssem Maghrebi. Orthogonal Direct Sum Masking - A Smartcard Friendly Computation Paradigm in a Code, with Builtin Protection against Side-Channel and Fault Attacks. In David Naccache and Damien Sauveron, editors, *Information Security Theory and Practice. Securing the Internet of Things - 8th IFIP WG 11.2 International Workshop, WISTP 2014, Heraklion, Crete, Greece, June 30 - July 2, 2014. Proceedings*, volume 8501 of *Lecture Notes in Computer Science*, pages 40–56. Springer, 2014.

[BCO04] Éric Brier, Christophe Clavier, and Francis Olivier. Correlation power analysis with a leakage model. In Marc Joye and Jean-Jacques Quisquater, editors, *Cryptographic Hardware and Embedded Systems - CHES 2004: 6th International Workshop Cambridge, MA, USA, August 11-13, 2004. Proceedings*, volume 3156 of *Lecture Notes in Computer Science*, pages 16–29. Springer, 2004.

[BDF$^+$17] Gilles Barthe, François Dupressoir, Sebastian Faust, Benjamin Grégoire, François-Xavier Standaert, and Pierre-Yves Strub. Parallel Implementations of Masking Schemes and the Bounded Moment Leakage Model. In *Advances in Cryptology - EUROCRYPT 2017, Paris, France, April 30 - May 4, 2017, Proceedings, Part I*, pages 535–566, 2017.

[BFG15] Josep Balasch, Sebastian Faust, and Benedikt Gierlichs. Inner Product Masking Revisited. In Oswald and Fischlin [OF15], pages 486–510.

[BFG$^+$17] Josep Balasch, Sebastian Faust, Benedikt Gierlichs, Clara Paglialonga, and François-Xavier Standaert. Consolidating Inner Product Masking. In Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology - ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of*

Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I, volume 10624 of *Lecture Notes in Computer Science*, pages 724–754. Springer, 2017.

[BFGV12] Josep Balasch, Sebastian Faust, Benedikt Gierlichs, and Ingrid Verbauwhede. Theory and Practice of a Leakage Resilient Masking Scheme. In Xiaoyun Wang and Kazue Sako, editors, *Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings*, volume 7658 of *Lecture Notes in Computer Science*, pages 758–775. Springer, 2012.

[BGHR14] Nicolas Bruneau, Sylvain Guilley, Annelie Heuser, and Olivier Rioul. Masks Will Fall Off – Higher-Order Optimal Distinguishers. In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology – ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014, Proceedings, Part II*, volume 8874 of *Lecture Notes in Computer Science*, pages 344–365. Springer, 2014.

[Car03] Jean-François Cardoso. Dependence, Correlation and Gaussianity in Independent Component Analysis. *Journal of Machine Learning Research*, 4:1177–1203, 2003.

[Car10] Claude Carlet. Boolean Functions for Cryptography and Error Correcting Codes: Chapter of the monography Boolean Models and Methods in Mathematics, Computer Science, and Engineering. pages 257–397. Cambridge University Press, Y. Crama and P. Hammer eds, 2010. Preliminary version is available at http://www.math.univ-paris13.fr/~carlet/chap-fcts-Bool-corr.pdf.

[CDG+14] Claude Carlet, Jean-Luc Danger, Sylvain Guilley, Houssem Maghrebi, and Emmanuel Prouff. Achieving side-channel high-order correlation immunity with leakage squeezing. *J. Cryptographic Engineering*, 4(2):107–121, 2014.

[CG99] Claude Carlet and Philippe Guillot. A New Representation of Boolean Functions. In Marc P. C. Fossorier, Hideki Imai, Shu Lin, and Alain Poli, editors, *AAECC*, volume 1719 of *Lecture Notes in Computer Science*, pages 94–103. Springer, 1999.

[CG18] Claude Carlet and Sylvain Guilley. Statistical properties of side-channel and fault injection attacks using coding theory. *Cryptography and Communications*, 10(5):909–933, 2018.

[CG20] Wei Cheng and Sylvain Guilley. Open-source: Quantifying Information Leakages in GCM, September 2020. http://github.com/Qomo-CHENG/GeneralizedCM.

[CGC+21] Wei Cheng, Sylvain Guilley, Claude Carlet, Sihem Mesnager, and Jean-Luc Danger. Optimizing Inner Product Masking Scheme by a Coding Theory Approach. *IEEE Trans. Inf. Forensics Secur.*, 16:220–235, 2021.

[CGMÖ18] Claude Carlet, Cem Güneri, Sihem Mesnager, and Ferruh Özbudak. Construction of some codes suitable for both side channel and fault injection attacks. In Lilya Budaghyan and Francisco Rodríguez-Henríquez, editors, *Arithmetic of Finite Fields - 7th International Workshop, WAIFI 2018, Bergen, Norway, June 14-16, 2018, Revised Selected Papers*, volume 11321 of *Lecture Notes in Computer Science*, pages 95–107. Springer, 2018.

[CJRR99] Suresh Chari, Charanjit S. Jutla, Josyula R. Rao, and Pankaj Rohatgi. Towards Sound Approaches to Counteract Power-Analysis Attacks. In *CRYPTO*, volume 1666 of *LNCS*. Springer, August 15-19 1999. Santa Barbara, CA, USA. ISBN: 3-540-66347-9.

[CMP18] Hervé Chabanne, Houssem Maghrebi, and Emmanuel Prouff. Linear repairing codes and side-channel attacks. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2018(1):118–141, 2018.

[CRZ13] Guilhem Castagnos, Soline Renner, and Gilles Zémor. High-order masking by using coding theory and its application to AES. In Martijn Stam, editor, *Cryptography and Coding - 14th IMA International Conference, IMACC 2013, Oxford, UK, December 17-19, 2013. Proceedings*, volume 8308 of *Lecture Notes in Computer Science*, pages 193–212. Springer, 2013.

[CS21] Nicolas Costes and Martijn Stam. Redundant code-based masking revisited. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2021(1):426–450, 2021.

[DDF14] Alexandre Duc, Stefan Dziembowski, and Sebastian Faust. Unifying Leakage Models: From Probing Attacks to Noisy Leakage. In Phong Q. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, volume 8441 of *Lecture Notes in Computer Science*, pages 423–440. Springer, 2014.

[DFS15] Alexandre Duc, Sebastian Faust, and François-Xavier Standaert. Making Masking Security Proofs Concrete - Or How to Evaluate the Security of Any Leaking Device. In Oswald and Fischlin [OF15], pages 401–429.

[DGH$^+$18] Jean-Luc Danger, Sylvain Guilley, Annelie Heuser, Axel Legay, and Ming Tang. Physical Security Versus Masking Schemes. In Çetin Kaya Koç, editor, *Cyber-Physical Systems Security.*, pages 269–284. Springer, 2018.

[GM11] Louis Goubin and Ange Martinelli. Protecting AES with Shamir's Secret Sharing Scheme. In Preneel and Takagi [PT11], pages 79–94.

[GSF13] Vincent Grosso, François-Xavier Standaert, and Sebastian Faust. Masking vs. Multiparty Computation: How Large Is the Gap for AES? In Guido Bertoni and Jean-Sébastien Coron, editors, *Cryptographic Hardware and Embedded Systems - CHES 2013 - 15th International Workshop, Santa Barbara, CA, USA, August 20-23, 2013. Proceedings*, volume 8086 of *Lecture Notes in Computer Science*, pages 400–416. Springer, 2013.

[ISW03] Yuval Ishai, Amit Sahai, and David Wagner. Private Circuits: Securing Hardware against Probing Attacks. In *CRYPTO*, volume 2729 of *Lecture Notes in Computer Science*, pages 463–481. Springer, August 17–21 2003. Santa Barbara, California, USA.

[Mas93] James L Massey. Minimal codewords and secret sharing. In *Proceedings of the 6th joint Swedish-Russian international workshop on information theory*, pages 276–279. Citeseer, 1993.

[MP13] Gary L. Mullen and Daniel Panario. *Handbook of Finite Fields*. Chapman and Hall/CRC, June 17 2013. ISBN 9781439873786 - CAT# K13417.

[MS77] F. Jessie MacWilliams and Neil J. A. Sloane. *The Theory of Error-Correcting Codes*. Elsevier, Amsterdam, North Holland, 1977. ISBN: 978-0-444-85193-2.

[NS94] Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994.

[OF15] Elisabeth Oswald and Marc Fischlin, editors. *Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I*, volume 9056 of *Lecture Notes in Computer Science*. Springer, 2015.

[PGS$^+$17] Romain Poussier, Qian Guo, François-Xavier Standaert, Claude Carlet, and Sylvain Guilley. Connecting and Improving Direct Sum Masking and Inner Product Masking. In Thomas Eisenbarth and Yannick Teglia, editors, *Smart Card Research and Advanced Applications - 16th International Conference, CARDIS 2017, Lugano, Switzerland, November 13-15, 2017, Revised Selected Papers*, volume 10728 of *Lecture Notes in Computer Science*, pages 123–141. Springer, 2017.

[PR11] Emmanuel Prouff and Thomas Roche. Higher-Order Glitches Free Implementation of the AES Using Secure Multi-party Computation Protocols. In Preneel and Takagi [PT11], pages 63–78.

[PR13] Emmanuel Prouff and Matthieu Rivain. Masking against Side-Channel Attacks: A Formal Security Proof. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings*, volume 7881 of *Lecture Notes in Computer Science*, pages 142–159. Springer, 2013.

[PRR14] Emmanuel Prouff, Matthieu Rivain, and Thomas Roche. On the Practical Security of a Leakage Resilient Masking Scheme. In Josh Benaloh, editor, *Topics in Cryptology - CT-RSA 2014 - The Cryptographer's Track at the RSA Conference 2014, San Francisco, CA, USA, February 25-28, 2014. Proceedings*, volume 8366 of *Lecture Notes in Computer Science*, pages 169–182. Springer, 2014.

[PT11] Bart Preneel and Tsuyoshi Takagi, editors. *Cryptographic Hardware and Embedded Systems - CHES 2011 - 13th International Workshop, Nara, Japan, September 28 – October 1, 2011. Proceedings*, volume 6917 of *LNCS*. Springer, 2011.

[RP10] Matthieu Rivain and Emmanuel Prouff. Provably Secure Higher-Order Masking of AES. In Stefan Mangard and François-Xavier Standaert, editors, *CHES*, volume 6225 of *LNCS*, pages 413–427. Springer, 2010.

[Sha79] Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979.

[SVO$^+$10] François-Xavier Standaert, Nicolas Veyrat-Charvillon, Elisabeth Oswald, Benedikt Gierlichs, Marcel Medwed, Markus Kasper, and Stefan Mangard. The World Is Not Enough: Another Look on Second-Order DPA. In Masayuki Abe, editor, *Advances in Cryptology - ASIACRYPT 2010 - 16th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 5-9, 2010. Proceedings*, volume 6477 of *Lecture Notes in Computer Science*, pages 112–129. Springer, 2010.

[Uni] University of Sydney (Australia). Magma Computational Algebra System. http://magma.maths.usyd.edu.au/magma/, Accessed on 2014-08-22.

[WMCS20] Weijia Wang, Pierrick Méaux, Gaëtan Cassiers, and François-Xavier Standaert. Efficient and Private Computations with Code-Based Masking. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2020(2):128–171, 2020.

A Detailed Proofs of Lemmas

Before presenting these proofs, we recall below two well-known properties of *Fourier transform*. We omit the proofs for the sake of brevity and refer to [Car10] for details.

Lemma 7 (Involution Property)**.** $\widehat{\widehat{P}}(z) = |\mathbb{K}^{n\ell}|P(z) = 2^{n\ell}P(z)$, $\forall z \in \mathbb{K}^{n\ell}$.

Lemma 8 (Inverse *Fourier Transform*)**.** $P(z) = 2^{-n\ell}\sum_{y\in\mathbb{K}^{n\ell}} \widehat{P}(y)(-1)^{y\cdot z}$, $\forall z \in \mathbb{K}^{n\ell}$.

A.1 Proof of Lemma 3

In order to demonstrate Lemma 3, we clarify the computations in $\mathbb{V}[\mathbb{E}[P(Z)|X]]$ as follows. Let us consider Eqn. 1 in basefield $\mathbb{F}_2$, and thus let $\mathcal{X} = \mathbb{F}_2^{\ell}$, $\mathcal{Y} = \mathbb{F}_2^{t\ell}$ and $\mathcal{Z} = \mathbb{F}_2^{n\ell}$. Moreover, the $\mathcal{C}$ and $\mathcal{D}$ are expanded into $\mathbb{F}_2$ by using code expansion (Def. 5):

- $\mathbb{E}[P(Z)|X = x]$ for a given $x \in \mathcal{X}$ is:

$$\begin{aligned}\mathbb{E}[P(x\mathbf{G} + Y\mathbf{H})] &= \sum_{y\in\mathcal{Y}} \mathbb{P}(Y = y)P(x\mathbf{G} + y\mathbf{H}) = \frac{1}{|\mathcal{Y}|}\sum_{y\in\mathcal{Y}} P(x\mathbf{G} + y\mathbf{H}) \\ &= \frac{1}{|\mathcal{D}|}\sum_{d\in\mathcal{D}} P(x\mathbf{G} + d).\end{aligned}$$

- For any variable X, we have that:

$$\mathbb{V}[\mathbb{E}[P(Z)|X]] = \mathbb{E}\left[\mathbb{E}[P(Z)|X]^2\right] - \mathbb{E}[\mathbb{E}[P(Z)|X]]^2 .$$

Next, we derive formulas for both sub-terms $\mathbb{E}[\mathbb{E}[P(Z)|X]]$ and $\mathbb{E}\left[\mathbb{E}[P(Z)|X]^2\right]$ and their proofs are in Appendix A.2 and A.3, respectively.

Lemma 9. $\mathbb{E}[\mathbb{E}[P(Z)|X]] = \frac{1}{2^{n\ell}}\sum_{x\in(\mathcal{C}\oplus\mathcal{D})^{\perp}} \widehat{P}(x)$.

Lemma 10. $\mathbb{E}\left[\mathbb{E}[P(Z)|X]^2\right] = \frac{1}{2^{2n\ell}}\sum_{\substack{x,\, y\in\mathcal{D}^{\perp}\\ x+y\in\mathcal{C}^{\perp}}} \widehat{P}(x)\widehat{P}(y)$.

Therefore, relying on the two lemmas, the proof of Lemma 3 is as follows.

Proof of Lemma 3. From Lemma 9, we compute $\mathbb{E}[\mathbb{E}[P(Z)|X]]^2$ as follows:

$$\begin{aligned}\mathbb{E}[\mathbb{E}[P(Z)|X]]^2 &= \left(\frac{1}{2^{n\ell}}\sum_{x\in(\mathcal{C}^{\perp}\cap\mathcal{D}^{\perp})} \widehat{P}(x)\right)^2 = \frac{1}{2^{2n\ell}}\left(\sum_{x\in(\mathcal{C}^{\perp}\cap\mathcal{D}^{\perp})} \widehat{P}(x)\right)^2 \\ &= \frac{1}{2^{2n\ell}}\sum_{x,\, y\in(\mathcal{C}^{\perp}\cap\mathcal{D}^{\perp})} \widehat{P}(x)\widehat{P}(y).\end{aligned} \tag{22}$$

Finally, we obtain $\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right]$ by combining Lemma 10 and Eqn. 22 as follows.

$$\begin{aligned}\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right] &= \mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]^2\right] - \mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]\right]^2\\ &= \frac{1}{2^{2n\ell}} \sum_{\substack{x,\, y \in \mathcal{D}^\perp;\\ x+y\in\mathcal{C}^\perp}} \widehat{P}(x)\widehat{P}(y) - \frac{1}{2^{2n\ell}} \sum_{x,\, y\in(\mathcal{C}^\perp \cap \mathcal{D}^\perp)} \widehat{P}(x)\widehat{P}(y)\\ &= \frac{1}{2^{2n\ell}} \left(\sum_{\substack{x,\, y \in \mathcal{D}^\perp;\\ x+y\in\mathcal{C}^\perp}} \widehat{P}(x)\widehat{P}(y) - \sum_{\substack{x,\, y \in \mathcal{D}^\perp;\\ x,\, y\in\mathcal{C}^\perp}} \widehat{P}(x)\widehat{P}(y) \right) .\end{aligned} \tag{23}$$

Due to Lemma 1, we have $\mathcal{C}^\perp \cap \mathcal{D}^\perp = (\mathcal{C} \oplus \mathcal{D})^\perp$ in SSS-based polynomial masking, where $\oplus$ denotes the direct sum operation. Notice that $\{(x, y) \in \mathbb{K}^n \times \mathbb{K}^n | x, y \in \mathcal{D}^\perp,\ x + y \in \mathcal{C}^\perp\} \supseteq \{(x, y) \in (\mathcal{D}^\perp \cap \mathcal{C}^\perp) \times (\mathcal{D}^\perp \cap \mathcal{C}^\perp)\}$. This means that in Eqn. 23, the subtracted terms are already included in the first sum. Indeed, if $x \in \mathcal{D}^\perp$ also satisfies $x \in \mathcal{C}^\perp$, then $x + y \in \mathcal{C}^\perp$ in the first sum implies $y \in \mathcal{C}^\perp$. Therefore, Eqn. 23 can be rewritten as follows:

$$\begin{aligned}\mathbb{V}\left[\mathbb{E}\left[P(Z)|X\right]\right] &= \mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]^2\right] - \mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]\right]^2\\ &= \frac{1}{2^{2n\ell}} \sum_{x,\, y \in \mathcal{D}^\perp \setminus \mathcal{C}^\perp;\, x+y\in\mathcal{C}^\perp} \widehat{P}(x)\widehat{P}(y).\end{aligned} \tag{24}$$

□

A.2 Proof of Lemma 9

Proof. Note that $\mathcal{C} \cap \mathcal{D} = \{0\}$, while $(\mathcal{C} \oplus \mathcal{D})^\perp = (\mathcal{C}^\perp \cap \mathcal{D}^\perp) \supseteq \{0\}$. We have

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]\right] &= \frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}}\left(\frac{1}{|\mathcal{Y}|}\sum_{d\in\mathcal{D}} P(x\mathbf{G}+d)\right) = \frac{1}{|\mathcal{C}|}\sum_{c\in\mathcal{C}}\left(\frac{1}{|\mathcal{D}|}\sum_{d\in\mathcal{D}} P(c+d)\right)\\ &= \frac{1}{|\mathcal{C}||\mathcal{D}|}\sum_{c\in\mathcal{C},\, d\in\mathcal{D}} P(c+d)\\ &= \frac{1}{|\mathcal{C}||\mathcal{D}|}\cdot\frac{1}{2^{n\ell}}\sum_{c\in\mathcal{C},\, d\in\mathcal{D}}\sum_{x\in\mathbb{F}_2^{n\ell}} \widehat{P}(x)(-1)^{(c+d)\cdot x} && \triangleright \textit{ By Lemma 8}\\ &= \frac{1}{|\mathcal{C}||\mathcal{D}|}\cdot\frac{1}{2^{n\ell}}\sum_{x\in\mathbb{F}_2^{n\ell}} \widehat{P}(x)\left(\sum_{c\in\mathcal{C}}(-1)^{c\cdot x}\right)\left(\sum_{d\in\mathcal{D}}(-1)^{d\cdot x}\right)\\ &= \frac{1}{2^{n\ell}}\sum_{x\in\mathbb{F}_2^{n\ell}} \widehat{P}(x)\mathbb{1}_{\mathcal{C}^\perp}(x)\mathbb{1}_{\mathcal{D}^\perp}(x) = \frac{1}{2^{n\ell}}\sum_{x\in\mathcal{C}^\perp,\, x\in\mathcal{D}^\perp}\widehat{P}(x)\\ &= \frac{1}{2^{n\ell}}\sum_{x\in(\mathcal{C}+\mathcal{D})^\perp}\widehat{P}(x)\ . && \triangleright \textit{ By Lemma 1}\end{aligned} \tag{25}$$

□

A.3 Proof of Lemma 10

Proof. By definition,

$$\mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]^2\right] = \frac{1}{|\mathcal{C}|}\sum_{c\in\mathcal{C}}\left(\frac{1}{|\mathcal{D}|}\sum_{d\in\mathcal{D}}P(c+d)\right)^2 = \frac{1}{|\mathcal{C}||\mathcal{D}|^2}\sum_{c\in\mathcal{C}}\left(\sum_{d\in\mathcal{D}}P(c+d)\right)^2. \tag{26}$$

We have:

$$\sum_{c\in\mathcal{C}}\left(\sum_{d\in\mathcal{D}}P(c+d)\right)^2=\frac{1}{2^{n\ell}}\cdot\frac{1}{2^{n\ell}}\sum_{\substack{c\in\mathcal{C},\,d,d'\in\mathcal{D}\\ x,y\in\mathbb{F}_2^{n\ell}}}\widehat{P}(x)\widehat{P}(y)(-1)^{x\cdot(c+d)+y\cdot(c+d')},\tag{27}$$

since, according to the inverse *Fourier transform* (by using Lemma 8), we have:

$$P(a)=2^{-n\ell}\sum_{x\in\mathbb{F}_2^{n\ell}}\widehat{P}(x)(-1)^{x\cdot a}\ .$$

Hence we obtain

$$\begin{aligned}\text{Eqn. 27}&=\frac{1}{2^{n\ell}}\cdot\frac{1}{2^{n\ell}}\sum_{\substack{c\in\mathcal{C},\,d,d'\in\mathcal{D}\\ x,\,y\in\mathbb{F}_2^{n\ell}}}\widehat{P}(x)\widehat{P}(y)(-1)^{(x+y)\cdot c+x\cdot d+y\cdot d'}\\&=\frac{1}{2^{n\ell}}\cdot\frac{1}{2^{n\ell}}\sum_{\substack{c\in\mathcal{C},\,d,d'\in\mathcal{D}\\ x,\,y\in\mathbb{F}_2^{n\ell}}}\widehat{P}(x)\widehat{P}(y)(-1)^{(x+y)\cdot c}(-1)^{x\cdot d}(-1)^{y\cdot d'}\\&=\frac{1}{2^{2n\ell}}\cdot|\mathcal{C}|\cdot|\mathcal{D}|^2\sum_{x,\,y\in\mathcal{D}^\perp;\,x+y\in\mathcal{C}^\perp}\widehat{P}(x)\widehat{P}(y),\end{aligned}\tag{28}$$

where $\mathcal{C}$, $\mathcal{D}$ are not necessary to be complementary codes and $|\mathcal{C}||\mathcal{D}|=2^{t\ell}\le 2^{n\ell}$. Indeed, since $\mathcal{C}$ is linear, $\sum_{c\in\mathcal{C}}(-1)^{(x+y)\cdot c}$ is null when $x+y$ does not belong to $\mathcal{C}^\perp$ and equals the size of $\mathcal{C}$ if it does, and the same with $\mathcal{D}$. Note that $x,\,y\in\mathcal{D}^\perp$ and $x+y\in\mathcal{C}^\perp$ which implies $x+y\in\mathcal{C}^\perp\cap\mathcal{D}^\perp$. In summary, we have the following result for $\mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]^2\right]$.

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[P(Z)|X\right]^2\right]&=\frac{1}{|\mathcal{C}||\mathcal{D}|^2}\cdot\frac{1}{2^{2n\ell}}\cdot|\mathcal{C}|\cdot|\mathcal{D}|^2\sum_{x,\,y\in\mathcal{D}^\perp;\,x+y\in\mathcal{C}^\perp}\widehat{P}(x)\widehat{P}(y)\\&=\frac{1}{2^{2n\ell}}\sum_{x,\,y\in\mathcal{D}^\perp;\,x+y\in\mathcal{C}^\perp}\widehat{P}(x)\widehat{P}(y).\end{aligned}\tag{29}$$

□

B Further Results

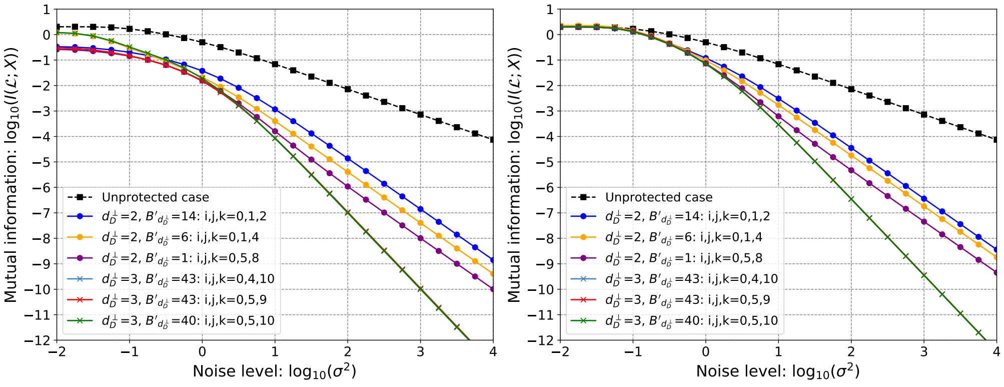
B.1 $(3,1)$-SSS based Masking on 4-bit Variables

The information-theoretic evaluations of $(3,1)$-SSS based masking over $\mathbb{F}_{2^4}$ are shown in Fig. 6, which are similar with the results over $\mathbb{F}_{2^8}$ as in Fig. 3.

B.2 Comparison of MI on 1-D and n-D Leakages

We add more results on MI to compare the efficiency of different combination functions φ_P in exploiting information leakages. In Fig. 3, we show the advantages to use joint distribution in trivariate leakages. In addition, we compare the two combination function in 2-share cases by plotting MI curves together. As shown in Fig. 7, the combination by using joint distribution is more efficient than the one by using sum in bivariate leakages scenarios. Moreover, this is true for n-variate leakages.

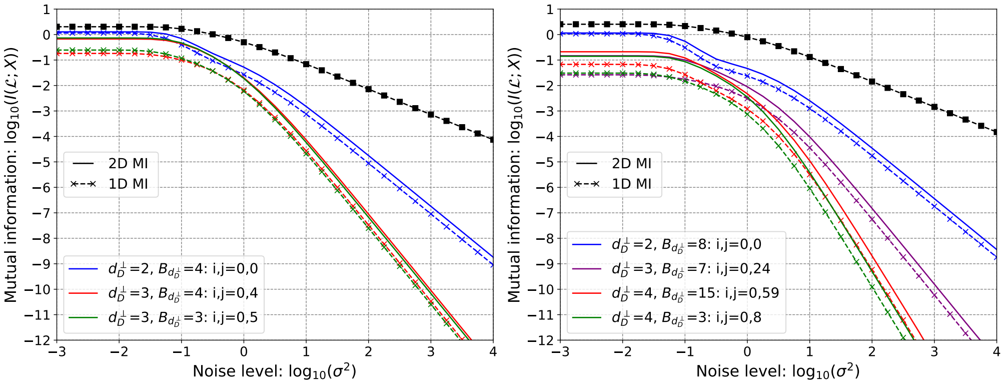
More importantly, the superiority of GCM can be fully unleashed by choosing appropriate codes. In this respect, our leakage quantitation approach is a simple, generic and effective way to choose the optimal codes for GCM.



(a) 1-D MI on 4-bit case.

(b) 3-D MI on 4-bit case.

Figure 6: An information-theoretic evaluation of the leakages $\mathcal{L}$ and the sensitive variable $X \in \mathbb{F}_{2^4}$. Six codes are chosen with different $d^\perp_{\mathcal{D}_2}$ and/or $B'_{d^\perp_{\mathcal{D}_2}}$.



(a) 1-D vs. 2-D MI on 4-bit case.

(b) 1-D vs. 2-D MI on 8-bit case.

Figure 7: Comparing 1-D and 2-D MI on different linear codes where the sum and joint distribution are used to combine the bivariate leakages, respectively. Note that the blue curves are for the Boolean masking.

B.3 Optimal Codes for $(3, 1)$-SSS based Masking

As shown in Tab. 1, the generator matrix of $\mathcal{D}$ is $\mathbf{H} = \begin{pmatrix}\alpha_1 & \alpha_2 & \alpha_3\end{pmatrix}$. From an exhaustive study on 32131 candidates, the three optimal codes for $(3, 1)$-SSS based masking are: $(\alpha_1, \alpha_2, \alpha_3) \in \{(\alpha^0, \alpha^{72}, \alpha^{80}), (\alpha^0, \alpha^{175}, \alpha^{247}), (\alpha^0, \alpha^8, \alpha^{183})\}$. Note that permutation on three public points does not change the codes due to equivalence.

The generator matrices of the three optimal codes are shown below.

$$\mathbf{H}_1 = \begin{pmatrix}\alpha^0 & \alpha^{72} & \alpha^{80}\end{pmatrix} = \begin{pmatrix}
1\,0\,0\,0\,0\,0\,0\,0 & 1\,0\,1\,0\,0\,1\,1\,0 & 1\,0\,1\,1\,1\,1\,1\,1 \\
0\,1\,0\,0\,0\,0\,0\,0 & 0\,1\,0\,1\,0\,0\,1\,1 & 1\,1\,1\,0\,0\,1\,1\,1 \\
0\,0\,1\,0\,0\,0\,0\,0 & 1\,0\,0\,1\,0\,0\,0\,1 & 1\,1\,0\,0\,1\,0\,1\,1 \\
0\,0\,0\,1\,0\,0\,0\,0 & 1\,1\,1\,1\,0\,0\,0\,0 & 1\,1\,0\,1\,1\,1\,0\,1 \\
0\,0\,0\,0\,1\,0\,0\,0 & 0\,1\,1\,1\,1\,0\,0\,0 & 1\,1\,0\,1\,0\,1\,1\,0 \\
0\,0\,0\,0\,0\,1\,0\,0 & 0\,0\,1\,1\,1\,1\,0\,0 & 0\,1\,1\,0\,1\,0\,1\,1 \\
0\,0\,0\,0\,0\,0\,1\,0 & 0\,0\,0\,1\,1\,1\,1\,0 & 1\,0\,0\,0\,1\,1\,0\,1 \\
0\,0\,0\,0\,0\,0\,0\,1 & 0\,0\,0\,0\,1\,1\,1\,1 & 1\,1\,1\,1\,1\,1\,1\,0
\end{pmatrix} \in \mathbb{F}_2^{8\times 24}$$

$$\mathbf{H}_2 = \begin{pmatrix}\alpha^0 & \alpha^{175} & \alpha^{247}\end{pmatrix} = \begin{pmatrix}
1&0&0&0&0&0&0&0 & 1&1&1&1&1&1&1&1 & 1&1&0&0&0&0&0&1\\
0&1&0&0&0&0&0&0 & 1&1&0&0&0&1&1&1 & 1&1&0&1&1&0&0&0\\
0&0&1&0&0&0&0&0 & 1&1&0&1&1&0&1&1 & 0&1&1&0&1&1&0&0\\
0&0&0&1&0&0&0&0 & 1&1&0&1&0&1&0&1 & 0&0&1&1&0&1&1&0\\
0&0&0&0&1&0&0&0 & 1&1&0&1&0&0&1&0 & 0&0&0&1&1&0&1&1\\
0&0&0&0&0&1&0&0 & 0&1&1&0&1&0&0&1 & 1&0&1&1&0&1&0&1\\
0&0&0&0&0&0&1&0 & 1&0&0&0&1&1&0&0 & 1&1&1&0&0&0&1&0\\
0&0&0&0&0&0&0&1 & 0&1&0&0&0&1&1&0 & 0&1&1&1&0&0&0&1
\end{pmatrix} \in \mathbb{F}_2^{8\times 24}$$

$$\mathbf{H}_3 = \begin{pmatrix}\alpha^0 & \alpha^{8} & \alpha^{183}\end{pmatrix} = \begin{pmatrix}
1&0&0&0&0&0&0&0 & 1&0&1&1&1&0&0&0 & 0&0&1&0&0&0&1&1\\
0&1&0&0&0&0&0&0 & 0&1&0&1&1&1&0&0 & 1&0&1&0&1&0&0&1\\
0&0&1&0&0&0&0&0 & 0&0&1&0&1&1&1&0 & 1&1&1&0&1&1&0&0\\
0&0&0&1&0&0&0&0 & 0&0&0&1&0&1&1&1 & 0&1&1&1&0&1&1&0\\
0&0&0&0&1&0&0&0 & 1&0&1&1&0&0&1&1 & 0&0&1&1&1&0&1&1\\
0&0&0&0&0&1&0&0 & 1&1&1&0&0&0&0&1 & 1&0&1&0&0&1&0&1\\
0&0&0&0&0&0&1&0 & 1&1&0&0&1&0&0&0 & 1&1&1&0&1&0&1&0\\
0&0&0&0&0&0&0&1 & 0&1&1&0&0&1&0&0 & 0&1&1&1&0&1&0&1
\end{pmatrix} \in \mathbb{F}_2^{8\times 24}$$

B.4 Different codes for $(3,1)$-SSS and $(5,2)$-SSS based masking

We present further results for both $(3,1)$-SSS and $(5,2)$-SSS based masking schemes which are supplementary to Tab. 2.

Note that in Tab. 4 we fix both α_1 and α_2 since there are too many candidates for enumeration (more accurately, $\binom{255}{5}$= $8,637,487,551$ candidates in total). In addition, the reason for taking $\alpha_2 = \alpha^8$ is that $(1 \;\; \alpha^8) \in \mathbb{F}_{2^8}^2$ is one of the optimal code for $(2,1)$-SSS based masking.

Table 3: Exhibiting different codes in $(3,1)$-SSS scheme over $\mathbb{F}_{2^4}$ generated by Eqn. 19. Note that we take $\alpha_1 = \alpha^i = 1$, $\alpha_2 = \alpha^j$ and $\alpha_3 = \alpha^k$.

	$j=1$ $k=2$	$j=1$ $k=3$	$j=3$ $k=7$	$j=4$ $k=8$	$j=5$ $k=10$
Minimum distance $d_{\mathcal{D}}$	3	3	3	3	3
Dual distance (word) $d_{\mathcal{D}}^{\perp}$	2	2	2	2	2
Dual distance (bit) $d_{\mathcal{D}_2}^{\perp}$	2	2	2	**3**	**3**
Coefficient (bit) $B_{d_{\mathcal{D}_2}^{\perp}}$	8	6	**1**	17	**16**
Coefficient (bit) $B'_{d_{\mathcal{D}_2}^{\perp}}$	14	6	**1**	45	**40**

Table 4: Exhibiting different codes in $(5,\,2)$-SSS scheme over $\mathbb{F}_{2^8}$. Note that we fix $\alpha_1 = \alpha^i = 1$, $\alpha_2 = \alpha^8$ and enumerate all possible $\alpha_3 = \alpha^k$, $\alpha_4 = \alpha^l$ and $\alpha_5 = \alpha^r$.

	$k=116$ $l=169$ $r=214$	$k=1$ $l=3$ $r=184$	$k=139$ $l=172$ $r=225$	$k=1$ $l=3$ $r=12$	$k=18$ $l=52$ $r=219$	$k=1$ $l=5$ $r=51$	$k=14$ $l=111$ $r=219$	$k=90$ $l=92$ $r=192$
Minimum distance $d_{\mathcal{D}}$	4	4	4	4	4	4	4	4
Dual distance (word) $d_{\mathcal{D}}^{\perp}$	3	3	3	3	3	3	3	3
Dual distance (bit) $d_{\mathcal{D}_2}^{\perp}$	3	3	4	4	5	5	**6**	**6**
Coefficient (bit) $B_{d_{\mathcal{D}_2}^{\perp}}$	19	**1**	29	**1**	43	**1**	115	**30**
Coefficient (bit) $B'_{d_{\mathcal{D}_2}^{\perp}}$	35	**1**	39	**1**	55	**1**	215	**32**

Table 5: Exhibiting different codes in (5, 2)-SSS scheme over $\mathbb{F}_{2^4}$. Note that we take $\alpha_1 = \alpha^i = 1$, $\alpha_2 = \alpha^j$, $\alpha_3 = \alpha^k$, $\alpha_4 = \alpha^l$ and $\alpha_5 = \alpha^r$.

	$j=1$, $k=4$ $l=6$, $r=12$	$j=1$, $k=4$ $l=6$, $r=11$	$j=1$, $k=2$ $l=3$, $r=11$	$j=3$, $k=6$ $l=9$, $r=12$	$j=1$, $k=3$ $l=5$, $r=8$
Minimum distance $d_{\mathcal{D}}$	4	4	4	4	4
Dual distance (word) $d_{\mathcal{D}}^{\perp}$	3	3	3	3	3
Dual distance (bit) $d_{\mathcal{D}_2}^{\perp}$	3	3	3	**4**	**4**
Coefficient (bit) $B_{d_{\mathcal{D}_2}^{\perp}}$	12	11	**1**	25	**17**
Coefficient (bit) $B'_{d_{\mathcal{D}_2}^{\perp}}$	20	19	**1**	225	**39**

B.5 A Special Example from [WMCS20]

As shown in Remark 2, there are some cases of GCM in which the side-channel security order can be greater than the dual distance of $\mathcal{D}$ minus one. In particular, Wang et al. [WMCS20] presented an example where the generator matrices of $\mathcal{C}$ and $\mathcal{D}$ as follows, respectively,

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix} \in \mathbb{F}_2^{2\times 8},$$

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{4\times 8}.$$

We can compute the generator matrices of the dual codes $\mathcal{C}^{\perp}$ and $\mathcal{D}^{\perp}$ as follows, respectively,

$$\mathbf{G}^{\perp} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{F}_2^{6\times 8},$$

$$\mathbf{H}^{\perp} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \in \mathbb{F}_2^{4\times 8},$$

where $\mathcal{C}^{\perp}$ is a code with parameters $[8, 6, 1]$ and $\mathcal{D}^{\perp}$ is of parameters $[8, 4, 2]$. We have $d_{\mathcal{D}}^{\perp} = d_{\mathcal{D}^{\perp}} = 2$ and $B_2 = 1$ for $\mathcal{D}^{\perp}$. Therefore, there is only one codeword $u = [1, 1, 0, 0, 0, 0, 0, 0] \in \mathcal{D}^{\perp}$ such that $w_H(u) = 2$. Since u is also in $\mathcal{C}^{\perp}$, which indicates that B'_2 equals 0. As a consequence, applying Theorem 2 gives that SNR equals 0 for $\deg(P) = d_{\mathcal{D}}^{\perp} = 2$ under Hamming weight leakages (e.g., $P(Z) = w_H(Z)$) and then the security order is at least equal to $d_{\mathcal{D}}^{\perp}$, rather than $d_{\mathcal{D}}^{\perp} - 1$. More generally, taking Theorem 1 gives the same conclusion for any leakage function P with $\deg(P) = 2$.

In particular, we checked that the first nonzero $B'_{d_{\mathcal{D}}^{\perp}}$ for nonzero codewords is $B'_3 = 3$. Therefore the security order is exactly 2 in above example.