Scalable Information Theoretic Evaluation of the Rank Statistics in Side-Channel Attacks

Julien Béguinot[1], Olivier Rioul[1], Loïc Masure[2], François-Xavier Standaert[3], Wei Cheng[1,4,5] and Sylvain Guilley[1,5]

[1] Télécom Paris, Palaiseau, France, firstname.lastname@telecom-paris.fr
[2] LIRMM, CNRS, Univ. Montpellier, firstname.lastname@lirmm.fr
[3] ICTEAM/ELEN/Crypto Group, Université Catholique de Louvain, Belgium, firstname.lastname@uclouvain.be
[4] School of Cyber Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
[5] Secure-IC S.A.S, firstname.lastname@secure-ic.com

Abstract. Evaluating the security of a device against side-channel attacks is a difficult task. One prominent strategy for this purpose is to characterize the distribution of the *rank* of the correct key among the different key hypotheses produced by a maximum likelihood attack, depending on the number of measured traces. In practice, evaluators can estimate some statistics of the rank that are used as security indicators—e.g., the arithmetic and geometric mean rank, the median rank, the α-marginal guesswork, or the success rate of level L. Yet, a direct estimation becomes time-consuming as security levels increase.

In this work, we provide new bounds on these figures of merit in terms of the mutual information between the secret and its side-channel leakages. These bounds provide theoretical insights on the evolution of the figures of merit in terms of noise level, computational complexity (how many keys are evaluated) and data complexity (how many side-channel traces are used for the attack). To the best of our knowledge, these bounds are the first to formally characterize security guarantees that depend on the computational power of the adversary, based on a measure of their informational leakages. It follows that our results enable fast shortcut formulas for the certification laboratories, potentially enabling them to speed up the security evaluation process. We demonstrate the tightness of our bounds on both synthetic traces (in a controlled environment) and real-world traces from two popular datasets (Aisylab/AES_HD and SMAesH).

Keywords: Side-Channel Analysis · Security Evaluations · Success Rate of Level L · Guessing Entropy · α-Marginal Guesswork · Mutual Information.

1 Introduction

Side-channel attacks are a powerful class of attacks that aim to recover the secret key used by a crypto-system using some side information obtained via physical measurements. The power consumption and electromagnetic emanations of an implementation are typical examples [KJJ99, GMO01].

Given a cryptographic device, its security level against such attacks is essentially characterized from the probability distribution of the rank of the correct key in the ranking of the different key hypotheses produced by the attack [VGS13].The worst-case evaluation of security shall ideally leverage an optimal side-channel attack procedure in order to best exploit the side-information [HRG14].

However, it is both computationally and data intensive to estimate this distribution completely. Therefore, certifying laboratories generally evaluate some figures of merit of the best side-channel attack instead. The most widespread figures of merit are the success rate of level[1] L and the guessing entropy (GE) of the key [Mas94], initially proposed as metrics in the side-channel context in [SMY09], and the α-marginal guesswork [Pli00b]. These figures of merit can be seen as the cumulative distribution function, the expectation, and the quantiles of the rank, respectively. Nevertheless, Martin et al. observed that the rank of the key has a very large variance in a divide-and-conquer setting [MMOS16]. As a consequence, the relevance of the guessing entropy as a security metric alone can be questioned, and combining the median and the geometric mean of the rank appears to be a more meaningful approach. Besides, the direct estimation of these figures of merit can remain computationally intensive, especially as the security levels of the target cryptographic implementations increase.

1.1 State of the Art and Contributions

Context. Mangard [Man04] first connected the probability of success of a *differential power analysis* (DPA) with the number of traces used for the attack depending on the *signal-to-noise ratio* (SNR) of the side-channel traces. Mangard's analysis was based on Pearson's correlation coefficient, which is limited to monovariate leakages. However, attacks based solely on correlation are suboptimal; in particular, they can miss some of the non-linear leakages. For this reason, mutual information has been introduced as a more reliable and conservative measure of side-channel leakages that is also better suited for multivariate leakages [SMY09]. Heuser et al. [HRG14] modeled the side-channel attack as an unintended communication channel of the key from the hardware to the attacker. They derived the *success exponent* (SE) of the *difference of mean* (DoM) [KJJ99], *correlation power analysis* (CPA) [BCO04] and *mutual information analysis* (MIA) [GBTP08] based on the *confusion coefficient* [GHR15]. While improving Mangard's derivations, these attacks are still suboptimal, which may lead to a false sentiment of security. De Chérisey et al. [dCGRP19a, dCGRP19b] bridged this gap to the maximum likelihood attack performance by exploiting Fano's inequality [Fan52] to bound the success rate (of level $L = 1$) of a maximum likelihood attack using the Shannon entropy. This corresponds to a worst-case attack scenario for the defender, where it is assumed that the adversary perfectly knows the leakage model. This model can be profiled offline with a clone device, for instance using *template attacks* [CRR02], leading to the question of how accurate is the model estimation—a problem that is orthogonal to our investigations and is sometimes referred to as *leakage certification* [DSV14, BHM+19, CRBO24].

Key Enumeration. In the divide-and-conquer setting, the side-channel attacker needs to recombine each key chunk to perform a guess of the full key. She can enumerate a given number of key hypotheses until she succeeds. This is the idea of DPA combined with computational power from the seminal work of Veyrat-Charvillon et al. [VGRS12]. This methodology is termed *key enumeration* in the side-channel literature and has been investigated thoroughly [VGRS12, MOOS15, DW17, BKM+15, PSG16, LMM+16, Gro18]. Key enumeration becomes computationally and memory intensive when the key rank increases.

Key Rank Estimation. A laboratory evaluator can benefit from the knowledge of the actual secret key to facilitate the estimation of the correct key rank. This methodology is called *key rank estimation* and was introduced in [VGS13]. It can be efficiently performed[2] even for large key using various approaches [GGP+15, MOOS15, PGS15, Gro18, DW19a, DW19b, DW21]. The MCRank probabilistic algorithm [CDS22] is another more efficient

[1] Often termed success rate of *order* L. We say "of *level* L" to avoid confusion with the masking order.
[2] An open-source implementation of the histogram methods is available on SCALib [CB23].

(yet probabilistic) way to estimate the rank even with very large keys with possibly non-independent key chunks.

Rank estimation algorithms can be used to build the *security graphs* introduced in [VGS13, Figs. 7 & 8]. Essentially, the problem of efficient (in terms of computational complexity and memory) key rank estimation can be considered as solved given the state-of-the-art algorithm. Yet, as already mentioned, this approach is computationally and data intensive because it requires evaluating the rank of the correct key for a significant number of independent attacks, for each number of traces.

Attack Score Vector as a Multivariate Gaussian Distribution. Rivain [Riv08] made the assumption/approximation that the *score vector* produced at the output of an attack is multivariate Gaussian. Under this assumption he obtained formulas to bound the success rate of level L for a key byte. Lomné et al. [LPR+14] extended this latter work in the setting of masked encodings. Finally, Zhang et al. [ZDF20] introduced the *GEEAA* and *PS-TH-GE* algorithms based on the same assumption to bound the guessing entropy of a full AES key. However, both the *PS-TH-GE* and *GEAA* reliabilities and relevance have been questioned in [YMO22]. This assumption cannot hold when the rank approaches 1. We do not rely on this type of assumptions in this article. Hence, since we do not need it within our derivations we prefer not to use it to be conservative in our analysis.

Distinction between GM and GE in the literature. In the side-channel literature a distinction has been introduced between two estimators GM and GE of the guessing entropy which can be confusing [RPC22]. It differentiates between two empirical *estimators* of the guessing entropy. Both quantities assume that N independent attacks are mounted. Note that it is not required to know the secret key to compute GM, while it is necessary to compute GE [3]. While their values can differ if the profiling is inaccurate or with a too few samples both estimators estimate the same quantity which is the guessing entropy i.e. the average rank for the optimal attack. For a full key, the value of M becomes prohibitively large (e.g. $M = 2^{128}$) to compute exactly GE and GM. For this reason, instead of computing the rank directly, the method of histograms [VGS13] enables to approximate efficiently rank_i.

Metric Based Evaluation. Our method is *metric-based*, it requires only to evaluate the leakages on a given dataset. It does not predict the rank for a given attack but rather bounds some chosen figure of merits (chosen summary statistics of the rank). The main advantage for this method is that it provides tendencies. It predicts the behavior of the attack if the noise level, the masking order, the number of traces or the enumeration power of the adversary changes.

We detail these motivations:

1. **Scalability:** Many side-channel attacks fall in a divide-and-conquer framework. The key of the crypto-system is split into several key chunks and side-channel information is gathered about each key chunk independently. This is typically the case of many attacks against the AES-128 where the 16 S-boxes' outputs after the initial AddRoundKey are targeted. While the figures of merit can be easily evaluated for one key chunk, it is much more challenging to evaluate them for the full key. Fortunately, informational metrics, such as entropy or mutual information, scale easily to the whole key. Hence, bounds on the figures of merit in terms of informational metrics are very relevant since they provide scalable shortcut formulas to the evaluators.

[3] Let $\widehat{p_{K_i|Y_i}}$ be the posterior likelihood of the different key hypotheses for the i-th attack with side information y_i and correct key k_i. Let τ_i be a permutation of the key space so that $\widehat{p_{K_i|Y_i}}(\tau_i(1)|y_i) \geq \ldots \geq \widehat{p_{K_i|Y_i}}(\tau_i(M)|y_i)$ where M is the size of the key-space. In particular, the rank of the correct key for the i-th attack is given by $\text{rank}_i = \tau_i^{-1}(k_i)$. On the one hand, the GM is defined by $\text{GM} = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{M} k\widehat{p_{K_i|Y_i}}(\tau_i(k)|y_i)$. On the other hand, the GE is defined by $\text{GE} = \frac{1}{N}\sum_{i=1}^{N}\text{rank}_i$.

2. **Dependency on the Parameters:** Bounds on the figures of merit provide insights on the impact of the noise, number of traces and computational power of the adversary. For instance, they permit to analyze the trade-off between data complexity (number of traces) and computational complexity (number of keys enumerated at the end of the attack). This is in contrast with a direct evaluation process, which should then be repeated for each number of traces and computational power of the adversary.
3. **Combination with Security Proofs and Countermeasures:** The bounds on the figures of merit can be combined with formal security proofs that bound the side-channel adversaries' success when a countermeasure is implemented [DFS15, MRS22, BCG+23, MS23, BCGR24]. They can be used to assess the impact of the masking order on the adversaries' success.

De Chérisey et al. [dCGRP19a, dCGRP19b] bounded the success rate of an attack without enumeration in terms of mutual information using Fano's inequality. De Chérisey et al. [dCGRP19a, § 6] also showed how to bound the guessing entropy of the key in terms of mutual information, using Massey's inequlity. Choudary et al. [CP17, TCRP21, RPC22] provided tight bounds on the guessing entropy based on Rényi entropy of order $\frac{1}{2}$ and Shannon entropy. Note that while the order $\frac{1}{2}$ for the Rényi entropy makes sense in terms of tightness given Arikan's inequalities [Ari96], there does not exist yet an operational shortcut formula such as *Mrs. Gerbers Lemma* for $H_{\frac{1}{2}}$ to facilitate the evaluation for masked implementations. Note that the bounds used in this approach have been recently slightly improved in [BR24]. As pointed out by Martin et al [MMOS16] the guessing entropy alone can be a misleading figure of merit. In this work, we derive metric-based shortcut formula for the first time for the median rank and to the log guessing entropy of the key. Hence, our extension to the log-guessing entropy and the median rank is an important complementary contribution for the evaluators.

Our method features the exact same computational and memory costs as the previous metric-based approach from the state of the art. It cannot be compared in terms of tightness since it bounds other figures of merit (it does not compete but rather complements the state-of-the-art).

Contributions. Including computational power with informational metrics has been a recurring issue for the side-channel analysis community. We lack expressions for:

1. the impact of the adversary's computational power. We bridge this gap by bounding the success rate of level L of the maximum likelihood attack and, consequently, the α-marginal guesswork and the median of the rank.
2. the *geometric mean* rank of the key, which was observed to be very close to its median rank [MMOS16]. We bridge this gap by using the *log-guessing entropy* and deriving the optimal bound between the log-guessing entropy and the entropy. This provides an alternative bound on the median rank of the key.

1.2 Outline and Technical Overview

The main notations are presented in Subsection 1.3 together with some useful, known preliminary results.

Section 2 derives a bound on the success rate of level L of a side-channel attack in terms of mutual information using a variation of Fano's inequality tailored for the list decoding problem [Eli57]. As a byproduct, we show how it provides a bound on the α-marginal guesswork, i.e., on the quantiles of the rank. As a particular case, we find a simple and explicit bound on the median rank of the key.

Section 3 derives a bound on the log-guessing entropy, i.e., the logarithm of the geometric mean rank of the key, in terms of mutual information using the variational characterization of entropy given by Gibbs' inequality.

The efficiency of our methods are assessed in Section 4 via simulations and in Section 5 on real-world measurements. Section 6 concludes by discussing the practical gain of

our methodology for evaluators, particularly with respect to the evaluation score in the Common Criteria [Joi20].

1.3 Notations

Throughout this article, M and L are strictly positive integers such that $L \leq M$. A random variable is denoted by an upper-case letter (e.g., X) while its realization is denoted by the corresponding lower-case letter (e.g., x). The support of the random variable is denoted by a calligraphic letter (e.g., $\mathcal{X}$). When $\mathcal{X}$ is discrete, its cardinality is denoted by $|\mathcal{X}|$. The probability distribution of a random variable X is denoted by an upper-case P_X while its probability mass function (pmf) or probability density function (pdf) is denoted by a lower-case p_X.

The joint distribution of a couple of random variable (X, Y) is denoted by $P_{X,Y}$ which factorizes as $P_X P_{Y|X}$, where $P_{Y|X}$ characterizes the *channel* (a.k.a. Markov kernel) [BCGR24], that is, the random transformation that maps a random variable X to a random variable Y. We then write $X \to \boxed{P_{Y|X}} \to Y$ or simply $X \to Y$ for short. The conditional transitional probability density/mass function is denoted by $p_{Y|X}$. Some usual probability distributions are as follows:

- $\mathcal{U}(\mathcal{X})$ is the uniform distribution on $\mathcal{X}$;
- $\mathcal{B}(p)$ is the Bernoulli distribution with parameter p;
- $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and standard deviation σ.

In the sequel, we rely on the conditional entropy (equivocation) $H(X|Y)$ and the mutual information $I(X;Y)$ for two random variables X, Y as defined in [PW23]. The Kullback-Leibler divergence [PW23] between two distributions P, R is denoted by $D_{\mathrm{KL}}(P\|R)$. The binary Kullback-Leibler divergence $d_{\mathrm{KL}}(p\|r) \triangleq D_{\mathrm{KL}}(\mathcal{B}(p)\|\mathcal{B}(r)) = p \log \frac{p}{r} + (1-p) \log \frac{1-p}{1-r}$ is defined as the Kullback-Leibler divergence between two Bernoulli distributions.

1.3.1 Side Channel as an Unintended Communication Problem

We consider a secret key $K = (K_1, \ldots, K_r) \in \mathcal{K} \triangleq \mathcal{X}^r$ composed of r key chunks where each key chunk is composed of n bits, i.e., $\mathcal{X} = \mathbb{F}_{2^n}$. For the AES-128, one has $r = 16$ and $n = 8$ which yield a 128-bit long key and a key space of cardinality $M \triangleq |\mathcal{K}| = |\mathcal{X}|^r = 2^{128}$.

We consider q uniformly distributed but publicly known (plain)-text T^q and the corresponding sensitive values X^q which are assumed i.i.d. given the secret K. We assume that the adversary obtains q leakages Y^q about the sensitive variables X^q, through the memoryless and stationary channel $X \to \boxed{P_{Y|X}} \to Y$. More precisely, in the divide-and-conquer setting, with a plaintext $T = (T_1, \ldots, T_r)$ we assume that the sensitive values decompose as $X = (X_1, \ldots, X_r)$ where X_i depends only on K_i, T_i and each chunk of the key leak independently, so that the overall channel factorizes as $P_{Y|X} = \prod_{i=1}^{r} P_{Y_i|X_i}$. A typical channel example is $X_i \to Y_i = w_H(X_i) + N_i$ where w_H denotes the Hamming weight function and N_i is an additive white Gaussian noise.

The adversary processes the side information with a distinguisher $\mathcal{D}$ to recover the secret K, as illustrated by the Markov chain

$$K \to (X^q, T^q) \to \boxed{P^q_{Y|X}, \mathrm{Id}^q} \to (Y^q, T^q) \to \boxed{\mathcal{D}} \to Z \tag{1}$$

where Z is typically either a key *guess* $\hat{K}$, a *list* of L guesses by decreasing order of likelihood $\hat{K}_L$, or a *full* ranking (i.e., a permutation of the key space.) R and Id is the identity channel. Since the full ranking information encompasses the other cases, we shall always assume in the following that the distinguisher outputs a full ranking R [4].

[4] Note that in practice, the side-channel analysis does not output the full ranking when the key space becomes large for computational and memory constraints.

In particular, de Chérisey et al. [dCGRP19a, dCGRP19b] proved the following linear bound on the mutual information of the adversary using q traces in terms of the "single letter" mutual information of the adversary (with one trace):

Proposition 1 (Linear Bound)**.** $I(K; Y^q, T^q) \leq I(K, T^q; Y^q) \leq I(X^q; Y^q) \leq qI(X; Y)$.

It can also be observed that the mutual information is bounded above by the entropy of K. This shows that the linear bound becomes loose when either q or $I(X;Y)$ is large.

In our setting, we observe that the mutual information for the full key "tensorizes" for each key chunk: $I(K; Y^q, T^q) = \sum_{i=1}^{r} I(K_i; Y_i^q, T_i^q)$, so that we may apply the linear bound to each key chunk as follows:

$$I(K; Y^q, T^q) = \sum_{i=1}^{r} I(K_i; Y_i^q, T_i^q) \leq \sum_{i=1}^{r} \min\{H(K_i), qI(X_i; Y_i)\}. \tag{2}$$

Applying the inequalities in this order improves the bound when the leakages differ for the various key chunks. Contrarily to the *linear* bound from Proposition 1, this approach can be easily seen to yield a concave and piecewise linear bound on the mutual information. This significantly improves the bound for uneven leakages on the different bytes.

Throughout the paper we regularly consider the common example of Hamming weight leakages (of the variable or of its S-Box) with Gaussian noise. In this case the SNR is $\frac{n}{4\sigma^2}$ and the corresponding mutual information is approximately

$$I(X; \mathrm{HW}(\mathrm{Sbox}(X)) + \sigma N) \lesssim \frac{1}{2} \log\left(1 + \frac{n}{4\sigma^2}\right) \tag{3}$$

where $X \sim \mathcal{U}(\mathbb{F}_2^n)$, it can be checked numerically that for $n = 8$ when $\sigma \geq 0.6$ then the relative error of this approximation is lower than 10^{-3}. This is an upper bound on mutual information. While being asymptotically equivalent for high noise to [BCPZ16, Section. 4.2] $\frac{n \log e}{8\sigma^2}$, this is always more precise especially for lower noise.

1.3.2 Definitions and Rationales of the Figures of Merit

We assume that the side-channel attack sorts the key hypotheses by decreasing order of likelihood. The success of the attack is characterized by the rank of the correct key in this ordering. Given the conditional probabilities $p_{K|Y^q,T^q}(\widehat{K}|Y^q, T^q)$, the *rank* of the correct key K, denoted by $\mathrm{rank}(K|Y^q, T^q)$ [IUH22] is the position of K in the probability mass function $p_{K|Y^q,T^q}(\widehat{K}|Y^q, T^q)$ arranged in descending order[5]. We use the following figures of merit:

- The *success rate of level* L SR_L is the probability that the rank is at most L:

$$\mathbb{P}_{s,L}(K|Y^q, T^q) \triangleq \mathbb{P}(\mathrm{rank}(K|Y^q, T^q) \leq L) = F(L) \tag{4}$$

  where F is the cumulative distribution function of $\mathrm{rank}(K|Y^q, T^q)$ [SMY09]. When no side-information is available, we say that there is a *blind guess* and we write $\mathbb{P}_{s,L}(K)$. For a uniformly distributed key, $\mathbb{P}_{s,L}(K) = L/M$. Operationally, an adversary that enumerates the list of at most L key hypotheses can succeed in her attack with probability at most $\mathbb{P}_{s,L}$.
- The *guessing entropy* [Mas94] GE (a.k.a. *guesswork* [Pli00a]) is the arithmetic mean of the rank of the key

$$G(K|Y^q, T^q) \triangleq \mathbb{E}[\mathrm{rank}(K|Y^q, T^q)]. \tag{5}$$

  We write $G(K)$ for a blind guess. For a uniformly distributed key, $G(K) = (M+1)/2$. Operationally, if an adversary enumerates all the key hypothesis until the secret is found, its average running time scales as $G(K|Y^q, T^q)$.

[5] Ties are resolved randomly, which does not affect the value of the figures of merit.

- Martin et al. [MMOS16] observed that the geometric mean of the rank is a more relevant security criterion than its arithmetic mean. This motivates us to use the *log-guessing entropy* (called *ranking entropy* by Martin et al [MMOS16]) that we formally define as

$$\mathrm{LG}(K|Y^q,T^q) \triangleq \mathbb{E}[\log \mathrm{rank}(K|Y^q,T^q)]. \tag{6}$$

  The exponential of the log-guessing entropy is the geometric mean of the rank. It is more conservative than the guessing entropy by Jensen's inequality:

$$\mathrm{LG}(K|Y^q,T^q) = \mathbb{E}[\log \mathrm{rank}(K|Y^q,T^q)] \leq \log \mathbb{E}[\mathrm{rank}(K|Y^q,T^q)] = \log G(K|Y^q,T^q). \tag{7}$$

  Equivalently, $\exp \mathrm{LG}(K|Y^q,T^q) \leq G(K|Y^q,T^q)$, which can also be seen as a consequence of the arithmetic-geometric inequality. We write $\mathrm{LG}(K)$ for a blind guess. For a uniformly distributed key, $\mathrm{LG}(K) = (\log M!)/M \approx \log(M/e)$ by Stirling's approximation.
- Pliam [Pli00a] introduced the *α-marginal guesswork* as a relevant security metric. It is defined as

$$L(\alpha)(K|Y^q,T^q) \triangleq \min_L\{L \mid \mathbb{P}_{s,L}(K|Y^q,T^q) \geq \alpha\} = F^{-1}(\alpha). \tag{8}$$

  This corresponds to the *pseudo*-inverse of the cumulative distribution function F i.e. the quantile function. In particular, $L(\frac{1}{2})(K|Y^q,T^q) \triangleq \mathrm{Med}(K|Y^q,T^q)$ is a *median* of the rank. We write $L(\alpha)(K)$ for a blind guess. For a uniformly distributed key, $L(\alpha)(K) = \lceil \alpha M \rceil$. Operationally, if an adversary wants to succeed in his attack with probability α, she needs to enumerate up to $L = L(\alpha)(K|Y)$ key hypotheses.

All these figures of merit can be applied at the scale of a byte of the key. For example, $\mathbb{P}_{s,L}(K_i|Y_i^q,T_i^q)$ is the probability of success of level L to guess the i-th chunk of the key.

2 Success Rate of Level L and α-marginal Guesswork

In this section, we derive bounds on the success rate of level L and α-marginal guesswork in terms of equivocation (equivalently, in terms of mutual information).

2.1 Success Rate of Level L

In this subsection, we derive bounds on the probability of success of level L. First, the probability of success of level L of a M-ary secret can always be lower bounded as follows

Proposition 2. *Let K be a M-ary secret, then $\mathbb{P}_{s,L}(K|Y,T) \geq \mathbb{P}_{s,L}(K) \geq L/M$.*

Proof. Without loss of generality we assume that $p_K(1) \geq p_K(2) \geq \ldots \geq p_K(M)$ so that $\mathbb{P}_{s,L}(K) = \sum_{i=1}^{L} p_K(i)$. First, observe that

$$\mathbb{P}_{s,L}(K) = \sum_{k=1}^{L} p_K(k) = \sum_{k=1}^{L} \mathbb{E}_{y,t} p_{K|Y,T}(k|y,t) = \mathbb{E}_{y,t} \sum_{k=1}^{L} p_{K|Y,T}(k|y,t) \tag{9}$$

$$\leq \mathbb{E}_{y,t} \mathbb{P}_{s,L}(K|Y=y,T=t) = \mathbb{P}_{s,L}(K|Y,T). \tag{10}$$

Second, for all $i > L$ we have

$$p_K(i) \leq p_K(L) \leq \sum_{i=1}^{L} p_K(i)/L = \mathbb{P}_{s,L}(K)/L. \tag{11}$$

But then $1 = \mathbb{P}_{s,L}(K) + \sum_{i=L+1}^{M} p_K(i) \geq \mathbb{P}_{s,L}(K) + \mathbb{P}_{s,L}(K)(M-L)/L = \mathbb{P}_{s,L}(K)M/L$. □

The next ingredients are so-called *Fano inequalities*. Ahlswede, Gács, and Körner [AGK76] used Fano's inequality for the *list decoding problem* where the decoder is allowed to test L values [Eli57]. While de Chérisey et al. [dCGRP19a, dCGRP19b] used Fano's inequality to bound the success rate of a maximum likelihood attack ($L = 1$). The following generalization of Fano's inequality to list decoding seems to have never been used for side-channel attacks with enumerations (which can be seen as a type of list decoding [PPS12]).

Lemma 1 (Fano Inequality for Success Rate of level L)**.** *Let K be a M-ary secret. The extension of Fano's inequality to the success rate of level L is*

$$I(K; Y^q, T^q) \geq d_{\mathrm{KL}}\big(\mathbb{P}_{s,L}(K|Y^q, T^q) \| \mathbb{P}_{s,L}(K)\big) \geq d_{\mathrm{KL}}\big(\mathbb{P}_{s,L}(K|Y^q, T^q) \| \tfrac{L}{M}\big) \tag{12}$$

where $\mathbb{P}_{s,L}(K) = L/M$ for a uniformly distributed key.

Proof. Consider the channel $K \to Y^q, T^q \to \widehat{K_L}$ where $K \in \mathcal{K}$ and $\widehat{K_L} \in \mathcal{K}^L$, then

$$I(K; Y^q, T^q) \overset{(a)}{\geq} I(K; \widehat{K_L}) \overset{(b)}{=} D_{\mathrm{KL}}(P_{K,\widehat{K_L}} \| P_K P_{\widehat{K_L}}) \tag{13}$$

$$\overset{(c)}{\geq} D_{\mathrm{KL}}(\mathcal{B}(\mathbb{P}_{s,L}(K|Y^q, T^q)) \| \mathcal{B}(\mathbb{P}_{s,L}(K))) \tag{14}$$

$$\overset{(d)}{=} d_{\mathrm{KL}}(\mathbb{P}_{s,L}(K|Y^q, T^q) \| \mathbb{P}_{s,L}(K)) \overset{(e)}{\geq} d_{\mathrm{KL}}(\mathbb{P}_{s,L}(K|Y^q, T^q) \| \tfrac{L}{M}) \tag{15}$$

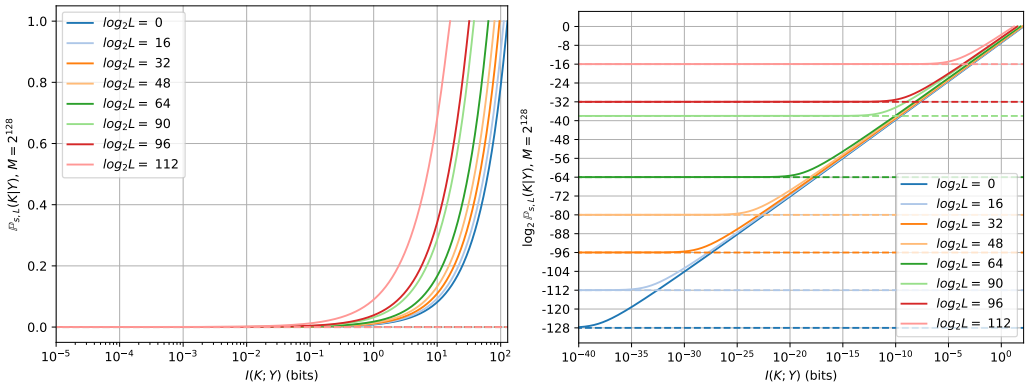
which proves the lemma. (a) is the data processing inequality for mutual information applied to the Markov chain $K \to Y^q, T^q \to \widehat{K_L}$. (b) holds by definition of the mutual information. (c) is the data processing inequality for the Kullback-Leibler divergence. On the left hand side the couple $(K, \widehat{K_L})$ is data processed into to the Bernoulli distribution defined by $\mathbf{1}_{K \in \widehat{K_L}}$ whose parameter is $\mathbb{P}_{s,L}(K|Y^q, T^q)$. On the right hand side the couple $(K', \widehat{K_L})$ where K' is an independent copy of K is data processed into to the Bernoulli distribution defined by $\mathbf{1}_{K' \in \widehat{K_L}}$ whose parameter is $\mathbb{P}_{s,L}(K)$. (d) holds by definition of the binary divergence. (e) holds by the data processing inequality for the binary divergence since by Proposition 2, $\mathbb{P}_{s,L}(K|Y^q, T^q) \geq \mathbb{P}_{s,L}(K) \geq \frac{L}{M}$. □

Figure 1 shows the bound from Lemma 1 for alphabet size $M = 2^{128}$. The bound allows to analyze the computational complexity (as measured by L) versus data complexity (as measured by q) of any side-channel attacks. The curves in Figure 1 have a noticeable elbow, we can distinguish two specific regimes in this respect, which aligns with the intuition:

For $L = 1$ one recovers the classical Fano inequality which is known to be obtained from the data processing inequality [PW23, Thm. 3.12]. As explained by Sakai [Sak20, Eqn. 11] Fano's inequality can be easily generalized to the list decoding problem. The following proof (for completeness since it is known) relies on the data processing inequalities for the Kullback-Leibler divergence and mutual information. This bound on the probability of success of level L is especially interesting as it covers attacks were the adversary enumerates several key hypotheses. This is the first security bound accounting for the enumeration power of the adversary in the side channel literature.

The previous bound for $L = 1$ [dCGRP19a] established that when the mutual information tends to zero, an adversary with a single trial could not do better than a random guess. However, this did not prove security against adversaries that enumerate *several* key hypotheses. We bridge this gap using this generalized Fano inequality for any number L of enumerated key hypotheses.

- First, accumulating more traces only affects very mildly the success rate of level L of the attack which can be seen by the plateau on the left of Figure 1b. Furthermore, the effect of the computational complexity L is linear in the probability of success of the attack. In the bound, enumerating $L = 2^l$ key candidates amounts to apply the



(a) Linear Scale (b) Log Scale

Figure 1: Fano for List Decoding, with $M = 2^{128}$ and various values of L. The curve have a noticeable elbow shape in both the linear and log scale. The X-axis corresponds to the mutual information expressed in bits. If we assume that the linear bound is valid then this axis is proportional to the number traces. In any case the mutual information increases with the number of traces so that the impact of the number of traces can be read from the graph.

usual Fano's inequality with $L = 1$ but where the size of the key space $M = 2^{nr}$ is replaced by an effective size of the key space 2^{nr-l}. In this sense, enumerating 2^l keys reduces roughly the entropy of the key by l bits which explains this factor.

- Second, the impact of the number of traces is almost linear with the probability of success. Increasing the number of traces is the best way to improve the attack performance in this regime. However, the impact of the number of enumerated key hypothesis L is notable but milder as can be observed in Figure 1a.

The COPACOBANA FPGA clusters [KPP+06] from 18 years ago could enumerate 2^{35} keys per second with a setup costing less than 10k dollars. In SMAesH challenge[6] from CHES2023, up to $L = 2^{68}$ keys were enumerated which is the number of blocks hashed by the Bitcoin mining network in only one second. The choice of L in the theorem should be guided by the expected computational power of the adversary. In the figure, we choose for illustrative purposes to plot it for conventional values of L such that $\log_2 L$ is a multiple of 16. Given the current enumeration power of modern adversary the choices $L = 2^{64}$ or more conservatively $L = 2^{90}$ seems relevant.

2.2 Lower Bound on the Number of Traces for the Probability of Success of Level L

We can leverage Lemma 1 to obtain a lower bound on the minimum number of traces q needed to achieve a given success rate of level L:

Theorem 1 (Lower Bound on the Number of Traces)**.** *The minimum number of side channel queries q which are necessary to reduce the probability of success of level of the key $\mathbb{P}_{s,L}(K|Y^q, T^q)$ to α is lower bounded as follows:*

$$d_{\mathrm{KL}}(\alpha \| L/M) I(X;Y)^{-1} \leq q. \tag{16}$$

Theorem 1 generalizes the lower bound on q for the success rate of level $L = 1$ [dCGRP19a, Eqn. 15] to an arbitrary value of L hence taking into account the computational

[6] https://smaesh-challenge.simple-crypto.org/

power as used in side-channel attacks with key enumeration [VGRS12]. If a countermeasure such as masking is implemented it can be combined with upper bound on $I(X;Y)$ depending on the masking order [BCG+23, MS23].

Proof. By Lemma 1 with the linear bound (Proposition 1) $qI(X;Y) \geq I(K;Y^q,T^q) \geq d_{\mathrm{KL}}(\mathbb{P}_{s,L}(K|Y^q,T^q)\|L/M)$. □

2.3 Bound on the α-marginal Guesswork

We first begin by a preliminary upper bound on the α-marginal guesswork:

Proposition 3. *For any M-ary secret K (not necessarily uniform) and $\alpha \in (0,1)$:*

$$L(\alpha)(K|Y,T) \leq L(\alpha)(K) \leq \lceil \alpha M \rceil. \tag{17}$$

In particular, for any y,t, $L(\alpha)(K|Y=y,T=t) \leq \lceil \alpha M \rceil$ e.g. $\mathrm{Med}(K|Y,T) \leq \lceil M/2 \rceil$.

Proof. Let $\alpha \in (0,1)$. By definition, $L(\alpha)(K) = \min E$ where $E = \{L\|\mathbb{P}_{s,L}(K) \geq \alpha\}$ and $L(\alpha)(K|Y,T) = \min E_{Y,T}$ where $E_{Y,T} = \{L\|\mathbb{P}_{s,L}(K|Y,T) \geq \alpha\}$. Now by Proposition 2, $\mathbb{P}_{s,L}(K|Y,T) \geq \mathbb{P}_{s,L}(K)$ so that $E \subseteq E_{Y,T}$. Since the minimum of a set is at most equal to the minimum of one of its subset, the inequality $L(\alpha)(K|Y,T) \leq L(\alpha)(K)$ follows. If $\frac{L}{M} \geq \alpha$ then by Proposition 2, $L \in E$ and $L(\alpha)(K) \leq L$. This is valid for all integer L such that $L \geq \alpha M$ i.e. $L \geq \lceil \alpha M \rceil$. We obtain the result by taking $L = \lceil \alpha M \rceil$. □

We now derive a lower bound on $L(\alpha)(K|Y)$ which we expect to approach $\lceil \alpha M \rceil$ for a uniform key as $I(K;Y) \to 0$. The α-marginal guesswork $L(\alpha)$ can be lower bounded using Lemma 1 using the following lemma:

Lemma 2. *Let $\phi : L \mapsto [0,1]$ be an upper bound on $\mathbb{P}_{s,L}$ then,*

$$L(\alpha) = \min\{L|\mathbb{P}_{s,L} \geq \alpha\} \geq W(\alpha) \triangleq \min\{L|\phi(L) \geq \alpha\}. \tag{18}$$

Proof. Since $\mathbb{P}_{s,L} \leq \phi(L)$ for any $\alpha \in [0,1]$ we have the inclusion $\{L|\mathbb{P}_{s,L} \geq \alpha\} \subseteq \{L|\phi(L) \geq \alpha\}$. Hence, the results since the minimum on a subset of a set is always larger than the minimum on the set itself. Lemma 2 can be visually understood with the illustration of Figure 2. □

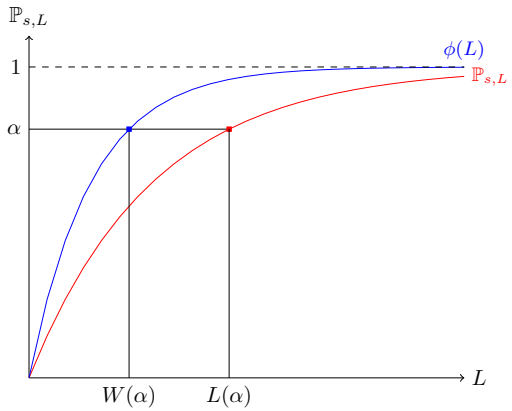


Figure 2: Visual Proof of Lemma 2.

As a consequence the upper bound on the success rate of level L (Lemma 1) yields a bound on the α-marginal guesswork:

Theorem 2. *Let* $\alpha \in (0,1)$,

$$L(\alpha)(K|Y^q,T^q) \geq \min\left\{L|I(K;Y^q,T^q) \geq d_{\mathrm{KL}}\big(\alpha\|\mathbb{P}_{s,L}(K)\big)\right\}. \tag{19}$$

For the median rank of an a priori uniformly distributed key we obtain:

$$\mathrm{Med}(K|Y^q,T^q) \geq \frac{M}{2}\left(1-\sqrt{1-\exp(-2I(K;Y^q,T^q))}\right) \tag{20}$$

Proof. We combine Lemma 1 with Lemma 2. The median is obtained for $\alpha = \frac{1}{2}$ and

$$d_{\mathrm{KL}}(\frac{1}{2}\|\mathbb{P}_{s,L}(K)) = \frac{1}{2}(\log\frac{1}{2\mathbb{P}_{s,L}(K)} + \log\frac{1}{2(1-\mathbb{P}_{s,L}(K))}) = -\frac{1}{2}\log\frac{4L(M-L)}{M^2}. \tag{21}$$

Hence, we have to solve the degree polynomial in L,

$$I(K;Y^q,T^q) = -\frac{1}{2}\log\frac{4L(M-L)}{M^2}. \tag{22}$$

The discriminant is always positive so that we obtain two roots

$$L = \frac{M}{2}(1 \pm \sqrt{1-\exp(-2I(K;Y^q,T^q))}). \tag{23}$$

Now by Proposition 3 we know that $\mathrm{Med}(K|Y^q,T^q) \leq M/2$ hence we can remove one of the two roots to obtain Eqn. (20). □

A bound on the median rank can be very interesting for the evaluator. It is less sensitive to large deviation than the average rank (as measured by the guessing entropy). Furthermore, it answers the following question: *How many key hypotheses should be enumerated to achieve a success rate of* $\frac{1}{2}$ *?*

A dichotomic search over L can be used to compute $L(\alpha)$ in Theorem 2. As the leakage $I(K;Y^q,T^q)$ approaches zero then necessarily $\mathbb{P}_{s,L}(K)$ should approach α for L to satisfy $I(K;Y^q,T^q) \geq d_{\mathrm{KL}}\big(\alpha\|\mathbb{P}_{s,L}(K)\big)$. In other words L should approach αM. This is in line with the upper bound from Proposition 3. Theorem 2 yields the bounds from Figure 3.

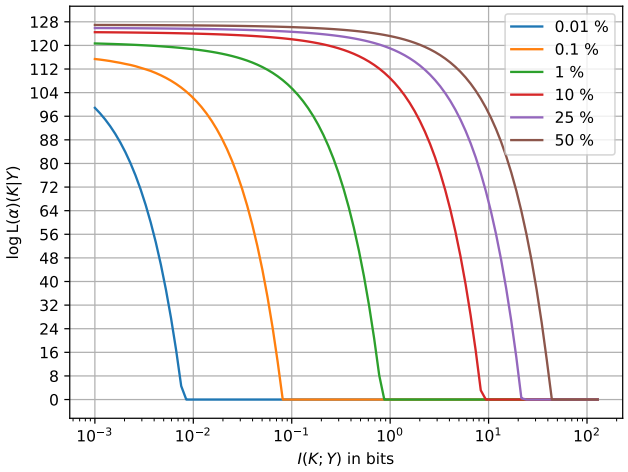


Figure 3: Illustration of Theorem 2 for a 128 bits key.

Remark 1 (Tightness). When the mutual information becomes large Theorem 2 becomes vacuous. This happens when

$$\frac{M}{2}(1-\sqrt{1-\exp(-2I(K;Y^q,T^q))}) \leq 1 \tag{24}$$

i.e.

$$I(K;Y^q,T^q) = -\frac{1}{2}\log(\frac{4}{M}(1-\frac{1}{M})) \approx \frac{nr-2}{2} \tag{25}$$

bits. In other words, our bound on the median becomes vacuous when approximately half of the secret is leaked.

Example 1 (Behavior For Hamming Weight Leakages)**.** For Hamming weight leakages, using Eqn. (3), we have approximately

$$\mathrm{Med}(K|Y^q,T^q) = \frac{M}{2}\Big(1-\sqrt{1-\Big(1+\frac{n}{4\sigma^2}\Big)^{-qr}}\Big) \tag{26}$$

Now if $\sigma \to +\infty$,

$$\mathrm{Med}(K|Y^q,T^q) \geq \frac{M}{2}\Big(1-\frac{\sqrt{qrn}}{2\sigma}\Big). \tag{27}$$

2.4 Lower Bound on the Number of Traces for the Median Rank

Leveraging Theorem 2, we lower bound the minimum number of traces q that are needed to achieve a given median rank of the key:

Theorem 3 (Lower Bound on the Number of Traces)**.** *The minimum number of side channel queries q which are necessary to reduce the median rank of the key* $\mathrm{Med}(K|Y^q,T^q)$ *to* Med *is lower bounded as follows:*

$$q \geq -\frac{1}{2I(X;Y)}\log\Big(4\frac{\mathrm{Med}}{M}\big(1-\frac{\mathrm{Med}}{M}\big)\Big). \tag{28}$$

Proof. Theorem 2 implies that

$$\mathrm{Med}(K|Y^q,T^q) \geq \frac{M}{2}(1-\sqrt{1-\exp(-2I(K;Y^q,T^q))}). \tag{29}$$

Hence,

$$1-\exp(-2I(K;Y^q,T^q)) \geq (1-\frac{2}{M}\mathrm{Med}(K|Y^q,T^q))^2. \tag{30}$$

That is

$$4\frac{\mathrm{Med}(K|Y^q,T^q)}{M}(1-\frac{\mathrm{Med}(K|Y^q,T^q)}{M}) \geq \exp(-2I(K;Y^q,T^q)). \tag{31}$$

Using the linear bound (Proposition 1),

$$qI(X;Y) \geq I(K;Y^q,T^q) \geq -\frac{1}{2}\log(4\frac{\mathrm{Med}(K|Y^q,T^q)}{M}(1-\frac{\mathrm{Med}(K|Y^q,T^q)}{M})) \tag{32}$$

which concludes the proof. □

3 Log Guessing Entropy or Geometric Mean Rank

The logarithm of the geometric mean rank of the key is an interesting security metric for an evaluator as it represents the average of the number security bits of the key. As observed by Martin et al. [MMOS16] this is more reliable metric than the logarithm of the guessing entropy. This quantity is captured by the log-guessing guessing entropy. In this section, we derive both a lower and an upper bound on the log-guessing entropy in terms of equivocation (or mutual information) whose evaluation scales to the full key.

3.1 Lower Bound on Log-Guessing Entropy

We derive the optimal lower bound on conditional log-guessing entropy in terms of conditional entropy:

Theorem 4. *Let* $\phi_M(a) \triangleq \ln(\sum_{i=1}^{M} i^{-a})$. *For all* $a \in \mathbb{R}^+$,

$$H(K|Y^q, T^q) \leq a\mathrm{LG}(K|Y^q, T^q) + \phi_M(a)(\log e). \tag{33}$$

The optimal lower bound on $\mathrm{LG}(K|Y^q, T^q)$ *in terms of* $H(K|Y^q, T^q)$ *is given by a parametric curve* $\{(H(a), \mathrm{LG}(a))|a \in \mathbb{R}^+\}$, *where*

$$\begin{aligned} H(a) &\triangleq (\phi_M(a) - a\frac{\mathrm{d}\phi_M}{\mathrm{d}a}(a))(\log e), \\ \mathrm{LG}(a) &\triangleq -\frac{\mathrm{d}\phi_M}{\mathrm{d}a}(a)(\log e). \end{aligned} \tag{34}$$

Proof. First, observe that Eqn. (33) is linear in H and LG hence it is enough to prove it in the unconditional setting. The conditional version is obtained from the unconditional one by averaging the inequality for each $Y^q, T^q = y^q, t^q$.

Let p be the secret's pmf, without loss of generality we assume that

$$p(1) \geq p(2) \geq \ldots \geq p(M). \tag{35}$$

For any pmf r we have by Gibbs' inequality (Thm. 2.6.3 in[CT01]) that

$$H(K) \leq -\sum_{k=1}^{M} p(k) \log r(k). \tag{36}$$

The upper bound should depend on the secret only through it log-guessing entropy $\mathrm{LG}(K)$.

This is the case when $\log r(k) = -a \log k + b$ where b is a normalization factor i.e.

$$r(k) = k^{-a}(\sum_{k=1}^{M} k^{-a})^{-1} \tag{37}$$

is a Zipf's law (a.k.a. discrete Pareto distribution) of parameter $a \in \mathbb{R}^+$. We obtain,

$$H(K) \leq -\sum_{k=1}^{M} p(k)(-a \log k - \phi_M(a)(\log e)) = a\mathrm{LG}(K) + \phi_M(a)(\log e) \tag{38}$$

which proves Eqn. (33). Furthermore, Gibbs' inequality is achieved with equality when $p = r$ for which

$$\mathrm{LG}(K|Y^q, T^q) = \mathrm{LG}(a) = -\frac{\mathrm{d}\phi_M}{\mathrm{d}a}(a)(\log e) \tag{39}$$

and

$$H(K|Y^q, T^q) = H(a) = (\phi_M(a) - a\frac{\mathrm{d}\phi_M}{\mathrm{d}a}(a))(\log e) \tag{40}$$

which proves Eqn. (34). □

When $a = 1$, $\phi_M(1) \approx \ln(H_M)$ where $H_M = \ln M + \gamma + \frac{1}{2M} + o(\frac{1}{M})$ is the M-th harmonic number and $\gamma \approx 0.577$ is Euler–Mascheroni's constant. When $a \neq 1$ we suggest to use comparison with integral techniques such as [Nal15, Eqn. 16] to approximate tightly $\phi_M(a)$ and its derivative to avoid the naive computation of the sum in their definition whose complexity would be linear in M which can be intractable.

This bound on the log-guessing entropy is useful for the evaluator. It is less sensitive to large deviation than the guessing entropy. In a sense it gives the average security parameter of the implementation[7].

[7] while the log of the guessing entropy would be the security parameter associated to the average security

Remark 2 (Tightness). David and Wool [DW15, § 5] experimentally observed that the distribution of the rank is long-tailed and well approximated by a Pareto distribution. Since the rank is discrete and bounded, their observation can be refined by saying that the rank is well approximated by a truncated Zipf law (which is the discrete analog of a truncated Pareto distribution). Martin et al. [MMOS16, § 4.2] reported that the rank is well approximated by a delta-log-normal distribution which is another type of "heavy tailed" distribution. As Theorem 4 is achieved with equality for a truncated Zipf law we expect our bound to be very tight on practical side-channel traces.

3.2 Upper Bound on Log-Guessing Entropy

McEliece and Yu derived an optimal upper bound on the guessing entropy in terms of entropy [MY95, Theorem 1, Eqn. 5]. We modify their derivations to obtain an upper bound on the log-guessing entropy in terms of entropy:

Theorem 5 (Upper Bound on LG)**.** *The following upper bound on the log-guessing entropy holds*

$$\mathrm{LG}(K|Y^q,T^q) \leq \log(M!)(M\log M)^{-1}H(K|Y^q,T^q). \tag{41}$$

Note that $\frac{\log M!}{M\log M} \approx 1-\frac{\log e}{\log M}$ *approaches* 1 *as* M *increases. This inequality is tight in the sense that if* K *is uniformly distributed and the channel* $K \to Y^q,T^q$ *is an erasure channel with erasure probability* ϵ *then the bound is matched everywhere as* ϵ *varies from* 0 *to* 1*.*

Proof. We slightly modify the derivations of McEliece and Yu [MY95] that derived an upper bound on the guessing entropy in terms of entropy to obtain a bound on the log-guessing entropy in terms of entropy. Let

$$K_M = \mathrm{ConvHull}(\{a_1,\ldots,a_M\}) = \{p : 0 \leq p_M \leq \ldots \leq p_1 \text{ and } \sum_{i=1}^{M} p_i = 1\} \tag{42}$$

be a convex compact subset of the M-ary probability simplex where $a_i = (\frac{1}{i},\ldots,\frac{1}{i},0,\ldots,0)$ for $1 \leq i \leq M$ and $f : p \in K_M \mapsto H(a_M)\mathrm{LG}(p) - \mathrm{LG}(a_M)H(p) \in \mathbb{R}$. Since the entropy $H(p)$ is concave in p and the log-guessing entropy is linear in p, f is convex in p. As a result, f reaches its maximum value in one of the corners a_i whose convex hull spans K_M i.e. $\max_{P\in K_M} f(P) = \max_{i=1,\ldots,M} f(a_i)$. Now observe that $f(a_1) = 0$ and for $i \geq 2$,

$$f(a_i) = H(a_i)H(a_M)(\mathrm{LG}(a_i)H(a_i)^{-1} - \mathrm{LG}(a_M)H(a_M)^{-1}) \leq 0 \tag{43}$$

since the sequence $\mathrm{LG}(a_i)H(a_i)^{-1} = \log(i!)(i\log i)^{-1}$ is monotonically increasing for $i \geq 2$. As a consequence we have shown that if K is a M-ary random variable then

$$\mathrm{LG}(K) \leq H(K)\log(M!)(M\log M)^{-1}. \tag{44}$$

Now, by averaging for each $(K|Y^q = y^q, T^q = t^q)$ we obtain Eqn. (41) as desired. □

An upper bound on the log-guessing entropy in terms of entropy is useless for security evaluations as we access to upper bound on the mutual information and not lower bounds. However, it is interesting as it allows us to assess the worst case gap between our lower bound on the log-guessing entropy and its actual value. The combination of Theorem 5 with Theorem 4 (with $a = 1$) yields the following sandwiching bound,

$$H(K|Y^q,T^q) - \log(H_M) \leq \mathrm{LG}(K|Y^q,T^q) \leq H(K|Y^q,T^q) \tag{45}$$

where $H_M = \ln M + \gamma + \frac{1}{2M} + o(\frac{1}{M})$ is the M-th harmonic number. This shows that the log-guessing entropy can be accurately bounded in terms of entropy as the sandwiching bound roughly indicates that $\mathrm{LG}(K|Y^q,T^q) \approx H(K|Y^q,T^q)$. The maximal gap between the lower and upper bound is very moderate as it is approximately $\log(\ln M)$ as M grows.

Example 2 (Behavior For Hamming Weight Leakages)**.** For Hamming weight leakages, using Eqn. (3), we have approximately

$$nr - \frac{qr}{2}\log(1+\frac{n}{4\sigma^2}) - \log(nr\ln 2+\gamma) \lessapprox \mathrm{LG}(K|Y^q,T^q) \lessapprox nr - \frac{qr}{2}\log(1+\frac{n}{4\sigma^2}). \quad (46)$$

3.3 Lower Bound on the Number of Traces for Log-Guessing Entropy

Combining Theorem 4 with the linear bound (Proposition 1) we obtain a lower bound on the minimum number of side channel traces needed by an adversary to reduce the log-guessing entropy of the key to a targeted value LG:

Theorem 6 (Lower Bound on the Number of Traces)**.** *The minimum number of side channel queries q which are necessary to reduce the log-guessing entropy of the key* $\mathrm{LG}(K|Y^q,T^q)$ *to* LG *is lower bounded by* $(\log(M/H_M)-\mathrm{LG})I(X;Y)^{-1}$.

3.4 Discussion on the Bound Tightness

How tight are our bounds? Since we upper-bound the mutual information with a bound linear in the number of traces we cannot sandwich the log-guessing entropy with our method. However, we can plot in Figure 4 both the upper and lower bound for a known value of the mutual information in order to assess the gap between the upper and lower bound. This shows that the maximal gap between the upper and lower bound is at most $\log H_M \approx \log(\log M) \approx \log(nr) = 8$ bits. Furthermore, this gap decreases when the log-guessing entropy approaches its minimal or maximal value. Typically, when the mutual information exceeds 124 bits then the maximal gap is at most 2.5 bits. If the mutual information does not exceed 4 bits then the maximal gap is at most 2.2 bits. Depending on the application this worst-case precision may be sufficient or not but it is the best we can achieve using mutual information and without further assumptions. Actually the main source of gap in our final result (as we will see on the simulations) does not come from our bound between the mutual information and a figure of merit but rather from the linear bound on mutual information which becomes too conservative for a large number of traces. Improving upon the linear bound is orthogonal to this article but would as corollary directly

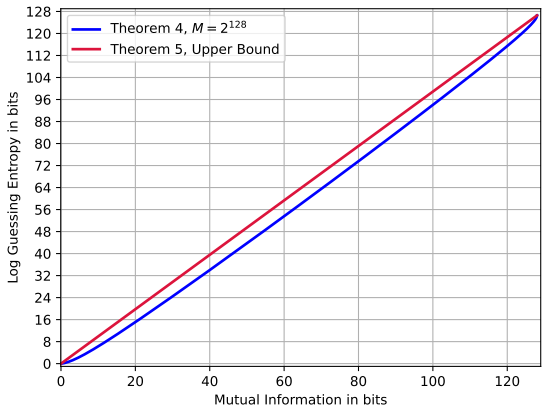


Figure 4: Upper and Lower Bound on $\mathrm{LG}(K|Y)$ in terms of mutual information $I(K;Y)$ for $M = 2^{128}$.

4 Synthetic Data Evaluation

We evaluated the performance of the a maximum likelihood attack on an AES where the Hamming weight of the 16 S-boxes of the first round is leaking perturbed by additive Gaussian noise of standard deviation σ. We plot the result for $\sigma \in \{1, 4, 16, 32\}$ i.e. SNR $\in \{2, 2^{-3}, 2^{-7}, 2^{-11}\}$. Results are given in Figure 5 and Figure 6.

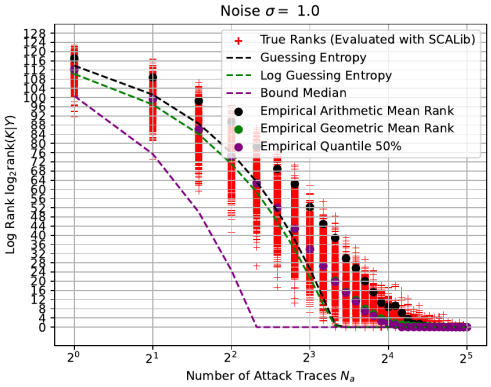
We estimate the figure of merits using the ranks evaluated with histograms as implemented in SCALib [CB23] and then plugging these values in the different estimators. Let $\widehat{\text{rank}}_i$ be the estimation of the rank for the i-th attack. We estimated the guessing entropy, log guessing entropy and median rank by the respective expressions $\frac{1}{N}\sum_{i=1}^{N} \widehat{\text{rank}}_i$, $\frac{1}{N}\sum_{i=1}^{N} \log \widehat{\text{rank}}_i$ and $\text{median}(\{\widehat{\text{rank}}_i, i = 1, \ldots, N\})$. In the simulations $N = 1000$, for the real dataset (in the next section) N is the minimum between 1000 and the number of available traces divided by the number of traces used to mount the attack. For the simulations, we computed the exact mutual information with a Monte-Carlo estimator using $N_s = 10^4$ samples.

We confirm the observation of [MMOS16] that the log-guessing entropy (geometric mean rank) is close to the median rank of the correct key. Our bound on the log-guessing entropy (Theorem 4, green dashed line) is close to the empirical geometric mean and median rank of the key. Our bound on the median rank (Theorem 2, purple dashed line) is tight for small leakages but becomes looser as the number of traces increases. In this case, it is better to use our bound on the log-guessing entropy as a bound on the median, under the assumption that the geometric mean of the rank and its median are close.

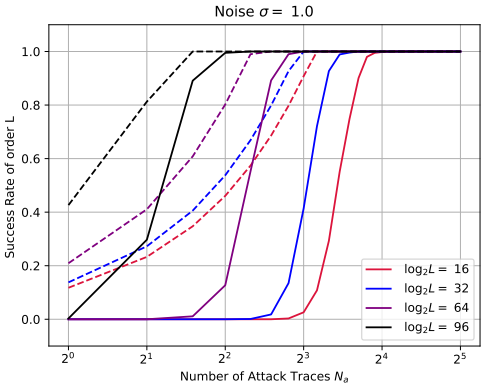
The experiments also confirm that the arithmetic mean rank of the key can be a misleading security parameters as it is drifted above by the worst attack performance. The blacks dots are at the very top of the cloud of red crosses. This shows that the rank of the correct key in most attacks can be order of magnitude bellow the average rank of the correct key. For instance, with $q = 2^{10}$ and $\sigma = 32$ we observed that approximately 95% of the attack ranks are bellow the arithmetic mean rank for these parameters. Hence, it is more meaningful to consider the geometric mean rank of the correct key, which is the average security parameter in a sense, as a security parameter (as already advocated by Martin et al. [MMOS16]). We also plot the improved bound on the guessing entropy [BR24, Theorem 1] (black dashed line) which is close to the empirical arithmetic mean of the rank for small leakages (black dots), but the gap widens as the number of traces increases (the bound seems closer to the geometric mean rank than the arithmetic mean rank).

We also confirm the tightness of Lemma 1 on the success rate of level for different values of L. We observe a multiplicative gap with respect to the number of traces between the bound and the empirical values. Interestingly, this gap seems approximately constant with respect to L. We observe on Figure 5 and 6 that our bounds on the median rank (purple dashed line) and log-guessing entropy (green dashed line) are close to the actual values (purple and green dots respectively) for rank larger or equal to 2^{40}. This holds in particular for the range we are interested in as an evaluator, namely range larger than $2^{70}, 2^{808}$. When the number of traces becomes too large the linear bound (Proposition 1) becomes looser which impacts the tightness of the final results. We attribute this gap not to our bound but to the non-tightness of the linear bound for large number of traces. In particular, for ranks lower than 2^{40} our bounds on the median rank and log-guessing entropy becomes looser. However, from an evaluator perspective the implementation can be considered as broken for ranks in this order of magnitude. For this reason, the gap with our bounds does not matter since proving security for an insecure implementation does not make sense. In all our experiments (Sections 4 and 5) we estimate the figures of merits by first estimating the rank using the histogram method implemented in SCALib [CB23] and then using this value in the different estimator.

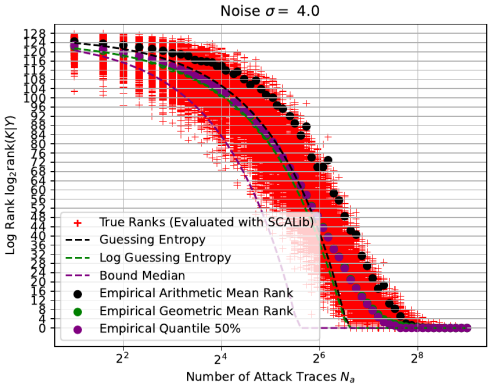
[8] As discussed in Section 2.1



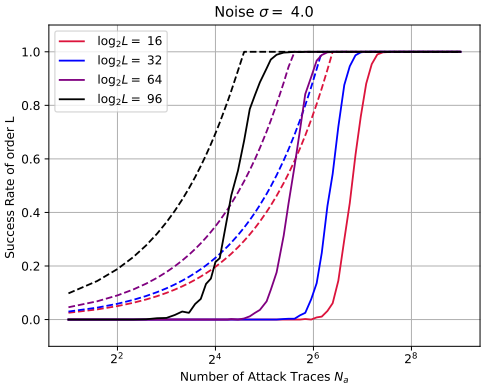
(a) Rank and $\sigma = 1$



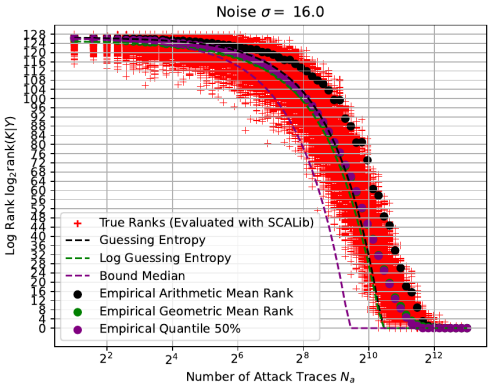
(b) Success Rate of Level L and $\sigma = 1$



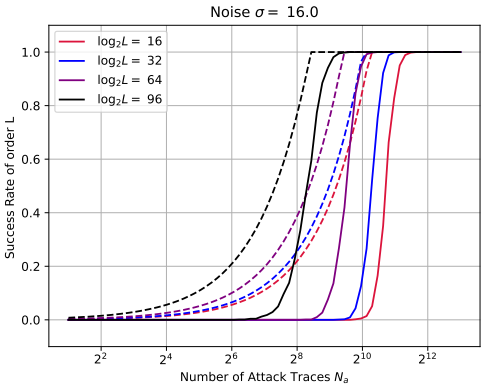
(c) Rank and $\sigma = 4$



(d) Success Rate of Level L and $\sigma = 4$

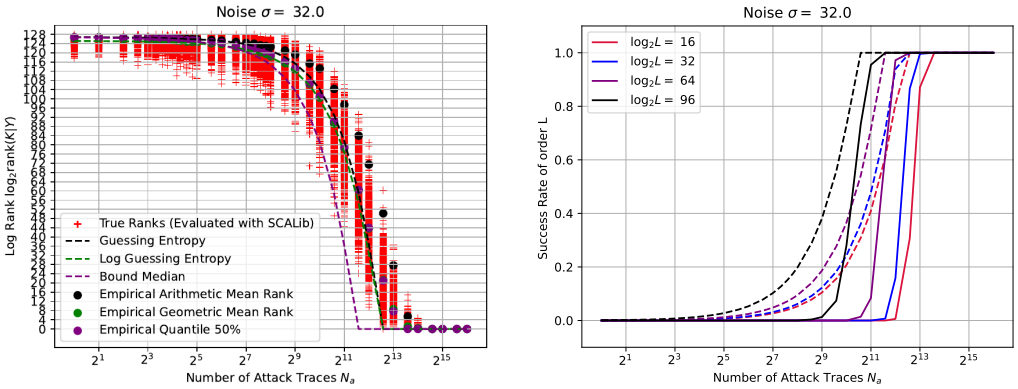


(e) Rank and $\sigma = 16$



(f) Success Rate of Level L and $\sigma = 16$

Figure 5: Success Rate of Level L and rank evaluated on 1000 attacks for different number of traces and various noise levels. The ranks evaluated using histograms on the attacks are plotted as red crosses. The dots represent the empirical values for the figures of merit while the dashed lines are the informational bounds that we derived in terms of mutual information.



(a) Rank and $\sigma = 32$ (b) Success Rate of Level L and $\sigma = 32$

Figure 6: Success Rate of Level L and rank evaluated on 1000 attacks for different number of traces and various noise levels. The ranks evaluated using histograms on the attacks are plotted as red crosses. The dots represent the empirical values for the figures of merit while the dashed lines are the informational bounds that we derived in terms of mutual information.

5 Real Data Evaluation

Experimental Setup and Inaccurate Mutual Information. We profiled a Regression-Based Linear Discriminant Analysis (RLDA) template using SCAlib and evaluated the corresponding mutual information as the average cross-entropy [CDSU23]. When using our methodology on real data where the ground-truth mutual information is unknown, a natural question is to assess the impact of an inaccurate estimation of the mutual information plugged into our bound. To show the impact of this limitation we added the inverse of the square root of the number of sample samples to the evaluation of the mutual information of each share. As a result, we obtain a region whose width shows the impact of an inaccurate mutual information evaluation. With real traces the ground-truth mutual information is unknown and so is the value of the estimation error. However, in [MCHS23] it is reported that the estimation error is generally decaying as the inverse square root of the number of sample. For this reason, we selected an illustrative error term which is decaying as the inverse square root of the number of sample. We do not claim that it corresponds to the actual error term which is unknown. This illustrates that confidence intervals on the mutual information evaluation can be automatically translated into confidence intervals on the figures of merit with our bounds. As we will see on both datasets, this impact is mild on the final evaluation.

5.1 AES HD Dataset

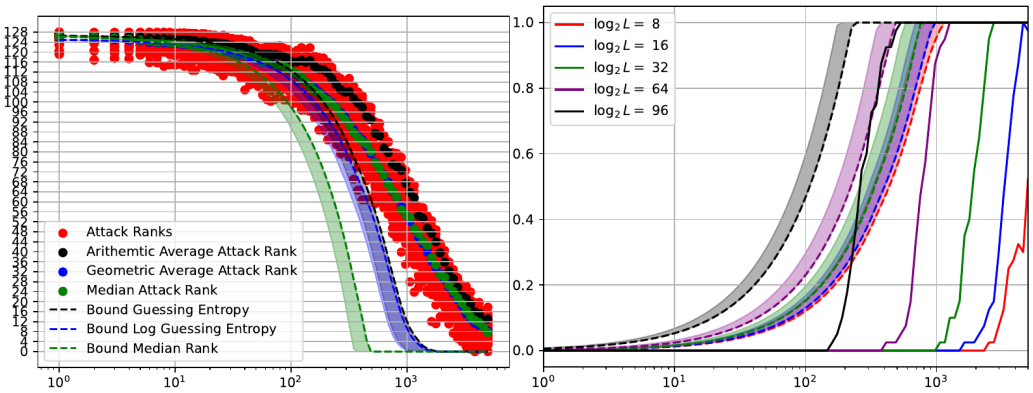
We confirm our analysis with real traces from the AES HD_ext dataset [BJP20]: `http://aisylabdatasets.ewi.tudelft.nl/`. We targeted the 16 bytes of the first round of the AES. We checked using the SNR that the leakage of the first round occurs in the 200 first points of the traces, hence we restricted our attack window to these points. We profiled a *Regression Based Linear Discriminant Analysis* (RLDA) for each byte using SCALib framework [CB23] using the first 300.000 traces. We divided the remaining 200.000 traces into 40 subsets of 5.000 traces on which we performed $N = 40$ independent attacks (one on each subset) using the profiled RLDAs and an increasing number of traces (ranging from 1 trace to 5.000 traces). We estimated the rank of the correct key for these attacks using the rank estimation technique based on histograms from SCALib [CB23] and plotted the

results. We estimated the training information using the *RLDA Information Estimator* implemented in SCALib to apply our bounds as reported in Table 1.

The results of the evaluations are given in Figure 7. The results are perfectly in line with the observations we made on synthetic data in Section 4.

Table 1: Estimated Informational Leakage on Each Byte.

Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
S-box ($bits \times 10^{-3}$)	7.53	7.77	5.98	5.48	4.97	7.37	9.34	8.85	9.39	4.41	8.26	8.83	10.80	22.43	6.81	11.49



(a) Empirical Ranks and Bounds (b) Success Rate of level L and Bounds

Figure 7: Results on AES-HD Dataset.

5.2 SMAesH Dataset

We also confirm our analysis with real traces from SMAesH dataset [CM24] available at https://smaesh-challenge.simple-crypto.org/datasets.html. The dataset is composed in three parts. A first part with variable keys is used for the profiling step and is composed of 2^{24} traces. The uncompressed archive represents 141 GB of traces. The two other parts are made for attack evaluations with fixed keys $fk0$ and $fk1$ each composed with 2^{24} traces. The uncompressed archives are both composed with 146 GB of traces. This is a realistic evaluation scenario with a target which is known to be hard to attack and the set of traces that are very long to acquire. Furthermore, the processing time for the attack evaluation is non-negligible and requires computational resources even with optimized libraries such as SCALib. We repeated $N = \min\{\frac{2^{24}}{q}, 1000\}$ independent attacks with q traces for both $fk0$ and $fk1$. This results into 1000 attacks with one trace but only two attacks with 2^{23} traces. We could not perform more attacks since they are only 2^{24} traces available for each fixed keys (which justifies our attack). For this reason, the empirical evaluation of the figures of merit on the right of the figure have a larger variance since they are estimated on few samples.

In this setting, the implementation is protected by state-of-the-art *Hardware Private Circuit* (HPC) masking scheme with two shares [CGLS21]. We consider the reference attack of the SMAesH challenge which is targeting the two shares of each byte at the output of the 16 S-boxes of the first round. This attack setting falls directly in the setting of our analysis. We can evaluate the informational leakage on each share as measured by mutual information and then use *Mrs. Gerber's Lemma* [BCG+23] to upper bound the informational leakage on the shared bytes. We can also directly evaluate the informational leakage on the unprotected key byte to avoid the small overhead due to Mrs. Gerber's Lemma. However, this second approach is harder to perform reliably as it requires to evaluate a much smaller informational leakage (in the order of the square of informational leakage on the shares because of the masking countermeasure). The estimated leakage

using both methodologies are reported in Table 2. In particular, both methodologies are consistent and indicates that the informational leakage on the unmasked S-boxes is in order of magnitude of 10^{-6} bits per trace. The results are given in Figure 8 using the evaluation with Mrs. Gerber's Lemma and shows the practicality of our approach for practical settings with state-of-the-art target that are hard to attack. We applied our bounds by keeping the i-th byte of the key that were leaking the least for $i = 0, \ldots, 15$ and kept the best bound. This slightly improves the bound because the different bytes of the key leak differently.

Table 2: Estimated informational leakage. Respectively on both shares, on the unmasked S-boxes using Mrs. Gerber's Lemma and the unmasked S-boxes with direct evaluation.

Byte Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Share 0 (bits $\times 10^{-3}$)	2.91	1.22	0.38	0.77	2.99	1.22	0.37	0.81	2.97	1.22	0.41	0.71	2.33	0.94	0.33	0.47
Share 1 (bits $\times 10^{-3}$)	1.24	3.72	0.54	0.77	0.97	3.71	0.51	0.75	1.03	3.73	0.53	0.70	0.71	2.81	0.33	0.36
S-box MGL (bits $\times 10^{-6}$)	4.98	6.27	0.29	0.81	4.00	6.28	0.26	0.84	4.24	6.31	0.30	0.69	2.30	3.67	0.15	0.23
S-box Direct (bits $\times 10^{-6}$)	4.68	4.86	1.29	1.41	4.37	5.06	1.00	1.20	4.93	4.85	1.07	1.10	3.48	3.65	0.71	0.57

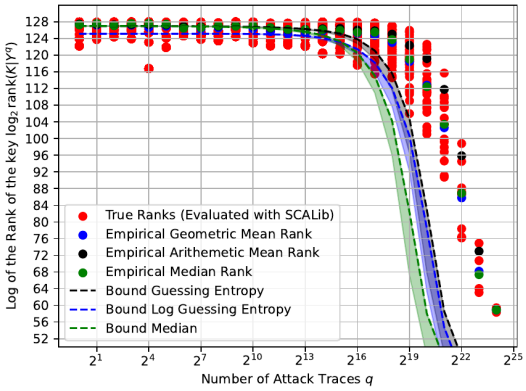


Figure 8: Results on SMAesH Dataset: Empirical Ranks and Bounds.

Results Analysis. We observe on both the *AES HD* and *SMAesH* datasets that the observations made on simulated data from Section 4 extend to real traces. The SMAesH dataset is a protected implementation, harder to attack, and we confirm that our methodology scales to harder targets. Also, it shows that the methodology combines well with other shortcut formula, such as *Mrs. Gerbers Lemma* used to tightly upper bounds leakages in the presence of masking. Compared to the simulation setting, we observe on Figure 8 and Figure 7 a larger evaluation gap for the log-guessing entropy and median rank using our methodology. While our method is more conservative, the dashed curves are bellow the dotted values, in terms of evaluation of the number of traces required to mount the attack this only implies a factor 4 gap. It is possible that due to the masking countermeasure our bound is less tight and using *Mrs. Gerbers Lemma* also adds some gap in the estimation of mutual information. In terms of sensitivity to inaccurate mutual information estimation we confirm on Figure 8 and Figure 7 that the effect is mild on the final results. It only modifies the evaluation of the number of traces to mount the attack by a moderate multiplicative factor. In our setting, we consider a divide-and-conquer setting on the 16-Sboxs after the first AddRoundKey. Hence, our bounds apply to this type of adversaries. The state-of-the-art attack on the SMAesH dataset, leverages more leakages in the diffusion layer of the AES to mount its attack and hence as expected outperforms our bound. Covering this type of attacks goes beyond the scope of this study.

6 Discussion, Limitations and Conclusion

6.1 Discussion and Limitations

Our methodology comes with both advantages and limitations highlighted in this section:

Advantages. One advantage of our methodology is that it bounds the performance of the information-theoretical optimal adversary. Hence, it is valid for any other adversary so that our bounds are agnostic to the attacker. In particular, it is valid for DL-SCA irrespectively of the architecture, loss function and training procedure used by the attacker. However, since we bound the performance of an optimal attack, our bound may appear as too conservative with respect to the performance of a simple non profiled correlation power analysis. Another advantage of our method is that it provides tendencies for the evaluator, we can evaluate the impact of an increased enumeration power, increased number of traces or variation of the noise level. Finally, our methodology does not need to measure traces with a fixed key to evaluate the performance of the attack. The dataset with variable keys required for the profiling step is sufficient to evaluate the mutual information and use the bounds. Furthermore, we use the value of the mutual information $I(X;Y)$ in our analysis but we do not make any assumption on the nature of the channel $X \to Y$. In particular, we never assume that the noise is Gaussian or that a specific function of X such as its Hamming weight is leaking.

Limitations. Our methodology also comes with drawbacks.

The main issue is that it may be challenging to obtain a reliable evaluation of the mutual information. For instance, it has been reported that the perceived information estimator is negatively biased with respect to the actual mutual information [IUH25, Section 4.2]. Hence, using our bounds with perceived information may provide a false sentiment of security if it underestimates the leakages as measured by mutual information. As an example we evaluated the mutual information on the SMAesH dataset by considering only the leakages from the SBoxes of the first round. However, if an adversary also exploits leakages from other computations from the circuit then our leakage evaluation can underestimate the leakages available to the attacker and hence its attack performances.

Also while our analysis in terms of mutual information is valid even in the presence of physical default such as glitches, this should be taken into account during the mutual information estimation which, to our knowledge, is still an open question.

Finally, the direct evaluation of the mutual information between a sensitive value and its corresponding leakages is too hard in the presence of high dimensional leakages. For this reason, the mutual information is rather evaluated between the sensitive value and the leakages processed with a dimensionality reduction. Again, if the mutual information estimator use the assumption that the noise is Gaussian (with a Gaussian template) and this assumption is not met in practice then the estimator could return an inaccurate estimation of the mutual information. Overall, the main limitation of our work is that it relies on an accurate estimation of mutual information which depends on the side channel modeling quality (leakages from all the circuit, physical defaults, assumption on the noise distribution, dimension of the leakages and number of samples).

Real-World Security Evaluations. In this paragraph, we emphasize how our bounds can be use for practical usage. The first step is to evaluate the side channel leakages via mutual information. For a protected implementation, we suggest evaluating the leakage on each share to obtain an upper on the mutual information of a protected sensitive value via Mrs. Gerber's Lemma. Then using the linear bound for mutual information we can obtain an upper bound on the side channel leakages using multiple traces. Then the evaluator needs to choose a security criterion such as a value log-guessing entropy or a value of probability of success of level L for a given value L. Finally, combining our inequality with the upper bound on mutual information the evaluator obtain a lower on the minimum number of

traces required to achieve the targeted security criterion. We obtain a security guarantee for attack using up to this number of traces. This can also be converted into a time needed to mount the attack assuming the measurement campaign length is proportional to the number of traces to acquire.

6.2 Conclusion

We provided informational bounds on the figures of merit of side-channel attacks. Most notably we provided bounds on the success rate of level L of a side-channel attack bridging the important gap of taking into account the computational power of the adversary in security bounds based on informational metrics. As a byproduct, we obtained bounds on the α-marginal guesswork and in particular a simple explicit bound for the median rank of the key. Furthermore, we obtained for the first time an informational bound on the geometric mean rank of the key (log-guessing entropy).

These formulas are of theoretical interest as they provide the dependency of the figure of merits with the number of traces (data complexity), the computational complexity and the level of noise in the traces.

Besides their theoretical interest the inequalities can be of practical interest for security evaluators. We confirm our findings on both synthetic data and real side-channel measurements. On the AES-HD dataset the full evaluation can be performed in a few minutes on a laptop. However, the evaluation on SMAesH dataset [CM24] had to performed on a server with 96 processors (*AMD EPYC 7F72 24-Core Processor*) and 500 GB of random access memory. While the evaluation of mutual information on the full variable key dataset can be performed fast (approximately 56 minutes) even on a non-threaded implementation, the construction of the security graph takes several hours even when the attacks are parallelized in twelve different processes each leveraging multi-threading capability from SCALib. This evaluation process is still manageable on the SMAesH dataset. However, the largest benefit of our method is that building the security graph requires taking measurements with fixed keys and a large number of traces which can be extremely time-consuming. For SMAesH, only two fixed keys datasets are available because of the huge memory usage (150 GB per dataset) and time that would be required to acquire more traces. In particular, the security graph is based on only two different fixed keys while it should be averaged with many keys. Furthermore, the attacks with 2^{24} traces can only be performed once for each key while they should be repeated. On the other hand, our methodology does not require obtaining measurements with fixed keys. Finally, we emphasize that if the masking order was 3 with the SMAesH implementation while maintaining a similar leakage per share then approximately 10^3 times more traces would be required to mount a successful attack according to our informational bounds. On the one hand, the evaluation of informational leakage would still be fast as its running time is proportional with the number of shares (because we rely on Mrs. Gerbers Lemma as a shortcut formula for masked encodings). On the other hand, building the security graph would be extremely hard as the attack time (repeated $100\times$ for a success rate estimation) would require 100×500 hours assuming that the running time of the attack is linear in the number of traces and since the attack with 2^{24} runs in 30 minutes with our implementation (i.e., 5.7 years if not performed in parallel). Not to mention the required 1000×21 hours (i.e., 2.4 years) required for collection of the dataset that would be of 150 TB size (the SMAesH datasets took 21 hours each to be measured [CM24]). These orders of magnitude would be again multiplied with masking at order 4 with an even more intractable setting.

In terms of *Common Criteria Methodology* quotation system our results directly translates into (JIL rating [Joi20, Table 12]) the score for "elapsed time". The score ranges from 0 for an attack that can be performed in less than one hour to 10 points for an attack requiring more than four months to be performed. For SMAesH [CM24] each trace is measured in 4.46 ms, and we provide lower bounds on the rank of the correct key in the

key ranking at the output of a maximum likelihood attack in terms of the number of traces used for the attack. Assuming that each key is enumerated in, say 1 ms, we obtain a score for the Common Criteria in terms of elapsed time as a sum of the two terms. Since we do not have to mount the attack our methodology offers a significant advantage for attacks that hard to mount.

Another practical usage of our lower bounds is that it can be used within fresh re-keying policy [AB00] to estimate when the ephemeral keys should be refreshed.

As an open question it would be interesting to obtain security guarantees when the adversary does not use its computational power to enumerate the key hypothesis but try to solve a cryptographically hard problem using the side-channel leakage as a trapdoor.

Acknowledgments François-Xavier Standaert is a research director of the Belgian Fund for Scientific Research (F.R.S.-FNRS). This work has been funded in part by the European Research Council (ERC) Advanced Grant BRIDGE (number 101096871). Views and opinions expressed are those of the authors and do not necessarily reflect those of the European Union or the ERC. Neither the European Union nor the granting authority can be held responsible for them.

This work received funding from the France 2030 program, managed by the French National Research Agency under grant agreement No. ANR-22-PETQ-0008 PQ-TLS.

Secure-IC acknowledges funding by the SOITEC-SIC project, from the IPCEI/PIIEC MICROÉLECTRONIQUE framework, under contract DOS0220776 - DOS0220774

Julien Béguinot's PhD is funded by the Institut Mines-Télécom (IMT) through the funding Futur & Ruptures.

References

[AB00] Michel Abdalla and Mihir Bellare. Increasing the lifetime of a key: A comparative analysis of the security of re-keying techniques. In Tatsuaki Okamoto, editor, *Advances in Cryptology - ASIACRYPT 2000, 6th International Conference on the Theory and Application of Cryptology and Information Security, Kyoto, Japan, December 3-7, 2000, Proceedings*, volume 1976 of *Lecture Notes in Computer Science*, pages 546–559. Springer, 2000.

[AGK76] Rudolf Ahlswede, Peter Gács, and János Körner. Bounds on conditional probabilities with applications in multi-user communication. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 34(2):157–177, 1976.

[Ari96] Erdal Arikan. An inequality on guessing and its application to sequential decoding. *IEEE Trans. Inf. Theory*, 42(1):99–105, 1996.

[BCG+23] Julien Béguinot, Wei Cheng, Sylvain Guilley, Yi Liu, Loïc Masure, Olivier Rioul, and François-Xavier Standaert. Removing the field size loss from duc et al.'s conjectured bound for masked encodings. In Elif Bilge Kavun and Michael Pehl, editors, *Constructive Side-Channel Analysis and Secure Design - 14th International Workshop, COSADE 2023, Munich, Germany, April 3-4, 2023, Proceedings*, volume 13979 of *Lecture Notes in Computer Science*, pages 86–104. Springer, 2023.

[BCGR24] Julien Béguinot, Wei Cheng, Sylvain Guilley, and Olivier Rioul. Formal security proofs via doeblin coefficients: - optimal side-channel factorization from noisy leakage to random probing. In Leonid Reyzin and Douglas Stebila, editors, *Advances in Cryptology - CRYPTO 2024 - 44th Annual*

International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2024, Proceedings, Part VI, volume 14925 of *Lecture Notes in Computer Science*, pages 389–426. Springer, 2024.

[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.

[BCPZ16] Alberto Battistello, Jean-Sébastien Coron, Emmanuel Prouff, and Rina Zeitoun. Horizontal side-channel attacks and countermeasures on the ISW masking scheme. In Benedikt Gierlichs and Axel Y. Poschmann, editors, *Cryptographic Hardware and Embedded Systems - CHES 2016 - 18th International Conference, Santa Barbara, CA, USA, August 17-19, 2016, Proceedings*, volume 9813 of *Lecture Notes in Computer Science*, pages 23–39. Springer, 2016.

[BHM+19] Olivier Bronchain, Julien M. Hendrickx, Clément Massart, Alex Olshevsky, and François-Xavier Standaert. Leakage certification revisited: Bounding model errors in side-channel security evaluations. In *CRYPTO (1)*, volume 11692 of *Lecture Notes in Computer Science*, pages 713–737. Springer, 2019.

[BJP20] Shivam Bhasin, Dirmanto Jap, and Stjepan Picek. AES HD dataset - 500 000 traces. AISyLab repository. 2020.

[BKM+15] Andrey Bogdanov, Ilya Kizhvatov, Kamran Manzoor, Elmar Tischhauser, and Marc Witteman. Fast and memory-efficient key recovery in side-channel attacks. In Orr Dunkelman and Liam Keliher, editors, *Selected Areas in Cryptography - SAC 2015 - 22nd International Conference, Sackville, NB, Canada, August 12-14, 2015, Revised Selected Papers*, volume 9566 of *Lecture Notes in Computer Science*, pages 310–327. Springer, 2015.

[BR24] Julien Béguinot and Olivier Rioul. What can Information Guess? Guessing Advantage vs. Rényi Entropy for Small Leakages. *arXiv preprint arXiv:2401.17057*, 2024.

[CB23] Gaëtan Cassiers and Olivier Bronchain. SCALib: A Side-Channel Analysis Library. *Journal of Open Source Software*, 8(86):5196, 2023.

[CDS22] Giovanni Camurati, Matteo Dell'Amico, and François-Xavier Standaert. Mcrank: Monte carlo key rank estimation for side-channel security evaluations. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2023:277–300, 2022.

[CDSU23] Gaëtan Cassiers, Henri Devillez, François-Xavier Standaert, and Balazs Udvarhelyi. Efficient regression-based linear discriminant analysis for side-channel security evaluations towards analytical attacks against 32-bit implementations. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2023(3):270–293, 2023.

[CGLS21] Gaëtan Cassiers, Benjamin Grégoire, Itamar Levi, and François-Xavier Standaert. Hardware private circuits: From trivial composition to full verification. *IEEE Trans. Computers*, 70(10):1677–1690, 2021.

[CM24] Gaëtan Cassiers and Charles Momin. The SMAesH dataset: Power leakage of a masked AES hardware implementation. 2024.

[CP17] Marios Omar Choudary and Pantelimon George Popescu. Back to Massey: Impressively fast, scalable and tight security evaluation tools. In *Proc. 19th Workshop on Cryptographic Hardware and Embedded Systems (CHES 2017)*, volume LNCS 10529, pages 367–386, 2017.

[CRBO24] Aakash Chowdhury, Arnab Roy, Carlo Brunetta, and Elisabeth Oswald. Leakage certification made simple. In *CRYPTO (6)*, volume 14925 of *Lecture Notes in Computer Science*, pages 427–460. Springer, 2024.

[CRR02] Suresh Chari, Josyula R. Rao, and Pankaj Rohatgi. Template attacks. In *CHES*, volume 2523 of *Lecture Notes in Computer Science*, pages 13–28. Springer, 2002.

[CT01] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2001.

[dCGRP19a] Éloi de Chérisey, Sylvain Guilley, Olivier Rioul, and Pablo Piantanida. Best information is most successful: Mutual information and success rate in side-channel analysis. *IACR Transactions on Cryptographic Hardware and Embedded Systems (CHES 2019)*, pages 49–79, February 2019.

[dCGRP19b] Éloi de Chérisey, Sylvain Guilley, Olivier Rioul, and Pablo Piantanida. An information-theoretic model for side-channel attacks in embedded hardware. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 310–315, 2019.

[DFS15] Alexandre Duc, Sebastian Faust, and François-Xavier Standaert. Making masking security proofs concrete (or how to evaluate the security of any leaking device), extended version. *Journal of Cryptology*, 32:1263 – 1297, 2015.

[DSV14] François Durvaux, François-Xavier Standaert, and Nicolas Veyrat-Charvillon. How to certify the leakage of a chip? In *EUROCRYPT*, volume 8441 of *Lecture Notes in Computer Science*, pages 459–476. Springer, 2014.

[DW15] Liron David and Avishai Wool. A bounded-space near-optimal key enumeration algorithm for multi-dimensional side-channel attacks. *IACR Cryptol. ePrint Arch.*, page 1236, 2015.

[DW17] Liron David and Avishai Wool. A bounded-space near-optimal key enumeration algorithm for multi-subkey side-channel attacks. In *CT-RSA*, volume 10159 of *Lecture Notes in Computer Science*, pages 311–327. Springer, 2017.

[DW19a] Liron David and Avishai Wool. Fast analytical rank estimation. In Ilia Polian and Marc Stöttinger, editors, *Constructive Side-Channel Analysis and Secure Design - 10th International Workshop, COSADE 2019, Darmstadt, Germany, April 3-5, 2019, Proceedings*, volume 11421 of *Lecture Notes in Computer Science*, pages 168–190. Springer, 2019.

[DW19b] Liron David and Avishai Wool. Poly-logarithmic side channel rank estimation via exponential sampling. In Mitsuru Matsui, editor, *Topics in Cryptology - CT-RSA 2019 - The Cryptographers' Track at the RSA Conference 2019, San Francisco, CA, USA, March 4-8, 2019, Proceedings*, volume 11405 of *Lecture Notes in Computer Science*, pages 330–349. Springer, 2019.

[DW21] Liron David and Avishai Wool. Rank estimation with bounded error via exponential sampling. *Journal of Cryptographic Engineering*, 12:151 – 168, 2021.

[Eli57] Peter Elias. List decoding for noisy channels. In *Technical Report 335, Research Laboratory of Electronics, MIT*, 1957.

[Fan52] Robert Mario Fano. Class notes for course 6.574: Transmission of information. *Lecture Notes*, 1952.

[GBTP08] Benedikt Gierlichs, Lejla Batina, Pim Tuyls, and Bart Preneel. Mutual information analysis. In Elisabeth Oswald and Pankaj Rohatgi, editors, *Cryptographic Hardware and Embedded Systems - CHES 2008, 10th International Workshop, Washington, D.C., USA, August 10-13, 2008. Proceedings*, volume 5154 of *Lecture Notes in Computer Science*, pages 426–442. Springer, 2008.

[GGP+15] Cezary Glowacz, Vincent Grosso, Romain Poussier, Joachim Schüth, and François-Xavier Standaert. Simpler and more efficient rank estimation for side-channel security assessment. In Gregor Leander, editor, *Fast Software Encryption - 22nd International Workshop, FSE 2015, Istanbul, Turkey, March 8-11, 2015, Revised Selected Papers*, volume 9054 of *Lecture Notes in Computer Science*, pages 117–129. Springer, 2015.

[GHR15] Sylvain Guilley, Annelie Heuser, and Olivier Rioul. A key to success - success exponents for side-channel distinguishers. In Alex Biryukov and Vipul Goyal, editors, *Progress in Cryptology - INDOCRYPT 2015 - 16th International Conference on Cryptology in India, Bangalore, India, December 6-9, 2015, Proceedings*, volume 9462 of *Lecture Notes in Computer Science*, pages 270–290. Springer, 2015.

[GMO01] Karine Gandolfi, Christophe Mourtel, and Francis Olivier. Electromagnetic analysis: Concrete results. In *CHES*, volume 2162 of *Lecture Notes in Computer Science*, pages 251–261. Springer, 2001.

[Gro18] Vincent Grosso. Scalable key rank estimation (and key enumeration) algorithm for large keys. In Begül Bilgin and Jean-Bernard Fischer, editors, *Smart Card Research and Advanced Applications, 17th International Conference, CARDIS 2018, Montpellier, France, November 12-14, 2018, Revised Selected Papers*, volume 11389 of *Lecture Notes in Computer Science*, pages 80–94. Springer, 2018.

[HRG14] Annelie Heuser, Olivier Rioul, and Sylvain Guilley. Good is not good enough - deriving optimal distinguishers from communication theory. In Lejla Batina and Matthew Robshaw, editors, *Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings*, volume 8731 of *Lecture Notes in Computer Science*, pages 55–74. Springer, 2014.

[IUH22] Akira Ito, Rei Ueno, and Naofumi Homma. Perceived Information Revisited New Metrics to Evaluate Success Rate of Side-Channel Attacks. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2022(4):228–254, 2022.

[IUH25] Akira Ito, Rei Ueno, and Naofumi Homma. Perceived information revisited II information-theoretical analysis of deep-learning based side-channel attacks. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2025(1):450–474, 2025.

[Joi20] Joint Interpretation Library. Application of Attack Potential to Smartcards and Similar Devices (Version 3.1). Technical report, Senior Officials Group Information Systems Security,

June 2020. https://www.sogis.eu/documents/cc/domains/sc/JIL-Application-of-Attack-Potential-to-Smartcards-v3-1.pdf.

[KJJ99] Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael J. Wiener, editor, *Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 1999, Proceedings*, volume 1666 of *Lecture Notes in Computer Science*, pages 388–397. Springer, 1999.

[KPP+06] Sandeep S. Kumar, Christof Paar, Jan Pelzl, Gerd Pfeiffer, and Manfred Schimmler. Breaking ciphers with COPACOBANA - A cost-optimized parallel code breaker. In Louis Goubin and Mitsuru Matsui, editors, *Cryptographic Hardware and Embedded Systems - CHES 2006, 8th International Workshop, Yokohama, Japan, October 10-13, 2006, Proceedings*, volume 4249 of *Lecture Notes in Computer Science*, pages 101–118. Springer, 2006.

[LMM+16] Jake Longo, Daniel P. Martin, Luke Mather, Elisabeth Oswald, Benjamin Sach, and Martijn Stam. How low can you go? using side-channel data to enhance brute-force key recovery. Cryptology ePrint Archive, Paper 2016/609, 2016. https://eprint.iacr.org/2016/609.

[LPR+14] Victor Lomné, Emmanuel Prouff, Matthieu Rivain, Thomas Roche, and Adrian Thillard. How to estimate the success rate of higher-order side-channel attacks. In Lejla Batina and Matthew Robshaw, editors, *Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings*, volume 8731 of *Lecture Notes in Computer Science*, pages 35–54. Springer, 2014.

[Man04] Stefan Mangard. Hardware countermeasures against DPA ? A statistical analysis of their effectiveness. In Tatsuaki Okamoto, editor, *Topics in Cryptology - CT-RSA 2004, The Cryptographers' Track at the RSA Conference 2004, San Francisco, CA, USA, February 23-27, 2004, Proceedings*, volume 2964 of *Lecture Notes in Computer Science*, pages 222–235. Springer, 2004.

[Mas94] James L Massey. Guessing and entropy. In *Proceedings of 1994 IEEE International Symposium on Information Theory*, page 204, 1994.

[MCHS23] Loïc Masure, Gaëtan Cassiers, Julien M. Hendrickx, and François-Xavier Standaert. Information bounds and convergence rates for side-channel security evaluators. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2023(3):522–569, 2023.

[MMOS16] Daniel P. Martin, Luke Mather, Elisabeth Oswald, and Martijn Stam. Characterisation and estimation of the key rank distribution in the context of side channel evaluations. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part I*, volume 10031 of *Lecture Notes in Computer Science*, pages 548–572, 2016.

[MOOS15] Daniel P. Martin, Jonathan F. O'Connell, Elisabeth Oswald, and Martijn Stam. Counting keys in parallel after a side channel attack. In *ASIACRYPT (2)*, volume 9453 of *Lecture Notes in Computer Science*, pages 313–337. Springer, 2015.

[MRS22] Loïc Masure, Olivier Rioul, and François-Xavier Standaert. A Nearly Tight Proof of Duc et al.'s Conjectured Security Bound for Masked Implementations. In *IACR Cryptology ePrint Archive*, 2022.

[MS23] Loïc Masure and François-Xavier Standaert. Prouff and Rivain's Formal Security Proof of Masking, Revisited - Tight Bounds in the Noisy Leakage Model. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology - CRYPTO 2023 - 43rd Annual International Cryptology Conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings, Part III*, volume 14083 of *Lecture Notes in Computer Science*, pages 343–376. Springer, 2023.

[MY95] Robert J. McEliece and Zhong Yu. An inequality on entropy. In *Proceedings of 1995 IEEE International Symposium on Information Theory*, page 329, 1995.

[Nal15] Maurizio Naldi. Approximation of the truncated Zeta distribution and Zipf's law. *CoRR*, abs/1511.01480, 2015.

[PGS15] Romain Poussier, Vincent Grosso, and François-Xavier Standaert. Comparing approaches to rank estimation for side-channel security evaluations. In Naofumi Homma and Marcel Medwed, editors, *Smart Card Research and Advanced Applications - 14th International Conference, CARDIS 2015, Bochum, Germany, November 4-6, 2015. Revised Selected Papers*, volume 9514 of *Lecture Notes in Computer Science*, pages 125–142. Springer, 2015.

[Pli00a] John O. Pliam. Guesswork and variation distance as measures of cipher security. In Howard Heys and Carlisle Adams, editors, *Selected Areas in Cryptography*, pages 62–77, 2000.

[Pli00b] John O. Pliam. On the incomparability of entropy and marginal guesswork in brute-force attacks. In Bimal K. Roy and Eiji Okamoto, editors, *Progress in Cryptology - INDOCRYPT 2000, First International Conference in Cryptology in India, Calcutta, India, December 10-13, 2000, Proceedings*, volume 1977 of *Lecture Notes in Computer Science*, pages 67–79. Springer, 2000.

[PPS12] Kenneth G. Paterson, Antigoni Polychroniadou, and Dale L. Sibborn. A coding-theoretic approach to recovering noisy RSA keys. 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 386–403. Springer, 2012.

[PSG16] Romain Poussier, François-Xavier Standaert, and Vincent Grosso. Simple key enumeration (and rank estimation) using histograms: An integrated approach. In Benedikt Gierlichs and Axel Y. Poschmann, editors, *Cryptographic Hardware and Embedded Systems - CHES 2016 - 18th International Conference, Santa Barbara, CA, USA, August 17-19, 2016, Proceedings*, volume 9813 of *Lecture Notes in Computer Science*, pages 61–81. Springer, 2016.

[PW23] Yury Polyanskiy and Yihong Wu. *Information Theory, From Coding to Learning (1. ed.)*. Cambridge University Press, 2023.

[Riv08] Matthieu Rivain. On the Exact Success Rate of Side Channel Analysis in the Gaussian Model. In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *Selected Areas in Cryptography, 15th International Workshop, SAC 2008, Sackville, New Brunswick, Canada, August 14-15, Revised Selected Papers*, volume 5381 of *Lecture Notes in Computer Science*, pages 165–183. Springer, 2008.

[RPC22] Anca Radulescu, Pantelimon George Popescu, and Marios O. Choudary. GE vs GM: efficient side-channel security evaluations on full cryptographic keys. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2022(4):886–905, 2022.

[Sak20] Yuta Sakai. Generalizations of Fano's Inequality for Conditional Information Measures via Majorization Theory. *Entropy*, 22(3):288, 2020.

[SMY09] François-Xavier Standaert, Tal G. Malkin, and Moti Yung. A unified framework for the analysis of side-channel key recovery attacks. In Antoine Joux, editor, *Advances in Cryptology - EUROCRYPT 2009*, pages 443–461, 2009.

[TCRP21] Andrei Tănăsescu, Marios O. Choudary, Olivier Rioul, and Pantelimon George Popescu. Tight and scalable side-channel attack evaluations through asymptotically optimal Massey-like inequalities on guessing entropy. *Entropy*, 23(11):1–10, 2021.

[VGRS12] Nicolas Veyrat-Charvillon, Benoît Gérard, Mathieu Renauld, and François-Xavier Standaert. An optimal key enumeration algorithm and its application to side-channel attacks. In Lars R. Knudsen and Huapeng Wu, editors, *Selected Areas in Cryptography, 19th International Conference, SAC 2012, Windsor, ON, Canada, August 15-16, 2012, Revised Selected Papers*, volume 7707 of *Lecture Notes in Computer Science*, pages 390–406. Springer, 2012.

[VGS13] Nicolas Veyrat-Charvillon, Benoît Gérard, and François-Xavier Standaert. Security evaluations beyond computing power. 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 126–141. Springer, 2013.

[YMO22] Rebecca Young, Luke Mather, and Elisabeth Oswald. Comparing key rank estimation methods. In Ileana Buhan and Tobias Schneider, editors, *Smart Card Research and Advanced Applications - 21st International Conference, CARDIS 2022, Birmingham, UK, November 7-9, 2022, Revised Selected Papers*, volume 13820 of *Lecture Notes in Computer Science*, pages 188–204. Springer, 2022.

[ZDF20] Ziyue Zhang, A. Adam Ding, and Yunsi Fei. A fast and accurate guessing entropy estimation algorithm for full-key recovery. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2020(2):26–48, 2020.