Connecting Leakage-Resilient Secret Sharing to Practice: Scaling Trends and Physical Dependencies of Prime Field Masking

Sebastian Faust, Loïc Masure, Elena Micheli, Maximilian Orlt, François-Xavier Standaert

▶ To cite this version:

Sebastian Faust, Loïc Masure, Elena Micheli, Maximilian Orlt, François-Xavier Standaert. Connecting Leakage-Resilient Secret Sharing to Practice: Scaling Trends and Physical Dependencies of Prime Field Masking. EUROCRYPT 2024 - 43rd International Conference on the Theory and Applications of Cryptographic Techniques, May 2024, Zürich, Switzerland. lirmm-04484194

HAL Id: lirmm-04484194
https://hal-lirmm.ccsd.cnrs.fr/lirmm-04484194v1

Submitted on 29 Feb 2024

Connecting Leakage-Resilient Secret Sharing to Practice: Scaling Trends and Physical Dependencies of Prime Field Masking

Sebastian Faust[1], Loïc Masure[2], Elena Micheli[1],
Maximilian Orlt[1], François-Xavier Standaert[3]

[1] Department of Computer Science, TU Darmstadt, Darmstadt, Germany
[2] LIRMM, Univ. Montpellier, CNRS, France
[3] Crypto Group, ICTEAM Institute, UCLouvain, Louvain-la-Neuve, Belgium

Abstract. Symmetric ciphers operating in (small or mid-size) prime fields have been shown to be promising candidates to maintain security against low-noise (or even noise-free) side-channel leakage. In order to design prime ciphers that best trade physical security and implementation efficiency, it is essential to understand how side-channel security evolves with the field size (i.e., scaling trends). Unfortunately, it has also been shown that such scaling trends depend on the leakage functions and cannot be explained by the standard metrics used to analyze Boolean masking with noise. In this work, we therefore initiate a formal study of prime field masking for two canonical leakage functions: bit leakages and Hamming weight leakages. By leveraging theoretical results from the leakage-resilient secret sharing literature, we explain formally why (1) bit leakages correspond to a worst-case and do not encourage operating in larger fields, and (2) an opposite conclusion holds for Hamming weight leakages, where increasing the prime field modulus p can contribute to a security amplification that is exponential in the number of shares, with $\log(p)$ seen as security parameter like the noise variance in Boolean masking. We combine these theoretical results with simulated experiments and show that the interest masking in larger prime fields can degrade gracefully when leakage functions slightly deviate from the Hamming weight abstraction, motivating further research towards characterizing (ideally wide) classes of leakage functions offering such guarantees.

1 Introduction

Security against differential side-channel analysis [26], where an adversary continuously accumulates leakage about a long-term secret, is needed for any symmetric authentication or encryption scheme with embedded security guarantees [5]. Masking is the main countermeasure to mitigate such attacks [16,23].[1] It can be viewed as multi-party computation on silicon, where the (e.g., symmetric) cryptographic algorithm is executed on d shares. Informally, masking forces

[1] The only known alternative is to rely on fresh re-keying with a leakage-resilient PRF [4], which is only exploitable in tailored designs such as ISAP [17].

the adversary to combine the leakage of different shares in order to gain information on the long-term secret. A broad sequence of works has investigated the theoretical security guarantees that such schemes provide and connected them to practical (e.g., noisy leakage) models [25,40,18,19,39,34]: they show that under some noise and independence assumptions, masking can lead to exponential security amplification. The independence assumption has been the focus of significant research efforts over the last decade and established design and proof techniques enable to ensure it to a sufficient extent [38,21,36,14]. The noise condition has been more investigated from the adversarial viewpoint and several works showed that the requirement is strict for binary ciphers [2,24,35,11], hence calling for masked operations that manipulate each share minimally [1,13].

Concretely, the weakness of masking in low-noise settings is due to the strong algebraic compatibility between leakage functions observed in practice, such as the Hamming Weight (HW) leakage function, and operations performed in binary fields. For example, say an adversary is able to observe the noise-free HW of Boolean shares (processed in serial or in parallel). Then, just observing the parity of the leakages provides easily exploitable information about the secret, regardless of the number of shares [43]. By adding noise to the leakages, designers essentially ensure that this algebraic compatibility is sufficiently hidden so that the only remaining attack path is statistical (i.e., requires to estimate a high-order moment of the leakage distribution [41,37]). A bit more formally, the reason of this weakness is that the finite group over which masking operates has non-trivial subgroups in the binary case. If the support of the Probability Mass Function (PMF) of each share given the leakage is contained in a coset of a non-trivial subgroup, then the PMF of the corresponding secret is also contained in a coset of the same subgroup [44]. As a consequence, the support of the secret PMF cannot be full, which results in an amount of informative leakage about the secret that cannot be arbitrarily low, regardless of the number of shares.

In order to circumvent this issue, Dziembowski et al. showed that the finite group in which the masking operates should not have any non-trivial subgroup, which characterizes prime fields [20]. This seed result has recently triggered an interest for prime-field masking in symmetric cryptography. Early works in this direction show that ciphers that natively operate in prime fields have a good potential to leverage the excellent properties of prime-field masking. They could in turn enable better implementation security vs. efficiency tradeoffs than binary ciphers, especially in low-noise settings, and with only mild overheads when side-channel attacks are not a concern [33,12]. Yet, these results also show that taking full advantage of this potential requires understanding the scaling trends of prime-field masking. In particular, one central open question of which the answer could guide the design of new prime ciphers is whether increasing the prime modulus is beneficial to side-channel security (and by how much)?

Both the empirical evaluations in [33] and the theoretical results in [20] suggest that answering this question is non-trivial. On the empirical side, Masure et al. showed that the interest of increasing the prime modulus depends on the leakage function. For example, a larger modulus improves the security amplification of masking for (noise-free) HW leakages while it does not for (noise-free) LSB leakages. Such a dependency on the (deterministic part of) the leakage function implies that the standard tools and metrics used to characterize Boolean masking are unlikely to explain the scaling trends of prime-field masking. This is because Boolean masking is only effective if leakages are sufficiently noisy, so that the deterministic part of the leakage function is essentially lifted in this case. As a result, the security amplification of Boolean masking only depends on the informativeness of the shares' leakages, classically measured with the Mutual Information (MI) or Statistical Distance (SD). But the HW of a value is more informative than a single-bit leakage according to these metrics, which does not back up the observations of [33]. As for the theoretical side, the noise amplification bounds provided in [20] are not tight for our purpose and do not suggest that increasing the size of the field in which masking operates is beneficial.

Based on this state of the art, the main goal of this paper is to provide theoretical explanations for previous empirical observations on prime-field masking, in order to establish foundations on which prime ciphers could be designed. Interestingly, it turns out that the case of bit (e.g., LSB) leakages has been the topic of (for now mostly theoretical) investigations in the context of leakage-resilient secret sharing, and extended towards any deterministic leakage model with a bounded range [7,27,31,8,28,30,29]. Among others, these works show that bit leakages are in some sense the most powerful leakage functions with bounded range, which therefore raises the question whether more positive results could be obtained for other, ideally more realistic, leakage functions.

In order to provide a complete analysis, we study the leakage resilience with respect to average-case and worst-case metrics. The average-case metric considers leakage from a masked *random* secret and is prominently used by the physical security community (since choosing plaintexts otherwise than uniformly at random has been shown to bring limited gains in this context [45]). On the other hand, worst-case security considers the resistance of masking for a worst-case choice of the secret. The latter is a standard notion in the cryptographic theory community, and, e.g., is used by the aforementioned results on leakage-resilient secret sharing. Interestingly, we show that for the LSB leakage function the analysis can be tightened when considering the weaker, yet realistic, average-case metric. To sum up, we achieve the following theoretical results:

- *For bit leakages*, we show that for both the average-case and worst-case metric, increasing the prime modulus in prime-field masking cannot lead to increased security. This confirms formally the experimental observations from [33]. On the positive side, our analysis for the average-case improves existing bounds from the worst-case setting by a constant factor.
- *For Hamming weight leakages* (which, to the best of our knowledge, were not formally studied so far), we show that for both the average-case and worst-case metric, increasing the prime modulus p can contribute to a security amplification that is exponential in the number of shares, with $\log(p)$ serving as a security parameter like the noise variance in Boolean masking.

In contrast to prior work on leakage-resilient secret sharing, our analysis takes into account that the range of the HW leakage function scales with the underlying field size. Concretely, we show that the distinguishing advantage approaches 0, while naively applying bounds from the literature on leakage-resilient secret sharing only gives a trivial upper bound of 1.

In Table 1, we provide a summary of our contribution within the existing literature, with references to the corresponding theorems in this paper. We remark that, for Hamming weight leakage, this work only considers upper bounds. The rationale behind this decision is our primary focus in arguing about how the security improves when increasing the field size. Nonetheless, we recognize the interest in quantifying the minimal information derived from a Hamming-weight leakage attack, which we leave as an open research question.

	Bit Leakage	Hamming Weight
Worst-case	*upper bound:* Theorem 1 [7] *lower bound:* Theorem 4 [29]	*upper bound:* Theorem 6 *lower bound:* open problem
Average-case	*upper bound:* Theorem 1 [7] *lower bound:* Theorem 5	*upper bound:* Theorem 8 *lower bound:* open problem

Table 1: Overview of the results for the security of additive secret sharing against bit and HW leakage, for both the worst-case and the average-case metric.

Hence, our results indicate that for the Hamming weight leakage function, there is theoretical support to design ciphers that operate in larger prime fields (which should then be weighted with the performance overheads this increase leads to). Quite naturally, they also raise the question whether concrete leakage functions that are close to but not exactly equal to the Hamming weight function maintain this interest. In order to stimulate research in this direction, we combine our theoretical analyzes with a simulated experiment, where we evaluate linear leakage functions that gradually deviate from the Hamming weight function. While extreme deviations lead to bit-like leakages where increasing p does not help (e.g., if a single bit leaks with such a high contribution to the overall leakage that it can be isolated), we show that this loss is gradual and that a broad class of leakage functions maintains the interest of Hamming weight leakages.

Finally, and despite we primarily use techniques in the context of leakage-resilient secret sharing to improve the understanding of masking in prime fields, our results come with observations that could be relevant for (more theoretical) research on leakage-resilient secret sharing as well. For example, the average-case security notion allows us to obtain tighter bounds for LSB leakages, and it would be interesting to study whether similar gains can be obtained for other practically-relevant leakage functions. In this respect, we note that for many applications of leakage-resilient secret sharing (e.g., in threshold cryptography for sharing a random secret key), the average-case notion suffices. In addition, our work highlights the importance of achieving good bounds when the range of a leakage function increases with the underlying field. We believe this is a quite natural generalization that could also be relevant in the domain of leakage-resilient secret sharing and opens up avenues for future research.

2 Background

Notations. In this paper, calligraphic letters like $\mathcal{S}$ denote sets, small letters like x denote elements of a given set, and capital letters like X denote random variables over a given set. The notation $\mathrm{X} \leftarrow \mathcal{S}$ means that X is uniformly drawn from $\mathcal{S}$. If $f : \mathcal{A} \to \mathcal{B}$ denotes a function mapping a set $\mathcal{A}$ to an image set $\mathcal{B}$, and $y \in \mathcal{B}$, $f^{-1}(y)$ denotes the pre-image set of all values $x \in \mathcal{A}$ such that $f(x) = y$. For a set $\mathcal{S} \subseteq \mathcal{A}$, we denote by $\mathbf{1}_{\mathcal{S}} : \mathcal{A} \to \{0, 1\}$ the function mapping any $x \in \mathcal{A}$ to 1 if and only if (i.f.f.) $x \in \mathcal{S}$ and to 0 otherwise. In particular, when considering characteristic functions over pre-image sets, we may use the shortcut notation $\mathbf{1}_y$ to denote $\mathbf{1}_{f^{-1}(y)}$ as long as there is no ambiguity on f.

Masking. For a finite field $\mathcal{Y} = \mathbb{F}_p$ of prime size p, let $\mathrm{Y} \in \mathcal{Y}$ be a *sensitive* value — *i.e.*, depending on a chunk of secret. To protect Y against a too much informative leakage, let $\mathrm{Y}_1, \ldots, \mathrm{Y}_d$ be d random variables uniformly drawn from $\mathcal{Y}$, such that $\mathrm{Y} = \mathrm{Y}_1 + \ldots + \mathrm{Y}_d$. This encoding is commonly referred to as masking, and the corresponding random variables are called shares. The adversary is then given access to a random vector $\mathbf{L} = (\mathrm{L}_1, \ldots, \mathrm{L}_d)$ such that each random variable L_i, also known as *leakage*, solely depends on the realization of Y_i. In the remaining of this paper, we make the additional assumption that we are in a *low-noise* setting, *i.e.* that each leakage L_i is a deterministic (non-injective) function randomized by its input Y_i.[2]

2.1 Quantifying the Distance to Uniform

To quantify the distance to the uniform distribution over $\mathcal{Y}$, we will use different metrics. Let p, m be two Probability Mass Functions (PMFs) over the finite set $\mathcal{Y}$. We denote by $\mathsf{TV}(\mathsf{p}; \mathsf{m})$ the Total Variation (TV) between p and m:

$$\mathsf{TV}(\mathsf{p}; \mathsf{m}) = \frac{1}{2} \sum_{y \in \mathcal{Y}} |\mathsf{p}(y) - \mathsf{m}(y)| \quad . \tag{1}$$

We will sometimes denote with $\mathsf{TV}\left(\mathrm{X}^{(0)}; \mathrm{X}^{(1)}\right)$ the total variation for the PMFs corresponding to the random variables $\mathrm{X}^{(0)}, \mathrm{X}^{(1)}$. Adapting the terminology from

[2] The literature of noise amplification bounds does not directly assume a noise-free setting, as it can be encompassed in the noisy leakage framework. On the opposite, to the best of our knowledge it is more common in the literature about the leakage-resilience of linear secret sharing schemes to rely on this assumption [8,29].

Prouff & Rivain [40], we define the *statistical bias* as Duc et al. [18]:

$$\beta(\mathrm{Y}|\mathbf{L}) = \mathop{\mathbb{E}}_{\mathbf{L}}\left[\mathsf{TV}\Big(\mathsf{p}_{\mathrm{Y}\,\mid\,\mathbf{L}}; \mathsf{p}_{\mathrm{Y}}\Big)\right] = \mathsf{TV}\big(\mathsf{p}_{\mathrm{Y},\mathbf{L}}; \mathsf{p}_{\mathrm{Y}} \otimes \mathsf{p}_{\mathbf{L}}\big) \enspace , \tag{2}$$

where $\otimes$ denotes the Cartesian product between two marginal probability distributions.[3] Notice that β is symmetric in its arguments, *i.e.*, $\beta(\mathrm{Y}|\mathbf{L}) = \beta(\mathbf{L}|\mathrm{Y}) = \mathop{\mathbb{E}}_{\mathrm{Y}}\left[\mathsf{TV}\Big(\mathsf{p}_{\mathbf{L}\,\mid\,\mathrm{Y}}; \mathsf{p}_{\mathbf{L}}\Big)\right]$. In the particular case of masking — which will be the main focus in this paper — the latter one can be rephrased as

$$\mathop{\mathbb{E}}_{y \leftarrow \mathcal{Y}}\left[\mathsf{TV}(\mathbf{L}\left(\mathsf{AddEnc}\left(y\right)\right); \mathbf{L}\left(\mathsf{AddEnc}\left(\mathrm{Y}\right)\right))\right] \enspace ,$$

where $\mathsf{AddEnc}\left(\cdot\right)$ is a random function that maps a secret value with one of its additive encodings. The bias is actually the *average* Total Variation (TV) between two PMFs. This definition is usual in papers dealing with masking against side-channel analysis. Note that this metric depends not only on the chosen leakage model, but also on the underlying distribution of the secret. That is why some related works [7] also considered a variant of this metric that we call *worst-case bias*:

$$\mathsf{M}_{\infty}\left(\mathbf{L}\right) = \max_{y^{(0)}, y^{(1)} \in \mathcal{Y}} \mathsf{TV}\Big(\mathbf{L}\left(\mathsf{AddEnc}\left(y^{(0)}\right)\right); \mathbf{L}\left(\mathsf{AddEnc}\left(y^{(1)}\right)\right)\Big) \enspace .$$

Here, the expectation is replaced by a maximum over the product set of secrets $\mathcal{Y}$. As a result, it follows that for all random variable Y:

$$\beta(\mathrm{Y}|\mathbf{L}) \leq \mathsf{M}_{\infty}\left(\mathbf{L}\right) \leq p \cdot \beta(\mathrm{Y}|\mathbf{L}) \enspace . \tag{3}$$

2.2 The Limits of Generic Noise Amplification Bounds

So far, the literature of masking security proofs provides bounds on the statistical bias of the following shape:

$$\beta(\mathrm{Y}|\mathbf{L}) \leq f(\delta_1, \ldots, \delta_d), \quad \text{if } \delta_i < t \text{ for all } i \enspace ,$$

where δ_i stands for the statistical bias between one share Y_i and its corresponding leakage L_i, t is a threshold for which the bound is valid, and f is a decreasing function with its arguments and converging exponentially fast with d towards 0 [20,3]. These so-called *noise amplification* bounds in the literature have the main strength of being *tight*, *i.e.*, there exists a leakage model such that the inequality becomes an equality. They also have the advantage of being *universal*, which means that they do not depend on the nature of the underlying leakage

[3] The first equality is used in noise amplification papers [40,20,34] and aims at quantifying how the distribution of the secret deviates from the prior knowledge whenever some side information $\mathbf{L}$ is available. The second equality is used in simulation-based proofs [18], and rather aims at quantifying how the side information $\mathbf{L}$ changes whenever the secret is known.

model, provided that the latter one verifies $\beta(\mathrm{Y}_i|\mathrm{L}_i) \leq \delta_i$ for all $i \in [\![0, d]\!]$. This would strongly suggest the intuitive idea that when comparing two leakage models applied on each share, the higher the δ_i for all i, the higher the bias $\beta(\mathrm{Y}|\mathbf{L})$. In other words, the leakier each share, the leakier the resulting secret.

To discuss this intuition, let us take as an example two well-known leakage models. The first one is the Hamming Weight (HW), the leakage function returning the sum of bits of the underlying leaky variable, *i.e.*, $\mathsf{HW}(y) = \sum_{i=0}^{n} y_i$.

Lemma 1 (Bias of Hamming weight). *For $n \in \mathbb{N}^\star$ and $p = 2^n - 1$, let $Y \leftarrow \mathbb{Z}_p$ and let $L = \mathsf{HW}(Y)$. Then the statistical bias between Y and L verifies*

$$\beta(Y|L) = 1 - \frac{\binom{2n}{n} - 1}{(2^n - 1)^2} \approx 1 - \frac{1}{\sqrt{\pi n}} \ . \tag{4}$$

We also consider a second leakage model where a proportion α of the n bits leaks. In other words, the leakage function ℓ^S returns the bits $y_{s_1}, \ldots, y_{s_{\alpha n}}$ where $S = \{s_1, \ldots, s_{\alpha n}\}$ is a set of αn indices.

Lemma 2 (Proportion of leaky bits). *For $n \in \mathbb{N}^\star$ and $p = 2^n - 1$, let $Y \leftarrow \mathbb{Z}_p$ and let $L = (Y_{s_1}, \ldots, Y_{s_{\alpha n}})$. Then the statistical bias between Y and L verifies*

$$\beta(Y|L) = 1 - \frac{2^{n(1-\alpha)}}{2^n - 1} + \frac{2^{n(1-\alpha)} - 1}{(2^n - 1)^2} \ . \tag{5}$$

In particular, for the LSB leakage model, $\alpha = \frac{1}{n}$, we have $\beta(Y|L) \approx \frac{1}{2} - 2^{-(n+1)}$.

Lemma 1 and Lemma 2, proven in Appendix B, tell us that a share leaking in Hamming Weight (HW) is leakier than the same share leaking in Least Significant Bit (LSB).[4] Even more, for the HW leakage model, the amount of leakage increases with the field size, whereas it remains nearly constant in the LSB leakage model. We would therefore expect a target device protected with masking and leaking the LSB of each share to be more secure than the same device leaking in HW. Furthermore, we would expect that in the latter case, increasing the field size could be harmful. But the observations of Masure *et al.* [33, Fig. 3] — measured in terms of MI — contradict both intuitions: not only the HW behaves not worse than the LSB leakage with masking in $\mathbb{F}_p$, but increasing the field size seems helpful to get a better leakage-resilience.

2.3 Refined Bounds through Fourier Analysis

A recent line of works have studied the so-called *local leakage resilience* of secret-sharing schemes designed over prime fields [7]. As a perhaps unexpected application of their framework (initially devoted to the security of MPC protocols),

[4] This is true regardless of the choice of the metric, *i.e.*, similar trends hold when considering the Mutual Information (MI).

Benhamouda *et al.*'s framework can be used to refine the noise amplification bounds prone to the limitations emphasized at the previous subsection.

Their framework relies on the Fourier analysis of leakage functions of each share. That is why we first recall a few facts about Fourier analysis.

Definition 1 (Discrete Fourier Transform). *Let $f : \mathbb{Z}_p \to \mathbb{C}$ be a function over the cyclic group $\mathbb{Z}_p$. Then the Discrete Fourier Transform (DFT) of f of harmonic α — or the α-th Fourier coefficient of f for short — is defined as*

$$\widehat{f}(\alpha) = \frac{1}{p} \sum_{y \in \mathbb{Z}_p} f(y) e^{-\frac{2\pi\alpha y}{p} i} \ . \tag{6}$$

Next, we list some interesting properties of the discrete Fourier transform.

Proposition 1 (Parseval). *Let $f, g : \mathbb{Z}_p \to \mathbb{C}$. Then*

$$\frac{1}{p} \sum_{y \in \mathbb{Z}_p} f(y) \cdot g(y) = \sum_{\alpha \in \mathbb{Z}_p} \widehat{f}(\alpha) \cdot \overline{\widehat{g}(\alpha)} \ .$$

In particular, this implies

$$\frac{1}{p} \sum_{y \in \mathbb{Z}_p} |f(y)|^2 = \sum_{\alpha \in \mathbb{Z}_p} \left| \widehat{f}(\alpha) \right|^2 \ . \tag{7}$$

Proposition 1 tells us that the DFT is an isometry for the Euclidean norm, up to a normalizing field-size factor. We next need another well-known formula from Fourier analysis, namely the Poisson summation formula.[5]

Proposition 2 (Poisson Summation Formula). *Let $\mathbb{F}$ be a finite field, and let $C \subseteq \mathbb{F}^d$ be a linear code with dual code $C^\perp$. Let $f_1, \ldots, f_d : \mathbb{F} \to C$ be functions. Then the following inequality holds:*

$$\mathop{\mathbb{E}}_{\boldsymbol{x} \leftarrow C} \left[\prod_{j=1}^{d} f_j(x_j) \right] = \sum_{\boldsymbol{\alpha} \in C^\perp} \prod_{j=1}^{d} \widehat{f_j}(\alpha_j) \ .$$

As observed in [7], the Poisson Summation Formula can be leveraged to link total variation and Fourier coefficients.

Proposition 3. *Let $\mathbb{F}_p$ be a prime field of size p. Then, for all $y^{(0)}, y^{(1)} \in \mathbb{F}_p$:*

$$\begin{aligned} &\mathsf{TV}\Big(\mathbf{L}\left(\mathsf{AddEnc}\left(y^{(0)}\right)\right); \mathbf{L}\left(\mathsf{AddEnc}\left(y^{(1)}\right)\right)\Big) \\ &\quad = \frac{1}{2} \sum_{\boldsymbol{\ell} \in \mathcal{L}^d} \left| \sum_{\alpha \in \mathbb{F}^\star} \left(\prod_{j=1}^{d} \widehat{\mathbf{1}_{L_j^{-1}(\ell_j)}}(\alpha) \right) \left(e^{-\frac{2i\pi\alpha y^{(0)}}{p}} - e^{-\frac{2i\pi\alpha y^{(1)}}{p}} \right) \right| \ . \end{aligned} \tag{8}$$

[5] For a more general statement of Proposition 2 concerning any linear code, see [7].

Furthermore, if Y follows the uniform distribution over $\mathbb{F}_p$, then for every $y \in \mathbb{F}_p$, $\mathsf{TV}(\mathbf{L}\left(\mathsf{AddEnc}\left(y\right)\right); \mathbf{L}\left(\mathsf{AddEnc}\left(Y\right)\right))$ *equals*

$$\frac{1}{2} \sum_{\boldsymbol{\ell} \in \mathcal{L}^d} \left| \sum_{\alpha \in \mathbb{F}^\star} \left(\prod_{j=1}^{d} \widehat{\mathbf{1}_{L_j^{-1}(\ell_j)}}(\alpha) \right) e^{-\frac{2i\pi\alpha y}{p}} \right| \ . \tag{9}$$

Stated as is, the right-hand sides of the equalities in Proposition 3 are not numerically tractable, as they sum over $\mathcal{L}^d$, which becomes quickly hard when the range of $\mathcal{L}$ or d increase. Monte-Carlo estimations can be used to partially circumvent the problem [33]. But the core idea of such simulations is to leverage the phenomenon of *concentration* of probability distributions in high-dimensional spaces [9]. In this respect, they are efficient whenever the dimensionality of the leakage space — d here — increases, but not necessarily whenever the range of $\mathcal{L}$ increases, which is the core question of this paper. That is why we leverage another corollary from Benhamouda *et al.*[6]

Corollary 1 (Cauchy-Schwarz [8, p. 30, restated]). *Let $\mathbf{L} = (L_1, \ldots, L_d)$ be any family of leakage. Then, $\mathsf{M}_\infty\left(\mathbf{L}\right)$ is upper bounded by*

$$\frac{1}{p} \left(\sum_{\ell_1} \left\| \mathbf{1}_{L_1^{-1}(\ell_1)} \right\|_2 \right) \cdot \left(\sum_{\ell_2} \left\| \mathbf{1}_{L_2^{-1}(\ell_2)} \right\|_2 \right) \cdot \prod_{j=3}^{d} \left(\sum_{\ell_j} \max_{\alpha \in \mathbb{F}^\star} \left| \widehat{\mathbf{1}_{L_j^{-1}(\ell_j)}}(\alpha) \right| \right) \ . \tag{10}$$

Corollary 1 provides an upper bound for the worst-case metric as a product of d sums over $\mathcal{L}$, whose complexity grows linearly with both the range of $\mathcal{L}$ and d. Hence, the right-hand side of Equation 10 is much more tractable and can be exactly computed — up to negligible numerical errors. Benhamouda *et al.* also leverage Corollary 1 to establish an upper bound of the worst-case metric for any noise-free m-bounded leakage function, *i.e.*, any function that can take up to 2^m different values, for some integer m.

Theorem 1 ([7, Thm. 4.7], restated). *For a secret $Y \leftarrow \mathbb{F}_p$ protected with additive secret sharing, let $\mathbf{L} = (L_1, \ldots, L_d)$ be any family of leakage functions where $L_i : \mathbb{F}_p \rightarrow \{0,1\}^m$. Let $c_m = \frac{2^m \sin(\pi/2^m)}{p \sin(\pi/p)} < 1$ (when $2^m < p$). Then,*

$$\mathsf{M}_\infty\left(\mathbf{L}\right) \leq 2^m \cdot c_m^{d-2} \ . \tag{11}$$

For the average-case metric with uniform Y, the upper bound stays the same, but with the additional multiplicative factor $\frac{1}{2}$.

Equation 11 provides a non-trivial upper bound on the worst-case bias whenever the field size p increases. This upper bound is asymptotically tight, as Maji *et al.* have exhibited a leakage function, namely the LSB, for which there is a lower bound asymptotically matching the right-hand side of Equation 11, up to a

[6] The Cauchy-Schwarz trick has already been used in the noise amplification bound of Prouff & Rivain [40, Thm. 1], although applied to another metric.

small constant factor. Theorem 1 thereby provides an upper bound on the statistical bias, as it can be trivially upper bounded by the worst-case bias. Therefore, the only missing inequality to get a comprehensive view of the leakage-resilience of m-bounded leakage functions is a lower bound on the statistical bias. This will be the focus of section 3. Furthermore, we note that in Theorem 1 the right-hand side of Equation 11 is non-trivial when p increases if and only if the range m of the leakage function is constant. This assumption does not always hold, and there are concrete counter-examples of leakage functions for which the range m depends on the field size, such as the HW — where $m \approx \log \log p$. As a result, one should go back to Corollary 1 to tighten the upper bound through a refined analysis of the Fourier coefficients of the specific leakage function under study. We will instantiate this approach for the HW leakage in section 4.

3 Bit leakages

In this section, we investigate how the information derived from bit leakage is affected by the field size. Our findings indicate that this leakage model exhibits limited sensitivity to changes in the field size, up to the point that we can lower-bound the amount of information by a constant in p. As an initial step, we provide an outlook on a prior result of Maji *et al.* [29]. In their work, they already showed that the information deriving from LSB leakage can be lower-bounded by a constant in p in the worst-case scenario. Following, we demonstrate that the weak noise amplification observed in the worst-case setting actually occurs for a large fraction of secrets. In other words, we provide a lower bound that is independent of p for the LSB in the average-case metricl. Eventually, we move out of the LSB leakage model and achieve the same lower bounds for the more general case of bit probing. We conclude that there is a barrier beyond which security against bit probing cannot be enhanced by increasing the field size.

In the first part of this section, we consistently work in the LSB leakage model. Henceforth, unless specified differently, we use the more general notation $\mathbf{L} : \mathbb{F}_p^{d+1} \to \{0,1\}^{d+1}$ for the function returning the LSB of every component.

3.1 Worst-case characterization

With the following theorem, we recall the LSB analysis in the worst-case metric of [29]. Their main observation is that the information obtained from LSB leakage can be lower-bounded by a constant in p in the worst-case scenario.

Theorem 2 ([29, Thm. 10], restated). *Let p be a prime ≥ 3. Then, for every number of shares $d \in \mathbb{N}$*

$$\mathsf{M}_\infty(\mathbf{L}) \geq \frac{1}{2} \cdot \left(\frac{2}{\pi}\right)^d \cdot \left[\frac{3}{2} - 4\left(\left(\frac{2}{3}\right)^d + \frac{1}{d+1}\left(\frac{1}{3}\right)^{d+1}\right)\right]. \qquad (12)$$

We note that the above result is non-trivial only when the right-hand side is non negative, which happens for every $d \geq 3$. We remark that Theorem 2 slightly differs from [29, Thm. 10]. This is because the latter is only stated asymptotically, while the actual parameters only appear inside the corresponding proof. In Theorem 2, we condense the information from [29] to provide a more comprehensive version of the same statement.

3.2 Average-case characterization

Given that $\mathsf{M}_\infty(\mathbf{L}) \leq p \cdot \beta(\mathrm{Y}|\mathbf{L})$ (from Eq. (1)), Theorem 2 already yields a lower bound to the average-case metric, that is

$$\beta(\mathrm{Y}|\mathbf{L}) \geq \frac{1}{2p} \cdot \left(\frac{2}{\pi}\right)^d \cdot \left[\frac{3}{2} - 4\left(\left(\frac{2}{3}\right)^d + \frac{1}{d+1}\left(\frac{1}{3}\right)^{d+1}\right)\right]. \tag{13}$$

In contrast with the worst-case result, this lower bound could suggest that increasing the field size may result in a more effective strategy against LSB attacks for random secrets. In other words, the possible tightness of Eq. (13) would reveal that only a handful of secrets yield the weak noise amplification observed in the worst-case scenario. In fact, we show that the lower bound of Eq. (13) can be improved to remove the dependence on p. This means that the weak noise amplification involves the majority of secrets, and thus must be considered in the randomized setting as well. We make this finding formal with Theorem 3. From a technical standpoint, our key observation is that Eq. (12) can be slightly modified to lower-bound the information derived from at least half of the possible secrets. Later in the section, we provide a discussion of our results, and explore their applicability within the broader bit-probing model.

Theorem 3. *Let p be a prime ≥ 3 and $d \in \mathbb{N}$ be any number of shares. Let Y be the uniform distribution over $\mathbb{F}_p$. Then $\beta(Y|\mathbf{L})$ is lower-bounded by*

$$\frac{1}{2} \cdot \left(\frac{2}{\pi}\right)^d \cdot \left[\frac{p+1}{2p} \cdot \sqrt{2} \cdot \sqrt{1 - \sin\left(\frac{\pi}{2p}\right)} - 2 \cdot \left(\left(\frac{2}{3}\right)^d + \frac{1}{d+1}\left(\frac{1}{3}\right)^{d+1}\right)\right].$$

As for the worst-case analysis of [29], this lower bound is non-trivial whenever $d \geq 3$.

We now outline the proof of Theorem 3. As in [29], the first step is to rewrite the Statistical Distance (SD) in the Fourier domain. Using Eq. (9) from Proposition 3, we get

$$\beta(\mathrm{Y}|\mathbf{L}) = \frac{1}{2p} \sum_{y \in \mathbb{F}_p} \sum_{\boldsymbol{\ell} \in \{0,1\}^d} \left| \sum_{\alpha \in \mathbb{F}^\star} \left(\prod_{j=1}^{d} \widehat{\mathbf{1}_{\mathrm{L}_j^{-1}(\ell_j)}}(\alpha) \right) e^{-\frac{2i\pi\alpha y}{p}} \right|. \tag{14}$$

As observed in [29, Claim 15], the LSB Fourier coefficients for $\alpha \in \mathbb{F}^*$ satisfy

$$\widehat{\mathbf{1}_{\mathrm{L}_j^{-1}(0)}}(\alpha) = -\widehat{\mathbf{1}_{\mathrm{L}_j^{-1}(1)}}(\alpha) = \frac{1}{2p} \cdot \frac{1}{\cos(\pi\alpha/p)} \cdot e^{\frac{i\pi\alpha}{p}}.$$

Therefore, Eq. (14) becomes

$$\beta(\mathrm{Y}|\mathbf{L}) = \frac{2^{d-1}}{p} \sum_{y \in \mathbb{F}_p} \left| \sum_{\alpha \in \mathbb{F}^\star} \left(\frac{1}{2p} \cdot \frac{1}{\cos(\pi\alpha/p)} \cdot e^{\frac{i\pi\alpha}{p}} \right)^d e^{-\frac{2i\pi\alpha y}{p}} \right| .$$

Similar to [29], we observe that the dominant terms are those corresponding to $\alpha = \left\{\frac{p-1}{2}, \frac{p+1}{2}\right\}$. However, since we consider an *average-case* metric, the above observation needs to hold for a large enough set of field elements $y \in \mathcal{Y}$. Therefore, the proof relies on the following lemmas.

Lemma 3. *Let $p \geq 3$. There exist a subset $\tilde{\mathcal{Y}} \subseteq \mathbb{F}_p$ such that every $y \in \tilde{\mathcal{Y}}$ satisfies*

$$\begin{aligned} &\left| \sum_{\alpha \in \left\{\frac{p-1}{2}, \frac{p+1}{2}\right\}} \left(\frac{1}{2p} \cdot \frac{1}{\cos(\pi\alpha/p)} \cdot e^{\frac{i\pi\alpha}{p}} \right)^d \cdot e^{-\frac{2i\pi\alpha y}{p}} \right| \\ &\qquad \geq \pi^{-d} \cdot \sqrt{2} \cdot \sqrt{1 - \sin\left(\frac{\pi}{2p}\right)} . \end{aligned} \tag{15}$$

Lemma 4. *Let $p \geq 3$. For every secret $y \in \mathbb{F}_p$, it holds*

$$\begin{aligned} &\left| \sum_{\alpha \in \mathbb{F}^\star \setminus \left\{\frac{p-1}{2}, \frac{p+1}{2}\right\}} \left(\frac{1}{2p} \cdot \frac{1}{\cos(\pi\alpha/p)} \cdot e^{\frac{i\pi\alpha}{p}} \right)^d \cdot e^{-\frac{2i\pi\alpha y}{p}} \right| \\ &\qquad \leq 2\pi^{-d} \cdot \left(\left(\frac{2}{3}\right)^d + \frac{1}{d+1} \left(\frac{1}{3}\right)^{d+1} \right) . \end{aligned}$$

Given that Lemma 3 holds for $\frac{p+1}{2}$ elements, and Lemma 4 is true for every field element, Theorem 3 follows by triangular inequality

Proof of Theorem 3.

$$\begin{aligned} \beta(\mathrm{Y}|\mathbf{L}) \geq & \frac{2^{d-1}}{p} \sum_{y \in \mathbb{F}_p} \left| \sum_{\alpha \in \left\{\frac{p-1}{2}, \frac{p+1}{2}\right\}} \left(\frac{1}{2p} \cdot \frac{1}{\cos(\pi\alpha/p)} \cdot e^{\frac{i\pi\alpha}{p}} \right)^d e^{-\frac{2i\pi\alpha y}{p}} \right| \\ & - \frac{2^{d-1}}{p} \sum_{y \in \mathbb{F}_p} \left| \sum_{\alpha \in \mathbb{F}_p^\star \setminus \left\{\frac{p-1}{2}, \frac{p+1}{2}\right\}} \left(\frac{1}{2p} \cdot \frac{1}{\cos(\pi\alpha/p)} \cdot e^{\frac{i\pi\alpha}{p}} \right)^d e^{-\frac{2i\pi\alpha y}{p}} \right| . \\ \geq & \frac{1}{2} \left(\frac{2}{\pi}\right)^d \left[\frac{p+1}{2p} \sqrt{2} \sqrt{1 - \sin\left(\frac{\pi}{2p}\right)} - 2 \left(\left(\frac{2}{3}\right)^d + \frac{1}{d+1} \left(\frac{1}{3}\right)^{d+1} \right) \right] \end{aligned}$$

□

To complete the analysis, we present the proofs of Lemma 3 and Lemma 4.

Proof of Lemma 3. We first restate the left-hand side of Eq. (15) as

$$\left|\frac{1}{2p}\cdot\frac{1}{\sin\left(\frac{\pi}{2p}\right)}\right|^d\cdot\left|1+(-1)^d e^{\frac{i\pi d}{p}}e^{-\frac{2i\pi y}{p}}\right|$$

and observe that it can be lower-bounded by $\pi^{-d}\cdot\left|1+e^{i\pi\left(d+\frac{d}{p}-\frac{2y}{p}\right)}\right|$ because $\frac{1}{x\sin(1/x)}$ is a decreasing function in x, and $\lim_{x\to\infty}\frac{1}{x\sin(1/x)}=1$. So it remains to find $\tilde{\mathcal{Y}}$ so that, for every $y\in\tilde{\mathcal{Y}}$, we can lower-bound $\left|1+e^{i\pi\left(d+\frac{d}{p}-\frac{2y}{p}\right)}\right|$.

We use the fact that, whenever x lies in $\left[-\frac{p+1}{2p},\frac{p+1}{2p}\right]$, then

$$\left|1+e^{i\pi x}\right|=\sqrt{2}\cdot\sqrt{1+\cos(\pi x)}\geq\sqrt{2}\cdot\sqrt{1-\sin\left(\frac{\pi}{2p}\right)},$$

as cos is symmetric in $\left[-\frac{p+1}{2p}\pi,\frac{p+1}{2p}\pi\right]$ and decreasing in $\left[0,\frac{p+1}{2p}\pi\right]$.

This means that, whenever y belongs to the interval

$$\mathcal{I}=\left[\frac{p}{2}\left(-\frac{p+1}{2p}+d+\frac{d}{p}\right),\frac{p}{2}\left(\frac{p+1}{2p}+d+\frac{d}{p}\right)\right]\mod p,$$

then

$$\left|1+e^{i\pi\left(d+\frac{d}{p}-\frac{2y}{p}\right)}\right|\geq\sqrt{2}\cdot\sqrt{1-\sin\left(\frac{\pi}{2p}\right)}.$$

Let $\tilde{\mathcal{Y}}=\mathbb{F}_p\cap\mathcal{I}$. Since $\mathcal{I}$ has length $\frac{p+1}{2}$, then $\tilde{\mathcal{Y}}$ has at least $\frac{p+1}{2}$ elements. This concludes the proof. □

Proof of Lemma 4. Using the triangular inequality on the left-hand side of Lemma 4, we get

$$\left|\sum_{\alpha\in\mathbb{F}^\star\setminus\left\{\frac{p-1}{2},\frac{p+1}{2}\right\}}\left(\frac{1}{2p}\cdot\frac{1}{\cos(\pi\alpha/p)}\cdot e^{\frac{i\pi\alpha}{p}}\right)^d\cdot e^{-\frac{2i\pi\alpha y}{p}}\right|$$
$$\leq\sum_{\alpha\in\mathbb{F}^\star\setminus\left\{\frac{p-1}{2},\frac{p+1}{2}\right\}}\left|\frac{1}{2p}\cdot\frac{1}{\cos(\pi\alpha/p)}\cdot e^{\frac{i\pi\alpha}{p}}\right|^d.$$

Then, the lemma follows from the following observation of [29] (proof of Claim 17)

$$\sum_{\alpha\in\mathbb{F}^\star\setminus\left\{\frac{p-1}{2},\frac{p+1}{2}\right\}}\left|\frac{1}{2p}\cdot\frac{1}{\cos(\pi\alpha/p)}\cdot e^{\frac{i\pi\alpha}{p}}\right|^d$$
$$\leq 2\pi^{-d}\cdot\left(\left(\frac{2}{3}\right)^d+\frac{1}{d+1}\left(\frac{1}{3}\right)^{d+1}\right).$$

□

3.3 Discussion

On the field size (in)dependence. Asymptotically in p, both the worst-case and the average-case metrics are lower-bounded by a value independent of p. The latter defines a barrier beyond which security cannot be enhanced by increasing the field size. This property is inherent to the LSB leakage model, as its maximal Fourier coefficients over $\mathbb{F}_p^*$ do not converge towards zero as p approaches infinity. This observation is already made formal in Theorem 2 and Theorem 3, but we can provide a more pictorial intuition of why it's true.

First note that, for every $\alpha \in \mathbb{F}_p$ and for every $S \subseteq \mathbb{F}_p$, we can restate

$$\widehat{\mathbf{1}_S}(\alpha) = \frac{|S|}{p} \cdot \frac{1}{|S|} \sum_{z \in \alpha S} e^{-\frac{2\pi z i}{p}}.$$

That is, $\widehat{\mathbf{1}_S}(\alpha)$ equals the barycenter of the roots of unity corresponding to αS up to the multiplicative factor $\frac{|S|}{p}$. When $S = \mathsf{lsb}^{-1}(0)$, then both $\frac{p-1}{2} S$ and $\frac{p+1}{2} S$ yield sets of $\frac{p+1}{2}$ consecutive field elements. That is, the corresponding Fourier coefficients converge to the barycenter of half of the unit circle up to a constant, which is not zero. Fig. 1 shows the Fourier spectrum of $\mathbf{1}_{\mathrm{L}_j^{-1}(0)}$ in the LSB model for different primes p.

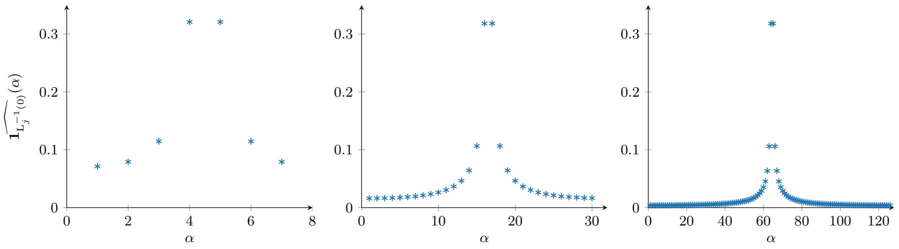

Fig. 1: α vs $\widehat{\mathbf{1}_{\mathrm{L}_j^{-1}(0)}}(\alpha)$ for $\alpha \in \mathbb{F}^*$ in the LSB leakage model, for $p = 7, 31, 127$.

On tightness. As observed in [29], the lower bounds for the LSB leakage model can be used to argue about the tightness of the upper bounds of [7]. In particular, Theorem 1 states that the information provided by any leakage function with range 1 satisfies

$$\mathsf{M}_\infty(\mathbf{L}) \leq 2 \cdot \left(\frac{2}{p \cdot \sin\left(\frac{\pi}{p}\right)} \right)^{d-2} ,$$

which converges to $\frac{\pi^2}{2} \cdot \left(\frac{2}{\pi}\right)^d$ asymptotically in p. When excluding the factor 2, the same upper bound extends to the average-case metric. As seen in this section, both the worst-case and the average-case metric in the LSB model are lower-bounded by $\left(\frac{2}{\pi}\right)^d$ modulo a factor that only depends on d. Therefore, these result witness that Theorem 1 is asymptotically tight for leakage functions with range 1, modulo a factor that only depends on d.

In Fig. 2, we compare the LSB lower and upper bounds with the corresponding numerical computation in both the average and worst-case metrics.

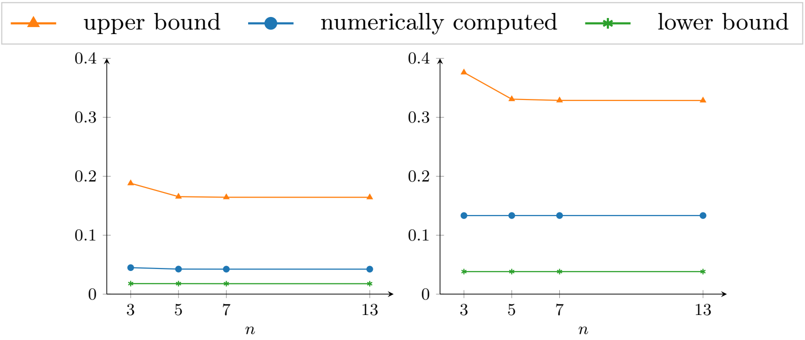

Fig. 2: Both images illustrate the comparison between the upper bound from [7] and the lower bounds discussed in this section with the corresponding numerical computation. The values displayed correspond to different field size values, computed as $p = 2^n - 1$, and fixed number of shares $d = 6$. The left image pertains to the average-case metric, while the right one refers to the worst-case metric.

On arbitrary bit leakage. As observed in [22], the above remark readily extends to the case where the adversary can probe an arbitrary bit location. This follows from the fact that shifting backwards $n - k$ times a string whose k-th significant bit is zero yields one whose LSB is zero. More formally, denote with ksb the function returning the k-th significant bit. Then

$$2^k \mathsf{ksb}^{-1}(0) = \mathsf{lsb}^{-1}(0).$$

This property defines a bijection between the Fourier coefficients of all single-bit probing models, *i.e.* for every $\alpha \in \mathbb{F}_p$ and $k \in [\![1, n]\!]$,

$$\widehat{\mathbf{1}_{\mathsf{lsb}^{-1}(0)}}(\alpha) = \widehat{\mathbf{1}_{\mathsf{ksb}^{-1}(0)}}(2^k \alpha).$$

As a consequence, the maximizing Fourier coefficients in the ksb leakage model are those for $\alpha \in \left\{2^{k-1}, p - 2^{k-1}\right\}$, and there is no convergence towards zero when the field size goes to infinity. In line with this observation, we can extend Theorem 2 and Theorem 3 to the ksb leakage model.

Theorem 4 (Generalizing Theorem 2 to the ksb **model).** *Let p be a prime ≥ 3 and let $\mathbf{L} : \mathbb{F}_p^d \to \{0,1\}^d$ be the function returning the* ksb *of every component. Then, for every $d \in \mathbb{N}$*

$$\mathsf{M}_\infty(\mathbf{L}) \geq \frac{1}{2} \cdot \left(\frac{2}{\pi}\right)^d \cdot \left[\frac{3}{2} - 4\left(\left(\frac{2}{3}\right)^d + \frac{1}{d+1}\left(\frac{1}{3}\right)^{d+1}\right)\right].$$

Theorem 5 (Generalizing Theorem 3 to the ksb **model).** *Let p be a prime ≥ 3, $\mathbf{L} : \mathbb{F}_p^d \to \{0,1\}^d$ be the function returning the* ksb *of every component, and $d \in \mathbb{N}$ be the number of shares. Let Y be the uniform distribution over $\mathbb{F}_p$. Then, $\beta(Y|\mathbf{L})$ is lower-bounded by*

$$\frac{1}{2} \cdot \left(\frac{2}{\pi}\right)^d \cdot \left[\frac{p+1}{2p} \cdot \sqrt{2} \cdot \sqrt{1 - \sin\left(\frac{\pi}{2p}\right)} - 2 \cdot \left(\left(\frac{2}{3}\right)^d + \frac{1}{d+1}\left(\frac{1}{3}\right)^{d+1}\right)\right].$$

Proof sketch. The proofs of Theorem 4 and Theorem 5 follow the same structure as those for the LSB leakage model, but instead of isolating $\alpha \in \left\{\frac{p-1}{2}, \frac{p+1}{2}\right\}$, we isolate the terms corresponding to the maximal Fourier coefficients in the ksb model, *i.e.* $\alpha \in \left\{2^{k-1}, p - 2^{k-1}\right\}$. We upper-bound the sum of the terms different from $\left\{2^{k-1}, p - 2^{k-1}\right\}$ using the same strategy as in the proof of Lemma 4. Namely, we upper-bound it with the sum of all the corresponding Fourier coefficients. Given the bijection with the non-dominant Fourier coefficients for LSB, the same upper bound holds. It remains to lower-bound the summands corresponding to $\left\{2^{k-1}, p - 2^{k-1}\right\}$. Note that by changing the variable y to $2^k y$, the problem reduces to the estimation of the dominant terms corresponding to $\left\{\frac{p-1}{2}, \frac{p+1}{2}\right\}$ in the LSB model. Therefore, the same statement holds. □

From Leaking a Single-bit to a Proportion of Bits. We close this section by discussing to which extent the results established for a single-bit leakage apply as well when several of the bits are revealed to the adversary. In this respect, we first emphasize that the upper bound from Theorem 1 already covers this case. As for the lower bound, we observe that a "single-bit" adversary can be trivially simulated by an adversary having access to several bits of each share. As a consequence, the lower bounds of Theorems 4 and 5 remain true.[7] Hence, leaking a proportion of bits keeps the conclusion of this section unchanged.

4 Hamming Weight Leakages

The previous section has focused on the LSB leakage function, as it is "a realistic and analytically-tractable leakage function" [29, p. 2]. From the physical viewpoint, observing such bit leakages is quite challenging though, and a more realistic leakage function, on which we focus in this section, is the HW, which maps a value to the sum of its bits [32]. As argued at the end of section 2, Theorem 1 does not cover such leakage functions, as their range increases with the

[7] This can be formalized by applying the data processing inequality to the TV.

field size leading ultimately to trivial bounds. To circumvent this issue, we need to go one step backwards by starting our analysis from Corollary 1, and trying to refine the Fourier analysis of our specific leakage function under scrutiny.

Hereupon, notice how the right-hand side of Equation 10 depends on the quantity $\sum_h \max_{\alpha\neq 0}\left|\widehat{\mathbf{1}_{\mathsf{HW}^{-1}(h)}}(\alpha)\right|$. This quantity illustrates how sensitive is the highest Fourier coefficient of the leakage function for the resulting security bound. In order to get some insights into the quantity behavior, Figure 3 plots the exemplary Fourier spectra for the bit leakage function studied in the previous section, and for the Hamming weight model that we study now.

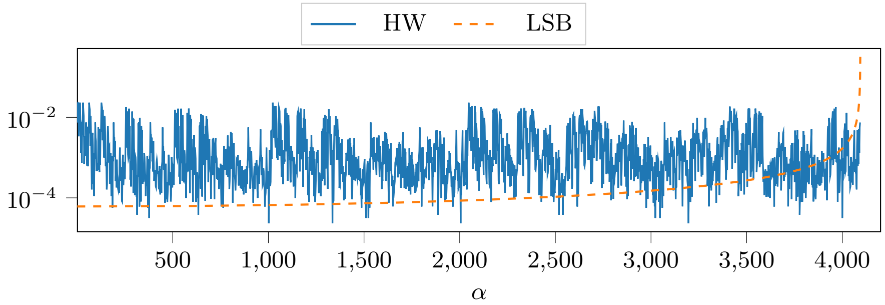

Fig. 3: First half of Fourier spectra of $\mathbf{1}_{\mathsf{HW}^{-1}(n/2)}$ (Hamming weight leakage) and $\mathbf{1}_{LSB^{-1}(0)}$ (bit leakage) for $n = 13, p = 2^n - 1$ (the second half is symmetric).

On the one hand, the orange curve denotes the spectrum of a pre-image set of the LSB. This spectrum is described by a very smooth curve with a peak close to the value $p/2$, as more formally discussed in the proof outline of section 3. On the other hand, the blue curve denotes the spectrum of a pre-image set of the HW leakage function. This curve is much less smooth than the orange curve, but even worse, there is no concentration of the spectrum around one peak as for the LSB leakage. The missing peak means that we cannot isolate dominant terms as in Lemma 3 and Lemma 4 with $\alpha \in \{\frac{p-1}{2}, \frac{p+1}{2}\}$. Hence, the proof technique used in section 3 does not provide sufficiently tight security bounds for the HW leakage function, and this section requires an alternative approach.

Nevertheless, the security bound is strongly related to the quantity

$$\sum_h \max_{\alpha\neq 0}\left|\widehat{\mathbf{1}_{\mathsf{HW}^{-1}(h)}}(\alpha)\right|.$$

That is why Figure 4 numerically plots this quantity of interest for an increasing Mersenne number. This will help us derive some insights before diving into a more formal result, as it shows that the quantity decreases with increasing field sizes. In particular, the blue curve denoting the quantity of interest is decreasing at a polynomial rate between $\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$ and $\mathcal{O}\left(\frac{1}{n}\right)$, represented as the green and

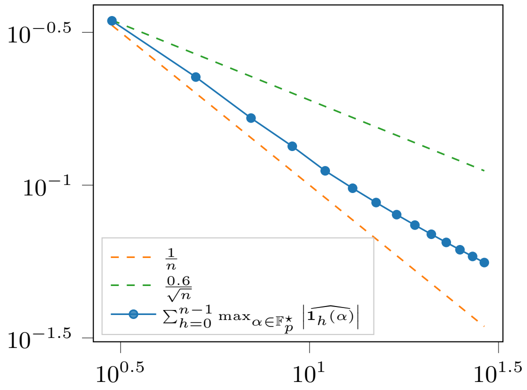

Fig. 4: $\sum_{h=0}^{n-1} \max_{\alpha \neq 0} \left|\widehat{\mathbf{1}}_h(\alpha)\right|$ vs. $n \in [\![3, 29]\!]$ odd, and $p = 2^n - 1$.

orange curve, respectively. In other words, the plot provides a numerical evidence of a bias in $\mathcal{O}\big(\log(p)^{-r \cdot d}\big)$, where $r \in \left[\frac{1}{2}, 1\right]$. In the remaining part of this section, we move towards the formalization of this observation, which in turn allows us to provide a proof for a more conservative upper bound.

4.1 Worst-Case Characterization

We start with an upper bound valid for the worst-case bias.

Theorem 6. *For n odd, let $p = 2^n - 1$ be a Mersenne number. Let $\mathbf{L} : \mathbb{F}_p^d \to [\![0, n-1]\!]^d$ be the function returning $\mathsf{HW}(\cdot)$ for each component. Then the following inequality is valid:*

$$\mathsf{M}_\infty(\mathbf{L}) \leq \frac{1}{2} \cdot \left(\sum_{h=0}^{n-1} \sqrt{\binom{n}{h}}\right)^d \cdot p^{-\frac{d}{2}} \cdot (2n)^{\frac{d}{2}} = \mathcal{O}\Big(n^{1-\frac{d}{4}}\Big) \quad . \tag{16}$$

Compared to the numerical computations depicted in Figure 4 suggesting a convergence in $\mathcal{O}\big(\log(p)^{-r \cdot d}\big)$ for $r \in \left[\frac{1}{2}, 1\right]$, Theorem 6 only guarantees that $r \geq \frac{1}{4}$, and would require at least $d \geq 5$ to be non-trivial. So our provable bound is not completely tight. Still, it proves that the HW leakage model is more leakage-resilient with the help of masking in Mersenne prime fields.

The remaining of this subsection gives an outline of the proof of Theorem 6. It is derived from Corollary 1 thanks to the following theorem bounding the highest Fourier coefficient of each pre-image set of the HW leakage model.

Theorem 7. *For n odd, $p = 2^n - 1$ a Mersenne number and $0 \leq h \leq n - 1$,*

$$\max_{\alpha \neq 0} \left|\widehat{\mathbf{1}}_h(\alpha)\right|^2 \leq \frac{\binom{n}{h}}{p} \cdot \frac{1 - \frac{\binom{n}{h}}{p}}{2 \cdot n} \quad . \tag{17}$$

The full proof of Theorem 7 is deferred to Appendix B. In essence, it starts from Equation 7 in Parseval's theorem applied to the characteristic function $\mathbf{1}_h$, and leverages the following lemma, stating that each term in the right-hand side of Equation 7 applied to $\mathbf{1}_{\mathsf{HW}^{-1}(h)}$ has $2 \cdot n$ similar values in other terms.

Lemma 5. *Let $p = 2^n - 1$. Then, for all $\alpha \in \mathbb{F}_p^\star$ and for all $k \in \mathbb{N}$,*

$$\left|\widehat{\mathbf{1}_h}\left(2^k\alpha\right)\right| = \left|\widehat{\mathbf{1}_h}(\alpha)\right| \quad . \tag{18}$$

Proof of Lemma 5. We leverage the property spotted in [33]: multiplying x by a power of two modulo a Mersenne number, *i.e.* of the shape, $p = 2^n - 1$, is just a rotation of the bits of x. Hence it is invariant for the Hamming weight. Let us then express the left hand-side of Equation 18. For any $k \in \mathbb{N}$,

$$\widehat{\mathbf{1}_h}\left(2^k\alpha\right) = \frac{1}{p}\sum_{x:\mathsf{HW}(x)=h} e^{-\frac{2i\pi 2^k \alpha x}{p}} \quad . \tag{19}$$

We make the following change of variable: let $x' = 2^k x$. Then the sum goes over the same values for x' as for x since multiplying by a power of two modulo a Mersenne prime keeps the Hamming weight unchanged. Meanwhile, the exponent in each term becomes $-\frac{2i\pi\alpha x'}{p}$. We thus identify the Fourier coefficient of harmonic α. □

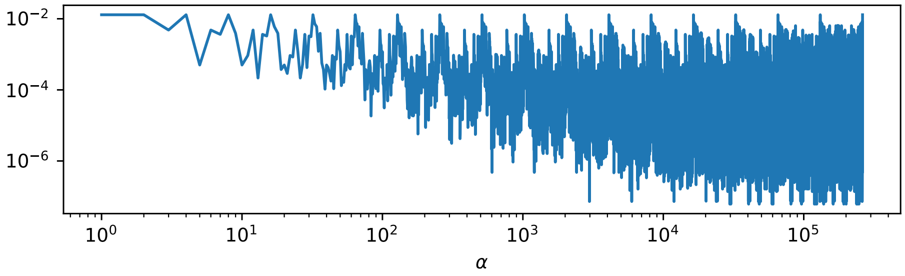

Fig. 5: First half of the Fourier spectrum of $\mathbf{1}_{\lfloor \frac{n}{2} \rfloor}$, $n = 19, p = 2^n - 1$. Notice some regular patterns in the peaks, when displaying the x-axis in log-scale.

As a result of Lemma 5 (illustrated in Figure 5), the sum of squared Fourier coefficients can be factorized by $2 \cdot n$, hence the denominator in Equation 17, making the upper-bound non-trivial for all the observed leakage values h.

4.2 Average-case characterization

As discussed in subsection 4.1, we have established an upper bound for the worst-case bias, which approaches zero as the field size p tends to infinity. Since the average-case bias can be trivially upper bounded by the worst-case bias with $\beta(\mathrm{Y}|\mathbf{L}) \leq \mathsf{M}_\infty(\mathbf{L})$ (Eq. 3), the upper bound derived in the right-hand side of Theorem 6 (Eq. 16) is also a valid upper bound of the average-case bias.

Theorem 8. *For n odd, let $p = 2^n - 1$ be a Mersenne number. Let $\mathbf{L} : \mathbb{F}_p^d \to [\![0, n-1]\!]^d$ be the function returning $\mathsf{HW}(\cdot)$ for each component. Then the following inequality is valid:*

$$\beta(Y|\mathbf{L}) \leq \frac{1}{2} \cdot \left(\sum_{h=0}^{n-1} \sqrt{\binom{n}{h}} \right)^d \cdot p^{-\frac{d}{2}} \cdot (2n)^{\frac{d}{2}} = \mathcal{O}\left(n^{1-\frac{d}{4}}\right) \quad . \tag{20}$$

4.3 Discussion

On the field size dependence. Asymptotically with the field size p, the worst-case bias is upper bounded by a value following a poly-log trend, whereas asymptotically with the number of shares, the bias goes to zero exponentially fast. In a sense, the $n = \log(p)$ value in Equation 16 and Equation 20 may be seen as a surrogate of the Gaussian noise parameter in the early security analysis of Chari *et al.* [16], hence the interest of turning masking into prime fields in the particular setting of noise-free Hamming weight leakages.

On the Tightness. So far in this section, we have explained why the techniques used for the LSB analysis could not be applied for the HW, before providing upper bounds for the latter leakage using another approach through Parseval's identity, and we have said that the derived upper bounds are not completely tight with what is observed on numerical computation. We may therefore wonder where this gap comes from. Essentially, the core idea of our proof was to upper bound the maximum Fourier coefficient — in squared absolute value — by a sum provably containing this maximum coefficient. This approach provides a tight bound only if the Fourier spectrum is concentrated in its maximum coefficient — like for the bit leakage. Looking at Figure 5, it is clear that many more Fourier coefficients are dominant, *i.e.*, of the same order of magnitude as the maximum one, without being exactly equal to the maximum. Therefore, one possible way to improve the inequalities could be to first prove which value of α maximizes the Fourier transform.[8] Then, one could find other values of α for which the Fourier transform is close in absolute value to the maximum coefficient, and to bound them by a quantity depending the latter one. Hence the hope would be to increase the $2 \cdot n$ denominator of the right-hand side of Equation 17.

5 Empirical Evaluation

The results in Sections 3 and 4 provide formal confirmation that increasing the size of the modulus used in prime-field masking leads to significant security gains

[8] We conjecture based on the numerical calculations of the spectrum that for $n \geq 13$, $\left|\widehat{\mathbf{1}_h}(1)\right|$ maximizes the Fourier transform.

for the HW leakage function and does not lead to any gains for the LSB one. While the HW function is a more realistic abstraction for the power consumption of actual implementations [32], it remains a quite abstract one and in practice, implementations may show up leakage models that are highly correlated with the HW function without exactly matching it [10]. This raises the question whether such small deviations from the HW leakage function directly bring us back to the LSB case, or whether a more graceful degradation takes place. Formally answering this question will require to characterize classes of leakage functions for which similar results as in Sections 3 and 4 can be obtained. As a first step in this direction, we next extend the simulated experiments of Masure et al. from [33] towards linear leakage functions that generalize the HW one. Precisely, and inspired by [42], we consider a leakage function that outputs a weighted sum of bits, like the HW one, but with less constraints on the weights:

$$\mathrm{L}(\mathrm{Y}) = \sum_{i=1}^{\lceil \log(p) \rceil} \omega_i \cdot Y(i),$$

with $Y(i)$ the ith bit of Y. In the HW case, $\omega_i = 1$ for all i's. We propose two generalizations: the *Skewed Hamming Weight (SHW)* function where only the LSB gets a higher weight s, and the the *Random Linear (Rlin)* functions where all coefficients are picked up uniformly at random between 1 and s.

We then ran Monte-Carlo simulations: we uniformly drew $N = 1,000$ additive encodings $\mathsf{AddEnc}\left(\mathrm{Y}\right)_i \leftarrow \mathbb{F}_p^d$ for which we applied the leakage model L under study to each share. It leads to a dataset of leakages $\{\mathbf{L}_1, \ldots, \mathbf{L}_N\}$ and their corresponding PMF. The statistical bias is therefore estimated as follows:

$$\beta(\mathrm{Y}|\mathbf{L}) = \mathop{\mathbb{E}}_{\mathbf{L}}\left[\mathsf{TV}\Big(\mathsf{p}_{\mathrm{Y}\,|\,\mathbf{L}}; \mathsf{p}_{\mathrm{Y}}\Big)\right] \approx \frac{1}{N}\sum_{i=1}^{N} \mathsf{TV}\Big(\mathsf{p}_{\mathrm{Y}\,|\,\mathbf{L}=\boldsymbol{\ell}_i}; \mathsf{p}_{\mathrm{Y}}\Big) \quad .$$

The simulations are repeated for different field sizes, and for different number of shares d. The masked PMF takes the shape of a discrete convolution product that can be seen as an instance of a so-called SASCA attack, which is efficiently implemented in the SCALib library for estimating this statistical bias up to $d = 6$ shares [15]. All simulations depicted in Figure 6 assume noise-free leakages.

Starting with the upper plots, which correspond to the SHW case, we can see that increasing s gradually decreases the interest of increasing the prime modulus p. This can be explained by the "isolating effect" of increasing s (i.e., with sufficiently large s, a single-bit is leaked on top of the HW information). There is a single curve per size of p in this case since we only vary the single coefficient ω_1. A similar gradual degradation effect can be observed for the lower plots, which correspond to the RLin case. Here we have multiple curves per size of p since coefficients ω_i are picked up uniformly at random.

Both sets of plots motivate the further characterization of these generalized leakages. For the SHW case, it is for example questionable whether the "isolating

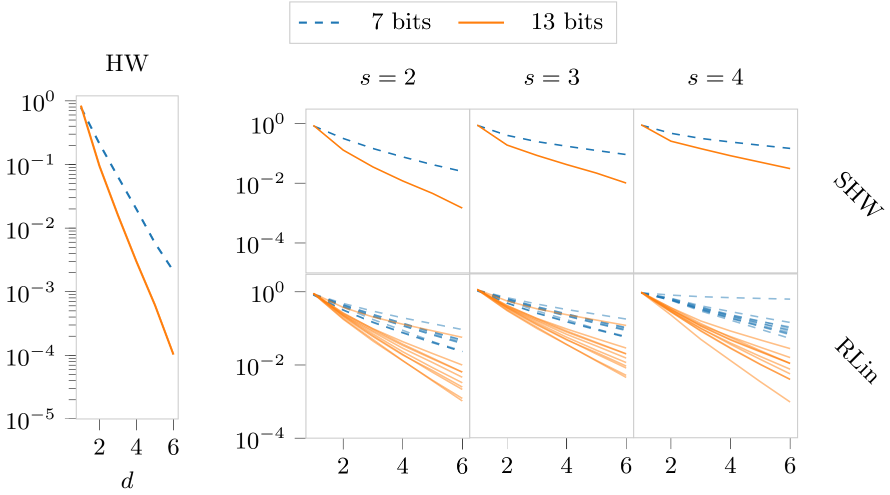

Fig. 6: Estimation of the average-case bias vs. number of shares d.

effect" that is caused by an increase of s can be mitigated by an increase of the field size. But experimenting with large fields is computationally-intensive and putting forward a theoretical trend would be handy. For the Rlin case, we observe a significant variation for different functions with coefficients having the same range. So it suggests a need to characterize the leakage functions with another parameter than s in order to capture our empirical observations.

6 Conclusions and open problems

By combining theoretical advances from the leakage-resilient secret sharing literature with a formal analysis of two representative leakage functions, our results make an important step towards understanding the potential of prime-field masking and motivating the design of dedicated ciphers for this purpose. In particular, they show that for practically-relevant leakage functions like the Hamming weight one (and small variations thereof), increasing the size of the prime modulus p can lead to important security gains, so that $\log(p)$ can serve as a security parameter like the noise variance in Boolean masking. Such increases should in turn be compared with the (possibly limited [12]) performance overheads that computations in prime field imply. We conjecture that using small primes (e.g., 8- to 16-bit) can lead to excellent results for hardware implementations and that software implementations (with efficient multipliers) will benefit from prime sizes up to 32-bit. The careful investigations of symmetric designs enabling to back up this claim is an important scope for further investigations.

Besides, our formal results are for now limited to two canonical leakage functions, and we rely on our empirical evaluations to gain confidence that the good properties of the Hamming weight leakage function can be generalized. This leads to the other important open problem of formalizing this generalization and providing theoretical evidence that large classes of practically-relevant leakage functions can actually benefit from prime-field masking in larger fields. We believe the investigation of linear leakage functions that we initiate in work are natural candidates for this purpose, as they provide a good characterization of the physical measurements met in practice [42]. In this respect, another important problem will be to investigate how much the increase of the prime modulus p can be combined with an increase of the noise level. On the one hand, both have been shown to combine gracefully in the experiments of [33]. On the other hand, linear leakage functions with increasing granularity (i.e., larger s parameter in the experiments of Section 5) will inevitably end up being bijective (and therefore trivially insecure) in the noise-free setting. Yet, it is unlikely that it poses an insurmountable security issue, since for many relevant leakage functions, we expected that a mild noise addition will be enough to lift this granularity.

We hope these questions and observations stimulate further research towards analyzing prime-field masking under increasingly realistic assumptions.

Acknowledgments. François-Xavier Standaert is a senior research associate of the Belgian Fund for Scientific Research (FNRS-F.R.S.). This work has been funded in parts by the German Research Foundation (DFG) CRC 1119 CROSSING (project S7), by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE, by the ERC Consolidator Grant 101044770 (CRYPTOLAYER) and by the ERC Advanced Grant 101096871 (BRIDGE). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

A Proofs of section 2

Proof of Lemma 1. By definition we have

$$\begin{aligned}
\Pr(\mathrm{Y} = y \mid \mathrm{L} = h) &= \frac{1_{\mathsf{HW}\,y=h}}{\binom{n}{h}}, h \in [\![0, n-1]\!] \ , \\
\Pr(\mathrm{Y} = y) &= \frac{1}{2^n - 1} \ , \\
\Pr(\mathrm{L} = h) &= \frac{\binom{n}{h}}{2^n - 1}, h \in [\![0, n-1]\!] \ .
\end{aligned}$$

It follows that the statistical distance for Hamming weight is

$$\begin{aligned}\beta(\mathrm{Y}|\mathrm{L}) &= \frac{1}{2}\sum_{h=0}^{n-1}\sum_{y\in\mathcal{Y}}\left|\frac{1_{\mathsf{HW}\,y=h}}{2^n-1} - \frac{\binom{n}{h}}{(2^n-1)^2}\right| \\ &= \frac{1}{2^n-1}\sum_{h=0}^{n-1}\binom{n}{h}\left(1-\frac{\binom{n}{h}}{2^n-1}\right) \quad .\end{aligned}$$

We may then leverage Vandermonde's convolution equality to replace the sum of squared binomial coefficients by $\binom{2n}{n}-1$. Finally, using Stirling's formula, we get the desired result.

As per the MI, notice that it is equal to the entropy of a binomial distribution, truncated from its maximum value n, and re-normalized by a factor $\frac{2^n}{2^n-1} \approx 1 + 2^{-n}$. We argue hereafter that this slight change to the binomial distribution does not change the entropy approximation. First, the re-normalization factor does not change much the non-truncated terms of the entropy. As per the truncated term, it initially contributed to the entropy at a value $\frac{n}{2^n} \ll \log(n)$. □

Proof of Lemma 2. The MI of this leakage model is trivially $\alpha \cdot n$. As per the bias, first, assume $\alpha < 1$. Then, for every $h \in \{0,1\}^{\alpha n}$

$$\begin{aligned}\Pr(\mathrm{Y}=y) &= \frac{1}{2^n-1} \quad , \\ \Pr(\mathrm{Y}=y \mid \mathrm{L}=h) &= \begin{cases} 0 & \text{if } h \neq \ell^S(y) \\ \frac{1}{2^{n(1-\alpha)}} & \text{else if } h \neq 1^{\alpha n} \\ \frac{1}{2^{n(1-\alpha)}-1} & \text{else if } h = 1^{\alpha n} \end{cases} \quad , \\ \Pr(\mathrm{L}=h) &= \begin{cases} \frac{2^{n(1-\alpha)}}{2^n-1} & \text{if } h \neq 1^{\alpha n} \\ \frac{2^{n(1-\alpha)}-1}{2^n-1} & \text{if } h = 1^{\alpha n}. \end{cases}\end{aligned}$$

Therefore

$$\beta(\mathrm{Y}|\mathrm{L}) = \frac{1}{2}\sum_{\substack{h\in\{0,1\}^{\alpha n}\\ h\neq 1^{\alpha n}}}\sum_{y\in\mathcal{Y}}\left|\frac{1_{\ell^S(y)=h}}{2^n-1} - \frac{2^{n(1-\alpha)}}{(2^n-1)^2}\right| + \frac{1}{2}\sum_{y\in\mathcal{Y}}\left|\frac{1_{\ell^S(y)=1^{\alpha n}}}{2^n-1} - \frac{2^{n(1-\alpha)}-1}{(2^n-1)^2}\right| .$$

The result follows from further algebraic computation.

If $\alpha = 1$, then the event $\{\mathrm{L} = 1^n\}$ has probability 0, and thus we cannot directly use the above analysis. On the other hand, $\alpha = 1$ describes the case where the secret is completely leaked, that is

$$\beta(\mathrm{Y}|\mathrm{L}) = 1 - \frac{1}{|\mathbb{F}_p|}.$$

Since the above equation corresponds to Eq. (5) for $\alpha = 1$, the lemma follows. □

B Proofs of section 4

The following lemma tells us that the symmetry property of the HW leakage function, coming from the symmetry of the binomial distribution, also translates in the Fourier transform.

Lemma 6 (Symmetry). *Let $n \in \mathbb{N}$ and $p = 2^n - 1$. Then, for all $h \in [\![0, n]\!]$, and for all $\alpha \in \mathbb{Z}_p$,*

$$\left|\widehat{\mathbf{1}_h}(\alpha)\right| = \left|\widehat{\mathbf{1}_{n-h}}(\alpha)\right| \quad . \tag{21}$$

Proof. We start by recalling the identity $\mathsf{HW}(x \oplus y) = \mathsf{HW}(x) + \mathsf{HW}(y) - 2\,\mathsf{HW}(x \wedge y)$. In particular, for $y = 2^n - 1 = 1...1$, and by observing that $1...1 \oplus x = 2^n - 1 - x$, we get that

$$\mathsf{HW}(x) = n - \mathsf{HW}(2^n - 1 - x) \quad .$$

Therefore, applying the change of variable $x \mapsto 2^n - 1 - x$ in Equation 6 gives us that

$$\widehat{\mathbf{1}_{n-h}}(\alpha) = \overline{\widehat{\mathbf{1}_h}(\alpha)} \quad .$$

Finally, taking the absolute value in both side gives us the desired result. □

Proof of Theorem 7. Using the Parseval formula, and the fact that $\widehat{\mathbf{1}_h}(0) = \frac{\binom{n}{h}}{p}$ by definition of a PMF, we get that

$$\left(\frac{\binom{n}{h}}{p}\right)^2 + \sum_{\alpha=1}^{p-1} \left|\widehat{\mathbf{1}_h}(\alpha)\right|^2 = \frac{1}{p} \cdot \sum_{i=0}^{p-1} |\mathbf{1}_h(i)|^2 = \frac{\binom{n}{h}}{p} \quad . \tag{22}$$

We will use the fact that the maximum squared absolute value of the Fourier coefficients (for non-zero harmonic) is upper bounded by the sum in the left hand-side of Equation 22. However, stated as such, Equation 22 would imply trivial upper bounds. Nevertheless, we will show that the sum of the left hand-side contains many identical terms. Thus by factoring the sum, we will tighten the bound.

Let $1 \leq \alpha \leq p-1$. Lemma 5 tells us that we can find at least n other Fourier coefficients sharing the same absolute value as $\left|\widehat{\mathbf{1}_h}(\alpha)\right|$ — they can be derived by cyclically shifting the bits of α. Therefore, we can partition the set $[\![1, p-1]\!]$ of harmonics into classes of harmonics α sharing the same absolute value of Fourier coefficients, each classes containing at least n elements. We shall prove that each class actually contains at least $2 \cdot n$ elements.

Claim. Each class contains at least $2 \cdot n$ elements.

Proof. Since $\mathbf{1}_h$ is real-valued, we have for all $\alpha \neq 0$, $\left|\widehat{\mathbf{1}_h}(\alpha)\right| = \left|\widehat{\mathbf{1}_h}(p - \alpha)\right|$. In other words, for each of the n harmonics that are in the same class, we can derive another harmonic having the same Fourier coefficient in absolute value, hence in the class as well. We shall prove that this *conjugate* harmonic does not coincide with any of the first n harmonics emphasized. To see why, observe that

$$\mathsf{HW}(p-\alpha) = n - \mathsf{HW}(\alpha)$$

(*cf*. proof of Lemma 6). Moreover, by assumption n is odd, so the parity of both hand-sides differ. It implies that the Hamming weight of α is always different from the Hamming weight of $p-\alpha$. As a consequence, none of the harmonics derived by shifting the bits of $p-\alpha$ can never coincide with any of the harmonics derived by shifting the bits of α. Hence, there is at least $2 \cdot n$ elements in each class. □

We can now conclude the proof. Let $\mathcal{F}_\alpha = \left\{ \alpha' \in \mathbb{F}_p : \left| \widehat{\mathbf{1}_h}(\alpha') \right| = \left| \widehat{\mathbf{1}_h}(\alpha) \right| \right\}$. Then for all class $\mathcal{F}_i$, let $\alpha_{\mathcal{F}_\alpha} = \frac{|\mathcal{F}_i|}{2n}$. Observe that $\alpha_{\mathcal{F}_\alpha} \geq 1$ by virtue of the claim.[9] Then we can rephrase Equation 22 as follows:

$$\left(\frac{\binom{n}{h}}{p} \right)^2 + 2n \cdot \sum_{\mathcal{F}_\alpha} \alpha_i \left| \widehat{\mathbf{1}_h}(\alpha) \right|^2 = \frac{\binom{n}{h}}{p} \quad , \tag{23}$$

therefore for any α, we get that

$$\left| \widehat{\mathbf{1}_h}(\alpha) \right|^2 \leq \sum_{\mathcal{F}_\alpha} \alpha_{\mathcal{F}_\alpha} \left| \widehat{\mathbf{1}_h}(\alpha) \right|^2 \leq \left(\frac{\binom{n}{h}}{p} - \left(\frac{\binom{n}{h}}{p} \right)^2 \right) \cdot \frac{1}{2n} \quad .$$

□

Remark 1. We may even prove that for Mersenne primes, the $\alpha_{\mathcal{F}_\alpha}$ coefficients in Equation 23 are integer values. To see why, observe that if p is a Mersenne prime, then n is necessarily an odd prime. It follows from Fermat's little theorem that $2n$ divides $p-1$, *i.e.* the number of terms in Equation 22.

Proof of Theorem 6. We start from Equation 17 from which we derive the following inequality.

$$\sum_{\ell} \max_{\alpha \in \mathbb{F}^\star} \left| \widehat{\mathbf{1}_{\mathsf{HW}^{-1}(\ell)}} \right| \leq \frac{1}{\sqrt{2np}} \sum_{h=0}^{n-1} \sqrt{\binom{n}{h}} \quad .$$

The latter inequality is injected into Equation 10, where we also use the following equality, by definition of the Hamming weight leakage model:

$$\sum_{\ell} \left\| \mathbf{1}_{\mathsf{HW}^{-1}(\ell)} \right\|_2 = \frac{1}{\sqrt{p}} \sum_{h=0}^{n-1} \sqrt{\binom{n}{h}} \quad .$$

[9] We empirically observe that all the α_i are equal to one, however this stronger claim would not seem to improve the upper bound.

It then comes that

$$\mathsf{M}_{\infty}^{\mathrm{Y}}\left(\mathbf{L},\mathrm{Y}\right) \leq \frac{1}{2}\cdot\left(\sum_{h=0}^{n-1}\sqrt{\binom{n}{h}}\right)^{d+1}\cdot\frac{1}{2^{\frac{d+1}{2}-1}}\cdot\frac{1}{p^{\frac{d+1}{2}}}\cdot\frac{1}{n^{\frac{d+1}{2}-1}} \;,$$

hence the inequality in Equation 16. The asymptotic estimation of the right hand-side is then derived from the following claim [6, Sec. 3.1]:

$$\sum_{h=0}^{n}\sqrt{\binom{n}{h}} \sim 2^{\frac{n}{2}+\frac{1}{4}}\cdot(\pi n)^{1/4} \;.$$

□

References

1. Marcin Andrychowicz, Stefan Dziembowski, and Sebastian Faust. Circuit compilers with o(1/\log (n)) leakage rate. In *EUROCRYPT (2)*, volume 9666 of *Lecture Notes in Computer Science*, pages 586–615. Springer, 2016. 2
2. Alberto Battistello, Jean-Sébastien Coron, Emmanuel Prouff, and Rina Zeitoun. Horizontal side-channel attacks and countermeasures on the ISW masking scheme. In *CHES*, volume 9813 of *Lecture Notes in Computer Science*, pages 23–39. Springer, 2016. 2
3. 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 *COSADE*, volume 13979 of *Lecture Notes in Computer Science*, pages 86–104. Springer, 2023. 6
4. Sonia Belaïd, Fabrizio De Santis, Johann Heyszl, Stefan Mangard, Marcel Medwed, Jörn-Marc Schmidt, François-Xavier Standaert, and Stefan Tillich. Towards fresh re-keying with leakage-resilient prfs: cipher design principles and analysis. *J. Cryptogr. Eng.*, 4(3):157–171, 2014. 1
5. Davide Bellizia, Olivier Bronchain, Gaëtan Cassiers, Vincent Grosso, Chun Guo, Charles Momin, Olivier Pereira, Thomas Peters, and François-Xavier Standaert. Mode-level vs. implementation-level physical security in symmetric cryptography - A practical guide through the leakage-resistance jungle. In *CRYPTO (1)*, volume 12170 of *Lecture Notes in Computer Science*, pages 369–400. Springer, 2020. 1
6. Edward A Bender. Asymptotic methods in enumeration. *SIAM review*, 16(4):485–515, 1974. 27
7. Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. In *CRYPTO (1)*, volume 10991 of *Lecture Notes in Computer Science*, pages 531–561. Springer, 2018. 3, 4, 6, 7, 8, 9, 14, 15
8. Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. *J. Cryptol.*, 34(2):10, 2021. 3, 5, 9
9. S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford, 2013. 9
10. Eric Brier, Christophe Clavier, and Francis Olivier. Optimal statistical power analysis. *IACR Cryptol. ePrint Arch.*, page 152, 2003. 21

11. Olivier Bronchain and François-Xavier Standaert. Breaking masked implementations with many shares on 32-bit software platforms or when the security order does not matter. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2021(3):202–234, 2021. 2
12. Gaëtan Cassiers, Loïc Masure, Charles Momin, Thorben Moos, and François-Xavier Standaert. Prime-field masking in hardware and its soundness against low-noise SCA attacks. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2023(2):482–518, 2023. 2, 22
13. Gaëtan Cassiers and François-Xavier Standaert. Towards globally optimized masking: From low randomness to low noise rate or probe isolating multiplications with reduced randomness and security against horizontal attacks. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2019(2):162–198, 2019. 2
14. Gaëtan Cassiers and François-Xavier Standaert. Provably secure hardware masking in the transition- and glitch-robust probing model: Better safe than sorry. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2021(2):136–158, 2021. 2
15. Gaëtan Cassiers and Olivier Bronchain. Scalib: A side-channel analysis library. *Journal of Open Source Software*, 8(86):5196, 2023. 21
16. Suresh Chari, Charanjit S. Jutla, Josyula R. Rao, and Pankaj Rohatgi. Towards sound approaches to counteract power-analysis attacks. In *CRYPTO*, volume 1666 of *Lecture Notes in Computer Science*, pages 398–412. Springer, 1999. 1, 20
17. Christoph Dobraunig, Maria Eichlseder, Stefan Mangard, Florian Mendel, Bart Mennink, Robert Primas, and Thomas Unterluggauer. Isap v2.0. *IACR Trans. Symmetric Cryptol.*, 2020(S1):390–416, 2020. 1
18. Alexandre Duc, Stefan Dziembowski, and Sebastian Faust. Unifying leakage models: From probing attacks to noisy leakage. In *EUROCRYPT*, volume 8441 of *Lecture Notes in Computer Science*, pages 423–440. Springer, 2014. 2, 6
19. Alexandre Duc, Sebastian Faust, and François-Xavier Standaert. Making masking security proofs concrete - or how to evaluate the security of any leaking device. In *EUROCRYPT (1)*, volume 9056 of *Lecture Notes in Computer Science*, pages 401–429. Springer, 2015. 2
20. Stefan Dziembowski, Sebastian Faust, and Maciej Skórski. Optimal amplification of noisy leakages. In *TCC (A2)*, volume 9563 of *Lecture Notes in Computer Science*, pages 291–318. Springer, 2016. 2, 3, 6
21. Sebastian Faust, Vincent Grosso, Santos Merino Del Pozo, Clara Paglialonga, and François-Xavier Standaert. Composable masking schemes in the presence of physical defaults & the robust probing model. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2018(3):89–120, 2018. 2
22. Steven D. Galbraith, Joel Laity, and Barak Shani. Finding significant fourier coefficients: Clarifications, simplifications, applications and limitations, 2018. 15
23. Louis Goubin and Jacques Patarin. DES and differential power analysis (the "duplication" method). In *CHES*, volume 1717 of *Lecture Notes in Computer Science*, pages 158–172. Springer, 1999. 1
24. Vincent Grosso and François-Xavier Standaert. Masking proofs are tight and how to exploit it in security evaluations. In *EUROCRYPT (2)*, volume 10821 of *Lecture Notes in Computer Science*, pages 385–412. Springer, 2018. 2
25. Yuval Ishai, Amit Sahai, and David A. Wagner. Private circuits: Securing hardware against probing attacks. In *CRYPTO*, volume 2729 of *Lecture Notes in Computer Science*, pages 463–481. Springer, 2003. 2
26. Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In *CRYPTO*, volume 1666 of *Lecture Notes in Computer Science*, pages 388–397. Springer, 1999. 1

27. Hemanta K. Maji, Hai H. Nguyen, Anat Paskin-Cherniavsky, Tom Suad, and Mingyuan Wang. Leakage-resilience of the shamir secret-sharing scheme against physical-bit leakages. In *EUROCRYPT (2)*, volume 12697 of *Lecture Notes in Computer Science*, pages 344–374. Springer, 2021. 3
28. Hemanta K. Maji, Hai H. Nguyen, Anat Paskin-Cherniavsky, Tom Suad, Mingyuan Wang, Xiuyu Ye, and Albert Yu. Leakage-resilient linear secret-sharing against arbitrary bounded-size leakage family. In *TCC (1)*, volume 13747 of *Lecture Notes in Computer Science*, pages 355–383. Springer, 2022. 3
29. Hemanta K. Maji, Hai H. Nguyen, Anat Paskin-Cherniavsky, Tom Suad, Mingyuan Wang, Xiuyu Ye, and Albert Yu. Tight estimate of the local leakage resilience of the additive secret-sharing scheme & its consequences. In *ITC*, volume 230 of *LIPIcs*, pages 16:1–16:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. 3, 4, 5, 10, 11, 12, 13, 14, 16
30. Hemanta K. Maji, Hai H. Nguyen, Anat Paskin-Cherniavsky, and Mingyuan Wang. Improved bound on the local leakage-resilience of shamir's secret sharing. In *ISIT*, pages 2678–2683. IEEE, 2022. 3
31. Hemanta K. Maji, Anat Paskin-Cherniavsky, Tom Suad, and Mingyuan Wang. Constructing locally leakage-resilient linear secret-sharing schemes. In *CRYPTO (3)*, volume 12827 of *Lecture Notes in Computer Science*, pages 779–808. Springer, 2021. 3
32. Stefan Mangard, Elisabeth Oswald, and Thomas Popp. *Power analysis attacks - revealing the secrets of smart cards*. Springer, 2007. 16, 21
33. Loïc Masure, Pierrick Méaux, Thorben Moos, and François-Xavier Standaert. Effective and efficient masking with low noise using small-mersenne-prime ciphers. In *EUROCRYPT (4)*, volume 14007 of *Lecture Notes in Computer Science*, pages 596–627. Springer, 2023. 2, 3, 7, 9, 19, 21, 23
34. Loïc Masure and François-Xavier Standaert. Prouff and Rivain's formal security proof of masking, revisited - tight bounds in the noisy leakage model. In *CRYPTO (3)*, volume 14083 of *Lecture Notes in Computer Science*, pages 343–376. Springer, 2023. 2, 6
35. Thorben Moos. Static power SCA of sub-100 nm CMOS asics and the insecurity of masking schemes in low-noise environments. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2019(3):202–232, 2019. 2
36. Thorben Moos, Amir Moradi, Tobias Schneider, and François-Xavier Standaert. Glitch-resistant masking revisited or why proofs in the robust probing model are needed. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2019(2):256–292, 2019. 2
37. Amir Moradi and François-Xavier Standaert. Moments-correlating DPA. In *TIS@CCS*, pages 5–15. ACM, 2016. 2
38. Svetla Nikova, Vincent Rijmen, and Martin Schläffer. Secure hardware implementation of nonlinear functions in the presence of glitches. *J. Cryptol.*, 24(2):292–321, 2011. 2
39. Thomas Prest, Dahmun Goudarzi, Ange Martinelli, and Alain Passelègue. Unifying leakage models on a rényi day. In *CRYPTO (1)*, volume 11692 of *Lecture Notes in Computer Science*, pages 683–712. Springer, 2019. 2
40. Emmanuel Prouff and Matthieu Rivain. Masking against side-channel attacks: A formal security proof. In *EUROCRYPT*, volume 7881 of *Lecture Notes in Computer Science*, pages 142–159. Springer, 2013. 2, 6, 9
41. Emmanuel Prouff, Matthieu Rivain, and Régis Bevan. Statistical analysis of second order differential power analysis. *IEEE Trans. Computers*, 58(6):799–811, 2009. 2

42. Werner Schindler, Kerstin Lemke, and Christof Paar. A stochastic model for differential side channel cryptanalysis. In Josyula R. Rao and Berk Sunar, editors, *Cryptographic Hardware and Embedded Systems - CHES 2005, 7th International Workshop, Edinburgh, UK, August 29 - September 1, 2005, Proceedings*, volume 3659 of *Lecture Notes in Computer Science*, pages 30–46. Springer, 2005. 21, 23
43. François-Xavier Standaert. How (not) to use welch's t-test in side-channel security evaluations. In *CARDIS*, volume 11389 of *Lecture Notes in Computer Science*, pages 65–79. Springer, 2018. 2
44. Karl Stromberg. Probabilities on a compact group. *Transactions of the American Mathematical Society*, 94(2):295–309, 1960. 2
45. Nicolas Veyrat-Charvillon and François-Xavier Standaert. Adaptive chosen-message side-channel attacks. In *ACNS*, volume 6123 of *Lecture Notes in Computer Science*, pages 186–199, 2010. 3