A Decomposition Approach for Evaluating Security of Masking

Vahid Jahandideh, Bart Mennink, and Lejla Batina

Radboud University, Nijmegen, The Netherlands
{v.jahandideh, b.mennink, lejla}@cs.ru.nl

Abstract. Masking is a common countermeasure against side-channel attacks that encodes secrets into multiple shares, each of which may be subject to leakage. A key question is under what leakage conditions, and to what extent, does increasing the number of shares actually improve the security of these secrets. Although this question has been studied extensively in low-SNR regimes, scenarios where the adversary obtains substantial information—such as on low-noise processors or through static power analysis—have remained underexplored.

In this paper, we address this gap by deriving *necessary and sufficient* noise requirements for masking security in both standalone encodings and linear gadgets. We introduce a decomposition technique that reduces the relationship between an extended-field variable and its leakage into subproblems involving linear combinations of the variable's bits. By working within binary subfields, we derive optimal bounds and then lift these results back to the extended field.

Beyond binary fields, we also present a broader framework for analyzing masking security in other structures, including prime fields. As an application, we prove a conjecture by Dziembowski et al. (TCC 2016), which states that for an additive group $\mathbb{G}$ with its largest subgroup $\mathbb{H}$, a δ-noisy leakage satisfying $\delta < 1 - \frac{|\mathbb{H}|}{|\mathbb{G}|}$ ensures that masking enhances the security of the secret.

1 Introduction

Masking to Mitigate Side-Channel Threats. Side-channel information refers to unintended leakages that an adversary can obtain from the implementation of a cryptographic algorithm. A *leakage model* provides an abstraction for describing such leakages. One widely studied model is the *noisy leakage model*, introduced by Prouff and Rivain [28] and subsequently explored in several works [9, 10, 12, 13, 27]. In this model, for each intermediate value $X \in \mathbb{F}_q$ in the execution of a cryptographic algorithm, the adversary learns a function $\mathsf{L}(X)$, such as the noisy Hamming weight of X.

A primary countermeasure against side-channel leaks is *masking*. In this approach, a secret X is split into random *shares* $X_1, \ldots, X_n \in \mathbb{F}_q$ such that $X = X_1 + \cdots + X_n$. Rather than processing X directly, the implementation manipulates these shares, and the adversary only observes leakages $\mathsf{L}(X_i)$ from each

share. In the most basic encoding, the only intermediate values are the shares themselves, but in protected circuits, additional intermediates and secrets may also be introduced. The effectiveness of masking is typically assessed by how a *security metric*—for instance, the adversary's success rate—degrades as the sharing order n increases. Since the seminal work of Chari et al. [7], this line of investigation has remained a central focus for both standalone encodings and protected gadgets and circuits.

Open Challenge. If $\mathsf{L}(X)$ fully reveals X, then masking offers no protection. Thus, $\mathsf{L}(X)$ must introduce some form of *noise*. Determining the minimal noise level required to make masking effective in protecting secrets, and understanding how a chosen security metric scales with the sharing order n under borderline leakage conditions, remain open problems. This paper addresses these challenges.

Practical Relevance. Low-noise (high-SNR) conditions arise when $\mathsf{L}(X)$ reveals a substantial amount of information about X. Such scenarios have been reported in various contexts. For instance, *low-noise processors*—particularly small embedded devices such as the ARM Cortex-M0—exhibit inherently lower noise levels in their power consumption [6]. Likewise, *static power analysis*, unlike dynamic power analysis, measures a stable leakage signal over an extended period, resulting in highly precise side-channel observations [25]. Lastly, *averaging* or *horizontal attacks* can combine multiple leakage samples corresponding to the same or related intermediates to produce a clearer, aggregated leakage trace [3].

Security Metrics. A widely used security metric is the *success rate* (SR), which measures the probability that the leakage $\mathsf{L}(X)$ correctly identifies X [30]. Another common metric is the *guessing entropy* (GE), defined as the average rank of the correct X in the adversary's list of hypotheses [30]. A more technical measure is the *statistical distance* (SD) between the prior distribution of X and its posterior distribution $X \mid \mathsf{L}(X)$. We denote this distance by δ; a larger value of δ indicates less noise. Additionally, a more mathematically rigorous metric is the *mutual information* (MI) between X and $\mathsf{L}(X)$, a measure well-studied in information theory and used in various papers [4,10,20]. The quantity $\mathsf{MI}(X\,;\,\mathsf{L}(X))$ lies in $[0, \log(q)]$, where lower values correspond to higher noise. Security metrics are not mutually independent. For instance, MI and SD are linked through Pinsker's inequality [10,15], and SR also has a relationship with SD (see Lemma 1). However, these connections are not tight in general.

In this work, we focus on the success rate metric, because it is more intuitive and directly indicates how many leakage traces are required for a successful attack. In typical divide-and-conquer scenarios where a secret is split into multiple chunks, checking multiple guesses for each chunk is impractical; the adversary usually needs to identify each chunk correctly in a single guess. This practical consideration further motivates our choice of the SR metric.

1.1 Evaluating the Security of Single Encoding

Building on the reduction proposed in [9], Duc et al. [10] proved that $q\delta < 1$ is sufficient to ensure the effectiveness of masking. Improving this bound, Dziembowski et al. [13] derived an optimal threshold $\delta < \frac{1}{2}$ for binary extended fields ($q = 2^u$). Their result states that if, for a given leakage function $\mathsf{L}(X)$, the corresponding δ is less than one-half, masking will be effective. However, the case $\delta \geq \frac{1}{2}$ is not covered by their analysis.

Seeking more concrete guidelines and using mutual information (MI) as the security metric, Ito et al. [20] proposed a threshold of $\mathsf{MI}(X\,;\,\mathsf{L}(X)) < 0.72$ for all shares. Béguinot et al. [4] relaxed the requirement to $\mathsf{MI}(X\,;\,\mathsf{L}(X)) \leq 1$ for at least some of shares, under the assumption that different shares may have distinct leakage functions. Despite these contributions, the question of whether masking is effective for a given leakage $\mathsf{L}(X)$ remains only partially resolved.

Our Contribution to the Problem. We address this gap by relaxing the noise requirements and showing that masking can improve security *if and only if* $\mathsf{L}(X)$ does not completely determine any bit combination of X. Specifically, for a u-bit X, the leakage $\mathsf{L}(X)$ must not fully reveal the binary bitwise inner product

$$\langle X, h\rangle \;=\; \bigoplus_{j=0}^{u-1} x_j\, h_j$$

for any $h \in [1, 2^u - 1]$, where x_j and h_j are the individual bits of X and h, respectively. In terms of mutual information, this requirement is

$$\mathsf{MI}\big(\langle X, h\rangle\;;\;\mathsf{L}(X)\big) \;<\; 1.$$

We first prove a tight security bound for binary fields. Building on this result, we analyze the security of various binary variables $\langle X, h\rangle$ and derive a new security bound for binary extension fields. This allows us to accurately compute the adversary's success probability, thereby relaxing the noise constraints previously imposed on $\mathsf{L}(X)$.

Moreover, we introduce a general framework for quantifying masking security (via the SR metric) in other algebraic structures, including prime fields and additive groups. Employing this framework, we confirm a conjecture by Dziembowski et al. [13]: for an additive group $\mathbb{G}$ with its largest subgroup $\mathbb{H}$, if a given leakage satisfies

$$\delta \;<\; 1 \;-\; \frac{|\mathbb{H}|}{|\mathbb{G}|},$$

then masking can indeed enhance the security of X.

1.2 Evaluating the Security of Protected Circuits

The security requirements for an encoded secret do not immediately translate into those for a protected circuit. Prouff and Rivain [28] pioneered the study of

noisy leakage in masked circuits; this work was later extended by Masure and Standaert [24]. However, analyzing noisy leakages in complete protected circuits remains challenging, often requiring unrealistic assumptions such as leak-free refresh gadgets.

Duc et al. [9] tackled this issue by introducing a *reduction* from noisy leakage to the *random probing* model. This reduction allows security proofs in the random probing model to carry over to the noisy leakage setting. While subsequent works [12,26,27] have refined the approach, gaps remain, limiting the reduction's applicability in certain scenarios.

Our Contributions for Security of Circuits. We demonstrate that our decomposition strategy can yield improved reductions from noisy to random probing under certain assumptions, specifically for linear protected circuits. This focus is motivated by recent results of Jahandideh et al. [21], who showed that linear circuits can provide side-channel security bounds even in settings involving some non-linear gadgets.

1.3 Outline

Section 2 introduces security metrics and relations among them for single variables. Section 3 focuses on masking, presenting our decomposition approach and a general framework for additive groups. Section 4 extends the discussion to the broader context of linear circuits.

2 Security of a Single Variable

In this section, we formalize the noisy leakage model and introduce the adversary's *advantage*, which is a normalized version of the success rate. In Lemma 1, we establish a connection between two key metrics, δ and this advantage. Subsection 2.2 presents a reduction from the noisy leakage model to the random probing model, and Lemma 2 demonstrates that this reduction is tight in the binary-field setting. Finally, Lemma 3 applies the reduction to derive a bound on the adversary's advantage.

2.1 Preliminaries

Let X be a random variable uniformly distributed over $\mathbb{F}_q$, representing an intermediate value in a cryptographic implementation. This implementation might leak side-channel information modeled by a probabilistic function $\mathsf{L}(X) \in \mathbb{R}^m$. Given $\mathsf{L}(X)$, the adversary's goal is to determine the actual value of X. If the joint distribution of $(X, \mathsf{L}(X))$ is known, the optimal strategy is *maximum a posteriori* (MAP) estimation, also called *Bayesian estimation* [19].

Upon observing $l \leftarrow \mathsf{L}(X)$, the MAP estimator outputs:

$$\hat{X} \xleftarrow{\$} \Big\{ \underset{\alpha \in \mathbb{F}_q}{\operatorname{argmax}} \Pr\big(X = \alpha \mid l\big) \Big\}.$$

The probability that $\hat{X} = X$ depends on the particular realization l. For example, if $\mathsf{L}(X)$ is the Hamming weight $\mathsf{HW}(X) = \sum x_i$, where x_i are the bits of X, then upon observing $l = 0$, the adversary can correctly deduce that $X = 0$.

To account for different leakage realizations, we define P_c as the expected success probability, averaged over all possible outputs of $\mathsf{L}(X)$:

$$P_c \triangleq \mathbb{E}_{l \leftarrow \mathsf{L}(X)} \Big[\Pr\big(\hat{X} = X \mid l\big) \Big] = \sum_{l \in \mathbb{R}^m} \Pr\big(\mathsf{L}(X) = l\big) \Pr\big(\hat{X} = X \mid l\big).$$

A more practical metric for the adversary's effectiveness is the *advantage* over random guessing:

$$\mathsf{Adv}_X \triangleq P_c - \frac{1}{q}.$$

The advantage Adv_X indicates how informative $\mathsf{L}(X)$ is about a single random variable X. However, masking schemes involve multiple variables and their joint distributions. In the remainder of this section, we briefly review other metrics that quantify the information content of $\mathsf{L}(X)$.

Class of δ-Noisy Leakages [9]. The *statistical distance* (SD) between X and $X \mid \mathsf{L}(X)$ is a measure of the informativeness of the leakage. It quantifies how much the distribution of X changes when the leakage $\mathsf{L}(X)$ is observed. Concept of this metric was first introduced in the work of Prouff and Rivain [28], and it aligns well with practical side-channel analysis experiences.

The SD between X and $X \mid \mathsf{L}(X)$ is an *expected value*, defined over all possible leakage values as:

$$\mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) \triangleq \sum_{l \in \mathbb{R}^m} \Pr\left(\mathsf{L}(X) = l\right) \mathsf{TV}(X\,;\, X \mid l), \tag{1}$$

where the *total variation distance* (TV) between the (uniformly distributed) X and $X \mid l$, for a fixed leakage instance l, is defined as follows:

$$\begin{aligned} \mathsf{TV}\left(X\,;\, X \mid l\right) &\triangleq \frac{1}{2} \sum_{\alpha \in \mathbb{F}_q} \left| \Pr(X = \alpha \mid l) - \frac{1}{q} \right| \\ &= \sum_{\alpha \in \mathbb{F}_q, \Pr(X=\alpha \mid l) > \frac{1}{q}} \left(\Pr(X = \alpha \mid l) - \frac{1}{q} \right) \\ &= \sum_{\alpha \in \mathbb{F}_q, \Pr(X=\alpha \mid l) < \frac{1}{q}} - \left(\Pr(X = \alpha \mid l) - \frac{1}{q} \right). \end{aligned} \tag{2}$$

We say that the leakage $\mathsf{L}(X)$ is δ-noisy if $\mathsf{SD}\left(X; X \mid \mathsf{L}(X)\right) = \delta$. From (2), we deduce that the range of TV is $0 \leq \mathsf{TV}\left(X; X \mid l\right) \leq 1 - \frac{1}{q}$. Consequently, by taking the expectation over $l \leftarrow \mathsf{L}(X)$, we conclude that δ lies within the range $[0, 1 - \frac{1}{q}]$.

Relation Between δ and Adv_X. A lower δ indicates that the leakage $\mathsf{L}(X)$ is noisier, meaning the adversary gains less advantage from it. More concretely, we have the following lemma.

Lemma 1. *For a random variable X over $\mathbb{F}_q$ with leakage function $\mathsf{L}(X)$, if $\mathsf{SD}\left(X\,;\, X \mid \mathsf{L}(X)\right) = \delta$, then the adversary's advantage in guessing the value of X from learning $\mathsf{L}(X)$ is bounded as:*

$$\frac{\delta}{q-1} \leq \mathsf{Adv}_X \leq \delta. \tag{3}$$

As a corollary, in the special case of binary fields (i.e., $q = 2$), we have $\mathsf{Adv}_X = \delta$.

Proof. For each instance $l \leftarrow \mathsf{L}(X)$, we have:

$$\begin{aligned}\Pr\left((\hat{X} = X) \mid l\right) - \frac{1}{q} &= \max_{\alpha \in \mathbb{F}_q} \Pr(X = \alpha \mid l) - \frac{1}{q} \\ &\leq \sum_{\alpha \in \mathbb{F}_q, \Pr(X=\alpha|l) > \frac{1}{q}} \left(\Pr(X = \alpha \mid l) - \frac{1}{q}\right) = \mathsf{TV}(X\,;\, X \mid l).\end{aligned}$$

Taking the expectation over $l \leftarrow \mathsf{L}(X)$ on both sides of the inequality proves the right-hand side of the lemma. The proof of the left-hand side follows similarly, noting that the cardinality of the set $\left\{\alpha \in \mathbb{F}_q \mid \Pr(X = \alpha \mid l) > \frac{1}{q}\right\}$ is at most $q - 1$. □

2.2 Leakage Simulation

The probabilistic mapping $\phi^\epsilon \colon \mathbb{F}_q \to \{\perp, \mathbb{F}_q\}$ is known as an *erasure channel* [8, 18], and it is computed as follows:

$$\phi^\epsilon(X) = \begin{cases} X & \text{with probability } \epsilon, \\ \perp & \text{otherwise.} \end{cases}$$

Duc et al. [9] demonstrated that for sufficiently noisy leakage $\mathsf{L}(X)$, one can construct a (probabilistic) function L' such that for any value of X, $\mathsf{L}'(\phi^\epsilon(X))$ and $\mathsf{L}(X)$ are statistically identical, i.e.,

$$\forall \alpha \in \mathbb{F}_q, \quad \mathsf{TV}\left(\mathsf{L}'(\phi^\epsilon(X)) \mid X = \alpha\,;\, \mathsf{L}(X) \mid X = \alpha\right) = 0.$$

This holds only when $\epsilon \geq \epsilon_{\min}$, where:

$$\epsilon_{\min} \triangleq 1 - \sum_l \min_{\alpha \in \mathbb{F}_q} \Pr(l \mid X = \alpha) \leq_{(\mathtt{I})} q\delta. \tag{4}$$

The summation is over all possible leakage values. This result is particularly useful for deriving security bounds, especially when considering masked encoding and circuits. The value of $\epsilon_{\min}$, also known as *Doeblin coefficient* [5], lies in

$[0,1]$, where $\epsilon_{\min} = 0$ indicates that $\mathsf{L}(X)$ and X are independent, and $\epsilon_{\min} = 1$ indicates that L' always depends on X, rendering the technique ineffective.

The right-hand side of inequality (I) in (4) was proved in [9]. Here, we prove that for the specific case of $q = 2$, the inequality is in fact an equality.

Lemma 2. *For a joint distribution $(X, \mathsf{L}(X))$ where $\mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) = \delta$, let $\epsilon_{\min}$ be defined as in (4). If X is a uniform binary random variable, then $\epsilon_{\min} = 2\delta$.*

Proof.

$$\begin{aligned}
\epsilon_{\min} &= 1 - \sum_l \min_{\alpha\in\{0,1\}} \Pr(l \mid X = \alpha) \\
&= \sum_l \Pr(\mathsf{L}(X) = l) - \sum_l \min_{\alpha\in\{0,1\}} \Pr(l \mid X = \alpha) \\
&=_{(\mathtt{I})} \sum_l \Pr(\mathsf{L}(X) = l) - 2\sum_l \Pr(\mathsf{L}(X) = l) \min_{\alpha\in\{0,1\}} \Pr(X = \alpha \mid l) \\
&= \sum_l \Pr(\mathsf{L}(X) = l) \left[1 - 2 \min_{\alpha\in\{0,1\}} \Pr(X = \alpha \mid l)\right] \\
&= \sum_l \Pr(\mathsf{L}(X) = l) \left[\max_{\alpha\in\{0,1\}} \Pr(X = \alpha \mid l) - \min_{\alpha\in\{0,1\}} \Pr(X = \alpha \mid l)\right] \\
&= \sum_l \Pr(\mathsf{L}(X) = l) \left[\left|\Pr(X = 1 \mid l) - \frac{1}{2}\right| + \left|\Pr(X = 0 \mid l) - \frac{1}{2}\right|\right] \\
&= 2\mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) = 2\delta.
\end{aligned}$$

The second summation in (I) follows from applying Bayes' rule. □

In the binary case, we have $\delta \leq \frac{1}{2}$, and Lemma 2 gives $\epsilon_{\min} = 2\delta$, which leads to $\epsilon_{\min} < 1$ for $\delta \neq \frac{1}{2}$. The case $\delta = \frac{1}{2}$ only occurs when $\mathsf{L}(X)$ uniquely identifies X. We conclude that, if there remains some uncertainty about X after observing the leakage, it will be reflected in the $\epsilon_{\min}$ metric, and this tight phenomenon only occurs when $q = 2$.

Example 1. Let $X \in \mathbb{F}_{2^u}$, and let the leakage function $\mathsf{L}(X)$ be defined as $x_0 \oplus e$, where x_0 is the least significant bit (LSB) of X, and $\Pr(e = 1) = \mathsf{e}$, with $\mathsf{e} \leq \frac{1}{2}$. The leakage $\mathsf{L}(X)$ provides noisy information about the LSB of X, while revealing no information about the remaining bits, from bit 1 to bit $u-1$. Upon receiving leakage l, the posterior distribution of X becomes:

$$\begin{aligned}
&\forall\ \alpha \in \{0,1\}^{u-1} || l, \quad \Pr(X = \alpha \mid l) = \frac{1-\mathsf{e}}{2^{u-1}}, \\
&\forall\ \alpha \in \{0,1\}^{u-1} || (1 \oplus l), \quad \Pr(X = \alpha \mid l) = \frac{\mathsf{e}}{2^{u-1}}.
\end{aligned}$$

This leads to the following results:

$$\mathsf{Adv}_X = \frac{1-2\mathsf{e}}{2^u}, \quad \mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) = \frac{1}{2} - \mathsf{e}, \quad \text{and} \quad \epsilon_{\min} = 1 - 2\mathsf{e}.$$

Notably, as u increases, the gap for the upper bounds provided in equations (3) and (4) widens, while the lower bound $\frac{\frac{1}{2}-\mathsf{e}}{2^u-1} \leq \mathsf{Adv}_X$ in equation (3) becomes tighter. □

A Security Reduction. The joint distributions $(X, \mathsf{L}'(\phi^\epsilon(X)))$ and $(X, \mathsf{L}(X))$ are identical. Therefore, an adversary cannot distinguish between samples drawn from these distributions. This implies that the adversary's advantage, denoted by Adv_X, when leakage l is sampled from $\mathsf{L}(X)$, is equal to the advantage when l is sampled from $\mathsf{L}'(\phi^\epsilon(X))$. Otherwise, the adversary could distinguish between the two leakage functions. We express this equality as:

$$\mathsf{Adv}_X\left[l \leftarrow \mathsf{L}(X)\right] = \mathsf{Adv}_X\left[l \leftarrow \mathsf{L}'(\phi^\epsilon(X))\right].$$

We will write $\mathsf{Adv}_X[.]$ to specify the leakage source and avoid ambiguity.

The random variables (RVs) X, $\phi^\epsilon(X)$, $\mathsf{L}'(\phi^{\epsilon_{\min}}(X))$, and $\mathsf{L}(X)$ form a *Markov chain*:

$$X \to \phi^\epsilon(X) \to \phi^{\epsilon_{\min}}(X) \to \mathsf{L}'(\phi^{\epsilon_{\min}}(X)) \to \mathsf{L}(X),$$

where $\epsilon \geq \epsilon_{\min}$. As we move along the direction of the chain, we receive increasingly *degraded* versions of X. Consequently, the success probability of computing any metric related to X decreases as the available information is taken from links farther from X. For instance:

$$\mathsf{Adv}_X\left[l \leftarrow \mathsf{L}(X)\right] \leq \mathsf{Adv}_X\left[l \leftarrow \phi^\epsilon(X)\right]. \tag{5}$$

Similarly:

$$\mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) \leq \mathsf{SD}(X\,;\, X \mid \phi^\epsilon(X)). \tag{6}$$

In general, to prove security with leakage $\mathsf{L}(X)$, it suffices to prove it with leakage $\phi^{\epsilon_{\min}}(X)$. This result from Duc et al. [9] is known as the *reduction* of δ-noisy leakage to ϵ-random probing leakage. The following lemma demonstrates an application of this reduction.

Lemma 3 ([5] Proposition 1). *For a random variable $X \in \mathbb{F}_q$, let $\mathsf{L}(X)$ be a leakage with Doeblin coefficient $\epsilon_{\min}$. The adversary's advantage when observing this leakage is bounded by:*

$$\mathsf{Adv}_X \leq \frac{q-1}{q}\epsilon_{\min}. \tag{7}$$

This result is a simplified version of Proposition 1 from the work of Béguinot et al. [5].

$\epsilon_{\min}$ Metric is not a Tight Indicator. The reduction from noisy leakage to random probing leakage is highly useful. However, it is not always tight. For some leakage models, such as when $\mathsf{L}(X) = \mathsf{HW}(X)$ (Hamming weight), we compute $\epsilon_{\min} = 1$. Intuitively, this occurs because, for any leakage value l, there are some values of X that cannot produce that l. From the perspective of the reduction,

this implies there is no difference between $\mathsf{L}(X) = \mathsf{HW}(X)$ and $\mathsf{L}(X) = X$, even though the Hamming weight function does not fully reveal X.

It may seem that this result occurs because $\mathsf{HW}(X)$ is too informative. To illustrate this point, we define a leakage function, $\mathsf{ZV}(X)$, where the amount of information conveyed in the leakage is less than one bit, yet we still have $\epsilon_{\min} = 1$.

Inspired by the zero-value leakage model [22], $\mathsf{ZV}(X)$ only distinguishes when $X = 0$ and returns a constant value for all other inputs. Specifically, for some real values ν_a and ν_b, we define:

$$\mathsf{ZV}(X) = \begin{cases} \nu_a & \text{if } X = 0, \\ \nu_b \neq \nu_a & \text{otherwise.} \end{cases} \tag{8}$$

Our refined reduction approach, introduced in the next section, will be able to distinguish the noise embedded in leakage functions like $\mathsf{ZV}(X)$.

3 Security of Mask Encoding

We begin by introducing the basics of masking and defining the adversary's advantage in this setting. In Lemma 4, we use the leakage parameter δ to bound the statistical distance associated with mask encoding. Next, Lemma 5 shows that security metrics become simpler to compute in binary fields, laying the groundwork for our main contribution in Subsection 3.2, which is formalized in Theorem 1. Lemma 6 further refines the result of that theorem. Finally, we conclude with an examination of prime fields and additive groups in Subsection 3.3.

3.1 Mask Encoding

A standard technique for improving the side-channel security of a variable X is *masking*. Under this approach, X is encoded via a random n-tuple of *shares*, $\boldsymbol{X} = (X_1, \ldots, X_n)$, such that

$$X \;=\; \sum_{i=1}^{n} X_i.$$

In this context, the side-channel adversary observes the leakage vector

$$\boldsymbol{L}(\boldsymbol{X}) \;=\; \big[\mathsf{L}_1(X_1), \ldots, \mathsf{L}_n(X_n)\big].$$

For simplicity, we assume that each leakage function L_i is identical (i.e., $\mathsf{L}_i = \mathsf{L}$) and that all are independent in their internal randomness. A central question is how $\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})]$ is influenced by both the number of shares n and the structure of the field $\mathbb{F}_q$.

Impracticality of Exact Computation. Because the distribution space of $\boldsymbol{L}(\boldsymbol{X})$ grows exponentially with n, exact calculations of security metrics quickly become infeasible. An alternative is to approximate the metrics of $\boldsymbol{L}(\boldsymbol{X})$ using known properties of $\mathsf{L}(X)$. While this strategy circumvents exponential complexity, it also introduces a potential gap between the estimated and the exact metrics.

A Not-So-Tight Bound. Duc et al. [10] investigated the practical relevance of the noisy-to-random probing reduction, which assumes that each share is revealed to the adversary with probability $\epsilon_{\min} \leq q\delta$. Under this assumption, the adversary learns all shares with probability $(q\delta)^n$. Denote the corresponding leakage vector by

$$\boldsymbol{\phi^{\epsilon}}(\boldsymbol{X}) \;=\; \big[\phi^{\epsilon}(X_1),\, \ldots,\, \phi^{\epsilon}(X_n)\big].$$

Generalizing from (6), we have

$$\Delta \;=\; \mathsf{SD}\big(X\,;\, X \;\mid\; \boldsymbol{L}(\boldsymbol{X})\big) \;\leq\; \mathsf{SD}\big(X\,;\, X \;\mid\; \boldsymbol{\phi^{\epsilon}}(\boldsymbol{X})\big). \tag{9}$$

If any one share does not leak, then $\mathsf{SD}\big(X\,;\, X \mid \boldsymbol{\phi^{\epsilon}}(\boldsymbol{X})\big)$ is zero. Conversely, if all shares leak, it equals $1 - \frac{1}{q}$. Combining these observations yields

$$\Delta \;\leq\; \mathsf{SD}\big(X\,;\, X \;\mid\; \boldsymbol{\phi^{\epsilon}}(\boldsymbol{X})\big) \;=\; \Big(1 - \tfrac{1}{q}\Big)\, q^n\, \delta^n. \tag{10}$$

A Tighter Bound. The bound in (10) grows exponentially with the field size q. However, Duc et al. [10] noted that experimental evidence does not exhibit such a factor, leading them to conjecture that dependence on q might be a proof artifact. A subsequent result by Masure et al. [23] removed the factor of q from (10):

Lemma 4 ([23], Proposition 4). *Let $X = (X_1, \ldots, X_n)$ be a masking of $X \in \mathbb{F}_q$, and assume the leakage function satisfies $\mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) = \delta$. Then, for $\Delta = \mathsf{SD}\big(X\,;\, X \mid \boldsymbol{L}(\boldsymbol{X})\big)$, we have*

$$\Delta \;\leq\; 2^{\,n-1}\, \delta^n. \tag{11}$$

Dziembowski et al. [13] showed—reaffirmed here—that as long as $\delta < \frac{1}{2}$, the posterior distribution of X (after observing the leakages) becomes increasingly uniform as n grows, independent of the underlying field structure. Consequently, the highest point in the posterior distribution converges to the uniform value $\frac{1}{q}$. Applying Lemma 1 to this scenario leads to

$$\mathsf{Adv}_X\big[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})\big] \;\leq\; 2^{\,n-1}\, \delta^n. \tag{12}$$

Case of $q = 2$. For binary fields, we show via direct computation that the inequalities in (11) and (12) become equalities. In fact, it suffices to verify this for (12), as the other inequality follows as a corollary of Lemma 1, which states that $\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})] = \Delta$.

Lemma 5. *In the setting of Lemma 4, if the underlying field is $\mathbb{F}_2$, we have* $\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})] = 2^{n-1}\delta^n$.

Proof. We prove this lemma by extending a technique originally due to Wyner [32] in the context of *wire-tap channels*.

To recover $X = X_1 \oplus \cdots \oplus X_n$ from the leakage vector $\boldsymbol{L}(\boldsymbol{X})$, the adversary must estimate each X_i. Let $\hat{X}_i$ represent the estimate of X_i after observing $\mathsf{L}(X_i)$, and let $\mathsf{e}_i = \Pr(\hat{X}_i \neq X_i)$ be the average probability of error. The averaging is done over both the uniform choice of $X_i \in \{0,1\}$ and the leakage space of $\mathsf{L}(X_i)$. That is,

$$\mathsf{e_i} = \mathop{\mathbb{E}}_{\substack{X_i \xleftarrow{\$} \{0,1\} \\ l \leftarrow \mathsf{L}(X_i)}} \left[\Pr\left((\hat{X}_i \neq X_i) \mid l\right)\right].$$

We assume $\mathsf{e}_i \leq \frac{1}{2}$; the case $\mathsf{e}_i > \frac{1}{2}$ can be handled similarly. By subtracting, the random guessing factor $\frac{1}{2}$, we express the adversary's advantage as $\mathsf{Adv}_{X_i} = (1 - \mathsf{e}_i) - \frac{1}{2} = \frac{1}{2} - \mathsf{e}_i$. According to Lemma 1, for a single random variable X_i, we have $\mathsf{Adv}_{X_i}[l \leftarrow \mathsf{L}(X_i)] = \delta$, which implies $\mathsf{e}_i = \frac{1}{2} - \delta$. We can simplify by dropping the subscript i and setting $\mathsf{e} = \frac{1}{2} - \delta$.

The adversary can estimate X from the values $\hat{X}_1, \ldots, \hat{X}_n$ using $\hat{X} = \hat{X}_1 \oplus \cdots \oplus \hat{X}_n$. The estimate $\hat{X}$ will match X if an even number of errors occur. Thus, the average probability of success is:

$$\begin{aligned}
\Pr(\hat{X} = X) &= \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2j} \mathsf{e}^{2j} (1-\mathsf{e})^{n-2j} \\
&= \frac{1}{2}\left[\sum_{i=0}^{n} \binom{n}{i} \mathsf{e}^i (1-\mathsf{e})^{n-i} + \sum_{i=0}^{n} \binom{n}{i} (-\mathsf{e})^i (1-\mathsf{e})^{n-i}\right] \\
&=_{(\mathtt{I})} \frac{1}{2}[(\mathsf{e} + (1-\mathsf{e}))^n + (-\mathsf{e} + (1-\mathsf{e}))^n] \\
&= \frac{1}{2}[1^n + (1-2\mathsf{e})^n] = \frac{1}{2} + 2^{n-1}(\frac{1}{2} - \mathsf{e})^n = \frac{1}{2} + 2^{n-1}\delta^n,
\end{aligned}$$

where step (I) follows from the binomial expansion:

$$(\pm\mathsf{e} + (1-\mathsf{e}))^n = \sum_{i=0}^{n} \binom{n}{i} (\pm\mathsf{e})^i (1-\mathsf{e})^{n-i}.$$

Subtracting the random guessing contribution $\frac{1}{2}$, we obtain:

$$\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})] = \Pr(\hat{X} = X) - \frac{1}{2} = 2^{n-1}\delta^n. \qquad \square$$

The binary case will serve as the foundation for our reasoning when working with extended fields. Before proceeding, we pause to make another useful observation.

Optimality of the Reduction at $q = 2$. Lemma 2 proves that the δ-noisy to ϵ-random reduction for $q = 2$ is possible with $\epsilon_{\min} = 2\delta$. Substituting this into (10), we get:

$$\Delta \leq \mathsf{SD}(X\,;\, X \mid \boldsymbol{\phi}^{\boldsymbol{\epsilon_{\min}}}(\boldsymbol{X})) = (1 - \frac{1}{2})2^n\delta^n = 2^{n-1}\delta^n$$

On the other hand, from Lemma 5, we know that $\Delta = 2^{n-1}\delta^n$, which implies that $\Delta = \mathsf{SD}(X\,;\, X \mid \boldsymbol{\phi}^{\boldsymbol{\epsilon_{\min}}}(\boldsymbol{X}))$. This equality indicates the tightness of the reduction in the broader context of multiple random variables (RVs), at least for the metrics Δ and Adv_X.[1] If the leakages functions for the shares are different, the corresponding δ_i will also vary among them. In this case, we can show that $\Delta = 2^{n-1}\prod_{i=1}^{n}\delta_i$.

For $q > 2$, the Bound in Lemma 3 is Loose. We work out an example at $q = 4$, for which we are able to determine exact metrics at order n. The result illustrates that the bound is not necessarily tight for non-binary fields.

Example 2. Suppose for $X \in \mathbb{F}_{2^2}$, the leakage function is $\mathsf{L}(X) = (x_1 \oplus e_1)\|(x_0 \oplus e_0)$, where x_i denotes the ith bit of X, and e_0 and e_1 are independent binary RVs with $\Pr(e_0 = 1) = \Pr(e_1 = 1) = \mathsf{e}$ for some $\mathsf{e} < \frac{1}{2}$. For this leakage function, we compute:

$$\delta = \mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) = (\frac{1}{2} - \mathsf{e})(\frac{3}{2} - \mathsf{e}). \tag{13}$$

The leakage function is essentially a concatenation of two independent binary leakages, each corresponding to one bit of X.

In a masked encoding, using a method similar to the proof of Lemma 5, we can show that the adversary estimates each bit of X with an error probability of $\mathsf{e}_n = \frac{1}{2}\left[1 - (1 - 2\mathsf{e})^n\right]$ from the leakage information. Using this, we can replace the leakage observations $\mathsf{L}(X_1)$ to $\mathsf{L}(X_n)$ with an equivalent leakage function $\mathsf{L}'(X) = (x_1 \oplus e'_1)\|(x_0 \oplus e'_0)$, where e'_0 and e'_1 are independent binary RVs such that $\Pr(e'_0 = 1) = \Pr(e'_1 = 1) = \mathsf{e}_n$. Since L' and L only differ in their parameters, we can use (13) to write:

$$\begin{aligned}\Delta &= \mathsf{SD}(X\,;\, X \mid \boldsymbol{L}(\boldsymbol{X})) = \mathsf{SD}(X\,;\, X \mid \mathsf{L}'(X)) = (\frac{1}{2} - \mathsf{e}_n)(\frac{3}{2} - \mathsf{e}_n)\\ &= 2^{n-1}(\frac{1}{2} - \mathsf{e})^n(1 + \frac{1}{2}(1 - 2\mathsf{e})^n).\end{aligned}$$

Lemma 3 gives the bound $\Delta \leq 2^{n-1}\delta^n$. However, in this example, we obtain:

$$\begin{aligned}\Delta &= 2^{n-1}(\frac{1}{2} - \mathsf{e})^n(1 + \frac{1}{2}(1 - 2\mathsf{e})^n)\\ &<_{(\mathrm{I})} 2^{n-1}(\frac{1}{2} - \mathsf{e})^n(1 + \frac{1}{2}(1 - 2\mathsf{e}))^n = 2^{n-1}\delta^{n-1},\end{aligned}$$

[1] For other metrics, such as the *mutual information* between $\boldsymbol{L}(\boldsymbol{X})$ and X, this tightness may not necessarily hold.

where step (I) follows because, for any $0 \leq t < 1$, $(1 + \frac{1}{2}t^n)$ decreases, whereas $(1 + \frac{1}{2}t)^n$ increases by n while they are equal at $n = 1$. □

The conclusion of this example that $\Delta < 2^{n-1}\delta^n$, implies that the bound in Lemma 4 is likely not tight for $q > 2$.

Need for More Fine-Tuned Analysis. Our discussion thus far has made it clear that for $q = 2^u$ with $u > 1$, there exists a gap in both the noisy-to-random probing reduction (as illustrated with $\mathsf{L}(X) = \mathsf{ZV}(X)$ instance) and in indirect metric estimation (as pointed out in Example 2). The main contribution of this paper is to narrow these gaps by introducing a *decomposition* approach that works for $q = 2^u$ fields.

3.2 Decomposition into Binary Subfields

The observation that in binary fields exact metrics are easy to compute and the reduction is tight, motivates us to decompose relations in a $\mathbb{F}_{2^u}$ field into binary relations, where metrics can be efficiently calculated, and then translate the results back. In this section, we demonstrate the applicability of this approach for masked encoding, and in the following section, we extend it to linear circuits. Below, we provide the foundational concepts of such a decomposition.

Consider two u-bit integers, A and B, and define their bitwise *inner product* as:

$$\langle A, B\rangle = \oplus_{i=0}^{u-1} a_i b_i,$$

where a_i and b_i are bits of A and B. For an RV $X \in \mathbb{F}_{2^u}$, with masked encoding as $\boldsymbol{X} = [X_1, \ldots, X_n]$, any integer $h \in [1, 2^u - 1]$, in its u-bit representation, can be deployed to map $X = X_1 \oplus \ldots \oplus X_n$ equation into a binary equation as:

$$\langle X, h\rangle = \langle X_1, h\rangle \oplus \ldots \oplus \langle X_n, h\rangle. \tag{14}$$

We will later establish the validity of this mapping in a broader context, specifically within Boolean systems of equations. To illustrate the practical usefulness of this mapping, in Theorem 1 will prove that if an adversary fails to learn any of the $2^u - 1$ binary random variables $\langle X, h\rangle$ from the leakage $\mathsf{L}(X)$, they cannot successfully deduce X from this leakage.

Theorem 1. *Given the leakage* $\mathsf{L}(X)$ *for* $X \in \mathbb{F}_{2^u}$*, the adversary's advantage in learning* X *is limited as:*

$$\frac{1}{2^{u-1}} \max_h \mathsf{Adv}_{\langle X,h\rangle} \leq \mathsf{Adv}_X \leq \frac{1}{2^{u-1}} \sum_{h=1}^{2^u-1} \mathsf{Adv}_{\langle X,h\rangle} < 2 \max_h \mathsf{Adv}_{\langle X,h\rangle}. \tag{15}$$

Here, $\mathsf{Adv}_{\langle X,h\rangle}$ *denotes the adversary's advantage in learning the binary random variable* $\langle X, h\rangle$*.*

Proof. For ease of notation, let us denote $\mathsf{Adv}_{\langle X,h\rangle}$ by μ_h. Given the $2^u - 1$ values $\{\mu_1, \mu_2, \ldots, \mu_{2^u-1}\}$, we aim to study the behavior of Adv_X. Specifically, we will show that the values μ_h are sufficient for determining the maximum value that Adv_X can attain.

Proof Overview. The proof follows a structured sequence of steps: (A) Formulating the problem as a system of equations involving the probability distributions of $X \mid l$ and the values μ_i. (B) Transforming this system into the *Hadamard* representation, which corresponds to computing the Walsh transform. (C) Leveraging the properties of Hadamard matrices, particularly the inverse Walsh transform, to characterize the solution space. (D) Incorporating constraints expressed in terms of expectation values. (E) Establishing bounds on the advantage term Adv_X.

Proof Detail. Let $\{p_0^l, p_1^l, \ldots, p_{2^u-1}^l\}$ be the posterior distribution of X given an instance l of the leakage. Using this notation, for a given h, μ_h is computed as follows:

$$\begin{aligned}\mu_h &= \mathop{\mathbb{E}}_{l}\left[\max_{\alpha\in\{0,1\}} \Pr(\langle X,h\rangle = \alpha \mid l)\right] - \frac{1}{2} \\ &= \mathop{\mathbb{E}}_{l}\left[\max_{\alpha\in\{0,1\}} \sum_{\langle\beta,h\rangle=\alpha} \Pr(X=\beta \mid l)\right] - \frac{1}{2} \\ &= \mathop{\mathbb{E}}_{l}\left[\max_{\alpha\in\{0,1\}} \sum_{\langle\beta,h\rangle=\alpha} p_\beta^l\right] - \frac{1}{2}.\end{aligned} \tag{16}$$

Since the values p_β^l represent probability masses, we have the following constraint:

$$\sum_{\langle\beta,h\rangle=0} p_\beta^l + \sum_{\langle\beta,h\rangle=1} p_\beta^l = 1,$$

where β enumerates over $\mathbb{F}_{2^u}$. With reordering the terms, we derive:

$$\sum_{\langle\beta,h\rangle=1} p_\beta^l - \frac{1}{2} = -\Big(\sum_{\langle\beta,h\rangle=0} p_\beta^l - \frac{1}{2}\Big) \Rightarrow \max_{\alpha\in\{0,1\}}\left[\sum_{\langle\beta,h\rangle=\alpha} p_\beta^l - \frac{1}{2}\right] = \left|\sum_{\langle\beta,h\rangle=1} p_\beta^l - \frac{1}{2}\right|.$$

Since $\langle\beta,h\rangle \in \{0,1\}$, we can use the following simplification:

$$\sum_{\langle\beta,h\rangle=1} p_\beta^l = \sum_{\beta} \langle\beta,h\rangle \cdot p_\beta^l.$$

Now, we define a new set of random variables as $\theta_h^l \triangleq \sum_\beta \langle\beta,h\rangle \cdot p_\beta^l - \frac{1}{2}$. The steps of the proof verify that $\mathbb{E}_l\left[|\theta_h^l|\right] = \mu_h$. We can also reorder the terms as:

$$\sum_{\beta} \langle\beta,h\rangle \cdot p_\beta^l = \frac{1}{2} + \theta_h^l. \tag{17}$$

By collecting the relations in (17) for all values of $h \in [1, 2^u - 1]$, we obtain the following system of 2^u equations in terms of the variables p_β^l and θ_h^l.

$$\begin{cases} \sum_{\beta\in\mathbb{F}_{2^u}} p^l_\beta = 1, \\ \sum_{\beta\in\mathbb{F}_{2^u}} \langle\beta,1\rangle \cdot p^l_\beta = \frac{1}{2} + \theta^l_1, \\ \sum_{\beta\in\mathbb{F}_{2^u}} \langle\beta,2\rangle \cdot p^l_\beta = \frac{1}{2} + \theta^l_2, \\ \quad\vdots \\ \sum_{\beta\in\mathbb{F}_{2^u}} \langle\beta,2^u-1\rangle \cdot p^l_\beta = \frac{1}{2} + \theta^l_{2^u-1}, \end{cases} \tag{18}$$

The first equation is sum of all p^l_β values, which is 1.

We update the given system of equations by applying a simple row operation to every row r_i except for the first row r_0. For $1 \le i \le 2^u - 1$, each row r_i is multiplied by the scalar -2 and then added to r_0. In other words, each row r_i is updated as $r_i \leftarrow -2r_i + r_0$. This operation results in a new system of equations in a familiar format without altering the solution space. Using a matrix representation, and observing that $(-1)^{\langle i,j\rangle} = -2\langle i,j\rangle + 1$, the updated system is expressed as:

$$\begin{bmatrix} (\text{-}1)^{\langle 0,0\rangle} & (\text{-}1)^{\langle 1,0\rangle} & \dots & (\text{-}1)^{\langle 2^u-1,0\rangle} \\ (\text{-}1)^{\langle 0,1\rangle} & (\text{-}1)^{\langle 1,1\rangle} & \dots & (\text{-}1)^{\langle 2^u-1,1\rangle} \\ (\text{-}1)^{\langle 0,2\rangle} & (\text{-}1)^{\langle 1,2\rangle} & \dots & (\text{-}1)^{\langle 2^u-1,2\rangle} \\ \vdots & \vdots & \ddots & \vdots \\ (\text{-}1)^{\langle 0,2^u-1\rangle} & (\text{-}1)^{\langle 1,2^u-1\rangle} & \dots & (\text{-}1)^{\langle 2^u-1,2^u-1\rangle} \end{bmatrix} \times \begin{bmatrix} p^l_0 \\ p^l_1 \\ p^l_2 \\ \vdots \\ p^l_{2^u-1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2\theta^l_1 \\ 2\theta^l_2 \\ \vdots \\ 2\theta^l_{2^u-1} \end{bmatrix} \tag{19}$$

For convenience, we denote the matrices participating in (19) as $\boldsymbol{M}$, $\boldsymbol{P}$, $\boldsymbol{C}_1$, and $\boldsymbol{C}_2$, respectively. Thus, we can write:

$$\boldsymbol{M} \times \boldsymbol{P} = \boldsymbol{C}_1 + \boldsymbol{C}_2.$$

The description of the solution(s) for this system depends on the structure of $\boldsymbol{M}$. For some initial values, we compute $\boldsymbol{M}$. At $u = 1$, $\boldsymbol{M}$ is:

$$\boldsymbol{M} = \begin{bmatrix} (-1)^{\langle 0,0\rangle} & (-1)^{\langle 0,1\rangle} \\ (-1)^{\langle 1,0\rangle} & (-1)^{\langle 1,1\rangle} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & \text{-}1 \end{bmatrix}.$$

At $u = 2$, $\boldsymbol{M}$ is:

$$\boldsymbol{M} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \text{-}1 & 1 & \text{-}1 \\ 1 & 1 & \text{-}1 & \text{-}1 \\ 1 & \text{-}1 & \text{-}1 & 1 \end{bmatrix}.$$

Upon closer inspection, it turns out that $\boldsymbol{M}$ is equivalent to a Hadamard matrix of size $2^u \times 2^u$, denoted $\boldsymbol{H}_{2^u}$. The Hadamard matrix is recursively defined, with $\boldsymbol{H}_1 = [1]$, and $\boldsymbol{H}_{2^k}$ defined as:

$$\boldsymbol{H}_{2^k} = \begin{bmatrix} \boldsymbol{H}_{2^{k-1}} & \boldsymbol{H}_{2^{k-1}} \\ \boldsymbol{H}_{2^{k-1}} & -\boldsymbol{H}_{2^{k-1}} \end{bmatrix}. \tag{20}$$

The equivalence of $\boldsymbol{M}$ and $\boldsymbol{H}_{2^u}$ is based on the following property of the $(-1)^{\langle i,j\rangle}$ mapping. Let A and B be u-bit variables, and let A' and B' be $u+1$-bit variables constructed from A and B by prepending bits a_u and b_u as $A' = a_u \| A$ and $B' = b_u \| B$. We have:

$$(-1)^{\langle A',B'\rangle} = \begin{cases} -(-1)^{\langle A,B\rangle} & \text{if both } a_u \text{ and } b_u \text{ are 1,} \\ (-1)^{\langle A,B\rangle} & \text{otherwise.} \end{cases}$$

Returning to our main discussion, by the properties of Hadamard matrices, the inverse of $\boldsymbol{M}$ is computed as:

$$\boldsymbol{M}^{-1} = \frac{1}{2^u}\boldsymbol{M}.$$

Since $\boldsymbol{M}$ is full rank, there is a unique solution for $\boldsymbol{P}$ that satisfies:

$$\boldsymbol{M} \times \boldsymbol{P} = \boldsymbol{C}_1.$$

We denote this solution by $\boldsymbol{P}_1$. Finding $\boldsymbol{P}_1$ is straightforward, and it can be readily verified that $\boldsymbol{P}_1$ is the uniform distribution, i.e., all entries of $\boldsymbol{P}_1$ are $\frac{1}{2^u}$.

It remains to study the structure of the solutions for:

$$\boldsymbol{M} \times \boldsymbol{P} = \boldsymbol{C}_2.$$

We denote these solutions by $\boldsymbol{P}_2$, and they can be derived as:

$$\boldsymbol{M} \times \boldsymbol{P}_2 = \boldsymbol{C}_2 \quad \Rightarrow \quad \boldsymbol{P}_2 = \frac{1}{2^u}\boldsymbol{M} \times \boldsymbol{C}_2.$$

Hence, each entry $\boldsymbol{P}_2(\beta)$ for $0 \le \beta \le 2^u - 1$ is given by:

$$\boldsymbol{P}_2(\beta) = \frac{1}{2^u}\sum_{h=1}^{2^u-1} 2\boldsymbol{M}(\beta,h)\theta_h^l = \frac{1}{2^{u-1}}\sum_{h=1}^{2^u-1} \boldsymbol{M}(\beta,h)\theta_h^l. \tag{21}$$

By adding the solutions corresponding to $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$, we can write:

$$\forall \beta \in \mathbb{F}_{2^u}, \quad p_\beta^l = \frac{1}{2^u} + \frac{1}{2^{u-1}}\sum_{h=1}^{2^u-1} \boldsymbol{M}(\beta,h)\theta_h^l.$$

Finally, we have:

$$\begin{aligned}
\mathsf{Adv}_X &= \mathop{\mathbb{E}}_l\left[\max_{\beta\in\mathbb{F}_{2^u}} \Pr(X=\beta|l) - \frac{1}{2^u}\right] = \mathop{\mathbb{E}}_l\left[\max_\beta p_\beta^l - \frac{1}{2^u}\right] \\
&= \mathop{\mathbb{E}}_l\left[\max_\beta \frac{1}{2^{u-1}}\sum_{h=1}^{2^u-1} \boldsymbol{M}(\beta,h)\theta_h^l\right] \\
&\le_{(\mathrm{I})} \mathop{\mathbb{E}}_l\left[\max_\beta \frac{1}{2^{u-1}}\sum_{h=1}^{2^u-1} \left|\boldsymbol{M}(\beta,h)\theta_h^l\right|\right] \\
&=_{(\mathrm{II})} \mathop{\mathbb{E}}_l\left[\frac{1}{2^{u-1}}\sum_{h=1}^{2^u-1} \left|\theta_h^l\right|\right] \\
&=_{(\mathrm{III})} \frac{1}{2^{u-1}}\sum_{h=1}^{2^u-1} \mu_h < 2\max_h \mu_h,
\end{aligned} \tag{22}$$

where,

- step (I) is because sum of elements is less than sum of their absolute values.
- step (II) follows since all entries of $\boldsymbol{M}$ are either $+1$ or -1.
- step (III) follows from $\mathbb{E}_l\left[\left|\theta_h^l\right|\right] = \mu_h$, shown at the initial parts of the proof.

The right side of (III) equals the upper bound claimed in the lemma. For the lower bound, assuming that $h^* = \text{argmax}_h\, \mu_h$, we revisit (22) and write as follows:

$$\begin{aligned}
\mathsf{Adv}_X &= \mathop{\mathbb{E}}_{l}\left[\max_{\beta} \frac{1}{2^{u-1}} \sum_{h=1}^{2^u-1} \boldsymbol{M}(\beta,h)\theta_h^l\right] \\
&\geq_{(\mathtt{I})} \mathop{\mathbb{E}}_{l,\boldsymbol{M}(\beta,h^*)=\mathsf{Sign}(\theta_{h^*}^l)}\left[\frac{1}{2^{u-1}} \sum_{h=1}^{2^u-1} \boldsymbol{M}(\beta,h)\theta_h^l\right] \\
&=_{(\mathtt{II})} \frac{1}{2^{u-1}}\mathop{\mathbb{E}}_{l}\left[\left|\theta_{h^*}^l\right|\right] + \frac{1}{2^{u-1}}\mathop{\mathbb{E}}_{l}\left[\sum_{h\neq h^*}\mathop{\mathbb{E}}_{\boldsymbol{M}(\beta,h^*)=\mathsf{Sign}(\theta_{h^*}^l)}[\boldsymbol{M}(\beta,h)]\theta_h^l\right] \\
&=_{(\mathtt{III})} \frac{1}{2^{u-1}}\max_h \mu_h,
\end{aligned} \tag{23}$$

where,

- step (I) follows since instead of choosing β that maximizes the sum, β is chosen randomly from those that satisfy $\boldsymbol{M}(\beta,h^*) = \mathsf{Sign}(\theta_{h^*}^l)$. For completeness, we also assume that $\mathsf{Sign}(0) = 1$.
- step (II) follows since $\mathsf{Sign}(\theta_{h^*}^l)\theta_{h^*}^l = |\theta_{h^*}^l|$. Also, for any a, b in $\{-1,1\}$ and $h \neq h^*$, we can show that $\Pr(\boldsymbol{M}(\beta,h) = a \mid \boldsymbol{M}(\beta,h^*) = b) = \frac{1}{2}$. Hence, we have $\mathbb{E}_{\boldsymbol{M}(\beta,h^*)=b}[\boldsymbol{M}(\beta,h)] = 0$.
- step (III) follows from $\mathbb{E}_l\left[\left|\theta_h^l\right|\right] = \mu_h$. □

Properties of the Upper Bounds in Theorem 1. We continue with the convention established in the proof and denote $\mathsf{Adv}_{\langle X,h\rangle}$ as μ_h.

Theorem 1 limits the adversary's advantage Adv_X by stating that:

$$\mathsf{Adv}_X \leq \frac{1}{2^{u-1}}\sum_{h\neq 0}\mu_h.$$

Our first objective is to show that, for certain leakage models, this bound is achievable. Then, we will prove that for any δ-noisy leakage function, $\frac{1}{2^{u-1}}\sum_{h\neq 0}\mu_h$ is less than δ, which aligns with the bound $\mathsf{Adv}_X \leq \delta$ as shown in Lemma 1.

Achievability. Consider the extreme case where $\mathsf{L}(X) = X$. In this case, we have $\mu_h = \frac{1}{2}$ for all $h \in [1, 2^u - 1]$. Substituting this into the inequality $\mathsf{Adv}_X \leq \frac{1}{2^{u-1}}\sum_{h\neq 0}\mu_h$, we get:

$$\mathsf{Adv}_X \leq \frac{1}{2^{u-1}}\left((2^u-1)\frac{1}{2}\right) = 1 - \frac{1}{2^u}.$$

On the other hand, from the definition of Adv_X, we have:

$$\mathsf{Adv}_X = \Pr(\hat{X} = X) - \frac{1}{2^u} = 1 - \frac{1}{2^u}.$$

The equivalence of these two expressions demonstrates that the upper bound is indeed achievable in this case.

Tightness. In the following, we prove a more interesting property of the given upper bound, demonstrating its closeness to the actual value of Adv_X.

Lemma 6. *For $X \in \mathbb{F}_{2^u}$ with a leakage function such that $\mathsf{SD}(X\,;\,X \mid \mathsf{L}(X)) = \delta$, we have:*

$$\frac{1}{2^{u-1}} \sum_{h=1}^{2^u-1} \mu_h \leq \delta.$$

Proof. Let $\{p_0^l, p_1^l, \ldots, p_{2^u-1}^l\}$ be the posterior distribution of X given a leakage instance l. Since $\langle X, h\rangle$ is a binary random variable, we can compute μ_h as:

$$\mu_h = \frac{1}{2}\,\mathop{\mathbb{E}}_{l}\left[\left|\sum_{\langle\beta,h\rangle=0} p_\beta^l - \frac{1}{2}\right| + \left|\sum_{\langle\beta,h\rangle=1} p_\beta^l - \frac{1}{2}\right|\right], \tag{24}$$

where β ranges over $\mathbb{F}_{2^u}$. Using this notation, we also have:

$$\delta = \frac{1}{2}\,\mathop{\mathbb{E}}_{l}\left[\sum_{\beta}\left|p_\beta^l - \frac{1}{2^u}\right|\right].$$

Define q_β^l as:

$$q_\beta^l = p_\beta^l - \frac{1}{2^u}.$$

Then, we have $\sum_\beta q_\beta^l = 0$, and since for any $h \neq 0$, $\langle\beta, h\rangle$ is 1 for half of the β values and 0 for the other half, the computations of δ and μ_h become:

$$\mu_h = \frac{1}{2}\,\mathop{\mathbb{E}}_{l}\left[\left|\sum_{\langle\beta,h\rangle=0} q_\beta^l\right| + \left|\sum_{\langle\beta,h\rangle=1} q_\beta^l\right|\right], \quad \delta = \frac{1}{2}\,\mathop{\mathbb{E}}_{l}\left[\sum_{\beta}\left|q_\beta^l\right|\right]. \tag{25}$$

By applying the triangle inequality to absolute values, it is straightforward to see that for any h, we have $\mu_h \leq \delta$. However, this result alone does not prove the lemma, so we require a more detailed analysis.

We first observe that if the signs of a and b differ in a sum like $|a+b+c|$, then:

$$|a+b+c| \leq |a+b| + |c| = |a| + |b| - 2\min\{|a|, |b|\} + |c|.$$

We will use this inequality successively in the proof. However, direct application poses a challenge: we need to know the outcome of many pairs of the form

$\min\{q_\beta^l, q_{\beta'}^l\}$. To address this, we propose decomposing the q_β^l values to avoid this ordering issue.

For each $q_\beta^l > 0$, there exists a decomposition as $q_\beta^l = a_{\beta,0} + a_{\beta,1} + \cdots + a_{\beta,2^u-1}$, where each $a_{\beta,i}$ is positive. For any $q_\alpha^l \leq 0$, we can write:

$$q_\alpha^l = -\left(\sum_{\beta=0}^{2^u-1} a_{\beta,\alpha}\right).$$

When $q_\beta^l > 0$ and $q_\alpha^l \leq 0$ appear in an absolute value sum, applying the inequality yields:

$$|q_\beta^l + q_\alpha^l + c| \leq |q_\beta^l| + |q_\alpha^l| - 2a_{\beta,\alpha} + |c|.$$

This rule can be generalized to sums involving multiple positive and negative terms. Suppose $q_{\beta_i}^l$ are positive and $q_{\alpha_j}^l$ are negative. Generalizing the inequality, we can write:

$$\left|\sum_{i\in I} q_{\beta_i}^l + \sum_{j\in J} q_{\alpha_j}^l\right| \leq \sum_{i\in I} |q_{\beta_i}^l| + \sum_{j\in J} |q_{\alpha_j}^l| - 2\sum_{i\in I}\sum_{j\in J} a_{\beta_i,\alpha_j}.$$

Now, applying this to the computation of μ_h in (25), we can show that each pair $q_\beta^l > 0$ and $q_\alpha^l \leq 0$ appears in the same absolute sum for half of the $h \in [1, 2^u - 1]$. For each such pair, there will be a *loss factor* of $-2a_{\beta,\alpha}$. Corresponding to each such pair inside $\sum_{h=1}^{2^u-1} \mu_h$, there will be a loss factor of $-\frac{2^u-1}{2}(2a_{\beta,\alpha})$. The total amount of loss for all pairs is:

$$-\sum_\beta \sum_\alpha \frac{2^u-1}{2}(2a_{\beta,\alpha}) = -(2^u-1)\sum_\alpha \left(\sum_\beta a_{\beta,\alpha}\right) = (2^u-1)\sum_{\alpha, q_\alpha^l \leq 0} (-q_\alpha^l).$$

Since $\sum_i q_i^l = 0$, we have:

$$\sum_{\alpha, q_\alpha^l \leq 0} -q_\alpha^l = \sum_{\beta, q_\beta^l > 0} q_\beta^l = \frac{1}{2}\sum_{i=0}^{2^u-1} |q_i^l|.$$

Putting everything together, we obtain:

$$\sum_{h=1}^{2^u-1} \mu_h \leq \frac{2^u-1}{2}\mathop{\mathbb{E}}_l\left[\sum_{i=0}^{2^u-1} |q_i^l|\right] - \frac{2^u-1}{4}\mathop{\mathbb{E}}_l\left[\sum_{i=0}^{2^u-1} |q_i^l|\right] = \frac{2^u-1}{4}\mathop{\mathbb{E}}_l\left[\sum_{i=0}^{2^u-1} |q_i^l|\right] = \frac{2^u-1}{2}\delta. \tag{26}$$

Consequently, we have:

$$\frac{1}{2^{u-1}}\sum_{h=1}^{2^u-1} \mu_h \leq \frac{2^u-1}{2^u}\delta < \delta.$$

This proves the lemma. □

Application of Theorem 1. This theorem expresses the side-channel security of X in terms of the security of the binary random variables $\langle X, h\rangle$. This approach simplifies the relatively difficult task of estimating $\mathsf{Adv}_X[l \leftarrow \mathsf{L}(X)]$ by reducing it to the more straightforward computation of the terms $\mu_h = \mathsf{Adv}_{\langle X,h\rangle}[l \leftarrow \mathsf{L}(X)]$. The power of this technique becomes especially evident when dealing with more complex structures. In this subsection, we apply it to the case of single mask encoding.

We recall that in the masking domain, given the leakage $\boldsymbol{L(X)}$ for an encoding $\boldsymbol{X} = \{X_1, \ldots, X_n\}$ of a secret X, our goal is to estimate $\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L(X)}]$. Additionally, for each value of h, we have the following binary equation:

$$\langle X, h\rangle = \langle X_1, h\rangle \oplus \cdots \oplus \langle X_n, h\rangle.$$

This equation can be interpreted as the masked encoding of $\langle X, h\rangle$ using the shares $\langle X_i, h\rangle$. Moreover, from Lemma 5, we deduce that for a binary secret $\langle X, h\rangle$, the following holds:

$$\mathsf{Adv}_{\langle X,h\rangle}[\boldsymbol{l} \leftarrow \boldsymbol{L(X)}] = 2^{n-1}\left(\mathsf{Adv}_{\langle X_i,h\rangle}[l \leftarrow \mathsf{L}(X_i)]\right)^n.$$

Now, applying Theorem 1, we obtain:

$$\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L(X)}] \leq \frac{1}{2^{u-1}} \sum_{h=1}^{2^u-1} 2^{n-1}\mu_h^n = 2^{n-u} \sum_{h=1}^{2^u-1} \mu_h^n \leq \Delta \leq_{(\mathtt{I})} 2^{n-1}\delta^n, \quad (27)$$

where $\Delta = \mathsf{SD}(X; X \mid \boldsymbol{L(X)})$, and $(\mathtt{I})$ follows from Lemma 4.

Interpretation in Terms of Mutual Information. The inequality

$$\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(X)] \;\leq\; \frac{1}{2^u} \sum_{h=1}^{2^u-1} (2\mu_h)^n$$

shows that if $(2\mu_h) < 1$ for all h, then Adv_X converges to zero as n increases. Furthermore, the condition $\mu_h < \frac{1}{2}$ is satisfied whenever $\mathsf{MI}\big(\langle X, h\rangle; \mathsf{L}(X)\big) < 1$.

From information theory [8], for the binary random variable $\langle X, h\rangle$, we have:

$$\mathsf{MI}\big(\langle X, h\rangle; \mathsf{L}(X)\big) \;=\; 1 \;-\; \mathsf{H}\big(\tfrac{1}{2} \pm \mu_h\big), \quad (28)$$

where $\mathsf{H}(\cdot)$ denotes the binary entropy function.[2] The mutual information attains its maximum value of 1 only when $\frac{1}{2} \pm \mu_h$ is 0 or 1, which requires $\mu_h = \frac{1}{2}$. In this scenario, the leakage $\mathsf{L}(X)$ fully reveals $\langle X, h\rangle$.

[2] Recall that the capacity of a Binary Symmetric Channel (BSC) between A and B is $\mathsf{MI}(A; B) = \mathsf{H}(A) - \mathsf{H}(A \mid B) = \log_2\big(|A|\big) - \mathsf{H}(P_e) = 1 - \mathsf{H}(P_e)$, where P_e is the probability of incorrectly estimating A given B. In our setting, $\mu_h = |(1 - P_e) - \frac{1}{2}| = |P_e - \frac{1}{2}|$, so $P_e = \frac{1}{2} \pm \mu_h$.

Example 3. We revisit Example 2 to demonstrate the power of our decomposition approach. In that example, for $X \in \mathbb{F}_{2^2}$ with $\mathsf{L}(X) = (x_1 \oplus e_1) \| (x_0 \oplus e_0)$ and $\Pr(e_0 = 1) = \Pr(e_1 = 1) = \mathsf{e}$, we had:

$$\delta = \mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) = \left(\frac{1}{2} - \mathsf{e}\right)\left(\frac{3}{2} - \mathsf{e}\right).$$

For masked encoding, we derived:

$$\Delta = \mathsf{SD}(X\,;\, X \mid \boldsymbol{L}(\boldsymbol{X})) = 2^{n-1}\left(\frac{1}{2} - \mathsf{e}\right)^n \left(1 + \frac{1}{2}(1 - 2\mathsf{e})^n\right).$$

For this leakage function, from Example 1 (by setting $u = 1$), we know $\mu_1 = \mu_2 = \frac{1}{2} - \mathsf{e}$. Recall that μ_1 and μ_2 represent the adversary's advantage in estimating the first and second bits of X, respectively. Similarly, μ_3 represents the adversary's advantage in estimating $x_0 \oplus x_1$, for which we can show that:

$$\mu_3 = \frac{1}{2} - 2\mathsf{e}(1 - \mathsf{e}).$$

Alternatively, by deploying Theorem 1, we get:

$$\mathsf{Adv}_X[l \leftarrow \mathsf{L}(X)] \leq \frac{1}{2}(\mu_1 + \mu_2 + \mu_3) = \left(\frac{1}{2} - \mathsf{e}\right)\left(\frac{3}{2} - \mathsf{e}\right) = \delta.$$

For masked encoding, using (27), we write:

$$\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})] \leq 2^{n-2}\left(\mu_1^n + \mu_2^n + \mu_3^n\right) = 2^{n-1}\left(\frac{1}{2} - \mathsf{e}\right)^n \left(1 + \frac{1}{2}(1 - 2\mathsf{e})^n\right) = \Delta. \tag{29}$$

This result underscores the effectiveness of our approach. For the given leakage function, setting $\mathsf{e} = 0.1$ yields $\delta = 0.56$. Because $\delta > \frac{1}{2}$, Lemma 4 does not determine whether secure masking is achievable. Meanwhile, using the MI-based criterion with $\mathsf{e} = 0.1$ gives $\mathsf{MI}(X; \mathsf{L}(X)) = 1.06$, which exceeds the $\mathsf{MI}(X; \mathsf{L}(X)) < 0.72$ threshold required by Ito et al. [20] (see Subsection 1.1). Thus, neither approach can confirm or deny secure masking in this setting. In contrast, our method confirms that secure masking is indeed possible. □

Example 4. We previously introduced the leakage model $\mathsf{ZV}(X)$, defined as:

$$\mathsf{ZV}(X) = \begin{cases} \nu_a & \text{if } X = 0, \\ \nu_b \neq \nu_a & \text{otherwise.} \end{cases}$$

For this model, we derive $\epsilon_{\min} = 1$ using (4). Thus, the noisy-to-random probing reduction cannot be used to analyze the security of masking in the presence of this leakage function.

Alternatively, using our decomposition approach, we can compute μ_h as $\mu_h = \frac{1}{2^u}$ (using relation (24)) and obtain $\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})]$ (using relation (27)) as:

$$\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})] \leq 2^{n-u} \sum_{h=1}^{2^u - 1} \mu_h^n = (2^u - 1)2^{n-u-nu},$$

which indicates that the adversary's advantage decreases with increasing masking order for $u > 1$. □

Application to Leakage Certification. Leakage certification laboratories evaluate a given device and its cryptographic implementation to assess leakage and robustness against side-channel attacks (see [11, 29]). In doing so, they typically compute metrics such as $\mathsf{MI}(X; \mathsf{L}(X))$ and $\mathsf{SD}(X; X \mid \mathsf{L}(X))$. Estimating these metrics requires knowledge of the distribution of $(X, \mathsf{L}(X))$, which can be obtained either via parametric methods (e.g., assuming a Gaussian distribution and estimating its parameters) or non-parametric methods (e.g., histogram-based) [2, 16].

Our work proposes an additional metric for evaluating a device's leakage. Specifically, for a u-bit value X, we require that

$$\mathsf{MI}\big(\langle X, h\rangle; \mathsf{L}(X)\big) < 1 \quad \text{for all} \quad h \in [1, 2^u - 1].$$

Masking provides side-channel protection if and only if this condition holds for every h.

Using Equation (28), we can restate the results of Theorem 1 (and the masked variant in Equation (27)) in terms of mutual information as follows:

$$\frac{1}{2^u} \max_h \left| \tfrac{1}{2} - \mathsf{H}^{-1}\big(I_h\big) \right|^n \;\leq\; \mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(X)] \;\leq\; \Big(2 \max_h \big| \tfrac{1}{2} - \mathsf{H}^{-1}\big(I_h\big) \big| \Big)^n, \quad (30)$$

where $I_h = 1 - \mathsf{MI}\big(\langle X, h\rangle; \mathsf{L}(X)\big)$, and H^{-1} is the inverse of the binary entropy function.

3.3 Masking in Odd Prime Fields

Grassi et al. [17], building on earlier results from Dziembowski et al. [13], demonstrated that when the leakage function is too informative, such as when $\mathsf{L}(X) = \mathsf{HW}(X)$, masking is only effective in odd prime fields. This finding has motivated efforts to better understand the aspects of prime field masking [14].

In this section, we contribute to this line of research by presenting new findings. As a preliminary result, we show that for a specific class of leakage functions, the adversary's advantage Adv_X decays more rapidly with increasing masking order n in prime fields (Lemma 7). We then establish a general condition under which masking effectively reduces the Adv_X metric (Lemma 8), and in doing so, we prove a conjecture posed by Dziembowski et al. [13] (Theorem 2).

Class of Symmetric Leakages. We define a leakage class as *symmetric* if, for a uniform $X \in \mathbb{F}_q$, observing an instance l of the leakage transforms the distribution of $X \mid l$ to $(p_{e_0}, p_{e_1}, \ldots, p_{e_{q-1}})$, where $\sum_{i=0}^{q-1} p_{e_i} = 1$, and p_{e_i} denotes the probability mass on the element $X + i$. [3] In this notation, p_{e_0} represents the

[3] We refer to this as a symmetric leakage class because it generalizes the binary symmetric channel.

probability of a correct estimation, and we have:

$$\mathsf{Adv}_X = p_{e_0} - \frac{1}{q}.$$

In a masked encoding with $n = 2$, let X_1 and X_2 represent the shares of X, such that $X = X_1 + X_2$. After observing the leakages, the adversary forms estimates $\hat{X_1}$ and $\hat{X_2}$. We assume that the best estimate of X is derived as $\hat{X_1} + \hat{X_2}$. The adversary correctly recovers X if there are no errors in either X_1 or X_2, or if the errors in estimating X_1 and X_2 are i and $q - i$, respectively. Consequently, the adversary's success probability, denoted p'_{e_0}, is updated as:

$$p'_{e_0} = (p_{e_0})^2 + \sum_{i=1}^{q-1} p_{e_i} p_{e_{q-i}}. \tag{31}$$

Lemma 7. *For the defined symmetric leakage class, when the field order is prime,* Adv_X *decays faster with increasing* n*.*

Proof. In a prime field, i and $q-i$ are distinct elements, and we can rewrite (31) as:

$$p'_{e_0} = (p_{e_0})^2 + \sum_{i=1}^{q-1} p_{e_i} p_{e_{q-i}} = (p_{e_0})^2 + 2\sum_{i=1}^{\frac{q-1}{2}} p_{e_i} p_{e_{q-i}}.$$

However, when $q = 2^u$, $q - i$ and i are equal, and we can rewrite (31) as:

$$p'_{e_0} = (p_{e_0})^2 + \sum_{i=1}^{q-1} p_{e_i}^2.$$

By applying the inequality $2ab \leq a^2 + b^2$, we can deduce that:

$$(p_{e_0})^2 + 2\sum_{i=1}^{\frac{q-1}{2}} p_{e_i} p_{e_{q-i}} \leq (p_{e_0})^2 + \sum_{i=1}^{q-1} p_{e_i}^2.$$

This proves that p'_{e_0}, and consequently the adversary's advantage, is smaller in prime fields than in fields where $q = 2^u$. □

Condition Under Which Adv_X Decreases with Masking. For a masked encoding of the secret X, we show that if there is no *hole* in the posterior distribution of the shares after receiving the leakage vector, then Adv_X will decrease.

Definitions. For later reference, we define a leakage instance l as *dummy* if it causes no change in the distribution of $X \mid l$.[4] With a dummy leakage, the peak point of the posterior distribution will be $\frac{1}{q}$. A *hole* in a distribution is an element of its domain with zero probability mass, and the *support* of a random variable is the number of elements in its domain with non-zero probability mass. We denote the support of X as $|X|$.

[4] We refer to it as dummy because any random variable that is independent of X will have a similar effect.

Problem Statement. Let X_1 and X_2 be shares of X in $\mathbb{F}_q$, and suppose the adversary receives leakage instances $l_1 \leftarrow \mathsf{L}(X_1)$ and $l_2 \leftarrow \mathsf{L}(X_2)$ corresponding to these two shares. The probability distributions of $X_1 \mid l_1$ and $X_2 \mid l_2$ are denoted by $\mathcal{P}^1 = (p_0^1, p_1^1, \ldots, p_{q-1}^1)$ and $\mathcal{P}^2 = (p_0^2, p_1^2, \ldots, p_{q-1}^2)$, respectively. Let $p_{i^*}^1$ and $p_{j^*}^2$ be the peak points of the distributions $\mathcal{P}^1$ and $\mathcal{P}^2$.

To estimate the value of X, the maximum a posteriori (MAP) adversary computes the distribution of the sum $(X_1 \mid l_1) + (X_2 \mid l_2)$. We denote the resultant distribution by $\mathcal{P} = (p_0, p_1, \ldots, p_{q-1})$. The adversary declares the index of the peak point of $\mathcal{P}$ as $\hat{X}$, and their advantage is denoted as $\mathsf{Adv}_X[l_1, l_2 \leftarrow \mathsf{L}(X_1), \mathsf{L}(X_2)]$. We seek conditions that guarantee:

$$\mathsf{Adv}_X[l_1, l_2 \leftarrow \mathsf{L}(X_1), \mathsf{L}(X_2)] < \mathsf{Adv}_{X_i}[l_i \leftarrow \mathsf{L}(X_i)],$$

which implies that masking has *strictly* improved the side-channel security of X.

Lemma 8. *If, for at least one non-dummy instance of leakage, there is no hole in the posterior distributions $\mathcal{P}^1$ and $\mathcal{P}^2$, that is, $\min \mathcal{P}^1 > 0$ and $\min \mathcal{P}^2 > 0$, then the adversary's advantage will strictly decrease.*

Proof. For $0 \le i \le q-1$, define $\zeta_i = \frac{p_i^1}{p_{i^*}^1}$ and $\xi_i = \frac{p_i^2}{p_{j^*}^2}$. Since $p_{i^*}^1$ and $p_{j^*}^2$ are the peak values of the probability distributions, we have $\zeta_i \le 1$ and $\xi_i \le 1$. Substituting into the normalization conditions $\sum_{i=0}^{q-1} p_i^1 = 1$ and $\sum_{i=0}^{q-1} p_i^2 = 1$, we obtain:

$$p_{i^*}^1 = \frac{1}{\zeta_0 + \zeta_1 + \cdots + \zeta_{q-1}}, \quad \text{and} \quad p_{j^*}^2 = \frac{1}{\xi_0 + \xi_1 + \cdots + \xi_{q-1}}.$$

Let k^* be the peak point of the distribution $\mathcal{P}$ (introduced earlier in the problem). We now prove that $p_{k^*} \le \min\{p_{i^*}^1, p_{j^*}^2\}$. Without loss of generality, assume $\min\{p_{i^*}^1, p_{j^*}^2\} = p_{i^*}^1$. Hence, we aim to show that $p_{k^*} \le p_{i^*}^1$, which implies:

$$p_{k^*} = \sum_{i=0}^{q-1} p_{k^*-i}^1 \, p_i^2 = p_{i^*}^1 \, p_{j^*}^2 \sum_{i=0}^{q-1} \zeta_{k^*-i} \, \xi_i \le p_{i^*}^1.$$

Since $p_{i^*}^1 > 0$, the above inequality holds if:

$$\sum_{i=0}^{q-1} \zeta_{k^*-i} \, \xi_i \le \sum_{i=0}^{q-1} \xi_i,$$

which is always true because $\zeta_i \le 1$.

Moreover, equality $p_{k^*} = p_{i^*}^1$ holds only if:

$$\forall i \in [0, q-1], \quad \zeta_{k^*-i} \, \xi_i = \xi_i. \tag{32}$$

Since $\mathcal{P}^2$ has no holes (i.e., $\xi_i > 0$ for all i), the condition in (32) is only possible if $\zeta_i = 1$ for all i. This implies that $p_{i^*}^1 = \frac{1}{q}$, meaning the leakage is dummy,

which contradicts our assumptions. Thus, $p_{k^*} = \min\{p^1_{i^*}, p^2_{j^*}\}$ cannot hold, and we must have $p_{k^*} < \min\{p^1_{i^*}, p^2_{j^*}\}$.

For all instances of leakage, $p_{k^*} \leq \min\{p^1_{i^*}, p^2_{j^*}\}$, and for at least one instance, $p_{k^*} < \min\{p^1_{i^*}, p^2_{j^*}\}$. Therefore, taking expectations, we have:

$$\mathop{\mathbb{E}}_{l}[p_{k^*}] < \mathop{\mathbb{E}}_{l}[p^1_{i^*}] = \mathop{\mathbb{E}}_{l}[p^2_{j^*}] \quad \Rightarrow \quad \mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(\boldsymbol{X})] < \mathsf{Adv}_X[l \leftarrow \mathsf{L}(X)]. \qquad \square$$

Reaching a Hole-Free Posterior Distribution. When the leakage function is less noisy, such as with $\mathsf{L}(X) = \mathsf{HW}(X)$, there will be holes in the posterior distribution of $X_i \mid \mathsf{L}(X_i)$, preventing the application of Lemma 8. However, increasing the number of shares resolves this issue, as we explain below.

Although Lemma 8 is confined to the simple case of $n = 2$, the results generalize easily to higher values of n. Let $X_1, \ldots, X_{2n}$ be shares of X, and let $\mathsf{L}(X_1), \ldots, \mathsf{L}(X_{2n})$ represent the corresponding leakage functions. The problem of estimating X from the leakage vector can be decomposed into two steps: first, estimating $X_1 + \cdots + X_n$ and $X_{n+1} + \cdots + X_{2n}$ from their respective leakages, and second, estimating X from the distribution of their sum.

If, from a certain threshold order n_0 onward, the distribution of $(X_1 + \cdots + X_n) \mid (\mathsf{L}(X_1), \ldots, \mathsf{L}(X_n))$ has no holes, then by applying Lemma 8, we know that $\mathsf{Adv}_X[l_1, \ldots, l_{2n} \leftarrow \mathsf{L}(X_1), \ldots, \mathsf{L}(X_{2n})]$ will decrease with n, demonstrating the security of the mask encoding.

By the independence of shares and the internal randomness of leakage functions, the probability distribution $(X_1 + \cdots + X_n) \mid (\mathsf{L}(X_1), \ldots, \mathsf{L}(X_n))$ simplifies to $X_1 \mid \mathsf{L}(X_1) + \cdots + X_n \mid \mathsf{L}(X_n)$. At any instance of the leakage vector, $X_i \mid l_i$ are probability distributions, and we want to know how large the support of $X_1 \mid l_1 + \cdots + X_n \mid l_n$ will be. Specifically, if the support reaches $|\mathbb{F}_q|$, by definition, there will be no holes in the summed probability distribution.

Lemma 9 (Generalized Cauchy-Davenport Theorem). *Let $Z_1, \ldots, Z_t$ be independent random variables with supports $|Z_1|, \ldots, |Z_t|$ defined in the same prime field $\mathbb{F}_q$. For the support of their sum, we have:*

$$|Z_1 + Z_2 + \cdots + Z_t| \geq \min\{|Z_1| + |Z_2| + \cdots + |Z_t| - t, q\}.$$

Proof. This is a direct generalization of the Cauchy-Davenport theorem. The original statement is for $t = 2$. In the side-channel literature, the $t = 2$ case was used in the work of Dziembowski et al. [13]. $\square$

Using this lemma, we deduce that if $X_i \mid l_i$ places probability mass on more than one element of $\mathbb{F}_q$ (i.e., if $|X_i \mid l_i| > 1$), then for some $n \geq n_0$, the distribution $X_1 \mid l_1 + \cdots + X_n \mid l_n$ will have no holes. Consequently, from that point onward, Adv_X will be decreasing.

Our discussion here is intended to outline the asymptotic behavior of the adversary's advantage; it is not meant to provide an efficient way of obtaining this advantage. Nonetheless, the methodology helps to prove the following conjecture for the Adv_X security metric. The original conjecture was made for the statistical distance Δ metric.

Theorem 2 (Conjectured in [13]). *Let X be a secret in an additive group $\mathbb{G}$ with largest subgroup $\mathbb{H}$, and let X_i be its shares. If the leakage satisfies $\delta = \mathsf{SD}(X_i\,;\, X_i \mid \mathsf{L}(X_i)) < 1 - \frac{|\mathbb{H}|}{|\mathbb{G}|}$, then $\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(X)]$ decreases asymptotically as n increases.*

Proof. We must show that under the given leakage, the summed posterior distribution

$$\big(X_1 \mid \mathsf{L}(X_1)\big) \;+\; \ldots \;+\; \big(X_n \mid \mathsf{L}(X_n)\big)$$

eventually has no *holes*, i.e., no elements in $\mathbb{G}$ that are assigned zero probability. Once we establish this, Lemma 8 immediately implies that the adversary's advantage decreases with n.

Step 1: Showing the existence of mass in $\mathbb{G} - \mathbb{H}$. We first prove that there exists at least one leakage instance l for which the conditional distribution $X \mid \mathsf{L}(X) = l$ has nonzero mass on some element of $\mathbb{G} - \mathbb{H}$. Let A be the set of all $\alpha \in \mathbb{G}$ such that $\Pr(X = \alpha \mid l) < \frac{1}{|\mathbb{G}|}$. Suppose, for the sake of contradiction, that the support of $X \mid l$ is contained entirely in $\mathbb{H}$, i.e., it assigns zero probability to every element in $\mathbb{G} - \mathbb{H}$. Then, the statistical distance can be evaluated as:

$$\begin{aligned}
\delta \;=\; \mathsf{SD}(X\,;\, X \mid \mathsf{L}(X)) &= \sum_l \Pr(l)\, \mathsf{TV}\big(X,\, X \mid l\big) \\
&= \sum_l \Pr(l) \sum_{\alpha \in A} \Big[\tfrac{1}{|\mathbb{G}|} - \Pr(X = \alpha \mid l)\Big] \\
&= \sum_l \Pr(l) \sum_{\substack{\alpha \in A \\ \alpha \in \mathbb{H}}} \Big[\tfrac{1}{|\mathbb{G}|} - \Pr(X = \alpha \mid l)\Big] \;+\; (|\mathbb{G}| - |\mathbb{H}|)\left(\tfrac{1}{|\mathbb{G}|} - 0\right).
\end{aligned} \tag{33}$$

Because $\mathbb{G} - \mathbb{H}$ is excluded from the distribution, the second term in (33) simplifies to $1 - \frac{|\mathbb{H}|}{|\mathbb{G}|}$. Hence, if the support were entirely in $\mathbb{H}$, we would get $\delta \geq 1 - \frac{|\mathbb{H}|}{|\mathbb{G}|}$, contradicting the assumption $\delta < 1 - \frac{|\mathbb{H}|}{|\mathbb{G}|}$. Therefore, there must be a leakage instance l such that $X \mid \mathsf{L}(X) = l$ has support on at least one element of $\mathbb{G} - \mathbb{H}$.

Step 2: Covering all of $\mathbb{G}$ through summation. A result by Stromberg [31] shows that sums of elements in $\mathbb{G} - \mathbb{H}$ will eventually cover the entire group $\mathbb{G}$. In our setting, this implies that the combined posterior distribution over $X_1 + \cdots + X_n$, conditioned on $\mathsf{L}(X_1), \ldots, \mathsf{L}(X_n)$, will place nonzero probability on every element of $\mathbb{G}$ once n is large enough. Thus, the distribution becomes *hole-free*.

Step 3: Concluding the proof. By Lemma 8, when the distribution is hole-free, the adversary's advantage $\mathsf{Adv}_X[\boldsymbol{l} \leftarrow \boldsymbol{L}(X)]$ decreases with n. This completes the proof. □

4 Security of Linear Gadgets

Our discussion so far has been focused on standalone secrets and their masked encodings. In this section, we extend the analysis to study the adversary's ad-

vantage in more complex structures, specifically linear *gadgets*. Our primary goal is to demonstrate the applicability of the proposed decomposition approach for security evaluation in these gadgets.

Gates and Gadgets. A gadget is a *family* of circuits (one for each order n) designed to compute the masked counterpart of a gate. Let $\mathsf{G}\colon (\mathbb{F}_{2^u})^t \to \mathbb{F}_{2^u}$ be a gate with fan-in (number of input variables) t and fan-out 1. For example, XOR and AND gates have $t = 2$. A gadget for gate G, denoted as $\mathbb{S}\mathsf{G}\colon (\mathbb{F}_{2^u}^n)^t \to \mathbb{F}_{2^u}^n$, accepts masked encodings as inputs and produces masked encodings as outputs.

A *refresh gadget*, denoted as $\boldsymbol{X}' = \mathbb{S}\mathsf{R}(\boldsymbol{X})$, has fan-in and fan-out of 1. It updates the encoding of input $\boldsymbol{X}$ while preserving the secret, i.e., $(\oplus_{i=1}^n X_i) = (\oplus_{i=1}^n X_i')$, where X_i and X_i' represent the input and output shares. An example of an $\mathbb{F}_2$-linear refresh gadget is $\mathbb{S}\mathsf{R}\text{-}\mathsf{SNI}$ [1], described in Algorithm 1.

Algorithm 1 $\mathbb{S}\mathsf{R}\text{-}\mathsf{SNI}$

Input $\boldsymbol{X} = (X_1, \ldots, X_n)$
Output $\boldsymbol{X}' = (X_1', \ldots, X_n')$
1: **for** $i = 1$ **to** n **do**
2: **for** $j = i + 1$ **to** n **do**
3: $r \xleftarrow{\$} \mathbb{F}_{2^u}$
4: $X_i = X_i \oplus r$
5: $X_j = X_j \oplus r$
6: **return** $\boldsymbol{X}' = \boldsymbol{X}$

4.1 Problem Statement

Let $\boldsymbol{\Sigma}_n = \{V_1, \ldots, V_{\mathsf{T}(n)}\}$ represent the set of intermediate variables in an $\mathbb{F}_2$-linear masked gadget processing a secret X. We assume that X and the elements of $\boldsymbol{\Sigma}_n$ are all within the same field $\mathbb{F}_{2^u}$. In addition to the shares of X, $\boldsymbol{\Sigma}_n$ includes other random variables (RVs) whose leakage might assist the adversary in estimating X. Our goal is to determine the extent to which these leakages empower the adversary (via MAP estimation).

The assumption that the circuit is $\mathbb{F}_2$-linear implies the existence of a matrix $\boldsymbol{P}_n \in \mathbb{F}_2^{\mathsf{P}(n)\times(\mathsf{T}(n)+1)}$ such that:

$$\boldsymbol{P}_n \times [X, V_1, \ldots, V_{\mathsf{T}(n)}]^\dagger = \mathbf{0}_{\mathsf{P}(n)\times 1}. \tag{34}$$

$\boldsymbol{P}_n$ fully describes the circuit, as it can be used to compute outputs given inputs.[5] The $\mathsf{P}(n)$ rows of $\boldsymbol{P}_n$ represent the *parity relations* among the random variables $[X, \boldsymbol{\Sigma}_n]$, with any other dependencies between $[X, \boldsymbol{\Sigma}_n]$ expressible as linear combinations of these rows. The adversary is assumed to know $\boldsymbol{P}_n$ or any equivalent linear form of it.[6] Additionally, by conducting side-channel measurements, the adversary obtains $\mathsf{T}(n)$ leakages as:

$$\boldsymbol{L}_n = [\mathsf{L}(V_1), \ldots, \mathsf{L}(V_{\mathsf{T}(n)})]. \tag{35}$$

[5] We consider randomness variables as inputs.
[6] The results do not depend on the specific representation of $\boldsymbol{P}_n$.

The adversary's ultimate objective is to use the leakage information $\boldsymbol{L}_n$ and the parity relations $\boldsymbol{P}_n$ to estimate the realized value of the secret X. Our aim is to determine their advantage, denoted as $\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n]$.

MAP Adversary and Exact $\mathsf{Adv}_{\boldsymbol{X}}[\boldsymbol{l_n} \leftarrow \boldsymbol{L_n}]$. Let $\mathcal{S}_n$ denote the set of all solutions that satisfy the system of equations described by $\boldsymbol{P}_n$ (given in (34)). Each element $\boldsymbol{S} \in \mathcal{S}_n$ is a $(\mathsf{T}(n)+1)$-length tuple. Let $\boldsymbol{S}(0)$ represent the value of X in this solution. Given a leakage instance $\boldsymbol{l}_n$, the MAP adversary's estimation of X, denoted as $\hat{X}$ (the most probable value of X), is computed as:

$$\hat{X} = \underset{\alpha \in \mathbb{F}_{2^u}}{\operatorname{argmax}} \sum_{\boldsymbol{S} \in \mathcal{S}_n, \boldsymbol{S}(0)=\alpha} \Pr(\boldsymbol{S} \mid \boldsymbol{l}_n). \tag{36}$$

The adversary's advantage $\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n]$, by definition, is computed as:

$$\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n] = \mathop{\mathbb{E}}_{\boldsymbol{l}_n} \left[\Pr\left((\hat{X} = X) \mid \boldsymbol{l}_n \right) \right]. \tag{37}$$

A Non-Tight Upper Bound for $\mathsf{Adv}_{\boldsymbol{X}}[\boldsymbol{l_n} \leftarrow \boldsymbol{L_n}]$. By employing the noisy-to-random probing reduction, we can derive an upper bound for $\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n]$ that is more practical to compute. The key idea is to replace the leakage $\mathsf{L}(V_i)$ with $\phi^\epsilon(V_i)$, where $\phi^\epsilon(.)$ is the erasure function defined in Section 2.2 with an erasure probability of $1-\epsilon$. Here, $\epsilon \geq \epsilon_{\min}$, with $\epsilon_{\min}$ calculated from the leakage function L via (4). Using similar reasoning to that in (5), we can apply the reduction to obtain the following upper bound:

$$\mathsf{Adv}_X\left[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n\right] \leq \mathsf{Adv}_X\left[l_1, \ldots, l_{\mathsf{T}(n)} \leftarrow \phi^\epsilon(V_1), \ldots, \phi^\epsilon(V_{\mathsf{T}(n)})\right]. \tag{38}$$

For ease of notation, we denote $\mathsf{Adv}_X[l_1, \ldots, l_{\mathsf{T}(n)} \leftarrow \phi^\epsilon(V_1), \ldots, \phi^\epsilon(V_{\mathsf{T}(n)})]$ as $\mathsf{Adv}(q, n, \epsilon)$, which represents the adversary's advantage in estimating the secret X after receiving the random probing leakage of all intermediates. A lower ϵ indicates less information for the adversary, and hence a lower advantage. Thus, we can express:

$$\mathsf{Adv}_X\left[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n\right] \leq \mathsf{Adv}(q, n, \epsilon_{\min}) \leq \mathsf{Adv}(q, n, \epsilon). \tag{39}$$

Jahandideh et al. [21] recently developed a framework to estimate $\mathsf{Adv}(q, n, \epsilon)$ in linear circuits. For example, they estimated $\mathsf{Adv}(q, n, \epsilon)$ for the $\mathbb{S}$R-SNI gadget in Algorithm 1, for $n < 30$ and $\epsilon < 0.15$, as:

$$\mathsf{Adv}(q, n, \epsilon) \leq \frac{q-1}{q} \epsilon^{0.6n}. \tag{40}$$

However, the gap in the noisy-to-random reduction for fields $\mathbb{F}_q$ with $q > 2$ can result in a loose or even trivial upper bound. For instance, with $\mathsf{L}(X) = \mathsf{ZV}(X)$ (defined in (8)), we have $\epsilon_{\min} = 1$. Given that $\mathsf{Adv}(q, n, 1)$ equals 1, the upper bound becomes trivial: $\mathsf{Adv}\left[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n\right] \leq 1$.

In the following section, we demonstrate how our decomposition approach can address this issue for certain leakage functions, leading to a tighter upper bound for $\mathsf{Adv}\left[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n\right]$ in $\mathbb{F}_{2^u}$ fields.

4.2 Decomposition into Binary Systems

Our first observation is that in an $\mathbb{F}_2$-linear system such as

$$\boldsymbol{P}_n \times [X, V_1, \ldots, V_{\mathsf{T}(n)}]^\dagger = \mathbf{0}, \tag{41}$$

the variables $\langle X, h\rangle$ and $\langle V_i, h\rangle$ are linearly related, and their dependencies are described by the same matrix $\boldsymbol{P}_n$. More concretely, we present the following lemma.

Lemma 10. *Since the entries of $\boldsymbol{P}_n$ are binary, for any $h \in [1, 2^u - 1]$, the system $\boldsymbol{P}_n \times [X, V_1, \ldots, V_{\mathsf{T}(n)}]^\dagger = \mathbf{0}$ implies that*

$$\boldsymbol{P}_n \times [\langle X, h\rangle, \langle V_1, h\rangle, \ldots, \langle V_{\mathsf{T}(n)}, h\rangle]^\dagger = \mathbf{0}. \tag{42}$$

Proof. For the inner product of u-bit integers h, V_1, and V_2, we can write:

$$\langle V_1 \oplus V_2, h\rangle = \langle V_1, h\rangle \oplus \langle V_2, h\rangle. \tag{43}$$

For a binary scalar b, by testing both possible values of b, we can verify that:

$$\langle bV, h\rangle = b\langle V, h\rangle. \tag{44}$$

With iterative application of these rules, for binary coefficients $\{b_1, b_2, \ldots, b_{\mathsf{T}(n)}\}$ and u-bit variables $\{X, V_1, \ldots, V_{\mathsf{T}(n)}\}$, we can show that:

$$\langle (b_1 X \oplus b_2 V_1 \oplus \ldots \oplus b_{\mathsf{T}(n)} V_{\mathsf{T}(n)}), h\rangle = b_1\langle X, h\rangle \oplus b_2\langle V_1, h\rangle \oplus \ldots \oplus b_{\mathsf{T}(n)}\langle V_{\mathsf{T}(n)}, h\rangle. \tag{45}$$

Since $\boldsymbol{P}_n$ consists of $\mathsf{P}(n)$ equations, each with coefficients as in equation (45), the lemma follows directly by applying these results to all $\mathsf{P}(n)$ equations. □

A Tighter Upper Bound for $\mathsf{Adv}_{\boldsymbol{X}}[\boldsymbol{l_n} \leftarrow \boldsymbol{L_n}]$. The adversary also has access to side-channel information for each intermediate. From Theorem 1, given the leakage $\mathsf{L}(V)$, we know that:

$$\mathsf{Adv}_V[l \leftarrow \mathsf{L}(V)] < 2 \max_h \mathsf{Adv}_{\langle V,h\rangle}[l \leftarrow \mathsf{L}(V)] = 2\mu_{h^*}, \tag{46}$$

where h^* is the index of the maximum value. Recall that we defined μ_h as the adversary's advantage in recovering the binary $\langle V, h\rangle$ from $\mathsf{L}(V)$, and from Lemma 1 and 2, the corresponding $\epsilon_{\min}$ for a binary RV $\langle V, h\rangle$ is twice its advantage, i.e.,

$$\epsilon^h_{\min} = 2\mu_h. \tag{47}$$

The systems for $\langle X, h\rangle$ and X are the same. Therefore, we assume that the adversary's advantage in attacking the system for X is less than their advantage in attacking the subsystem defined by $\langle X, h^*\rangle$, which is supplied with a twice-informative leakage (in the advantage metric)—that is, a leakage with a binary advantage of $2\mu_{h^*}$. We denote this hypothetical leakage as $\mathsf{L}'(\langle V, h^*\rangle)$. This assumption is valid, at least in cases where the distribution of non-correct

elements is uniform, such as with $\mathsf{ZV}(V)$ leakage (defined in (8)). Consequently, we can write:

$$\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n] \leq \mathsf{Adv}_{\langle X,h^*\rangle}[\boldsymbol{l}'_n \leftarrow \boldsymbol{L}'_n]. \tag{48}$$

By applying the noisy-to-random probing reduction to the binary $\langle V_i, h^*\rangle$ intermediates, as in (39), we obtain:

$$\mathsf{Adv}_{\langle X,h^*\rangle}[\boldsymbol{l}'_n \leftarrow \boldsymbol{L}'_n] \leq \mathsf{Adv}(2, n, 2\epsilon^{h^*}_{\min}), \tag{49}$$

where $\epsilon^{h^*}_{\min} = 2\mu_{h^*}$. Putting it all together, we have:

$$\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n] \leq \mathsf{Adv}(2, n, 4\mu_{h^*}). \tag{50}$$

Example 5. To illustrate the usefulness of the upper bound in (50), we apply it to the leakage function $\mathsf{ZV}(V)$ and the $\mathbb{S}$R-SNI gadget. From Example 4, for $\mathsf{ZV}(V)$, we know that $\mu_h = \frac{1}{2^u}$, and from (40), we have:

$$\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n] \leq \mathsf{Adv}(2, n, 4\mu_{h^*}) \leq \frac{1}{2}\left(\frac{1}{2^{u-2}}\right)^{0.6n},$$

which is valid if $4\mu_{h^*} < 0.15$, a condition satisfied when $u \geq 5$.

Without the decomposition approach, we had $\epsilon_{\min} = 1$, and the bound for $\mathsf{Adv}_X[\boldsymbol{l}_n \leftarrow \boldsymbol{L}_n]$ was trivial. □

We leave the question of under what broader conditions the assumption made in (48) holds for future research.

5 Conclusion

In this work, we identified the necessary and sufficient noise requirements for ensuring the security of masked encodings in binary extended fields. Our findings show that the leakage must not reveal *any* linear combination of bits of an intermediate value, thereby resolving a longstanding open question regarding the minimum noise needed for secure masking. This result is especially relevant in high-SNR settings, where state-of-the-art noise assumptions are too restrictive.

We further demonstrated the applicability of our decomposition approach for analyzing the security of masked gadgets and circuits, focusing on linear gadgets. By decomposing into binary subfields, we showed that security metrics can be efficiently computed, enabling more precise noise thresholds for secure implementations.

While our analysis concentrated on linear gadgets, extending this framework to a broader range of protected circuits remains an open problem. In future work, we plan to investigate non-linear gadgets and complete circuits, pushing the limits of noise-based side-channel countermeasures.

References

1. Barthe, G., Belaïd, S., Dupressoir, F., Fouque, P.A., Grégoire, B., Strub, P.Y., Zucchini, R.: Strong Non-Interference and Type-Directed Higher-Order Masking. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. pp. 116–129 (2016)
2. Batina, L., Gierlichs, B., Prouff, E., Rivain, M., Standaert, F., Veyrat-Charvillon, N.: Mutual information analysis: a comprehensive study. J. Cryptol. **24**(2), 269–291 (2011). https://doi.org/10.1007/S00145-010-9084-8, https://doi.org/10.1007/s00145-010-9084-8
3. Battistello, A., Coron, J.S., Prouff, E., Zeitoun, R.: Horizontal side-channel attacks and countermeasures on the isw masking scheme. In: Gierlichs, B., Poschmann, A.Y. (eds.) Cryptographic Hardware and Embedded Systems – CHES 2016. pp. 23–39. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)
4. Béguinot, J., Cheng, W., Guilley, S., Liu, Y., Masure, L., Rioul, O., Standaert, F.: Removing the Field Size Loss from Duc et al.'s Conjectured Bound for Masked Encodings. In: Kavun, E.B., Pehl, M. (eds.) Constructive Side-Channel Analysis and Secure Design - 14th International Workshop, COSADE 2023, Munich, Germany, April 3-4, 2023, Proceedings. Lecture Notes in Computer Science, vol. 13979, pp. 86–104. Springer (2023). https://doi.org/10.1007/978-3-031-29497-6"˙5, https://doi.org/10.1007/978-3-031-29497-6_5
5. Béguinot, J., Cheng, W., Guilley, S., Rioul, O.: Formal Security Proofs via Doeblin Coefficients: - Optimal Side-Channel Factorization from Noisy Leakage to Random Probing. In: Reyzin, L., Stebila, D. (eds.) Advances in Cryptology - CRYPTO 2024 - 44th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2024, Proceedings, Part VI. Lecture Notes in Computer Science, vol. 14925, pp. 389–426. Springer (2024). https://doi.org/10.1007/978-3-031-68391-6"˙12, https://doi.org/10.1007/978-3-031-68391-6_12
6. Bronchain, O., Standaert, F.: 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). https://doi.org/10.46586/TCHES.V2021.I3.202-234
7. Chari, S., Jutla, C.S., Rao, J.R., Rohatgi, P.: Towards sound approaches to counteract power-analysis attacks. In: Wiener, M. (ed.) Advances in Cryptology — CRYPTO' 99. pp. 398–412. Springer Berlin Heidelberg, Berlin, Heidelberg (1999)
8. Cover, T.M., Thomas, J.A.: Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing). Wiley-Interscience (July 2006)
9. Duc, A., Dziembowski, S., Faust, S.: Unifying leakage models: From probing attacks to noisy leakage. J. Cryptol. **32**(1), 151–177 (2019). https://doi.org/10.1007/S00145-018-9284-1, https://doi.org/10.1007/s00145-018-9284-1
10. Duc, A., Faust, S., Standaert, F.: Making Masking Security Proofs Concrete (Or How to Evaluate the Security of Any Leaking Device), Extended Version. J. Cryptol. **32**(4), 1263–1297 (2019). https://doi.org/10.1007/S00145-018-9277-0, https://doi.org/10.1007/s00145-018-9277-0
11. Durvaux, F., Standaert, F., Veyrat-Charvillon, N.: How to certify the leakage of a chip? In: Nguyen, P.Q., Oswald, E. (eds.) Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark,

May 11-15, 2014. Proceedings. Lecture Notes in Computer Science, vol. 8441, pp. 459–476. Springer (2014). https://doi.org/10.1007/978-3-642-55220-5“¨26, https://doi.org/10.1007/978-3-642-55220-5_26

12. Dziembowski, S., Faust, S., Skorski, M.: Noisy leakage revisited. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology - EUROCRYPT 2015. pp. 159–188. Springer Berlin Heidelberg, Berlin, Heidelberg (2015)
13. Dziembowski, S., Faust, S., Skórski, M.: Optimal Amplification of Noisy Leakages. In: Kushilevitz, E., Malkin, T. (eds.) Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part II. Lecture Notes in Computer Science, vol. 9563, pp. 291–318. Springer (2016). https://doi.org/10.1007/978-3-662-49099-0“¨11, https://doi.org/10.1007/978-3-662-49099-0_11
14. Faust, S., Masure, L., Micheli, E., Orlt, M., Standaert, F.: Connecting Leakage-Resilient Secret Sharing to Practice: Scaling Trends and Physical Dependencies of Prime Field Masking. In: Joye, M., Leander, G. (eds.) Advances in Cryptology - EUROCRYPT 2024 - 43rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zurich, Switzerland, May 26-30, 2024, Proceedings, Part IV. Lecture Notes in Computer Science, vol. 14654, pp. 316–344. Springer (2024). https://doi.org/10.1007/978-3-031-58737-5“¨12, https://doi.org/10.1007/978-3-031-58737-5_12
15. Fedotov, A., Harremoes, P., Topsoe, F.: Refinements of pinsker's inequality. IEEE Transactions on Information Theory **49**(6), 1491–1498 (2003). https://doi.org/10.1109/TIT.2003.811927
16. Gierlichs, B., Batina, L., Tuyls, P., Preneel, B.: Mutual information analysis. In: Oswald, E., Rohatgi, P. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2008, 10th International Workshop, Washington, D.C., USA, August 10-13, 2008. Proceedings. Lecture Notes in Computer Science, vol. 5154, pp. 426–442. Springer (2008). https://doi.org/10.1007/978-3-540-85053-3“¨27, https://doi.org/10.1007/978-3-540-85053-3_27
17. Grassi, L., Masure, L., Méaux, P., Moos, T., Standaert, F.: Generalized Feistel Ciphers for Efficient Prime Field Masking. In: Joye, M., Leander, G. (eds.) Advances in Cryptology - EUROCRYPT 2024 - 43rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zurich, Switzerland, May 26-30, 2024, Proceedings, Part III. Lecture Notes in Computer Science, vol. 14653, pp. 188–220. Springer (2024). https://doi.org/10.1007/978-3-031-58734-4“¨7, https://doi.org/10.1007/978-3-031-58734-4_7
18. Guo, Q., Grosso, V., Standaert, F., Bronchain, O.: Modeling soft analytical side-channel attacks from a coding theory viewpoint. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2020**(4), 209–238 (2020). https://doi.org/10.13154/TCHES.V2020.I4.209-238, https://doi.org/10.13154/tches.v2020.i4.209-238
19. Heuser, A., Rioul, O., Guilley, S.: Good is not good enough. In: Batina, L., Robshaw, M. (eds.) Cryptographic Hardware and Embedded Systems – CHES 2014. pp. 55–74. Springer Berlin Heidelberg, Berlin, Heidelberg (2014)
20. Ito, A., Ueno, R., Homma, N.: On the Success Rate of Side-Channel Attacks on Masked Implementations: Information-Theoretical Bounds and Their Practical Usage. In: Yin, H., Stavrou, A., Cremers, C., Shi, E. (eds.) Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11,

2022. pp. 1521–1535. ACM (2022). https://doi.org/10.1145/3548606.3560579, https://doi.org/10.1145/3548606.3560579
21. Jahandideh, V., Mennink, B., Batina, L.: An Algebraic Approach for Evaluating Random Probing Security With Application to AES. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2024**(4), 657–689 (2024). https://doi.org/10.46586/TCHES.V2024.I4.657-689
22. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks: Revealing the Secrets of Smart Cards (Advances in Information Security). Springer-Verlag, Berlin, Heidelberg (2007)
23. Masure, L., Rioul, O., Standaert, F.: A Nearly Tight Proof of Duc et al.'s Conjectured Security Bound for Masked Implementations. In: Buhan, I., Schneider, T. (eds.) Smart Card Research and Advanced Applications - 21st International Conference, CARDIS 2022, Birmingham, UK, November 7-9, 2022, Revised Selected Papers. Lecture Notes in Computer Science, vol. 13820, pp. 69–81. Springer (2022). https://doi.org/10.1007/978-3-031-25319-5"˙4, https://doi.org/10.1007/978-3-031-25319-5_4
24. Masure, L., Standaert, F.: Prouff and Rivain's Formal Security Proof of Masking, Revisited - Tight Bounds in the Noisy Leakage Model. In: Handschuh, H., Lysyanskaya, A. (eds.) Advances in Cryptology - CRYPTO 2023 - 43rd Annual International Cryptology Conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings, Part III. Lecture Notes in Computer Science, vol. 14083, pp. 343–376. Springer (2023). https://doi.org/10.1007/978-3-031-38548-3"˙12, https://doi.org/10.1007/978-3-031-38548-3_12
25. Moos, T.: 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). https://doi.org/10.13154/TCHES.V2019.I3.202-232, https://doi.org/10.13154/tches.v2019.i3.202-232
26. Obresmki, M., Ribeiro, J., Roy, L., Standaert, F.X., Venturi, D.: Improved Reductions from Noisy to Bounded and Probing Leakages via Hockey-Stick Divergences. In: Reyzin, L., Stebila, D. (eds.) Advances in Cryptology – CRYPTO 2024. pp. 461–491. Springer Nature Switzerland, Cham (2024)
27. Prest, T., Goudarzi, D., Martinelli, A., Passelègue, A.: Unifying Leakage Models on a Rényi Day. In: Boldyreva, A., Micciancio, D. (eds.) Advances in Cryptology - CRYPTO 2019 - 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2019, Proceedings, Part I. Lecture Notes in Computer Science, vol. 11692, pp. 683–712. Springer (2019). https://doi.org/10.1007/978-3-030-26948-7"˙24, https://doi.org/10.1007/978-3-030-26948-7_24
28. Prouff, E., Rivain, M.: Masking against side-channel attacks: A formal security proof. In: Johansson, T., Nguyen, P.Q. (eds.) Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings. Lecture Notes in Computer Science, vol. 7881, pp. 142–159. Springer (2013). https://doi.org/10.1007/978-3-642-38348-9"˙9, https://doi.org/10.1007/978-3-642-38348-9_9
29. Renauld, M., Standaert, F., Veyrat-Charvillon, N., Kamel, D., Flandre, D.: A formal study of power variability issues and side-channel attacks for nanoscale devices. In: Paterson, K.G. (ed.) Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings. Lecture Notes in Computer Science, vol. 6632,

pp. 109–128. Springer (2011). https://doi.org/10.1007/978-3-642-20465-4“¨8, https://doi.org/10.1007/978-3-642-20465-4_8
30. Standaert, F.X., Malkin, T.G., Yung, M.: A unified framework for the analysis of side-channel key recovery attacks. In: Joux, A. (ed.) Advances in Cryptology - EUROCRYPT 2009. pp. 443–461. Springer Berlin Heidelberg, Berlin, Heidelberg (2009)
31. Stromberg, K.: Probabilities on a compact group. Transactions of the American Mathematical Society **94**(2), 295–309 (1960), http://www.jstor.org/stable/1993313
32. Wyner, A.D.: The wire-tap channel. The Bell System Technical Journal **54**(8), 1355–1387 (1975). https://doi.org/10.1002/j.1538-7305.1975.tb02040.x