Leakage Bounds for Gaussian Side Channels

Thomas Unterluggauer[1], Thomas Korak[1], Stefan Mangard[1], Robert Schilling[1], Luca Benini[2], Frank K. Gürkaynak[2], and Michael Muehlberghuber[2]

[1] Graz University of Technology, Austria
{firstname.lastname}@iaik.tugraz.at
[2] Integrated Systems Laboratory, ETH Zürich, Switzerland

Abstract. In recent years, many leakage-resilient schemes have been published. These schemes guarantee security against side-channel attacks given bounded leakage of the underlying primitive. However, it is a challenging task to reliably determine these leakage bounds from physical properties.
In this work, we present a novel approach to find reliable leakage bounds for side channels of cryptographic implementations when the input data complexity is limited such as in leakage-resilient schemes. By mapping results from communication theory to the side-channel domain, we show that the channel capacity is the natural upper bound for the mutual information (MI) to be learned from multivariate side-channels with Gaussian noise. It shows that this upper bound is determined by the device-specific signal-to-noise ratio (SNR). We further investigate the case when attackers are capable of measuring the same side-channel leakage multiple times and perform signal averaging. Our results here indicate that the gain in the SNR obtained from averaging is exponential in the number of points of interest that are used from the leakage traces. Based on this, we illustrate how the side-channel capacity gives a tool to compute the minimum attack complexity to learn a certain amount of information from side-channel leakage. We then show that our MI bounds match with reality by evaluating the MI in multivariate Gaussian templates built from practical measurements on an ASIC. We finally use our model to show the security of the KECCAK-f[400]-based authenticated encryption scheme ISAP on this ASIC against power analysis attacks.

Keywords: leakage-resilient cryptography, leakage model, mutual information, channel capacity

1 Introduction

Side-channel attacks are a serious threat to cryptographic implementations as they allow attackers to learn secret information processed inside a device from observing its physical behavior, e.g., the power consumption. In order to protect implementations from such attacks, one approach actively being researched for several years now is leakage-resilient cryptography. Leakage-resilient schemes are designed such as to resist a certain amount of side-channel leakage.

This means that if every invocation of the underlying primitive leaks λ bits of information, leakage-resilient schemes guarantee that their overall leakage stays within predefined bounds. In addition, the side-channel leakage λ is commonly bounded by limiting the input data complexity to the internal primitives, as for example in leakage-resilient encryption using the 2PRG primitive [15]. However, it is an ongoing topic of research to specify concrete leakage bounds λ based on the implementation and its physical properties.

For example, Medwed et al. [9] evaluated a set of practical differential power analysis (DPA) attacks on simulated leakages from parallel implementations with unknown in- and outputs. Their resulting success probabilities indicate that even for identity leakage of the secret state, its exploitation is practically hard once enough processes happen in parallel. While their specific results also suggest security for limited data complexities, it is hard to derive a concrete leakage bound λ in bits. On the other hand, Standaert et al. [14] suggested using the mutual information (MI) from information theory as a general tool to concretely state the amount of information learned from side-channel leakage in bits. While the MI can only be exactly computed once the actual leakage distribution of an implementation is known, Duc et al. [4] mention an upper bound for the MI for univariate leakages that solely depends on the device- and measurement-specific signal-to-noise ratio (SNR). It, however, remains unclear how this bound scales for multivariate leakages that are exploited in practice.

For a single measurement of the side-channel leakage, physical constraints such as the SNR will typically bound the MI to suit leakage-resilient schemes. While most of these schemes indeed confine the attacker to a single measurement by requiring a fresh initial state on every invocation, there are also schemes allowing attackers to observe the same execution using the same data multiple times, e.g., as for multiple decryptions in ISAP [3]. However, multiple measurements of the same decryption process allow an attacker to perform signal averaging to increase the SNR. This can allow unbounded side-channel attackers to distinguish tiny variances in the signal to learn the complete secret state. However, in practice, side-channel attackers are bounded by physical and computational resources. This gives the interesting question of the actual attack complexity when the side-channel attacker is capable of observing the same execution multiple times and performing signal averaging.

Our Contribution. In this work, we present a new approach to give reliable upper bounds for the leakage from side channels of cryptographic implementations under a single data input. For this purpose, we map results from communication theory to the side-channel domain. In particular, we show that the channel capacity of n-to-m communication channels is the natural upper bound for the MI in multivariate side-channel leakages with Gaussian noise. Without any further leakage assumptions, we show that this bound depends on a device- and measurement-specific SNR that is uniquely determined by the device's statistical leakage behavior in the m points of interest (POIs) in the leakage trace. In a second step, we investigate the effect of signal averaging on this SNR and show that averaging N leakage traces increases the SNR by a factor N^m. Our results

provide both attackers and implementers with a tool for computing the expected minimum attack complexity, i.e., the number of leakage traces required to learn a certain amount of the processed state from side-channel information. We then show that our model and results fit the reality by evaluating the MI in multivariate Gaussian templates. For this purpose, we used power measurements from a real system on chip (SoC) that features a KECCAK-f[400] engine that computes three rounds per cycle. Last, we use our model to demonstrate the security of the scheme ISAP implemented on this SoC w.r.t. power analysis attacks.

Outline. This paper is organized as follows. Section 2 gives bounds for the information leakage of multivariate side channels with Gaussian noise. We analyze the case of signal averaging and provide a tool to compute the expected minimum attack complexity for side-channel attackers in Section 3. The soundness of our leakage model is shown in Section 4 based on power measurements of an ASIC, and we finally conclude in Section 5.

2 Modeling Side-Channel Leakage as a Communication Channel

In this section, we consider the case of leakage-resilient cryptography where an attacker can use the side-channel information in a single leakage trace to learn the secret state of a device. In particular, we adapt the results from communication theory to fit side-channel leakages and use the channel capacity of n-to-m wireless channels to give a leakage upper bound for multivariate side channels with Gaussian noise independent of the underlying leakage function.

2.1 Attack Model

We consider an attacker trying to recover the secret state x from a single leakage trace l_x generated by an implementation $\mathcal{I}$ with input complexity $q = 1$. This implies that the attacker is unable to perform multi-input attacks such as DPA. Moreover, attackers are allowed to observe the operation using the secret state x only a single time, i.e., they are not allowed to average traces to improve their SNR. However, we consider a profiled attack setting, i.e., the attacker has the opportunity to build templates before performing the actual attack.

2.2 Mutual Information

A common metric to assess the amount of information about a secret x contained in the leakage l_x is the mutual information (MI) [4,14]. We therefore introduce the random variables X and L_x to denote the distributions of x and l_x, respectively. The mutual information is then defined as

$$MI(X; L_x) = H[X] - H[X|L_x]. \tag{1}$$

Hereby, $H[X]$ and $H[X|L_x]$ denote the entropy of the random variable X and the conditional entropy of X given the leakage L_x, respectively. Note however

that the (conditional) entropy (and thus the MI) is an average metric depending on the actual distribution of values $x_i \in X$ and $l_x \in L_x$. This means that the actual information learned from a side-channel leakage depends on the actually processed value and might thus for some events even be higher than the MI. Yet, the MI is a good metric to give bounds on the expected leakage behavior.

2.3 Linear Channel Model

For giving bounds on the MI of side channels, we consider an implementation that transmits the single bits of a secret state to the attacker via a side channel. Hereby, the physical leakage behavior and measurement effects define the mapping of the single bits to the output samples of the side channel. We model this multivariate side channel as an n-to-m linear communication channel with Gaussian noise, i.e., it transfers linear combinations of the bits of the secret state. While this linear channel model allows to adapt results from communication theory, the resulting bounds are yet independent from the concrete leakage behavior and Gaussian noise is the sole assumption. Namely, our final bounds will only depend on the side-channel signal observed by the attacker. Further note that non-linear mappings can easily be added to this model similar as for regression techniques [13].

In our linear channel model, the attacker observes an $m \times 1$ leakage trace $\mathbf{l}_x$ from the processing of the secret state x in the implementation $\mathcal{I}$. Let $\mathbf{x}$ denote the $n \times 1$ vector consisting of the n bits of the secret state x. We then model the leakage trace $\mathbf{l}_x$ as the multiplication of the secret state vector $\mathbf{x}$ with a $m \times n$ side-channel matrix $\mathbf{H}$ plus an $m \times 1$ noise vector $\boldsymbol{\nu}$:

$$\mathbf{l}_x = \mathbf{H}\mathbf{x} + \boldsymbol{\nu}. \tag{2}$$

The i-th row of $\mathbf{H}$ specifies how the n bits of the secret state x map to the i-th point of the measured leakage $\mathbf{l}_\mathbf{x}$. The maximum MI that an attacker can learn from the side-channel leakage according to Eq. 2 depends on the maximum number of states that are distinguishable at the receiver of this channel. This upper bound on the MI is typically called the channel capacity. In particular, Telatar [16] states the channel capacity C as the maximum average mutual information between in- and output over the choice of the input distribution, i.e.,

$$C = \max_{p(X)} \; MI(X, L_x). \tag{3}$$

We observe that the side-channel leakage given by Eq. 2 bears some familiarity with the notion of multi-input multi-output (MIMO) channels as used in wireless communication. For a constant, known channel $\mathbf{H}$, Goldsmith et al. [7] state the channel capacity for signals in the domain of complex numbers as follows:

$$C = \max_{\Sigma_\mathbf{x}: tr(\Sigma_\mathbf{x})=P} \log_2 |\mathbf{I}_m + \mathbf{H}\Sigma_\mathbf{x}\mathbf{H}^H| \tag{4}$$

Hereby, $\mathbf{I}_m$ and $\Sigma_{\mathbf{x}}$ denote the $m \times m$ identity matrix and $n \times n$ signal covariance matrix, respectively. P is the total power constraint of the transmitter, $\mathbf{H}^H$ the complex conjugate of $\mathbf{H}$, $|\cdot|$ the determinant, and $tr(\cdot)$ the trace of a matrix. For Eq. 2 to hold true, the noise vector $\boldsymbol{\nu}$ must consist of independent samples of Gaussian white noise with variance $\sigma_\nu^2 = 1$, i.e., the $m \times m$ noise covariance matrix $\Sigma_{\boldsymbol{\nu}}$ is the identity matrix $\mathbf{I}_m$.

We can use the channel capacity of MIMO channels as an upper bound for the MI in side-channel leakages according to Eq. 2. However, there are different constraints for side channels than in wireless communication, requiring some modifications of Eq. 4. For example, an attacker cannot influence the signal covariance $\Sigma_{\mathbf{x}}$ such as to optimize the capacity C. Moreover, side-channel attacks typically exploit real-valued information like the power consumption, whereas signals in communication channels are represented in the domain of complex numbers. This effectively halves the capacity for the side-channel case. In practice, we also observe that the samples in the noise vector $\boldsymbol{\nu}$ are not necessarily independent and have different variances. According to [7], dependent samples in the noise $\boldsymbol{\nu}$ can be modeled via a modified channel matrix $\tilde{\mathbf{H}} = \Sigma_{\boldsymbol{\nu}}^{-1/2}\mathbf{H}$ given the noise covariance matrix $\Sigma_{\boldsymbol{\nu}}$. By adapting Eq. 4 according to these considerations, we extract the special case of linear side channels as in Eq. 2 and state their leakage upper bound:

$$C = \max_{p(X)} \; MI(X, L_x) = \frac{1}{2} \log_2 |\mathbf{I}_m + \Sigma_{\boldsymbol{\nu}}^{-1}\mathbf{H}\Sigma_{\mathbf{x}}\mathbf{H}^H|. \tag{5}$$

2.4 Leakage Bound for Gaussian Side Channels

The side-channel matrix $\mathbf{H}$ is typically unknown but fixed. An interesting question thus is how to determine the channel capacity if $\mathbf{H}$ is unknown. A common approach to characterize a side channel are multivariate Gaussian templates. Hereby, for each secret state $\mathbf{x}$, the respective side-channel leakage $\mathbf{l}_{\mathbf{x}}$ is described as a multivariate Gaussian distribution. This characterization gives a set of templates $(\mu_i, \Sigma_{\boldsymbol{\nu},i})$ with mean μ_i and noise covariance $\Sigma_{\boldsymbol{\nu},i}$ for all states $\mathbf{x}_i$. The means μ_i give an estimation of the $n \times n$ covariance matrix $\Sigma_{\mathbf{y}}$ of the side-channel signal $\mathbf{y} = \mathbf{H}\mathbf{x}$. This covariance matrix $\Sigma_{\mathbf{y}}$ equals $\mathbf{H}\Sigma_{\mathbf{x}}\mathbf{H}^H$ from Eq. 5. Similarly, assuming that the noise is independent from the signal and thus has constant covariance (as in [11]), the single noise covariances $\Sigma_{\boldsymbol{\nu},i}$ give an estimation of $\Sigma_{\boldsymbol{\nu}}$.[3] Putting this together, we adapt Eq. 5 to derive our main result. Namely, we use the signal and noise covariance matrices $\Sigma_{\mathbf{y}}, \Sigma_{\boldsymbol{\nu}}$ to state the capacity of a side channel characterized via multivariate Gaussian templates, or more generally, of multivariate leakages with Gaussian noise.

[3] The constant covariance assumption is invalid in case the covariance carries information as, e.g., in masked implementations. However, leakage-resilient cryptography aims to bound the leakage without the use of countermeasures like masking, and thus noise will typically be independent from the signal.

Main Result. (Leakage Bound of a Gaussian Side Channel) *The mutual information of a multivariate side channel with signal covariance $\Sigma_{\mathbf{y}}$ and Gaussian noise $\Sigma_{\boldsymbol{\nu}}$ is bounded by*

$$C = \frac{1}{2} \log_2 |\mathbf{I}_m + \Sigma_{\boldsymbol{\nu}}^{-1} \Sigma_{\mathbf{y}}|. \tag{6}$$

Interestingly, the term $\Sigma_{\boldsymbol{\nu}}^{-1} \Sigma_{\mathbf{y}}$ is an SNR taking noise and signal covariances between the POIs into account. The capacity of the side channel is thus determined by the actual power of signal and noise, and correlations in the samples of $\boldsymbol{\nu}$ and $\mathbf{y}$. Such correlations typically mark redundancies that effectively reduce the side-channel capacity. Moreover, note that the side-channel capacity given here depends on the side-channel signal $\mathbf{y}$ only. This means that our result applies to any leakage function/model having the properties given by $\Sigma_{\mathbf{y}}$.

For univariate leakages or when the same leakage is observed in multiple POIs within the leakage trace, the leakage bound in Eq. 6 can further be simplified.

Univariate Leakage. An attacker exploiting univariate leakage is confined to the leakage in a single point of the execution of an implementation $\mathcal{I}$. This means that the side channel degenerates to

$$l_{\mathbf{x}} = \mathbf{h}\mathbf{x} + \nu, \tag{7}$$

where $l_{\mathbf{x}}$ and ν are scalars and the $1 \times n$ channel vector $\mathbf{h}$ specifies the leakage of the single bits of the state $\mathbf{x}$. Let us now assume the channel vector $\mathbf{h}$ maps the n bits in $\mathbf{x}$ to y according to the identity of the respective state variable x. Intuitively, the MI between the secret state $\mathbf{x}$ and its leakage $l_{\mathbf{x}}$ is here bounded by the number of different states that an attacker can distinguish in the single leakage point $l_{\mathbf{x}}$. This number depends both on the distance between the different states along the measured signal range and the noise. When adapting Eq. 6 for univariate leakage, we can observe exactly this dependence:

$$C = \frac{1}{2} \log_2 \left(1 + \frac{\sigma_y^2}{\sigma_\nu^2}\right) = \frac{1}{2} \log_2 \left(1 + SNR\right), \tag{8}$$

where σ_y^2 is the variance of the signal $y = \mathbf{h}\mathbf{x}$ and σ_ν^2 is the variance of the noise ν. As also noted in [4, 10], this upper bound for the MI in univariate leakages solely depends on the SNR and is better known as the Shannon-Hartley theorem [2].

Identical Leakage in Multiple Points. In many cases, an attacker will try to exploit the leakage in multiple points of the execution to increase their success rate. If these points are chosen to be in close vicinity within the leakage trace, these POIs will often carry highly redundant information. An example where this case occurs are attackers sampling the side channel at a very high rate and using several consecutive sampling points in their attack. In such situation, one can assume the

leakage to be the same for all points of the leakage trace. This case is equivalent to single-input multiple-output (SIMO) channels in wireless communication. The side-channel matrix is then expressed as the vector multiplication $\mathbf{H} = \mathbf{h}_{gain} \cdot \mathbf{h}_l$, where $\mathbf{h}_l$ states the $1 \times n$ side-channel vector mapping the n bits of x to a scalar value and $\mathbf{h}_{gain}$ is the $m \times 1$ gain vector over the m POIs used by the attacker. The capacity formula in Eq. 5 degenerates for such leakage behavior, but can simply be expressed using the vector $\mathbf{h}_{gain}$ only [6]:

$$C = \frac{1}{2} \log_2 \left(1 + \sigma_z^2 \mathbf{h}_{gain}^H \Sigma_{\boldsymbol{\nu}}^{-1} \mathbf{h}_{gain}\right), \tag{9}$$

where σ_z^2 is the variance of the signal $z = \mathbf{h}_l \mathbf{x}$ such that $\mathbf{l}_\mathbf{x} = \mathbf{h}_{gain} z + \boldsymbol{\nu}$.

2.5 Description of Common Leakage Models

Our leakage model in Eq. 2 allows to easily describe linear side-channel leakages. We now give several examples on how to map existing power models to Eq. 2. Note that we give these examples without consideration of the effective signal range in the leakage $\mathbf{l}_\mathbf{x}$.

Identity Leakage. In identity leakage, the n-bit secret state $\mathbf{x}$ leaks linear to the value x it represents. If $\mathbf{x}$ leaks the identity in the i-th sample of $\mathbf{l}_\mathbf{x}$, the i-th row in the side-channel matrix $\mathbf{H}$ takes the form $\mathbf{h} = \left(2^0\ 2^1\ 2^2\ \ldots\ 2^{n-2}\ 2^{n-1}\right)$.

Hamming Weight Leakage. In Hamming Weight (HW) leakage, the secret state $\mathbf{x}$ leaks the number of bits set to one. HW leakage in the i-th sample of $\mathbf{l}_\mathbf{x}$ results in the i-th row of $\mathbf{H}$ to take the form $\mathbf{h} = \left(1\ 1\ 1\ \ldots\ 1\ 1\right)$. Hamming Distance (HD) leakage is modeled in the same way by setting the secret x to be the xor of the leaking state before and after it toggles.

Time-Serialized Leakage. In time-serialized implementations, an attacker collecting the side-channel leakage at different points in time will be able to learn different information in the different POIs. One prominent example are byte-oriented cryptographic implementations, where in each clock cycle a different byte of the n-bit state $\mathbf{x}$ is processed and leaks. For simplicity, let us assume an 8-bit state and HW leakage of a 2-bit chunk processed in the respective clock cycle. This will give a side-channel matrix of the form

$$\mathbf{H} = \begin{pmatrix} 1\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1\,1\,0\,0 \\ 0\,0\,0\,0\,0\,0\,1\,1 \end{pmatrix}.$$

Localized Leakage. Localized electromagnetic emanation (EM) attacks are a powerful way to extract information from parts of the secret state. Such localized EM leakage can in principal be modeled the same way as time-varying leakage. For example, consider the leakages $\mathbf{l}_{\mathbf{x},1}$ and $\mathbf{l}_{\mathbf{x},2}$ observed in two different EM positions. Moreover, assume that $\mathbf{l}_{\mathbf{x},1}, \mathbf{l}_{\mathbf{x},2}$ consist each of two samples leaking the identity of the first or second half of a 4-bit state, respectively. Concatenating the

two leakages $\mathbf{l}_{\mathbf{x}}^T = (\mathbf{l}_{\mathbf{x},1}^T \mathbf{l}_{\mathbf{x},2}^T)$ means concatenating the respective channel matrices $\mathbf{H}_1, \mathbf{H}_2$ to a combined side-channel matrix of the form

$$\mathbf{H} = \begin{pmatrix} 2^0 & 2^1 & 0 & 0 \\ 2^0 & 2^1 & 0 & 0 \\ 0 & 0 & 2^0 & 2^1 \\ 0 & 0 & 2^0 & 2^1 \end{pmatrix}.$$

This model underlines the intuition that gathering additional leakage from observing a parallel implementation in different locations and measuring a serial implementation at different times is equivalent. In particular, it shows that side-channel leakage becomes optimal if the leakages in the side-channel signal $\mathbf{y} = \mathbf{Hx}$ are independent. In the best case, the signal covariance matrix becomes a diagonal matrix, i.e, $\Sigma_{\mathbf{y}} = diag(\sigma_{y_1}^2, \sigma_{y_2}^2, ..., \sigma_{y_m}^2)$. In the same way, noise effects are canceled out the best if the noise samples in $\boldsymbol{\nu}$ are independent, i.e., $\Sigma_{\boldsymbol{\nu}}$ is a diagonal matrix as well.

3 Complexity of State Recovery

The side-channel capacity is an upper bound on the MI to be learned via a side channel. This bound essentially depends on the implementation's SNR. While in most leakage-resilient schemes an attacker is restricted to a single leakage trace for a specific state, there are schemes, e.g., Isap [3], that allow attackers to observe the execution of an implementation processing the same data multiple times. This gives attackers the option to perform signal averaging, which improves the side-channel SNR and thus side-channel capacity.

In this section, we therefore consider an attacker capable of averaging multiple leakage traces. We show how averaging improves the side-channel capacity in multivariate attacks and provide attackers and implementers with a tool to compute the expected minimum complexity to learn the secret state of a device.

3.1 Attack Model

As in Section 2, we assume an attacker trying to recover a secret state x from side-channel leakages l_x generated by an implementation $\mathcal{I}$ with input complexity $q = 1$ and thus preclude multi-input attacks. However, the attacker is capable of observing the same execution of $\mathcal{I}$ multiple times. This attack setting is observed when a ciphertext, e.g., a firmware image, must be decrypted multiple times using a leakage-resilient scheme like Isap.

3.2 Averaging Attacker

An attacker that observes the same processing of the secret state x multiple times is capable of averaging the side-channel leakage l_x to yield a better SNR

and thus higher side-channel capacity. In general, averaging N observations gives the averaged noise covariance matrix

$$\overline{\Sigma}_{\boldsymbol{\nu}} = \frac{1}{N}\Sigma_{\boldsymbol{\nu}}, \tag{10}$$

where $\Sigma_{\boldsymbol{\nu}}$ is the noise covariance matrix valid for a single leakage trace. This means that the noise (co-)variances reduce linearly with the number of averaged traces. Note here that for the univariate case Eq. 10 simplifies to the well-known relation $\overline{\sigma}_{\nu}^2 = \frac{\sigma_{\nu}^2}{N}$. Given the noise covariance matrix after averaging $\overline{\Sigma}_{\boldsymbol{\nu}}$, we can now investigate the effect of averaging on the side-channel capacity. Inserting Eq. 10 into the generic side-channel capacity given in Eq. 4 yields

$$C = \frac{1}{2}\log_2\left|\mathbf{I}_m + N \cdot \Sigma_{\boldsymbol{\nu}}^{-1}\Sigma_{\mathbf{y}}\right|. \tag{11}$$

Note that the SNR term $N \cdot \Sigma_{\boldsymbol{\nu}}^{-1}\Sigma_{\mathbf{y}}$ is an $m \times m$ matrix and its determinant behaves proportionally to N^m. This means that the side-channel capacity increases stronger with the number of averaged traces the more POIs are used in an attack. This is because each POI can potentially transfer completely independent data as, e.g., for time-serialized and localized EM leakages.

Identical Leakage in Multiple Points. For identical leakage in all POIs, the side-channel capacity behaves differently. Inserting Eq. 10 into the SIMO channel capacity given in Eq. 9 yields

$$C = \frac{1}{2}\log_2\left(1 + N \cdot \sigma_z^2 \cdot \mathbf{h}_{gain}^H \Sigma_{\boldsymbol{\nu}}^{-1}\mathbf{h}_{gain}\right). \tag{12}$$

It shows that the number of traces N used for averaging has a linear influence on the SNR and is independent of the number of POIs m.

3.3 Expected Minimum Attack Complexity

In the worst case, physical attackers have unbounded complexity. This means they can measure and average an unlimited number of leakage traces $N \to \infty$, leading to zero noise and virtually unlimited channel capacity and MI. This can be thought of state differences causing vanishingly small differences in the side-channel signal being distinguishable if the noise is eliminated completely. It thus seems reasonable to set the side-channel capacity in relation with the actual attack complexity, i.e., the number of leakage traces N, to learn a certain amount of bits. This is also the common approach when assessing the security of masked implementations.

It is yet difficult to determine such attack complexity since it is strongly influenced by the implementation's leakage behavior, which is commonly unknown. For example, it is unknown to what extent information and noise in the single points of a leakage trace correlate, and as shown in Section 2, these effects strongly influence the side-channel capacity. However, the device- and measurement-specific multivariate $SNR = \Sigma_{\mathbf{y}} \cdot \Sigma_{\boldsymbol{\nu}}^{-1}$ takes exactly these effects into account and can thus

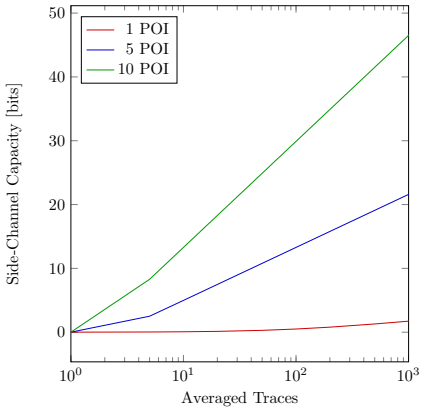


Fig. 1: Expected side-channel capacity given the number of averaged leakage traces for different numbers of POIs and $SNR_m = 0.01$.

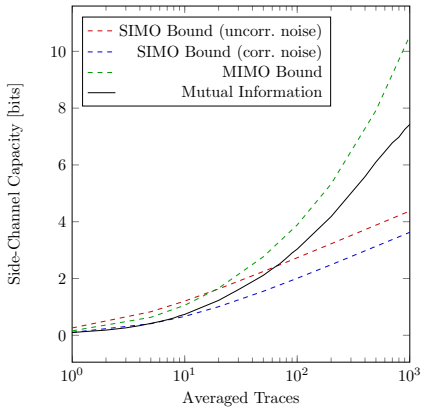


Fig. 2: Mutual information of KECCAK-$f[400]$ on FULMINE and side-channel capacity of different channel models (256 classes, 10 POIs).

be used to generically express the expected minimum complexity of a side-channel attacker without any concrete leakage assumptions. In particular, we can rewrite the multivariate channel capacity for averaging attackers (Eq. 11) as follows:

$$C = \frac{1}{2} \log_2 N^m \left| \frac{1}{N} \mathbf{I}_m + \Sigma_{\boldsymbol{\nu}}^{-1} \Sigma_{\mathbf{y}} \right| . \tag{13}$$

For a large number of averaged traces N, Eq. 13 can be further approximated to give the side-channel capacity in dependence of a scalar device SNR.

$$C \approx \frac{1}{2} \log_2 \left(1 + N^m \left| \Sigma_{\boldsymbol{\nu}}^{-1} \Sigma_{\mathbf{y}} \right| \right) = \frac{1}{2} \log_2 (1 + N^m \cdot SNR_m) \tag{14}$$

An implementation will in practice give some side-channel $SNR_m = |\Sigma_{\mathbf{y}} \cdot \Sigma_{\boldsymbol{\nu}}^{-1}|$ that is observed in m POIs in the leakage traces. This SNR takes into account all kinds of correlations in both noise and side-channel leakage. For an implementation that is expected to give a certain SNR_m, designers and implementers can thus compute the expected minimum attack complexity in terms of traces to measure and average.

Fig. 1 gives an overview on the expected side-channel capacity for $m = 1, 5, 10$ POIs given the number of averaged traces. It shows that the side-channel capacity rises quickly with the number of averaged traces for multivariate leakages. In particular, it shows that if SNR_m is not sufficiently low, a state of virtually any size can theoretically be recovered with practical complexity. However, this effect is also limited by the available POIs with sufficiently low signal correlations.

4 Experimental Verification and Security Analysis

The previous sections introduced theoretical leakage upper bounds for multivariate side channels with Gaussian noise. In this section, we show that these bounds match the real leakage behavior by evaluating the MI on a hardware implementation of the KECCAK-f[400]-based scheme ISAP [3] on the real system on chip FULMINE. Our experiments further show the security of this implementation of ISAP in terms of power analysis attacks.

4.1 Evaluation Hardware: Fulmine

At FSE 2017, Dobraunig et al. [3] presented the sponge-based authenticated encryption scheme ISAP to inherently prevent DPA during both en-/decryption. This is achieved by limiting the number of inputs processed under a single key by one. To further express their scheme's capability to cope with side-channel leakage from a single data input, the authors proposed using the sponge parameters themselves. However, in the view of ISAP allowing for the multiple decryption of the same ciphertext and tag, it is an open question how much information an attacker can learn when averaging multiple leakage traces.

To verify the soundness of our leakage bounds and to evaluate the side-channel resistance of ISAP, we developed and fabricated the multi-core SoC FULMINE, a prototype ASIC in the UMC 65 nm LL 1P8M technology. FULMINE, as shown in Fig. 3, is based on the PULP platform [12] including four general purpose processing elements (enhanced OpenRISC cores with DSP extensions [5, 8]) and two dedicated hardware accelerators: the Hardware Cryptography Engine (HWCRYPT) and the Hardware Convolution Engine (HWCE). All processing elements share the same 64 kB level-1 Tightly-Coupled Data Memory (TCDM) to support a fast and efficient communication and to avoid single point-to-point channels.

HWCRYPT is a flexible, software-programmable hardware accelerator supporting various cryptographic primitive functions such as the KECCAK-f[400] permutation [1]. Moreover, the accelerator supports high-level encryption schemes such as ISAP. The accelerator is designed to achieve maximum throughput. To achieve that goal, the KECCAK-f[400] permutation utilizes three fully parallel round instances to maximize the throughput but to also match the length of the critical path of other parts of the accelerator. When using ISAP, HWCRYPT supports a flexible configuration of the rate (from 1 bit to 128 bits in powers of two) and the number of permutation rounds in multiples of three including 20 to flexibly trade-off between throughput and security. HWCRYPT is configured and monitored via a set of status registers. A flexible event and interrupt system indicates other processing elements when an operation has finished.

4.2 Soundness of Model and Bounds

To verify the soundness of our model and the bounds in Section 2, we analyzed the leakage behavior of the KECCAK-f[400] permutation on FULMINE. For this purpose, we constructed multivariate Gaussian templates for the power consumption of

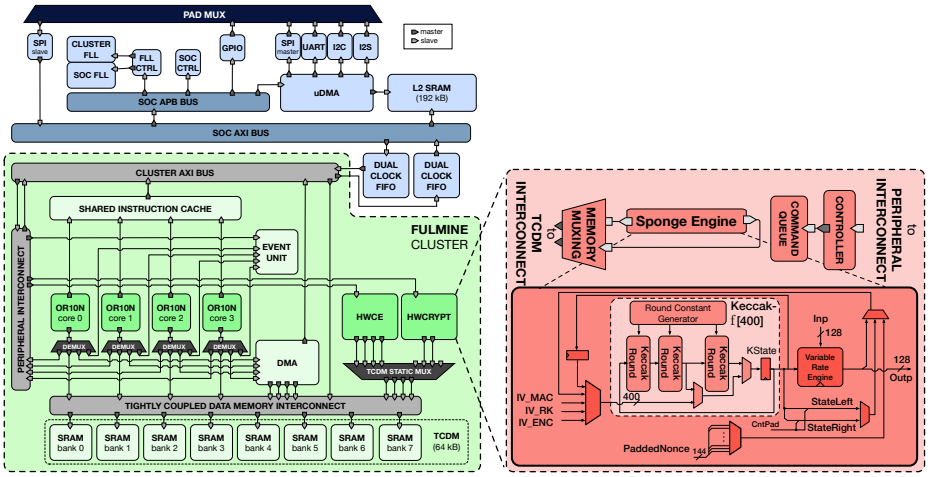
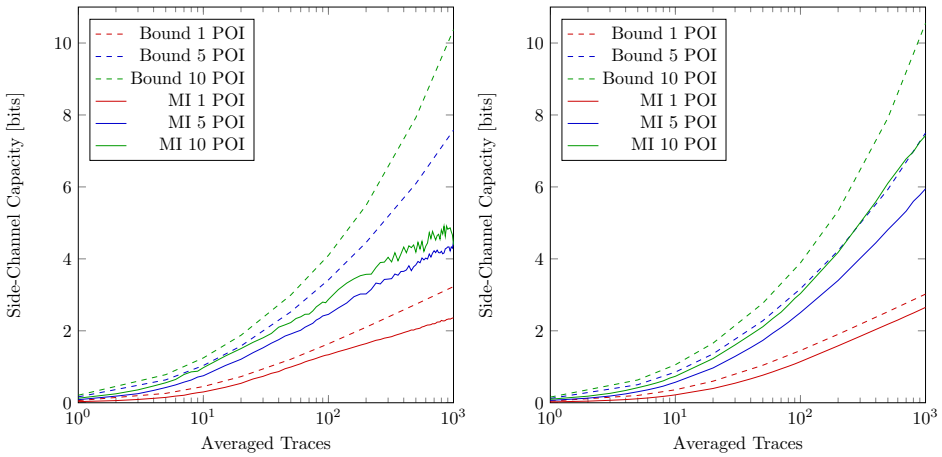


Fig. 3: FULMINE SoC and HWCRYPT architecture.

FULMINE for 5- and 8-bit parts of the 400-bit state of KECCAK-f[400]. More concretely, we target the intermediate state *KState* of KECCAK-f[400], depicted in Fig. 3, such that FULMINE computes three rounds of the permutation before and after the target state to preclude load-time leakages. The remaining state not covered by our templates, i.e., 395 and 392 bits respectively, was held constant. For each class, we used 1400 power measurements in the training phase. The POIs were chosen as the points of highest variance fulfilling a certain minimum distance within the leakage trace and include both register and combinatorial activity. Based on these templates, we computed the side-channel capacity and evaluated both the MI and the 1st-order success rate of classification. The evaluations were done in dependence of the number of leakage traces used for signal averaging.

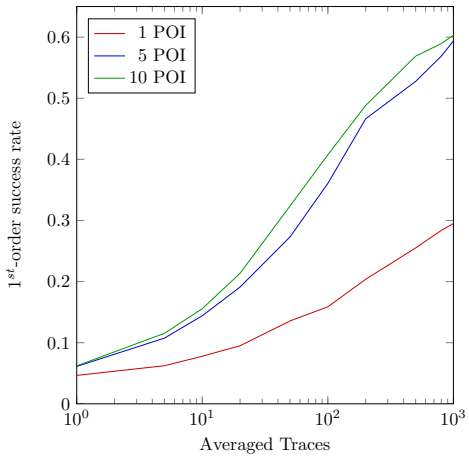
Our evaluation results in Fig. 4 suggest that the channel model used to compute the side-channel capacity of multivariate leakages is sound. In particular, for both 5-bit and 8-bit templates the MI between leakage and secret state stays within the bounds given by the side-channel capacity. While there is a gap between the MI and the channel capacity, the MI follows the shape of the side-channel bound well. Moreover, the first-order classification rate rises accordingly. However, Fig. 4a also shows that for higher numbers of averaged traces the MI goes into saturation, and thus the gap between capacity and the learned information gets bigger. In particular, it shows that once the MI converges to the maximum number of bits that could be recovered using the trained template set, i.e., 5 or 8 bits respectively, the increase in learned information for additional numbers of averaged traces gets successively smaller. This indicates that the side-channel information is not distributed to perfectly use the channel.

We further investigated how different channel models suit the actual leakage behavior. We therefore compared the MIMO channel model used in the previous evaluation with the SIMO channel model, which assumes identical leakages in the POIs of a leakage trace, e.g., within a clock cycle. For the SIMO channel

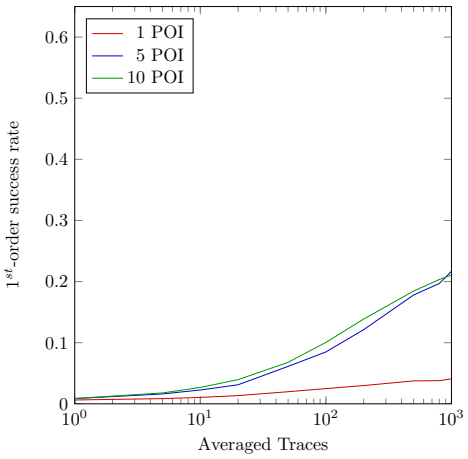


(a) 32 classes.

(b) 256 classes.



(c) 32 classes.



(d) 256 classes.

Fig. 4: Side-channel capacity, mutual information and success rate for the KECCAK-$f[400]$ permutation given the number of averaged traces and different numbers of POIs and number of classes. The remaining state was held constant.

model, we analyzed two cases: one taking noise correlations into account, and one assuming independent noise. The channel capacities of the different channel models were computed based on the 8-bit templates constructed in the previous evaluation. In particular, for the SIMO model we used the signal variance in each POI, but neglected signal covariances. The results of our evaluations are shown in Fig. 2. These suggest that the leakages in the single POIs are not identical and thus the MIMO channel model suits the leakage behavior clearly better than the SIMO channel model. Moreover, from the plots using the SIMO model one can observe that there is some noise correlation that lowers the channel capacity.

4.3 Security of ISAP

In most situations, designers and implementers want to assess the security of a complete cryptographic implementation. However, the state sizes involved in a cryptographic scheme like ISAP are typically large and the channel capacity computed from a low number of templates cannot be directly used since more hardware will be active. On the other hand, it is impossible to build templates for a 400-bit state that would allow to compute the channel capacity exactly. Yet, we can use the experiments on the KECCAK-$f[400]$ permutation to estimate leakage bounds for the full state of ISAP.

As we can see from Fig. 4, the channel capacity is practically the same for both 5- and 8-bit templates. The reason for this is that the SNR we observed on FULMINE using our measurement setup is the same. This gives the question whether and how the SNR would change for 400-bit templates. Now if the same measurement setup was used for constructing 400-bit templates, we can safely say that the range of the measured noise will not decrease by orders of magnitude. In the same way, the range of the side-channel signal will definitely not rise by orders of magnitude using the same setup, especially since the diffusion of three rounds of KECCAK-$f[400]$ already causes large parts of the logic to become active within the profiled clock cycle.

On the other hand, the side-channel capacity from a single power measurement of FULMINE is very low, and thus, even if the channel SNR was 100 times higher, the channel capacity would hardly rise. We thus scale the SNR with a factor γ to get a security margin that allows to estimate how many traces an attacker will at least require to recover the complete state or to exceed the leakage bounds. Using the SNR_m of the m-variate leakage from the 8-bit templates, we compute the minimum number of traces needed to learn the state of size S:

$$N = \left(\frac{2^{2S} - 1}{\gamma \cdot SNR_m} \right)^{1/m} . \tag{15}$$

The authors of ISAP state concrete leakage bounds for their re-keying function and encryption scheme to still provide 128-bit security. We thus evaluated Eq. 15 on FULMINE for three different state sizes: the full state of KECCAK-$f[400]$, the leakage bound for the ISAP re-keying function (272 bits), and the leakage bound for the ISAP encryption itself (128 bits). The results in Fig. 5 indicate that the

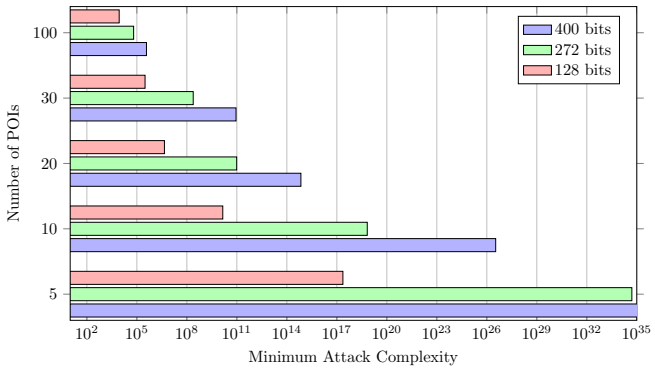


Fig. 5: Minimum attack complexity as the number of measurements needed to average to recover (parts) of the ISAP state from FULMINE. As a security margin we set $\gamma = 100$.

minimum attack complexity in terms of measurement traces is impracticable for less than 20 POIs and all mentioned state sizes. However, for higher numbers of POIs the minimum attack complexities tend towards practically feasible. Namely, when using 100 POIs, 10 000 measurements can be enough to learn 128 bits of the state, and 500 000 measurements are the minimum to recover the full state.

However, using that many POIs often hampers template building or leads to overfitting effects reducing the classification rate. Besides, side-channel leakage is not distributed such as to perfectly use the channel. This becomes visible in the gap between channel capacity and MI in Fig. 4. While this might allow an attacker to recover a few states more easily, in consideration of all possible states the attack complexity yet stays above the bounds in Fig. 5. Namely, for non-ideal distributions of the leakage, an attacker will, in general, require even more measurements to learn the specified amount of information.

From a practical perspective, conducting such powerful attack would require an attacker to successfully build templates on the respective state. In many cases, this is however not possible, e.g., when the attacker does not have control over the state on a suitable device. Even further, the complexity to build, measure, and evaluate such large set of templates is clearly impractical. In this respect, the implementation of ISAP on FULMINE can for the used measurement setup be considered secure against power analysis attacks also above the bounds in Fig. 5.

5 Conclusion

In this work, we presented a novel approach to determine leakage upper bounds for side channels of cryptographic implementations under a single data input. Without any further leakage assumptions we showed that the channel capacity of transmission channels with multiple in- and outputs gives the natural upper bound for information leakage in multivariate side channels with Gaussian noise.

We then considered the case where attackers are capable of performing multiple measurements of the same execution in order to improve their SNR. We showed that the gain in the SNR of multivariate leakages resulting from signal averaging is exponential in the number of POIs. This observation gives a tool for attackers to learn about the feasibility of an attack and for implementors to assess the minimum attack complexity of state recovery in leakage-resilient schemes allowing for multiple decryptions like Isap. We verified the soundness of our model and our bounds using the ASIC Fulmine implementing Isap and the keccak-f[400] permutation. Finally, we gave lower bounds on the complexity for recovering the Isap state using power analysis. The results indicate that recovery of the Isap state on Fulmine is practically infeasible with power analysis and the used measurement setup.

Acknowledgements. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 681402) and from the Austrian Research Promotion Agency (FFG) under grant number 845589 (SCALAS).

References

1. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Keccak sponge function family main document. Submission to NIST (Round 2) 3, 30 (2009)
2. Cover, T.M., Thomas, J.A.: Elements of information theory. John Wiley & Sons (2012)
3. Dobraunig, C., Eichlseder, M., Mangard, S., Mendel, F., Unterluggauer, T.: Isap - towards side-channel secure authenticated encryption. IACR Transactions on Symmetric Cryptology 2017(1), 80–105 (2017), http://tosc.iacr.org/index.php/ToSC/article/view/585
4. Duc, A., Faust, S., Standaert, F.: Making masking security proofs concrete - or how to evaluate the security of any leaking device. In: Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I. pp. 401–429 (2015), http://dx.doi.org/10.1007/978-3-662-46800-5_16
5. Gautschi, M., Schiavone, P.D., Traber, A., Loi, I., Pullini, A., Rossi, D., Flamand, E., Gürkaynak, F.K., Benini, L.: Near-threshold risc-v core with dsp extensions for scalable iot endpoint devices. IEEE Transactions on Very Large Scale Integration (VLSI) Systems PP(99), 1–14 (2017)
6. Goldsmith, A.: Wireless communications. Cambridge university press (2005)
7. Goldsmith, A., Jafar, S.A., Jindal, N., Vishwanath, S.: Capacity limits of MIMO channels. IEEE Journal on Selected Areas in Communications 21(5), 684–702 (2003), http://dx.doi.org/10.1109/JSAC.2003.810294
8. Lampret, D., Chen, C.M., Mlinar, M., Rydberg, J., Ziv-Av, M., Ziomkowski, C., McGary, G., Gardner, B., Mathur, R., Bolado, M.: Openrisc 1000 architecture manual. Description of assembler mnemonics and other for OR1200 (2003)

9. Medwed, M., Standaert, F., Nikov, V., Feldhofer, M.: Unknown-input attacks in the parallel setting: Improving the security of the CHES 2012 leakage-resilient PRF. In: Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part I. pp. 602–623 (2016), http://dx.doi.org/10.1007/978-3-662-53887-6_22
10. Mizuno, H., Iwai, K., Tanaka, H., Kurokawa, T.: Information theoretical analysis of side-channel attack. In: Information Systems Security - 9th International Conference, ICISS 2013, Kolkata, India, December 16-20, 2013. Proceedings. pp. 255–269 (2013), https://doi.org/10.1007/978-3-642-45204-8_20
11. Rivain, M.: On the exact success rate of side channel analysis in the gaussian model. In: Selected Areas in Cryptography, 15th International Workshop, SAC 2008, Sackville, New Brunswick, Canada, August 14-15, Revised Selected Papers. pp. 165–183 (2008), http://dx.doi.org/10.1007/978-3-642-04159-4_11
12. Rossi, D., Conti, F., Marongiu, A., Pullini, A., Loi, I., Gautschi, M., Tagliavini, G., Capotondi, A., Flatresse, P., Benini, L.: Pulp: A parallel ultra low power platform for next generation iot applications. In: Hot Chips 27 Symposium (HCS), 2015 IEEE. pp. 1–39. IEEE (2015)
13. Schindler, W., Lemke, K., Paar, C.: A stochastic model for differential side channel cryptanalysis. In: Cryptographic Hardware and Embedded Systems - CHES 2005, 7th International Workshop, Edinburgh, UK, August 29 - September 1, 2005, Proceedings. pp. 30–46 (2005), http://dx.doi.org/10.1007/11545262_3
14. Standaert, F., Malkin, T., Yung, M.: A unified framework for the analysis of side-channel key recovery attacks. In: Advances in Cryptology - EUROCRYPT 2009, 28th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cologne, Germany, April 26-30, 2009. Proceedings. pp. 443–461 (2009), http://dx.doi.org/10.1007/978-3-642-01001-9_26
15. Standaert, F., Pereira, O., Yu, Y., Quisquater, J., Yung, M., Oswald, E.: Leakage resilient cryptography in practice. In: Towards Hardware-Intrinsic Security - Foundations and Practice, pp. 99–134 (2010)
16. Telatar, E.: Capacity of Multi-antenna Gaussian Channels. European Transactions on Telecommunications 10, 585–595 (1999), https://doi.org/10.1002/ett.4460100604