Leak Me If You Can: Does TVLA Reveal Success Rate?

Abstract. Test Vector Leakage Assessment Methodology (*TVLA*) has emerged as a popular side-channel testing methodology as it can detect the presence of side-channel information in leakage measurements. However, in its current form, *TVLA* results cannot be used to quantify side-channel vulnerability. In this paper, we extend the *TVLA* testing beyond its current scope. Precisely, we derive concrete relationship between *TVLA* and signal to noise ratio (*SNR*). The linking of the two metrics, allows direct computation of success rate (*SR*) from *TVLA*, and thus unify these popular side channel detection and evaluation metrics. This, to our knowledge, is the first work in this direction. An end-to-end methodology is proposed, which can be easily automated, to derive attack *SR* starting from *TVLA* testing. The proposed methodology can take leakage model as a input and report attack *SR* which is validated on simulated and practical measurements. Not to surprise, the methodology performs better when the leakage model is accurately profiled. The methodology, although still limited to first-order leakage, is also further extended to (first order) multivariate setting.

1 Introduction

Since the seminal work by Kocher et al. [1], side channels have emerged as a serious threat to implementations of cryptographic algorithms in the past two decades, with the ability to render even mathematically robust cryptographic algorithms vulnerable. A side-channel adversary observes the physical properties of a cryptographic implementation, such as timing, power or electromagnetic emanations, and tries to infer the secret key by modelling a sensitive intermediate state of the design which then depends on these physical properties. Cryptographic designs must therefore provide security guarantees against such threats. In this context, efficient validation and evaluation methodology for testing side channel vulnerability has gathered significant interest in the research community. In particular, there exist today, two popular security certification programs - Common Criteria (CC) [2] and FIPS [3] that recommend crypto-implementations to be secure against side channel attacks. Each of these programs follows two distinct testing methodologies, namely *evaluation-style testing* and *conformance-style testing*.

Evaluation-Style Testing. The Common Criteria (CC) certification is a prime example of evaluation-style testing. CC is essentially a set of security guidelines (ISO-15408) that define a common framework for evaluating crypto-implementations using a standard set of pre-defined evaluation assurance levels. From the point of view of detecting side channel vulnerabilities, it recommends evaluating the system against all state-of-the-art attack strategies, with the knowledge of the threat model. An ever-increasing list of attack strategies, together with a large number of models characterizing different leakage profiles of

the device, often renders such a testing methodology cumbersome, costly and limited by the testing expertise available at hand. Additionally, the success of evaluation-style testing methodologies depends strongly on appropriate choices of the leakage models, and an error of judgement in this regard could cause a potentially vulnerable crypto-implementation to pass the test. This makes evaluation style testing mechanisms less favourable for testing crypto-implementations against side channel vulnerability.

Conformance-Style Testing. Unlike CC, FIPS [3] certification is an example of conformance-style testing that uses a cryptographic module validation program (CMVP) to validate a design in terms of whether it meets the necessary security levels or not, rather than an exact evaluation of its vulnerability. With respect to side channels, it employs a simplified approach of merely detecting the presence of *any* leakage, independent of attack methodologies and leakage models. This makes it possible to have structured conformance-style testing methodologies that are cost-effective and consistent across different testing labs with varied testing expertise. Fortifications with precise security specifications and test plan coverage have the potential to make this style of testing against side-channel vulnerabilities highly efficient and suitable for wide-scale use.

Test Vector Leakage Assessment (*TVLA*) [4] which was proposed at NIST sponsored NIAT workshop 2011, is one of such conformance style testing mechanism which has gained popularity among the researchers and specially the practitioners due to its robustness, applicability to different crypto-implementations and easy integrability with the exiting testing methodologies. Multiple research papers on side channel attacks have used this tool to show the effectiveness of their proposed attacks and countermeasures. *TVLA* uses well known *Welch's t-test*. It was proposed as a *PASS/FAIL* test, which checks if t-value crosses the predefined threshold (proposed as ± 4.5 [4]). If the t-value crosses the threshold, the measurement are considered to carry data dependant information, which could be potentially exploited.

TVLA can be classified into: *non-specific* and *specific* [4]. Non-specific *TVLA* partitions traces on basis of public inputs. Specific *TVLA* partitions based on intermediate key-dependent variables and thus can provide intuitions on source of leakage. It has been shown in [5] that non-specific *TVLA* outperforms specific *TVLA* as the number of false positives will be less in case of non-specific *TVLA*. Both methods are discussed in details in section 2.

One demerit of *TVLA* methodology is that, failed t-test may or may not lead to successful key extraction. It may happen that key extraction procedure fails due to wrong assumption of or high complexity of the hypothetical power model in-spite of having high *TVLA* leakage [6]. Moreover *TVLA* does not quantifies the side channel vulnerability. In some cases, it would be useful to know how unsafe the design is, which demands the need of quantification of side channel vulnerability. However, in current form, *TVLA* fails to report side-channel vulnerabilities and evaluation based testing are too costly and expertise dependent to be deployed for this objective.

1.1 Motivation

Conformance based testing based on *TVLA* is gaining popularity due to its simplicity and ease of computation. Although *TVLA* itself is not sufficient for a comprehensive security evaluation, it is often considered as a first test to guide further evaluations. However, in its current form it is mostly used for detecting presence of leakage and sometimes to derive the order of implementation security. In this work, we attempt to develop a *hybrid* methodology to extract more information from the initial *TVLA* testing. The information extraction is oriented towards expressing the *TVLA* results in terms of other metrics commonly used in *evaluation-based* testing. More precisely, we derive relationships and develop methodologies to utilize information from the *TVLA* test for computation of signal-to-noise ratio (SNR) and attack success rate (SR).

1.2 Context

Countermeasures against side-channel attack are advancing every year [7]. Along side there are comprehensive evaluation methodologies which are also developed [8]. Such evaluations are all the more important when basic T-test can be misleading. A recent work [9] shows the limitations of T-test in security evaluation of a higher-order masking scheme. However, conducting a comprehensive and detailed security evaluation can be a time-taking task. Time is a limiting factor for evaluation process and for the same reason *CC* evaluations contain time spend for evaluation as a metric. Some work deal with further simplifying the evaluation process [10].

Most, if not all, real implementation are currently considering basic countermeasures due to the cost of security attached. This scenario might change in the future. Thus, evaluation laboratories are still often dealing with unprotected or low-order protected cryptographic implementations, which might also suffer from accidental first order leakage. In such scenarios, a simple testing methodology like *TVLA* can be a good start.

In this paper, we develop a methodology that starts with specific *TVLA* testing to get some information on leaking variables. Thereafter, by plugging in a leakage model, the evaluator can reuse the *TVLA* results to compute attack *SR*. This computation first converts *TVLA* results to *SNR*. Next, it applies previously developed techniques like [11,12,13], to compute *SR* from *SNR* and the leakage model.

A further extension of this proposed technique to multivariate setting is also discussed. In its present form the developed methodology is limited to unprotected targets or protected targets with accidental first order leakage.

1.3 Related Work

A unified framework to evaluate side-channel attack was proposed by Standaert et al. [14]. It put forwards two key metrics success rate (SR) and guessing entropy (GE) as main attack metrics. Success rate of a specific side channel attack is defined as the probability of successful secret key retrieval. In simple mathematical notation, success rate (SR) of a side channel attack (A) is presented as follows:

$$SR = Pr[A(E_{k_0}, L) = k_0] \tag{1}$$

where k_0 is the correct key used in the encryption process E_{k_0}, L is the leakage obtained from side channel traces. In CHES 2012, Fei et al. [11] introduce the notion of confusion coefficient which can be used to compute theoretical success rate of a mono-bit differential power analysis (i.e. difference of mean) given the *SNR*. This work was further improved and extended to correlation power analysis by Thillard et al. [12]. Fei et al. [13] also extended the initial work on success rate estimation for monobit DPA to CPA and beyond.

On the other hand, to simplify the evaluation process, simple and model-agnostic techniques were also developed in parallel. The main technique of this class being the previously mentioned *TVLA* [4] was proposed as a FIPS 140-3 candidate. Another simple method to detect point of leakage in a univariate first-order setting was proposed in [15], termed as *Normalized Inter Class Variance (NICV)*. Authors show that NICV is an estimate of *SNR* and approaches (squared) Pearson's correlation coefficient in absence of noise. *NICV* is actually output of statistical F-test (also known as ANOVA (ANalysis Of VAriance)). Owing to its relationship to *SNR*, *NICV* was also used to derive *SR* for monobit DPA using formulation from [11]. In this work, we work on connecting the individual techniques to develop the whole chain. The main missing link in the above techniques is the relationship between *TVLA* and *SNR*. By developing that link, we are able to develop a methodology that can be automated end to end to estimate attack *SR* right from computation of *TVLA*.

1.4 Contribution

The main contributions of this paper are as follows:

- *SNR* of side-channel measurement and *TVLA* are independently developed metrics. We derive the relationship between *SNR* and *TVLA*. We formally show that the two metrics are equivalent and one can be easily computed from the other.
- Next, we devise a methodology to estimate the theoretical bounds for success rate of an attack from the *TVLA* results. This, to our knowledge, is the first attempt to extend *TVLA* results for quantification of side channel vulnerability through *SR*. The methodology uses theoretical success rate formulation for CPA by Fei et al. [13]. In other words, the developed methodology attempts to bridge the gap between conformance and evaluation based testing by setting the following chain: $TVLA \rightarrow SNR \rightarrow SR$.
- We also show that non-specific *TVLA* actually captures only a fraction of the total *SNR*. On the other hand, from specific *TVLA*, we can compute the total *SNR* from *TVLA*.
- The developed methodology is extended to multivariate setting under first-order leakage setting.

The rest of the paper is organized as follows: section 2 briefly describes the mathematics behind different metrics for validation and evaluation of side channel vulnerabilities. Next, section 3, derives the relationship between *Welch's t-test* based *TVLA* and *ANOVA* based *NICV* (and *SNR*). The derived relationship is experimentally validated in section 4 followed by application to AES in section 5. The extension of the proposed methodology to multivariate setting is discussed in section 6 followed by final conclusions in section 7.

2 Preliminaries

In this section we will introduce our notations and provide a brief description of *TVLA*, *NICV*, *SNR*. Finally, the previously proposed relationship between *SR* and *SNR* is discussed.

We denote X, k as the plaintext and key bytes. Let $L = l(X, k)$ denote the normalized leakage model with $\mathbb{E}(L) = 0$ and $\mathsf{Var}(L) = \mathbb{E}(L^2) = 1$ and let Y denote the leakage measurements such that

$$Y = \epsilon L + N \tag{2}$$

where ϵ is the scaling coefficient and $N \sim \mathcal{N}(0, \sigma^2)$ is the noise, which is independent of X. A common example for $l(X, k)$ is the Hamming weight leakage model on n bits:

$$l(X, k) = \frac{2}{\sqrt{n}} \left(HW(X \oplus k) - \frac{n}{2} \right).$$

Definition 1. SNR *[16, § 4.3.2, page 73] The Signal-to-Noise Ratio (*SNR*) is defined as:*

$$\mathit{SNR} = \frac{\mathsf{Var}(\mathbb{E}(Y|X))}{\mathbb{E}(\mathsf{Var}(Y|X))}.$$

Lemma 1 (*SNR* in the case of leakage model (2)).

$$\mathit{SNR} = \frac{\epsilon^2}{\sigma^2}.$$

Proof. Let x a plaintext, and $l = l(x, k)$. Then $\mathbb{E}(Y|X = x) = \mathbb{E}(\epsilon L + N|L = l) = \epsilon l$, by expression of the model (2) and noise independence from the L. Therefore, $\mathsf{Var}(\mathbb{E}(Y|X)) = \mathsf{Var}(\epsilon L) = \epsilon^2$. Besides, $\mathbb{E}(\mathsf{Var}(Y|X)) = \mathbb{E}(\sigma^2) = \sigma^2$. Hence, $\mathsf{SNR} = \frac{\mathsf{Var}(\mathbb{E}(Y|X))}{\mathbb{E}(\mathsf{Var}(Y|X))} = \frac{\epsilon^2}{\sigma^2}$.

2.1 Normalized Inter Class Variance

Normalized Inter-Class Variance (*NICV*) is a technique which was designed to detect relevant point of interest (PoI) in an SCA trace [15]. It has application in side channel trace compression and dimensionality reduction. *NICV* is based on *ANOVA (ANalysis Of VAriance)* or *F-test* [17]. The main advantage of *NICV* is that, it is leakage model agniostic and can be applied with the knowledge of only plain-text or cipher-text and does not require knowledge of target implementation or secret key.

Definition 2 (*NICV* [15, Eqn. (4) of Sec. 3.1] or [18, Eqn. (4) of Sec. 3.1]). *The Normalized Inter-Class Variance (*NICV*) is defined as:*

$$\mathit{NICV} = \frac{\mathsf{Var}(\mathbb{E}(Y|X))}{\mathsf{Var}(Y)} \tag{3}$$

Lemma 2 (*NICV* in the case of leakage model (2)).

$$\mathit{NICV} = \frac{1}{1+\frac{\sigma^2}{\epsilon^2}}.$$

In particular, $0 \leq \mathit{NICV} \leq 1$.

Proof. The numerator has already been proven to be equal to ϵ^2. Besides, $\mathsf{Var}(Y) = \mathsf{Var}(\epsilon L) + \mathsf{Var}(N) = \epsilon^2 + \sigma^2$, by independence of X and N. Hence $\mathsf{NICV} = \frac{\mathsf{Var}(\mathbb{E}(Y|X))}{\mathsf{Var}(Y)} = \frac{\epsilon^2}{\epsilon^2+\sigma^2} = \frac{1}{1+\frac{\sigma^2}{\epsilon^2}}$.

Proposition 1 (Link between *NICV* and *SNR*, [15, Eqn. (5) of Sec. 3.1] or [18, Eqn. (5) of Sec. 3.1]). *We have:*

$$\mathit{NICV} = \frac{1}{\frac{1}{\mathit{SNR}}+1} \qquad \textit{and, conversely,} \qquad \mathit{SNR} = \frac{1}{\frac{1}{\mathit{NICV}}-1}. \tag{4}$$

Proof. Direct application of Lemmas 1 and 2.

2.2 Test Vector Leakage Assessment (*TVLA*)

*Test Vector Leakage Assessment (*TVLA*)* [4] is direct application of *Welch's t-test* on side channel traces for validation of side channel vulnerabilities. *TVLA* methodology can be classified in to two different categories: *non-specific TVLA* and *specific TVLA*. For both the cases, one must acquire two sets of traces. In case of *non-specific TVLA*, one set corresponds to a fixed key and fixed plaintext as input to the cryptographic IP, the second set collects traces corresponding to same fixed key and random plaintext. The captured side channel traces are then partitioned into two different sets: Y^f (fixed plaintext as input) and Y^r (random plaintext as input). Thereafter a hypothesis testing performed by assuming a null hypothesis that the these two sets of traces have identical means and variance. If the null hypothesis is accepted, it signifies that the traces carry no sensitive information. On the other hand, a rejected null hypothesis indicates presence of exploitable leakage. This can be expressed as:

$$\mathsf{TVLA} = \frac{\mu_r - \mu_f}{\sqrt{\frac{\sigma_r^2}{n_r} + \frac{\sigma_f^2}{n_f}}} \ , \tag{5}$$

where n_r, n_f signifies the number of traces in set Y^r, Y^f respectively. The mean and standard deviation of set Y_r is denoted by μ_r and σ_r. Similarly, μ_f and σ_f refer to mean and standard deviation of Y^f. The testing also commonly known as fixed vs random (FVR) test. The null hypothesis of two equal means is rejected when the $TVLA$ exceeds a threshold of ± 4.5, which ensures with degrees of freedom > 1000, $P[|TVLA| > 4.5] < 0.00001$, this threshold leads to a confidence of 0.99999. Thus, if the $TVLA$ value is within ± 4.5, the traces are considered to not contain data-dependant leakage. Otherwise, it reject the null hypothesis and declare the crypto-implementation to leak exploitable side-channel information.

Now, connecting $TVLA$ with previous derivations we have:

Definition 3 (*TVLA* [4, page 7]). *The* non-specific TVLA *is defined for Q queries as:*

$$\widehat{TVLA}_x = \frac{\left(\frac{1}{\sum_{q/x_q=x} 1}\sum_{q/x_q=x} y_q\right) - \left(\frac{1}{\sum_q 1}\sum_q y_q\right)}{\sqrt{\frac{1}{\sum_{q/x_q=x} 1}\left(\frac{1}{\sum_{q/x_q=x} 1} y_q^2 - \left(\frac{1}{\sum_{q/x_q=x} 1} y_q\right)^2\right) + \frac{1}{\sum_q 1}\left(\frac{1}{\sum_q 1} y_q^2 - \left(\frac{1}{\sum_q 1} y_q\right)^2\right)}},$$

where we used $\sum_q$ for $\sum_{q=1}^{Q}$ and $\sum_{q/t_q=t}$ for $\sum_{\substack{1\le q\le Q, \\ s.t.\ t_q=t}}$.

We notice that this test is consistent, in that, asymptotically,

$$\widehat{\mathsf{TVLA}}_x \xrightarrow[Q\to+\infty]{} \begin{cases} +\infty & \text{if } \mathbb{E}(Y|X=x) \neq \mathbb{E}(Y), \\ 0 & \text{otherwise.} \end{cases}$$

More precisely, according to the law of large numbers (LLN), we have that:

$$\widehat{\mathsf{TVLA}}_x \underset{Q\to+\infty}{\approx} \sqrt{Q}\frac{\mathbb{E}(Y|X=x) - \mathbb{E}(Y)}{\sqrt{\mathsf{Var}(\mathbb{E}(Y|X=x)) + \mathsf{Var}(\mathbb{E}(Y))}}.$$

We therefore define the asymptotic constant $\lim_{Q\to+\infty} \frac{1}{\sqrt{Q}}\widehat{\mathsf{TVLA}}_x = \mathsf{TVLA}_x$ as:

Definition 4. *Asymptotic constant for Test Vector Leakage Assessment (*TVLA*) for Fixed versus Random is:*

$$TVLA_x = \frac{\mathbb{E}(Y|X=x) - \mathbb{E}(Y)}{\sqrt{\mathsf{Var}(\mathbb{E}(Y|X=x)) + \mathsf{Var}(\mathbb{E}(Y))}},$$

where the fixed plaintext is x. In this definition, the test is non-specific*, since one does not need to know the key.*

Lemma 3 (*TVLA* in the case of leakage model (2)).

$$TVLA_x = \frac{\epsilon l(x,k)}{\sigma}.$$

Proof. Indeed, we have $\mathbb{E}(Y) = 0$, hence the result follows.

For *specific TVLA*, knowledge of secret key is required as in this case the traces are partitioned depending upon the value of some intermediate data of crypto-execution [4]. Depending upon the choice of intermediate data, there could be multiple way to do this partitioning. In [5], the superiority of *non-specific TVLA* over *specific TVLA* is established. *TVLA* is compared with mutual information based analysis techniques in [19] and comparative analysis between them is presented. In [20], authors have focussed on applicability of *TVLA*. They have extended application of *TVLA* to higher order attacks. Moreover, they have presented efficient algorithms for on-line computation of *TVLA*. A modified paired T-test based *TVLA* methodology is presented in [21].

2.3 *SNR* and *SR*

A closed-form expression for DPA and CPA has been derived in [11,12,13] that depends on three factors: number of measurements Q, SNR, confusion coefficient vector $\boldsymbol{\kappa}$, and confusion matrices $\boldsymbol{K}, \boldsymbol{K}^{**}$.

Definition 5 (Confusion vector and matrices for CPA [13][1]). *Let k_c denote the secret key and k_{g_i} with $1 \leq i \leq 2^{n-1}$ a key guess where $k_{g_i} \neq k_c$, then the confusion vector $\boldsymbol{\kappa}$ and the confusion matrices $\boldsymbol{K}$,$\boldsymbol{K}^{**}$ are defined as*

$$\boldsymbol{\kappa} = (\kappa(k_c, k_{g_1}), \ldots, \kappa(k_c, k_{g_{2^{n-1}}})^T$$

$$\boldsymbol{K} = \begin{pmatrix} \kappa(k_c, k_{g_1}, k_{g_1}) & \kappa(k_c, k_{g_1}, k_{g_2}) & \cdots & \kappa(k_c, k_{g_1}, k_{g_{2^n-1}}) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa(k_c, k_{g_{2^n-1}}, k_{g_1}) & \kappa(k_c, k_{g_{2^n-1}}, k_{g_2}) & \cdots & \kappa(k_c, k_{g_{2^n-1}}, k_{g_{2^n-1}}) \end{pmatrix}$$

$$\boldsymbol{K}^{**} = \begin{pmatrix} \kappa^{**}(k_c, k_{g_1}, k_{g_1}) & \kappa^{**}(k_c, k_{g_1}, k_{g_2}) & \cdots & \kappa^{**}(k_c, k_{g_1}, k_{g_{2^n-1}}) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa^{**}(k_c, k_{g_{2^n-1}}, k_{g_1}) & \kappa^{**}(k_c, k_{g_{2^n-1}}, k_{g_2}) & \cdots & \kappa^{**}(k_c, k_{g_{2^n-1}}, k_{g_{2^n-1}}) \end{pmatrix}$$

with

$$\begin{aligned} \kappa(k_c, k_g) &= E((l(X, k_c) - l(X, k_g))^2) \\ \kappa(k_c, k_{g_i}, k_{g_j}) &= E((l(X, k_c) - l(X, k_{g_i})(l(X, k_c) - l(X, k_{g_j})) \\ \kappa^{**}(k_c, k_{g_i}, k_{g_j}) &= 4E((l(X, k_c) - E(l(X, k_c)))^2 \\ &\quad (l(X, k_c) - l(X, k_{g_i}))(l(X, k_c) - l(X, k_{g_j}))). \end{aligned}$$

Note that, in case of no-weak keys $\boldsymbol{\kappa}, \boldsymbol{K}, \boldsymbol{K}^{**}$ are not key dependent and thus can be determined without knowing the correct key by setting w.l.o.g $k_c = 0$. Now, considering a leakage model as in Eq. (2), the theoretical success rate is given by

$$\mathsf{SR} = \Phi_{[\boldsymbol{K} + (\frac{\epsilon}{2\sigma})^2(\boldsymbol{K}^{**} - \boldsymbol{\kappa}\boldsymbol{\kappa}^T)]}(\sqrt{Q}\frac{\epsilon}{2\sigma}\boldsymbol{\kappa}) \tag{6}$$

where $\Phi_{[C]}(\boldsymbol{\mu})$ is the cumulative distributive function of the multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance C. Now as $\mathsf{SNR} = \frac{\epsilon^2}{\sigma^2}$ a direct relation between *SNR* and *SR* is given by

$$\mathsf{SR} = \Phi_{[\boldsymbol{K} + (\frac{1}{4})\mathsf{SNR}(\boldsymbol{K}^{**} - \boldsymbol{\kappa}\boldsymbol{\kappa}^T)]}(\sqrt{Q}\frac{1}{2}\sqrt{\mathsf{SNR}}\boldsymbol{\kappa}). \tag{7}$$

Note that, Eqs. (6) and (7) hold for Eq. (2) and thus assume that $l(X, k)$ is known. However, which has not been mentioned in previous works, is that in a practical scenario one may use an approximation of $l(X, k)$ (e.g., $HW(Sbox^{-1}(X \oplus k))$.

[1] Note that, the formula for the theoretical success rate in [12] should yield equivalent results.

This approximation may influence the goodness of the estimation of the theoretical *SR* in two different ways. First, it may influence the values of $\boldsymbol{\kappa}, \boldsymbol{K}, \boldsymbol{K}^{**}$ as the approximation may not have the same (less or more) "distinguishing ability" as $l(X, k)$. Second, the error made in the approximation of $l(X, k)$ introduces additional noise (epistemic noise from the leakage model) which is not captured when estimating the *SNR* on the traces. From previous experiments we observed that the second aspect is more crucial than the first one.

To take a global look on the previous work, *NICV* is shown directly related with the *SNR*, which in turn is a key input for computing the minimum number of side channel traces required for performing successful *CPA*. However, no such formulation exist in case of *TVLA*. In the subsequent section, we will establish the relationship between *TVLA* and *SNR* so that we can extend the testing mechanism of *TVLA* based conformance standards.

3 Equivalence of *TVLA* and *NICV*

The objective of this section is to establish relationship between *TVLA* and *NICV*, which will be the first step in connecting *TVLA* with *SNR*. We follow the same methodology as *TVLA* i.e. dividing data into two groups followed by application of *NICV* (and *SNR*) to it.

Let us assume that an adversary has collected n side channel traces. The entire set of side channel traces is designated as Y and individual side channel trace is denoted as Y_i, where $i \in [1, n]$ is the index of the corresponding side channel trace. Next following the *TVLA* approach, the traces are partitioned into two groups: Y^{G1} and Y^{G2}, having cardinality n_1 and n_2 ($n = n_1 + n_2$) respectively. Mean and variance of group Y^{G1} and group Y^{G2} are denoted by μ_1, σ_1^2 and μ_2, σ_2^2 respectively. Moreover, mean and variance of the entire set Y are denoted as μ and σ^2. The objective is to derive the relationship between *TVLA* and *NICV* metric. Since, we are dealing with only two groups in this case, the corresponding two group *NICV* is denoted as $NICV_2$. This $NICV_2$ will be generalized in the following subsection.

Theorem 1. *Consider two group of side channel traces Y_1 and Y_2 with cardinality n_1 and n_2. The computation of* TVLA *and NICV_2 on these two groups are related by the following formula*

$$\mathsf{NICV}_2 = \frac{1}{\frac{n}{\mathsf{TVLA}^2} + \frac{n}{C}(\sigma_1^2 - \sigma_2^2)\left(\frac{1}{n_2} - \frac{1}{n_1}\right) + 1} \tag{8}$$

where $C = \left(\mu_1^2 - \mu_2^2\right)^2$

Proof. The derivation is provided in appendix A

Corollary 1. *If both the group have same number of side channel traces ($n_1 = n_2 = \frac{n}{2}$), Eqn.* (8) *transforms into*

$$\mathit{NICV}_2 = \frac{1}{\frac{n}{\mathit{TVLA}^2} + 1}. \tag{9}$$

Remark 1. It must be noticed that $TVLA$ needs to be evaluated for a finite number of traces (n), otherwise it diverges to $+\infty$. However, TVLA^2/n tends to a finite value when n tends to $+\infty$, which bounds the value of $\mathsf{NICV} \in [0, 1]$.

3.1 Generalizing the *NICV* Computation

The relationship between $TVLA$ and NICV_2 (2-class $NICV$) was derived previously. However, the general application of $NICV$ (or SNR) is not restricted to two classes. In this section, the relation between $TVLA$ is extented from NICV_2 to a generic k-class NICV (NICV_k).

Let us now assume that n number of side channel traces can be partitioned into k number of groups where i^{th} group contains n_i number of traces. A generic example in case of ciphers like AES, where byte-wise computation is performed and the desired value k is 256. NICV_k can be directly computed from NICV_2 by following an iterative approach. For the derived k groups, pairwise computation of $(k-1)$ different NICV_2 is performed and the results are combined as follows:

- $\forall i \in \mathbb{Z}_k$, create two groups: the first group contains the side channel traces with particular byte of the plain-text equal to i, the other group will contain the side channel traces with that particular byte value not equal to i. The mean of these two groups are denoted as μ_i and $\mu_{\bar{i}}$ respectively.
- Compute NICV_2 for each of these two groups. We denote this as NICV_2^i.

Theorem 2. *The computation of NICV_k and NICV_2^i are related by the following formula if all k groups have same number of side channel traces*

$$\mathit{NICV}_k = \frac{k-1}{k} \sum_{i=1}^{k} \mathit{NICV}_2^i. \tag{10}$$

Proof. The derivation is provided in appendix B

3.2 Extension to Non-Specific *TVLA*

In this part, we establish the relationship between SNR and non-specific $TVLA$. A first hint of link between SNR and $TVLA$ was qualitatively discussed in [18]. The formal relationship is derived as follows.

Proposition 2 (Link between *SNR* and *TVLA*). *The* SNR *is the variance of the* TVLA *values in the Fixed versus Random (or non-specific) setup, the variance being computed over all possible fixed values:*

$$\mathit{SNR} = \mathsf{Var}(\mathit{TVLA}_X).$$

Algorithm 1: Computing *SNR* and *SR* from *TVLA*

Input: Side channel traces and corresponding intermediate state
Output: *SNR*, *SR* for chosen sub-key

1 **for** $i = 0$ *to* k **do**
2 Partition the side channel traces into two groups: G_1 and G_2
3 G_1: Side channel traces where j^{th} byte of the intermediate data $= i$
4 G_2: Side channel traces where j^{th} byte of the intermediate data $\neq i$
5 Apply *TVLA* on groups G_1 and G_2
6 Compute NICV_2^i from the *TVLA* value by using Eqn. (8)
7 Compute $\mathsf{NICV}_k = \frac{k-1}{k} \sum_{i=1}^{k} \mathsf{NICV}_2^i$
8 Compute $\mathsf{SNR} = \frac{1}{\frac{1}{\mathsf{NICV}_k} - 1}$
9 Compute $\mathsf{SR} = \Phi_{[\boldsymbol{K} + (\frac{1}{4})\mathsf{SNR}(\boldsymbol{K}^{**} - \boldsymbol{\kappa}\boldsymbol{\kappa}^T)]}(\sqrt{Q}\frac{1}{2}\sqrt{\mathsf{SNR}}\boldsymbol{\kappa})$
10 Return SNR, SR

Proof. As $\mathsf{TVLA}_X = \frac{\epsilon l(X,k)}{\sigma}$, we have: $\mathsf{Var}(\mathsf{TVLA}_X) = \frac{\epsilon^2}{\sigma^2}\mathsf{Var}(L) = \frac{\epsilon^2}{\sigma^2} = \mathsf{SNR}$.

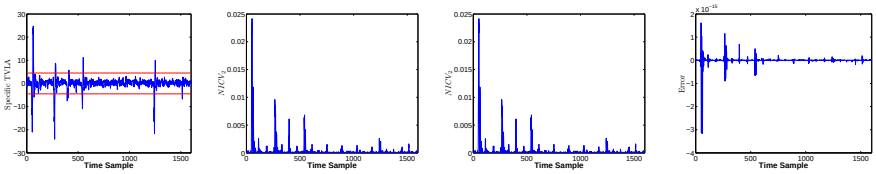
For *non-specific TVLA*, the traces are partitioned depending upon the entire plaintext value, where one group contains traces with fixed plaintext and other contains traces with random plaintext. If we want to extend our approach to *non-specific TVLA* to compute *SNR*, we need to compute *TVLA* for each plaintext value, which is computationally infeasible. Thus, in the following, we stick to specific *TVLA* only.

3.3 Extending *TVLA* flow to Side-Channel Analysis

Side channel analysis works using divide and conquer approach. For instance, *SPN* cipher where each $b \times b$ S-box handle b bits of the entire key bits, the attack focuses on each of these b bit groups separately. In case of AES-128, $b = 8$ which means that the attack is applied on 8-bits or one byte of the secret key, also known as sub-key. The attack is repeated 16 times to recover all the key bytes in AES-128. This reduces the complexity of the attack significantly. The same applies to *SNR* and *NICV*. One can compute *SNR* or *NICV* byte-wise to zero down the leakage zone of each key byte and apply the attack.

Now we present the methodology to extend the *TVLA* computation to recover *SNR* and there after compute success rate with a given attack model. From *TVLA*, NICV_2 can be computed by Eqn. (8), which further leads to NICV_k by Eq. (10). NICV_k (or just *NICV*) can directly provide the *SNR* by Eq. (4). Finally, *SNR* leads to *SR* by Eq. (7). The methodology is presented in Algorithm 1. The algorithm is repeated for each sub-key to recover the whole secret key.

It must be noted that partitioning the side channel traces, depending upon a particular byte value of the intermediate state was deployed for *specific TVLA* also. Steps 1 and 2 of algorithm 1 are actually application of *specific TVLA*. Thus using the formalization approach presented in this and previous sections, we can compute *SNR* of the crypto-system from *specific TVLA* computation. As stated above, the methodology cannot be applied to non-specific *TVLA* due to computational infeasibility.



(a) Single specific $TVLA$ (b) Predicted NICV_2 from plot (a) (c) Computed NICV_2 from traces (d) Prediction Error in Eq. (8) (plot (b) - (c))

Fig. 1: Equivalence of $TVLA$ and NICV_2

4 Experimental Verification of Derived *TVLA* and *NICV* Relation

The derived relation between *specific* $TVLA$ and SNR (or $NICV$) is experimentally validated in this section on an AES-128 implementation (without side-channel countermeasures) running on an $ATMEGA8515$ smart-card.

4.1 Experimental Setup

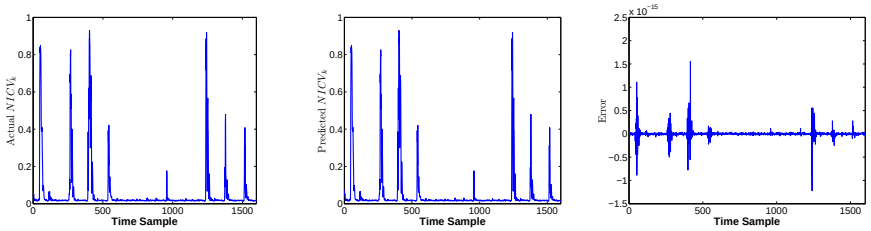
The AES design is implemented on a SAKURA-GW platform [22]. SAKURA-GW platform consists of two boards: SAKURA-G and SAKURA-W. SAKURA-G board contains a SPARTAN-6 FPGA which controls the communication and SAKURA-W board contains the smart-card containing implementation of AES-128. The power measurements are taken using a Tektronix MSO4034B mixed signal oscilloscope with sampling frequency 500 MHz. Being an unprotected implementation, it is obvious that the AES implementation must have exploitable leakage and its $TVLA$ value should be more than the threshold of 4.5.

4.2 Validation of *TVLA* and NICV_2 Relationship

The relationship between $TVLA$ and NICV_2 was established in Eq. (8). It is verified on the collected power measurement for AES on ATMEGA-8515 smart-card. We start with partitioning the traces based on the first byte value ($k = 256$) of the round 9 output as intermediate state, following step 1 of Algo. 1. Next we compute $TVLA$ and NICV_2 from the partitions again following Algo. 1. The results are shown in Fig. 1. An example of specific $TVLA$ trace is shown in Fig. 1 (a). Next the $TVLA$ trace in Fig. 1 (a) is used to compute NICV_2 using Eq. (8) and shown in Fig. 1 (b). We also compute $NICV_2$ from power measurement as shown in Fig. 1 (c). The error between predicted and computed $NICV_2$ is in the order of 10^{-15} i.e. negligible (Fig. 1 (d)), which confirms Eq. (8).

4.3 Validation of NICV_k and NICV_2 relationship

Similar validation is also done for Eq. (10) that relates NICV_2 and NICV_k. Using the same set of traces and no. of partitions ($k = 256$), we compute NICV_k from



(a) Computed NICV_k (b) Predicted NICV_k (c) Error in Eq. (10) ((a) - (b))

Fig. 2: Prediction of NICV_k

the traces and predict it from previously computed NICV_2. The results are shown in Fig. 2. As the computed NICV_k(Fig. 2 (a)) follows closely the predicted NICV_k (Fig. 2 (b)), the prediction error (Fig. 2 (c)) also stays in the range of 10^{-15}.

5 Case Study: Application to AES

The equivalence of *TVLA* and *SNR* was theoretically derived and experimentally verified in the previous sections. The step by step procedure to compute *SNR* (and *SR*) from the *specific TVLA* value was presented in Algo. 1. In this section, we focus on the application of these relations towards testing an unprotected AES-128 design. First results are shown on simulated power traces, followed by application of the evaluation methodology on actual power traces acquired from AES implementation running on an ATMEGA-8515 smart-card.

5.1 Under Simulated Setting

Simulated traces are generated for an 8 bit micro-controller, assuming perfect Hamming weight leakage and added zero mean Gaussian noise ($\mathcal{N}(0, \sigma)$), where σ denotes the standard deviation of the noise distribution. The side channel trace can be represented as $Y = HW(v) + \mathcal{N}$, where v is the chosen intermediate value, which in this case is first 8-bits of round 9 output. We have generated side channel traces for different *SNR* values ranging from 0.03 to 2.

Next, we directly apply Algo. 1 to first derive *SNR* and then compute the theoretical success rate *SR*. A practical CPA attack is also performed on the set of the traces to compare practical success rate with the theoretical estimation. The corresponding result is shown in Fig. 3, where we compare the practical *SR* computation with the computation of theoretical *SR*. It can be observed that under perfect HW model assumption, the estimated theoretical estimation and practical computation of *SR* fits quite closely. A minor overshoot for practical *SR* is seen for high *SNR* (> 0.5). This overshoot is a approximation glitch in the theoretical formulation under central limit theorem and law of large numbers, which needs few dozen traces to converge. Otherwise, the approximation overshoot remains constant even for extremely high *SNR* (tested up to *SNR=20*). The

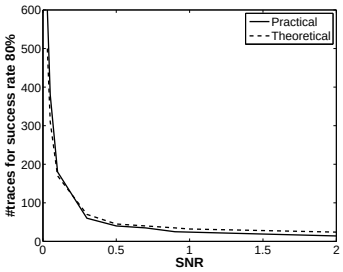


Fig. 3: Comparison Between Theoretical *SR* and Practical *SR* for different *SNRs*

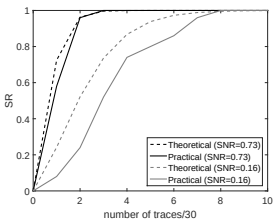
overshoot can be seen in real traces as well for high *SNR* scenarios in the next subsection.

5.2 On Real Power Traces

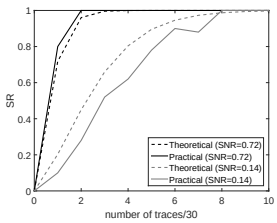
The experimental setup for the acquisition of power traces is equivalent to the one described in section 4.1. Further white Gaussian noise is added to experiment in low-*SNR* scenarios. The experiments were performed with 20,000 traces. For practical *SR*, a CPA was mounted on a randomly chosen set of 300 traces, repeated 50 times. Following Algo. 1 and assuming that the ATMEGA-8515 smart-card leaks in HW model, we generate plots for estimated theoretical success rate. The results are shown in Fig. 4 for two distinct points on the trace.

Finding a device with perfect HW is a very strong assumption. The two distinct points chosen are as such that one point has leakage very close to HW model while the other deviates from the model. An closer estimation to the actual model is computed using profiling based on Stochastic modelling [23] of leakage into 9 dimensions as $\Sigma_{i=1}^{8}\beta_i v_i$. The β weights of different points are shown in Fig. 5. While Fig. 5(a) shows a point where the leakage model deviates from HW model, Fig. 5(b) stays close to HW model. Referring back to Fig. 4, when the *SNR* is high, the practical *SR* for both near perfect and imperfect model closely matches the theoretical prediction. However, as the *SNR* reduces, the deviation between theoretical and practical *SR* increases. This deviation is even worse when the model is imperfect (see Fig. 4 (b)).

We repeat the experiments by taking the actual model into the account and rerunning Algo. 1. Precisely it is only the last step of Algo. 1 which is affected by the leakage model as stated in Eq. (7). The results are shown in Fig. 6. Again under high *SNR*, the practical attack results matches the theoretical estimation. However, by taking the correct leakage model into the account, the theoretical estimation and practical also matches closely for leakage sample with imperfect HW as well as near perfect HW leakage sample. The matching is a result of two fold impact: firstly, the theoretical estimation of *SR* becomes less optimistic than perfect HW and the practical *SR* is more realistic than perfect HW assumption. This experiment confirms the importance of leakage modelling in a side-channel

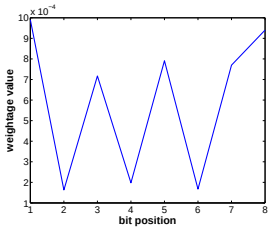


(a) sample point with imperfect HW leakage model

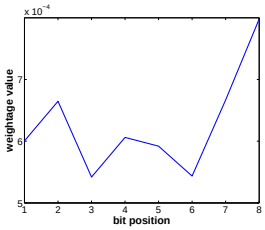


(b) sample point with nearly perfect HW leakage model

Fig. 4: Comparison Between Theoretical *SR* and Practical *SR* for different *SNRs* using Hamming weight model at different sample points



(a) sample point with imperfect HW leakage model



(b) sample point with nearly perfect HW leakage model

Fig. 5: Sample points with perfect and imperfect HW leakage model

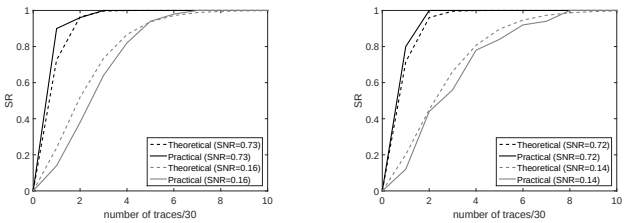
attack. From the methodology aspect, it shows that the better profiled the model is, the more realistic prediction of *SR* can be made from the *TVLA* results.

6 Multivariate Analysis

In its current form, *TVLA* metric can not be applied in multivariate analysis without modifying its formulation. Recently in [9], the limitations of *TVLA* in detection of multivariate side channel vulnerabilities was addressed in details for higher order analysis. In [20], the authors have focussed on extending *TVLA* methodology to higher order leakage detection. Consequently, a strategy for applying *d-th order d-variate TVLA test* is given. A typical application for such analysis can be a software implementation of d^{th} order masking, where shares are executed sequentially. Our approach in this section is different from them as we focus on *1st order d-variate TVLA test* where d denotes the dimension of a single side channel trace. We investigate the extension of proposed methodology for unprotected implementation in multivariate setting for side-channel vulnerability quantification. Therefore, the weaknesses pointed out in [9], do not apply to our setting. Moreover, in this section we try to extend applicability of *TVLA* from univariate to multivariate settings to address one of the shortcoming of traditional *TVLA* [9].

6.1 Proposed Formulation

To obtain *SR* for multivariate side channel analysis, we can follow two different approaches. We can either compute *TVLA* on each sample and then combine



(a) sample point with imperfect HW leakage model

(b) sample point with nearly perfect HW leakage model

Fig. 6: Comparison Between Theoretical *SR* and Practical *SR* for different *SNRs* using first order stochastic model at different sample points

those values to get the corresponding *SR* in multivariate settings, or combine the different sample points using an optimal dimensionality reduction formulation to convert the multivariate side channel traces into a single point. For latter, we use the framework of [24]. In particular, the traces Y arise from a single leakage model L, which depend on the correct key $k = k^*$, and which is taken standard (i.e., $\mathbb{E}(L) = 0$, $\mathsf{Var}(L) = 1$), through the relationship:

$$Y_d = \alpha_d L(k^*) + N_d,$$

where d is the dimensionality ($1 \leq d \leq D$).

Remark 2. This equation implies $\mathbb{E}(Y) = 0$. When computing a t-test, using non-specific or specific, the evaluator also has to evaluate $\mathbb{E}(Y|X = x_0)$ for a given plaintext (or a given byte value of the plaintext) x_0. Let's assume that $\mathbb{E}(Y|X = x_0) = c \neq 0$. The condition $\neq 0$ is here to avoid having $\mathbb{E}(Y) = \mathbb{E}(Y|X = x_0)$, in which case the attacker would conclude the device is secure whereas in practice it is not (e.g. for a different value of x_0', we would have $\mathbb{E}(Y) \neq \mathbb{E}(Y|X = x_0')$).

In matrix form, for Q number of side channel traces, we can write the above equation as below:

$$Y^{D,Q} = \alpha^D L^Q(k^*) + N^D,$$

Here α_D is a non-zero vector of length D, and can be calculated as follows [24]:

$$\alpha^D = \frac{Y^D (L^Q(k^*))^T}{L^Q(k^*) L^Q(k^*)^T}. \tag{11}$$

We assume that the noise N^D is multivariate normal, and we denote by $\boldsymbol{\Sigma}$ its $D \times D$ covariance matrix. The value of $\boldsymbol{\Sigma}$ can be computed as below [24]:

$$\Sigma = \frac{1}{Q-1}(Y^{D,Q} - \alpha^D L^Q(k^*))(Y^{D,Q} - \alpha^D L^Q(k^*))^T. \tag{12}$$

With the knowledge of α^D and Σ, we can now calculate the optimal reduction formulation as the optimal dimensionality reduction is $\frac{(\boldsymbol{\alpha^D})^T \boldsymbol{\Sigma^{-1}} \boldsymbol{Y^{D,Q}}}{(\boldsymbol{\alpha^D})^T \boldsymbol{\Sigma^{-1}} \boldsymbol{\alpha^D}}$ [24].

SNR* and *TVLA* in multivariate settings** To compute the *SNR* and *TVLA* in multivariate settings, we propose following pre-processing steps. Here by ***boldface we denote multivariate trace of dimension D.

- Step 1: Compute $\boldsymbol{\Sigma}$,
- Step 2: Standardize the measurements, that is: $\boldsymbol{Y}$ becomes $\boldsymbol{Y}' = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{Y}$.

Notice that $\boldsymbol{Y}' = (\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\alpha})L + \boldsymbol{N}'$, where $\boldsymbol{N}'$ is now an isotropic standard noise (all D samples of noise are i.i.d., of mean 0 and variance 1). Indeed,

$$\mathbb{E}(\boldsymbol{N}'(\boldsymbol{N}')^{\mathsf{T}}) = \mathbb{E}(\boldsymbol{\Sigma}^{-1/2}\boldsymbol{N}\boldsymbol{N}^{\mathsf{T}}\boldsymbol{\Sigma}^{-1/2}) = \boldsymbol{\Sigma}^{-1/2}\mathbb{E}(\boldsymbol{N}\boldsymbol{N}^{\mathsf{T}})\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{I}, \quad (13)$$

where $\boldsymbol{I}$ is the $D \times D$ identity matrix.

On step 2, we can now re-estimate $\boldsymbol{\mu}'_1$, as $\mathbb{E}(\boldsymbol{Y}')$. For the sake of clarity, we drop index 1 and 2 in $\boldsymbol{\mu}$ (when it is clear given the context). We see that the optimal dimensionality reduction is (theorem 1 of [24])

$$\frac{(\boldsymbol{\mu}')^{\mathsf{T}}\boldsymbol{Y}'}{(\boldsymbol{\mu}')^{\mathsf{T}}\boldsymbol{\mu}'} = \|\boldsymbol{\mu}'\|^{-2}\ (\boldsymbol{\mu}')^{\mathsf{T}}\boldsymbol{Y}'. \quad (14)$$

Consequently, we can define multivariate *SNR* and multivariate *TVLA* as follow:

$$\mathsf{SNR} = (\boldsymbol{\mu}')^{\mathsf{T}}\boldsymbol{\mu}' = \sum_{d=1}^{D}(\mu'_d)^2. \quad (15) \qquad \mathsf{TVLA}^2 = \sum_{d=1}^{D}\frac{(\mu'_{1,d} - \mu'_{2,d})^2}{\frac{1}{n_1} + \frac{1}{n_2}} \quad (16)$$

because $\sigma'_{1,d} = \sigma'_{2,d} = 1$ (by (13)).

Remark 3. This is equal to (up to an irrelevant $\frac{1}{4}$ proportionality factor) the Hotelling's T-Square [25]). Indeed, let us consider that $n_1 = n_2 = n/2$. We have:

$$\mathsf{TVLA}^2 = \sum_{d=1}^{D}\frac{(\mu'_{1,d} - \mu'_{2,d})^2}{\frac{1}{n_1} + \frac{1}{n_2}} = \frac{1}{4}n(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\mathsf{T}}\Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2). \quad (17)$$

The definition of multivariate *SNR* (15) and multivariate *TVLA* (16) remains consistent with the dimensionality reduction (14). Namely we have:

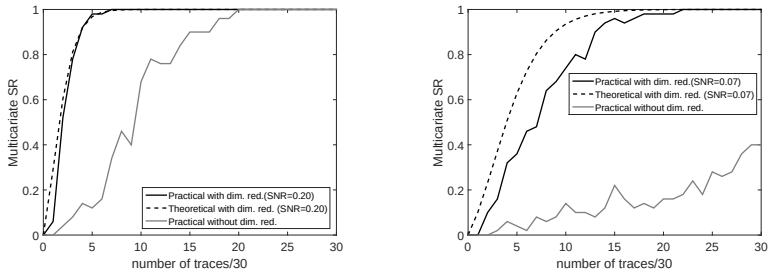
Proposition 3. *The application of univariate* SNR *(resp* TVLA*) of reduced trace* (14) *yield multivariate* SNR (15) *(resp. multivariate* TVLA (16)*)*

Proof. The derivation is provided in appendix C

6.2 Analysis

Multivariate setting of the proposed methodology is now experimentally validated. We first apply *optimal dimension reduction* technique on the acquired traces to convert them to univariate traces from multivariate one. As shown in proposition 3 that multivariate *SNR* computed on the multivariate traces is equivalent to the

univariate *SNR* computed on the dimension reduced traces. Hence, we can use our proposed methodology for univariate traces on the dimension reduced traces and can compute the theoretical *SR* and practical *SR* (see Fig. 7). The theoretical predictions are compared with practical attacks on reduced dimension traces. Fig. 7 shows that the proposed formulation for computation of theoretical *SR* follows practical *SR* which successfully validates our proposed methodology for computation of *SR* in first order multivariate settings. It must be noted that the *SNR* shown in Fig. 7 is computed after applying dimension reduction.



(a) (b)

Fig. 7: Comparison Between Theoretical *SR* and Practical *SR* for different *SNRs* in multivariate settings

7 Conclusion

Though *TVLA* based testing methodology is becoming popular due to its simplicity and integrability with standard testing mechanism, it does not give much information about the side-channel resistance of the target. In this paper, we make a first attempt to extend the *TVLA* based testing methodology beyond its current scope. Analytic relationship between *TVLA* and *SNR* derived, which allows to directly compute *SR* from *TVLA* test. By connecting *TVLA* with *SR*, an attempt is made to bridge the gap between conformance based testing and evaluation based testing, addressing both side channel leakage detection and side channel leakage quantification. The methodology is successfully verified on an unprotected AES smart-card implementation in a simulated setting as well as practical measurements. The theoretical and practical results are show to match, specially under a well profiled model. The proposed methodology is further extended to address unprotected implementation with multivariate leakage, with supporting results. Further extension of this approach to protected implementation, specially using the formulation of [20] would be an interesting direction.

References

1. Paul Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In *Annual International Cryptology Conference*, pages 388–397. Springer, 1999.

2. The Common Criteria. https://www.commoncriteriaportal.org/. Accessed: 2016-09-25.
3. FIPS 1403 DRAFT Security Requirements for Cryptographic Modules (Revised Draft). http://csrc.nist.gov/publications/drafts/fips1403/reviseddraftfips1403_PDFzip_documentannexAtoannexG.zip.
4. Jaffe J. Goodwill G., Jun B. and Rohatgi P. A testing methodology for side-channel resistance validation. http://csrc.nist.gov/news_events/non-invasive-attack-testing-workshop/papers/08_Goodwill.pdf, 2011.
5. G. Becker, J. Cooper, E. DeMulder, G. Goodwill, J. Jaffe, G. Kenworthy, T. Kouzminov, A. Leiserson, M.Marson, P. Rohatgi, and S. Saab. Test Vector Leakage Assessment (TVLA) methodology in practice. http://icmc-2013.org/wp/wp-content/uploads/2013/09/Rohatgi_Test-Vector-Leakage-Assessment.pdf, 2013.
6. François Durvaux and François-Xavier Standaert. From improved leakage detection to the detection of points of interests in leakage traces. *IACR Cryptology ePrint Archive*, 2015:536, 2015.
7. Jean-Sébastien Coron, Aurélien Greuet, Emmanuel Prouff, and Rina Zeitoun. Faster evaluation of sboxes via common shares. In *International Conference on Cryptographic Hardware and Embedded Systems*, pages 498–514. Springer, 2016.
8. François Durvaux, François-Xavier Standaert, and Nicolas Veyrat-Charvillon. How to certify the leakage of a chip? In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 459–476. Springer, 2014.
9. Franois-Xavier Standaert. How (not) to use welch's t-test in side-channel security evaluations. Cryptology ePrint Archive, Report 2017/138, 2017. http://eprint.iacr.org/2017/138.
10. François Durvaux, François-Xavier Standaert, and Santos Merino Del Pozo. Towards easy leakage certification. In *International Conference on Cryptographic Hardware and Embedded Systems*, pages 40–60. Springer, 2016.
11. Yunsi Fei, Qiasi Luo, and A. Adam Ding. A statistical model for DPA with novel algorithmic confusion analysis. In *Cryptographic Hardware and Embedded Systems - CHES 2012 - 14th International Workshop, Leuven, Belgium, September 9-12, 2012. Proceedings*, pages 233–250, 2012.
12. Adrian Thillard, Emmanuel Prouff, and Thomas Roche. Success through confidence: Evaluating the effectiveness of a side-channel attack. *IACR Cryptology ePrint Archive*, 2015:402, 2015.
13. Yunsi Fei, A. Adam Ding, Jian Lao, and Liwei Zhang. A statistics-based success rate model for DPA and CPA. *J. Cryptographic Engineering*, 5(4):227–243, 2015.
14. François-Xavier Standaert, Tal G Malkin, and Moti Yung. A unified framework for the analysis of side-channel key recovery attacks. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 443–461. Springer, 2009.
15. Shivam Bhasin, Jean-Luc Danger, Sylvain Guilley, and Zakaria Najm. Side-channel leakage and trace compression using normalized inter-class variance. *IACR Cryptology ePrint Archive*, 2014:1020, 2014.
16. Stefan Mangard, Elisabeth Oswald, and Thomas Popp. *Power Analysis Attacks: Revealing the Secrets of Smart Cards*. Springer, December 2006. ISBN 0-387-30857-1, http://www.dpabook.org/.
17. Sébastien Tiran, Guillaume Reymond, Jean-Baptiste Rigaud, Driss Aboulkassimi, Benedikt Gierlichs, Mathieu Carbone, Gilles R. Ducharme, and Philippe Maurine. Analysis of variance and CPA in SCA. *IACR Cryptology ePrint Archive*, 2014:707, 2014.

18. Shivam Bhasin, Jean-Luc Danger, Sylvain Guilley, and Zakaria Najm. Side-channel Leakage and Trace Compression Using Normalized Inter-class Variance. In *Proceedings of the Third Workshop on Hardware and Architectural Support for Security and Privacy*, HASP '14, pages 7:1–7:9, New York, NY, USA, 2014. ACM.
19. Luke Mather, Elisabeth Oswald, Joe Bandenburg, and Marcin Wójcik. Does my device leak information? an a priori statistical power analysis of leakage detection tests. In *Advances in Cryptology - ASIACRYPT 2013 - 19th International Conference on the Theory and Application of Cryptology and Information Security, Bengaluru, India, December 1-5, 2013, Proceedings, Part I*, pages 486–505, 2013.
20. Tobias Schneider and Amir Moradi. Leakage assessment methodology - extended version. *J. Cryptographic Engineering*, 6(2):85–99, 2016.
21. A. Adam Ding, Cong Chen, and Thomas Eisenbarth. Simpler, faster, and more robust t-test based leakage detection. Cryptology ePrint Archive, Report 2015/1215, 2015. `http://eprint.iacr.org/2015/1215`.
22. SASEBO-GII. `satoh.cs.uec.ac.jp/SAKURA/hardware/SASEBO-GII.html`. Accessed: 2016-09-25.
23. Werner Schindler, Kerstin Lemke, and Christof Paar. A stochastic model for differential side channel cryptanalysis. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pages 30–46. Springer, 2005.
24. Nicolas Bruneau, Sylvain Guilley, Annelie Heuser, Damien Marion, and Olivier Rioul. Less is More - Dimensionality Reduction from a Theoretical Perspective. In *Cryptographic Hardware and Embedded Systems - CHES 2015 - 17th International Workshop, Saint-Malo, France, September 13-16, 2015, Proceedings*, pages 22–41, 2015.
25. Harold Hotelling. The generalization of student's ratio. *Ann. Math. Statist.*, 2(3):360–378, 08 1931.

A Proof of Theorem 1

Proof. From Eqn. (3) we can write NICV_2 as below:

$$\begin{aligned}\mathsf{NICV}_2 &= \frac{\frac{1}{n}\sum_{i=1}^{2} n_i(\mu_i-\mu)^2}{\frac{1}{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i}(Y_{i,j}-\mu)^2} \\ &= \frac{\frac{1}{n}\sum_{i=1}^{2} n_i(\mu_i-\mu)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2} \end{aligned} \tag{18}$$

From Eqn. (5) we can write $TVLA$ as follows:

$$\begin{aligned}\mathsf{TVLA} &= \frac{\mu_1-\mu_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \\ \mathsf{TVLA}^2 &= \frac{(\mu_1-\mu_2)^2}{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\end{aligned}$$

$$= \frac{C}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{19}$$

where $C = (\mu_1 - \mu_2)^2$. Now we will consider only the numerator part of the NICV_2 formulation which is

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{2} n_i \left(\mu_i - \mu\right)^2 \\
&= \frac{1}{n}\left(n_1\left(\mu_1 - \frac{n_1\mu_1 + n_2\mu_2}{n}\right)^2 + n_2\left(\mu_2 - \frac{n_1\mu_1 + n_2\mu_2}{n}\right)^2\right) \\
&= \frac{1}{n}\left(\frac{n_1 n_2^2}{n^2}\left(\mu_1 - \mu_2\right)^2 + \frac{n_1^2 n_2}{n^2}\left(\mu_1 - \mu_2\right)^2\right) \\
&= \frac{n_1 n_2 (n_1 + n_2)}{n^3} C \\
&= \frac{n_1 n_2}{n^2} C
\end{aligned} \tag{20}$$

Next we will consider the denominator part of the NICV computation which is as follows:

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} \left(Y_i - \mu\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(Y_i^2 - \frac{2Y_i\left(n_1\mu_1 + n_2\mu_2\right)}{n} + \frac{\left(n_1\mu_1 + n_2\mu_2\right)^2}{n^2}\right) \\
&= \frac{1}{n}\sum_{Y_i \in Y^{G1}}\left(Y_i^2 - \frac{2Y_i\left(n_1\mu_1 + n_2\mu_2\right)}{n}\right) + \frac{1}{n}\sum_{Y_i \in Y^{G2}}\left(Y_i^2 - \frac{2Y_i\left(n_1\mu_1 + n_2\mu_2\right)}{n}\right) + \frac{\left(n_1\mu_1 + n_2\mu_2\right)^2}{n^2} \\
&= \frac{1}{n}\sum_{Y_i \in Y^{G1}}\left(Y_i^2 - 2Y_i\mu_1 + \mu_1^2 + \left(\frac{2Y_i n_2\left(\mu_1 - \mu_2\right)}{n} - \mu_1^2\right)\right) \\
&+ \frac{1}{n}\sum_{Y_i \in Y^{G2}}\left(Y_i^2 - 2Y_i\mu_2 + \mu_2^2 + \left(\frac{2Y_i n_1\left(\mu_2 - \mu_1\right)}{n} - \mu_1^2\right)\right) + \frac{\left(n_1\mu_1 + n_2\mu_2\right)^2}{n^2} \\
&= \frac{n_1}{n}\sigma_1^2 + \frac{n_2}{n}\sigma_2^2 + \frac{n_1 n_2}{n} C
\end{aligned} \tag{21}$$

We can now combine Eqn. (18), (19), (20) and (21) to achieve the desired formulation

$$\begin{aligned}
\mathsf{NICV}_2 &= \frac{\frac{n_1 n_2}{n^2} C}{\frac{n_1}{n}\sigma_1^2 + \frac{n_2}{n}\sigma_2^2 + \frac{n_1 n_2}{n^2} C} \\
&= \frac{C}{n\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} + \sigma_1^2\left(\frac{1}{n_2} - \frac{1}{n_1}\right) + \sigma_2^2\left(\frac{1}{n_1} - \frac{1}{n_2}\right)\right) + C}
\end{aligned}$$

$$= \frac{1}{n\frac{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}{C} + \frac{n}{C}(\sigma_1^2-\sigma_2^2)\left(\frac{1}{n_2}-\frac{1}{n_1}\right)+1}$$

Thus we can write NICV_2 as

$$\mathsf{NICV}_2 = \frac{1}{\frac{n}{\mathsf{TVLA}^2} + \frac{n}{C}(\sigma_1^2-\sigma_2^2)\left(\frac{1}{n_2}-\frac{1}{n_1}\right)+1}.$$

B Proof of Theorem 2

Proof. From Eqn. (18), we can compute NICV_2^i as below

$$\begin{aligned}
\mathsf{NICV}_2^i &= \frac{\frac{1}{n}\left(n_i\left(\mu_i-\mu\right)^2 + \left(n-n_i\right)\left(\mu_{\overline{i}}-\mu\right)^2\right)}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2} \\
&= \frac{\frac{1}{n}\left(n_i\left(\mu_i-\mu\right)^2 + \frac{1}{n-n_i}\left(\frac{n_i\sum_{j=1}^{j=k} n_j\mu_j - nn_i\mu_i}{n}\right)^2\right)}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2} \\
&= \frac{\frac{n_i}{\overline{n_i}}\left(\mu_i-\mu\right)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2}, where \quad \overline{n_i} = n-n_i
\end{aligned} \tag{22}$$

Now if we add each NICV_2^i, we will get the following relationship

$$\begin{aligned}
\sum_{i=1}^{k}\mathsf{NICV}_2^i &= \frac{\sum_{i=1}^{k}\frac{n_i}{\overline{n_i}}(\mu_i-\mu)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2} \\
&= \frac{\sum_{i=1}^{k}\frac{n}{\overline{n_i}}\frac{n_i}{n}(\mu_i-\mu)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2} = \frac{\sum_{i=1}^{k}(1+\frac{n_i}{\overline{n_i}})\frac{n_i}{n}(\mu_i-\mu)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2} \\
&= \frac{\sum_{i=1}^{k}\frac{n_i}{n}(\mu_i-\mu)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2} + \frac{\sum_{i=1}^{k}\frac{n_i^2}{n\overline{n_i}}(\mu_i-\mu)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j-\mu)^2}
\end{aligned}$$

(23)

From Eqn. (18), we can write NICV_k as follows

$$\mathsf{NICV}_k = \frac{\sum\limits_{i=1}^{k} \frac{n_i}{n}(\mu_i - \mu)^2}{\frac{1}{n}\sum\limits_{j=1}^{n}(Y_j - \mu)^2} \tag{24}$$

Combining Eqn. (23) and (24), we arrive at the following relation

$$\sum_{i=1}^{k} \mathsf{NICV}_2^i = \mathsf{NICV}_k + \frac{\sum\limits_{i=1}^{k} \frac{n_i^2}{n\overline{n}_i}(\mu_i - \mu)^2}{\frac{1}{n}\sum\limits_{j=1}^{n}(Y_j - \mu)^2} \tag{25}$$

Using the assumption of uniform setting, we presume that each group has same number of side channel traces. Then, Eqn. (23) becomes

$$\begin{aligned}\sum_{i=1}^{k} \mathsf{NICV}_2^i &= \frac{\frac{1}{k}\sum\limits_{i=1}^{k}(\mu_i - \mu)^2 + \frac{1}{k(k-1)}\sum\limits_{i=1}^{k}(\mu_i - \mu)^2}{\frac{1}{n}\sum\limits_{j=1}^{n}(Y_j - \mu)^2} \\ &= \frac{\frac{k}{k-1}\frac{1}{k}\sum\limits_{i=1}^{k}(\mu_i - \mu)^2}{\frac{1}{n}\sum\limits_{j=1}^{n}(Y_j - \mu)^2} = \frac{k}{k-1}\mathsf{NICV}_k \end{aligned} \tag{26}$$

Thus we arrive at the desired formulation

$$NICV_k = \frac{k-1}{k}\sum_{i=1}^{k} NICV_2^i.$$

It must be noted that NICV_k is actually the generalized *NICV* which was introduced in [15].

C Proof of Proposition 3

Proof. After dimensionality reduction, we get:

$$Y'' = L + \frac{1}{\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}} \boldsymbol{\mu}'^T \boldsymbol{N}'$$

For the *SNR*, we thus have:

- **signal**: $\mathsf{Var}(L) = 1$;
- **noise**:

$$\frac{1}{(\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})^2} \mathsf{Var}(\boldsymbol{\mu}'^T \boldsymbol{N}') = \frac{1}{\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}} \tag{27}$$

Hence *SNR* is $\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$, which is is equal to (15).

Regarding *TVLA*, we will assume that $\mathbb{E}(Y) = \boldsymbol{\mu_1} = 0$, and $\mathbb{E}(Y|X = x_0) = \boldsymbol{\mu_2} = c\boldsymbol{\mu}$. Hence, after dimensionality reduction (14), one gets

- reduced average for random plaintext: 0,
- reduced average for fixed plaintext $= x_0$: c,
- reduced noise has variance (27).

Hence the univariate (squared) *TVLA* on reduced traces is

$$c^2(\boldsymbol{\mu^T} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}).$$

Now, the multivariate (squared) *TVLA* (16) is (using Hotteling formula (17):

$$\frac{1}{4} n(\mathbf{0} - c\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{0} - c\boldsymbol{\mu}),$$

which also match with the *TVLA* expression obtained after dimensionality reduction. It must be noted that this formulation is applicable to both specific and non-specific *TVLA* test.