Security Analysis of Higher-Order Boolean Masking Schemes for Block Ciphers

(with Conditions of Perfect Masking)

Gilles Piret[1], François-Xavier Standaert[2⋆]

[1]Ecole Normale Supérieure, Département d'Informatique,
45, Rue d'Ulm, 75230 Paris cedex 05, France

[2]UCL Crypto Group, Université Catholique de Louvain
Place du Levant, 3, B-1348 Louvain-La-Neuve, Belgium

Gilles.Piret@ens.fr,standaert@dice.ucl.ac.be

Abstract. Side-channel attacks are an important class of cryptanalytic techniques against cryptographic implementations and masking is a frequently considered solution to improve the resistance of a cryptographic implementation against side-channel attacks. In this paper, we consequently analyze the security of higher-order Boolean masking schemes in various contexts. Our results are twofold. First, we formalize the definitions of higher-order side-channel attacks with the related security notions and put forward certain security weaknesses in recently proposed masking schemes. Second, we investigate the conditions upon which a substitution box in a block cipher can be perfectly masked by boolean values in order to counteract side-channel attacks. That is, can the leakages statistical distributions at a masked S-box output (over all possible masks) be independent of the secret key targeted in the attacks? We study the consequences of this requirement in two commonly considered leakage models, namely the Hamming weight and distance models and derive conditions on the substitution boxes. As a result of our analysis, it appears that these conditions are not achievable as they lead to evident cryptanalytic weaknesses. Thus this work formally confirms that masking cannot be used as a stand-alone countermeasure and cannot offer provable security against side-channel attacks.

⋆ Postdoctoral researcher funded by the Belgian Fund for Scientific Research (FNRS).

1 Introduction

Side-channel attacks are an important class of cryptanalytic techniques in which an adversary attempts to take advantage of some physical leakages obtained from a cryptographic implementation in order to recover secret information. Power consumption and electromagnetic radiation in smart cards are typical examples of such leakages. Since their apparition in the late 1990s [KJJ99], they have been shown extremely efficient to defeat a variety of implementations of secret and public key cryptosystems. Among various other proposals, the idea of masking the intermediate values inside a cryptographic algorithm was consequently suggested in several papers, *e.g.* [GP99,AG03,OMPR05] as a possible countermeasure to side-channel attacks. The technique is generally applicable if all the fundamental operations used in a given algorithm can be rewritten in a masked domain. For example, in a Boolean masking scheme, one (or more) random value(s) is (are) initially XORed to the plaintext and the encryption algorithm is modified in such a way that the running data is always hidden by a known mask. Masking can be applied at the gate level [FG05,PM05] or algorithmic level (we focus on the latter one). Unfortunately, none of these solutions allow to perfectly protect an implementation. For example, higher-order side-channel attacks that have been discussed in a number of articles are often capable to break these countermeasures, *e.g.* in [Mes00b,WW04,SPQ05,PSDQ05,OMHT06]. As a consequence, this paper discusses the conditions upon which one can prevent such higher-order techniques. We note that higher-order attacks are not the only concern about masking but we believe they receive a sufficient attention in the recent years to benefit from a more formal treatment.

For this purpose, we first describe two settings of masking schemes, respectively denoted as the hardware and software approaches[1]. The first one keeps the cipher's S-box **S** unchanged, and uses a mask update function **S'** to compute the output mask from the masked input to the original S-box and the input mask. The drawback of this solution is that **S'** takes a $2n$-bit input, which can make it costly to evaluate in

[1] In present block ciphers, it is generally assumed that the most critical components to mask are the non-linear S-boxes. For simplicity and without loss of generality (since we aim to demonstrate an impossibility result), our investigations will be focused on the combination of a key addition and the application of a layer of S-boxes.

some implementations. The second one initially picks up the S-boxes input and output masks and re-computes a modified S-box **S*** from the original S-box and this pair of random masks. The drawback of this solution is that **S*** has to be re-computed for every new pair of masks and S-box which is a highly time-consuming process. Both approaches are pictured in Figure 1. It is interesting to note that, from a security point of view, these solutions differ by the mask dependencies. In the hardware approach, the output mask is deterministically specified by its inputs. On the other hand, in the software approach, the input and output masks are picked up independently. We note that the names "*hardware vs. software*" approaches were chosen since the schemes in Figure 1 will typically take advantage of hardware *vs.* software implementation facilities. However, one could think about various other masking schemes. We believe these two examples are representative of most countermeasures acting at the algorithmic level.

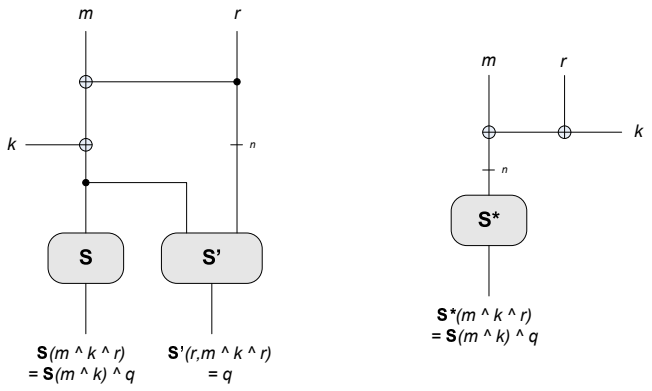

Fig. 1. Masking schemes: the hardware and software approaches.

From a theoretical point of view, Prouff showed in [Pro05] the link between the intrinsic resistance of *unmasked* S-boxes to power analysis and their cryptographic properties. In this paper, we aim to complement this analysis and derive the conditions upon which an S-box can be *perfectly masked* to counteract side-channel attacks. That is, can the leakages statistical distributions (taken over all possible masks) at a masked S-box output be independent of the secret key targeted in the attacks? As a first (but practically meaningful) case study, we study the consequences of this condition in two commonly considered leakage models: the Hamming weight and distance models.

Our following analysis first considers the hardware setting. Our results show that (1) in the Hamming weight model, perfect masking requires the original S-box outputs to have a constant Hamming weight; (2) in the Hamming distance model, perfect masking requires the S-boxes to be extremely weak with respect to linear and differential cryptanalysis attacks. In both cases, these conditions are not achievable since they lead to evident theoretical cryptanalytic weaknesses in the block ciphers structure.

Second, we investigate the software setting. We start by putting forward security weaknesses, assuming that an adversary not only observes the computations in Figure 1 but also the re-computation procedure of the S-box **S***. By combining both measurements, we show that it is possible to build attacks based on the construction of temporary templates corresponding to a re-computation leakage table. Then, we show that, in the Hamming weight leakage model, perfect masking requires S-boxes with constant outputs, which is not an acceptable condition for secure ciphers.

Finally, we briefly discuss attacks that do not take advantage of non-linear S-boxes and directly target mask additions. We underline that, in certain contexts, such attacks can be particularly powerful and have to be taken into account.

We note that the aim of this paper is mainly to derive theoretical conditions on the block cipher components in order to reach *perfect* security. Therefore, the practical implementation of higher-order attacks and their efficiency evaluation are not investigated. We refer to the previously mentioned publications for these purposes. In addition, our analysis is not based on particular flaws in the investigated schemes by contrast, *e.g.* with the recent work in [CPR07]. They only show that masking cannot totally prevent side-channel attacks in two reasonable leakage models. In addition, we do not claim that actual implementations are uniquely protected with the schemes of Figure 1. As a matter of fact, there is a large distance between perfect security and practical security. We finally acknowledge that a part of the presented results have been intuitively put forward in previous works. Beyond the work of Prouff [Pro05], Carlet *et al.* [Car05] and Guilley *et al.* [GHP04] illustrated the relations between the cryptographic strength of an S-box and its resistance against side-channel attacks. The analysis of ghost peaks can also be seen as related to the same issues, *e.g.* in [BCO04,CC05]. We believe that the analysis presented in the following of this paper (that extends this intuition to masking) is useful since it allows deriving formal statements from these known intuitions.

The rest of the paper is structured as follows. Section 2 defines our model for higher-order side-channel attacks as well as the notion of perfect masking. Section 3 describes higher-order masking schemes in the hardware approach and section 4 discusses the conditions of perfect masking in this setting. Section 5 evaluates the security of a higher-order masking technique in the software approach and describes an attack based on the combined observation of the encryption algorithm and the S-box re-computation algorithm. The conditions of perfect masking against this latter attack are also given. Attacks that directly target the mask additions are briefly discussed in section 6. Finally, our conclusions are in section 7.

2 Model and definitions

We model our side-channel attacks following the principles introduced in [SMY06], assuming that an adversary encrypts a number N of plaintext messages m_i's under the same secret key k. For each of the encrypted messages, it measures the leakage (*e.g.* power consumption) of a target device and obtains observations $O_i = \mathsf{L}(\Sigma, R)$'s, where L is a leakage function depending on some secret machine state Σ (typically, the target of the attack) and randomness R. We note that our investigations only consider simple leakage models. As a first (but practically meaningful) step in the analysis of masking schemes, we consider the Hamming weight and distance leakage functions that have been used in a number of practical attacks. By contrast, template attacks [CRR02] are not considered in our analysis. Additionally, our proofs consider perfect models in the sense that the leakages are not affected by random noise. Typically, given some secret state Σ or pair of consecutive secret states Σ_1, Σ_2, the adversary obtains either the Hamming weight of Σ or the Hamming distance between Σ_1 and Σ_1. Since our results show that even these simple leakage models give rise to unrealistic conditions of perfect masking, they constitute an interesting first step in the formal understanding of higher-order side-channel attacks. Also, we initially *only consider side-channel attacks that target the S-box outputs in a block cipher*. As frequently mentioned in the literature (*e.g.* [BCO04]), side-channel attacks are usually performed after a non linear component because they better discriminate the different key candidates. However, it is worth noting that targeting the addition of the masks directly may be feasible as well, depending on the leakage model considered (as discussed in Section 6).

We define a higher-order side-channel attack as follows:

Definition 1. *A side-channel attack of order d against a n-bit target is a side-channel attack based on the observation of the activity and leakage of at most $n \cdot d$ bits in at most d different instants within an observable implementation.*

Example: Let us consider the hardware (*i.e.* left) scheme of Figure 1 with bit size $n = 8$ and target the key k. A first-order attack is an attack based on the observation of the activity and leakage of only 8 bits. For example, one could observe the leakage of $\mathbf{S}(m \oplus k) \oplus q$ or the one of q (that are both 8-bit wide). It is straightforward that *when used separately*, those leakages do not reveal anything on k as long as the random generation of r is not biased. Indeed, q is independent of k and $\mathbf{S}(m \oplus k) \oplus q$ is uniformly distributed over the q values. By contrast, a second-order attack is an attack based on the observation of the activity and leakage of 16 bits. For example, one could observe the *combined* leakage of $\mathbf{S}(m \oplus k) \oplus q$ and of q to mount such an attack.

One important point of this definition is that an attack of order d is *not* defined with respect to the product of d leakage traces (*e.g.* as in [WW04]) but with respect to the size of the statistical distribution required to perform an attack. Indeed, it has been shown in [PSDQ05] that the product of the leakage traces leads to a suboptimal distinguisher. Similarly, [CKN00] argued that perfect security against side-channel attacks requires the statistical independence between the target data and the physical observations. Another straightforward point is that an attack of order d includes the attacks of lower orders. However, in practice an adversary is only interested in the attack of lowest possible degree since it determines an implementation security.

Definition 2. *A masking scheme of order*[2] *d is a masking scheme for which the lowest degree of a successful side-channel attack is $d+1$.*

Definition 3. *An r-mask scheme on a n-bit target is a masking scheme using $n \cdot r$ random bits to mask any set of n bits in an implementation.*

[2] [CPR07] use a slightly different terminology. They consider that a *masking scheme* is of order d if every sensitive variable in the algorithm is split into d shares. Therefore, any masking scheme of order d can be defeated by an attack of order d. For example, a single mask countermeasure is of order 1 in our terminology and of order 2 in [CPR07]. But (most importantly) our definitions for an *attack* of order d are similar.

Definition 4. *A perfect masking scheme is a masking scheme secure against attacks of any possible order, i.e. for which the leakages are statistically independent of the data processed within the target physically observable device.*

With respect to this latter definition, it is again important to have in mind that perfect masking is a theoretically relevant notion. But not reaching this level of security (as it will be shown in the next sections) does not involve that an implementation is insecure. As a matter of fact, a masking scheme combined with a sufficient amount of noise in the physical observations (or combined with other countermeasures) may lead to reasonable levels of security. That is, the practical efficiency of an attack has to be measured too.

3 Higher-order masking in the HW approach

In this section we discuss d-mask schemes of order d ($d \geq 1$) suitable for implementations in hardware. In its basic version, such a scheme first XORs the d masks with the input; then the key is XORed as well before the application of $\mathbf{S}$. So we have:

$$\boxed{\begin{array}{c} m_1 = m \oplus r_1 \\ m_2 = m_1 \oplus r_2 \\ \cdots \\ m_d = m_{d-1} \oplus r_d \\ m^* = m_d \oplus k \end{array}}$$

And finally:

$$m^* = m \oplus k \oplus \bigoplus_{i=1}^{d} r_i \tag{1}$$

The output masks $q_1, ..., q_d$ corresponding to $r_1, ..., r_d$ are equal to:

$$\begin{aligned} q_1 &= \mathbf{S}(m \oplus k) \oplus \mathbf{S}(m \oplus k \oplus r_1), \\ q_j &= \mathbf{S}\left(m \oplus k \oplus \bigoplus_{i=1}^{j-1} r_i\right) \oplus \mathbf{S}\left(m \oplus k \oplus \bigoplus_{i=1}^{j} r_i\right) \quad (j = 2 \ldots d), \end{aligned} \tag{2}$$

so that we have:

$$\mathbf{S}(m^*) \oplus q_1 \oplus ... \oplus q_d = \mathbf{S}(m \oplus k) \tag{3}$$

We denote the mask update functions as: $q_j = \mathbf{S}^{[j]}(m^*, r_j, ..., r_d),\ (j = 1 \ldots d)$. Figure 2 illustrates the whole circuit in the case $d = 2$. It is easy to see that such a d-mask scheme

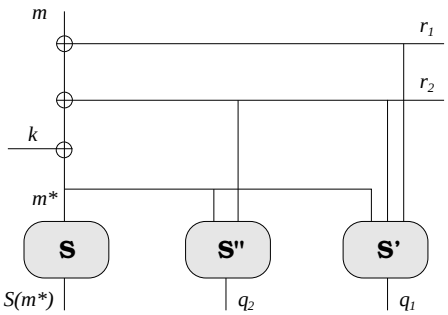

Fig. 2. Higher-order hardware masking scheme: basic version.

is indeed of order d, if only the S-box outputs are observed. Since the r_i's are chosen to be independent random variables with uniform distribution,

$$(\mathbf{S}(m \oplus k \oplus r_1), \mathbf{S}(m \oplus k \oplus r_1 \oplus r_2), ..., \mathbf{S}(m \oplus k \oplus r_1 \oplus ... \oplus r_d)) \tag{4}$$

is uniformly random[3], and so is $(q_1, ..., q_d)$ as it is in bijection with (4). We can conclude that observing the leakage of d n-bit words among:

$$(\mathbf{S}(m \oplus k) \oplus q_1 \oplus ... \oplus q_d, q_1, ..., q_d) \tag{5}$$

in the Hamming weight model (*resp.* among $(\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k) \oplus q_1 \oplus ... \oplus q_d \oplus q'_1 \oplus ... \oplus q'_d, q_1 \oplus q'_1, ..., q_d \oplus q'_d)$ in the Hamming distance model, where the data m, m' and the masks q_i, q'_i correspond to two consecutive clock cycles) does not provide any information on the key. The drawback of this scheme (more precisely, of the way we suggest to implement it) is that a d-mask scheme requires the computation of functions with input space of size up to $(d+1) \cdot n$ bits, which in some implementations becomes rapidly infeasible as d grows. It is why another implementation is worth considering. It relies on doing key addition after the *first* mask has been XORed:

[3] provided $\mathbf{S}$ is surjective with all elements of the output space having the same number of predecessors, which is usually the case for actual block ciphers.

$$
\boxed{
\begin{array}{c}
m_1 = m \oplus r_1 \\
m^* = m_1 \oplus k \\
m_2 = m^* \oplus r_2 \\
\cdots \\
m_d = m_{d-1} \oplus r_d
\end{array}
}
$$

In this case, the output masks can be computed as:

$$
\begin{aligned}
q_1 &= \mathbf{S'}(r_1, m^*), \\
q_j &= \mathbf{S'}(r_j, m_j) \qquad (j = 2 \ldots d),
\end{aligned}
\tag{6}
$$

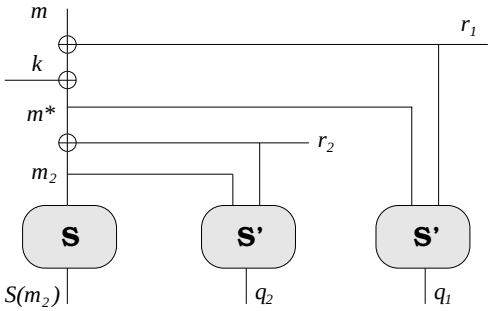

Fig. 3. Higher-order hardware masking scheme: more efficient version.

where the same function $\mathbf{S'}(a, b) := \mathbf{S}(b) \oplus \mathbf{S}(a \oplus b)$ (with two inputs only) is used for all masks. Figure 3 illustrates the circuit in the case $d = 2$. The drawback of this approach is that if an attacker is able to observe the leakage relative to m^* in addition to the one of q_1, it can deduce key information. For example, in the Hamming weight model, the probability distribution (over r_1) of:

$$
WH[m \oplus k \oplus r_1] + WH[\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m \oplus k \oplus r_1)] \tag{7}
$$

is key-dependent. However, in the context of a Hamming distance leakage model, this is not true anymore since $WH(m \oplus k \oplus r_1 \oplus m' \oplus k \oplus r'_1)$ does not depend on the key. As in the hardware approach the latter leakage model is generally in use, this solution may be relevant. In general, if we assume (due to our preliminary hypothesis) that the only observable leakage is the one corresponding to the data after the S-box layer, then the masking scheme is of order d (the output of this layer is actually the same as in

the first scheme). Attacks relaxing this hypothesis will be discussed in Section 6. We note that in general, the Hamming weight model is typically applicable to pre-charged data buses in microprocessors while the Hamming distance model is rather applicable to attacks on registers in hardware implementations. However, this is not an absolute statement, *e.g.* see [SMPQ06]. In the following and for theoretical completeness, we consider both leakage models for our two masking schemes settings.

4 Conditions of perfect masking in the HW approach

4.1 Conditions of perfect 1-mask schemes

In this section, we consider the 1-mask scheme represented in the left part of Figure 1 and discuss the conditions upon which it could lead to perfect masking. For this purpose, we will investigate the two previously mentioned leakage functions, *i.e.* the Hamming weight and distance models. Perfect masking is achieved if the probability distribution of the leakage function is independent of the key. The theorems of this section will make use of the following technical lemma:

Lemma 1. *Let Σ, q be n-bit words and assume the $2n$-bit word $V = (\Sigma \oplus q, q)$ is computed in a hardware device. We denote its Hamming weight as:*

$$O(\Sigma, q) = WH[V] = WH[\Sigma \oplus q] + WH[q] \tag{8}$$

Then the probability distribution of O, computed over q, only depends on $WH[\Sigma]$. Moreover, two different values for $WH[\Sigma]$ always imply two different distributions.

Proof. For each $\Sigma = \sigma$, we are interested in the probability distribution (computed over q) of the observation O, that is $\Pr_q[O|\Sigma = \sigma]$. First, we note that:

$$WH[\sigma \oplus q] = WH[\sigma] + WH[q] - 2WH[\sigma \wedge q] \tag{9}$$

where $\wedge$ denotes the logical and. Thus we have:

$$O(\sigma, q) = WH[\sigma] + 2WH[q] - 2WH[\sigma \wedge q] \tag{10}$$

Let $h_\sigma = WH[\sigma]$. We obtain:

$$\Pr_q[O = h_\sigma + 2p|\Sigma = \sigma] = \Pr_q[WH[q] - WH[\sigma \wedge q] = p] \tag{11}$$

We have to compute the number of values of q satisfying:

$$WH[q] - WH[\sigma \wedge q] = p \tag{12}$$

The value of bit q_i does not matter at the h_σ bit positions i with $\sigma_i = 1$. Amongst the $n - h_\sigma$ positions i with $\sigma_i = 0$, exactly p bits q_i must satisfy $q_i = 1$. Therefore there are $2^{h_\sigma} \cdot \binom{n-h_\sigma}{p}$ values of q satisfying (12). So, we obtain:

$$\Pr[O = h_\sigma + 2p | \Sigma = \sigma] = 2^{h_\sigma} \cdot \binom{n - h_\sigma}{p} / 2^n \text{ for } p \in \{0, 1, ..., n - h_\sigma\} \tag{13}$$

We observe that this probability distribution only depends on the Hamming weight h_σ, and that two different Hamming weights imply two different distributions. □

Perfect 1-mask schemes in the Hamming weight model

Theorem 1. *In the Hamming weight model, the probability distribution of the Hamming weight at the output of a masked S-box* ($\boldsymbol{S}, \boldsymbol{S'}$) *is independent of the key if and only if (iff) all possible outputs of* $\boldsymbol{S}$ *have the same Hamming weight.*

Proof. Let $V = (V_1, V_2) = (\mathbf{S}(b \oplus k) \oplus q, q)$ be the $2n$-bit word considered in the side-channel attack, where $\mathbf{S}(b \oplus k) =: \Sigma$ is the secret state and q is the mask. The corresponding Hamming weight is:

$$O(\Sigma, q) = WH[V] = WH[\Sigma \oplus q] + WH[q] \tag{14}$$

We can apply lemma 1. We conclude that in order to obtain always the same distribution whatever the key, all the S-box's outputs must have the same Hamming weight. It is also straightforward that this condition is sufficient. □

Perfect 1-mask schemes in the Hamming distance model

We prove that perfect masking is possible only for S-boxes of which the best linear approximation has correlation 1, and the best differential has probability 1. Before giving the proof, we recall the definitions of the λ-parameter and δ-parameter which characterize the intrinsic resistance of the S-box to linear and differential cryptanalysis.

Definition 5. *The* λ***-parameter*** *of an S-box* $\boldsymbol{S}$*:*$\{0,1\}^p \to \{0,1\}^q$ *is defined as:*

$$\lambda_S = \max_{\substack{\alpha \in \{0,1\}^p \\ 0 \neq \beta \in \{0,1\}^q}} |2^{1-p} \cdot \#\{x \in \{0,1\}^p | \alpha \bullet x = \beta \bullet \boldsymbol{S}(x)\} - 1|$$

where $\bullet$ *denotes the scalar product.*

Definition 6. *The* δ***-parameter*** *of an S-box* $\boldsymbol{S}$*:*$\{0,1\}^p \rightarrow \{0,1\}^q$ *is defined as:*

$$\delta_S = 2^{-p} \cdot \max_{\substack{0 \neq a \in \{0,1\}^p \\ b \in \{0,1\}^q}} \#\{x \in \{0,1\}^p | \boldsymbol{S}(x \oplus a) \oplus \boldsymbol{S}(x) = b\}$$

Regarding the linear parameter, our result is as follows:

Theorem 2. *Consider an S-box* $\boldsymbol{S} : \{0,1\}^p \rightarrow \{0,1\}^t$*. In the Hamming distance model, if the probability distribution of the leakages at the output of its masked implementation* ($\boldsymbol{S}, \boldsymbol{S'}$) *is independent of the key, then* $\lambda_{\boldsymbol{S}} = 1$*.*

Proof. In this case, the leakage measurement corresponds to a $2t$-bit word:

$$V = (\mathbf{S}(m \oplus k) \oplus q \oplus \mathbf{S}(m' \oplus k) \oplus q', q \oplus q'), \tag{15}$$

where the data m, m' (*resp.* the masks q, q') correspond to two consecutive clock cycles. So we are interested in the probability distribution of:

$$WH[\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k) \oplus q \oplus q'] + WH[q \oplus q'] \tag{16}$$

Let $q^{\Delta} := q \oplus q'$, $\mathbf{S}^{\Delta}(k) := \mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k)$. Then (16) becomes:

$$WH[\mathbf{S}^{\Delta}(k) \oplus q^{\Delta}] + WH[q^{\Delta}] \tag{17}$$

Again we can apply Lemma 1. Thus $\mathbf{S}$ must be such that:

$$\forall m, m', k : WH[\mathbf{S}(m) \oplus \mathbf{S}(m')] = WH[\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k)] \tag{18}$$

Let us define $WH_2[x] := WH[x] \mod 2$. Moreover $\mathbf{s}_i(x)$ denotes the i^{th} output bit of $\mathbf{S}(x)$ ($i = 1, ..., t$). Then (18) implies:

$$\begin{aligned} &\forall m, m', k : WH_2[\mathbf{S}(m) \oplus \mathbf{S}(m')] = WH_2[\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k)] \\ \Leftrightarrow &\forall m, m', k : \bigoplus_{i=1}^{t} (\mathbf{s}_i(m) \oplus \mathbf{s}_i(m')) = \bigoplus_{i=1}^{t} (\mathbf{s}_i(m \oplus k) \oplus \mathbf{s}_i(m' \oplus k)) \end{aligned} \tag{19}$$

If we define $\mathbf{s}^*(x) := \bigoplus_{i=1}^{t} \mathbf{s}_i(x)$, (19) can be rewritten as:

$$\forall m, m', k : \mathbf{s}^*(m) \oplus \mathbf{s}^*(m') = \mathbf{s}^*(m \oplus k) \oplus \mathbf{s}^*(m' \oplus k) \tag{20}$$

And by setting $m' = 0$ we obtain:

$$\mathbf{s}^*(m \oplus k) = \mathbf{s}^*(m) \oplus \mathbf{s}^*(k) \oplus \mathbf{s}^*(0), \tag{21}$$

which shows that s^* is affine. Therefore $\exists \alpha = (\alpha_1, ..., \alpha_p) \in \{0,1\}^p$ such that

$$\forall x \in \{0,1\}^p : \bigoplus_{i=1}^{t} \mathbf{s}_i(x) = \alpha \bullet x \oplus \mathbf{s}^*(0) \tag{22}$$

This implies that $\lambda_{\mathbf{S}} = 1$. □

Its counterpart for the differential parameter is:

Theorem 3. *Consider* $\boldsymbol{S} : \{0,1\}^p \rightarrow \{0,1\}^q$ *with* $p \geq q$.[4] *In the Hamming distance model, if the probability distribution of the leakages at the output of the masked implementation* $(\boldsymbol{S}, \boldsymbol{S'})$ *is independent of the key, then* $\delta_{\boldsymbol{S}} = 1$.

Proof. The proof begins like the one of Theorem 2. We use Lemma 1 to obtain:

$$\forall m, m', k : WH[\mathbf{S}(m) \oplus \mathbf{S}(m')] = WH[\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k)] \tag{18}$$

We consider two cases:

- First, we assume there exist $m, m' \in \{0,1\}^p$ such that $\mathbf{S}(m) \oplus \mathbf{S}(m') = 2^q - 1$. Therefore we have

  $$WH[\mathbf{S}(m) \oplus \mathbf{S}(m')] = q \tag{23}$$

  By using (18), we obtain:

  $$\begin{aligned} &\forall k : WH[\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k)] = q \\ \Leftrightarrow &\forall k : \mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k) = 2^q - 1 \\ \Leftrightarrow &\forall k : \mathbf{S}(k) \oplus \mathbf{S}(m \oplus m' \oplus k) = 2^q - 1 \end{aligned} \tag{24}$$

  Therefore $m \oplus m' \rightarrow 2^q - 1$ is a differential of probability 1 and $\delta_{\mathbf{S}} = 1$.
- Now suppose there is no $m, m' \in \{0,1\}^p$ such that $\mathbf{S}(m) \oplus \mathbf{S}(m') = 2^q - 1$. It implies that the function is not injective; as a matter of fact, such an injective function should satisfy $\mathbf{S}(0) \neq \mathbf{S}(a) \neq \mathbf{S}(0) \oplus 2^q - 1$ for all $a \neq 0$. But it would imply a collision elsewhere, as $\#\{0 \neq a \in \{0,1\}^p\} = 2^p - 1 > 2^q - 2$. So there exist $m \neq m'$ with $WH[\mathbf{S}(m) \oplus \mathbf{S}(m')] = 0$. Using (18), we have:

  $$\begin{aligned} &\forall k : WH[\mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k)] = 0 \\ \Leftrightarrow &\forall k : \mathbf{S}(m \oplus k) \oplus \mathbf{S}(m' \oplus k) = 0 \\ \Leftrightarrow &\forall k : \mathbf{S}(k) \oplus \mathbf{S}(m \oplus m' \oplus k) = 0 \end{aligned} \tag{25}$$

  Therefore $m \oplus m' \rightarrow 0$ is a differential of probability 1 and $\delta_{\mathbf{S}} = 1$. □

[4] We already made this hypothesis in section 3 (see footnote 3).

As a matter of fact, none of these conditions of perfect masking (in the Hamming weight and distance leakage models) make sense from a cryptanalytic point of view.

4.2 Conditions for perfect n-mask schemes

In this section we extend our analysis to higher-order schemes presented in Section 3. We show that necessary conditions on $\mathbf{S}$ in order to achieve perfect masking are actually the same as in the 1-mask case.

Perfect n-mask schemes in the Hamming weight model

We consider the probability distribution of:

$$O(\Sigma, q_1, ..., q_d) = WH[\Sigma \oplus q_1 \oplus \ldots \oplus q_d] + WH[q_1] + \ldots + WH[q_d], \tag{26}$$

with $\Sigma := \mathbf{S}(m \oplus k)$. That is, we are interested in $\Pr_{q_1,\ldots,q_d}[O|\Sigma = \sigma]$. Let us consider the smallest value of O having non-zero probability. It is easy to see that it is equal to $WH[\sigma]$. Therefore if $\mathbf{S}(m \oplus k)$ and $\mathbf{S}(m \oplus k')$ have different Hamming weights, the corresponding probability distributions will be different. So a necessary condition on $\mathbf{S}$ is that all its outputs have the same Hamming weight. It is a sufficient condition as well, as permuting bits of σ obviously does not change the probability distribution.

Perfect n-mask schemes in the Hamming distance model

In this case we are interested in the probability distribution of:

$$\begin{aligned} WH[\mathbf{S}(m_1 \oplus k) \oplus \mathbf{S}(m_2 \oplus k) \oplus (q_1 \oplus q'_1) \oplus \ldots \oplus (q_d \oplus q'_d)] \\ + WH[q_1 \oplus q'_1] + \ldots + WH[q_d \oplus q'_d] \end{aligned} \tag{27}$$

Let $\mathbf{S}^{\Delta}(k) := \mathbf{S}(m_1 \oplus k) \oplus \mathbf{S}(m_2 \oplus k)$, $q_i^{\Delta} := q_i \oplus q'_i$ $(i = 1 \ldots d)$.
Then (27) can be rewritten as:

$$WH[\mathbf{S}^{\Delta}(k) \oplus q_1^{\Delta} \oplus \ldots \oplus q_d^{\Delta}] + WH[q_1^{\Delta}] + \ldots + WH[q_d^{\Delta}] \tag{28}$$

We showed in the previous paragraph that this distribution is independent of the key only if all elements of $\mathrm{Im}(S^{\Delta})$ have the same Hamming weight.
So $\mathbf{S}$ must satisfy:

$$\forall m_1, m_2, k : WH[\mathbf{S}(m_1) \oplus \mathbf{S}(m_2)] = WH[\mathbf{S}(m_1 \oplus k) \oplus \mathbf{S}(m_2 \oplus k)]. \tag{29}$$

We already showed in the proofs of theorems 2 and 3 that this condition implies $\lambda_{\mathbf{S}} = \delta_{\mathbf{S}} = 1$. Therefore, our conclusions for 1-mask schemes hold.

Remark: Equation (18) does not imply that $\mathbf{S}$ is linear. A counterexample for $p{=}t{=}4$ is $(\mathbf{S}(0), \mathbf{S}(1), ..., \mathbf{S}(15)) = (0, 12, 3, 15, 10, 9, 6, 5, 9, 10, 5, 6, 12, 0, 15, 3)$, for which $\mathbf{S}(2) \oplus \mathbf{S}(4) \neq \mathbf{S}(6)$. There exist bijective counterexamples for S-boxes of larger dimensions.

5 Security of higher-order masking in the SW approach

The previous section demonstrated that perfect masking is only possible in the hardware approach under very restrictive (and practically unrealistic) conditions for the S-boxes. The reason behind these conditions is that there always remains a statistical relation between the cipher inputs and outputs: given one input plaintext and mask, the masked output and output mask are deterministically specified. Otherwise said, the efficiency weakness (*i.e.* the need of an S-box **S'** implemented in the circuit) is also the source of a security weakness. As already mentioned, in the software approach, such dependencies do not exist. Therefore, it is natural to investigate if similar weaknesses can be found. In particular, does the efficiency weakness (*i.e.* the need to recompute **S*** for every mask) affect the security of the countermeasure? For illustration, we take the

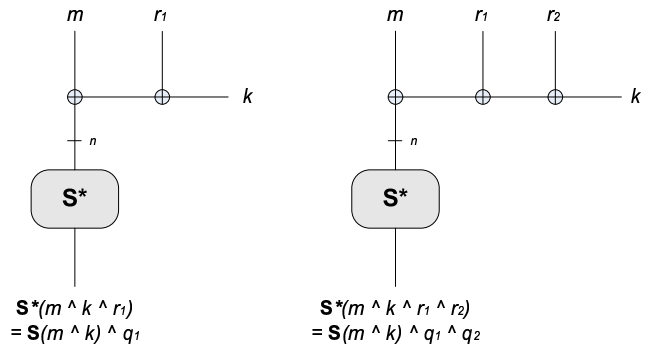

Fig. 4. Masking schemes of order 1 and 2 in the software approach.

higher-order masking scheme of the AES described in [SP06] as our running example. It is represented in Figure 4 for the cases of first and second-order masking. As such, these schemes cannot be the target of any attack of degree lower than 2 (for the first-order scheme) and 3 (for the second-order scheme), as predicted in the original paper. However, our aim is now to analyze these schemes, under the reasonable assumption that the adversary not only observes the computation of $\mathbf{S}(m \oplus k) \oplus q_1 \oplus ...$ but also

the re-computation procedure of **S*** that has to be performed before any new encryption. For illustration purposes, we first consider the following simple re-computation algorithm that was originally proposed in [Mes00a], for an 8-bit S-box:

for $i = 0 : 255$

$A = \mathbf{S}(i)$;

$\mathbf{S^*}(i \oplus r_1) = A \oplus q_1$;

end

More complex re-computation solutions will be discussed later in the section.

5.1 Attack against the first-order masking scheme

Let us consider a leakage function L and the first-order masking scheme of Figure 4. The attack proceeds as follows:

% *initialization of the key candidates likelihoods:*

$\mathbf{L}(k_c)$=1, $\forall k_c$'s;

% *attack:*

FOR various messages m_j's and random masks r_{1j}'s

% *storage of a re-computation leakage table:*

FOR i=0:255

$k_c = m_j \oplus i$;

$\mathbf{R}(j, k_c) = \mathsf{L}(\mathbf{S^*}(i \oplus r_{1j}))$;

END

% *storage of the observed measurements:*

$\mathbf{M}(j) = \mathsf{L}(\mathbf{S^*}(m_j \oplus k \oplus r_{1j}))$;

% *computation of the key candidates likelihoods:*

FOR all k_c's

$\mathbf{L}(k_c) = \mathbf{L}(k_c) * \Pr(\mathbf{M}(j)|\mathbf{R}(j, k_c))$;

END

% *normalize the likelihood vector* $\mathbf{L}$;

END

The attack is successful as soon as $\max(\mathbf{L}(k_c)) = \mathbf{L}(k)$.

Roughly explained, we store the leakages corresponding to the S-box re-computation for all possible inputs of the S-box and compare the leakage of a real message encryption under an unknown key with these re-computations. The attack is possible because we can compare the real output of the masked S-box $\mathbf{S}(m \oplus k) \oplus q_1$ and the re-computations $S(i) \oplus q_1$. Therefore, we can identify $m \oplus k$ with the most likely i. Note that the attack can be viewed as a first-order side-channel attack based on the construction of temporary templates (*i.e.* the re-computation leakage table is then derived offline) or as a second-order attack. This is because during the computation of the key candidates likelihoods, we compare the measured leakage with only one re-computation table. A very similar technique is applied in the collision-based attacks of [SLFP04]. Note also that the use of Bayesian key recovery is of particular interest in this context since it allows to limit the memory requirements of the attack. Indeed, it is not necessary to keep the matrix $\mathbf{R}$ and vector $\mathbf{M}$ in memory for all messages but for the running one.

5.2 Attack against higher-order masking schemes

Let us now consider a second-order masking scheme with the naive S-box recomputation procedure that follows:

for $i = 0 : 255$
$A = \mathbf{S}(i)$;
$\mathbf{S_1}(i \oplus r_1) = A \oplus q_1$;
end

for $i = 0 : 255$
$A = \mathbf{S_1}(i)$;
$\mathbf{S^*}(i \oplus r_2) = A \oplus q_2$;
end

The re-computation is sound since we have (see Figure 4):

$$\mathbf{S^*}(m \oplus k \oplus r_1 \oplus r_2) = \mathbf{S_1}(m \oplus k \oplus r_1) \oplus q_2 = \mathbf{S}(m \oplus k) \oplus q_1 \oplus q_2$$

As a consequence, an attack similar to the previous one holds. That is, in a first step, we identify $\mathbf{S^*}(m \oplus k \oplus r_1 \oplus r_2) = \mathbf{S_1}(m \oplus k \oplus r_1) \oplus q_2$ with the second re-computation. It yields the likelihood vector corresponding to all the candidates for the value $\mathbf{S_1}(m \oplus k \oplus r_1) = \mathbf{S}(m \oplus k) \oplus q_1$. Then, *by computing the leakage corresponding to these candidates*, we identify $\mathbf{S}(m \oplus k) \oplus q_1$ with the first re-computation. It yields the likelihood vector for the key candidates. It is important to remark that this attack requires to evaluate the leakage $\mathsf{L}(\mathbf{S_1}(m \oplus k \oplus r_1))$ which requires either that the

adversary has access to a nearly perfect leakage model or that it can perform this computation within a device under its control. This makes the technique less practical than for the first-order masking. In addition, the likelihood of the keys is evaluated in two steps: first by evaluating the likelihood of $\mathbf{S_1}(m \oplus k \oplus r_1)$, then by evaluating the one of the target key candidates:

$$\mathbf{L}(k_c) = \sum_{\mathbf{S_1}(m \oplus k \oplus r_1)} \mathbf{L}(k_c | \mathbf{S_1}(m \oplus k \oplus r_1)) \cdot \mathbf{L}(\mathbf{S_1}(m \oplus k \oplus r_1)) \tag{30}$$

That is, the attack can now be viewed as combination of two first-order side-channel attack based on the construction of two temporary templates or as a third order attack. The generalization of this attack to higher-orders is straightforward: any d-mask scheme is susceptible to an attack of order $n_r + 1$, where n_r is the number of re-computations performed for the S-boxes. Or it can be viewed as a combination of first-order template attacks based on the construction of n_r temporary templates. Note that in practice, the complete re-computation procedure is usually only done for the first S-box. Once a first S-box is masked with respective input and output masks $r_1, r_2, ...$ and $q_1, q_2, ...$, and a second S-box has to be masked with $u_1, u_2, ...$ and $v_1, v_2, ...$, one can simply re-compute the first masked S-box once, with masks $u_1 \oplus r_1 \oplus u_2 \oplus r_2 ...$ and $v_1 \oplus q_1 \oplus v_2 \oplus q_2 ...$ This allows improving the efficiency of the masked implementation.

Importantly, as far as the security of the countermeasure is concerned, these attacks do not exhibit/exploit particular flaws in the masking countermeasure. The aim of describing them is mainly to derive the conditions of perfect masking in the next section. By contrast, the recent work of [CPR07] showed that the countermeasure in [SP06] are the target of a 3rd-order side-channel attack, whatever the order of the masking scheme. These attacks have been successfully experimented and raise the open question of designing a higher-order side-channel resistant scheme.

Note that a straightforward improvement of the countermeasure would be to randomize the re-computation of the S-boxes. That is, rather than using a regular loop for $i \in [0 : 255]$, use a loop for $i \in \text{PERM}[0 : 255]$ where PERM is a random permutation of the vector. Different fast re-computation algorithms are proposed in [SP06]. Such solutions increase the order of a successful side-channel attack, but still in a limited way with respect to the order of the masking scheme.

5.3 Perfect S-box masking in the software approach

For a n-bit S-box, let us consider the probability distributions of the leakages associated with the 2^n re-computation steps. If any two of these distributions were different, an attacker would be able to guess the differences of leakages between two different keys. On the contrary, if all leakages $\mathsf{L}(\mathbf{S}^*(i \oplus r_1))$ are equal, all measurements always come from the same probability distribution, which prevents attacks. The following theorem gives one condition to obtain such perfect masking. Once again, it makes perfect masking practically impossible.

Theorem 4. *Consider an S-box $\boldsymbol{S}$, and its software-masked implementation $\boldsymbol{S}^*$. In the Hamming weight model, the probability distribution of the leakages at the output of $\boldsymbol{S}^*$ is independent of its input iff all outputs of $\boldsymbol{S}$ are equal.*

Proof. As $S^*(i \oplus r_1) = S(i) \oplus q_1$, the condition amounts to:

$$\forall q_1, i, i' : WH[S(i) \oplus q_1] = WH[S(i') \oplus q_1] \tag{31}$$

Assume that for some i, i', $S(i)$ and $S(i')$ differ in their b^{th} bit (at least), with $s_b(i) = 0$ and $s_b(i') = 1$ (their other bits can be either equal or different). Let $q_1 = \delta_b$ which is equal to 1 in its b^{th} bit and to 0 in all the others. Then we have:

$$WH[S(i)] + 1 = WH[S(i) \oplus \delta_b] = WH[S(i') \oplus \delta_b] = WH[S(i')] - 1, \tag{32}$$

where the second equality comes from (31). Because by considering $q_1 = 0$ we also have $WH[S(i)] = WH[S(i')]$, we obtain a contradiction. □

6 Higher-order attacks before the S-boxes

As already mentioned, all our previous analysis rely on the assumption that one only exploits the leakages after the application of the non-linear S-boxes. However, depending on the leakage functions considered, it may be possible to target the XOR of the masks directly. For example, it is typically the case in the context of a Hamming weight-based model. Importantly, in the latter context, the masks and masked values have to be manipulated with care. For example, in a naive implementation of the masking scheme of Section 3, an attack of order 2 is possible: indeed, the distribution of $(m \oplus r_1 \oplus r_2 \oplus ...,$ $m \oplus k \oplus r_1 \oplus r_2 \oplus ...)$ is always key dependent.

As a consequence, it is finally worth noting that, if the degree of the masking scheme increases, an attack that directly targets the mask addition may be more efficient than if using the masked S-boxes, despite the fact that such an attack is usually more difficult to mount (because they do not exploit the substitution non-linearities). Importantly, even in the context of hardware implementations where the Hamming distance model is prominent, Hamming weight dependencies may be exploited [SMPQ06].

7 Conclusion

This paper discusses the security of higher-order masking schemes for block ciphers. It is shown that perfect masking implies unrealistic conditions on the block cipher components (*e.g.* the constant Hamming weight of their outputs or weak linear/differential parameters). Additionally, our results underline the similarities between different implementation approaches for masking. We demonstrate that the computation of a masked S-box leads to weaknesses, no matter if it is done once and stored in memory (*i.e.* in a hardware approach) or re-computed for every new encryption (*i.e.* in a software approach). In the context of the software approach, we also describe higher-order attacks based on the monitoring of the re-computation process. We note that, with respect to the presented results, increasing the degree of a masking scheme improves its security, but such a trend has to be tempered by the possibility to carry out an attack directly on the mask addition. In addition, particular flaws have been exhibited for a number of masking schemes, *e.g.* in [CPR07]. The practical impact of higher-order masking schemes consequently has to be quantified carefully for different side-channel adversaries, *e.g.* based on a difference-of-mean test [OMPR05], correlation analysis [SP06] or Bayesian approach [PSDQ05]. Although a number of recent research work have tackled this question, certain contexts still require further investigation.

Acknowledgements: The authors would like to thank anonymous reviewers of IET Information Security for their meaningful comments and interesting references.

References

[AG03] Mehdi-Laurent Akkar and Louis Goubin, *A Generic Protection Against High-Order Differential Power Analysis.*, in Johansson [Joh03], pp. 192–205.

[BCO04] Eric Brier, Christophe Clavier, and Francis Olivier, *Correlation Power Analysis with a Leakage Model.*, in Joye and Quisquater [JQ04], pp. 16–29.

[Car05] Claude Carlet, *On Highly Nonlinear S-Boxes and their Inability to Thwart DPA Attacks.*, in Maitra et al. [MMV05], pp. 49–62.

[CC05] Cecile Canovas, Jessy Clediere, *What Do S-Boxes Say in Differential Side-Channel Attacks?*, Cryptology ePrint Archive, Rep. 2005/311, `http://eprint.iacr.org/`.

[cKKP99] Çetin Kaya Koç and Christof Paar (eds.), *Cryptographic Hardware and Embedded Systems, first international workshop, CHES'99, Worcester, MA, USA, August 12-13, 1999, proceedings*, LNCS, vol. 1717, Springer, 1999.

[cKKP00] Çetin Kaya Koç, Christof Paar (eds.), *Cryptographic Hardware and Embedded Systems - CHES 2000, second international workshop, Worcester, MA, USA, August 17-18, 2000, proceedings*, LNCS, vol. 1965, Springer, 2000.

[CKN00] Jean-Sébastien Coron, Paul C. Kocher, and David Naccache, *Statistics and Secret Leakage.*, in Frankel [Fra01], pp. 157–173.

[CPR07] Jean-Sébastien Coron, Emmanuel Prouff, and Matthieu Rivain, *Side Channel Cryptanalysis of a Higher Order Masking Scheme*, in the proceedings of CHES'07.

[CRR02] Suresh Chari, Josyula R. Rao, and Pankaj Rohatgi, *Template Attacks.*, in Jr. et al. [JcKKP03], pp. 13–28.

[DBL05] *International Symposium on Information Technology: Coding and Computing (ITCC), vol. 1, 4-6 April 2005, Las Vegas, USA*, IEEE Computer Society, 2005.

[FG05] Wieland Fischer and Berndt M. Gammel, *Masking at Gate Level in the Presence of Glitches.*, in Rao and Sunar [RS05], pp. 187–200.

[Fra01] Yair Frankel (ed.), *Financial Cryptography, 4th international conference, FC 2000, Anguilla, British West Indies, February 20-24, 2000, proceedings*, Lecture Notes in Computer Science, vol. 1962, Springer, 2001.

[GH05] Henri Gilbert and Helena Handschuh (eds.), *Fast Software Encryption: 12th international workshop, FSE 2005, Paris, France, February 21-23, 2005*, Lecture Notes in Computer Science, vol. 3557, Springer, 2005.

[GHP04] Sylvain Guilley, Philippe Hoogvorst, and Renaud Pacalet, *Differential Power Analysis Model and Some Results.*, in Quisquater et al. [QPDK04], pp. 127–142.

[GM06] Louis Goubin and Mitsuru Matsui (eds.), *Cryptographic Hardware and Embedded Systems - CHES 2006, 8th international workshop, Yokohama, Japan, October 10-13, 2006, proceedings*, LNCS, vol. 4249, Springer, 2006.

[GP99] Louis Goubin and Jacques Patarin, *Des and Differential Power Analysis (the "Duplication" Method).*, in Çetin Kaya Koç and Paar [cKKP99], pp. 158–172.

[JcKKP03] Burton S. Kaliski Jr., Çetin Kaya Koç, and Christof Paar (eds.), *Cryptographic Hardware and Embedded Systems - CHES 2002, 4th international workshop, Redwood Shores, CA, USA, August 13-15, 2002, revised papers*, Lecture Notes in Computer Science, vol. 2523, Springer, 2003.

[Joh03] Thomas Johansson (ed.), *Fast Software Encryption, 10th international workshop, FSE 2003, Lund, Sweden, February 24-26, 2003, revised papers*, Lecture Notes in Computer Science, vol. 2887, Springer, 2003.

[JQ04] Marc Joye, Jean-Jacques Quisquater (eds.), *Cryptographic Hardware and Embedded Systems - CHES 2004: 6th international workshop, Cambridge, MA, USA, August 11-13, 2004. proceedings*, LNCS, vol. 3156, Springer, 2004.

[KJJ99] Paul C. Kocher, Joshua Jaffe, and Benjamin Jun, *Differential Power Analysis.*, in Wiener [Wie99], pp. 388–397.

[Mes00a] Thomas S. Messerges, *Securing the AES Finalists Against Power Analysis Attacks.*, in Schneier [Sch01], pp. 150–164.

[Mes00b] ———, *Using Second-Order Power Analysis to Attack DPA Resistant Software.*, in Çetin Kaya Koç and Paar [cKKP00], pp. 238–251.

[MMV05] Subhamoy Maitra, C. E. Veni Madhavan, and Ramarathnam Venkatesan (eds.), *Progress in Cryptology - Indocrypt 2005, 6th international conference on cryptology in India, Bangalore, India, December 10-12, 2005, proceedings*, Lecture Notes in Computer Science, vol. 3797, Springer, 2005.

[OMHT06] Elisabeth Oswald, Stefan Mangard, Christoph Herbst, and Stefan Tillich, *Practical Second-Order DPA Attacks for Masked Smart Card Implementations of Block Ciphers.*, in Pointcheval [Poi06], pp. 192–207.

[OMPR05] Elisabeth Oswald, Stefan Mangard, Norbert Pramstaller, and Vincent Rijmen, *A Side-Channel Analysis Resistant Description of the AES S-box.*, in Gilbert and Handschuh [GH05], pp. 413–423.

[PM05] Thomas Popp and Stefan Mangard, *Masked Dual-Rail Pre-Charge Logic: Dpa-Resistance Without Routing Constraints.*, in Rao and Sunar [RS05], pp. 172–186.

[Poi06] David Pointcheval (ed.), *Topics in Cryptology - CT-RSA 2006, the Cryptographers' Track at the RSA Conference 2006, San Jose, CA, USA, February 13-17, 2006, proceedings*, Lecture Notes in Computer Science, vol. 3860, Springer, 2006.

[Pro05] Emmanuel Prouff, *DPA Attacks and S-Boxes*, in [GH05], pp. 424–441.

[PSDQ05] Eric Peeters, François-Xavier Standaert, Nicolas Donckers, and Jean-Jacques Quisquater, *Improved Higher-Order Side-Channel Attacks with FPGA Experiments.*, in Rao and Sunar [RS05], pp. 309–323.

[QPDK04] Jean-Jacques Quisquater, Pierre Paradinas, Yves Deswarte, and Anas Abou El Kalam (eds.), *Smart Card Research and Advanced Applications VI, IFIP 18th World Computer Congress, TC8/WG8.8 & tcTC11/WG11.2? sixth international conference on Smart Card Research and Advanced Applications (CARDIS), 22-27 August 2004, Toulouse, France*, Kluwer, 2004.

[RS05] Josyula R. Rao and Berk Sunar (eds.), *Cryptographic Hardware and Embedded Systems - CHES 2005, 7th international workshop, Edinburgh, UK, August 29 - September 1, 2005, proceedings*, LNCS, vol. 3659, Springer, 2005.

[Sch01] Bruce Schneier (ed.), *Fast Software Encryption, 7th international workshop, FSE 2000, New York, NY, USA, April 10-12, 2000, proceedings*, Lecture Notes in Computer Science, vol. 1978, Springer, 2001.

[SLFP04] Kai Schramm, Gregor Leander, Patrick Felke, Christof Paar, *A Collision-Attack on AES: Combining Side Channel- and Differential-Attack.*, in [JQ04], pp. 163–175.

[SMPQ06] François-Xavier Standaert, François Macé, Eric Peeters, and Jean-Jacques Quisquater, *Updates on the Security of FPGAs against Power Analysis Attacks.*, Reconfigurable Computing Architectures and Applications, 2006, pp. 335–346.

[SMY06] François-Xavier Standaert, Tal G. Malkin, and Moti Yung, *A Formal Practice-Oriented Model for the Analysis of Side-Channel Attacks*, Cryptology ePrint Archive, Report 2006/139, 2006, `http://eprint.iacr.org/`.

[SP06] Kai Schramm and Christof Paar, *Higher Order Masking of the AES.*, in Pointcheval [Poi06], pp. 208–225.

[SPAQ06] François-Xavier Standaert, Eric Peeters, Cédric Archambeau, and Jean-Jacques Quisquater, *Towards Security Limits in Side-Channel Attacks.*, in Goubin and Matsui [GM06], pp. 30–45.

[SPQ05] François-Xavier Standaert, Eric Peeters, and Jean-Jacques Quisquater, *On the Masking Countermeasure and Higher-Order Power Analysis Attacks.*, in *ITCC (1)* [DBL05], pp. 562–567.

[Wie99] Michael J. Wiener (ed.), *Advances in Cryptology - CRYPTO '99, 19th annual international cryptology conference, Santa Barbara, California, USA, August 15-19, 1999, proceedings*, LNCS, vol. 1666, Springer, 1999.

[WW04] Jason Waddle and David Wagner, *Towards Efficient Second-Order Power Analysis.*, in Joye and Quisquater [JQ04], pp. 1–15.