A relation calculus for reasoning about t -probing security

Maria Chiara Molteni^{1,3} · Vittorio Zaccaria²

Received: 18 April 2020 / Accepted: 18 January 2022 / Published online: 26 February 2022
© The Author(s) 2022

Abstract

In the context of side-channel attacks against cryptographic circuits, t -probing security characterizes the amount of information derivable about sensitive values (e.g., keys) by observing t output/internal values. Non-interference is a useful mathematical tool used by researchers to assess the probing security of a circuit which employs Boolean masking to protect itself from attacks. However, reasoning about non-interference still requires either difficult ratiocination or complex automatic tools. In this work, we propose a novel point of view to reason about non-interference, by exploiting the Walsh transform of a Boolean function. To this end, we introduce a calculus for mechanically reasoning about the shares of a variable and show that this formalism provides a lean algebraic explanation of known compositional patterns allowing for the discovery of new ones. Eventually, we show how this formalism can be applied to study the probing security of known cryptographic gadgets.

Keywords Boolean functions · Random variables · Correlation immunity · Spectral characterization · Walsh transform

Mathematics Subject Classification 94C10 · 94A60

1 Introduction

An attacker can mount a side-channel attack against a cryptographic circuit by probing its internal nodes to derive information correlated with the secret. The first countermeasure against this type of attack appeared almost two decades ago [1] and it officially gave birth to a branch of research known today as t -probing security. Since then, it became clear that, while proving the security for small gadgets required a small effort, reasoning about their composition was (and still is) not trivial. In fact, one of the main problems addressed is the *composability* of security proper-

ties, i.e., determining, given two t -probing¹ secure gadgets², if their functional composition is still t -probing secure.

Over the years, it has been observed that composability depends on the amount of *refreshing*³ that is used [2] and on the non-interference property [3] of the circuit's implementation, which ensures that the probabilistic distribution of the probed values does not depend on any secret [4]. Subsequent research has shown that a stronger type of non-interference property, which dictates that such probabilistic distribution might vary only with the number of internal probes, could be useful [3].

Since its introduction, the cryptographic research community has worked toward putting non-interference to work in proof mechanization and optimization [3–5]. Later developments concerned the realistic application to circuits that might leak more than one share per probe [6] and proposed some new categorization (e.g., pseudo- t -NI/ t -SNI [7]) to be able to reason about them. Among these, one of the most stud-

---

✉ Vittorio Zaccaria
vittorio.zaccaria@polimi.it

Maria Chiara Molteni
maria.molteni@unimi.it

¹ Dipartimento di Informatica “Giovanni Degli Antoni”,
Università degli Studi di Milano, Via Celoria 18, 20133
Milano, Italy

² Department of Electronics, Information and Bioengineering,
Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133
Milano, Italy

³ Present Address: Security Pattern srl, Brescia, Italy

¹ With *probing*, we mean a useful technique through which one can recover information from a subset of nodes of a circuit, e.g., measuring power consumption or EM emissions.

² These are secure in the sense that given t probes, it is impossible to derive information about the secret values encoded in the masks/shares.

³ The term *refresh* indicates a procedure that aims to bring back the secret's shares into a uniformly random state, after a series of operations that might have invalidated uniformity.

ied is the nonlinear part of the Advanced Encryption Standard (AES) [8,9], which is also one of the target applications of the current work.

1.1 Our contribution

Reasoning about non-interference still needs either difficult ratiocination or complex automatic tools [3,4]. In this work, we investigate an alternative formalization which, we argue, is simpler to reason with; in fact, it allowed us to prove some new properties of probing security (see Theorem 4). Our approach is based on the spectral theory of Boolean functions and correlation matrices [10,11], and represents, to some extent, an extension of the work presented in [12]. However, while the latter only addressed single output Boolean functions, we are the first to formalize multiple dependencies between outputs and inputs of a vector function, with the added benefit of being supported by matrix-based toolboxes. Note that the use of these tools in our area is not new; for example, in [13] the authors exploit the correlation matrix of shared functions to estimate some security bounds and define correlation between probed values, while other works use linear algebra to investigate resiliency against side-channel attacks [14] [15]. These works, however, address either monivariate attacks or other sharing schemes (inner product masking) while we aim at providing a larger and stronger foundational contribution allowing to concisely explain known general compositional patterns and potentially discover new ones.

This work is organized as follows: we present our new mathematical approach in Sect. 2 by introducing the concept of *shares' relation matrix*, which is a compact matrix representation of the (cor-)relation between a Boolean function's inputs and outputs. Methodologically, this new formalism allows to create a linear algebra view of t -probing security which we present in Sect. 3. This approach brings several benefits among which the ability to prove general probing security composition theorems through linear algebra. To this end, in Sect. 4 we put to work our formalism by proving the t -SNI property of the AES nonlinear part. We conclude by sketching the possible evolution of this work (Sect. 5) and by providing relevant mathematical background and proofs supporting our argument (Appendix).

2 A relation calculus for shares

Let us consider a circuit implementing a function:

$$f(x_1, \dots, x_n) : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$$

where the values x_i are *sensitive* (i.e., they have been computed using a secret). A side-channel attack consists of

measuring the power consumption of internal nodes of the circuit (through *probes*) and by searching through a set of guesses of the secret for the one that maximizes the correlation.

To design a mitigation against a side-channel attack, designers split each sensitive value x_i into d values $\alpha_i = \{\alpha_{i,j}\}_{j \in 1 \dots d}$ such that $\sum_j \alpha_{i,j} = x_i$; these d values are called *shares*. In principle, this is done by using $d - 1$ auxiliary, uniformly distributed random values (aka *masks*) and, unless one obtains all d shares $\alpha_{i,j}$, the correlation of each share with the sensitive value x_i is negligible [1]. The implementation of f must be changed so as to provide the result as a set of shares much like the original sensitive values. The computation of each output f_i is thus split into a set of d vector functions $\omega_i = \{\omega_{i,j}\}_{j \in 1 \dots d}$ such that:

$$f_i(x_1, \dots, x_n) = \sum_j \omega_{i,j}(A_1, \dots, A_n),$$

$$A_i \subseteq \{\alpha_{i,1}, \dots, \alpha_{i,d}\}$$

where each $\omega_{i,j}$ is called the j -th *output share* of f_i and it must be impossible to obtain information about f_i unless one obtains all d output shares.

In the probing security attack model, aside from regular output shares ω_i , attackers can observe (through probes) a group of the internal values of the circuit as additional outputs:

$$\Pi = \{\pi_1, \dots, \pi_{|\Pi|}\}$$

where each π_i is a function of the input shares. A mitigation against a probing attack ensures that none of the π_i are correlated with the original sensitive values. To design such countermeasures, besides the shares of the original sensitive values, designers use an additional group of inputs $P = \{\rho_1 \dots \rho_{|P|}\}$ which are uniformly random. These values are used to “refresh” the internally computed values of the function so as to make each π and ω not correlated with the sensitive values.

It is clear that correlation between each ω and π with any α and ρ is critical to determine whether the circuit is probing secure. A possible way to encode this information is to have a multi-dimensional matrix⁴ called the *shares' relation matrix*:

Definition 1 (*Shares' relation matrix*) Given a Boolean function $f : \mathbb{F}_2^{|A|+|P|} \rightarrow \mathbb{F}_2^{|Q|+|\Pi|}$, where A is the set of the function's input shares α_k , Q is the set of output shares ω_k , we define the *shares' relation matrix* of f as a multi-dimensional

⁴ Perhaps the most appropriate name for this type of object would be *tensor* but this name implies also some additional properties that are not used in this paper.

matrix F where each element:

$$F_{i_\pi i_{\omega|\Omega}| \dots i_{\omega_1}}^{j_\rho j_{\alpha|A}| \dots j_{\alpha_1}} \in \{0, 1\}, \quad (1)$$

is indexed by:

- $j_{\alpha_k} \in \{0, \dots, d\}, k \in \{1, \dots, |A|\}$
- $j_{\rho} \in \{0, \dots, |P|\}$
- $i_{\omega_p} \in \{0, \dots, d\}, p \in \{1, \dots, |\Omega|\}$
- $i_{\pi} \in \{0, \dots, |\Pi|\}$

and it is equal to 1 only if there exist a nonzero correlation between j_{α_k} shares of α_k and j_ρ randoms with i_{ω_p} output shares of ω_p and i_π probes, for all k, p .

A formal definition of such type of matrices is presented in Appendix B where, in particular, we consider multiple random P_l and probe Π_z groups instead of a single P and Π as above; however, for the rest of the paper, it is only necessary to get an intuitive understanding of it which we will develop in the following paragraphs. A practical way to compute a shares' relation matrix for a function f is deriving it from the Walsh matrix of f . Indeed, correlation matrices are useful to determine whether a set of output shares is *vulnerable*, i.e., correlated with one or more sensitive variables [16–18]. In particular, for a circuit f , any combination of outputs (encoded with the spectral coordinate ϕ) is correlated with a set of inputs (encoded with the spectral coordinate ψ) if $W_f(\phi, \psi) \neq 0$. It is possible to see a correlation matrix as an incidence matrix which encodes a *dependency relation* between the inputs and outputs of f . Such relation matrices (which are built over a Boolean semiring $K = \{(0, 1), \vee, \wedge\}$) are the fundamental building block for the *calculus of relations*⁵, an algorithmic device that allows the substitution of computation for a sometimes difficult ratiocination [19]. We can derive a relation matrix $\widehat{W}_f$ from the correlation matrix W_f of a vectorial Boolean function element-wise:⁶

$$\widetilde{W}_f(i, j) := (W_f(i, j) \neq 0). \quad (2)$$

Once a relation matrix $\tilde{W}_f$ has been found, one can derive easily the shares' relation matrix in Eq. 1 by inspection. Thus, we derive the shares' relation matrix in a two-step process

⁵ A relation matrix element $R_{i,j} \in K$ represents the absence (0) or presence (1) of a relationship iRj between entities encoded through row index i and column index j and the logic composition of relations can be encoded into a linear algebra expression and analyzed with conventional tools. In particular, logical disjunction is represented as matrix sum $((R+S)_{i,j} = R_{i,j} \vee S_{i,j})$, logical conjunction as the Hadamard product $((R \circ S)_{i,j} = R_{i,j} \wedge S_{i,j})$, and “Relative product” as conventional matrix multiplication $((RS)_{i,j} = \exists k R_{i,k} \wedge S_{k,j})$.

⁶ In this paper, we consider all indices starting from 0.

by starting from the correlation matrix of f :

$$W_f \rightarrow \tilde{W}_f \rightarrow F$$

Example 1 (Taken from [20] and reported here for simplicity)

Consider a function $f : \mathbb{F}_2^4 \rightarrow \mathbb{F}_2^3$

$$f(a_0, a_1, r_0, r_1) = \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_0 + r_0 + r_1 \\ a_1 + r_0 + r_1 \\ a_1 + r_0 \end{bmatrix}$$

such that

- a_0 and a_1 are two shares of a single sensitive input a ,
- r_0 and r_1 are two random values,
- f_0 and f_1 are two shares of a single output o , and
- f_2 is the value associated with a potential internal probe p within the circuit realization of f .

From its correlation matrix, we can derive through Eq. 2 the following relation matrix $\tilde{W}_f(\phi, \psi)$ ($\phi = [\gamma_{f_2} \gamma_{f_1} \gamma_{f_0}]$, $\psi = [\gamma_{r_1} \gamma_{r_0} \gamma_{a_1} \gamma_{a_0}]$):

$$\begin{array}{cccccccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \gamma_{r_1} \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \gamma_{r_0} \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & \gamma_{a_1} \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & \gamma_{a_0} \\
\gamma_{f_2} & \gamma_{f_1} & \gamma_{f_0} & & & & & & & & & & & & & & & \\
0 & 0 & 0 & 1 & & & & & & & & & & & & & & \\
0 & 0 & 1 & & & & & & & & & & & 1 & & & & \\
0 & 1 & 0 & & & & & & & & & & & & 1 & & & \\
0 & 1 & 1 & & & 1 & & & & & & & & & & & & \\
1 & 0 & 0 & & & & & 1 & & & & & & & & & & \\
1 & 0 & 1 & & & & & & & & & & & 1 & & & & \\
1 & 1 & 0 & & & & & & & & & & 1 & & & & & \\
1 & 1 & 1 & & & & & & & 1 & & & & & & & & \\
\end{array} \quad (3)$$

Note that we have labeled columns (rows) with the corresponding combination of inputs (outputs) in binary form. For example, element $\tilde{W}_f([011], [0011])$ (which is 1) represents an existing dependency between $f_0 \oplus f_1$ and $a_0 \oplus a_1$; note that in this specific case $W_f = \tilde{W}_f = 1$ but in general $\tilde{W}_f(i, j)$ is 1 whenever $W_f(i, j)$ is different from zero.

From the original correlation matrix $\tilde{W}_f$ (Eq. 3), it is thus possible to derive the corresponding *shares' relation matrix* F which accounts only for the amount of shares of the output and probes whose combination is correlated with a specified amount of shares of the input and of randoms:

$$\begin{array}{cccccccc}
0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & j_r \\
0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & j_a \\
i_p & i_o & & & & & & & & \\
0 & 0 & 1 & & & & & & & \\
0 & 1 & & & & & 1 & & & \\
0 & 2 & & 1 & & & & & & \\
1 & 0 & & & & & 1 & & & \\
1 & 1 & & & 1 & & 1 & & & \\
1 & 2 & & & & & 1 & & &
\end{array} \quad (4)$$

Note that the coordinates of this new relation matrix are computed by the Hamming weights of the spectral coordinates of $\tilde{W}_f$ split by their type (r the randoms, a the inputs, o the outputs, p the probes). This allows us to index in an alternative way any element $F(i, j)$:

$$F_{i_p i_o}^{j_r j_a}$$

where i_p, i_o, j_r, j_a are exactly the mixed-radix representation of i, j and carry additional information, i.e., the distinction between the related input-random (output-probe) composition. The mixed-radix representation $mr_{\rho}(n)$ of a number n over the vector of parts $\rho = [\rho_N, \dots, \rho_1]$ is a vector $b = [b_{N+1}, \dots, b_1]$ such that:

$$n = \sum_{i=1}^{N+1} b_i \prod_{j=1}^{i-1} (\rho_j + 1) \text{ where } 0 \leq b_i < \rho_i$$

The vector $[i_p, i_o]$ (resp., $[j_r, j_a]$) is just the mixed-radix representation of i (resp. j) over the vector of parts $[f_p, f_o]$ (resp. $[f_r, f_a]$) where f_a is the number of shares of the input of function f , f_r is the number of refresh values, f_o is the number of shares for function f 's outputs and f_p is the number of probes:

$$i = i_p \cdot (f_o + 1) + i_o \quad (5)$$

$$j = j_r \cdot (f_a + 1) + j_a \quad (6)$$

Example 2 The fifth row with index $i = 4$ of the matrix in Eq. (4) has a mixed-radix representation $[i_p, i_o] = [1, 1]$ over the vector of parts $[f_p, f_o] = [1, 2]$ because:

$$i = 1 \cdot (2 + 1) + 1$$

Same reasoning goes for column index 3 which corresponds to the representation $[1, 0]$ over the vector of parts $[f_r, f_a] = [2, 2]$, i.e.,

$$j = 1 \cdot (2 + 1) + 0$$

Thus, we have that $F_{1,1}^{1,0}$ is the element (4, 3) of matrix F and its indexes carries the fact that it corresponds the correlation of both probes and outputs ($[i_p, i_o] = [1, 1]$) with one of the random values ($[j_r, j_a] = [1, 0]$).

This work concerns itself with deriving security properties associated with the composition of functions. In the following we consider the function $h(x) = g(f(x))$ as a horizontal composition of g with f while the vector function:

$$v(x_1, x_2) = [f(x_1); g(x_2)]$$

as the vertical composition of f and g . We will show that the shares' relation matrix of a function distributes over vertical composition while, concerning horizontal composition, we can assert a weaker rule (see below) which will be still valid for inferring probing security. With regard to the proofs of following theorems, the reader is invited to refer to Appendix B.1.

Note that the definition of the shares' relation matrix is different from the Probe Distribution Table (PDT) introduced in [21] because the latter does not account for the potential compression of information that is obtained by encoding the Hamming weights of the spectral coordinates. With respect to [21], we show that it is possible to work with such minimal objects without resorting to encoding explicitly all possible input/output relationships. Note also that the goal of our work is more related to explaining how the composition of primitive gadgets works rather than determining inner properties for such primitives through their correlation matrices. However, other work [20] provides some deduction on the complexity required for deriving from scratch the above shares' relation matrix matrices.

Theorem 1 (Identity) *Given $id : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ the identity function, its shares' relation matrix is $\mathbf{I}_{n+1}$, where $\mathbf{I}$ is the identity matrix (see Appendix for the proof).*

The horizontal compositionality of the shares' relation matrices is determined by a weaker rule with respect to the conventional correlation matrix (see Theorem 7); in particular, as long as we look at the constituent parts of a horizontal composition of shares' relation matrices, their product will be always *conservatively more* than the original shares' relation matrix, as stated in the following theorem:

Theorem 2 (Shares' relation matrices pseudo-horizontal composition) *Given two functions f and g , and F, G, FG the shares' relation matrices of f, g and $g \circ f$, respectively, the following dominance holds:*

$$(FG)_{I_{\pi} I_{\omega}}^{J_{\rho} J_{\alpha}} \leq F_{I_{\pi} I_{\omega}}^{K L} G_{K L}^{J_{\rho} J_{\alpha}}$$

(see Appendix for the proof).

Practically speaking, if the product of two shares' relation matrices does not imply a dependency between variables, this will be absent from the whole shares' relation matrix as well. Vertical composition, however, still holds as the following theorems show:

Theorem 3 (Shares' relation matrices pseudo-distributivity over tensor product) *Given two functions f and g , and $F, G, F|G$ the shares' relation matrices of f, g and the vertical juxtaposition of g above f , respectively, the following holds:*

$$(F|G)_{(I_{\pi_f}|I_{\pi_g})(I_{\omega_f}|I_{\omega_g})}^{(J_{\rho_f}|J_{\rho_g})(J_{\alpha_f}|J_{\alpha_g})} = F_{I_{\pi_f}I_{\omega_f}}^{J_{\rho_f}J_{\alpha_f}} \otimes G_{I_{\pi_g}I_{\omega_g}}^{J_{\rho_g}J_{\alpha_g}}$$

where $\otimes$ is the Kronecker (or tensor) product, see Appendix for the proof.

Corollary 1 (Tensor product with identity) *From the previous theorems, it follows that the following equalities hold:*

$$\begin{aligned} (\mathbf{I}_{n+1}|G)_{(I_{\pi_d}|I_{\pi_g})(I_{\omega_d}|I_{\omega_g})}^{(J_{\rho_d}|J_{\rho_g})(J_{\alpha_d}|J_{\alpha_g})} &= \delta(J_{\rho_d}J_{\alpha_d}, I_{\pi_d}J_{\omega_d}) \cdot G_{I_{\pi_g}I_{\omega_g}}^{J_{\rho_g}J_{\alpha_g}} \\ (G|\mathbf{I}_{n+1})_{(I_{\pi_f}|I_{\pi_d})(I_{\omega_f}|I_{\omega_d})}^{(J_{\rho_f}|J_{\rho_d})(J_{\alpha_f}|J_{\alpha_d})} &= G_{I_{\pi_f}I_{\omega_f}}^{J_{\rho_f}J_{\alpha_f}} \cdot \delta(J_{\rho_d}J_{\alpha_d}, I_{\pi_d}J_{\omega_d}) \end{aligned} \quad (7)$$

where δ is the Kronecker's delta.

3 Application to t -probing security

In this section, we revisit and enhance known theorems about t -probing security by showing how they naturally descend from the relation calculus of shares based on shares' relation matrices. We recall that t -probing security centers around the concept of t -non-interfering function. A function f is t -NI if, when given a total of s outputs and internal probes, $s \leq t$ implies a dependency with maximum s input shares. A function f is t -SNI if $s \leq t$ implies a dependency with maximum i input shares, where i is the number of internal probes.

Much has been said about the composition rules of such functions and, unfortunately, their proofs are complex, long or require much expertise in type theoretical or formal validation area [3]; we will show that the relation calculus of shares allows to revisit and extend these proofs with conventional linear algebra tools, broadening the potential audience.

To talk about t -probing security, we've found useful to follow this general pattern: (i) we explicitly include random refresh values as inputs⁷ and (ii) we include in the signature of the function also the probes considered. This creates a natural

subdivision of the shares' relation matrix for the considered function. Before introducing some general results that can be derived with our formalism, however, we introduce an additional example that shows how one could identify a violation of compositionality in an existing gadget with our formalism.

Example 3 (Extended from [20]). In this example, we revisit through our formalism a case discovered in [2] that proves that, in general, the composition of t -NI and t -SNI functions is not t -NI.

Figure 1 shows the structure of a function $h(a)$ which is a composition of two functions f and g ; the assumptions are that f is t -NI and g is t -SNI. In particular, f refreshes its input a with two random bits r_f :

$$o_f(a_0, a_1, a_2, r_0, r_1) = [a_0 \oplus r_0 \oplus r_1, a_1 \oplus r_0, a_2 \oplus r_1]$$

and it is assumed to have been probed at location $p_f = a_0 \oplus r_0$. On the other hand, $g(a, b, r_g)$ is the ISW multiplication [1] which consumes 3 random bits r_g for the secret computation. Also in this case, it is assumed a single probe $p_g = a_2 \wedge b_1$. We will show that our method provides a sufficient precision to individuate the vulnerability spotted in [2]. To fit into our formalism however, we must consider the underlying correlation matrices that include explicitly *i*) the random values both f and g consume to refresh the data and *ii*) the probes that are present. The string diagram in Fig. 5 describes the composition pattern of correlation matrices as a mapping from the space of the Fourier transform of the input distribution $\mathbb{A} \otimes \mathbb{R}_f \otimes \mathbb{R}_g$ (i.e., the actual inputs plus the random values) to the one of the output distribution $\mathbb{O}_g \otimes \mathbb{P}_g \otimes \mathbb{P}_f$ (i.e., the actual output of g and the probes in both f and g). Still considering the string diagram of Fig. 5, one can derive one of the equivalent expressions of the correlation matrix of h as:

$$W_h = (\mathbf{I}_2 \otimes W_g)(W_q \otimes \mathbf{I}_{2^3} \otimes \mathbf{I}_{2^3})(\mathbf{I}_{2^3} \otimes W_f \otimes \mathbf{I}_{2^3})(\mathbf{I}_{2^3} \otimes \mathbf{I}_{2^2} \otimes W_s)$$

where W_s is the correlation matrix of the duplication function $s = (x) \mapsto (x, x)$ and W_q is the correlation matrix of

Fig. 1 The composition pattern of f (t -NI) and g (t -SNI) studied in Example 3 and derived from [2]. The composed function $h(a)$ is not t -NI as can be easily checked with our formalism

⁷ After all, these are values generated independently by a separated random number generator so it makes sense to include them in the signature of the function itself.

function $q = (x, y) \mapsto (y, x)$. We are interested in computing the potential dependencies between any combination of output/probes and inputs that are not masked by random values. Thus, computing the shares' relation matrices from all the previous correlation matrices, by Theorems 2 and 3, the following holds:

$$H \preceq (\mathbf{I}_2 \otimes G)(Q \otimes \mathbf{I}_4 \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes F \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \mathbf{I}_3 \otimes S) \quad (8)$$

where H, F, G, S and Q are the shares' relation matrices computed for functions h, f, g, s and q , respectively.

The value of the right-hand side of Eq. 8 is shown in Fig. 2. First of all, we are interested only in the first 4 columns, as these are the ones that represent relationships between the outputs and the shares of a not masked by any random value. We note that there is a potential dependency in row $[1, 1, 0]$, column $[0, 0, 3]$, exactly the one found in [2], which says that one needs only two probe values to get three shares; h is thus not even 2-NI, showing that t -NI and t -SNI do not compose into a t -NI function. This example shows that the proposed calculus of shares has sufficient precision to discover these cases. On one hand, these could be false positives because of the dominance relation in Eq. (8); on the other hand, however, this formalism rules out any false negative. We will show that the stronger concept of t -SNI naturally emerges, in our relation calculus, as a fundamental property to ensure compositionality.

Fig. 2 The shares' relation matrix of function h in Example 3 derived from [2] (we use Greek letters to indicate the spectral coordinate associated with each function variable, i.e., α is the spectral coordinate associated with variable a and so on). Gray areas indicate where h is allowed to have nonzero values in its shares' relation matrix to meet t -NI hypotheses. One can see that in row $[1, 1, 0]$, column $[0, 0, 3]$ there is a potential relation between two probes and the three shares of a , meaning that the composition is not even 2-NI

3.1 Proving general patterns of compositional security

The shares' relation matrix can be a reasonable way for exploring t -probing security, but there is more. In fact, it is possible to demonstrate that in order to rule out dependencies similar to Example 3, both f and g must be t -SNI. In this section, we will revisit some known composition patterns (e.g., Theorem 5 and Corollaries 2 and 3 appeared in [4]) and introduce a new one not known in literature (Theorem 4).

Here, we restate what it means for a function f to be t -NI/ t -SNI in terms of the shares' relation matrix F :

Definition 2 f is t -SNI iff, for any set of probes that could be introduced in it, the following predicate is true for any element (i, j) of its shares' relation matrix:

$$|i_\pi| + |i_\omega| \leq t \wedge (\exists a. j_{\alpha_a} > |i_\pi|) \implies \neg F_{i_\pi | \pi | \dots i_{\pi_1} | \pi_1 | \dots i_{\omega_1} | \omega_1 |}^{0 \dots 0 j_{\alpha_{|A|}} \dots j_{\alpha_1}}$$

Definition 3 f is t -NI iff, for any set of probes that could be introduced in it, the following predicate is true for any element (i, j) of its shares' relation matrix:

$$|i_\pi| + |i_\omega| \leq t \wedge (\exists a. j_{\alpha_a} > |i_\pi| + |i_\omega|) \implies \neg F_{i_\pi | \pi | \dots i_{\pi_1} | \pi_1 | \dots i_{\omega_1} | \omega_1 |}^{0 \dots 0 j_{\alpha_{|A|}} \dots j_{\alpha_1}}$$

where it is evident that t -NI corresponds to a weaker version of t -SNI.

Example 4 The Coron's linear-space variant [22] of the ISW multiplication [1] is t -SNI [3] and this can be easily seen through the shares' relation matrix. Let us consider its form for $t = 1$; in this case we have two shares for two inputs a and b , one random value r , two output shares o and six possible internal probes p :

$$\text{SecMult}(a_0, a_1, b_0, b_1, r) = [o_0, o_1, p_0, p_1, p_2, p_3, p_4, p_5]$$

where

$$\begin{bmatrix} o_0 \\ o_1 \\ p_0 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix} = \begin{bmatrix} a_0 b_0 + r \\ a_1 b_1 + ((a_0 b_1 + r) + a_1 b_0) \\ a_0 b_0 \\ a_1 b_1 \\ a_0 b_1 \\ a_1 b_0 \\ a_0 b_1 + r \\ (a_0 b_1 + r) + a_1 b_0 \end{bmatrix} \quad (10)$$

$$\begin{array}{cccccccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \dots \rho \\
0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & \dots \beta \\
0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & \dots \alpha
\end{array}$$

$$\begin{array}{cc}
\pi & \omega \\
0 & 0 \\
0 & 1 \\
\dots & \dots \\
1 & 0
\end{array}
\begin{array}{cc}
1 & 1 \\
1 & 1 \\
\dots & \dots
\end{array}
\begin{array}{cccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}
\quad (11)$$

Fig. 3 Part of the shares' relation matrix of SecMult function [22] (only interesting rows for t -SNI are shown). Note that α , β and ρ are the spectral coordinates associated with inputs a , b and r , while ω and π are the spectral coordinates for o and p . Gray areas indicate where SecMult is allowed to have nonzero values in its shares' relation matrix to meet t -SNI hypotheses

Part of the corresponding shares' relation matrix is shown in Fig. 3; it can be seen that for $\pi + \omega \leq 1$, $\rho = 0$ and $\alpha, \beta > 1$ (white areas) we have a null dependency, i.e., the function is 1-SNI. Note that, for this ISW implementation, the number of outputs and probes varies with t with the following law:

$$(t+1)^2 + 4 \cdot \sum_{i=1}^t i$$

where $t+1$ of these correspond to outputs while the others are internal probes.

The simplest composition pattern for which we can derive general rules is $l = g \circ f$. The corresponding map between the Fourier transforms of distributions is shown in Fig. 4. The question we address is if l (with the associated shares' relation matrix L) is t -SNI/ t -NI according to Definitions 2 and 3, by making assumptions on the probing security of the underlying functions f and g (whose shares' relation matrices are called F and G , respectively). Note that, to fit

Fig. 4 Map between Fourier transforms of probability distributions implied by a function composition $l = g \circ f$

within our formalism, we need to explicitly route the refresh values for g and probed value of f with a function q that just swaps those values. Note that, since matrix Q is the shares' relation matrix of $q : (x, y) \mapsto (y, x)$ function, it can be shown that the following holds:

$$Q_{i_{\rho_g}, i_{\pi_f}}^{j_{\pi_f}, j_{\rho_g}} = \delta(i_{\pi_f}, j_{\pi_f}) \cdot \delta(i_{\rho_g}, j_{\rho_g}).$$

Besides, by Theorem 2, we know that L is dominated by the product:

$$ABC = (I_{n_{\pi_f}} \otimes G) \cdot (Q \otimes I_{n_{\omega_f}}) \cdot (I_{n_{\rho_g}} \otimes F)$$

where $n_{\pi_f}(n_{\omega_f}, n_{\rho_g})$ is the number of probes in f (output's shares of f , randoms needed to refresh g) plus 1 (see Theorem 1).

The following lemma can be proved

Lemma 1 The product ABC is such that:

$$(ABC)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_\alpha} = \sum_r G_{i_{\pi_g} i_{\omega_g}}^{0,r} F_{i_{\pi_f},r}^{0,j_\alpha}$$

For a proof, see Appendix.

We are now able to derive formally whether and when l is t -SNI/ t -NI.

Theorem 4 If f is t -SNI and g is t -NI, then $l(x) = g(f(x))$ is t -SNI. Formally, the following three axioms:

- A1** $r + |i_{\pi_f}| \leq t \wedge v > |i_{\pi_f}| \implies \neg F_{i_{\pi_f},r}^{0,v}$
A2 $|i_{\omega_g}| + |i_{\pi_g}| \leq t \wedge r > |i_{\pi_g}| + |i_{\omega_g}| \implies \neg G_{i_{\pi_g},i_{\omega_g}}^{0,r}$
A3 $(|i_{\pi_g}| + |i_{\pi_f}| + |i_{\omega_g}| \leq t) \wedge (j_\alpha > |i_{\pi_g}| + |i_{\pi_f}|)$

$$\text{entail } (ABC)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_\alpha} = 0$$

Proof Exploiting above axioms and Lemma 1 we can derive that:

$$\begin{aligned}
(ABC)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_\alpha} &\stackrel{\text{Lem. 1}}{=} \sum_r G_{i_{\pi_g} i_{\omega_g}}^{0,r} F_{i_{\pi_f},r}^{0,j_\alpha} \\
&\stackrel{\text{A1,A2}}{\leq} \sum_{t-i_{\pi_f} < r \leq i_{\pi_g} + i_{\omega_g}} G_{i_{\pi_g} i_{\omega_g}}^{0,r} \stackrel{\text{A3}}{\leq} 0
\end{aligned}$$

□

Corollary 2 If f and g are t -SNI functions then also $l(x) = g(f(x))$ is t -SNI.

Proof Assuming g is t -SNI, then it is also t -NI and the thesis follows from Theorem 4. □

We already saw an example of another composition pattern studied in the literature, whose circuit diagram is shown in Fig. 1. The diagram associated with its correlation matrices is the one shown in Fig. 5. With our formalism, it is possible to identify some general rules to determine if such a composed function is t -NI/ t -SNI (according to Definitions 2 and 3) by making assumptions on the probing security of the underlying functions f and g . Note that, to reconcile with our model of function, we explicitly split the whole function l into a composition $a \circ b \circ c \circ d$. In particular, d contains the duplication function s that sends a copy of the shared input to both f and g , while b contains q as in the pattern that we previously studied. The shares relation matrix S associated with $s : x \mapsto (x, x)$ function is characterized by the following lemma:

Lemma 2 For any $i_{\alpha_1}, i_{\alpha_2}, j_{\alpha}$ indices, the following holds:

$$|i_{\alpha_1}| + |i_{\alpha_2}| < |j_{\alpha}| \implies S_{i_{\alpha_1}, i_{\alpha_2}}^{j_{\alpha}} = 0$$

For a proof, see Appendix.

From the point of view of the shares' relation matrix involved, we know that whole function is dominated by the product (see Theorem 2):

$$ABCD = (\mathbf{I}_{n_{\pi_f}} \otimes G) \cdot (Q \otimes \mathbf{I}_{n_{\omega_f}} \otimes \mathbf{I}_{n_{\alpha_1}}) \cdot (\mathbf{I}_{n_{\rho_g}} \otimes F \otimes \mathbf{I}_{n_{\alpha_1}}) \cdot (\mathbf{I}_{n_{\rho_g}} \otimes \mathbf{I}_{n_{\rho_f}} \otimes S)$$

where n_{α_1} (n_{ρ_f}) is the number of shares of the first g 's input (randoms needed to refresh f) plus 1.

Fig. 5 Map between Fourier transforms of probability distributions implied by the second composition pattern studied in this paper

Lemma 3 The complete relation matrix $ABCD$ computed in Figure 5 is such that:

$$(ABCD)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_{\alpha}} = \sum_{v,z} \left(\sum_r G_{i_{\pi_g} i_{\omega_g}}^{0,r,z} F_{i_{\pi_f},r}^{0,v} \right) S_{v,z}^{j_{\alpha}}$$

For a proof, see Appendix.

We are now able to derive formally when l is t -SNI/ t -NI.

Theorem 5 If f is t -SNI function and g is t -NI, then $l(x) = g(f(x), x)$ is t -NI. Formally, the following three axioms:

$$\mathbf{A4} \quad r + |i_{\pi_f}| \leq t \wedge v > |i_{\pi_f}| \implies \neg F_{i_{\pi_f},r}^{0,v}$$

$$\mathbf{A5} \quad |i_{\omega_g}| + |i_{\pi_g}| \leq t \wedge (r > |i_{\pi_g}| + |i_{\omega_g}| \vee z > |i_{\pi_g}| + |i_{\omega_g}|) \implies \neg G_{i_{\pi_g},i_{\omega_g}}^{0,r,z}$$

$$\mathbf{A6} \quad (|i_{\pi_g}| + |i_{\pi_f}| + |i_{\omega_g}| \leq t) \wedge (|j_{\alpha}| > |i_{\pi_g}| + |i_{\pi_f}| + |i_{\omega_g}|)$$

$$\text{entail } (ABCD)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_{\alpha}} = 0$$

Proof Exploiting above axioms and Lemmas 2 and 3:

$$(ABCD)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_{\alpha}} \stackrel{\text{Lem. 3}}{=} \sum_{v,z} \left(\sum_r G_{i_{\pi_g} i_{\omega_g}}^{0,r,z} F_{i_{\pi_f},r}^{0,v} \right) S_{v,z}^{j_{\alpha}} \quad (12)$$

$$\stackrel{\mathbf{A4}}{\leq} \sum_{v \leq i_{\pi_f}, z} \sum_{r > t - i_{\pi_f}} G_{i_{\pi_g} i_{\omega_g}}^{0,r,z} S_{v,z}^{j_{\alpha}} \quad (13)$$

$$\stackrel{\mathbf{A5}, \mathbf{A6}}{\leq} \sum_{v \leq i_{\pi_f}, z \leq i_{\pi_g} + i_{\omega_g}} S_{v,z}^{j_{\alpha}} \stackrel{\text{Lem. 2}, \mathbf{A6}}{\leq} 0 \quad (14)$$

□

Remark 1 Note that the case handled in Theorem 5 concerns f t -SNI and g t -NI; vice versa, Example 3 concerns the inverted case f t -NI and g t -SNI.

Corollary 3 If f and g are t -SNI functions then also $l(x) = g(f(x), x)$ is t -SNI. Formally, the following three axioms:

$$\mathbf{A7} \quad r + |i_{\pi_f}| \leq t \wedge v > |i_{\pi_f}| \implies \neg F_{i_{\pi_f},r}^{0,v}$$

$$\mathbf{A8} \quad |i_{\omega_g}| + |i_{\pi_g}| \leq t \wedge (r > |i_{\pi_g}| \vee z > |i_{\pi_g}|) \implies \neg G_{i_{\pi_g},i_{\omega_g}}^{0,r,z}$$

$$\mathbf{A9} \quad (|i_{\pi_g}| + |i_{\pi_f}| + |i_{\omega_g}| \leq t) \wedge (|j_{\alpha}| > |i_{\pi_g}| + |i_{\pi_f}|)$$

$$\text{entail } (ABCD)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_{\alpha}} = 0$$

Proof The initial part of the proof is the same of Theorem 5 up to Equation (13); then the different axioms apply:

$$\sum_{v \leq i_{\pi_f}, z, r > t - i_{\pi_f}} G_{i_{\pi_g} i_{\omega_g}}^{0, r, z} S_{v, z}^{j_{\alpha} A8, A9} \sum_{v \leq i_{\pi_f}, z \leq i_{\pi_g}} S_{v, z}^{j_{\alpha} Lem. 2, A9} \leq 0$$

□

4 Extending the approach to $\mathbb{F}_{2^k}^n$: the AES inversion

In this section, we present an extension of the proposed formalism to address the case where shares encode values over k bits, i.e., they belong to $\mathbb{F}_{2^k}^n$. Let us thus consider a function $f : \mathbb{F}_{2^k}^n \rightarrow \mathbb{F}_{2^k}^m$; we can extend Eq. 2 as follows:

$$\widetilde{W}_f(i, j) = \exists u, v. W_f(u, v) \neq 0 \wedge (u^{(k, n) \triangleright n} = i) \wedge (v^{(k, m) \triangleright m} = j) \quad (15)$$

where $u^{(k, n) \triangleright n}$ is a reduction operation over the binary encoding of the spectral coordinate u (see Fig. 6). It can be shown that the shares' relation matrix for the relation matrices computed as in Eq. 15 still complies with Definitions 2 and 3 and Theorems 2 and 3. In this setting, affine functions have a nice representation that will be useful to extend the application of previous theorems.

Definition 4 A function $f : \mathbb{F}_{2^k}^n \rightarrow \mathbb{F}_{2^k}^m$ is a (multi-share) *affine function* if:

$$\forall x \in \mathbb{F}_{2^k}^n, \forall i \in \{0, \dots, n-1\} \exists g. f(x)_i = g(x_i)$$

where g is an affine function, x_i is the i -th share of x and $f(x)_i$ is the i -th share of $f(x)$ (see [4]). For conciseness, we will refer to f as an affine function as well.

The relation matrix of an affine function (as well as its shares' relation matrix) is an identity, as the following lemma shows.

Lemma 4 Let $f : \mathbb{F}_{2^k}^n \rightarrow \mathbb{F}_{2^k}^m$ be an affine function; then $\widetilde{W}_f = I_{2^n}$.

Fig. 6 Example of reduction operation $u^{(k, n) \triangleright i}$. The new spectral coordinate binary encoding $u^{(4, 2) \triangleright 2}$ is the result of OR'ing k -bit wide blocks of the original encoding u

Proof The affine function f can be seen as the parallel application of n functions g_i such that $f(x)_i = g_i(x_i)$ with $0 \leq i \leq n-1$; this implies that:

$$W_f = \bigotimes_{i=n-1}^0 W_{g_i}.$$

Since each g_i is an affine (and balanced) function, then $\widetilde{W}_{g_i} = I_2$ and:

$$\widetilde{W}_f = \bigotimes_{i=n-1}^0 I_2 = I_{2^n}.$$

□

When using our formalism to determine if a function over $\mathbb{F}_{2^k}^n$ is t -NI/ t -SNI, we can thus treat affine functions as identities because their shares' relation matrix is the same as the one of an identity function.

4.1 AES inversion function

A function that has been widely studied in the probing security framework is the inversion function in AES algorithm; finding a gadget that implements it in a probing secure way, also when it is composed with previous and following gadgets, is an important research cornerstone.

Let us consider the t -SNI gadget proposed in [9] as the AES inversion in $\mathbb{F}_{2^8}$. A formal demonstration for the strong security of this implementation has been introduced in [4]. Here we show how this could be proven with our formalism, exploiting only patterns that we have presented and proved in this work.

(a) The m_x block

(b) Inversion algorithm proposed in [9]

Fig. 7 An example application of the proposed formalism to functions over $\mathbb{F}_{2^k}^n$. Blocks m_4 and m_2 in (b) are structured as in (a)

We report the inversion gadget in Fig. 7b. Note that we have slightly modified the algorithm presented in [9] by moving two power computation blocks across duplication points; semantically it is always the same circuit but it is easier to see how previously introduced patterns can still be used to show that it is t -SNI.

First of all, we note that there is a recurring pattern in that particular algorithm, i.e., the circuit in Fig. 7a. The block is composed of a mask refresh Refresh (t -SNI), the ISW multiplication SecMult (t -SNI), and $\cdot^x$, an affine power function parameterized over the exponent x (which is a multiple of two). It is possible to demonstrate that m_x is t -SNI following the same line of reasoning of Theorem 3 because, by Lemma 4, the relation matrix of the power function can be interpreted as an identity, thus the same case as the one shown in Fig. 5 applies. Considering the overall algorithm in Fig. 7b, we observe that this is t -SNI if $b \circ m_2$ is t -SNI (by Theorem 3). By Corollary 2, $b \circ m_2$ is t -SNI if b is t -SNI and the latter is true by Theorem 3 and by Lemma 4.

5 Conclusion

We originally started this research to extend our understanding of t -probing security. We have discovered a new relation calculus of shares which exploits the conventional Walsh transform. This calculus is precise enough to prove and extend known compositional properties without much semi-formal or verbal ratiocination. We believe that the underlying linear algebra, while providing a more intuitive understanding, but will allow for an easier mechanization of probing security proofs.

We also believe that a similar approach can be used to address vulnerabilities associated with circuit glitches. In this sense, we have made a preliminary proposal that shows that the approach is viable [20]. Indeed, more work must still be done toward a unifying approach that encompasses circuit glitches and new composability definitions such as the t -PINI condition [23].

Funding Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence,

unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Properties of the Walsh transform

This section recaps the important properties of the Walsh transform of a vectorial Boolean function and introduces the concept of tensor product for the resulting matrices [24].

Definition 5 (Walsh transform of a vectorial function) Given a vectorial Boolean function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$, we define its Walsh transform as a $2^m \times 2^n$ matrix $\hat{f}$ whose elements are:

$$\hat{f}_{\omega, \alpha} = \sum_{x \in \mathbb{F}_2^n} (-1)^{\omega^\top f(x) \oplus \alpha^\top x} \quad (16)$$

$\omega \in \mathbb{F}_2^m$, $\alpha \in \mathbb{F}_2^n$ being the binary encoding of the row and column indices, called *spectral coordinates* (or sometimes *masks*).

These matrices encode the correlation information between input variables' XOR-combinations and the corresponding output ones. For this reason, they sometimes appear in the literature, scaled by a coefficient 2^{-n} , as *correlation matrices* [10]:

$$W_f = 2^{-n} \hat{f}$$

For correlation matrices, the following theorem holds:

Theorem 6 (Correlation matrix as a map of probability distributions) Given a function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ and a probability distribution $p_X : \mathbb{F}_2^n \rightarrow \mathbb{R}$ for its input variable, the following relation holds:

$$W_f T_{p_X} = T_{p_Y}$$

where p_Y is the distribution of the output values while T_g is the Fourier transform of any pseudo-Boolean function $g : \mathbb{F}_2^n \rightarrow \mathbb{R}$ and defined as the following:

$$T_g(\gamma) = \sum_{x \in \mathbb{F}_2^n} g(x) (-1)^{\gamma^\top x}.$$

For a proof, see [10, 25].

Interpreting the Fourier transform of a probability distribution of a variable in $\mathbb{F}_2^n$ as a vector in a subset⁸ $\mathbb{P}_n$ of

⁸ Note that these are not sub-spaces as the set $\mathbb{P}_n$ is not closed under addition.

$\mathbb{R}^{2^n}$, we find that the correlation matrix W_f of a function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ is just a linear map $\mathbb{P}_n \rightarrow \mathbb{P}_m$. These maps are endowed with composition:

Theorem 7 (Composition of correlation matrices) *Given two functions $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ and $g : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^q$, the following holds:*

$$W_{g \circ f} = W_g W_f.$$

Moreover, if f is a bijection, $W_{f^{-1}} = W_f^{-1}$. For a proof, see [24, 26].

Given two independent variables $x_f \in \mathbb{F}_2^{n_f}$ and $x_g \in \mathbb{F}_2^{n_g}$, one can form the probability distribution of the vector $[x_f, x_g]$ with the product of distributions. From the point of view of its Fourier transform, this is a mapping $\mathbb{P}_{n_f} \times \mathbb{P}_{n_g} \rightarrow \mathbb{P}_{n_f+n_g}$. The following theorem holds:

Theorem 8 (Tensor product of correlation matrices) *Given two functions $f : \mathbb{F}_2^{n_f} \rightarrow \mathbb{F}_2^{m_f}$ and $g : \mathbb{F}_2^{n_g} \rightarrow \mathbb{F}_2^{m_g}$, the correlation matrix of the function $h([x_f, x_g]) = [f(x_f), g(x_g)]$ is $W_h = W_g \otimes W_f$ where the symbol $\otimes$ is the Kronecker product (or tensor product) of matrices (proof in the appendix). It is customary to say that W_h is a mapping from the space $\mathbb{P}_{n_f} \otimes \mathbb{P}_{n_g}$ to the space $\mathbb{P}_{m_f} \otimes \mathbb{P}_{m_g}$.*

Theorem 8 is informally proven in [10] and it is applied in several works, as in [27]; taking this into account, in this appendix we try to give to it a formal proof. For this scope, we define the quotient and remainder operators as follows, to remind ourselves of the structure of the indices:

$$i_{\uparrow p} = \left\lfloor \frac{i}{p} \right\rfloor, i_{\downarrow p} = i - pi_{\uparrow p}.$$

When $p = 2^n$ and i is a number that can be encoded over $k > n$ bit, $i_{\uparrow p}$ corresponds to the value encoded by the upper $k - n$ bits, while $i_{\downarrow p}$ corresponds to the value encoded by the lower n bits.

Definition 6 (Kronecker product of matrices) The tensor product of two matrices. X (of $n \times m$ elements) and Y (of $p \times q$ elements) can be defined as:

$$(X \otimes Y)_{i,j} = X_{i_{\uparrow p}, j_{\uparrow q}} Y_{i_{\downarrow p}, j_{\downarrow q}}.$$

Proof (Theorem 8) Note that ω (α) can be treated as a decimal number or as the corresponding (vector) binary encoding; moreover, the encoding of ω (α) can in turn be decomposed into two parts $[\omega_g, \omega_f]$ ($[\alpha_g, \alpha_f]$) of m_g (n_g) and m_f (n_f) bits, respectively. We start by rewriting the definition of $\hat{h}$:

Fig. 8 Example of compositional equality derived through a string diagram. The diagram on the left corresponds to the product $(W_g \otimes W_f)$ while the one on the right corresponds to $(1 \otimes W_f)(W_g \otimes 1)$ (each factor is highlighted with a dotted box). The fact that the second can be derived simply by moving boxes without crossing wires implies (because we are in monoidal category) that the underlying formulas are equivalent, i.e., $(W_g \otimes W_f) = (1 \otimes W_f)(W_g \otimes 1)$.

$$\begin{aligned} \hat{h}(\omega, \alpha) &= \frac{1}{2^{n_f+n_g}} \sum_{x \in \mathbb{F}_2^{n_f+n_g}} (-1)^{\omega^\top h(x) + \alpha^\top x} \\ &= \frac{1}{2^{n_f+n_g}} \sum_{[x_f, x_g] \in \mathbb{F}_2^{n_f+n_g}} (-1)^{\omega_f^\top f(x_f) + \omega_g^\top g(x_g) + \alpha_f^\top x_f + \alpha_g^\top x_g} \\ &= \hat{g}(\omega_g, \alpha_g) \hat{f}(\omega_f, \alpha_f) \\ &= \hat{g}\left(\left\lfloor \frac{\omega}{2^{m_g}} \right\rfloor, \left\lfloor \frac{\alpha}{2^{n_g}} \right\rfloor\right) \hat{f}\left(\omega - 2^{m_g} \left\lfloor \frac{\omega}{2^{m_g}} \right\rfloor, \alpha - 2^{n_g} \left\lfloor \frac{\alpha}{2^{n_g}} \right\rfloor\right) \\ &= \hat{g}(\omega_{\uparrow 2^{m_g}}, \alpha_{\uparrow 2^{n_g}}) \hat{f}(\omega_{\downarrow 2^{m_g}}, \alpha_{\downarrow 2^{n_g}}) \end{aligned}$$

and conclude (using Definition 6) that the last equation represents the generic element in position (ω, α) of the Kronecker product $\hat{g} \otimes \hat{f}$. $\square$

Reasonings on the effect of composing correlation matrices can be intuitively allowed through diagrams. Each correlation matrix is drawn as a box (except for identities which are drawn as simple wires), composition is the horizontal juxtaposition while tensor product is the vertical one (see Fig. 8 for an example). We note that there is a remarkable correspondence between a diagram and the underlying circuit diagram to the point that we could talk about “the” diagram of the circuit. Moreover, there always exist two mappings $B_{a,b} : \mathbb{P}_a \otimes \mathbb{P}_b \rightarrow \mathbb{P}_b \otimes \mathbb{P}_a$ and $B_{b,a}$ such that $B_{a,b} B_{b,a} = I$. $B_{a,b}$ is exactly the Walsh transform of a function that permutes variables a and b (this is typically drawn with crossing wires, such as block Q in Fig. 5).

B Formal definition of shares’ relation matrix

To produce precise proofs of the theorems introduced in this appendix, we need to slightly modify the notation given in Eqs. 5 and 6, to show the actual vector of parts over which the multi-radix representation is computed. For the sake of

generality, we will consider a generic vector of parts $\underline{\Omega}$ ($\underline{A}$) for the matrix rows (columns). Given a function f , we thus talk about a shares' relation matrix in the following form:

$$H_{\underline{\Omega}, \underline{A}}[\widetilde{W}_f]$$

With the understanding that each element (i, j) of $H_{\underline{\Omega}, \underline{A}}[\widetilde{W}_f]$ is such that

$$F_{mr_{\underline{\Omega}}(i)}^{mr_{\underline{A}}(j)} = H_{\underline{\Omega}, \underline{A}}[\widetilde{W}_f]_{i,j} \quad (17)$$

We will also use $|i_\xi|$ to indicate $\sum_k i_{\xi_k}$, i.e., the sum of the mixed-radix components of index i associated with the vector of parts Ξ . With this notation, $|i_\pi|$ practically means the number of probes associated with a specific value of index i .

Example 5 Considering Example 2, we have the following notational equivalence

$$F_{1,1}^{1,0} = H_{[1,2],[2,2]}[\widetilde{W}_f]_{4,3}.$$

The shares' relation matrix can be seen as the encoding of a predicate over the original relation matrix; this fact will be used to prove the remaining theorems in this appendix and corresponds to an equivalent definition of the matrix itself, as the following theorem shows.

Alternative Definition 1 (*Shares' relation matrix*) The shares' relation matrix computed from a relation matrix $\widetilde{Q} \in K^{2^\Omega \times 2^A}$ is a matrix $H_{\underline{\Omega}, \underline{A}}[\widetilde{Q}] \in K^{\pi_{\underline{\Omega}} \times \pi_{\underline{A}}}$ where

$$\pi_{\underline{\Omega}} = \prod_i (\omega_i + 1), \pi_{\underline{A}} = \prod_i (\alpha_i + 1) \quad (18)$$

and such that each element $H[\widetilde{Q}]_{i,j}$ is 1 iff:

$$\exists \widetilde{Q}_{r,s}. \widetilde{Q}_{r,s} \wedge wt_{\underline{\Omega}}(r) = mr_{\underline{\Omega}}(i) \wedge wt_{\underline{A}}(s) = mr_{\underline{A}}(j) \quad (19)$$

where $wt_V : \mathbb{N} \rightarrow \mathbb{N}^V$ is the hamming weight of each of the v binary parts according to the vector of parts V .

Remark 2 (Compact definition) We will sometimes use the notation $r \sim_{\underline{\Omega}} i$ to indicate the predicate $wt_{\underline{\Omega}}(r) = mr_{\underline{\Omega}}(i)$ (read r is a valid encoding for i). By construction the following predicate is thus true:

$$H_{\underline{\Omega}, \underline{A}}[\widetilde{Q}]_{i,j} \iff \exists r, s. \widetilde{Q}_{r,s} \wedge (r \sim_{\underline{\Omega}} i) \wedge (s \sim_{\underline{A}} j) \quad (20)$$

B.1 Relevant theorems and proofs: Section 2

Proof (Theorem 1) With the new notations, Theorem 1 can be rewritten as follows:

Given a vector of parts composed of a single part $\underline{\Pi} = [\pi]$ we have that $H_{\underline{\Pi}, \underline{\Pi}}[I_{2^\pi}] = I_{\pi+1}$.

To prove it, we elaborate the predicate in Eq. (20):

$$\begin{aligned} H_{\underline{\Pi}, \underline{\Pi}}[I_{2^\pi}]_{i,j} &\iff \exists r, s. I_{r,s} \wedge (r \sim_{\underline{\Pi}} i) \wedge (s \sim_{\underline{\Pi}} j) \\ \vdash H_{\underline{\Pi}, \underline{\Pi}}[I_{2^\pi}]_{i,j} &\iff \exists r, s. (r == s) \wedge (r \sim_{\underline{\Pi}} i) \wedge (s \sim_{\underline{\Pi}} j) \\ \vdash H_{\underline{\Pi}, \underline{\Pi}}[I_{2^\pi}]_{i,j} &\iff (i == j) \end{aligned}$$

and note that the implied condition means exactly that it must be an identity matrix. $\square$

Proof (Theorem 2) With the new notations, Theorem 2 can be rewritten as follows:

Given two correlation matrices $X \in K^{2^\Omega \times 2^Z}$ and $Y \in K^{2^Z \times 2^A}$, the following dominance holds between the shares' relation matrices:

$$H_{\underline{\Omega}, \underline{A}}[\widetilde{XY}] \preceq H_{\underline{\Omega}, \underline{Z}}[\widetilde{X}] H_{\underline{Z}, \underline{A}}[\widetilde{Y}]$$

for any choice of compatible⁹ parts $\underline{\Omega}$, $\underline{Z}$ and $\underline{A}$.

To prove it, note that

$$H_{\underline{\Omega}, \underline{A}}[\widetilde{XY}] \preceq H_{\underline{\Omega}, \underline{Z}}[\widetilde{X}] H_{\underline{Z}, \underline{A}}[\widetilde{Y}]$$

represents the following implication (assuming $\sum_{z_i \in \underline{Z}} z_i = Z$):

$$H_{\underline{\Omega}, \underline{A}}[\widetilde{XY}]_{i,j} \implies \exists \zeta. H_{\underline{\Omega}, \underline{Z}}[\widetilde{X}]_{i,\zeta} \wedge H_{\underline{Z}, \underline{A}}[\widetilde{Y}]_{\zeta,j}$$

which follows directly from the definition in Eq. (20) (note that once we have t, ζ exists since we assume Z is compatible):

$$\begin{aligned} H_{\underline{\Omega}, \underline{A}}[\widetilde{XY}]_{i,j} &\iff \exists r, s. (XY_{r,s} \neq 0) \wedge (r \sim_{\underline{\Omega}} i) \\ &\quad \wedge (s \sim_{\underline{A}} j) \\ \vdash H_{\underline{\Omega}, \underline{A}}[\widetilde{XY}]_{i,j} &\implies \exists r, s, t. (r \sim_{\underline{\Omega}} i) \wedge (s \sim_{\underline{A}} j) \\ &\quad \wedge (X_{r,t} \neq 0) \wedge (Y_{t,s} \neq 0) \\ \vdash H_{\underline{\Omega}, \underline{A}}[\widetilde{XY}]_{i,j} &\implies \exists r, s, t, \zeta. (r \sim_{\underline{\Omega}} i) \wedge (s \sim_{\underline{A}} j) \\ &\quad \wedge (\zeta \sim_{\underline{Z}} t) \wedge \widetilde{X}_{r,t} \wedge \widetilde{Y}_{t,s} \\ \vdash H_{\underline{\Omega}, \underline{A}}[\widetilde{XY}]_{i,j} &\implies \exists \zeta. H_{\underline{\Omega}, \underline{Z}}[\widetilde{X}]_{i,\zeta} \\ &\quad \wedge H_{\underline{Z}, \underline{A}}[\widetilde{Y}]_{\zeta,j} \end{aligned}$$

where, in the second step, we applied the following axiom:

$$XY_{r,s} \neq 0 \implies \exists t. X_{r,t} \neq 0 \wedge Y_{t,s} \neq 0$$

Note that the converse ($\Leftarrow$) does not hold since matrix multiplication between Walsh matrices is done over rational numbers which might have different signs so it may cancel out. $\square$

⁹ By compatible, we mean that the set of parts sums up to the specified size, e.g., $\sum_{\omega_i \in \underline{\Omega}} \omega_i = \Omega$.

Proof (Theorem 3) With the new notations, Theorem 3 can be rewritten as follows:

Given two correlation matrices $X \in K^{2^\Phi \times 2^\Psi}$ and $Y \in K^{2^\Omega \times 2^A}$, the following holds:

$$H_{\frac{\Omega}{\Phi}, \frac{A}{\Psi}}[\widetilde{X \otimes Y}] = H_{\Phi, \Psi}[\widetilde{X}] \otimes H_{\Omega, A}[\widetilde{Y}]$$

where $\frac{\Omega}{\Phi} = \Phi \parallel \Omega = \{\phi_{N_\Phi}, \dots, \phi_1, \omega_{N_\Omega}, \dots, \omega_1\}$, i.e., the concatenation of parts Φ and Ω

Before proceeding with the proof, we need to introduce the following lemma:

Lemma 5 (Equivalences over concatenations of vector of parts) When dealing with a concatenation of parts $\frac{\Omega}{\Phi}$, both the extended Hamming weight and the multi-radix representation comply with the following equivalences:

$$wt_{\frac{\Omega}{\Phi}}(r) = wt_{\Phi}(r \uparrow \pi_{\Omega}) \parallel wt_{\Omega}(r \downarrow \pi_{\Omega})$$

$$mr_{\frac{\Omega}{\Phi}}(i) = mr_{\Phi}(i \uparrow \pi_{\Omega}) \parallel mr_{\Omega}(i \downarrow \pi_{\Omega})$$

where $\parallel$ is the vector concatenation while $\pi_{\Omega} = \prod_i (\omega_i + 1)$. This means that we can split $r \sim_{\frac{\Omega}{\Phi}} i$ in the conjunction of two sub conditions:

$$r \sim_{\frac{\Omega}{\Phi}} i \iff (r \uparrow \pi_{\Omega} \sim_{\Phi} i \uparrow \pi_{\Omega}) \wedge (r \downarrow \pi_{\Omega} \sim_{\Omega} i \downarrow \pi_{\Omega}).$$

Theorem 3 is easily proved by expanding $H_{\frac{\Omega}{\Phi}, \frac{A}{\Psi}}[\widetilde{X \otimes Y}]$ through the definition in Eq. (20), and apply successively Lemma 5 and Definition 6:

$$\begin{aligned} H_{\frac{\Omega}{\Phi}, \frac{A}{\Psi}}[\widetilde{X \otimes Y}]_{i,j} &= \exists r, s. (X \otimes Y)_{r,s} \neq 0 \wedge (r \sim_{\frac{\Omega}{\Phi}} i) \wedge (s \sim_{\frac{A}{\Psi}} j) \\ &= \exists r, s. X_{r \uparrow \pi_{\Omega}, s \uparrow \pi_A} \neq 0 \wedge Y_{r \downarrow \pi_{\Omega}, s \downarrow \pi_A} \neq 0 \wedge (r \sim_{\frac{\Omega}{\Phi}} i) \wedge (s \sim_{\frac{A}{\Psi}} j) \\ &= \exists r, s. X_{r \uparrow \pi_{\Omega}, s \uparrow \pi_A} \neq 0 \wedge Y_{r \downarrow \pi_{\Omega}, s \downarrow \pi_A} \neq 0 \wedge (r \uparrow \pi_{\Omega} \sim_{\Phi} i \uparrow \pi_{\Omega}) \\ &\quad \wedge (r \downarrow \pi_{\Omega} \sim_{\Omega} i \downarrow \pi_{\Omega}) \wedge (s \uparrow \pi_A \sim_{\Psi} j \uparrow \pi_A) \wedge (s \downarrow \pi_A \sim_A j \downarrow \pi_A) \\ &= H_{\Phi, \Psi}[\widetilde{X}]_{i \uparrow \pi_{\Omega}, j \uparrow \pi_A} \wedge H_{\Omega, A}[\widetilde{Y}]_{i \downarrow \pi_{\Omega}, j \downarrow \pi_A} \\ &= (H_{\Phi, \Psi}[\widetilde{X}] \otimes H_{\Omega, A}[\widetilde{Y}])_{i,j}. \end{aligned}$$

Proof (Corollary 1) With the new notations, Corollary 1 can be rewritten as follows:

$$\begin{aligned} H_{\frac{\Omega}{[x]}, \frac{A}{[x]}}[I_{2^x} \otimes \widetilde{W}_f]_{i,j} &= \delta(i_x, j_x) \cdot F_{i_{\omega[\Omega]}, j_{\omega_1}}^{j_{\alpha[\Omega]} \dots j_{\alpha_1}} \\ H_{\frac{\Omega}{[x]}, \frac{A}{[x]}}[\widetilde{W}_f \otimes I_{2^x}]_{i,j} &= F_{i_{\omega[\Omega]}, j_{\omega_1}}^{j_{\alpha[\Omega]} \dots j_{\alpha_1}} \cdot \delta(i_x, j_x) \end{aligned}$$

For the first equality:

$$\begin{aligned} H_{\frac{\Omega}{[x]}, \frac{A}{[x]}}[I_{2^x} \otimes \widetilde{W}_f]_{i,j} &\iff [I_{x+1} \otimes H_{\Omega, A}(\widetilde{W}_f)]_{i,j} \\ &\iff \delta(i_x, j_x) \cdot F_{i_{\omega[\Omega]}, j_{\omega_1}}^{j_{\alpha[\Omega]} \dots j_{\alpha_1}}. \end{aligned}$$

The second equality can be deduced in the same way. $\square$

C Relevant theorems and proofs: Section 3

Proof (Lemma 1) The complete relation matrix ABC computed in Fig. 4 is such that

$$\begin{aligned} (ABC)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_{\alpha}} &= \sum_{p,q,r} AB_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{p,q,r} C_{p,q,r}^{0,0,j_{\alpha}} \\ &= \sum_{p,q,r} \sum_{l,m,n} A_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{l,m,n} B_{l,m,n}^{p,q,r} C_{p,q,r}^{0,0,j_{\alpha}} \stackrel{\text{Cor.}}{=} \sum_{p,q,r} A_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{p,q,r} C_{p,q,r}^{0,0,j_{\alpha}} \\ &\stackrel{\text{Cor.}}{=} \sum_{p,r} G_{i_{\pi_g} i_{\omega_g}}^{p,r} C_{p,i_{\pi_f},r}^{0,0,j_{\alpha}} \stackrel{\text{Cor.}}{=} \sum_r G_{i_{\pi_g} i_{\omega_g}}^{0,r} F_{i_{\pi_f},r}^{0,j_{\alpha}}. \end{aligned}$$

$\square$

Proof (Lemma 2) S is the shares' relation matrix computed from the relation matrix $\widetilde{W}_s$ of the duplication function $s = x \mapsto (x, x)$. It can be shown that elements of this relation matrix are such that:

$$\widetilde{W}_s(l, m) \iff l \uparrow 2^n \oplus l \downarrow 2^n = m.$$

To prove Lemma 2, we proceed with a reduction ad absurdum, i.e., we show that

$$\exists i_{\alpha_1}, i_{\alpha_2}, j_{\alpha}. (i_{\alpha_1} + i_{\alpha_2} < j_{\alpha}) \wedge S_{i_{\alpha_1}, i_{\alpha_2}}^{j_{\alpha}} = 1$$

is a contradiction. To derive it we expand in it the definition of shares' relation matrix S :

$$\begin{aligned} &\exists i_{\alpha_1}, i_{\alpha_2}, j_{\alpha}. (i_{\alpha_1} + i_{\alpha_2} < j_{\alpha}) \wedge S_{i_{\alpha_1}, i_{\alpha_2}}^{j_{\alpha}} = 1 \\ &\vdash \exists l, m. \widetilde{W}_s(l, m) \wedge wt_{\alpha_1}^{(l)}(l) \\ &= (i_{\alpha_1}) \wedge wt_{\alpha}(m) = j_{\alpha} \wedge (i_{\alpha_1} + i_{\alpha_2} < j_{\alpha}) \\ &\vdash \exists l, m. (l \uparrow 2^n \oplus l \downarrow 2^n = m) \wedge wt_{\alpha_1}(l \uparrow 2^n) = i_{\alpha_1} \wedge wt_{\alpha_2}(l \downarrow 2^n) \\ &= i_{\alpha_2} \wedge \wedge wt_{\alpha}(m) = j_{\alpha} \wedge (i_{\alpha_1} + i_{\alpha_2} < j_{\alpha}) \\ &\vdash \exists l. wt_{\alpha_1}(l \uparrow 2^n) = i_{\alpha_1} \wedge wt_{\alpha_2}(l \downarrow 2^n) = i_{\alpha_2} \wedge wt_{\alpha}(l \uparrow 2^n \oplus l \downarrow 2^n) \\ &= j_{\alpha} \wedge \wedge (i_{\alpha_1} + i_{\alpha_2} < j_{\alpha}) \\ &\vdash wt_{\alpha_1}(l \uparrow 2^n) + wt_{\alpha_2}(l \downarrow 2^n) < wt_{\alpha}(l \uparrow 2^n \oplus l \downarrow 2^n) \end{aligned}$$

the latter judgment is absurd because for all binary vectors a, b holds that $wt(a) \oplus wt(b) \geq wt(a \oplus b)$. $\square$

Proof (Lemma 3) The complete relation matrix $ABCD$ computed in Fig. 5 is such that

$$(ABCD)_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{0,0,j_\alpha} = \sum_{t,u,v,z} ABC_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{t,u,v,z} D_{t,u,v,z}^{0,0,j_\alpha}$$

$$\stackrel{\text{Lem. 6}}{=} \sum_{t,u,v,z} \sum_r G_{i_{\pi_g} i_{\omega_g}}^{t,r,z} F_{i_{\pi_f},r}^{u,v} D_{t,u,v,z}^{0,0,j_\alpha}$$

$$\stackrel{\text{Cor. 1}}{=} \sum_{v,z} \left(\sum_r G_{i_{\pi_g} i_{\omega_g}}^{0,r,z} F_{i_{\pi_f},r}^{0,v} \right) S_{v,z}^{j_\alpha}$$

where we used the following:

Lemma 6

$$ABC_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{t,u,v,z} = \sum_{p,q,r,s} AB_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{p,q,r,s} C_{p,q,r,s}^{t,u,v,z}$$

$$= \sum_{p,q,r,s} \sum_{l,m,n,o} A_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{l,m,n,o} B_{l,m,n,o}^{p,q,r,s} C_{p,q,r,s}^{t,u,v,z}$$

$$\stackrel{\text{Cor. 1}}{=} \sum_{p,q,r,s} A_{i_{\pi_f} i_{\pi_g} i_{\omega_g}}^{q,p,r,s} C_{p,q,r,s}^{t,u,v,z}$$

$$\stackrel{\text{Cor. 1}}{=} \sum_{p,r,s} G_{i_{\pi_g} i_{\omega_g}}^{p,r,s} C_{p,i_{\pi_f},r,s}^{t,u,v,z} \stackrel{\text{Cor. 1}}{=} \sum_r G_{i_{\pi_g} i_{\omega_g}}^{t,r,z} F_{i_{\pi_f},r}^{u,v}$$

□

References

- Ishai, Y., Sahai, A., Wagner, D.: Private Circuits: Securing Hardware against Probing Attacks. In: Boneh, D. (ed.) *Advances in Cryptology-CRYPTO 2003*. Lecture Notes in Computer Science, pp. 463–481. Springer, Berlin Heidelberg (2003)
- Coron, J., Prouff, E., Rivain, M., Roche, T.: Higher-Order Side Channel Security and Mask Refreshing. In: Moriai, S. (ed.) *Fast Software Encryption*. Lecture Notes in Computer Science, pp. 410–424. Springer, Berlin Heidelberg (2014)
- Barthe, G., Belaïd, S., Dupressoir, F., Fouque, P. A., Grégoire, B., Strub, P. Y., Zucchini, R.: Strong Non-Interference and Type-Directed Higher-Order Masking. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS '16*, pages 116–129, New York, NY, USA, (2016). ACM
- Barthe, G., Belaïd, S., Dupressoir, F., Fouque, P. A., Grégoire, B.: Compositional verification of higher-order masking: application to a verifying masking compiler. *IACR Cryptol. ePrint Arch.* **2015**, 506 (2015)
- Belaïd, S., Dahmun, G., and Matthieu R.: Tight private circuits: achieving probing security with the least refreshing. *Technical Report 439*, (2018)
- Barthe, G., Dupressoir, F., Faust, S., Grégoire, B., Standaert, F.X., Strub, P.Y.: Parallel implementations of masking schemes and the bounded moment leakage model. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10210 LNCS:535–566, (2017)
- Faust, S., Grosso, V., Del Pozo, S. M., Paglialonga, C., Standaert, F. X.: Composable Masking Schemes in the Presence of Physical Defaults and the Robust Probing Model. *Technical Report 711*, (2017). B
- Schmidt, J. M., Kim, C.: A probing attack on AES. pages 256–265, 02 (2009)
- Rivain, M., Prouff, E.: Provably Secure Higher-Order Masking of AES. *Technical Report 441*, (2010)
- Daemen, J., Govaerts, R., Vandewalle, J.: Correlation matrices. In: Preneel, B. (ed.) *Fast Software Encryption*. Lecture Notes in Computer Science, pp. 275–285. Springer, Berlin Heidelberg (1995)
- Parriaux, J., Guillot, P., Millérioux, G.: Towards a spectral approach for the design of self-synchronizing stream ciphers. *Cryptogr. Commun.* **3**(4), 259–274 (2011)
- Bloem, R., Gross, H., Iusupov, R., Könighofer, B., Mangard, S., Winter, J.: Formal Verification of Masked Hardware Implementations in the Presence of Glitches. In: Nielsen, J.B. and Rijmen, V., editors, *Advances in Cryptology - EUROCRYPT 2018*, Lecture Notes in Computer Science, pages 321–353. Springer International Publishing, (2018)
- Beyne, T., Dhooghe, S. and Zhenda, Z.: A not so random idea, *Cryptanalysis of masked ciphers* (2020)
- Bringer, J., Carlet, C., Chabanne, H., Guilley, S., Maghrebi, H.: Orthogonal direct sum masking. In David, N. and Damien, S., editors, *Information Security Theory and Practice. Securing the Internet of Things*, pages 40–56, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg
- Cheng, W., Guilley, S., Carlet, C., Mesnager, S., Danger, J.L.: Optimizing inner product masking scheme by a coding theory approach. *IEEE Trans. Inf. Forensics Secur.* **16**, 220–235 (2021)
- Zaccaria, V., Melzani, F., Bertoni, G.: Spectral features of higher-order side-channel countermeasures. *IEEE Trans. Comput.* **67**(4), 596–603 (2018)
- Xiao, G.Z., Massey, J.L.: A spectral characterization of correlation-immune combining functions. *IEEE Trans. Inf. Theory* **34**(3), 569–571 (1988)
- Carlet, C.: Boolean Functions for Cryptography and Error-Correcting Codes, pp 257–397. *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, (2010)
- Copilowish, I.M.: Matrix development of the calculus of relations. *J. Symbol. Logic* **13**(04), 193–203 (1948)
- Chiara, M. M., Zaccaria, V.: On the spectral features of robust probing security. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp 24–48, August (2020)
- Cassiers, G., Faust, S., Maximilian, O., and Standaert, F.-X.: Towards tight random probing security. *Technical Report 880*, (2021)
- Coron, J.: Higher Order Masking of Look-Up Tables. In: Nguyen, P.Q., Oswald, E. (eds.) *Advances in Cryptology – EUROCRYPT 2014*. Lecture Notes in Computer Science, pp. 441–458. Springer, Berlin Heidelberg (2014)
- Cassiers, G., Standaert, F.X.: Towards globally optimized masking: from low randomness to low noise rate: or probe isolating multiplications with reduced randomness and security against horizontal attacks. *IACR Trans. Cryptogr. Hardw. Embed. Syst.* **2019**(2), 162–198 (2019)
- Carlet, C.: Vectorial Boolean Functions for Cryptography. In: Crama, Y., Hammer, P.L. (eds.) *Boolean Models and Methods in Mathematics. Computer Science, and Engineering*, pp. 398–470. Cambridge University Press, Cambridge (2010)
- Dravie, B., Parriaux, J., Guillot, P., Millérioux, G.: Matrix representations of vectorial Boolean functions and eigenanalysis. *Cryptogr. Commun. Discret. Struct. Boolean Funct. Seq.* **8**(4), 555–577 (2016)
- Jérémy, P., Philippe, G., Gilles, M.: Towards a spectral approach for the design of self-synchronizing stream ciphers. *Cryptogr. Commun.* **3**(4), 259–274 (2011)
- Beyne, T.: Block cipher invariants as eigenvectors of correlation matrices (full version). *Cryptology ePrint Archive*, Report 2018/763, (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.