Strong Non-Interference and Type-Directed Higher-Order Masking

Gilles Barthe, Sonia Belaïd, François Dupressoir, Pierre-Alain Fouque, Benjamin Grégoire, Pierre-Yves Strub, Rébecca Zucchini

Strong Non-Interference and Type-Directed Higher-Order Masking*

Gilles Barthe[1], Sonia Belaïd[2], François Dupressoir[3†], Pierre-Alain Fouque[4], Benjamin Grégoire[5], Pierre-Yves Strub[6†], and Rébecca Zucchini[5,7]
(masking@projects.easycrypt.info)

[1] IMDEA Software Institute
[2] Thales Communications & Security
[3] University of Surrey
[4] Université de Rennes 1
[5] Inria Sophia Antipolis – Méditerranée
[6] École Polytechnique
[7] École Normale Supérieure de Cachan

Abstract. Differential power analysis (DPA) is a side-channel attack in which an adversary retrieves cryptographic material by measuring and analyzing the power consumption of the device on which the cryptographic algorithm under attack executes. An effective countermeasure against DPA is to *mask* secrets by probabilistically encoding them over a set of shares, and to run *masked* algorithms that compute on these encodings. Masked algorithms are often expected to provide, at least, a certain level of *probing security*.

Leveraging the deep connections between probabilistic information flow and probing security, we develop a precise, scalable, and fully automated methodology to verify the probing security of masked algorithms, and generate them from unprotected descriptions of the algorithm. Our methodology relies on several contributions of independent interest, including a stronger notion of probing security that supports compositional reasoning, and a type system for enforcing an expressive class of probing policies. Finally, we validate our methodology on examples that go significantly beyond the state-of-the-art.

1 Introduction

Differential power analysis, or DPA [25], is a class of side-channel attacks in which an adversary extracts secret data from the power consumption of the device on which a program manipulating the data executes. One practical countermeasure against DPA, called *masking* [11,22], transforms an algorithm that performs computations over a finite ring

* This work appears in the Proceedings of CCS 2016. This is the long version. A preliminary version, made public in 2015 under the title "Compositional Verification of Higher-Order Masking: Application to a Verifying Masking Compiler", can be found as revision 20150527:192221 of this report.

† The majority of the work presented here was performed while the author was working for the IMDEA Software Institute.

$\mathbb{K}$ into a randomized algorithm that manipulates probabilistic encodings.[8] At an abstract level, any masking transformation performs two tasks. First, it replaces every algebraic operation performed by the original algorithm by a call to a *gadget*, i.e. a probabilistic algorithm that simulates the behavior of algebraic operations on probabilistic encodings. Second, it inserts *refreshing* gadgets, i.e. gadgets that take a probabilistic encoding of v and rerandomizes its shares in order to produce another probabilistic encoding w of v. Inserting refreshing gadgets does not change the functional behavior of the masked algorithm, and increases the randomness complexity and execution time of the masked program. However, it is also compulsory for achieving security. Therefore, an important line of research is to find suitable trade-offs that ensure security while minimizing the performance overhead of masking; see [8] for recent developments in this direction.

The baseline notion of security for masked algorithms is t-probing security. Informally, an algorithm P is t-probing secure if the values taken by at most t intermediate variables of P during execution do not leak any information about secrets (held by its inputs). More formally, an algorithm P achieves t-probing security iff for every set of at most t intermediate variables, the joint distributions of the values taken by these intermediate variables coincide for any two executions initiated from initial inputs that agree on t shares of each input encoding. Stated in this form, probing security is an instance of probabilistic information flow, universally quantified over all position sets that meet a cardinality constraint, and is therefore potentially amenable to formal analysis using a well-developed body of work on language-based security and program verification. Indeed, the connection between probing security and information flow has been instrumental in a promising line of research, initiated in [27] and further developed in [7,20,19,4], which uses type systems, program logics, SMT solvers and other methods for verifying or synthesizing masked algorithms at small (≤ 5) orders. However, none of these works addresses the problem of composition, and all fail to scale either to higher orders or to larger algorithms.

Contributions We develop precise and scalable techniques for synthesizing masked algorithms that achieve probing security. Our techniques apply to a wide range of probing policies, including existing policies and new policies defined in this paper, and deliver masked algorithms that outperform (in terms of randomness complexity and computational efficiency) prior approaches. In more detail, we make the following broad contributions:

1. *Strong non-interference.* We introduce a stronger notion of probing security, which we call strong non-interference, and prove that it is in fact satisfied by many (but not all) gadgets from the literature. Furthermore, we justify that strong non-interference is the desired property for refreshing gadgets, by reconsidering known negative and positive results [15] for a simplified example extracted from Rivain and Prouff's inversion algorithm [31]. We first observe that the refreshing gadget used in the original, flawed, algorithm does not enjoy strong non-interference. Second, we note that the refreshing gadget used in the fixed, secure, algorithm is indeed strongly non-interfering, and we

[8] A t-encoding of an element $v \in \mathbb{K}$ is a $(t+1)$-tuple $\mathbf{v} = \langle \mathbf{v}^0, \ldots, \mathbf{v}^t \rangle$ such that $[\![\mathbf{v}]\!] \triangleq \mathbf{v}^0 \oplus \ldots \oplus \mathbf{v}^t = v$. Each of the $\mathbf{v}^i \in \mathbb{K}$ in an encoding $\mathbf{v}$ of v is called a *share*. Moreover, t is called the masking order. A probabilistic encoding of v is a distribution over encodings of v.

show that one can prove the probing security of the fixed algorithm, based simply on the assumption that the refreshing gadget is strongly non-interfering. Generalizing these observations, we prove that every non-interfering algorithm can be turned into a strongly non-interfering algorithm, by processing its inputs or its output with a strongly non-interfering refreshing gadget. We also provide more general results about the composition of strongly non-interfering gadgets.

2. *Formal proofs.* We develop and implement an automated method, inspired from [4], for checking strong non-interference. We apply our automated verifier for strong non-interference to several gadgets from the literature and some other interesting compositions, for orders $t \leq 6$. For several more widely-used gadgets, we further use EasyCrypt [5] to provide machine-checked proofs of t-probing security for all t.

3. *Type-based enforcement of probing security.* We define an expressive language for specifying a large class of non-interference properties with cardinality constraints. Our language can be seen as a variant of the first-order theory of finite sets with cardinality constraints [32,3], and can be used to specify baseline probing security and strong non-interference, among others. Then, we define a type system that enforces probing policies and prove its soundness. Furthermore, we show how to model in our language of probing policies the notion of affine gadget, and we show how it helps improve the precision of type-checking.

4. *Certifying Masking Transformation*. As a proof of concept, we implement a type inference algorithm and a certifying masking transformation that takes as input an arithmetic expression and returns a masked algorithm typable by our type system.[9] Our transformation improves over prior works by selectively inserting refreshing gadgets only at points where type-checking would otherwise fail. This strategy leads to improved efficiency while retaining provable soundness.

5. *Practical evaluation.* We evaluate our type system and masking transformation on complete algorithms at various orders, often achieving provable t-probing security levels far beyond the state-of-the-art for algorithms of those sizes, and with better performance than most known (provably secure) algorithms in terms of time, memory and randomness complexity.

Related work Section 9 discusses related work in more detail. Here we focus on recent work on automated tools for the verification of synthesis of masked algorithms, starting with Moss et al. [27], who point out and leverage connections between probing security and probabilistic information-flow for first-order boolean masking schemes. Subsequent works in this direction accommodate higher-order and arithmetic masking, using type systems and SMT solvers [7], or model counting and SMT solvers [20,19]. Although approaches based on model counting are more precise than early approaches based on type systems and can be extended to higher-order masking schemes, their algorithmic complexity constrains their applicability. In particular, existing tools based on model counting can only analyze first or second order masked implementations, and can only deal with round-reduced versions of the algorithms they consider (for instance, only

[9] The cryptography literature often refers to such transformations as *masking compilers*. We purposely avoid this terminology, since the terms is used in programming languages for transformations that output executable code

analyzing a single round of Keccak at order 1, and algorithms for field operations at orders 2 and higher). Breaking away from model counting, Barthe et al. [4] develop efficient algorithms for analyzing the security of masked algorithms in the probing model. Their approach outperforms previous work and can analyze a full block of AES at first-order, reduced-round (4 rounds) AES at the second-order, and several S-box computation algorithms masked at the third and fourth orders. However, their work does not readily scale either to higher orders or to larger algorithms, mainly due to the lack of composition results.

Our work also bears some connections with language-based security, and in particular with work on the specification and the enforcement of confidentiality policies using techniques from programming languages. For instance, our work has similarities with the work of Pettai and Laud [29], who develop techniques for proving security of multi-party computations in the presence of strong adversaries, and work by Zdancewic et al. [33], who propose a compiler that partitions programs for secure distributed execution.

Mathematical preliminaries A function $\mu : B \to \mathbb{R}^{\geq 0}$ is a (discrete) *distribution* over B if the subset $\mathsf{supp}(\mu)$ of B with non-zero weight under μ is discrete and moreover $\sum_{b \in \mathsf{supp}(\mu)} \mu(b) = 1$. We let $\mathcal{D}(B)$ denote the set of discrete distributions over B. Equality of distributions is defined as pointwise equality of functions. Distributions can be given a monadic structure with the two operators $\mathsf{munit}(\cdot)$ and $\mathsf{mlet}\ \cdot = \cdot$. For every $b \in B$, $\mathsf{munit}(b)$ is the unique distribution μ such that $\mu(b) = 1$. Moreover, given $\mu : \mathcal{D}(B)$ and $M : B \to \mathcal{D}(C)$, $\mathsf{mlet}\ x = \mu\ \mathsf{in} M\ x$ is the unique distribution μ' over C such that $\mu'(c) = \sum_b \mu(b)\ M(b)(c)$.

We often use the notion of marginals. The first and second marginals of a distribution $\mu \in \mathcal{D}(B_1 \times B_2)$ are the distributions $\pi_1(\mu) \in \mathcal{D}(B_1)$ and $\pi_2(\mu) \in \mathcal{D}(B_2)$ given by

$$\pi_1(\mu)(b_1) = \sum_{b_2 \in B_2} \mu(b_1, b_2) \qquad \pi_2(\mu)(b_2) = \sum_{b_1 \in B_1} \mu(b_1, b_2).$$

The notion of marginal readily extends to distributions over finite maps (rather than pairs).

2 A bird's eye view of strong non-interference

Before formalizing our definitions, we give an intuitive description of our language for gadgets and of our security notions, based on simple examples.

Gadgets and Positions Gadget $\mathsf{RefreshM}_2$ (Gadget 1) shows the description in our language of a mask refreshing gadget for $t = 2$. The gadget takes as input an encoding variable $\mathbf{a} \in \mathbb{K}^3$, where $\mathbb{K}$ is some finite ring and returns a new encoding $\mathbf{c} \in \mathbb{K}^3$ such that $[\![\mathbf{a}]\!] = [\![\mathbf{c}]\!]$. The gadget first makes local copies of individual input shares $\mathbf{a}^\imath$ (for $0 \leq \imath \leq 2$) of $\mathbf{a}$ into local variables $c_\imath$ (for $0 \leq \imath \leq 2$). After this first step, we sample uniform random elements from $\mathbb{K}$ into a local variable r and perform some ring operations. Finally, the algorithm returns a vector in $\mathbb{K}^3$, constructed from the final value of local variables c_0, c_1 and c_2.

Gadget 1 SNI Mask Refreshing with $t = 2$

function $\mathsf{RefreshM}_2(\mathbf{a})$
$c_{0,0} \leftarrow \mathbf{a}^0$; $c_{1,0} \leftarrow \mathbf{a}^1$; $c_{2,0} \leftarrow \mathbf{a}^2$;
$r_0 \xleftarrow{\$} \mathbb{K}$; $c_{0,1} \leftarrow c_{0,0} \oplus r_0$; $c_{1,1} \leftarrow c_{1,0} \ominus r_0$;
$r_1 \xleftarrow{\$} \mathbb{K}$; $c_{0,2} \leftarrow c_{0,1} \oplus r_1$; $c_{2,1} \leftarrow c_{2,0} \ominus r_1$;
$r_2 \xleftarrow{\$} \mathbb{K}$; $c_{1,2} \leftarrow c_{1,1} \oplus r_2$; $c_{2,2} \leftarrow c_{2,1} \ominus r_2$;
return $\langle c_{0,2}, c_{1,2}, c_{2,2} \rangle$

Note that the gadget is written in *single static assignment* (SSA) form, an intermediate representation in which each variable is defined exactly once. Having gadgets written in SSA form allows us to easily refer to the value of a particular variable at a particular point in the program–simply by referring to its name, which corresponds to a unique definition. In this paper, we refer to *positions* in gadgets and algorithms, which correspond exactly to intermediate variables. We distinguish between three different kinds of positions: *input positions*, which correspond to shares of the gadget's input (here, $\mathbb{I}_{\mathsf{RefreshM}_2} = \{\mathbf{a}^0, \mathbf{a}^1, \mathbf{a}^2\}$), *output positions*, which correspond to the variables that appear in the gadget's return vector (here, $\mathbb{O}^{\text{int}}_{\mathsf{RefreshM}_2} = \{c_{0,2}, c_{1,2}, c_{2,2}\}$), and *internal positions*, which refer to all other positions (here, $\mathbb{O}^{\text{ext}}_{\mathsf{RefreshM}_2} = \{c_{0,0}, c_{1,0}, c_{2,0}, c_{0,1}, c_{1,1}, c_{2,1}, r_0, r_1, r_2\}$). Intuitively, this separation allows us to distinguish between direct observations made by the adversary into a gadget (as *internal* positions), output shares about which the adversary may have learned some information by probing gadgets that use them as input (as *output* positions), and shares of the gadget's inputs (as *input* positions) about which the adversary is now learning information. In the following, we often write "the joint distribution of a set of positions" to discuss the joint distribution of the variables defined at these positions in the gadget (in order). For example, referring to $\mathsf{RefreshM}_2$, the joint distribution of the ordered set $\mathcal{O} = \langle c_{0,1}, c_{2,2} \rangle$ of positions can be described as the following function of $\mathbf{a}$, where we use $\$$ to denote a fresh uniform random sample in $\mathbb{K}$ (using indices to denote distinct samples): $[\![\mathsf{RefreshM}_2]\!]_{\mathcal{O}}(\mathbf{a}) \triangleq \langle \mathbf{a}^0 \oplus \$_0, (\mathbf{a}^2 \ominus \$_1) \ominus \$_2 \rangle$.

Probing Security and Non-Interference The $\mathsf{RefreshM}_2$ gadget is known to be 2-probing secure, or 2-non-interfering (2-NI) in our terminology, in the sense that the joint distribution of any set of at most 2 of its positions, corresponding to adversary *probes*, depends on at most 2 shares of the gadget's inputs. This guarantees, if the input encoding is uniform, that no information about it leaks through any 2 probes in the circuit.

Considering again the set $\mathcal{O} = \langle c_{0,1}, c_{2,2} \rangle$ of positions in $\mathsf{RefreshM}_2$ and its distribution $[\![\mathsf{RefreshM}_2]\!]_{\mathcal{O}}$, it is easy to see–purely syntactically–that it depends at most on shares $\mathbf{a}^0$ and $\mathbf{a}^2$ of the gadget's input encoding. Similarly considering all possible pairs of positions, we can prove that each of them has a joint distribution that depends on at most two shares of $\mathbf{a}$.

Strong Non-Interference Probing security is generally not *composable*: combining t-probing secure gadgets does not necessarily yield a t-probing secure algorithm [15]. Our main contribution is a new and stronger notion of security for gadgets, which we

dub *strong non-interference* (or SNI), which does support some compositional reasoning. SNI reinforces probing security by requiring that the number of input shares on which the distribution of a given position set may depend be determined only by the number of *internal* positions present in that set. For example, consider again position set $\mathcal{O} = \langle c_{0,1}, c_{2,2} \rangle$ in $\mathsf{RefreshM}_2$, and note that it contains only one internal position ($c_{0,1}$). We have seen that the joint distribution $[\![\mathsf{RefreshM}_2]\!]_{\mathcal{O}}$ of that position set syntactically depends on two shares of $\mathbf{a}$. However, it can be equivalently expressed as $[\![\mathsf{RefreshM}_2]\!]_{\mathcal{O}}(\mathbf{a}) = \langle \$_0, (\mathbf{a}^2 \ominus \$_1) \ominus \$_2 \rangle$ (since the ring addition $\oplus$ is a bijection of each of its arguments and $\$_0$ is a fresh and uniform ring element). This shows that the distribution in fact depends on at most one share of $\mathbf{a}$ (here $\mathbf{a}^2$). In fact, it can be shown that $\mathsf{RefreshM}_2$ is 2-SNI. More generally, surprisingly many gadgets from the literature achieve SNI.

However, and not unexpectedly, some gadgets from the literature do not satisfy SNI. Consider for instance $\mathsf{RefreshA}_2$ (Gadget 2). It is easy to see that the gadget is 2-NI (each position $c_{i,j}$ depends only on input share $\mathbf{a}^i$, and each position r_i is completely independent from the input encoding). Still, looking at position set $\mathcal{O}' = \langle c_{0,1}, c_{1,1} \rangle$, which is composed of one internal position and one external one, we see that the distribution $[\![\mathsf{RefreshA}_2]\!]_{\mathcal{O}'} \triangleq \langle \mathbf{a}^0 \oplus \$_0, \mathbf{a}^1 \oplus \$_0 \rangle$ does depend on more than one share of $\mathbf{a}$. $\mathsf{RefreshA}_2$ is therefore not 2-SNI.

Gadget 2 NI Mask Refreshing with $t = 2$

function $\mathsf{RefreshA}_2(\mathbf{a})$
$c_{0,0} \leftarrow \mathbf{a}^0$; $c_{1,0} \leftarrow \mathbf{a}^1$; $c_{2,0} \leftarrow \mathbf{a}^2$;
$r_0 \xleftarrow{\$} \mathbb{K}$; $c_{0,1} \leftarrow c_{0,0} \oplus r_0$; $c_{1,1} \leftarrow c_{1,0} \ominus r_0$;
$r_1 \xleftarrow{\$} \mathbb{K}$; $c_{0,2} \leftarrow c_{0,1} \oplus r_1$; $c_{2,1} \leftarrow c_{2,0} \ominus r_1$;
return $\langle c_{0,2}, c_{1,1}, c_{2,1} \rangle$

Compositional Probing Security This small difference between NI and SNI has a significant effect on security when used in larger circuits. Indeed, the output positions of a strongly non-interfering gadgets do not depend on any of its input positions: when considered independently from internal positions (in the absence of internal probes), their distribution is uniform; and in the presence of internal probes, their joint distribution is entirely determined by that of the probed internal positions. This is essential in supporting *compositional* reasoning about the probing security of larger algorithms. In particular, this makes algorithms of the form shown in Algorithm 3 (for some gadgets R and G of the appropriate types that work on 2-encodings) easy to prove t-NI if R is $\mathsf{RefreshM}_2$, and illustrates why composition might fail if R is instantiated with $\mathsf{RefreshA}_2$. A key observation to make is that an adversary that observes 2 positions internal to G may learn 2 shares of both $\mathbf{a}$ and $\mathbf{b}$. If R is instantiated with $\mathsf{RefreshA}_2$ (and is thus only 2-probing secure), the 2 shares of $\mathbf{b}$ can be used to infer information about 2 further shares of $\mathbf{a}$, which may give the adversary full knowledge of all 3 shares of $\mathbf{a}$. On the other hand, if R is instantiated with $\mathsf{RefreshM}_2$ (and is thus 2-SNI), the adversary's knowledge of 2 shares of $\mathbf{b}$ does not propagate any further back to $\mathbf{a}$, and the algorithm remains secure.

Alg. 3 An abstract algorithm

```
function Alg2(a)
  b := R(a);
  c := G(a, b);
  return c
```

Broader uses of SNI The notion of strong non-interference, and the masking transformation we define here have already found applications in follow-up work. Belaïd et al. [8] prove using our compositional techniques that their new non-interfering multiplication can be securely combined with the strongly non-interfering one of Rivain and Prouff [31] to build a strongly non-interfering AES S-box with reduced randomness complexity. Similarly, Goudarzi and Rivain [23] use our method to ensure the compositional security of their bitsliced software implementation of AES. Battistelo et al. [6] use and prove t-SNI for their $O(n \cdot \log n)$ mask refreshing gadget, allowing further randomness complexity reductions without loss of probing security. Coron et al. [16] use and prove t-SNI for their efficient parallel algorithms for the evaluation of SBoxes.

Outline The rest of the paper is organized as follows. Section 3 formalizes our two-tier language for masked gadgets and algorithms, the notion of position, and their semantics, as well as the joint distribution of a set of positions. Sections 4, and 5 formalize probing security as t-non-interference, and formally define our new notion of t-strong-non-interference before illustrating it more generally with simple examples. In Section 6, we define a language to describe probing policies, and define a simple type system for enforcing probing policies of algorithms, formalizing and generalizing the simple compositional arguments outlined here. In Section 7, we present an automated method to verify the strong non-interference of arbitrary gadgets at small fixed orders, that follows the approach used above in arguing that $\mathsf{RefreshM}_2$ is 2-SNI, and adapts algorithms by Barthe et al. [4] to reduce the number of position sets to consider. In Section 8, we extend our type system into a masking transformation which automatically builds a masked algorithm from an unprotected program, carefully choosing the proper locations for strongly non-interfering refreshing gadgets. We evaluate on full cryptographic algorithms the performance of the type system, of the resulting transformation, and of the transformed algorithms. Section 9 discusses related work on leakage models, composition for probing security, and other masking transformations. We interleave discussions of interesting leads for future work.

3 Masked algorithms

The formal development of this paper is based on a minimalist 2-tier language.[10] The lower tier models gagdets as sequences of probabilistic and (three-address code) deter-

[10] However, the verification tool supports richer settings to which the theory extends smoothly and our examples are written in a more general language, closer to our implementation, that supports static **for** loops, direct assignments to shares ($\mathbf{a}^i \leftarrow e$), arbitrary expressions on the right-hand side of assignments, and a broader return syntax. For example, Gadget 4 shows generic descriptions of the mask refreshing algorithms from Section 2.

$$
\begin{array}{lll}
\text{algorithm} & P(\mathbf{a}_1, \ldots, \mathbf{a}_n) ::= s; \mathsf{return}\ \mathbf{a} & \\
\text{alg. body} & s ::= \ \mathbf{b} :=_\ell G(\mathbf{a}_1, \ldots, \mathbf{a}_n) & \text{gadget call.} \\
 & \quad \mid s; s & \text{call seq.} \\
 & & \\
\text{gadget} & G(\mathbf{a}_1, \ldots, \mathbf{a}_n) ::= c; \mathsf{return}\ \vec{x} & \\
\text{gadget body} & c ::= \ x \xleftarrow{\$} \mathbb{K} & \text{prob. assign.} \\
 & \quad \mid x \leftarrow e & \text{det. assign.} \\
 & \quad \mid c; c & \text{assign. seq.} \\
 & & \\
\text{expressions} & e ::= \ x, y, \ldots & \text{variable} \\
 & \quad \mid \mathbf{a}^\imath & \imath^{\text{th}}\text{-share of } \mathbf{a} \\
 & \quad \mid x \star y & \text{ring operation}
\end{array}
$$

Fig. 1: Syntax of masked algorithms

ministic assignments, whereas the upper tier models algorithms as sequences of gadget calls (we assume that each gadget call is tagged with its instruction number $\ell \in \mathbb{N}$). The formal definition of the language is given in Figure 1, where we use vector notations ($\vec{x}$, $\ldots$) to denote $(t+1)$-tuples of scalar variables, $\imath$ to denote indices (such that $0 \leq \imath \leq t$) in such a tuple or in encoding variables, and exponents $\cdot^\imath$ to denote the projection of a component out of a $(t+1)$-tuple (for example $\mathbf{a}^\imath$, or $\vec{x}^\imath$). We require gadgets and algorithms to be well-formed, in the following sense. A gadget G is *well-formed* if its body is in SSA form, i.e. its scalar variables appear at most once on the left-hand side of an assignment. An algorithm P is well-formed if all its gadgets are defined and well-formed, and if, in all gadget calls $\mathbf{b} := G(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, variables $\mathbf{b}, \mathbf{a}_1, \ldots, \mathbf{a}_k$ are pairwise disjoint.

We now turn to the semantics of gadgets and algorithms. Crucially, the semantics of gadgets and algorithms is instrumented to keep track of the joint distribution of *all* intermediate values computed during execution. Formally, we assume that scalar and encoding variables take values in $\mathbb{K}$ and $\mathbb{K}^{t+1}$, where $\mathbb{K}$ is the carrier set of a finite ring $(\mathbb{K}, 0, 1, \oplus, \ominus, \odot)$. We let $\mathsf{Val} = \mathbb{K}^{t+1}$ denote the set of encoded values. Furthermore, we let $\mathcal{A}$ denote the set of encoding variables and define the set of global memories as $\mathsf{Mem} = \mathcal{A} \to \mathbb{K}^{t+1}$. Likewise, we let $\mathcal{V}$ denote the set of scalar variables and define the set of local memories as $\mathsf{LMem} = \mathcal{V} \rightharpoonup \mathbb{K}$ and extended local memories as $\mathsf{ELMem} = (\mathbb{N} \times \mathcal{V}) \rightharpoonup \mathbb{K}$. Then, the semantics of a gagdet G is a function $[\![G]\!]$ that takes as input a global memory and returns a distribution over pairs of local memories and values. Likewise, the semantics of an algorithm P is a function $[\![P]\!]$ that takes as input a global memory and returns a distribution over extended local memories and values. The semantics is outlined in Figure 2.

In order to define probing security, we first define a notion of position that corresponds to the intuition illustrated in Section 2. First, we define the set $\mathbb{I} \triangleq \{\mathbf{a}^\imath \mid \mathbf{a} \in \mathcal{A}, 0 \leq \imath \leq t\}$ of input positions (these correspond to shares of encodings used in the gadget or algorithm), the set $\mathbb{O} \triangleq \mathbb{I} \cup \mathcal{V}$ of positions (composed of input positions and scalara variables) and the set $\mathbb{O}^+ \triangleq \mathbb{I} \cup (\mathbb{N} \times \mathcal{V})$ of extended positions (where scalar variables are tagged with a label in $\mathbb{N}$ to differentiate between uses of a variable in

different gadgets). The input positions of a gadget G and of an algorithm P are denoted by $\mathbb{I}_G$ and $\mathbb{I}_P$ respectively and contain exactly those elements of $\mathbb{I}$ that correspond to encoding variables that occur in G or P. Likewise, the set of positions of a gadget G and of an algorithm P are denoted by $\mathbb{O}_G \subseteq \mathbb{O}$ and $\mathbb{O}_P \subseteq \mathbb{O}^+$ respectively and consist of all positions that occur in a gadget G, and all extended positions that occur in an algorithm P.

To capture the joint distribution of a set of positions $\mathcal{O}$ in a gadget G or an algorithm P (with $\mathcal{O} \subseteq \mathbb{O}_G$, resp. $\mathcal{O} \subseteq \mathbb{O}_P$), we take the marginal of the gadget or algorithm's semantics with respect to $\mathcal{O}$. These are denoted by $[\![G]\!]_{\mathcal{O}} : \mathsf{Mem} \to \mathcal{D}(\mathcal{O} \to \mathbb{K})$ and $[\![P]\!]_{\mathcal{O}} : \mathsf{Mem} \to \mathcal{D}(\mathcal{O} \to \mathbb{K})$ respectively.[11]

$$
\begin{array}{rl}
[\![e]\!](m, lm) & : \mathbb{K} \\
& \quad \text{with } m \in \mathsf{Mem} \text{ and } lm \in \mathsf{LMem} \\
[\![x]\!](m, lm) & = lm(x) \\
[\![\mathbf{a}^i]\!](m, lm) & = m(\mathbf{a})^i \\
[\![x \star y]\!](m, lm) & = lm(x) \star lm(y) \\
\hline
[\![c]\!](m, lm) & : \mathcal{D}(\mathsf{Mem} \times \mathsf{LMem}) \\
& \quad \text{with } m \in \mathsf{Mem} \text{ and } lm \in \mathsf{LMem} \\
[\![x \leftarrow e]\!](m, lm) & = \mathsf{munit}(m, lm\{x \leftarrow [\![e]\!](m, lm)\}) \\
[\![x \xleftarrow{\$} \mathbb{K}]\!](m, lm) & = \mathsf{mlet}\ v = \mathcal{U}_{\mathbb{K}}\ \mathsf{in}\ \mathsf{munit}(m, lm\{x \leftarrow v\}) \\
[\![c_1; c_2]\!](m, lm) & = \mathsf{mlet}\ (m_1, lm_1) = [\![c_1]\!](m, lm)\ \mathsf{in}\ [\![c_2]\!](m_1, lm_1) \\
\hline
[\![G]\!](m) & : \mathcal{D}(\mathsf{LMem} \times \mathsf{Val}) \\
& \quad \text{with } m \in \mathsf{Mem} \text{ and} \\
& \quad G(\mathbf{a}_1, \ldots, \mathbf{a}_n) ::= c; \mathsf{return}\ \vec{x} \\
[\![G]\!](m) & = \mathsf{mlet}\ (m_1, lm_1) = [\![c]\!](m, \emptyset)\ \mathsf{in}\ \mathsf{munit}(lm_1, lm_1(\vec{x})) \\
\hline
[\![s]\!](m, elm) & : \mathcal{D}(\mathsf{Mem} \times \mathsf{ELMem}) \\
& \quad \text{with } m \in \mathsf{Mem},\ elm \in \mathsf{ELMem} \text{ and} \\
& \quad G(\mathbf{a}_1, \ldots, \mathbf{a}_n) ::= c; \mathsf{return}\ \vec{x} \\
[\![\mathbf{b} :=_\ell G(\mathbf{c}_1, \ldots, \mathbf{c}_n)]\!](m, elm) & = \mathsf{mlet}\ (lm_1, v) = [\![G]\!](m\{(\mathbf{a}_i)_{1 \le i \le n} \leftarrow (m(\mathbf{c}_i))_{1 \le i \le n}\}) \\
& \quad \mathsf{in}\ \mathsf{munit}(m\{\mathbf{b} \leftarrow v\}, elm \uplus elm_1) \\
& \quad \text{where } elm_1 \text{ is the map that sets only} \\
& \quad elm_1(\ell, v) = lm(v) \text{ for all } v \in \mathsf{dom}(lm) \\
[\![s_1; s_2]\!](m, elm) & = \mathsf{mlet}\ (m_1, elm_1) = [\![s_1]\!](m, elm)\ \mathsf{in}\ [\![s_2]\!](m_1, elm_1) \\
\hline
[\![P]\!](m) & : \mathcal{D}(\mathsf{ELMem} \times \mathsf{Val}) \\
& \quad \text{with } m \in \mathsf{Mem} \text{ and } P(\mathbf{a}_1, \ldots, \mathbf{a}_n ::= s; \mathsf{return}\ \mathbf{b} \\
[\![P]\!](m) & = \mathsf{mlet}\ (m_1, elm_1) = [\![s]\!](m, \emptyset)\ \mathsf{in}\ \mathsf{munit}(elm_1, m_1(\mathbf{b})) \\
\hline
\end{array}
$$

where $m\{x_1, \ldots, x_n \leftarrow v_1, \ldots, v_n\}$ denotes the map m where x_i is updated with v_i for each i in increasing order, and $\uplus$ denotes the disjoint union of partial maps.

Fig. 2: Semantics of gadgets and algorithms

[11] In order to justify that the marginals have the required type, observe that one can refine the type of $[\![G]\!]$ given in Figure 2 to $\mathsf{Mem} \to \mathcal{D}(\mathsf{Val} \times (\mathbb{O}_G \to \mathbb{K}))$. Similarly, one can refine the type of $[\![P]\!]$ to $\mathsf{Mem} \to \mathcal{D}(\mathsf{Val} \times (\mathbb{O}_P \to \mathbb{K}))$.

4 Baseline probing security

We first review the basic notion of probabilistic non-interference and state some of its key properties. As usual, we start by introducing a notion of equivalence on memories.

Definition 1 *Let G be a gadget, and let $\mathcal{I} \subseteq \mathbb{I}_G$. Two memories $m, m' \in \mathsf{Mem}$ are $\mathcal{I}$-equivalent, written $m \sim_{\mathcal{I}} m'$, whenever $m(\mathbf{a})^i = m'(\mathbf{a})^i$ for every $\mathbf{a}^i \in \mathcal{I}$.*

Next, we define probabilistic non-interference.

Definition 2 *Let G be a gadget, and let $\mathcal{I} \subseteq \mathbb{I}_G$ and $\mathcal{O} \subseteq \mathbb{O}_G$. G is $(\mathcal{I}, \mathcal{O})$-non-interfering (or $(\mathcal{I}, \mathcal{O})$-NI), iff $[\![G]\!]_{\mathcal{O}}(m) = [\![G]\!]_{\mathcal{O}}(m')$ for every $m, m' \in \mathsf{Mem}$ s.t. $m \sim_{\mathcal{I}} m'$.*

For every gadget G and every position set $\mathcal{O}$, we define the dependency set of $\mathcal{O}$ as $\mathsf{depset}_G(\mathcal{O}) = \bigcap\{\, \mathcal{I} \mid G \text{ is } (\mathcal{I}, \mathcal{O})\text{-NI} \,\}$; thus, $\mathsf{depset}_G(\mathcal{O})$ is the smallest set $\mathcal{I} \subseteq \mathbb{I}_G$ such that G is $(\mathcal{I}, \mathcal{O})$-NI.

Lemma 1 *Let G be a gadget and $\mathcal{O} \subseteq \mathbb{O}_G$ be a set of positions in G. G is $(\mathsf{depset}_G(\mathcal{O}), \mathcal{O})$-NI.*

Proof (sketch). It suffices to show that there exists $\mathcal{I}$ such that G is $(\mathcal{I}, \mathcal{O})$-NI and that for every $\mathcal{I}$ and $\mathcal{I}'$, if G is $(\mathcal{I}, \mathcal{O})$-NI and $(\mathcal{I}', \mathcal{O})$-NI, then it is also $(\mathcal{I} \cap \mathcal{I}', \mathcal{O})$-NI. □

We conclude this section by providing an alternative definition of non-interference, in the style of simulation-based security.

Lemma 2 *A gadget G is $(\mathcal{I}, \mathcal{O})$-NI iff there exists a simulator $\mathbf{Sim} \in (\mathcal{I} \to \mathbb{K}) \to \mathcal{D}(\mathcal{O} \to \mathbb{K})$ such that for every $m \in \mathsf{Mem}$,*

$$[\![G]\!]_{\mathcal{O}}(m) = \mathbf{Sim}(m_{|\mathcal{I}})$$

where $m_{|\mathcal{I}}$ is the restriction of m to elements in $\mathcal{I}$.

Proof. For the direct implication, define $\mathcal{S}$ as follows: given a function $m \in \mathcal{I} \to \mathbb{K}$, we let $\mathcal{S}$ be $[\![G]\!]_{\mathcal{O}}(m')$, where m' is any memory that extends m. It is immediate to check that $\mathcal{S}$ satisfies the expected property. The reverse direction is immediate.

This observation is useful to connect the information-flow based formulation of probing security introduced below with the simulation-based formulations of probing security often used by cryptographers. Indeed, the dependency set $\mathsf{depset}_G(\mathcal{O})$ can be interpreted as a set of G's input shares that is sufficient to perfectly simulate the joint distribution of positions in $\mathcal{O}$ to an adversary.

Next we define our baseline notion of probing security, which we call t-non-interference, and state some of its basic properties. The notion of t-non-interference is based on the notion of degree of an input set, which we define first. Given an input set $\mathcal{I}$ and an encoding variable $\mathbf{a}$, we define the set $\mathcal{I}_{|\mathbf{a}} \triangleq \mathcal{I} \cap \overline{\mathbf{a}}$ of positions in $\mathcal{I}$ that correspond to shares of $\mathbf{a}$. Further, we define the degree of an input set $\mathcal{I}$ as $\|\mathcal{I}\| \triangleq \max_{\mathbf{a}} |\mathcal{I}_{|\mathbf{a}}|$ (where $|\cdot|$ is the standard notion of cardinality on finite sets). This notion captures the intuition that the adversary should not learn all shares of any single encoding variable, by bounding the information an adversary may learn about any of a gadget's shared inputs through positions probed internally to that gadget.

Definition 3 (Probing security) *A gadget G is* t-non-interfering *(or t-NI) whenever* $\|\mathsf{depset}_G(\mathcal{O})\| \leq |\mathcal{O}|$ *for every* $\mathcal{O} \subseteq \mathbb{O}_G$ *such that* $|\mathcal{O}| \leq t$.

The next lemma establishes that t-NI is already achieved under a weaker cardinality constraint on the dependency set. Variants of Lemma 3 in simulation-based settings appear in [10,8].

Lemma 3 *A gadget G is t-NI iff* $\|\mathsf{depset}_G(\mathcal{O})\| \leq t$ *for every* $\mathcal{O} \subseteq \mathbb{O}_G$ *s.t.* $|\mathcal{O}| \leq t$.

Proof. The direct implication follows from transitivity of $\leq$. The reverse implication proceeds by induction on $k = t - |\mathcal{O}|$. The case $k = 0$ is immediate by definition of t-NI. For the inductive case, let $\mathcal{O}$ be such that $t - |\mathcal{O}| = k + 1$. Let $\mathbf{a}$ be an encoding variable such that $|\mathsf{depset}(\mathcal{O})_{|\mathbf{a}}| = \|\mathsf{depset}(\mathcal{O})\|$. Since $\|\mathsf{depset}(\mathcal{O})\| \leq |\mathcal{O}| \leq t$, there necessarily exists an input position $\mathbf{a}^i$ such that $\mathbf{a}^i \notin \mathsf{depset}(\mathcal{O})$. Let $\mathcal{O}' = \mathcal{O} \cup \{\mathbf{a}^i\}$. We have $\|\mathsf{depset}_G(\mathcal{O}')\| \leq |\mathcal{O}'| = 1 + |\mathcal{O}|$ by induction hypothesis and $\|\mathsf{depset}_G(\mathcal{O}')\| = 1 + \|\mathsf{depset}_G(\mathcal{O})\|$ by construction, therefore we can conclude. □

The notion of t-non-interference extends readily to algorithms. In addition, one can prove that an algorithm is secure iff the gadget obtained by fully inlining the algorithm is secure.

Lemma 4 *A program P is t-NI iff the gadget* $\mathsf{inline}(P)$ *obtained by full inlining is t-NI.*

The lemma sheds some intuition on the definition of t-NI for algorithms. However, we emphasize that verifying fully inlined algorithms is a bad strategy; in fact, previous work indicates that this approach does not scale, and that composition results are needed.

5 Strong Non-Interference

We introduce strong non-interference, a reinforcement of probing security based on a finer analysis of cardinality constraints for dependency sets. Informally, strong non-interference distinguishes between internal and output positions, and requires that the dependency set of a position set $\mathcal{O}$ has degree $\leq k$, i.e. contains at most k shares of each encoding input, where k is the number of *internal* positions in $\mathcal{O}$. Formally, a local variable is an *output position* for G if it appears in the return tuple of G, and an internal position otherwise. Let $\mathcal{O}^{\mathrm{int}}$ (resp. $\mathcal{O}^{\mathrm{ext}}$) denote the subset of internal (resp. output) positions of a set $\mathcal{O}$. Strong t-non-interference requires that the degree of $\mathsf{depset}(\mathcal{O})$ is smaller than $|\mathcal{O}^{\mathrm{int}}|$, rather than $|\mathcal{O}|$. Intuitively, a t-SNI gadget information-theoretically hides dependencies between each of its inputs and its outputs, even in the presence of internal probes. This essential property is what supports compositional reasoning.

Definition 4 (Strong probing security) *A gadget G is* t-strongly non-interfering *(or t-SNI) if* $\|\mathsf{depset}_G(\mathcal{O})\| \leq |\mathcal{O}^{\mathrm{int}}|$ *for every position set $\mathcal{O}$ such that* $|\mathcal{O}| \leq t$.

Fortunately, many gadgets from the literature achieve strong non-interference (see Table 1 and Section 7). First, we note that gadget RefreshM (Gadget 4b) generalized from Ishai, Sahai and Wagner [24] is t-SNI for all t. (A proof sketch for this proposition is given in Appendix C.)

Gadget 4 Mask Refreshing Gadgets

```
0: function RefreshA(a)
1:   c^0 ← a^0
2:   for i = 1 to t do
3:     r ←$ K
4:     c^0 ← c^0 ⊕ r
5:     c^i ← a^i ⊖ r
6:   return c
```

(4a) Addition-Based Mask Refreshing

```
0: function RefreshM(a)
1:   for i = 0 to t do
2:     c^i ← a^i
3:   for i = 0 to t do
4:     for j = i + 1 to t do
5:       r ←$ K
6:       c^i ← c^i ⊕ r
7:       c^j ← c^j ⊖ r
8:   return c
```

(4b) Multiplication-Based Mask Refreshing

Proposition 1 RefreshM *(Gadget 4b) is* t*-SNI.*

On the contrary, the additive refreshing gadget RefreshA (Gadget 4a) achieves NI but fails to achieve SNI. Interestingly, Coron's linear-space variant of Ishai, Sahai and Wagner's multiplication [12, Alg. 6] (Gadget 7) and the MultLin gadget for securely multiplying linearly dependent inputs [15, Alg. 5] (Gadget 8) are both strongly non-interfering. The proof of SNI for Gadget 7 is easy to adapt to the more standard quadratic-space multiplication gadget, since they compute the same intermediate values in different orders.

Proposition 2 *The* SecMult *gadget (Gadget 7) is* t*-SNI.*

Proposition 3 *The* MultLin *gadget (Gadget 8) is* t*-SNI.*

The proofs of Propositions 1, 2 and 3 have been machine-checked using EasyCrypt [5]. We also provide more detailed game-based proof sketches in the full version of this paper.

Strong Non-Interference for Mask Refreshing We now show how choosing a t-SNI refreshing gadget over a t-NI refreshing gadget critically influences the security of algorithms. Concretely, we provide a separating example, which captures the essence of the flaw in the inversion algorithm of Rivain and Prouff [31]. The example considers two algorithms (Algorithm 5) which compute a cube in $\mathsf{GF}(2^8)$ by squaring and multiplying (using, for illustration purposes, some t-NI gadgets Square and Mult for squaring and multiplication). Both algorithms use a refreshing gadget between the two operations, but they differ in which gadget they use: BadCube (Gadget 5a) uses the additive refreshing gagdet RefreshA, which is t-NI but not t-SNI, and Cube (Gadget 5b) uses the RefreshM gadget, which is t-SNI. This simple difference is fundamental for the security of the two algorithms.

Alg. 5 Cubing procedures (with $\mathbb{K} = \mathsf{GF}(2^8)$)

function BadCube($\mathbf{x}$)	**function** Cube($\mathbf{x}$)
$\mathbf{y}_1 := \mathsf{Square}(\mathbf{x})$	$\mathbf{y}_1 := \mathsf{Square}(\mathbf{x})$
$\mathbf{y}_2 := \mathsf{RefreshA}(\mathbf{y}_1)$	$\mathbf{y}_2 := \mathsf{RefreshM}(\mathbf{y}_1)$
$\mathbf{z} := \mathsf{Mult}(\mathbf{x}, \mathbf{y}_2)$	$\mathbf{z} := \mathsf{Mult}(\mathbf{x}, \mathbf{y}_2)$
return $\mathbf{z}$	**return** $\mathbf{z}$
(5a) Insecure Cubing	**(5b)** Secure Cubing

Lemma 5 ([15]) BadCube *is not* t-NI *for any* $t \geq 2$. Cube *is* t-NI *for all* t.

Coron et al. [15] exhibit proofs for both statements. In Appendix C, we give a compact proof of t-NI for Cube that does not exhaustively consider all $(t+1)$-tuples of positions in Cube. The key argument is that RefreshM being t-SNI essentially renders useless any information on $\mathbf{y}_2$ the adversary may have learned from observing positions in Mult: those do not add any shares of $\mathbf{y}_1$ to the dependency set we compute for RefreshM, and therefore do not influence the shares of $\mathbf{x}$ that appear in the final dependency set for Cube. On the other hand, using a simple t-NI mask refreshing gadget (such as RefreshA) in its place breaks the proof by allowing us to deduce only that each position in the multiplication may depend on 2 shares of $\mathbf{x}$.

In Section 6, we show how the proof of Lemma 5 can be improved and extended into a compositional proof for the (repaired) inversion algorithm of Rivain and Prouff [31], and, in fact, outlines a general mehodology for proving algorithms t-NI or t-SNI.

A Generic Composition Result Before formalizing and automating this proof process to obtain precise probing security proofs for large circuits, we now give a coarse but simple composition result that illustrates the generality of SNI. Informally, an algorithm is t-NI if all its gadgets verify t-NI and every non-linear usage of an encoding variable is guarded by t-SNI refreshing gadgets. In addition, it shows that processing all inputs, or the output of a t-NI algorithm with a t-SNI gadget (here RefreshM) suffices to make the algorithm t-SNI.

Proposition 4 *An algorithm P is t-NI provided all its gadgets are t-NI, and all encoding variables are used at most once as argument of a gadget call other than* RefreshM. *Moreover P is t-SNI if it is t-NI and one of the following holds:*

- *its return expression is $\mathbf{b}$ and its last instruction is of the form $\mathbf{b} := \mathsf{RefreshM}(\mathbf{a})$;*
- *its sequence of encoding parameters is $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, its i^{th} instruction is $\mathbf{b} :=_i \mathsf{RefreshM}(\mathbf{a}_i)$ for $1 \leq i \leq n$, and $\mathbf{a}_i$ is not used anywhere else in the algorithm.*

6 Enforcing probing policies

We first define an expressive assertion language for specifying sets of position sets, and then introduce probing policies, which yield a convenient formalism for defining a large class of information flow policies with cardinality constraints.

Definition 5 (Probing policy)

1. *A* probing assertion *is a pair (Γ, ϕ), where Γ is a map from encoding variables to expressions in the theory of finite sets, and ϕ is a cardinality constraint. Each probing assertion (Γ, ϕ) defines a set of subsets of positions for a fixed algorithm P, denoted by $[\![(\Gamma, \phi)]\!]$. (The syntax and semantics of set expressions and cardinality constraints is explained below.)*
2. *A* probing policy *is a pair of assertions*

$$(\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$$

where $(\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$ is the post-assertion *and $(\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}})$ is the* pre-assertion*.*
3. *Algorithm P satisfies the policy $(\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$, written $P \models (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$, if for every position set $\mathcal{O} \in [\![(\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})]\!]$, P is $(\mathcal{I}, \mathcal{O})$-NI for some input position set $\mathcal{I} \in [\![(\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}})]\!]$.*

The syntax of set expressions and cardinality constraints is given by the following grammar:

$$\begin{array}{ll} \text{(set expr.)} & S := X \mid \emptyset \mid S \cup S \\ \text{(arith. expr.)} & l := |S| \mid |O^\ell| \mid t \mid l + l \\ \text{(cardinality constr.)} & \phi := l \leq l \mid \phi \wedge \phi \end{array}$$

The syntax distinguishes between variables X that are drawn from a set $\mathcal{X}$ of names–that we will use to represent sets of shares of an encoding variable, and variables O, annotated with a label ℓ, that are drawn from a disjoint set Ω of names–that we will use to represent sets of internal positions probed in the gadget used at instruction ℓ.

Remark 1 *Our syntax for set expressions and constraints is a fragment of the (decidable) theory of finite sets with cardinality constraints. It would be possible to include other set-theoretical operations, as in [32,3]. However, we have found our core fragment sufficient for our purposes.*

The semantics of assertions is defined using the notion of valuation. A valuation μ is a mapping from names in $\mathcal{X}$ and Ω to finite sets, such that $\forall\, X \in \mathcal{X}.\ \mu(X) \subseteq \{0, \ldots, t\}$ and $\forall\, O^\ell \in \Omega.\ \mu(O^\ell) \subseteq \mathbb{O}_{G_\ell}$, where G_ℓ is the gadget called at instruction ℓ. Every valuation μ defines, for every set expression S, a set of share indices $\mu(S) \subseteq \{0, \ldots, t\}$ and for every arithmetic expression l an interpretation $\mu(l) \in \mathbb{N}$, using the intended interpetation of symbols (i.e. $\cup$ is interpreted as set union, $+$ is interpreted as addition, $\ldots$).

Definition 6 (Interpretation of assertions)

1. *μ satisfies a cardinality constraint ϕ, written $\mu \models \phi$, if $\mu(l_1) \leq \mu(l_2)$ for every conjunct $l_1 \leq l_2$ of ϕ.*
2. *The interpretation of Γ under μ is the set*

$$\mu(\Gamma) = \bigcup_{\mathbf{a}} \{\mathbf{a}^\imath \mid \imath \in \mu(\Gamma(\mathbf{a}))\} \cup \bigcup_{O} \mu(O)$$

3. *The interpretation of (Γ, ϕ) is the set*

$$[\![(\Gamma, \phi)]\!] = \{\mu(\Gamma) \mid \mu \models \phi\}$$

We now turn to the definition of the type system.

Definition 7 *Algorithm* $P(\mathbf{a}_1, \ldots, \mathbf{a}_n) ::= s;$ return $\mathbf{r}$ *has type* $(\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$ *if the judgment* $\vdash s : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$ *is derivable using the typing rules from Figure 3. We denote this fact* $\vdash P : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$.

We briefly comment on the rules. Rule (SEQ) is used for typing the sequential composition of gadget calls and is as expected. The remaining rules are used for interpreting the non-interference properties of gadgets. We now detail them.

$$\frac{\vdash \mathbf{b} := G(\mathbf{a}_1, \ldots, \mathbf{a}_n) : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma, \phi) \qquad \vdash c : (\Gamma, \phi) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})}{\vdash \mathbf{b} := G(\mathbf{a}_1, \ldots, \mathbf{a}_n); c : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})} \ (\textsc{seq})$$

$$\frac{\begin{array}{c} G \text{ is } t\text{-}\mathsf{NI} \qquad \phi_{\mathsf{out}} \Rightarrow |\Gamma_{\mathsf{out}}(\mathbf{b})| + |O^\ell| \leq t \\ \Gamma_{\mathsf{in}} := \Gamma_{\mathsf{out}}\{\mathbf{b}, (\mathbf{a}_k)_{1 \leq k \leq n} \leftarrow \emptyset, (\Gamma_{\mathsf{out}}(\mathbf{a}_k) \cup X_k^\ell)_{1 \leq k \leq n}\} \end{array}}{\vdash \mathbf{b} :=_\ell G(\mathbf{a}_1, \ldots, \mathbf{a}_n) : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{out}} \wedge (\bigwedge_{1 \leq k \leq n} |X_k^\ell| \leq |\Gamma_{\mathsf{out}}(\mathbf{b})| + |O^\ell|)) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})} \ (\mathsf{NI}\text{-}\textsc{gadget})$$

$$\frac{\begin{array}{c} G \text{ is } t\text{-}\mathsf{SNI} \qquad \phi_{\mathsf{out}} \Rightarrow |\Gamma_{\mathsf{out}}(\mathbf{b})| + |O^\ell| \leq t \\ \Gamma_{\mathsf{in}} := \Gamma_{\mathsf{out}}\{\mathbf{b}, (\mathbf{a}_k)_{1 \leq k \leq n} \leftarrow \emptyset, (\Gamma_{\mathsf{out}}(\mathbf{a}_k) \cup X_k^\ell)_{1 \leq k \leq n}\} \end{array}}{\vdash \mathbf{b} :=_\ell G(\mathbf{a}_1, \ldots, \mathbf{a}_n) : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{out}} \wedge (\bigwedge_{1 \leq k \leq n} |X_k^\ell| \leq |O^\ell|)) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})} \ (\mathsf{SNI}\text{-}\textsc{gadget})$$

$$\frac{\begin{array}{c} G \text{ is affine} \qquad \Gamma_{\mathsf{in}} := \Gamma_{\mathsf{out}}\{\mathbf{b}, (\mathbf{a}_k)_{1 \leq k \leq n} \leftarrow \emptyset, (\Gamma_{\mathsf{out}}(\mathbf{a}_k) \cup \Gamma_{\mathsf{out}}(\mathbf{b}) \cup X^\ell)_{1 \leq k \leq n}\} \end{array}}{\vdash \mathbf{b} :=_\ell G(\mathbf{a}_1, \ldots, \mathbf{a}_n) : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{out}} \wedge |X^\ell| \leq |O^\ell|) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})} \ (\textsc{affine})$$

where $\Gamma\{\forall\, k.\ v_k \leftarrow \forall\, k.\ e_k\}$ stands for the map Γ where each v_k is updated to map to e_k and all other indices are left untouched.

Fig. 3: Typing rules

Rule (SNI-GADGET) is used for typing calls to a SNI-gadget with an arbitrary post-assertion and a pre-assertion in which the mapping Γ_{out} is updated to reflect the dependencies created by the call, and the constraint is strenghtened with the cardinality constraint imposed by strong non-interference. The rule has a side condition $|O^\ell| + |\Gamma_{\mathsf{out}}(\mathbf{b})| \leq t$ ensuring that the total number of positions whose dependency set by G we are considering is bounded by t, where O^ℓ is the name of the subset of positions that are observed in the current gadget (called at line ℓ), and $\Gamma_{\mathsf{out}}(\mathbf{b})$ is the set of shares of $\mathbf{b}$ the adversary has information about from positions probed in gadgets that use $\mathbf{b}$ later on in the algorithm. This side condition is verified under the condition ϕ_{out}. Note that the variables X_k^ℓ are fresh, and annotated with the label ℓ that identifies the current instruction, and an indice k that identifies the argument. Rule (NI-GADGET) is similar but deals with NI-gadgets, and therefore extends Γ_{in} with correspondingly weaker constraints on the X_k^ℓ.

We now turn to the rule for affine gadgets. Informally, we say that a gadget is affine if it manipulates its input encodings share by share; this includes standard implementations of ring addition, for example, but also of many other functions that are linear in $\mathbb{K}$ (for

example, multiplication by a constant–or public–scalar, or shifts in the representation when addition is bitwise). Formally, we say that a gadget G with parameters $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ is affine iff there exists a family of procedures $f_0, \ldots, f_t$ such that G is an inlining of

$$x_0 \leftarrow f_0(\mathbf{a}_1^0, \ldots, \mathbf{a}_n^0); \ldots; x_t \leftarrow f_t(\mathbf{a}_1^t, \ldots, \mathbf{a}_n^t);$$
$$\textbf{return } \langle x_0, \ldots, x_t \rangle;$$

Thus, one can define a mapping $\eta : \mathcal{O}_G \to \{0, \ldots, t\}$ such that for every position $\pi \in \mathcal{O}_G$, $\eta(\pi) = \imath$ if π occurs in the computation of the $\imath^{\text{th}}$ share (i.e. in the computation of $f_\imath(a_1^\imath, \ldots, a_n^\imath)$). The fine-grained information about dependencies given by this notion of affinity is often critical to proving the probing security of algorithms. Therefore, it is important to capture affinity in our type system. Let $\mathcal{O} = \mathcal{O}^{\text{int}} \uplus \mathcal{O}^{\text{ext}}$ be a position set, split between internal and output positions. The affine property ensures that the joint distribution of $\mathcal{O}$ depends only on input positions in $\eta(\mathcal{O}^{\text{int}} \cup \mathcal{O}^{\text{ext}})$, and furthermore that $|\eta(\mathcal{O}^{\text{int}} \cup \mathcal{O}^{\text{ext}})| = |\eta(\mathcal{O}^{\text{int}})| + |\eta(\mathcal{O}^{\text{ext}})| = |\eta(\mathcal{O}^{\text{int}})| + |\mathcal{O}^{\text{ext}}|$. Rule (AFFINE) interprets this affine property into our type system, using $\Gamma_{\mathsf{out}}(\mathbf{b})$ and a fresh O^ℓ for $\mathcal{O}^{\text{ext}}$ and $\mathcal{O}^{\text{int}}$, respectively, and encoding $\eta(O^\ell)$ into an abstract existential set X^ℓ. The condition $|X^\ell| \leq |O^\ell|$ precisely captures the fact that $|\eta(O)| \leq |O|$ for all O.

Definition 8 (Typing of an algorithm) *Let P be an algorithm with a definition of the form $P(\mathbf{a}_1, \ldots, \mathbf{a}_n) ::= s;$ return $\mathbf{r}$.*

P is well-typed for SNI, written $\vdash_{\mathsf{NI}} P$, whenever there exist Γ_{in}, ϕ_{in} such that $\vdash P : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\emptyset, \sum_{1 \leq \ell \leq |P|} |O^\ell| \leq t)$ and, for each $i \in \{1, \ldots, n\}$, $\phi_{\mathsf{in}} \Rightarrow |\Gamma_{\mathsf{in}}(\mathbf{a}_i)| \leq t$.

P is well-typed for SSNI, written $\vdash_{\mathsf{SNI}} P$, whenever there exist Γ_{in}, ϕ_{in} such that $\vdash P : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow ([\mathbf{r} \leftarrow O], |O| + \sum_{1 \leq \ell \leq |P|} |O^\ell| \leq t)$ and, for each $i \in \{1, \ldots, n\}$, we have $\phi_{\mathsf{in}} \Rightarrow |\Gamma_{\mathsf{in}}(\mathbf{a}_i)| \leq \sum_{1 \leq \ell \leq |P|} |O^\ell|$ (where $[v \leftarrow x]$ is the map that associates x to v and is everywhere else undefined).

When typing for NI, we start from the empty map for Γ_{out} and simply consider any output position observed as if they were internal. However, the same cannot be done when typing for SNI since we need to distinguish clearly between internal positions in one of the O^ℓ, used to type the gadget at instruction ℓ, and output positions in O, initially used as the set of position of the algorithm's return encoding.

Proposition 5 (Soundness of the Type System)
If $\vdash s : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$ then also $\models P : (\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}}) \Longleftarrow (\Gamma_{\mathsf{out}}, \phi_{\mathsf{out}})$
If $\vdash_{\mathsf{NI}} P$ then P is t-NI
If $\vdash_{\mathsf{SNI}} P$ then P is t-SNI

An Example: Rivain and Prouff's inversion algorithm We now illustrate the type system by describing a typing derivation on Rivain and Prouff's algorithm for computing inversion in $\mathsf{GF}(2^8)$ [31,15]. An algorithm implementing this operation securely is shown in Figure 4, with some information relevant to its typing derivation. We recall that the function $x \mapsto x^{2^n}$ is linear (for any n) in binary fields and rely on affine gadgets pow2, pow4, and pow16 to compute the corresponding functionalities.

We present the typing derivation in the slightly unusual form of a table, in Figure 4, which shows the code of the inversion algorithm along with the values of Γ_{in} and ϕ_{in} (ϕ_{in} shows only the part of the constraint that is *added* at that program point, not the entire constraint) at each program point. By the sequence rule, these serve as Γ_{out} and ϕ_{out} for the immediately preceding program point. The table also shows the side conditions checked during the application of gadget rules where relevant. It is easier to understand the type-checking process by reading the table from the bottom up.

As per the definition of well-typedness for SNI, we start from a state where the output position set O is associated to the algorithm's return encoding $\mathbf{r}_5$, and where the constraint contains only the global constraint that the whole position set $O \cup \bigcup_\ell O^\ell$ is of cardinality bounded by t. When treating line 9, we know that SecMult is t-SNI and try to apply rule (SNI-GADGET). We check that the number of positions observed in this instance of SecMult is bounded by t (which trivially follows from the global constraint), and construct the new value of $(\Gamma_{\mathsf{in}}, \phi_{\mathsf{in}})$ following the rule: since neither of the call's input encodings are used below, new sets X_1^9 and X_2^9 are associated to the call's inputs and the SNI constraints are added to ϕ_{in}. Applying the rules further until the top of the program is reached, and performing the appropriate set unions in Γ when an encoding variable is used more than once, we observe that the resulting pre-assertion is such that $|\Gamma_{\mathsf{in}}(\mathbf{a})| \leq |O^1| + |O^2| + |O^3| + |O^9| \leq \sum_\ell O^\ell$, and therefore proves that this inversion algorithm is t-SNI.

Γ_{in}	ϕ_{in}	Instructions	Side conditions
		function invert($\mathbf{a}$)	
$\mathbf{a} : X_2^3 \cup X_2^9 \cup X_1^2 \cup X^1$	$\lvert X^1\rvert \leq \lvert O^1\rvert$	$\mathbf{z}_1 :=_1$ pow2($\mathbf{a}$)	
$\mathbf{a} : X_2^3; \mathbf{z}_1 : X_2^9 \cup X_1^2$	$\lvert X_1^2\rvert \leq \lvert O^2\rvert$	$\mathbf{z}_2 :=_2$ Refresh($\mathbf{z}_1$)	$\lvert X_1^3\rvert + \lvert O^2\rvert \leq t$
$\mathbf{a} : X_2^3; \mathbf{z}_1 : X_2^9;\ \mathbf{z}_2 : X_1^3$	$\lvert X_k^3\rvert \leq \lvert O^3\rvert$	$\mathbf{r}_1 :=_3$ SecMult($\mathbf{z}_2, \mathbf{a}$)	$\lvert X_1^6 \cup X_2^8 \cup X_1^5 \cup X^4\rvert + \lvert O^3\rvert \leq t$
$\mathbf{z}_1 : X_2^9; \mathbf{r}_1 : X_1^6 \cup X_2^8 \cup X_1^5 \cup X^4$	$\lvert X^4\rvert \leq \lvert O^4\rvert$	$\mathbf{w}_1 :=_4$ pow4($\mathbf{r}_1$)	
$\mathbf{z}_1 : X_2^9; \mathbf{r}_1 : X_1^6;\ \mathbf{w}_1 : X_2^8 \cup X_1^5$	$\lvert X_1^5\rvert \leq \lvert O^5\rvert$	$\mathbf{w}_2 :=_5$ Refresh($\mathbf{w}_1$)	$\lvert X_2^6\rvert + \lvert O^5\rvert \leq t$
$\mathbf{z}_1 : X_2^9; \mathbf{r}_1 : X_1^6;\ \mathbf{w}_1 : X_2^8; \mathbf{w}_2 : X_2^6$	$\lvert X_k^6\rvert \leq \lvert O^6\rvert$	$\mathbf{r}_2 :=_6$ SecMult($\mathbf{r}_1, \mathbf{w}_2$)	$\lvert X_1^8 \cup X^7\rvert + \lvert O^6\rvert \leq t$
$\mathbf{z}_1 : X_2^9; \mathbf{w}_1 : X_2^8; \mathbf{r}_2 : X_1^8 \cup X^7$	$\lvert X^7\rvert \leq \lvert O^7\rvert$	$\mathbf{r}_3 :=_7$ pow16($\mathbf{r}_2$)	
$\mathbf{z}_1 : X_2^9; \mathbf{w}_1 : X_2^8; \mathbf{r}_3 : X_1^8$	$\lvert X_k^8\rvert \leq \lvert O^8\rvert$	$\mathbf{r}_4 :=_8$ SecMult($\mathbf{r}_3, \mathbf{w}_1$)	$\lvert X_1^9\rvert + \lvert O^8\rvert \leq t$
$\mathbf{z}_1 : X_2^9; \mathbf{r}_4 : X_1^9$	$\lvert X_k^9\rvert \leq \lvert O^9\rvert$	$\mathbf{r}_5 :=_9$ SecMult($\mathbf{r}_4, \mathbf{z}_1$)	$\lvert O\rvert + \lvert O^9\rvert \leq t$
$\mathbf{r}_5 : O$	$\lvert O\rvert + \sum_{1 \leq \ell \leq 9} \lvert O^\ell\rvert \leq t$	**return** $\mathbf{r}_5$	

Fig. 4: $\mathbf{a}^{-1}$ in $\mathsf{GF}(2^8)$

Finally, one can remark that the instances of SecMult at line 6 and 8 do not in fact need to be t-SNI. As pointed out by Belaïd et al. [8], using a t-NI multiplication gadget at these program points is sufficient to construct a type derivation for SNI.

7 SNI Checker for Gadgets

We present an automated method for proving that gadgets (or small algorithms, by inlining) are t-SNI at small fixed orders (up to $t = 6$ for ring multiplication). We then give some experimental results.

Verification algorithm We adapt to t-SNI the algorithmic contributions of Barthe et al. [4] that support the automated verification, on small to medium gadgets and for small orders, of Ishai, Sahai and Wagner's circuit privacy property [24], which is similar to our t-NI. Their work builds on two observations: first, every probabilistic program P taking input x and performing a (statically) bounded number (say q) of uniform samplings over $\mathbb{K}$ is equivalent, in the sense below, to composing a deterministic program $P^\dagger$ taking inputs x and r with random sampling over $\mathbb{K}^q$. Formally, for every x,

$$[\![P]\!](x) = \mathsf{mlet}\ r = \mathcal{U}_{\mathbb{K}^q}\ \mathsf{in}\ [\![P^\dagger]\!]_{\mathcal{O}}(x, r)$$

Second, P satisfies $(\mathcal{I}, \mathcal{O})$-NI iff there exists a function f such that for every x_1, x_2 and r, such that $x_1 \sim_{\mathcal{I}} x_2$

$$[\![P^\dagger]\!]_{\mathcal{O}}(x_1, r) = [\![P^\dagger]\!]_{\mathcal{O}}(x_2, f(x_2, r))$$

and moreover $f(x, \cdot)$ is a bijection for every x. The latter equality can be easily verified for all x and r using standard tools, therefore the key to proving non-interference is to exhibit a suitable function f. Their algorithm proceeds by incrementally defining bijections $f_1, \ldots, f_n$ satisfying the two conditions above until evenutally $[\![P^\dagger]\!]_{\mathcal{O}}(x, f_n(x, r))$ can be rewritten into an expression that does not depend syntactically on secrets.

However, even with efficient algorithms to prove that a program P is $(\mathcal{I}, \mathcal{O})$-NI for some position set $\mathcal{O}$, proving that P is t-NI remains a complex task: indeed this involves proving $(\mathcal{I}, \mathcal{O})$-NI for all $\mathcal{O}$ with $|\mathcal{O}| \leq t$. Simply enumerating all possible position sets quickly becomes untractable as P and t grow. Therefore, [4] uses the following fact: if P is $(\mathcal{I}, \mathcal{O}')$-NI then it is also $(\mathcal{I}, \mathcal{O})$-NI for all $\mathcal{O} \subseteq \mathcal{O}'$. Hence, checking that P is $(\mathcal{I}, \mathcal{O}')$-NI for some large set $\mathcal{O}'$ is sufficient to prove that P is $(\mathcal{I}, \mathcal{O})$-NI for every $\mathcal{O} \subseteq \mathcal{O}'$, and this using only one proof of non-interference. In particular, they exhibit algorithms that rely on the explicit construction of the bijection f_n to efficiently extend the set $\mathcal{O}$ from which it was constructed into a potentially much larger set $\mathcal{O}'$ for which that bijection still proves $(\mathcal{I}, \mathcal{O}')$-NI. Further, they also exhibit algorithms that rely on such extensions to prove the existence of $\mathcal{I}$ such that $(\mathcal{I}, \mathcal{O})$-NI for all position sets $\mathcal{O}$ much more efficiently than by considering all position sets individually.

We adapt their algorithms by changing the core bijection-finding algorithm in two ways: i. rather than being applied to a modified program that includes the initial uniform sampling of secret encodings, our core algorithm works directly on the gadget description (this is necessary to ensure that we prove t-SNI instead of alternative security notions); and ii. our search for a bijection stops when $[\![P^\dagger]\!]_{\mathcal{O}}(x, f_n(x, r))$ can be simplified into an expression that syntactically depends on at most d shares of the secret (for the desired bound d on $\|\mathcal{I}\|$, that is $d = |\mathcal{O}^{\mathrm{int}}|$ for SNI), rather than stopping when all syntactic dependencies on the secret input have been removed. We note that replacing the bound d from the second point with $d = t$ yields a verification algorithm for t-NI (by Lemma 3). Our full algorithm is given in App. B.

Evaluation We evaluate the performance of our SNI verifier on some medium and small gadgets: SecMult**,** Coron's linear-memory ring multiplication algorithm [12, Alg. 6]; MultLin**,** Coron et al.'s algorithm for the computation of functionalities of the form $\mathbf{x} \odot g(\mathbf{x})$ for some linear g [15, Alg. 5]; Add**,** the standard affine gadget for the addition

Gadget	Order 1		Order 2		Order 3		Order 4		Order 5		Order 6	
	1-SNI	Time	2-SNI	Time	3-SNI	Time	4-SNI	Time	5-SNI	Time	6-SNI	Time
SecMult	✓	0.07s	✓	0.08s	✓	0.09s	✓	0.86s	✓	36.40s	✓	37min
MultLin	✓	0.07s	✓	0.08s	✓	0.15s	✓	1.19s	✓	54.13s	✓	48min
RefreshA	✓	0.07s	✗	0.08s	✗	0.08s	–	–	–	–	–	–
RefreshIter^2	✓	0.08s	✓	0.08s	✓	0.08s	✓	0.08s	✓	0.13s	✗	.20s
RefreshIter^3	–	–	✓	0.09s	✓	0.08s	✓	0.09s	✓	0.14s	✓	.54s
WeakMult	✓	0.07s	✗	0.07s	✗	0.09s	–	–	–	–	–	–

Table 1: Experimental Results for the SNI Verifier

of two encodings; RefreshA**,** the weakly secure mask refreshing algorithm from Rivain and Prouff [31]; RefreshIter^k**,** the iterated additive refresh proposed by Coron [12, Alg. 4] for supporting more efficient composition in his full model (we make explicit the number of iterations k); WeakMult**,** the generic reduced-randomness multiplication algorithm proposed by Belaïd et al. [8]. Table 1 sums up our findings and some verification statistics.

8 Masking Transformation

As a proof of concept, we implement our type system for a comfortable subset of C that includes basic operators, static for loops, table lookups at public indices, and mutable secret state, and extended with libraries that implement core gadgets for some choices of $\mathbb{K}$. Moreover, we define a source-to-source *certifying masking transformation*, which takes an unprotected program and returns a masked algorithm accepted by our type system, selectively inserting refreshing gadgets as required for typing to succeed. We note that the transformation itself need not be trusted, since its result is the final program on which typing is performed.

Furthermore, the choice of C as a supporting language is for convenience, since many of the algorithms we consider have reference implementations written in C. In particular, we do not claim that compiling and executing the C programs produced by our masking transformation will automatically yield secure executables: our verification results are on *algorithms* described in the C language rather than on C programs in general. Making use of these verification results in practice still requires to take into account details not taken into account in the probing model. Although an important problem, this is out of the scope of this paper and a research area on its own: for example Balasch et al. [2] consider some of the issues involved in securely implementing probing secure algorithms.

8.1 Implementation

We now give an overview of the different passes performed by our masking transformation. The input programs use explicit typing annotations to distinguish public variables (for example, public inputs, or public loop indices) from sensitive or secret variables

that must be encoded. We call public type any type outside of those used for denoting variables that must be encoded.

Parsing and Pre-Typing This pass parses C code into our internal representation, checks that the program is within the supported subset of C, performs C type-checking and checks that variables marked as sensitive (variables given type $\mathbb{K}$) are never implicitly cast to public types. Implicit casts from public types to $\mathbb{K}$ (when compatible, for example, when casting a public `uint8_t` to a protected variable in $\mathsf{GF}(2^8)$) are replaced with public encoding gadgets (that set one share to the public value and all other shares to 0).

Gadget Selection and Generic Optimizations This pass heuristically selects optimal gadgets depending on their usage. For example, multiplication of a secret by a public value can be computed by an affine gadget that multiplies each share of the secret, whereas the multiplication of two secrets must be performed using the SecMult gadget. Further efforts in formally proving precise types for specialized core gadgets may also improve this optimization step. Since the encoding replaces scalar-typed variables (passed by value) with array-typed variables (passed by reference), it is also necessary to slightly transform the program to ensure the correctness of the resulting program. In addition, we also transform the input program into a form that more closely follows the abstract language from Figure 1, which makes it easier to type-check.

Type Inference and Refresh Insertion This is the core of our transformation. We implement a type inference algorithm for the type system of Section 6. The algorithm simplifies policies on the fly, supports inferred types on sub-algorithms as gadget-invocation types, and fails when the simplified policy is inconsistent. Failure arises exactly when a refreshing operation is needed. At the cost of tracking some more information and reinforcing the typing constraint on sub-algorithms, we use this observation to automatically insert Refresh gadgets where required. When type inference fails, the variable whose masks need to be refreshed is duplicated and one of its uses is replaced with the refreshed duplicate. To avoid having to re-type the entire program after insertion of a refresh gadget, our transformation keeps track of typing information for each program point already traversed and simply rewinds the typing to the program point immediately after the modification.

Code Generation Finally, once all necessary mask refreshing operations have been inserted and the program has been type-checked, we produce a masked C program. This transformation is almost a one-to-one mapping from the instructions in the type-checked programs to calls to a library of verified core gadgets or to newly defined gadgets. Some cleanup is performed on loops to clarify the final code whenever possible, and to remove initialization code on normalized gadgets. Interestingly, our transformation produces a (set of) C files that is parameterized by the masking order t. Producing executable versions of that algorithm at a particular order, for example to evaluate its performance, is as easy as defining a pre-processor macro at compile-time.

8.2 Practical Evaluation

To test the effectiveness of our transformation, we apply it to different algorithms, generating equivalent masked algorithms at various orders. We apply our transformation to the following programs: **AES ($\odot$),** a full computation (10 rounds including key schedule) of AES-128 masked using the multiplication gadget, and implemented in $\mathsf{GF}(2^8)$; **AES ($x \odot g(x)$),** a full computation (10 rounds including key schedule) of AES-128 masked using Coron et al.'s gadget for computing $x \odot g(x)$, and implemented in $\mathsf{GF}(2^8)$; **Keccak,** a full computation (24 rounds) of Keccak-f[1600], implemented in $\mathsf{GF}(2^{64})$; **Simon,** a block of Simon(128,128), implemented in $\mathsf{GF}(2^{64})$; **Speck,** a block of Speck(128,128), implemented in $\mathsf{GF}(2)^{64}$, and using one of the following modular addition algorithms; **AddLin,** Coron, Großschädl and Vadnala's algorithm [14] for the computation of modular addition on boolean-masked variables (in $\mathsf{GF}(2)^{64}$); **AddLog,** Coron et al.'s improved algorithm [13] for the computation of modular addition on boolean-masked variables (in $\mathsf{GF}(2)^{64}$). We first discuss the performance of our verifier and the verification results before discussing the practical significance, in terms of time, memory and randomness complexity of our masking transformation. Finally, we discuss examples on which our tool implementation could be improved.

Verification Performance and Results Table 2 shows resource usage statistics for generating the masked algorithms (at any order) from unprotected implementations of each algorithm. The table shows the number of mask refreshing operations inserted in the program[12], the compilation time, and the memory consumption. For Keccak, we show two separate sets of figures: the first, marked "no refresh", is produced by running our algorithm transformer on a bare implementation of the algorithm; the second, marked "refresh in χ", is produced by running our tool on an annotated implementation, where a mask refreshing operation is manually inserted in the χ function and the tool used for verification only. We discuss discrepancies between the numbers on these two lines in Section 9, and consider the "refresh in χ" set of statistics in all discussions until then. We first note the significant improvements these results represent over the state of the art in formal verification for probing security. Indeed, our closest competitor [4] report the verification of all 10 rounds of AES (including key schedule) at order 1 in 10 minutes, and could not verify all 10 rounds for higher orders. In contrast, our tool verifies the probing security of Rivain and Prouff's algorithm [31] as fixed by Coron et al. [15] *at all orders* in less than a second.[13] Further, we note that the masked algorithms our transformation produce for modular addition are the first such algorithms known to be t-probing secure using only $t + 1$ shares. Indeed, the original proofs [14,13] rely on the ISW framework and make use of $2t + 1$ shares to obtain t-probing security. We further note that Coron, Großschädl and Vadnala's algorithm [14] does not require the insertion of mask refreshing operations, and is thus t-probing secure with $t + 1$ shares as it was originally described. Finally, we note that, to the best of our knowledge, the

[12] Note that the number of mask refreshing operations executed during an execution of the algorithm may be much greater, since the sub-procedure in which the insertion occurs may be called multiple times.

[13] This excludes the once-and-forall cost of proving the security of core gadgets.

results obtained on Keccak, Simon and Speck constitute the first generic higher-order masking schemes for these algorithms.

Algorithm	# Refresh	Time	Mem.
AES ($\odot$)	2 per round	0.09s	4MB
AES ($x \odot g(x)$)	0	0.05s	4MB
AddLin	0	0.01s	4MB
AddLog	$\log_2(k) - 1$	0.01s	4MB
Keccak (no refresh)	1 per round	$\sim$20min	23GB
Keccak (refresh in χ)	0	18.20s	456MB
Simon	67 per round	0.38s	15MB
Speck (AddLin)	61 per round	0.35s	38MB
Speck (AddLog)	66 per round	0.21s	8MB

Table 2: Resource usage during masking and verification

Performance of Masked Algorithms Table 3 reports the time taken to execute the resulting programs 10,000 times at various orders on an Intel(R) Xeon(R) CPU E5-2667 0 @ 2.90GHz with 64GB of memory running Linux (Fedora). As an additional test to assess the performance of the generated algorithms at very high orders, we masked an AES computation at order 100: computation took $\sim$0.11 seconds per block. For AES and Speck, the figures shown in the "unmasked" column are execution times for the input to our transformation: a table-based implementation of AES or an implementation of Speck that uses machine arithmetic, rather than Coron, Großschädl and Vadnala's algorithm would be much faster, but cannot be masked directly using our transformation. Although these observations do highlight the cost of security, we note that using RefreshA when masking the AES SBox does not incur a significant timing gain for any of the masking orders we tested ($t \leq 20$). However, the randomness cost is greatly reduced, which may be significant in hardware or embedded software settings. Further research in reducing the randomness cost of SNI mask refreshing, or of other gadgets, serves to make security less costly [8,1,6]. We also confirm the 15% timing improvements reported by Coron et al. [15] when implementing the AES SBox using their gadget for computing $\mathbf{x} \odot g(\mathbf{x})$.

Algorithm	unmasked	Order 1	Order 2	Order 3	Order 5	Order 10	Order 15	Order 20
AES ($\odot$)	0.078s	2.697s	3.326s	4.516s	8.161s	21.318s	38.007s	59.567s
AES ($x \odot g(x)$)	0.078s	2.278s	3.209s	4.368s	7.707s	17.875s	32.552s	50.588s
Keccak	0.238s	1.572s	3.057s	5.801s	13.505s	42.764s	92.476s	156.050s
Simon	0.053s	0.279s	0.526s	0.873s	1.782s	6.136s	11.551s	20.140s
Speck (AddLin)	0.022s	4.361s	10.281s	20.053s	47.389s	231.423s	357.153s	603.261s
Speck (AddLog)	0.022s	0.529s	1.231s	2.258s	5.621	19.991s	42.032	72.358s

Table 3: Time taken by 10,000 executions of each program at various masking orders

We now look more closely at statistics for the modular addition algorithms AddLin and AddLog and their effects on the performance of masked algorithms for Speck. We first note that proving AddLog t-NI requires the addition of a mask refreshing gadget, whereas AddLin does not. Despite this additional cost, however, AddLog is better than AddLin when word size k grows, since it saves $k - \log(k)$ multiplications and replaces them with a single mask refreshing operation. These performance gains on modular addition become overwhelming when seen in the context of a masked algorithm for Speck, which computes one 64-bit modular addition per round. It would be interesting to consider using our transformer to produce masked algorithms for other efficient circuits for modular addition [26] and measure their performance impact in terms of randomness, time and memory when masked.

9 Discussions and Related Work

Here, we further discuss the relation between the definitions and results reported here and existing and future work in theoretical and practical cryptography. Our discussions focus mainly on: i. adversary and leakage models; ii. compositional security notions; iii. theoretical and practical masking transformations; and iv. limitations of our definitions and tools.

Adversary and Leakage Models for Masking We have considered security in the probing model of Ishai, Sahai and Wagner [24], which is particularly well-suited to automated analysis due to its tight relation to probabilistic non-interference. In particular, our notion of t-NI is equivalent to the notions of *t-probing security* and *perfect t-probing security* used by Carlet et al. [10] and others [31,15].

Despite its broad usage in the literature, the practical relevance of the probing model is not immediately obvious: in practice, side-channel adversaries observe *leakage traces*, which contain noisy information about all intermediate computations, rather than precise information about some. This threat model is much more closely captured by the *noisy leakage model*, first introduced by Chari et al. [11] and extended by Prouff and Rivain [30]. The noisy leakage model is much more complex and makes security proofs on masked algorithms significantly more involved, and much harder to verify.

Duc, Dziembowski and Faust [17] show that proving probing security allows one to estimate the practical (noisy leakage) security of a masked algorithm. While Duc, Faust and Standaert [18] empirically show that some of the factors of Duc *et al.*'s bound [17] are likely proof artefacts, the remainder of the bound, and in particular a factor that includes the size of the circuit, seems to be tight. Intuitively, Duc *et al.* [18] essentially show that the probing security order gives an indication of the smallest order moment of the distribution over leakage traces that contains information about the secret, whereas the size of the circuit the adversary can probe is an indicator of how easy it is to evaluate higher-order moments.

Composition, and Region and Stateful Probing This observation makes clear the importance of also considering more powerful probing adversaries that may place t probes in each of some (pre-determined) *regions* of an algorithm (the *t-region probing model*).

For example, each core gadget (field operations and mask refreshing operation) could be marked off as a separate region (as in [17]). More recently, and in work contemporary with that presented here, Andrychowicz, Dziembowski and Faust [1] consider a more general notion of region whose size must be linear in the security parameter (and masking order), and exhibit a mask refreshing gadget that is linear in size and fulfills, in the probing model, the *reconstructibility* and *re-randomization* properties from Faust et al. [21]. We now discuss the implications of reconstructibility and re-randomization, and their relation to our notion of SNI, based on the similarity of Prop. 4 with Ishai *et al.*'s remark on "Re-randomized outputs" [24], before discussing the applicability of SNI to security in the region and stateful probing models [24].

Intuitively, a gadget is t-reconstructible whenever any t of its positions can be simulated using only its (shared) inputs and outputs, and a gadget is re-randomizing whenever its output encoding is uniform and t-wise independent even if its input encoding is completely known. Our SNI notions combines both considerations. Formulating it in similar terms, a gadget is t-SNI whenever any t of its positions can be simulated using only its (shared) inputs, and if its output encoding is uniform and $(t-d)$-wise independent even if d shares of each of its inputs are known (for all d such that $0 \leq d < t$). Expressed in this way, it is clear that SNI is slightly weaker than "reconstructible and re-randomizable" in the probing model. This allows us to automatically verify that a gadget is SNI for some fixed t, whereas reconstructibility and re-randomization are more complex. In addition, the ability to combine the use of SNI and weaker (NI or affine) gadgets in a fine-grained way allows us to more precisely verify the security of large algorithms in models where the adversary can place t probes in the entire algorithm. We leave a formal investigation of the relation between SNI and "reconstructibility and re-randomization" as future work.

Based on reconstructibility and re-randomization, Faust et al. [21,1] prove elegant and powerful composition results that in fact apply in the more powerful region probing and stateful probing models [24], where the adversary may (adaptively) place t probes in each region (or in each subsequent iteration) of the algorithm. It is worth noting that our SNI notion also enables composition in these two models: indeed, it is easy to see that any two $2t$-SNI algorithms (our regions) can be composed securely when the adversary can place t probes in each of them. Further, our composition techniques also support elegant constructions that support compositional security proofs in the region and stateful probing models without doubling the number of shares computations are carried out on (instead, simply doubling the number of shares at region boundaries). We give details of these *robust composition* results in Appendix D. Depending on the size of regions that are considered, these robust composition results may bring significant performance gains in terms of randomness and time complexity.

Finally, our notion of SNI and the automated verification techniques presented allow the efficient, precise and automated verification of t-SNI inside each region, an issue which is not addressed by the works of Faust et al. [21,1].

Existing Masking Transformations Ishai, Sahai and Wagner [24] and others [17,1] also propose simple masking transformations that turn unprotected algorithms (or boolean or arithmetic circuits) into protected masked algorithms. Ishai, Sahai and Wagner [24] forgo the use of mask refreshing gadgets by doubling the number of shares on which masked

computations occur–with a quadratic impact on performance and randomness complexity. Faust et al. [17,1] rely on making sure that all gadgets used in the masked algorithm are reconstructible and re-randomizing. This guarantees security in a stronger probing model, but incurs an even greater loss of performance. By contrast, our transformation attempts to decide whether a mask refreshing operation is required to ensure security in the probing model, and our core contributions (the notion of SNI and the type-checker) do support composition in stronger probing models, whilst still allowing the proofs of security within regions to be handled precisely.

Coron [12] proposes schemes for masking lookups at secret or sensitive indices in public tables. We have not investigated whether or not the proposed algorithms are SNI or simply NI, and whether or not establishing these properties can be done by adapting our type-system or if it should be done in a different way (either as a direct proof or using the checker from Section 7). We note in passing that part of the result by Coron [12], namely that using $\mathsf{RefreshIter}_{2t}^{2t+1}$ between each query to the masked S-box supports security in the stateful probing model is subsumed and improved by the robust composition results described in the full version.

The security analysis of masking schemes in the t-probing model is connected to techniques from multi-party computation, exploited in parallel lines of research by threshold implementations [28,9]. In particular, higher-order threshold implementations are exposed to similar security issues due to composition, although they offer additional protection against practical considerations not captured in standard probing models, namely *glitches*. We believe that the results discussed here are in fact applicable to the compositional security analysis of threshold implementations but leave a formal investigation of these links as future work.

Refining SNI We now discuss some limitations of our current implementation, and leads for future theoretical work that may yield significant practical improvements.

Alg. 6 Semi Public Modular Addition in $\mathsf{GF}(2)^k$

function AddPub($\mathbf{x}, y$)
$\mathbf{w} := \mathbf{x} \odot y$
$\mathbf{a} := \mathbf{x} \oplus y$
$\mathbf{u} := \mathbf{w} \ll 1$
for $i = 2$ **to** $k - 1$ **do**
$\mathbf{a}' := \mathsf{RefreshM}(\mathbf{a})$
$\mathbf{ua} := \mathbf{u} \odot \mathbf{a}'$
$\mathbf{u} := \mathbf{ua} \oplus \mathbf{w}$
$\mathbf{u} := \mathbf{u} \ll 1$
$\mathbf{z} := \mathbf{a} \oplus \mathbf{u}$
return z

(6a) Masked algorithm produced by our tool

function AddPub($\mathbf{x}, y$)
$\mathbf{w} := \mathbf{x} \odot y$
$\mathbf{a} := \mathbf{x} \oplus y$
$\mathbf{w} := \mathsf{RefreshM}(\mathbf{w})$
$\mathbf{u} := \mathbf{w} \ll 1$
for $i = 2$ **to** $k - 1$ **do**
$\mathbf{ua} := \mathbf{u} \odot \mathbf{a}$
$\mathbf{u} := \mathbf{ua} \oplus \mathbf{w}$
$\mathbf{u} := \mathbf{u} \ll 1$
$\mathbf{z} := \mathbf{a} \oplus \mathbf{u}$
return z

(6b) Masked algorithm produced by hand

The first point we wish to discuss is the case of Keccak, for which algorithm transformation is prohibitively expensive. This issue is due to our handling of static for loops:

indeed, our tool unrolls them to perform type-checking and rolls them back up afterwards if possible (otherwise leaving them unrolled in the final algorithm). For smaller algorithms, this is not a problem, but unrolling all 24 rounds of Keccak-f, along with all the loops internal to each iteration, yields a very large program that is then backtracked over each time a mask refreshing operation is inserted. Refining our non-interference notions to multi-output gadgets and algorithms would allow us to significantly improve our tool's handling of loops and high-level composition, whilst gaining a better understanding of probing security in such scenarios. This improved understanding may in turn help inform the design of primitives that are easier to protect against higher-order probing.

Second, we discuss our greedy policy for the insertion of mask refreshing algorithms. In our experiments, we consider a version of the linear-time modular addition algorithm [14] whose second argument is a public (non-shared) value (for example, a round counter, as in Speck). We show its code, as produced by our masking transformer, in Gadget 6a, and display a hand-masked variant in Gadget 6b, slightly abusing notations by denoting simple gadgets with the symbol typically used for their unprotected versions. Notice that the variable $\mathbf{w}$ is used once per loop iteration, and that our tool refreshes each of them, while it is sufficient to mask only the first one. Improving our gadget selection algorithm to detect and implement this optimization—and others—would be an interesting avenue for future work, that could help improve our understanding of the effect on security of compiler optimizations.

Acknowledgements The work presented here was supported by projects S2013/ICE-2731 N-GREENS Software-CM, ANR-10-SEGI-015 PRINCE and ANR-14-CE28-0015 BRUTUS, and ONR Grants N000141210914 and N000141512750, as well as FP7 Marie Curie Actions-COFUND 291803.

References

1. Marcin Andrychowicz, Stefan Dziembowski, and Sebastian Faust. Circuit compilers with $O(1/\log(n))$ leakage rate. In *EUROCRYPT 2016*, LNCS, pages 586–615. Springer, Heidelberg, 2016.
2. Josep Balasch, Benedikt Gierlichs, Vincent Grosso, Oscar Reparaz, and François-Xavier Standaert. On the cost of lazy engineering for masked software implementations. In *Proceedings of the Smart Card Research and Advanced Application Conference (CARDIS)*, volume 8968 of *LNCS*, pages 64–81. Springer, Heidelberg, November 2014.
3. Kshitij Bansal, Andrew Reynolds, Clark Barrett, and Cesare Tinelli. A new decision procedure for finite sets and cardinality constraints in SMT. In *Proceedings of the 8th International Joint Conference on Automated Reasoning (IJCAR)*, volume 9706 of *LNCS*, pages 82–98, June 2016.
4. Gilles Barthe, Sonia Belaïd, François Dupressoir, Pierre-Alain Fouque, Benjamin Grégoire, and Pierre-Yves Strub. Verified proofs of higher-order masking. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 457–485. Springer, Heidelberg, April 2015.
5. Gilles Barthe, François Dupressoir, Benjamin Grégoire, César Kunz, Benedikt Schmidt, and Pierre-Yves Strub. EasyCrypt: A tutorial. In *Foundations of Security Analysis and Design VII - FOSAD 2012/2013 Tutorial Lectures*, pages 146–166, 2013.

6. Alberto Battistello, Jean-Sébastien Coron, Emmanuel Prouff, and Rina Zeitoun. Horizontal side-channel attacks and countermeasures on the ISW masking scheme. In *CHES 2016*, LNCS, pages 23–29. Springer, Heidelberg, 2016.
7. Ali Galip Bayrak, Francesco Regazzoni, David Novo, and Paolo Ienne. Sleuth: Automated verification of software power analysis countermeasures. In Guido Bertoni and Jean-Sébastien Coron, editors, *CHES 2013*, volume 8086 of *LNCS*, pages 293–310. Springer, Heidelberg, August 2013.
8. Sonia Belaïd, Fabrice Benhamouda, Alain Passelègue, Emmanuel Prouff, Adrian Thillard, and Damien Vergnaud. Randomness complexity of private circuits for multiplication. In *EUROCRYPT 2016*, LNCS, pages 616–648. Springer, Heidelberg, 2016.
9. Begül Bilgin, Benedikt Gierlichs, Svetla Nikova, Ventzislav Nikov, and Vincent Rijmen. Higher-order threshold implementations. In Palash Sarkar and Tetsu Iwata, editors, *ASIACRYPT 2014, Part II*, volume 8874 of *LNCS*, pages 326–343. Springer, Heidelberg, December 2014.
10. Claude Carlet, Emmanuel Prouff, Matthieu Rivain, and Thomas Roche. Algebraic decomposition for probing security. In Rosario Gennaro and Matthew J. B. Robshaw, editors, *CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 742–763. Springer, Heidelberg, August 2015.
11. Suresh Chari, Charanjit S. Jutla, Josyula R. Rao, and Pankaj Rohatgi. Towards sound approaches to counteract power-analysis attacks. In Michael J. Wiener, editor, *CRYPTO'99*, volume 1666 of *LNCS*, pages 398–412. Springer, Heidelberg, August 1999.
12. Jean-Sébastien Coron. Higher order masking of look-up tables. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 441–458. Springer, Heidelberg, May 2014.
13. Jean-Sébastien Coron, Johann Großschädl, Mehdi Tibouchi, and Praveen Kumar Vadnala. Conversion from arithmetic to boolean masking with logarithmic complexity. In Gregor Leander, editor, *FSE 2015*, volume 9054 of *LNCS*, pages 130–149. Springer, Heidelberg, March 2015.
14. Jean-Sébastien Coron, Johann Großschädl, and Praveen Kumar Vadnala. Secure conversion between boolean and arithmetic masking of any order. In Lejla Batina and Matthew Robshaw, editors, *CHES 2014*, volume 8731 of *LNCS*, pages 188–205. Springer, Heidelberg, September 2014.
15. Jean-Sébastien Coron, Emmanuel Prouff, Matthieu Rivain, and Thomas Roche. Higher-order side channel security and mask refreshing. In Shiho Moriai, editor, *FSE 2013*, volume 8424 of *LNCS*, pages 410–424. Springer, Heidelberg, March 2014.
16. Jean-Sébastien Coron, Aurélien Greuet, Emmanuel Prouff, and Rina Zeitoun. Faster evaluation of sboxes via common shares. In *CHES 2016*, LNCS, pages 498–514. Springer, Heidelberg, 2016.
17. Alexandre Duc, Stefan Dziembowski, and Sebastian Faust. Unifying leakage models: From probing attacks to noisy leakage. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 423–440. Springer, Heidelberg, May 2014.
18. Alexandre Duc, Sebastian Faust, and François-Xavier Standaert. Making masking security proofs concrete - or how to evaluate the security of any leaking device. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 401–429. Springer, Heidelberg, April 2015.
19. Hassan Eldib and Chao Wang. Synthesis of masking countermeasures against side channel attacks. In *Proceedings of the 26th International Conference on Computer Aided Verification.*, pages 114–130, 2014.
20. Hassan Eldib, Chao Wang, and Patrick Schaumont. SMT-based verification of software countermeasures against side-channel attacks. In *Proceedings of the 20th International*

Conference on Tools and Algorithms for the Construction and Analysis of Systems, pages 62–77, 2014.
21. Sebastian Faust, Tal Rabin, Leonid Reyzin, Eran Tromer, and Vinod Vaikuntanathan. Protecting circuits from leakage: the computationally-bounded and noisy cases. In Henri Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 135–156. Springer, Heidelberg, May 2010.
22. Louis Goubin and Jacques Patarin. DES and differential power analysis (the "duplication" method). In Çetin Kaya Koç and Christof Paar, editors, *CHES'99*, volume 1717 of *LNCS*, pages 158–172. Springer, Heidelberg, August 1999.
23. Dahmun Goudarzi and Matthieu Rivain. How fast can higher-order masking be in software? Cryptology ePrint Archive, Report 2016/264, 2016. `http://eprint.iacr.org/`.
24. Yuval Ishai, Amit Sahai, and David Wagner. Private circuits: Securing hardware against probing attacks. In Dan Boneh, editor, *CRYPTO 2003*, volume 2729 of *LNCS*, pages 463–481. Springer, Heidelberg, August 2003.
25. Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael J. Wiener, editor, *CRYPTO'99*, volume 1666 of *LNCS*, pages 388–397. Springer, Heidelberg, August 1999.
26. Thomas Walker Lynch. Binary adders, 1996.
27. Andrew Moss, Elisabeth Oswald, Dan Page, and Michael Tunstall. Compiler assisted masking. In Emmanuel Prouff and Patrick Schaumont, editors, *CHES 2012*, volume 7428 of *LNCS*, pages 58–75. Springer, Heidelberg, September 2012.
28. Svetla Nikova, Vincent Rijmen, and Martin Schläffer. Secure hardware implementation of nonlinear functions in the presence of glitches. *Journal of Cryptology*, 24(2):292–321, April 2011.
29. Martin Pettai and Peeter Laud. Automatic proofs of privacy of secure multi-party computation protocols against active adversaries. In Cédric Fournet, Michael W. Hicks, and Luca Viganò, editors, *IEEE 28th Computer Security Foundations Symposium*, pages 75–89. IEEE Computer Society, 2015.
30. Emmanuel Prouff and Matthieu Rivain. Masking against side-channel attacks: A formal security proof. In Thomas Johansson and Phong Q. Nguyen, editors, *EUROCRYPT 2013*, volume 7881 of *LNCS*, pages 142–159. Springer, Heidelberg, May 2013.
31. Matthieu Rivain and Emmanuel Prouff. Provably secure higher-order masking of AES. In Stefan Mangard and François-Xavier Standaert, editors, *CHES 2010*, volume 6225 of *LNCS*, pages 413–427. Springer, Heidelberg, August 2010.
32. Calogero G. Zarba. Combining sets with cardinals. *Journal of Automated Reasoning*, 34(1):1–29, 2005.
33. Steve Zdancewic, Lantian Zheng, Nathaniel Nystrom, and Andrew C. Myers. Untrusted hosts and confidentiality: Secure program partitioning. In Keith Marzullo and M. Satyanarayanan, editors, *Proceedings of the 18th ACM Symposium on Operating System Principles*, pages 1–14. ACM, 2001.

A Code for Gadgets Listed in the Paper

Gadget 7 Masked multiplication [12]

```
0:  function SecMult(a, b)
1:    for i = 0 to t do
2:      c^i ← a^i ⊙ b^i                                   /* line 2^i */
3:    for i = 0 to t do
4:      for j = i + 1 to t do
5:        r ←$ K                                          /* line 5^{i,j} */
6:        c^i ← c^i ⊖ r                                   /* line 6^{i,j} */
7:        t ← a^i ⊙ b^j                                   /* line 7^{i,j} */
8:        r ← r ⊕ t                                       /* line 8^{i,j} */
9:        t ← a^j ⊙ b^i                                   /* line 9^{i,j} */
10:       r ← r ⊕ t                                       /* line 10^{i,j} */
11:       c^j ← c^j ⊕ r                                   /* line 11^{i,j} */
12:   return c
```

Gadget 8 Masked multiplication between linearly dependent inputs with linear function g [15]

```
0:  function MultLin(a)
1:    for i = 0 to t do
2:      c^i ← a^i ⊙ g(a^i)                                /* line 2^i */
3:    for i = 0 to t do
4:      for j = i + 1 to t do
5:        r ←$ K                                          /* line 5^{i,j} */
6:        r' ←$ K                                         /* line 6^{i,j} */
7:        c^i ← c^i ⊖ r                                   /* line 7^{i,j} */
8:        t ← a^i ⊗ g(r') ⊕ r                             /* line 8^{i,j} */
9:        t ← t ⊕ (r' ⊗ g(a^i))                           /* line 9^{i,j} */
10:       t ← t ⊕ (a^i ⊗ g(a^j ⊖ r')                      /* line 10^{i,j} */
11:       t ← t ⊕ ((a^j ⊖ r') ⊗ g(a^i))                   /* line 11^{i,j} */
12:       c^j ← c^j ⊕ t                                   /* line 12^{i,j} */
13:   return c
```

B SNI Verification Algorithm

We now give a full description of our algorithm for the verification of SNI properties. Apart from notation changes, the algorithms are only slightly adapted from [4]. In particular, the proof of correctness for Algorithm 11 can be easily generalized from that given in [4].

In the following, we consider gadgets whose encoded inputs are all secret and denoted as a set of encodings $\{\mathbf{x}^i\}_i$. Given a core gadget, the set of all possible positions in that gadget can be represented as a set E of $\mathbb{K}$ expressions over the ${\mathbf{x_i}}^k$ and some *random variables* in $\mathcal{R}$ (corresponding to the random samplings occurring in the gadget). All algorithms in this subsection consider gadgets represented in this form. Given such a set E, we denote with $\mathsf{var}(E)$ the set of all variables that appear in E, and with E_{int} and E_{out} the subsets that correspond to internal and output positions, respectively.

Following Barthe et al. [4], our algorithm relies, at its core, on exhibiting an isomorphism between the distribution of each acceptable set of positions and some distribution that is *syntactically* non-interfering. In this case, we consider all sets of positions that are composed of t_i internal positions and t_o output positions (for any t_i and t_o such that $t_i + t_o \leq t$) and find a bijection with some distribution that depends on at most t_i shares of each of the gadget's inputs. Algorithm 9 takes an integer d and a set E of expressions over the ${\mathbf{x_i}}^k$ and random variables in $\mathcal{R}$ and finds, when successful, a sequence of substitutions $\vec{h}$ that construct a bijection between the distribution described by E and a distribution that syntactically depends on at most d shares of each one of the $\mathbf{x}_i$.

Algorithm 9 Given $\mathcal{O}$, d, find $\mathcal{I}$ such that $\|\mathcal{I}\| \leq d$ for $(\mathcal{I}, \mathcal{O})$-$\mathsf{NI}$

function $\mathsf{NI}^d_{\{\mathbf{x}^i\}_i}(E)$
  if $\forall i, |\mathsf{var}(E) \cap \mathbf{x}^i| \leq d$ **then**
    return DEGREE(d)
  $(E', e, r) \leftarrow \mathsf{choose}(\{(E', e, r) \mid e \text{ is invertible in } r \wedge r \in \mathcal{R} \wedge E = E'[e/r]\})$
  if $(E', e, r) \neq \bot$ **then**
    return OPT(e, r) : $\mathsf{NI}^d_{\{\mathbf{x}^i\}_i}(E')$
  return $\bot$

The fact that Algorithm 9 produces a witness sequence of substitutions can be leveraged, as in [4], to extend the position set E on which the corresponding bijection still proves the desired property. Algorithms 10a and 10b leverage this, and are only lightly adapted (for notation) from Barthe et al. [4]. An important thing to note is that $\mathsf{extend}^d_{\{\mathbf{x}^i\}_i}(X, E, \vec{h})$ always returns a set X' of expressions such that $X \subseteq X' \subseteq E$ and such that X' depends (by $\vec{h}$) on at most d shares of each one of the $\mathbf{x_i}$.

Finally, we adapt the second space-covering algorithm proposed by Barthe et al. [4] to the verification of the t-SNI property, by taking care to call NI^d with the value of d that corresponds to the number of internal positions being considered. Algorithm 11 shows the adapted algorithm for the core verification task, and Algorithms 12a and 12b show the instances used to verify that a gadget is t-SNI and t-NI, respectively. Here again, Algorithm 11 is only lightly adapted from [4] and their correctness proof, which establishes that whatever property is checked by NI^d is true for all sets in $\bigcup_j \mathcal{P}_{d_j}(E_j)$, also applies to the new property being verified. It is then easy to prove that the calls to check^d made by $\mathsf{check}^{t\text{-}\mathsf{SNI}}$ and $\mathsf{check}^{t\text{-}\mathsf{NI}}$ are sufficient to establish t-SNI and t-NI respectively (relying on Lemma 3 for t-NI).

Algorithm 10 Auxiliary algorithms for SNI verification

function $\mathsf{recheck}^{d}_{\{\mathbf{x}^i\}_i}(E, \vec{h})$
  if $\vec{h} = [\textsc{Degree}(d)]$ **then**
    return $\forall i, |\mathsf{var}(E) \cap \mathbf{x}^i| \leq d$
  if $\vec{h} = \textsc{Opt}(e, r) : \vec{h'}$ **then**
    $(E') \leftarrow \mathsf{choose}(\{E' \mid E = E'[e/r]\})$
    if $E' \neq \bot$ **then**
      return $\mathsf{recheck}^{d}_{\{\mathbf{x}^i\}_i}(E', \vec{h}')$
    else
      return false

(10a) Rechecking a derivation

function $\mathsf{extend}^{d}_{\{\mathbf{x}^i\}_i}(X, E, \vec{h})$
  $e \leftarrow \mathsf{choose}(E)$
  if $e \neq \bot$ **then**
    if $\mathsf{recheck}^{d}_{\{\mathbf{x}^i\}_i}(e, \vec{h})$ **then**
      return $\mathsf{extend}^{d}_{\{\mathbf{x}^i\}_i}((X, e), E \setminus \{e\}, \vec{h})$
    else return $\mathsf{extend}^{d}_{\{\mathbf{x}^i\}_i}(X, E \setminus \{e\}, \vec{h})$
  else return X

(10b) Extending the Position Set

Algorithm 11 Core Algorithm for $(\mathcal{I}, \mathcal{O})$-NI Verification

function $\mathsf{check}^{d}_{\{\mathbf{x}^i\}_i}(\{(d_j, E_j)\}_j)$
  if $\forall j,\ d_j \leq |E_j|$ **then**
    $Y_j \leftarrow \mathsf{choose}(\mathcal{P}_{d_j}(E_j))$
    $\vec{h} \leftarrow \mathsf{NI}^{d}_{\{\mathbf{x}^i\}_i}(\bigcup_j Y_j)$
    if $\vec{h} = \bot$ **then** raise CannotProve $(\bigcup_j Y_j)$
    $\widehat{Y_j} \leftarrow \mathsf{extend}_{\{\mathbf{x}^i\}_i}(Y_j, E_j \setminus Y_j, \vec{h})$
    $\mathsf{check}^{d}_{\{\mathbf{x}^i\}_i}(\{(d_j, E_j \setminus \widehat{Y_j})\}_j)$
    for $j; 0 < i_j < d_j$ **do**
      $\mathsf{check}^{d}_{\{\mathbf{x}^i\}_i}(\{(i_j, \widehat{Y_j}), (d_j - i_j, E_j \setminus \widehat{Y_j})\})$
  return true

Algorithm 12 Algorithms for t-SNI and t-NI verification

function $\mathsf{check}^{t\text{-}\mathsf{SNI}}_{\{\mathbf{x}^i\}_i}(E)$
  for $i; 0 \leq i \leq t$ **do**
    $\mathsf{check}^{i}_{\{\mathbf{x}^i\}_i}(\{(i, E_{\mathsf{int}}), (t - i, E_{\mathsf{out}})\})$
  return true

(12a) t-SNI Verification

function $\mathsf{check}^{t\text{-}\mathsf{NI}}_{\{\mathbf{x}^i\}_i}(E)$
  $\mathsf{check}^{t}_{\{\mathbf{x}^i\}_i}(\{(t, E)\})$
  return true

(12b) t-NI Verification

We also note that Algorithm 11 can be used in other ways, not shown in Algorithm 12, that would enforce other, perhaps more complex, probing policies. We do not claim that all probing policies could be easily verified in this manner (and in fact, the affine policy could not), and leave as future work an investigation of efficient (or semi-efficient) algorithms–in this style–for the verification of arbitrary probing.

C Proofs

In this Appendix, we give proofs for lemmas, theorems, and propositions whose proof is not given in the paper, and detail some of the proof sketches.

Proof (sketch for Proposition 1). Leveraging the equivalence between simulation and non-interference (Lemma 2), we prove t-SNI by constructing a simulator that uses at most $|\mathcal{O}^{\text{int}}|$ shares of the gadget's input to perfectly simulate the joint distribution of any position set $\mathcal{O}$ such that $|\mathcal{O}| \leq t$. The constructed simulator is very similar to those previously used in proofs of t-NI.

Let $\mathcal{O}$ be a set of positions such that $|\mathcal{O}| \leq t$, and let $d_1 = |\mathcal{O}^{\text{int}}|$ and $d_2 = |\mathcal{O}^{\text{ext}}|$. Note that $d_1 + d_2 \leq t$. Our goals are: i. to find an input set $\mathcal{I}$ such that $\|\mathcal{I}\| \leq d_1$, ii. to construct a perfect simulator that uses only input shares $\mathbf{a}^i \in \mathcal{I}$.

First, we identify which variables are internal (and therefore will be considered in $\mathcal{O}^{\text{int}}$) and which are outputs (in $\mathcal{O}^{\text{ext}}$). Internals are the $\mathbf{a}^i$, the $r_{i,j}$ (the value of r sampled at iteration i, j), and the $c_{i,j}$ (resp. $c_{j,i}$) which correspond to the value of the variable c_i (resp. c_j) at iteration i, j of the second loop. Outputs are the final values of $\mathbf{c}^i$ (i.e. $c_{i,t}$). Then, we define $\mathcal{I}$ as follows: for each position among $\mathbf{a}^i$, $r_{i,j}$ and $c_{i,j}$ (with $j < t$) we add share $\mathbf{a}^i$ to $\mathcal{I}$. It is clear that $\mathcal{I}$ contains at most d_1 positions, since each internal position adds at most one position to $\mathcal{I}$. We now construct the simulator. For clarity, observe that the RefreshM algorithm can be represented using the following matrix, observing that $c_{i,j}$ is the partial sum of the first $j + 2$ elements of line i.

$$\begin{pmatrix} a_0 & 0 & r_{0,1} & r_{0,2} & \cdots & r_{0,t} \\ a_1 & \ominus\, r_{0,1} & 0 & r_{1,2} & \cdots & r_{1,t} \\ a_2 & \ominus r_{0,2} & \ominus r_{1,2} & 0 & \cdots & r_{2,t} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_d & \ominus\, r_{0,t} & \ominus\, r_{1,t} & \ominus r_{2,t} & \cdots & 0 \end{pmatrix}.$$

For each $\imath_0$ such that $\mathbf{a}^{\imath_0} \in \mathcal{I}$ (that is, for each line $\imath_0$ of the matrix that contains at least one observed internal value), $\mathbf{a}^{\imath_0}$ is provided to the simulator (by definition of $\mathcal{I}$). Thus, the simulator can sample all $r_{\imath_0,j}$ and compute all partial sums $c_{\imath_0,j}$ and the $\imath_0{}^{\text{th}}$ output normally. At this point, all values (all internals and all outputs) on lines indexed by an $\imath$ with $\mathbf{a}^{\imath} \in \mathcal{I}$ follow the same distribution as they would in the real computation and are therefore perfectly simulated.

It remains to simulate output shares $\mathbf{c}^{\jmath}$ when $\mathbf{a}^{\jmath} \notin \mathcal{I}$. Remark that simulating the $\imath^{\text{th}}$ line as above also necessarily fixes the value of all random variables appearing in the $\imath^{\text{th}}$ column. After internal positions are simulated, at most d_1 lines of the matrix are fully filled. Therefore, each line $\jmath$ with $\mathbf{a}^{\jmath} \notin \mathcal{I}$ contains at least $t - d_1 \geq d_2$ holes

corresponding to random values that have not yet been fixed. For each of the output position made on one such line $\jmath$, we can therefore pick a different $r_{\jmath,k}$ that we choose so $\mathbf{c}^\jmath$ can be simulated by a freshly sampled uniform value.

Proof (sketch for Lemma 5). [15] exhibit proofs for both statements. We now sketch a proof of t-NI for Cube that does not exhaustively consider all $(t+1)$-tuples of positions in Cube, emphasizing the critical use of strong non-interference for the refreshing gadget. Recall that our goal is to upper-bound $\|\mathsf{depset}_{\mathsf{Cube}}(\mathcal{O})\|$ for all $\mathcal{O} \subseteq \mathbb{O}_{\mathsf{Cube}}$ such that $|\mathcal{O}| \leq t$. Given such a set, we first partition it as $\mathcal{O} \triangleq \mathcal{O}_{\mathsf{M}} \uplus \mathcal{O}_{\mathsf{R}} \uplus \mathcal{O}_{\mathsf{S}}$ following gadget boundaries (recall that positions in algorithms include the label of the gadget invocation they occur in). First, we consider the dependency set $\mathcal{I}_{\mathsf{M}} \triangleq \mathsf{depset}_{\mathsf{Mult}}(\mathcal{O}_{\mathsf{M}})$ of $\mathcal{O}_{\mathsf{M}}$ by Mult. We know $|\mathcal{O}_{\mathsf{M}}| \leq |\mathcal{O}| \leq t$, and by t-NI of Mult, we deduce that $\|\mathcal{I}_{\mathsf{M}}\| \leq |\mathcal{O}_{\mathsf{M}}|$. Considering now the invocation of RefreshM, we must establish cardinality properties of the dependency set $\mathcal{I}_{\mathsf{R}} \triangleq \mathsf{depset}_{\mathsf{RefreshM}}(\mathcal{O}_{\mathsf{R}} \cup \mathcal{I}_{\mathsf{M}|\mathbf{y}_2})$ of those direct *internal* positions observed by the adversary in RefreshM, *jointly* with those *output* positions she may have learned information about through positions probed in later parts of the circuit (here, in Mult). From previous inequalities, we know that $|\mathcal{I}_{\mathsf{M}|\mathbf{y}_2}| \leq \|\mathcal{I}_{\mathsf{M}}\| \leq |\mathcal{O}_{\mathsf{M}}|$, and thus we have $|\mathcal{O}_{\mathsf{R}} \cup \mathcal{I}_{\mathsf{M}|\mathbf{y}_2}| \leq t$. By t-SNI of RefreshM, we thus have $\|\mathcal{I}_{\mathsf{R}}\| \leq |\mathcal{O}_{\mathsf{R}}|$ (since positions in $\mathcal{I}_{\mathsf{M}|\mathbf{y}_2}$ are external to RefreshM). Finally, we consider the dependency set $\mathcal{I}_{\mathsf{S}} \triangleq \mathsf{depset}_{\mathsf{Square}}(\mathcal{O}_{\mathsf{S}} \cup \mathcal{I}_{\mathsf{R}|\mathbf{y}_1})$ of direct internal positions observed by the adversary in Square, jointly with those output positions she may have learned information about through positions probed in later parts of the circuit (here in RefreshM and Mult, propagated through the single use of $\mathbf{y}_1$ in the invocation of RefreshM). Since we have $|\mathcal{O}_{\mathsf{S}} \cup \mathcal{I}_{\mathsf{R}|\mathbf{y}_1}| \leq |\mathcal{O}_{\mathsf{S}}| + \|\mathcal{I}_{\mathsf{R}}\| \leq |\mathcal{O}_{\mathsf{S}}| + |\mathcal{O}_{\mathsf{R}}| \leq t$, and since Square is t-NI, we conclude that $\|\mathcal{I}_{\mathsf{S}}\| \leq |\mathcal{O}_{\mathsf{S}}| + |\mathcal{O}_{\mathsf{R}}|$. Overall, we have $\mathsf{depset}_{\mathsf{Cube}}(\mathcal{O}) \subseteq \mathcal{I}_{\mathsf{S}|\mathbf{x}} \cup \mathcal{I}_{\mathsf{M}|\mathbf{x}}$, and we can conclude (using some of the intermediate inequalities from above), that $\|\mathsf{depset}_{\mathsf{Cube}}(\mathcal{O})\| \leq |\mathcal{O}_{\mathsf{S}}| + |\mathcal{O}_{\mathsf{R}}| + |\mathcal{O}_{\mathsf{M}}| \leq t$.

C.1 Game-Based Proofs for SNI Gadgets

In this appendix, we give detailed proofs for Propositions 1 to 2.

Multiplication-Based Refreshing (Prop. 1) Following the proof sketch for Proposition 1, we give here a more detailed proof sketch, based on a sequence of games, that more closely follows the EasyCrypt formalization of the proof. This is only meant to illustrate the additional effort involved in obtaining formal proofs of such results, rather than convince the reader of the validity of the proof. The proof scripts can be obtained from the authors. The formal proof is done by constructing a simulator (R_3) and formally verifying its equivalence with an explicitly leaky version of the gadget (R_0).

Game 0. For clarity, we start, in Game 0, from the original refreshing gadget RefreshM, which computes the $t+1$ shares c_i of the gadget's output without leaking any intermediate observations. We recall its definition on the right.

function $\mathsf{RefreshM}(\mathbf{a})$:
for $i = 0$ **to** t **do** $\mathbf{c}^i \leftarrow \mathbf{a}^i$;
for $i = 0$ **to** t **do**
  for $j = i+1$ **to** t **do**
    $r_{i,j} \leftarrow \$$;
    $\mathbf{c}^i \leftarrow \mathbf{c}^i \oplus r_{i,j}$;
    $\mathbf{c}^j \leftarrow \mathbf{c}^j \ominus r_{i,j}$;
return $\mathbf{c}$

Game 1. In Game 1, we make the leakage explicit by taking as additional input from the adversary a set of observations, whose values are written, during execution, into a table visible to the adversary. We refer to this function as R_0 (shown left).
As explained in the proof sketch above, we distinguish three different kinds of internal observations a_i, $r_{i,j}$, $c_{i,j}$ and have the final definitions of c_i as output observations. Considering this first game, we define the set $I_{\mathcal{O}}$ as a function of $\mathcal{O}^{\text{int}}$, as:

$$\begin{aligned} I_{\mathcal{O}} = &\{i \mid a_i \in \mathcal{O}\} \\ &\cup \{i \mid \exists j. r_{i,j} \in \mathcal{O}\} \\ &\cup \{i \mid \exists j. c_{i,j} \in \mathcal{O}\} \end{aligned}$$

function $\mathsf{R}_0(\mathbf{a}, \mathcal{O})$:
for $i = 0$ **to** t **do** $\mathbf{c}^i \leftarrow \mathbf{a}^i$;
  if $(a_i \in \mathcal{O})$ **then** $\overline{a_i} \leftarrow \mathbf{a}^i$;
for $i = 0$ **to** t **do**
  for $j = i+1$ **to** t **do**
    $r_{i,j} \leftarrow \$$;
    if $(r_{i,j} \in \mathcal{O})$ **then** $\overline{r_{i,j}} \leftarrow r_{i,j}$;
    $\mathbf{c}^i \leftarrow \mathbf{c}^i \oplus r_{i,j}$;
    if $(c_{i,j} \in \mathcal{O})$ **then** $\overline{c_{i,j}} \leftarrow \mathbf{c}^i$;
    $\mathbf{c}^j \leftarrow \mathbf{c}^j \ominus r_{i,j}$;
    if $(c_{j,i} \in \mathcal{O})$ **then** $\overline{c_{j,i}} \leftarrow \mathbf{c}^j$;
for $i = 0$ **to** t **do**
  if $(c_i \in \mathcal{O})$ **then** $\overline{c_i} \leftarrow \mathbf{c}^i$;
return $\mathbf{c}$;

Game 2. We modify R_0 in order to i. pre-sample all fresh randomness using a random function Sample, ii. start by computing internal observations, and output shares $\mathbf{c}^i$ such that $i \in I_{\mathcal{O}}$, iii. then compute output shares such that $i \notin I_{\mathcal{O}}$. We refer to this new function as R_1 and we use EasyCrypt to prove that R_0 and R_1 produce equivalent distributions on the requested observations if their position sets are the same set $\mathcal{O}$ such that $|\mathcal{O}| \leq t$ and their input encodings agree for each $\mathbf{a}^i$ where $i \in I_{\mathcal{O}}$.

function SumCij(i, j) :
$s \leftarrow \mathbf{a}^i$;
for $k = 0$ **to** j **do**
 if $(i < k)$ **then**
 $s \leftarrow s \oplus \overline{r_{i,k}}$;
 elseif $(i > k)$ **then**
 $s \leftarrow s \ominus \overline{r_{k,i}}$;
return s;

function $\mathsf{R}_1(\mathbf{a}, \mathcal{O})$:
for $i = 0$ **to** t **do** **for** $j = i + 1$ **to** t **do** $\overline{r_{i,j}} \leftarrow$ Sample(i, j);
for $i = 0$ **to** t **do if** $(a_i \in \mathcal{O})$ **then** $\overline{a_i} \leftarrow \mathbf{a}^i$;
for $i = 0$ **to** t **do**
 if $(i \in I_{\mathcal{O}})$ **then**
 for $j = 0$ **to** t **do**
 if $(r_{i,j} \in \mathcal{O})$ **then** $r_{i,j} \leftarrow \overline{r_{i,j}}$;
 if $(c_{i,j} \in \mathcal{O})$ **then** $\overline{c_{i,j}} \leftarrow$ SumCij(i, j);
 if $(c_i \in \mathcal{O})$ **then** $\mathbf{c}^i \leftarrow$ SumCij(i, t);
for $i = 0$ **to** t **do**
 if $(i \notin I_{\mathcal{O}} \wedge c_i \in \mathcal{O})$ **then** $\mathbf{c}^i \leftarrow$ SumCij(i, t);
return $\mathbf{c}$

Game 3. We now delay the generation of random values as much as possible, sampling just before their first use. We refer to this new function as R_2. We prove the equivalence between Games 2 and 3 if the share the same $\mathcal{O}$ such that $|\mathcal{O}| \leq t$ and if they agree on input shares $\mathbf{a}^i$ such that $i \in I$. This equivalence leverages a generic argument equating lazy and eager sampling when independent from intermediate adversary views.

function SumCij(i, j) :
$s \leftarrow \mathbf{a}^i$;
for $k = 0$ **to** j **do**
 if $(i < k)$ **then**
 $s \leftarrow s \oplus$ Sample(i, k);
 elseif $(i > k)$ **then**
 $s \leftarrow s \ominus$ Sample(k, i);
return s;

function $\mathsf{R}_2(\mathbf{a}, \mathcal{O})$:
for $i = 0$ **to** t **do if** $(a_i \in \mathcal{O})$ **then** $\overline{a_i} \leftarrow \mathbf{a}^i$;
for $i = 0$ **to** t **do**
 if $(i \in I_{\mathcal{O}})$ **then**
 for $j = 0$ **to** d **do**
 if $(r_{i,j} \in O)$ **then** $r_{i,j} \leftarrow$ Sample(i, j);
 if $(c_{i,j} \in \mathcal{O})$ **then** $\overline{c_{i,j}} \leftarrow$ SumCij(i, j);
 if $(c_i \in \mathcal{O})$ **then** $\mathbf{c}^i \leftarrow$ SumCij(i, t);
for $i = 0$ **to** d **do**
 if $(i \notin I_{\mathcal{O}} \wedge c_i \in \mathcal{O})$ **then** $\mathbf{c}^i \leftarrow$ SumCij(i, t);
return $\mathbf{c}$;

Game 4 In this final game, we fully simulate the computation of the c_i when $i \notin I_{\mathcal{O}}$ by showing that there exists a non-empty set of indices $L_{\mathcal{O}}$ such that $\forall \ell \in L,\ r_{i,\ell}$ is not

assigned yet. Then, instead of computing the c_i (for $i \notin I$) as in Game 2:

$$\forall \ell \in L,\ r_{i,\ell} \leftarrow \$;\ \ c_i \leftarrow \mathbf{a}^i \oplus \bigoplus_{j=0}^{i-1} r_{i,j} \ominus \bigoplus_{j=i+1}^{t} r_{i,j},$$

we simulate it by setting the value of $r_{i,k}$ as follows:

$$\begin{aligned}
&c_i \leftarrow \$;\ \ \forall \ell \in L \backslash \{k\},\ r_{i,\ell} \leftarrow \$,\\
&\text{if } (i < k),\ \ r_{i,k} \leftarrow c_i \ominus \mathbf{a}^i \ominus \bigoplus_{j=0, j\neq k}^{i-1} r_{i,j} \oplus \bigoplus_{j=i+1, j\neq k}^{t} r_{i,j},\\
&\text{if } (i > k),\ \ r_{i,k} \leftarrow \ominus c_i \oplus \mathbf{a}^i \oplus \bigoplus_{j=0, j\neq k}^{i-1} r_{i,j} \ominus \bigoplus_{j=i+1, j\neq k}^{t} r_{i,j}.
\end{aligned}$$

We prove the equivalence between Game 3 and Game 4 with EasyCrypt when functions R_2 and R_3 share the same $\mathcal{O}$ such that $|\mathcal{O}| \leq t$ and agree on input shares $\mathbf{a}^i$ with $i \in I$. The most critical part of this step is undoubtedly to ensure that the subscript L contains at least one index. To do so, we need to show that the elements $r_{i,\ell}$ with $\ell \in L$ were not already used and won't be reused anywhere. Finally, we also formally prove that the results of R_3, which represents the final simulator, only depends on those input shares $\mathbf{a}^i$ such that $i \in I_{\mathcal{O}}$ for all $\mathcal{O}$ such that $|\mathcal{O}| \leq t$.

function SumCij(i, j) :
$s \leftarrow \mathbf{a}^i$;
for $k = 0$ **to** j **do**
 if $(i < k)$ **then** $s \leftarrow s \oplus$ Sample(i, k);
 else if $(i > k)$ **then** $s \leftarrow s \ominus$ Sample(k, i);
return s;
function SetCi(i);
$s \leftarrow \mathbf{a}^i;\ \ k \leftarrow 0$;
while $((k \leqslant t) \wedge ((i == k) \vee (k < i \wedge ((i, k) \in \text{dom } r)) \vee$
 $(i < k \wedge ((k, i) \in \text{dom } r)))$ **do**
 if $(i < k)$ **then** $s \leftarrow s \oplus r_{i,k}$;
 else if $(i > k)$ **then** $s \leftarrow s \ominus r_{i,k}$;
 $k \leftarrow k + 1$;
$k' \leftarrow k;\ \ r' \leftarrow \$$;
for $k = k'$ **to** t **do** $s \leftarrow s \oplus r_{i,k}$;
if $(i < k')$ **then** $r_{i,k'} \leftarrow s \oplus r'$; **else** $r_{k',i} \leftarrow s \ominus r'$;
return r';

```
function R3(a, O) :
for i = 0 to t do if (a_i ∈ O) then ā_i ← a^i;
for i = 0 to t; i ∈ I_O do
   for j = 0 to t do
     if (r_{i,j} ∈ O) then r_{i,j} ← Sample(i, j);
     if (c_{i,j} ∈ O) then c̄_{i,j} ← SumCij(i, j);
   if (c_i ∈ O) then c^i ← SetCi(i);
for i = 0 to t; i ∉ I_O do
   if (c_i ∈ O) then c^i ← SetCi(i);
return c;
```

Finally, we have formally proved that an adversary could not distinguish between two programs which share the same inputs $\{a_i\}_{i \in I}$ with at most t observations and that the cardinal of I was upper bounded by the number of internal observations d_1.

Secure Multiplication (Prop. 2) We now give an informal proof sketch that Gadget 7 is t-SNI for all t. The proof is formalized in EasyCrypt following the model described above for Proposition 1.

Proof (sketch for Proposition 2). Let $\mathcal{O}$ be a set of observations such that $|\mathcal{O}| \leq t$, and let $d_1 = |\mathcal{O}^{\text{int}}|$ and $d_2 = |\mathcal{O}^{\text{ext}}|$. Note that $d_1 + d_2 \leq t$. To prove t-SNI, we aim to find an input set $\mathcal{I}$ such that $\|\mathcal{I}\| \leq d_1$, and to construct a perfect simulator that uses only input shares from $\mathcal{I}$ to compute the same distribution over the observations $\mathcal{O}$ as the real gadget.

First, we identify which variables are internals and which are outputs. We directly split the internals in four groups for the needs of the proof:

1. the a_i, the b_i, and the $a_i \otimes b_i$,
2. the $c_{i,j}$ (resp. $c_{j,i}$) which correspond to the value of the variable c_i (resp. c_j) at iteration i, j,
3. the $r_{i,j}$ (the value of r on line $5^{i,j}$), and the $r_{j,i}$ (the value of r on line $8^{i,j}$),
4. the $a_i \otimes b_j$ and the $a_i \otimes b_j \ominus r_{i,j}$.

The output variables are the final values of $\mathbf{c}^i$ (that is $c_{i,t}$).

We construct $\mathcal{I}$. As the algorithm takes two inputs $\mathbf{a}$ and $\mathbf{b}$, we define two sets of indices $I, J \subsetneq \{0, \ldots, t\}$ such that $\mathcal{I} = \{\mathbf{a}^i \mid i \in I\} \cup \{\mathbf{b}^i \mid i \in J\}$. It is then sufficient to show that they are such that $|I|, |J| \leq |\mathcal{O}^{\text{int}}|$, and that the gadget is $(\mathcal{I}, \mathcal{O})$-NI. For each observation in the first or the second group, we add the index i to both I and J. For each observation in the third or in the fourth group: if the index i is already in I, we add the index j to I, otherwise we add the index i to I and similarly for J. It is clear that the final sets I and J contain each one at most d_1 indices, since each internal observation adds at most one index to each of them.

We now construct the simulator. For clarity, observe that the SecMult algorithm can be equivalently represented using the following matrix.

$$\begin{pmatrix} a_0 \otimes b_0 & 0 & r_{0,1} & r_{0,2} & \cdots & r_{0,t} \\ a_1 \otimes b_1 & r_{1,0} & 0 & r_{1,2} & \cdots & r_{1,t} \\ a_2 \otimes b_2 & r_{2,0} & r_{2,1} & 0 & \cdots & r_{2,t} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_t \otimes b_t & r_{t,0} & r_{t,1} & r_{t,2} & \cdots & 0 \end{pmatrix}.$$

In this setting, $c_{i,j}$ corresponds to the partial sum of the $j+2$ first elements of line i. For each variable $r_{i,j}$ $(i < j)$ entering in the computation of an observation, we assign it a fresh random value. Then, for each observation in the first group, a_i and b_i are provided to the simulator (by definition of I and J) thus the observation is perfectly simulated. For an observation in the third group, we distinguish two cases. If $i < j$, $r_{i,j}$ is already assigned to a fresh random value. If $i > j$, either $(i, j) \in I \times J$ and the observation is perfectly simulated from $r_{j,i}, a_i, b_i, a_j$ and b_j or $r_{j,i}$ does not enter in the computation of any internal variable that was observed and $r_{i,j}$ (line $5^{i,j}$) was assigned a fresh random value. Each observation made in the fourth group is perfectly simulated using $r_{i,j}$, a_i and b_j. As for an observation in the second group, the corresponding variable is a partial sum composed of a product $a_i \otimes b_i$ and of variables $r_{i,j}$. Since a_i and b_i are provided to the simulator in this case, we focus on each remaining $r_{i,j}$. Each one of them such that $i < j$ is already assigned to a fresh random value. For the others, if $r_{j,i}$ enters in the computation of any other internal observation, then $(i, j) \in I \times J$ and $r_{i,j}$ is simulated with $r_{j,i}, a_i, b_i, a_j$ and b_j. Otherwise, $r_{i,j}$ is assigned a fresh random value.

We still have to simulate output observations. We start with those for which an intermediate sum (group 2) is also observed. For each such variable c_i, the biggest partial sum which is observed is already simulated. Thus, we consider the remaining terms $r_{i,j}$. Each one of them such that $i < j$ is already assigned to a fresh random value. For the others, either $(i, j) \in I \times J$ and c_i is perfectly simulated from $r_{j,i}$, a_i, b_i, a_j and b_j or $r_{j,i}$ does not enter in the computation of any internal variable observed and c_i is assigned a fresh random value. We now consider output observations c_i for which none of the partial sums are observed. Each of them contains t random variables of the form $r_{i,j}$, at most one of which can enter in the computation of each one of the c_j with $i \neq j$. Since we already considered at most $t-1$ observations not counting the current one, at least one of the $r_{i,j}$ we need to consider does not enter in the computation of any other observed variable and is unassigned. Thus, we can simulate c_i using a fresh random value.

D Robust Composition

We consider the problem of securely composing algorithms or gadgets when the adversary may place t probes in each of them, capturing security in the region and stateful probing models [24]. In this informal discussion, we forbid wire duplications between regions (that is, each region's output can only be used once as input to another region). However, the same principles that allow us to compositional reason about security in

the simple probing model would allow us to reason about robust composition with wire duplication, using, in particular, $2t$-SNI mask refreshing gadgets where necessary.

As noted in Section 9, it is clear that any number of $2t$-SNI gadgets or algorithms can be securely composed in this setting. We first observe that, for any two $2t$-SNI algorithms F and G, and for any $\mathcal{O}_F$ and $\mathcal{O}_G$ such that $|\mathcal{O}_F|, |\mathcal{O}_G| \leq t$, we have $\|\mathsf{depset}_{F;G}(\mathcal{O}_F \cup \mathcal{O}_F)\| \leq t$. This allows us to conclude about the security of the composition of any number of gadgets by induction on that number.

This observation provides an elegant justification to Coron's result [12], which states that iterating $\mathsf{RefreshA}_{2t+1}$ $2t+1$ times allows the secure stateful composition of his masked table-lookup algorithms without further doubling the number of shares. Indeed, it is also easy to see that Coron's mask refreshing algorithm is $2t$-SNI.

However, the same composition could be obtained by slightly relaxing the security requirement on F and G. Indeed, we could instead use the following condition, which we call *robust non-interference*.

Definition 9 (Robust Non-Interference) *A gadget G is t-*robustly non-interfering *(or t-RNI) whenever, for any $\mathcal{O}_G$ such that $|\mathcal{O}_G| \leq 2t$ and $|\mathcal{O}_G^{\mathrm{int}}| \leq t$, we have $\|\mathsf{depset}_G(\mathcal{O}_G)\| \leq |\mathcal{O}_G^{\mathrm{int}}|$.*

The same argument that allows us to securely and robustly compose $2t$-SNI gadgets allows us to securely and robustly compose t-RNI gadgets. Intuitively, t-RNI allows the adversary to place $2t$ probes on the gadget, as long as at most t of them are internal. This captures exactly the scenario where an adversary can place t probes inside a region, and can also learn, by probing subsequent regions, information about t shares of the output.

Constructing RNI Algorithms We now show how any t-NI algorithm can be turned into a t-RNI algorithm without doubling its internal number of shares, simply by processing its inputs and output. To do so, we consider the Double and Half gadgets (Gadget 13).[14]

Gadget 13 The Double and Half gadgets

function $\mathsf{Double}_t(\mathbf{a} : \mathbb{K}^{t+1})$
for $i = 0$ **to** t **do**
$r_{2i} \xleftarrow{\$} \mathbb{K}$
$r_{2i+1} \leftarrow \mathbf{a}^i \ominus r_{2i}$
return $\langle r_0, \ldots, r_{2t+1} \rangle$

(13a) Doubling the number of shares

function $\mathsf{Half}_t(\mathbf{a} : \mathbb{K}^{2t+2})$
for $i = 0$ **to** t **do**
$r_i \xleftarrow{\$} \mathbf{a}^{2i} \oplus \mathbf{a}^{2i+1}$
return $\langle r_0, \ldots, r_t \rangle$

(13b) Halving the number of shares

We first note that Double_t is t-NI, and that Half_t is such that, for any $\mathcal{O}_\mathsf{H}$ such that $|\mathcal{O}_\mathsf{H}| \leq t$, we have $\|\mathsf{depset}_\mathsf{H}(\mathcal{O}_\mathsf{H})\| \leq 2t$. (Indeed, each internal position depends two shares of the input.)

[14] These are only gadgets for a slightly extended notion of gadget that allows the encoding size to change between inputs and outputs.

Alg. 14 Making a t-NI gadget t-RNI

function $\mathsf{Robust}_G(\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{K}^{2t+2})$
 for $i = 1$ **to** n **do**
  $\mathbf{a}_i := \mathsf{RefreshM}_{2t+1}(\mathbf{a}_i)$
  $\mathbf{x}_i := \mathsf{Half}_t(\mathbf{a}_i)$
 $\mathbf{y} := G(\mathbf{x}_1, \ldots, \mathbf{x}_n)$
 $\mathbf{r} := \mathsf{Double}(\mathbf{y})$
 $\mathbf{r} := \mathsf{RefreshM}_{2t+1}(\mathbf{r})$

Lemma 6 *Given a t-NI gadget G, the algorithm shown in Alg. 14 is t-RNI.*

Proof. The proof is by composition, leveraging the following facts (for some $\mathcal{O}$ as assumed by t-RNI): i. $\mathsf{RefreshM}_{2t+1}$ is t-SNI (so the dependency set of its position set is of degree at most $|\mathcal{O}_{\mathsf{R}_2}| \leq t$; ii. Double_t is t-NI (so the dependency set of its position set is of degree at most $|\mathcal{O}_{\mathsf{D}}| + |\mathcal{O}_{\mathsf{M}_2}| \leq t$); iii. G is t-NI (so the dependency set of its position set and output set is of degree at most $|\mathcal{O}_G| + |\mathcal{O}_{\mathsf{D}}| + |\mathcal{O}_{\mathsf{R}_2}| \leq t$); iv. Half_t has the property described above (so the dependency set of its position set and output set is of degree at most $2(|\mathcal{O}_{\mathsf{H}}| + |\mathcal{O}_G| + |\mathcal{O}_{\mathsf{D}}| + |\mathcal{O}_{\mathsf{R}_2}|) \leq 2t$); v. $\mathsf{RefreshM}_{2t+1}$ is $2t$-SNI (so the dependency set of its position set and output set if of degree at most $|\mathcal{O}_{\mathsf{R}_1}| \leq |\mathcal{O}^{\mathrm{int}}|$).

E Privacy vs probing security

In this appendix, we exhibit a simple example that separates Ishai, Sahai and Wagner's notion of *privacy* [24] and the widely used simulation-based notion of *t-probing security*. We recall that Ishai, Sahai and Wagner [24] define privacy for a masked circuit C as the fact that any t wires in C during an execution of $O \circ C \circ I$ (where I and O are the encoding and decoding circuits that sample uniform encodings of the secret inputs and recombine outputs from their encodings, away from adversary interference) can be simulated without access to any wire in the circuit.[15] On the other hand, as discussed earlier in the paper, t-probing security, as defined for example by Carlet et al. [10] states that a gadget or algorithm G is t-probing secure iff any t of its intermediate wires can be simulated using at most t shares of each of its inputs.

Note in particular that, unlike privacy, t-probing security makes no mention of secrets, or of uniform input encodings. In Gadget 15, we exhibit a small example gadget, which computes $\mathbf{a} \odot (\mathbf{a} \oplus \mathbf{b}$ for $t = 2$, and shows that this is indeed an important detail by separating the two notions. However, we note that even Ishai, Sahai and Wagner [24] prove that their transformers are private using a t-probing style simulation argument, so the separation we exhibit here makes little difference in practice.

[15] More accurately, they define privacy for a circuit transformer that includes definitions for I and O. This is not relevant here.

Gadget 15 A small separating example

```
1:  function separator(a, b : 𝕂^3)
2:    for i = 0 to 2 do
3:      r^i := a^i ⊕ b^i
4:    for i = 0 to 2 do
5:      c^i ← a^i ⊙ r^i
6:    for i = 0 to 2 do
7:      for j = i + 1 to 2 do
8:        r ←$ 𝕂
9:        c^i ← c^i ⊖ r
10:       t ← a^i ⊙ r^j
11:       r ← r ⊕ t
12:       t ← a^j ⊙ r^i
13:       r ← r ⊕ t
14:       c^j ← c^j ⊕ r
15:   return c
```

It is easy to see (and it can be checked automatically using, for example, Barthe et al.'s tool [4]) that this small algorithm is *private* for $t = 2$. Intuitively, this is because, since privacy requires simulation only when the input encodings are known to be uniform, and even though the inputs to SecMult are not independent, one of them is essentially a one-time-pad of the other.

However, simulating, say, lines $10^{0,1}$ (with value $\mathbf{a}^0 \odot (\mathbf{a}^1 \oplus \mathbf{b}^1)$) and $10^{1,2}$ (with value $\mathbf{a}^1 \odot (\mathbf{a}^2 \oplus \mathbf{b}^2)$) using only two shares of each of $\mathbf{a}$ and $\mathbf{b}$ is not possible.

We note that, as pointed out above, the tool by Barthe et al. [4] directly verify privacy, rather than going through t-NI and losing precision. However, they do not scale. When considering verification tools (or indeed even the verifiability of pen-and-paper proofs), there is therefore a tradeoff between precision and applicability.