Computationally sound implementations of equational theories against passive adversaries

Mathieu Baudet[1], Véronique Cortier[2], and Steve Kremer[1]

[1] LSV/ CNRS UMR 8643 & INRIA Futurs projet SECSI & ENS Cachan, France
{baudet,kremer}@lsv.ens-cachan.fr
[2] Loria UMR 7503 & INRIA Lorraine projet Cassis, France
cortier@loria.fr

Abstract. In this paper we study the link between formal and cryptographic models for security protocols in the presence of a passive adversary. In contrast to other works, we do not consider a fixed set of primitives but aim at results for an arbitrary equational theory. We define a framework for comparing a cryptographic implementation and its idealization *w.r.t.* various security notions. In particular, we concentrate on the computational soundness of static equivalence, a standard tool in cryptographic pi calculi. We present a soundness criterion, which for many theories is not only sufficient but also necessary. Finally, we establish new soundness results for the Exclusive Or, as well as a theory of ciphers and lists.

1 Introduction

Today's ubiquity of computer networks increases the need for theoretic foundations for cryptographic protocols. For more than twenty years now, two communities separately developed two families of models. Both views have been very useful in increasing the understanding and quality of security protocol design. On the one hand *formal* or *logical* models have been developed, based on the seminal work of Dolev and Yao [12]. These models view cryptographic operations in a rather abstract and idealized way. On the other hand *cryptographic* or *computational* models [13] are closer to implementations: cryptographic operations are modeled as algorithms manipulating bit-strings. Those models cover a large class of attacks, namely all those implementable by a probabilistic polynomial-time Turing machine.

The advantage of formal models is that security proofs are generally simpler and suitable for automatic procedures, even for complex protocols [10, 9, 1]. Unfortunately, the high degree of abstraction and the limited adversary power raise serious questions regarding the security offered by such proofs. Potentially, justifying symbolic proofs with respect to standard computational models has tremendous benefits: protocols can be analyzed using automated tools and still benefit from the security guarantees of the computational model.

Recently, a significant research effort has been directed at linking the two approaches. One of the first result is presented by Abadi and Rogaway [5]: they prove the computational soundness of formal (symmetric) encryption in the case a passive attacker. Since then, many results [7, 6, 14, 15] have been obtained. In particular, Backes et al. [7, 6]

prove the soundness of a rich language including digital signatures, public-key and symmetric key encryption in the presence of an active attacker. Laud [14] presents an automated procedure for computationally sound proofs of confidentiality in the case of an active attacker and symmetric encryption when the number of sessions is bounded.

Each of these results considers a fixed set of primitives, *e.g.* symmetric or public-key encryption. In this paper, we aim at presenting general results for arbitrary equational theories, such as encryption, but also less studied ones, *e.g.* groups or exclusive or. We concentrate on *static equivalence*, a notion of indistinguishability common in cryptographic pi calculi [4, 3]. Intuitively, static equivalence asks whether an attacker can distinguish between two tuples of terms, by exhibiting an equation which holds on one tuple but not on the other. This provides an elegant mean to express security properties against passive attackers. Moreover there exist exact [2] and approximate [11] algorithms to decide static equivalence for a large family of equational theories.

Our first contribution is a general framework for comparing formal and computational models in the presence of a passive attacker. We define the notions of *soundness* and *faithfulness* of a cryptographic implementation *w.r.t.* equality, static equivalence and deducibility. Soundness holds when each formal proof has a computational interpretation. Faithfulness is the converse, *i.e.* the formal model does not provide false attacks.

Our second contribution is a sufficient criterion for soundness *w.r.t* static equivalence: intuitively the usual computational semantics of terms has to be indistinguishable to an idealized one. We also provide a general definition of patterns for arbitrary equational theories that encompasses the notion usually defined for symmetric and public encryption. Those patterns allow us to characterize a large class of theories for which our soundness criterion is necessary.

Our third contribution consists in applying our framework to obtain two novel soundness results. The first theory deals with the Exclusive Or. Interestingly, our proof reflects the unconditional security (in the information-theoretic sense) of the One-Time Pad encryption scheme. Second we consider a theory of symmetric encryption and lists. In some sense, the result is similar to the one of Abadi and Rogaway [5]. However, we consider deterministic, length-preserving, symmetric encryption schemes *a.k.a.* ciphers. To the best of our knowledge, this is the first result on such schemes, whose specificity is that decryption always succeeds.

Outline of the paper. In the next section, we introduce our abstract and concrete models together with the notions of indistinguishability. We then define the notions of soundness and faithfulness and illustrate some consequences of soundness *w.r.t.* static equivalence on groups. In Section 4, we define the ideal semantics of abstract terms, present our soundness criterion and also show that for a large family of interesting equational theories, the soundness criterion is a necessary condition. As an illustration (Section 5), we prove the soundness for the theories modeling Exclusive Or, as well as ciphers and lists. We then conclude and give directions for future work.

2 Modeling cryptographic primitives with abstract algebras

In this section we introduce some notations and set our abstract and concrete models.

2.1 Abstract algebras

Our abstract models—which we call *abstract algebras*—consist of term algebras defined on a first-order signature with sorts and equipped with equational theories.

Specifically a *signature* $(\mathcal{S}, \mathcal{F})$ is made of a set of *sorts* $\mathcal{S} = \{s, s_1 \ldots\}$ and a set of *symbols* $\mathcal{F} = \{f, f_1 \ldots\}$ together with arities of the form $\mathrm{ar}(f) = s_1 \times \ldots \times s_k \to s$, $k \geq 0$. Symbols that take $k = 0$ arguments are called *constants*; their arity is simply written s. We fix an infinite set of *names* $\mathcal{N} = \{a, b \ldots\}$ and an infinite set of *variables* $\mathcal{X} = \{x, y \ldots\}$. We assume that names and variables are given with sorts. The set of *terms of sort* s is defined inductively by

$$\begin{array}{rll} T ::= & & \text{term of sort } s \\ & \mid \; x & \text{variable } x \text{ of sort s} \\ & \mid \; a & \text{name } a \text{ of sort s} \\ & \mid \; f(T_1, \ldots, T_k) & \text{application of symbol } f \in \mathcal{F} \end{array}$$

where for the last case, we further require that T_i is a term of some sort s_i and $\mathrm{ar}(f) = s_1 \times \ldots \times s_k \to s$. As usual, we write $\mathrm{var}(T)$ and $\mathrm{names}(T)$ for the set of variables and names occurring in T respectively. A term is *ground* or *closed* iff it has no variables.

Substitutions are written $\sigma = \{x_1 = T_1, \ldots, x_n = T_n\}$ with $\mathrm{dom}(\sigma) = \{x_1, \ldots, x_n\}$. We only consider *well-sorted* substitutions, that is substitutions $\sigma = \{x_1 = T_1, \ldots, x_n = T_n\}$ for which x_i and T_i have the same sort. σ is *closed* iff all of the T_i are closed. We extend the notation $\mathrm{names}(.)$ from terms to substitutions in the obvious way. The application of a substitution σ to a term T is written $\sigma(T) = T\sigma$.

Symbols in $\mathcal{F}$ are intended to model cryptographic primitives, whereas names in $\mathcal{N}$ are used to model nonces *i.e.* concretely random numbers. The abstract semantics of symbols is described by an equational theory E, that is an equivalence relation (also written $=_E$) which is stable by application of contexts and well-sorted substitutions of variables. We further require that E is stable by substitution of names. All the equational theories that we consider in this paper satisfy these properties. For instance, symmetric and deterministic encryption is modeled by the theory E_{enc} generated by the classical equation $E_{\mathsf{enc}} = \{\mathsf{dec}(\mathsf{enc}(x, y), y) = x\}$.

2.2 Frames, deducibility and static equivalence

Following [3, 2], a *frame* is an expression $\varphi = \nu\tilde{a}.\sigma$ where $\tilde{a}$ is a set of *bound (or restricted) names* and σ is a well-sorted substitution. Intuitively, frames represent sequences of messages learned by an attacker during the execution of a protocol.

For simplicity we only consider frames $\nu\tilde{a}.\sigma$ which restrict *every* names occurring in σ, that is $\tilde{a} = \mathrm{names}(\sigma)$. In other words, names a must be disclosed *explicitly* by adding a mapping $x_a = a$ to the substitution. Thus we tend to assimilate frames and their underlying substitutions.

A term T is *deducible* from a closed frame φ, written $\varphi \vdash_E T$ iff there exists a term M with $\mathrm{var}(M) \subseteq \mathrm{dom}(\varphi)$ and $\mathrm{names}(M) \cap (\mathrm{names}(\varphi) \cup \mathrm{names}(T)) = \emptyset$ such that $M\varphi =_E T$. Consider for instance the theory E_{enc} and the frame $\varphi_1 = \nu k_1, k_2, k_3, k_4. \{x_1 = \mathsf{enc}(k_1, k_2),\ x_2 = \mathsf{enc}(k_4, k_3),\ x_3 = k_3\}$: the name k_4 is deducible from φ_1 since $\mathsf{dec}(x_2, x_3)\varphi_1 =_{E_{\mathsf{enc}}} k_4$ but neither k_1 nor k_2 are deducible.

Deducibility is not always sufficient to account for the knowledge of an attacker. *E.g.* it lacks partial information on secrets. This is why the notion of static equivalence is used. Two closed frames φ_1 and φ_2 are *statically equivalent*, written $\varphi_1 \approx_E \varphi_2$, iff (i) $\mathrm{dom}(\varphi_1) = \mathrm{dom}(\varphi_2)$, (ii) for all terms M, N with variables included in $\mathrm{dom}(\varphi_i)$ and using no names occurring in φ_1 or φ_2, $M\varphi_1 =_E N\varphi_1$ is equivalent to $M\varphi_2 =_E N\varphi_2$.

For instance, the two frames $\nu k.\,\{x = \mathsf{enc}(0, k)\}$ and $\nu k.\,\{x = \mathsf{enc}(1, k)\}$ are statically equivalent with respect to E_{enc}, whereas the two frames $\nu k.\,\{x = \mathsf{enc}(0, k),\ y = k\}$ and $\nu k, k'.\{x = \mathsf{enc}(0, k'),\ y = k\}$ are not.

2.3 Concrete semantics

We now give terms and frames a concrete semantics, parameterized by an implementation of the primitives. Provided a set of sorts $\mathcal{S}$ and a set of symbols $\mathcal{F}$ as above, a *$(S, \mathcal{F})$-computational algebra A* consists of

- a non-empty set of bit-strings $[\![s]\!]_A \subseteq \{0, 1\}^*$ for each sort $s \in \mathcal{S}$;
- a function $f_A : [\![s_1]\!]_A \times \ldots \times [\![s_k]\!]_A \to [\![s]\!]_A$ for each $f \in \mathcal{F}$ with $\mathrm{ar}(f) = s_1 \times \ldots \times s_k \to s$, such that f_A is computable in polynomial time;
- a congruence $=_{A,s}$ for each sort s that is computable in polynomial time, in order to check the equality of elements in $[\![s]\!]_A$ (the same element may be represented by different bit-strings); by congruence, we mean a reflexive, symmetric, transitive relation such that $e_1 =_{A,s_1} e'_1, \ldots, e_k =_{A,s_k} e'_k \Rightarrow f_A(e_1, \ldots, e_k) =_{A,s} f_A(e'_1, \ldots, e'_k)$ (in the remaining we often omit s and write $=_A$ for $=_{A,s}$);
- a polynomial-time algorithm to draw random elements from $[\![s]\!]_A$; we denote such a drawing by $x \xleftarrow{R} [\![s]\!]_A$; the drawing may not follow a uniform distribution, but no $=_{A,s}$-equivalence class should have probability 0.

Assume a fixed $(S, \mathcal{F})$-computational algebra A. We associate to each closed frame $\varphi = \{x_1 = T_1, \ldots, x_n = T_n\}$ a distribution $\psi = [\![\varphi]\!]_A$, of which the drawings $\widehat{\psi} \leftarrow \psi$ are computed as follows:

1. for each name a appearing in $T_1, \ldots, T_n$, draw a value $\widehat{a} \xleftarrow{R} [\![s]\!]_A$;
2. for each x_i ($1 \leq i \leq n$) of sort s_i, compute $\widehat{T_i} \in [\![s_i]\!]_A$ recursively on the structure of terms: $f(\widehat{T'_1, \ldots, T'_m}) = f_A(\widehat{T'_1}, \ldots, \widehat{T'_m})$;
3. return the value $\widehat{\psi} = \{x_1 = \widehat{T_1}, \ldots, x_n = \widehat{T_n}\}$.

Such values $\phi = \{x_1 = e_1, \ldots, x_n = e_n\}$ with $e_i \in [\![s_i]\!]_A$ are called *concrete frames*. We extend the notation $[\![.]\!]_A$ to (sets of) closed terms in the obvious way. We also generalize the notation to terms or frames with variables, by specifying the concrete values for all of them: $[\![.]\!]_{A,\{x_1=e_1,\ldots,x_n=e_n\}}$. Notice that when a term or a frame contains no names, the translation is deterministic; in this case, we use the same notation to denote the distribution and its unique value.

(Families of) distributions over concrete frames benefit from the usual notion of cryptographic indistinguishability. Let us note $\eta \geq 0$ the complexity parameter. Two families (ψ_η) and (ψ'_η) of distributions over concrete frames are *indistinguishable*, written $(\psi_\eta) \approx (\psi'_\eta)$, iff for every probabilistic polynomial-time adversary $\mathcal{A}$, intuitively $\mathcal{A}$

cannot guess whether he is given a sample from ψ_η or ψ'_η with a probability significantly greater than $\frac{1}{2}$. Rigorously, we ask that the *advantage* of $\mathcal{A}$,

$$\mathrm{Adv}^{\mathrm{IND}}(\mathcal{A}, \eta, \psi_\eta, \psi'_\eta) = \mathbb{P}\left[\widehat{\psi} \leftarrow \psi_\eta; \mathcal{A}(\eta, \widehat{\psi}) = 1\right] - \mathbb{P}\left[\widehat{\psi} \leftarrow \psi'_\eta; A(\eta, \widehat{\psi}) = 1\right]$$

is a *negligible* function of η, that is, remains eventually smaller than any η^{-n} $(n > 0)$ as η tends to infinity.

3 Relating abstract and computational algebras

In the previous section we have defined abstract and computational algebras. We now relate formal notions such as equality, (non-)deducibility and static equivalence to their computational counterparts, *i.e.* equality, one-wayness and indistinguishability.

3.1 Soundness and faithfulness

We introduce the notions of sound, *resp.* faithful, computational algebras with respect to the formal relations studied here: equality, static equivalence and deducibility. In the remaining of the paper we only consider families of computational algebras (A_η) such that for any sort s, the probability of collisions of two random elements in $[\![s]\!]_{A_\eta}$, $\mathbb{P}\left[\, e_1, e_2 \leftarrow [\![s]\!]_{A_\eta};\ e_1 =_{A_\eta} e_2\right]$, is negligible.

Specifically a family of computational algebras (A_η) is

- $=_E$*-sound* iff for every closed terms T_1, T_2 of the same sort, $T_1 =_E T_2$ implies that $\mathbb{P}\left[\, e_1, e_2 \leftarrow [\![T_1,\ T_2]\!]_{A_\eta};\ e_1 \neq_{A_\eta} e_2\right]$ is negligible;
- $=_E$*-faithful* iff for every closed terms T_1, T_2 of the same sort, $T_1 \neq_E T_2$ implies that $\mathbb{P}\left[\, e_1, e_2 \leftarrow [\![T_1,\ T_2]\!]_{A_\eta};\ e_1 =_{A_\eta} e_2\right]$ is negligible;
- $\approx_E$*-sound* iff for every closed frames φ_1, φ_2 of the same domain, $\varphi_1 \approx_E \varphi_2$ implies that $([\![\varphi_1]\!]_{A_\eta}) \approx ([\![\varphi_2]\!]_{A_\eta})$;
- $\approx_E$*-faithful* iff for every closed frames φ_1, φ_2 of the same domain, $\varphi_1 \not\approx_E \varphi_2$ implies that there exists a polynomial-time adversary $\mathcal{A}$ for distinguishing concrete frames, such that $1 - \mathrm{Adv}^{\mathrm{IND}}(\mathcal{A}, \eta, [\![\varphi_1]\!]_{A_\eta}, [\![\varphi_2]\!]_{A_\eta})$ is negligible;
- $\not\vdash_E$*-sound* iff for every closed φ and T, $\varphi \not\vdash_E T$ implies that for all polynomial-time adversary $\mathcal{A}$, $\mathbb{P}\left[\phi, e \leftarrow [\![\varphi, T]\!]_{A_\eta}; \mathcal{A}(\phi) =_{A_\eta} e\right]$ is negligible;
- $\not\vdash_E$*-faithful* iff for every closed φ and T, $\varphi \vdash_E T$ implies that there exists a polynomial-time adversary $\mathcal{A}$ such that $1 - \mathbb{P}\left[\phi, e \leftarrow [\![\varphi, T]\!]_{A_\eta}; \mathcal{A}(\phi) =_{A_\eta} e\right]$ is negligible.

Sometimes, it is possible to prove stronger notions of soundness that hold without restriction on the computational power of adversaries. In particular, (A_η) is *unconditionally* $=_E$*-sound* iff for every closed terms T_1, T_2 of the same sort, $T_1 =_E T_2$ implies that $\mathbb{P}\left[\, e_1, e_2 \leftarrow [\![T_1,\ T_2]\!]_{A_\eta};\ e_1 =_{A_\eta} e_2\right] = 1$; *unconditionally* $\approx_E$*-sound* iff for every closed frames φ_1, φ_2 of the same domain, $\varphi_1 \approx_E \varphi_2$ implies $([\![\varphi_1]\!]_{A_\eta}) = ([\![\varphi_2]\!]_{A_\eta})$; *unconditionally* $\not\vdash_E$*-sound* iff for every closed φ and T *s.t.* $\varphi \not\vdash_E T$, the distributions for φ and T are independent: for all ϕ_0, e_0, $\mathbb{P}\left[\phi, e \leftarrow [\![\varphi, T]\!]_{A_\eta}; \phi = \phi_0 \text{ and } e = e_0\right] = \mathbb{P}\left[\phi \leftarrow [\![\varphi]\!]_{A_\eta}; \phi = \phi_0\right] \times \mathbb{P}\left[e \leftarrow [\![T]\!]_{A_\eta}; e = e_0\right]$.

Generally, (unconditional) $=_E$-soundness is given by construction. Indeed true formal equations correspond to the expected behavior of primitives and should hold in the concrete world with overwhelming probability. The other criteria are however more difficult to fulfill. Therefore it is often interesting to restrict frames to *well-formed* ones in order to achieve soundness or faithfulness: for instance Abadi and Rogaway [5] do forbid encryption cycles (*c.f.* Section 5.2).

It is worth noting that the notions of soundness and faithfulness introduced above are not independent.

Proposition 1. *Let (A_η) be a $=_E$-sound family of computational algebras. Then (A_η) is $\not=_E$-faithful. If moreover (A_η) is $=_E$-faithful, then it is also $\approx_E$-faithful.*

The proof is given in Appendix A.1. For many interesting theories, we have that $\approx_E$-soundness implies all the other notions of soundness and faithfulness. This emphasizes the importance of $\approx_E$-soundness and provides an additional motivation for its study. As an illustration, let us consider an arbitrary theory which includes keyed hash functions.

Proposition 2. *Let (A_η) be a family of $\approx_E$-sound computational algebras. Assume that free binary symbols $\mathsf{h}_s : s \times Key \to Hash$ are available for every sort s. Then (A_η) is $=_E$-faithful and $\not=_E$-sound. Besides if the implementations for the h_s are collision-resistant, then (A_η) is $=_E$-sound, $\approx_E$-faithful and $\not=_E$-faithful.*

The proof (Appendix A.2) is done by encoding the different problems with $\approx_E$ and h_s.

3.2 $\approx_E$-soundness implies classical assumptions on groups

In this section we present some interesting consequences of $\approx_E$-soundness. Inspired by the work of Rivest on pseudo-freeness [17], we show that several standard cryptographic assumptions are a direct consequence of the soundness of a theory representing *groups*. Let E_G be the equational theory modeling a free group G with exponents taken over a free commutative ring A. (Precise sets of symbols and equations are detailed in Appendix B.)

We now introduce several classical problems on groups, which in cryptography are considered to be *hard*, *i.e.* not feasible by any probabilistic polynomial-time adversary:

- *discrete logarithm* (DL) problem: given g and g', find a, such that $g^a = g'$;
- *computational Diffie-Hellman* (CDH) problem: given g, g^a and g^b, find g^{ab};
- *decisional Diffie-Hellman* (DDH) problem: given g, g^a and g^b, distinguish g^{ab} from a random element g^c;
- *RSA* problem: given elements a and g^a, find g.

Suppose that there is a family of computational algebras (A_η) which are $\approx_{E_\mathsf{G}}$-sound. Then no probabilistic polynomial-time adversary $\mathcal{A}$ can solve the DDH problem with non-negligible probability. Indeed consider the two frames

$$\varphi_1 = \nu g, a, b.\{x_1 = g, x_2 = g^a, x_3 = g^b, x_4 = g^{a \cdot b}\} \text{ and}$$
$$\varphi_2 = \nu g, a, b, c.\{x_1 = g, x_2 = g^a, x_3 = g^b, x_4 = g^c\}.$$

The question of distinguishing these two frames encodes exactly the DDH problem. Given the equational theory E_G, $\varphi_1 \approx_{E_\mathsf{G}} \varphi_2$. As we suppose $\approx_{E_\mathsf{G}}$-soundness, we have that the concrete semantics of those two frames cannot be distinguished, *i.e.* $([\![\varphi_1]\!]_{A_\eta}) \approx ([\![\varphi_2]\!]_{A_\eta})$. Hence, $\mathcal{A}$ cannot solve the DDH problem.

Clearly, if one can solve the DL problem, one can also solve the CDH problem, which itself allows us to solve the DDH problem. Therefore, the hardness of DDH implies the hardness of the two other problems.

In a similar way we show that $\approx_{E_\mathsf{G}}$-soundness implies the hardness of RSA. Instead of directly encoding the RSA problem, we introduce a slightly weaker decision problem, whose hardness implies the hardness of RSA. The encoding of this problem requires the extension of the signature by a one-way function $\mathsf{h} : G \to G$, adding no equation to the theory. Consider the two frames

$$\varphi_1 = \nu g, a.\{x_1 = g^a, x_2 = a, x_3 = \mathsf{h}(g)\} \text{ and } \varphi_2 = \nu g, g', a.\{x_1 = g^a, x_2 = a, x_3 = g'\}.$$

We have that $\varphi_1 \approx_{E_\mathsf{G}} \varphi_2$. As above, if we suppose $\approx_{E_\mathsf{G}}$-soundness of (A_η), we have that the RSA problem cannot be solved by a probabilistic polynomial-time adversary.

An interesting open question is whether $\approx_{E_\mathsf{G}}$-soundness implies or is implied by Rivest's notion of pseudo-free groups [17]. We conjecture that the two notions are in fact incomparable. Indeed, on the one hand, our notion implies the hardness of DDH, which remains an open question for pseudo-free groups. On the other hand pseudo-freeness deals with a form of adaptive attackers while our model is purely non-adaptive.

4 A sufficient (and often necessary) criterion for $\approx_E$-soundness

We now present useful results for proving $\approx_E$-soundness properties in general. Notably, we provide a sufficient criterion for $\approx_E$-soundness in Section 4.1 and prove it necessary under additional assumptions in Section 4.2.

4.1 Ideal semantics and $\approx_E$-soundness criterion

Given an implementation of the primitives, what we called the concrete semantics maps every closed frame φ to a distribution $[\![\varphi]\!]_{A_\eta}$ in the expected way. We now define the *ideal semantics* of a φ, intuitively as the uniform distribution over sequences of bit-strings (in the appropriate space) that pass all the formal tests verified by φ.

Given a closed frame φ, let us write $\mathrm{eq}_E(\varphi)$ for the set of tests that are true in φ: $\mathrm{eq}_E(\varphi) = \{(M,N) \mid \mathrm{var}(M) \cup \mathrm{var}(N) \subseteq \mathrm{dom}(\phi), (\mathrm{names}(M) \cup \mathrm{names}(N)) \cap \mathrm{names}(\varphi) = \emptyset$ and $M\varphi =_E N\varphi\}$. Notice that $\varphi \approx_E \varphi'$ iff $\mathrm{eq}_E(\varphi) = \mathrm{eq}_E(\varphi')$.

We say that (A_η) *has uniform distributions* iff for every η and every sort s, $[\![s]\!]_{A_\eta}$ is a finite set, $=_{A_\eta,s}$ is the usual equality and the distribution associated to s by A_η is the uniform one over $[\![s]\!]_{A_\eta}$.

Definition 1 (Ideal semantics). *Let (A_η) be an unconditionally $=_E$-sound family of computational algebras, having uniform distributions. Let $\varphi = \{x_1 = t_1, \ldots, x_n = t_n\}$ be a closed frame and s_i the sort of x_i. The* ideal semantics $[\![\varphi]\!]^{ideal}_{A_\eta}$ *of φ is the uniform distribution over the finite (non-empty) set of concrete frames:*

$$\Big\{\{x_1 = e_1, \ldots, x_n = e_n\} \;\mid\; (e_1, \ldots, e_n) \in [\![s_1]\!]_{A_\eta} \times \cdots \times [\![s_n]\!]_{A_\eta} \;\; \textit{and} \\ \forall (M,N) \in \mathrm{eq}_E(\varphi) \cdot [\![M]\!]_{A_\eta,\{x_1=e_1,\ldots,x_n=e_n\}} = [\![N]\!]_{A_\eta,\{x_1=e_1,\ldots,x_n=e_n\}}\Big\}$$

For instance, let $\varphi = \nu n_1, n_2.\{x_1 = n_1, x_2 = n_2\}$ with n_1 and n_2 of sort s. Then $\mathrm{eq}_E(\varphi) \subseteq \{(M,N) \mid M =_E N\}$ implies that $[\![\varphi]\!]^{ideal}_{A_\eta}$ is simply the uniform distribution over $[\![s]\!]_{A_\eta} \times [\![s]\!]_{A_\eta}$. A more general definition of the ideal semantics, which does not restrict (A_η) to have only uniform distributions is given in Appendix C.

We can now state our $\approx_E$-soundness criterion: intuitively, the two semantics, concrete and ideal, should be indistinguishable.

Theorem 1 ($\approx_E$-soundness criterion). *Let (A_η) be an unconditionally $=_E$-sound family of computational algebras. Assume that for every closed frame φ it holds that $([\![\varphi]\!]_{A_\eta}) \approx ([\![\varphi]\!]^{ideal}_{A_\eta})$. Then (A_η) is $\approx_E$-sound.*

Proof. Let $\varphi_1 \approx_E \varphi_2$. As $\mathrm{eq}_E(\varphi_1) = \mathrm{eq}_E(\varphi_2)$, for every η, by construction, the distributions $[\![\varphi_1]\!]^{ideal}_{A_\eta}$ and $[\![\varphi_2]\!]^{ideal}_{A_\eta}$ are equal. We use transitivity of the indistinguishability relation $\approx$ to conclude: $([\![\varphi_1]\!]_{A_\eta}) \approx ([\![\varphi_1]\!]^{ideal}_{A_\eta}) = ([\![\varphi_2]\!]^{ideal}_{A_\eta}) \approx ([\![\varphi_2]\!]_{A_\eta})$. □

4.2 Patterns revisited

Patterns have been introduced by Abadi and Rogaway [5] and used in subsequent work [15, 8] as a way to define computationally sound formal equivalences. Typically frames are mapped to patterns by replacing non-deducible subterms by boxes □. Two frames are then equivalent iff they yield the same pattern (up to renaming of names). For example, the pattern associated to the frame $\varphi_1 = \{x_1 = \mathsf{enc}(\mathsf{enc}(k_4, k_3), k_1),\ x_2 = \mathsf{enc}(k_1, k_2),\ x_3 = k_2\}$ is $\{x_1 = \mathsf{enc}(\square, k_1),\ x_2 = \mathsf{enc}(k_1, k_2),\ x_3 = k_2\}$.

In this section we propose a general, novel definition of patterns and study some of their properties. We then use these properties to prove that our soundness criterion is necessary in many cases.

Definition 2 (Pattern). *A closed frame φ is a* pattern *if each of its subterms is deducible from φ.*

Equivalently a pattern is a closed frame of the form $\varphi = \{x_1 = C_1[a_1, \ldots, a_m], \ldots, x_n = C_n[a_1, \ldots, a_m]\}$, where the $C_1 \ldots C_n$ are public (non necessarily linear) contexts and the $a_1 \ldots a_m$ are distinct deducible names: $\varphi \vdash_E a_i$. For example, the frame φ_1 considered above is not a pattern while $\{x_1 = \mathsf{enc}(n_1, k_1), x_2 = \mathsf{enc}(k_1, k_2), x_3 = k_2\}$ is.

The following proposition finitely characterizes the equations verified by a pattern.

Proposition 3. *Let φ be a pattern. For each a_i, let ζ_{a_i} be a public term such that $\mathrm{var}(\zeta_{a_i}) \subseteq \{x_1, \ldots, x_n\}$ and $\zeta_{a_i}\varphi =_E a_i$. Then every equation which holds in φ is a logical consequence (in the first-order theory of equality) of E and the equations $x_j = C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}]$,* i.e. *$E \cup \{x_j = C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}] \mid 1 \leq j \leq n\} \models \mathrm{eq}_E(\varphi)$.*

Interestingly the concrete and the ideal semantics of patterns often coincide.

Proposition 4. *Let (A_η) be an unconditionally $=_E$-sound family of computational algebras, having uniform distributions. Let φ be a pattern. The concrete and the ideal semantics of φ yield the same family of distributions: for all η, $[\![\varphi]\!]_{A_\eta} = [\![\varphi]\!]^{ideal}_{A_\eta}$.*

The idea of the proof (detailed in Appendix D.2) is that, using the finite characterization of $\mathrm{eq}_E(\varphi)$ (Proposition 3), one can draw a bijection between the drawing of nonces and the eligible values for the ideal semantics.

A theory E *admits patterns* iff for every closed frame φ, there exists a (not necessarily unique) pattern $\overline{\varphi}$ such that $\varphi \approx_E \overline{\varphi}$. In practice many theories useful in cryptography satisfy this property, *e.g.* the theories considered in Section 5. Note that we have proved *en passant* that $\approx_E$ is decidable for equational theories that admit patterns and for which $=_E$ is decidable, provided the construction of patterns is effective. Indeed, given two frames φ_1 and φ_2, we associate to each of them static equivalent patterns $\overline{\varphi_1}$ and $\overline{\varphi_2}$. It is then straightforward to check whether $\overline{\varphi_1}$ and $\overline{\varphi_2}$ are equivalent using the finite characterization of $\mathrm{eq}_E(\overline{\varphi}_i)$ by Proposition 3.

The following theorem states that our soundness criterion is actually very tight: whenever a theory admits patterns, our criterion is a necessary condition.

Theorem 2. *Assume that the theory E admits patterns. Let (A_η) be a family of computational algebras, such that (A_η) has uniform distributions, is $\approx_E$- and unconditionally $=_E$-sound. Then the soundness criterion of Theorem 1 is satisfied: for every closed frame φ, $(\llbracket\varphi\rrbracket_{A_\eta}) \approx (\llbracket\varphi\rrbracket^{ideal}_{A_\eta})$.*

Proof. By hypothesis, $\varphi \approx_E \overline{\varphi}$ implies $(\llbracket\varphi\rrbracket_{A_\eta}) \approx (\llbracket\overline{\varphi}\rrbracket_{A_\eta})$ and by Proposition 4, $(\llbracket\overline{\varphi}\rrbracket_{A_\eta}) = (\llbracket\overline{\varphi}\rrbracket^{ideal}_{A_\eta})$. We conclude since $\varphi \approx_E \overline{\varphi}$ implies $(\llbracket\varphi\rrbracket^{ideal}_{A_\eta}) = (\llbracket\overline{\varphi}\rrbracket^{ideal}_{A_\eta})$. □

5 Examples

We now apply the framework of Sections 3 and 4 to establish two novel $\approx_E$-soundness results, concerning the theory of Exclusive Or and that of ciphers and lists.

5.1 Exclusive Or

We study the soundness and faithfulness problems for the usual theory and implementation of the Exclusive Or (XOR).

The formal model consists of a single sort $Data$, an infinite number of names, the infix symbol $\oplus : Data \times Data \to Data$ and two constants $0, 1 : Data$. Terms are equipped with the equational theory $E_\oplus$ generated by:

$$\begin{array}{cc} x \oplus y = y \oplus x & x \oplus x = 0 \\ (x \oplus y) \oplus z = x \oplus (y \oplus z) & x \oplus 0 = x \end{array}$$

As an implementation, we define the computational algebras A_η, $\eta \geq 0$: the concrete domain $\llbracket Data\rrbracket_{A_\eta}$ is $\{0,1\}^\eta$ equipped with the uniform distribution; $\oplus$ is interpreted by the usual XOR function over $\{0,1\}^\eta$, $\llbracket 0\rrbracket_{A_\eta} = 0^\eta$, $\llbracket 1\rrbracket_{A_\eta} = 1^\eta$.

In this setting, statically equivalent frames enjoy an algebraic characterization. Indeed, let φ and φ' be two frames with $\mathrm{names}(\varphi) \cup \mathrm{names}(\varphi') \subseteq \{a_1, \dots, a_n\}$ and $\mathrm{dom}(\varphi) = \mathrm{dom}(\varphi') = \{x_1, \dots, x_m\}$. We associate to φ a $(m+1) \times (n+1)$-matrix $\alpha = (\alpha_{i,j})$ over the two element field $\mathbb{F}_2$: the 0-th row of α is $(1, 0 \dots 0)$ and for $1 \leq i \leq m$, $1 \leq j \leq n$ (*resp. j = 0*) $\alpha_{i,j}$ is the number of occurrences of a_j (*resp.* of 1)

in $\varphi(x_i)$, modulo 2. In the same way, a matrix α' is associated to φ'. Using classical manipulations on matrix, it is easy to show that $\varphi \approx_{E_\oplus} \varphi'$ iff $\mathrm{coker}(\alpha) = \mathrm{coker}(\alpha')$, where $\mathrm{coker}(\alpha)$ denotes the co-kernel of α (*i.e.* the set of rows β *s.t.* $\beta \cdot \alpha = 0$).

This characterization is the key point of our main result for the theory of XOR.

Theorem 3. *The usual implementation for the XOR theory is unconditionally $=_{E_\oplus}$-, $\approx_{E_\oplus}$- and $\not\models_{E_\oplus}$-sound. It is also $=_{E_\oplus}$-, $\approx_{E_\oplus}$- and $\not\models_{E_\oplus}$-faithful.*

The proof is detailed in Appendix E. We show that the concrete and ideal semantics coincide using the duality property $\mathrm{im}(\alpha) = \mathrm{coker}(\alpha)^\perp$. This result is comparable to the work of Bana [8], who shows the unconditional soundness of the One-Time Pad encryption in a setting similar to that of Abadi and Rogaway [5]. In some sense our result is more precise as we model the XOR symbol itself and not a particular use of it.

5.2 Symmetric, deterministic, length-preserving encryption and lists

We now detail the example of symmetric, deterministic and length-preserving encryption schemes. Such schemes, also known as *ciphers* [16], are widely used in practice, the most famous examples being DES and AES .

Our formal model consists of the set of sorts $\mathcal{S} = \{Data, List_0, List_1 \dots List_n \dots\}$, an infinite number of names for each sort and the symbols:

$\mathsf{enc}_n, \mathsf{dec}_n$	$: List_n \times Data \to List_n$	encryption, decryption
cons_n	$: Data \times List_n \to List_{n+1}$	list constructor
head_n	$: List_{n+1} \to Data$	head of a list
tail_n	$: List_{n+1} \to List_n$	tail of a list
nil	$: List_0 \qquad 0, 1 : Data$	empty list, constants

We consider the equational theory E_{sym} generated by (for every $n \geq 0$)

$$\begin{array}{rr}
\mathsf{dec}_n(\mathsf{enc}_n(x, y), y) = x & \mathsf{cons}_n(\mathsf{head}_n(x), \mathsf{tail}_n(x)) = x \\
\mathsf{enc}_n(\mathsf{dec}_n(x, y), y) = x & \mathsf{enc}_0(\mathsf{nil}, x) = \mathsf{nil} \\
\mathsf{head}_n(\mathsf{cons}_n(x, y)) = x & \mathsf{dec}_0(\mathsf{nil}, x) = \mathsf{nil} \\
\mathsf{tail}_n(\mathsf{cons}_n(x, y)) = y &
\end{array}$$

When oriented from left to right, the equations E_{sym} form an (infinite) convergent rewriting system, written $\mathcal{R}$. The concrete meaning of sorts and symbols is given by the computational algebras A_η, $\eta > 0$, defined as follows:

- the carrying sets are $[\![Data]\!]_{A_\eta} = \{0, 1\}^\eta$ and $[\![List_n]\!]_{A_\eta} = \{0, 1\}^{n\eta}$ equipped with the uniform distribution and the usual equality relation;
- $\mathsf{enc}_n, \mathsf{dec}_n$ are implemented by a cipher for data of size $n\eta$ and keys of size η (we discuss the required cryptographic assumptions later);
- $[\![\mathsf{nil}]\!]_{A_\eta}$ is the empty bit-string, $[\![\mathsf{cons}_n]\!]_{A_\eta}$ is the usual concatenation, $[\![0]\!]_{A_\eta} = 0^\eta$, $[\![1]\!]_{A_\eta} = 1^\eta$, $[\![\mathsf{head}_n]\!]_{A_\eta}$ returns the η first digits of bit-strings (of size $(n+1)\eta$) whereas $[\![\mathsf{tail}_n]\!]_{A_\eta}$ returns the last $n\eta$ digits.

We emphasize that no tags are added to messages. Tags—and in particular tags under encryption— would be harmful to the $\approx_{E_{\mathsf{sym}}}$-soundness. Indeed we expect that the formal equivalence $\nu a, b.\{x = enc(a, b),\ y = b\} \approx_{E_{\mathsf{sym}}} \nu a, b, c.\{x = enc(a, b),\ y = c\}$ also holds in the computational world; but this will not be the case if a is tagged before encryption.

For simplicity we assume that the encryption keys have the same size η as blocks of data. This is not a real restriction since smaller keys can always be padded with random digits and the additional digits ignored by the encryption algorithm. We also assume that keys are generated according to a uniform distribution. (This is the case for AES, and also for DES if we restrict keys to their 56 significant bits.)

Obviously, the above implementation is unconditionally $=_{E_{\mathsf{sym}}}$-sound. Before studying the $\approx_{E_{\mathsf{sym}}}$-soundness, we need to characterize statically equivalent frames. Specifically we show that this theory admits patterns, in the sense of Section 3.

Proposition 5. *Let φ be a closed frame. There exists a pattern $\overline{\varphi}$ such that $\varphi \approx_{E_{\mathsf{sym}}} \overline{\varphi}$.*

Proof (outline). We associate a pattern to any frame φ by the following procedure:

1. normalize φ using the rules $\mathcal{R}$ (the result is still denoted φ);
2. while φ is not a pattern, repeat: find any subterm T of the form $T = \mathsf{enc}_n(U, V)$, $T = \mathsf{dec}_n(U, V)$, $T = \mathsf{head}_n(V)$ or $T = \mathsf{tail}_n(V)$, with $\varphi \not\vdash_{E_{\mathsf{sym}}} V$ and replace T everywhere in φ by a fresh name a of the appropriate sort.

We prove in Appendix F.1 that this procedure always terminates on a pattern statically equivalent to the initial frame.

Notice that for any subterm W, $\varphi \not\vdash_{E_{\mathsf{sym}}} W$ implies $\varphi\{T \mapsto n\} \not\vdash_{E_{\mathsf{sym}}} W\{T \mapsto n\}$, where $\{T \mapsto n\}$ denotes the replacement of T by n. As a consequence, the procedure above yields a unique pattern (modulo renaming), no matter in what order the subterms T are replaced.

Provided that $\vdash_{E_{\mathsf{sym}}}$ is decidable, the procedure for associating patterns to frame is effective. Thus, as noticed in Section 4.2, we obtain another proof of the decidability of $\approx_{E_{\mathsf{sym}}}$ using Proposition 3. Notice that statically equivalent patterns may *not* be equal modulo renaming: consider *e.g.* $\{x = enc(a, b),\ y = b\} \approx_{E_{\mathsf{sym}}} \{x = c,\ y = b\}$.

We now study the $\approx_{E_{\mathsf{sym}}}$-soundness problem under realistic cryptographic assumptions. Classical assumptions on ciphers include the notions of super pseudo-random permutation (SPRP) and indistinguishability against lunchtime or adaptive, chosen-plaintext or chosen-ciphertext attacks (written IND-Pi-Cj, $i, j \in \{0, 1, 2\}$ depending on the different combinations). These notions and the relations between them have been studied notably in [16].

Initially, the SPRP and IND-Pi-Cj assumptions apply to ciphers specialized to plaintexts of a given size. Interestingly, this is not sufficient to imply the $\approx_{E_{\mathsf{sym}}}$-soundness for frames which contain plaintexts of heterogeneous sizes, encrypted under the same key. Thus we introduce strengthened versions of these assumptions, applying to a *collection* of ciphers $(\mathcal{E}_{\eta,n}, \mathcal{D}_{\eta,n})$, where η is the complexity parameter and $n \geq 0$ is the number of blocks of size η contained in plaintexts and ciphertexts.

We define the ω-IND-Pi-Cj assumption $i, j \in \{0, 1, 2\}$, by considering the following experience $\mathcal{G}_\eta^{ij}$ involving a 2-stage adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$:

- first a key k is randomly chosen from $\{0,1\}^\eta$;
- (Stage 1) $\mathcal{A}_1$ is given access—if $i \geq 1$—to the encryption oracles $\mathcal{E}_{\eta,n}(\cdot, k)$, and—if $j \geq 1$—to the decryption oracles $\mathcal{D}_{\eta,n}(\cdot, k)$. $\mathcal{A}_1$ outputs two plaintexts $m_0, m_1 \in \{0,1\}^{n_0\eta}$ for some n_0, and possibly some data d;
- (Stage 2) a random bit $b \in \{0,1\}$ is drawn. $\mathcal{A}_2$ receives the *challenge ciphertext* $c = \mathcal{E}_{\eta,n_0}(m_b, k)$, the data d and is given access—if $i \geq 2$—to the encryption oracles $\mathcal{E}_{\eta,n}(\cdot, k)$, and—if $j \geq 2$—to the decryption oracles $\mathcal{D}_{\eta,n}(\cdot, k)$. $\mathcal{A}_2$ then outputs a bit b';
- $\mathcal{A}$ *is successful in* $\mathcal{G}_\eta$ iff $b = b'$ and, during the two stages, it has not submitted m_0 or m_1 to an encryption oracle, nor c to a decryption oracle.

Define the *advantage* of $\mathcal{A}$ as: $\mathrm{Adv}_{\mathcal{A}}^{\omega\text{-IND-P}i\text{-C}j}(\eta) = 2 \times \mathbb{P}\left[\mathcal{A} \text{ is successful in } \mathcal{G}_\eta^{ij}\right] - 1$. *The ω-IND-Pi-Cj assumption holds for* $(\mathcal{E}_{\eta,n}, \mathcal{D}_{\eta,n})$ iff the advantage of any probabilistic polynomial-time adversary is negligible. It holds for the *inverse* of the encryption scheme, iff it holds for the collection of ciphers $(\mathcal{D}_{\eta,n}, \mathcal{E}_{\eta,n})$.

The usual IND-Pi-Cj assumption corresponds to the case where n is restricted to the value 1 in the above definition. A similar strengthening of the SPRP assumption, written ω-SPRP, is proposed in Appendix F.2.

As in previous work [5, 15, 6, 14], we restrict frames to those with only atomic keys and no encryption cycles. Specifically a closed frame φ *has only atomic keys* if for all subterms $\mathsf{enc}_n(u, v)$ and $\mathsf{dec}_n(u, v)$ of φ, v is a name. Given two (atomic) keys k_1 and k_2, we say that k_1 *encrypts* k_2 *in* φ, written $k_1 >_\varphi k_2$, iff there exists a subterm U of φ of the form $U = \mathsf{enc}_n(T, k_1)$ or $U = \mathsf{dec}_n(T, k_1)$ such that k_2 appears in T *not used as a key, i.e.* k_2 appears in T at a position which is not the right-hand argument of a $\mathsf{enc}_{n'}$ or a $\mathsf{dec}_{n'}$. An *encryption cycle* is a tuple $k_1 \ldots k_m$ such that $k_1 >_\varphi \ldots >_\varphi k_m >_\varphi k_1$.

The effect of the condition “not used as a key” is to allow considering more terms as free of encryption cycles, for instance $\mathsf{enc}_n(\mathsf{enc}_n(a, k), k)$. This improvement is already suggested in [5].

We now state our $\approx_{E_{\mathsf{sym}}}$-soundness theorem. A closed frame is *well-formed* iff its $\mathcal{R}$-normal form has only atomic keys, contains no encryption cycles and uses no head and tail symbols.

Theorem 4 **($\approx_{E_{\mathsf{sym}}}$-soundness).** *Let φ_1 and φ_2 be two well-formed frames of the same domain. Assume that the concrete implementations for the encryption and its inverse satisfy both the ω-IND-P1-C1 assumption. If $\varphi_1 \approx_{E_{\mathsf{sym}}} \varphi_2$ then $([\![\varphi_1]\!]_{A_\eta}) \approx ([\![\varphi_2]\!]_{A_\eta})$.*

The proof is detailed in Appendix F.3. The idea is to prove the computational soundness of each step of the procedure for mapping frames to patterns (Proposition 5). We conclude using Proposition 4 on the ideal semantics for patterns.

Note on the cryptographic assumptions. Cryptographic assumptions of Theorem 4 may appear strong compared to existing work on passive adversaries [5, 15]. This seems unavoidable when we allow frames to contain both encryption and decryption symbols. Nevertheless if φ_1 and φ_2 contain no decryption symbols, our proofs are easily adapted to work when the encryption scheme is ω-IND-P1-C0 only.

Also, it is possible to recover the classical assumptions IND-Pi-Cj by modeling the ECB mode (Electronic Code Book). Let us add two symbols $\mathsf{enc} : \mathit{Data} \times \mathit{Data} \rightarrow$

$Data$ and $\mathsf{dec} : Data \times Data \to Data$, and define the symbols enc_n and dec_n (formally and concretely) recursively by

$$\mathsf{enc}_{n+1}(x,y) = \mathsf{cons}_n(\mathsf{enc}(\mathsf{head}_n(x),y), \mathsf{enc}_n(\mathsf{tail}_n(x),y)) \quad \text{and}$$
$$\mathsf{dec}_{n+1}(x,y) = \mathsf{cons}_n(\mathsf{dec}(\mathsf{head}_n(x),y), \mathsf{dec}_n(\mathsf{tail}_n(x),y)).$$

Define well-formed frames as those of which the normal forms contain no encryption cycles. The $\approx_{E_{\mathsf{sym}}}$-soundness property for well-formed frames holds as soon as the implementations for enc and dec are both IND-P1-C1, or equivalently [16] enc is SPRP.

6 Conclusion and future work

In this paper we developed a general framework for relating formal and computational models of security protocols in the presence of a passive attacker. These are the first results on abstract models allowing arbitrary equational theories. We define the soundness and faithfulness of cryptographic implementations *w.r.t.* abstract models. We also provide a soundness criterion which for a large number of theories—those that admit a general notion of patterns—is not only sufficient but also necessary. Finally, we provide new soundness results for the Exclusive Or, as well as a theory of ciphers and lists.

As future work, we foresee to study the soundness of other theories. An interesting case would be the combination of the two theories considered in this paper: in a theory combining XOR, ciphers and lists, one can precisely model *cipher block chaining* commonly used with ciphers such as DES or AES. Another ambitious extension is to consider the case of an active attacker.

References

1. M. Abadi, B. Blanchet, and C. Fournet. Just fast keying in the pi calculus. In *Proc. 13th European Symposium on Programming (ESOP'04)*, volume 2986 of *LNCS*, pages 340–354, 2004.
2. M. Abadi and V. Cortier. Deciding knowledge in security protocols under equational theories. In *Proc. 31st International Colloquium on Automata, Languages and Programming (ICALP'04)*, volume 3142 of *LNCS*, pages 46–58, 2004.
3. M. Abadi and C. Fournet. Mobile values, new names, and secure communications. In *Proc. 28th Annual ACM Symposium on Principles of Programming Languages (POPL'01)*, pages 104–115, 2001.
4. M. Abadi and A. D. Gordon. A calculus for cryptographic protocols: The Spi calculus. In *Proc. 4th ACM Conference on Computer and Communications Security (CCS'97)*, pages 36–47, 1997.
5. M. Abadi and P. Rogaway. Reconciling two views of cryptography (the computational soundness of formal encryption). In *Proc. 1st IFIP International Conference on Theoretical Computer Science (IFIP–TCS'00)*, volume 1872 of *LNCS*, pages 3–22, 2000.
6. M. Backes and B. Pfitzmann. Symmetric encryption in a simulatable Dolev-Yao style cryptographic library. In *Proc. 17th IEEE Computer Science Foundations Workshop (CSFW'04)*, pages 204–218, 2004.

7. M. Backes, B. Pfitzmann, and M. Waidner. A composable cryptographic library with nested operations. In *Proc. 10th ACM Conference on Computer and Communications Security (CCS'03)*, 2003.
8. G. Bana. *Soundness and Completeness of Formal Logics of Symmetric Encryption*. PhD thesis, University of Pennsylvania, 2004.
9. K. Bhargavan, C. Fournet, A. D. Gordon, and R. Pucella. Tulafale: A security tool for web services. In *Proc. International Symposium on Formal Methods for Components and Objects (FMCO'03)*, volume 3188 of *LNCS*, pages 197–222, 2004.
10. B. Blanchet. An efficient cryptographic protocol verifier based on Prolog rules. In *Proc. 14th IEEE Computer Security Foundations Workshop (CSFW'01)*, pages 82–96, 2001.
11. B. Blanchet. Automatic proof of strong secrecy for security protocols. In *Proc. 25th IEEE Symposium on Security and Privacy (SSP'04)*, pages 86–100, 2004.
12. D. Dolev and A. C. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, IT-29(12):198–208, 1983.
13. S. Goldwasser and S. Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28:270–299, 1984.
14. P. Laud. Symmetric encryption in automatic analyses for confidentiality against active adversaries. In *Proc. IEEE Symposium on Security and Privacy (SSP'04)*, pages 71–85, 2004.
15. D. Micciancio and B. Warinschi. Completeness theorems for the Abadi-Rogaway logic of encrypted expressions. *Journal of Computer Security*, 12(1):99–129, 2004.
16. D. H. Phan and D. Pointcheval. About the security of ciphers (semantic security and pseudo-random permutations). In *Proc. Selected Areas in Cryptography (SAC'04)*, volume 3357 of *LNCS*, pages 185–200, 2004.
17. R. L. Rivest. On the notion of pseudo-free groups. In *Proc. 1st Theory of Cryptography Conference (TCC'04)*, volume 2951 of *LNCS*, pages 505–521, 2004.

A Soundness and faithfulness

A.1 Proof of Proposition 1

Proposition 1. *Let (A_η) be a $=_E$-sound family of computational algebras. Then*

1. *(A_η) is $\not\vdash_E$-faithful;*
2. *if (A_η) is also $=_E$-faithful, (A_η) is $\approx_E$-faithful.*

Proof.

1. Suppose $\varphi \vdash_E T$ *i.e.* there exists M such that $M\varphi =_E T$ and $\mathrm{var}(M) \subseteq \mathrm{dom}(\varphi)$ and the names of M do not occur neither in φ nor in T. We define the adversary $\mathcal{A}$ which can deduce $[\![T]\!]$ from $[\![\varphi]\!]$ as follows: $\mathcal{A}(\{x_i = e_i\}) = [\![M]\!]_{A_\eta,\{x_i=e_i\}}$. As $(A_\eta)_{\eta\geq 0}$ is $=_E$-sound, $\mathcal{A}$'s success probability is greater than 1 minus a negligible function.
2. Suppose $\varphi_1 \not\approx_E \varphi_2$: there exists a test M, N such that (for instance) $M\varphi_1 =_E N\varphi_1$ and $M\varphi_2 \neq_E N\varphi_2$. Let $\mathcal{A}$ be an adversary that, given η and $\hat{\psi}$, tests whether $[\![M]\!]_{A_\eta,\hat{\psi}} =_{A_\eta} [\![N]\!]_{A_\eta,\hat{\psi}}$ and returns the result of the test. $\mathcal{A}$ runs in polynomial-time and by hypothesis its advantage is 1 minus a negligible function. □

A.2 Proof of Proposition 2

Proposition 2. *Let (A_η) be a family of $\approx_E$-sound computational algebras. Assume that free binary symbols $\mathsf{h}_s : s \times \mathit{Key} \to \mathit{Hash}$ are available for every sort s. Then*

1. *(A_η) is $=_E$-faithful;*
2. *(A_η) is $\not\vdash_E$-sound;*
3. *if the implementations for the h_s are also collision-resistant, then (A_η) is $=_E$-sound, $\not\vdash_E$-faithful and $\approx_E$-faithful.*

Proof.

1. Let T_1, T_2 be two terms of sort s such that $T_1 \neq_E T_2$. Consider the frame $\varphi = \{x_1 = \mathsf{h}_s(T_1, k),\ x_2 = \mathsf{h}_s(T_2, k)\}$ where k is a fresh name of sort Key. As $T_1 \neq_E T_2$ formally and h_s has no equations, we have $\varphi \approx_E \{x_1 = n,\, x_2 = n'\}$ where n, n' are two distinct fresh names of sort Hash. Thus by hypothesis: $[\![\varphi]\!] \approx [\![\{x_1 = n,\, x_2 = n'\}]\!]$. The probability for the two names to be equal concretely is negligible, thus

$$\begin{aligned}\mathbb{P}\left[\, e_1, e_2 \leftarrow [\![T_1,\, T_2]\!]_{A_\eta};\ e_1 =_{A_\eta} e_2 \right] \\ \leq \mathbb{P}\left[\, e'_1, e'_2 \leftarrow [\![\mathsf{h}(T_1, k),\, \mathsf{h}(T_2, k)]\!]_{A_\eta};\ e'_1 =_{A_\eta} e'_2 \right]\end{aligned}$$

   is negligible.
2. Let φ be a frame and T a term of sort s. We let $\varphi_1 = \varphi \cup \{x = \mathsf{h}_s(T, k), y = k\}$ and $\varphi_2 = \varphi \cup \{x = n, y = k\}$ where x, y are fresh variables, k is a fresh name of sort Key, n is a fresh name of sort Hash. As $\varphi \not\vdash_E T$, we have $\varphi_1 \approx_E \varphi_2$. Thus by hypothesis, $[\![\varphi_1]\!] \approx_E [\![\varphi_2]\!]$. By contradiction, suppose there exists a polynomial-time adversary $\mathcal{A}$ to deduce $[\![T]\!]$ from $[\![\varphi]\!]$ concretely. We easily build an adversary $\mathcal{B}$ to distinguish between $[\![\varphi_1]\!]$, $[\![\varphi_2]\!]$ as follows: call $\mathcal{A}$ on the first part of φ_b and obtain $[\![T]\!]$. Using $y = [\![k]\!]$, compute $[\![\mathsf{h}_s(T, k)]\!]$ and compare it to the value of x. Hence, we contradict the hypothesis of $\approx_E$-soundness and conclude that no probabilistic polynomial-time adversary can deduce $[\![t]\!]$ from $[\![\varphi]\!]$ with significant probability, *i.e.* $(A_\eta)_\eta \geq 0$ is $\not\vdash_E$-sound.
3. Let T_1, T_2 be two terms of sort s such that $T_1 =_E T_2$. Consider the same frame as before: $\varphi = \{x_1 = \mathsf{h}_s(T_1, k),\ x_2 = \mathsf{h}_s(T_2, k)\}$. As $T_1 =_E T_2$ formally and h_s has no equations, we have $\varphi \approx_E \{x_1 = n,\, x_2 = n\}$ where n is a fresh name of sort Hash. Thus by hypothesis: $[\![\varphi]\!] \approx [\![\{x_1 = n,\, x_2 = n\}]\!]$ and

$$\mathbb{P}\left[\, e'_1, e'_2 \leftarrow [\![\mathsf{h}(T_1, k),\, \mathsf{h}(T_2, k)]\!]_{A_\eta};\ e'_1 =_{A_\eta} e'_2 \right] \geq 1 - \epsilon_\eta$$

   where ϵ_η is a negligible function. Thus if the implementation of h_s is collision-resistant,

$$\mathbb{P}\left[\, e_1, e_2 \leftarrow [\![T_1,\, T_2]\!]_{A_\eta};\ e_1 \neq_{A_\eta} e_2 \right]$$

   is negligible. Other properties follow from Proposition 1. □

B An abstract model of groups

We model a group G with exponents taken over a commutative ring A and define the following symbols:

$$\begin{array}{rlrl} * & : G \times G \to G & - & : A \to A \\ 1_G & : G & \cdot & : A \times A \to A \\ + & : A \times A \to A & 1_A & : A \\ 0 & : A & \mathsf{exp} & : G \times A \to G \end{array}$$

As usual, we use infix notation to denote the operators $*$, $\cdot$, $+$ and write g^a to denote $\mathsf{exp}(g, a)$. Consider the following equational theory E_{G}:

$$\begin{array}{ccc} x * 1_G = x & x + (-x) = 0 & (x + y) \cdot z = x \cdot z + y \cdot z \\ 1_G * x = x & x + (y + z) = (x + y) + z & (x^a)^b = x^{(a \cdot b)} \\ x * (y * z) = (x * y) * z & x \cdot 1_A = x & x^a * x^b = x^{a+b} \\ x + 0 = x & x \cdot y = y \cdot x & x^{1_A} = x \\ x + y = y + x & x \cdot (y \cdot z) = (x \cdot y) \cdot z & x^0 = 1_G \end{array}$$

Abelian groups are modeled by adding the equations $x*y = y*x$ and $(x*y)^a = x^a*y^a$.

C Ideal semantics

Definition 3 (Ideal semantics, general case). *Let (A_η) be a family of unconditionally $=_E$-sound computational algebras and $\varphi = \{x_1 = T_1, \ldots, x_n = T_n\}$ a closed frame.*

The ideal semantics *of φ is the family of distributions $(\llbracket\varphi\rrbracket^{ideal}_{A_\eta})$, where for each η, the drawing $\phi \leftarrow \llbracket\varphi\rrbracket^{ideal}_{A_\eta}$ is defined as follows:*

1. *for each i, draw $e_i \overset{R}{\longleftarrow} |T_i|_{A_\eta}$,*
2. *check that for all $(M, N) \in \mathrm{eq}_E(\varphi)$,*

$$|M|_{A_\eta,\{x_1=e_1,\ldots,x_n=e_n\}} =_{A_\eta} |N|_{A_\eta,\{x_1=e_1,\ldots,x_n=e_n\}}$$

3. *if this is the case, return $\phi = \{x_1 = e_1, \ldots, x_n = e_n\}$, otherwise go back to step 1.*

By unconditional $=_E$-soundness, the drawings ϕ which are ($=_{A_\eta}$-equivalent to values) in the image of the usual semantics of φ always pass the test of step 2. As we required that the drawing from $|T_i|_{A_\eta}$ gives no $=_{A_\eta}$-equivalent class a zero probability, the probability to succeed at each step 2 is thus greater than zero. Hence the loop "eventually" terminates and the definition makes sense. (Rigorously, $\llbracket\varphi\rrbracket^{ideal}_{A_\eta}$ is a conditional distribution.)

As $\mathrm{eq}_E(\varphi)$ is likely to be infinite, the definition may not be effective; this has no consequence here. Recent work of Abadi and Cortier [2] and results of Section 4 suggest that $\mathrm{eq}_E(\varphi)$ can be described finitely in a way that makes this definition effective, for many equational theories E.

D Patterns

D.1 Proof of Proposition 3

Proposition 3. *Let φ be a pattern. For each a_i, let ζ_{a_i} be a public term such that* $\mathrm{var}(\zeta_{a_i}) \subseteq \{x_1, \ldots, x_n\}$ *and $\zeta_{a_i}\varphi =_E a_i$. Then every equation which holds in φ is a logical consequence of E and the equations $x_j = C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}]$*, i.e.

$$\{x_j = C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}] \mid 1 \leq j \leq n\} \cup E \models \mathrm{eq}_E(\varphi).$$

By logical consequence, we mean that the equations are deducible from the hypotheses in the first-order theory of equality.

Proof. Let $(M = N) \in \mathrm{eq}_E(\varphi)$. By definition, we have $M\varphi =_E N\varphi$, *i.e.* $M[x_j \mapsto C_j[a_1, \ldots, a_m]] =_E N[x_j \mapsto C_j[a_1, \ldots, a_m]]$. Since E is stable by substitution of names, we get:

$$M[x_j \mapsto C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}]] =_E N[x_j \mapsto C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}]]$$

Using the equalities $x_j = C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}]$ and by transitivity, we obtain $\{x_j = C_j[\zeta_{a_1}, \ldots, \zeta_{a_m}] \mid 1 \leq j \leq n\} \cup E \models M = N$. □

D.2 Proof of Proposition 4

Proposition 4. *Let (A_η) be an unconditionally $=_E$-sound family of computational algebras having only uniform distributions. Let φ be a pattern. The concrete and the ideal semantics of φ yield the same family of distributions: for all η, $[\![\varphi]\!]_{A_\eta} = [\![\varphi]\!]^{ideal}_{A_\eta}$.*

Proof. Let $\varphi = \{x_1 = C_1[a_1, \ldots, a_m], \ldots, x_n = C_n[a_1, \ldots, a_m]\}$, with $\varphi\zeta_i =_E a_i$, $1 \leq i \leq m$ as above. Let s_i be the sort of a_i, s'_j be the sort of x_j, , and η a given complexity parameter.

For simplicity let us fix the order of variables $x_1, \ldots, x_n$ and see concrete frames merely as tuples of bit-strings. The concrete values for φ are then taken from the set:

$$F = [\![s'_1]\!]_{A_\eta} \times \cdots \times [\![s'_n]\!]_{A_\eta}$$

More precisely, the usual concrete semantics consists in mapping every drawing of names from the set $E = [\![s_1]\!]_{A_\eta} \times \cdots \times [\![s_m]\!]_{A_\eta}$ to a value in F. Let us note $\alpha : E \to F$ this function, defined by:

$$\begin{aligned} &\alpha(e_1, \ldots, e_m) \\ &= \Big([\![C_1[a_1, \ldots, a_m]]\!]_{\{a_1=e_1,\ldots,a_m=e_m\}}, \ldots, [\![C_n[a_1, \ldots, a_m]]\!]_{\{a_1=e_1,\ldots,a_m=e_m\}}\Big) \end{aligned}$$

Using the ζ_i, we can also define a function $\beta : F \to E$:

$$\beta(f_1, \ldots, f_n) = \Big([\![\zeta_1]\!]_{\{x_1=f_1,\ldots,x_n=f_n\}}, \ldots, [\![\zeta_m]\!]_{\{x_1=f_1,\ldots,x_n=f_n\}}\Big)$$

As $\varphi\zeta_i =_E a_i$ and (A_η) is unconditionally $=_E$-sound (and the distribution over E forgets no element), we have by construction: $\beta \circ \alpha = Id_E$. Thus α is injective and yields a bijection from E to its image $G = \alpha(E)$. Moreover G satisfies:

$$\begin{aligned}
G &= \{(f_1, \ldots, f_n) \mid \alpha(\beta(f_1, \ldots, f_n)) = (f_1, \ldots, f_n)\} \\
&= \{(f_1, \ldots, f_n) \mid \forall j, [\![C_j[a_1, \ldots, a_m]]\!]_{\{a_1=e_1,\ldots,a_m=e_m\}} = f_j \\
&\quad \text{where } e_i = [\![\zeta_i]\!]_{\{x_1=f_1,\ldots,x_n=f_n\}}\} \\
&= \{(f_1, \ldots, f_n) \mid \forall j, [\![C_j[\zeta_1, \ldots, \zeta_m]]\!]_{\{x_1=f_1,\ldots,x_n=f_n\}} = f_j\}
\end{aligned}$$

As φ is a pattern, by Proposition 3, $\mathrm{eq}_E(\varphi)$ is implied by the equations $C_j[\zeta_1, \ldots, \zeta_m] = x_j$ and E. By hypothesis equations in E are true in the concrete world with probability 1, thus G is precisely the set of values that pass all the tests in $\mathrm{eq}_E(\varphi)$.

Now recall that, by hypothesis, E and F are equipped with uniform distributions. Hence the concrete and the ideal semantics yield the same distribution for φ and A_η, that is the uniform distribution over G. □

E Soundness and faithfulness results for the theory of XOR

To begin with, let us detail the algebraic characterization of $\approx_{E_\oplus}$. We use the last two equations as a rewriting system

$$\begin{array}{c} x \oplus x \to 0 \\ x \oplus 0 \to x \end{array}$$

where we allow arbitrary AC-manipulations before (and after) each rewriting step. It is easy to show that this rewriting system is (AC-)convergent *i.e.* each term yields a unique (modulo AC) normal form. Specifically, a term T is in normal form iff each name or variable occurs at most once in T and either $T = 0$ or 0 does not occur in T.

Let $a_1 \ldots a_n$ be distinct names. Using the equations of XOR, each closed term T with $\mathrm{names}(T) \subseteq \{a_1 \ldots a_n\}$ can be written: $T =_{E_\oplus} \beta_0 \oplus \bigoplus_{j=1}^{n} \beta_j \, a_j$ where $\beta_j \in \{0, 1\}$ and we use the convention $0a_i = 0$ and $1a_i = a_i$. In the following, we see $\{0, 1\}$ as the two-element field $\mathbb{F}_2$; thus terms modulo $=_{E_\oplus}$ form a $\mathbb{F}_2$-vector space.

Similarly a closed frame φ with $\mathrm{names}(\varphi) \subseteq \{a_1 \ldots a_n\}$ is written

$$\varphi =_{E_\oplus} \left\{ x_1 = \alpha_{1,0} \oplus \bigoplus_{j=1}^{n} \alpha_{1,j} \, a_j \ , \ \ldots \ , \ x_m = \alpha_{m,0} \oplus \bigoplus_{j=1}^{n} \alpha_{m,j} \, a_j \right\}$$

where $\alpha_{i,j} \in \mathbb{F}_2$. Let us group the coefficients into a $(m+1) \times (n+1)$-matrix $\alpha = (\alpha_{i,j})$ over $\mathbb{F}_2$. φ is described by the relation:

$$\begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_m \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & \ldots & 0 \\ \alpha_{1,0} & \alpha_{1,1} & \ldots & \alpha_{1,n} \\ \vdots & & & \vdots \\ \alpha_{m,0} & \alpha_{m,1} & \ldots & \alpha_{m,n} \end{pmatrix}}_{\alpha} \cdot \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$$

We now characterize the set $eq_{E_\oplus}(\varphi)$ of equations valid in φ. Let M, N be such that $\mathrm{var}(M) \cup \mathrm{var}(N) \subseteq \mathrm{dom}(\phi)$, $(\mathrm{names}(M) \cup \mathrm{names}(N)) \cap \mathrm{names}(\varphi) = \emptyset$. First notice that: $M\varphi =_{E_\oplus} N\varphi$ iff $(M \oplus N)\varphi =_{E_\oplus} 0$. Therefore we only study the case where $N = 0$.

Assume M in normal form. $M\varphi =_{E_\oplus} 0$ and $\mathrm{names}(M) \cap \mathrm{names}(\varphi) = \emptyset$ implies $\mathrm{names}(M) = \emptyset$. Let $M =_{E_\oplus} \beta_0 \oplus \bigoplus_{i=1}^{m} \beta_i\, x_i$. The condition $M\varphi =_{E_\oplus} 0$ is equivalent to the vectorial equation:

$$(\beta_0 \dots \beta_m) \cdot \alpha = 0$$

that is, $(\beta_0 \dots \beta_m)$ belongs to the co-kernel of α, noted $coker(\alpha)$.

Finally let φ and φ' be two closed frames with $\mathrm{names}(\varphi) \cup \mathrm{names}(\varphi') \subseteq \{a_1 \dots a_n\}$ and $dom(\varphi) = dom(\varphi') = \{x_1 \dots x_m\}$. Let α and α' be the two corresponding $(m+1) \times (n+1)$-matrices defined as above. From the previous discussion, we deduce that:

$$\varphi \approx_{E_\oplus} \varphi' \text{ iff } \mathrm{coker}(\alpha) = \mathrm{coker}(\alpha')$$

Theorem 3. *The usual implementation for the XOR theory is unconditionally $=_{E_\oplus}$-, $\approx_{E_\oplus}$- and $\not\models_{E_\oplus}$-sound. It is also $=_{E_\oplus}$-, $\approx_{E_\oplus}$- and $\not\models_{E_\oplus}$-faithful.*

Proof (of Theorem 3). The unconditional $=_{E_\oplus}$-soundness is clear, hence the $\not\models_{E_\oplus}$-faithfulness (Proposition 1).

Let $T \neq 0$ a closed term in normal form. The semantics of T is either the constant 1^η (if $T = 1$) or the uniform distribution (if $T \neq 1$) on $\{0,1\}^\eta$. Thus $\mathbb{P}\left[[\![T]\!]_{A_\eta} = 0\right]$ is negligible. Hence the $=_{E_\oplus}$-faithfulness holds and by proposition 1, so does the $\approx_{E_\oplus}$-faithfulness.

We use the ideal semantics to address the unconditional $\approx_{E_\oplus}$-soundness. Indeed we shall prove that: for any frame φ, $([\![\varphi]\!]_{A_\eta}) = ([\![\varphi]\!]^{ideal}_{A_\eta})$. The result will follow from the proof of Theorem 1.

Let φ be a frame, and $\alpha = (\alpha_{i,j})$ its $(m+1) \times (n+1)$-matrix associated as before. Let us see α as a $\mathbb{F}_2$-linear function from $(\mathbb{F}_2)^{n+1}$ to $(\mathbb{F}_2)^{m+1}$.

The usual concrete semantics of φ consists in drawing a random vector from $(\mathbb{F}_2)^{(n+1)\eta}$ for the value of names, and then applying a $\mathbb{F}_2$-linear function $\widehat{\alpha} : (\mathbb{F}_2)^{(n+1)\eta} \rightarrow (\mathbb{F}_2)^{(m+1)\eta}$. Specifically, if we see $(\mathbb{F}_2)^{(n+1)\eta}$ as $\underbrace{\mathbb{F}_2{}^\eta \times \ldots \times \mathbb{F}_2{}^\eta}_{n+1}$, the initial distribution over $(\mathbb{F}_2)^{(n+1)\eta}$ is the uniform distribution over the affine space $A = (1^\eta, 0^\eta \dots 0^\eta) + \{0^\eta\} \times (\mathbb{F}_2)^{n\eta}$ made of vectors of which the first block is 1^η. (This first constant block is only a convenience for dealing with the constant term.) The function $\widehat{\alpha}$ is defined by

$$\widehat{\alpha}\,(f_0 \dots f_n) = \left(\bigoplus_{j=0}^{n} \alpha_{0,j}\, f_j, \ldots, \bigoplus_{j=0}^{n} \alpha_{m,j}\, f_j \right)$$

On the other hand, the ideal semantics of φ consists in drawing a random vector from the subset F of $(\mathbb{F}_2)^{(m+1)\eta} \approx \underbrace{\mathbb{F}_2{}^\eta \times \ldots \times \mathbb{F}_2{}^\eta}_{m+1}$ whose elements satisfy all the equations in $\mathrm{eq}_{E_\oplus}(\varphi)$ and have 1^η as first block. (Again the first block is not present in the real bit-strings.) From the previous discussion, F is written

$$F = \{(e_0, e_1 \ldots e_m) \mid \forall(\beta_0 \ldots \beta_m) \in coker(\alpha),\ \bigoplus_{i=0}^{m} \beta_i\, e_i = 0\} \cap B$$

where $B = (1^\eta, 0^\eta \ldots 0^\eta) + \{0^\eta\} \times (\mathbb{F}_2)^{m\eta}$.

Now let us change the basis of $(\mathbb{F}_2)^{(n+1)\eta}$ and see it as $\underbrace{{\mathbb{F}_2}^{n+1} \times \ldots \times {\mathbb{F}_2}^{n+1}}_{\eta}$, that is: we regroup the i-th bit of each block. Then $\widehat{\alpha}$ is simply the product application: $\underbrace{\alpha \times \ldots \times \alpha}_{\eta}$. Similarly, if $(\mathbb{F}_2)^{(m+1)\eta} \approx \underbrace{{\mathbb{F}_2}^{m+1} \times \ldots \times {\mathbb{F}_2}^{m+1}}_{\eta}$ then F is simply the product of all values that are orthogonal to $\mathrm{coker}(\alpha)$, restricted to B:

$$\begin{aligned} F &= (\underbrace{\mathrm{coker}(\alpha)^\perp \times \ldots \times \mathrm{coker}(\alpha)^\perp}_{\eta}) \cap B \\ &= (\underbrace{\mathrm{im}(\alpha) \times \ldots \times \mathrm{im}(\alpha)}_{\eta}) \cap B \\ &= \mathrm{im}(\widehat{\alpha}) \cap B \\ &= \alpha(A) \end{aligned}$$

Thus F is in fact the image of A by $\widehat{\alpha}$.

As $\widehat{\alpha}$ is linear, it transforms uniform distributions over affine spaces into uniform distributions. Therefore the ideal and the concrete semantics yield both the uniform distribution over F. Hence the unconditional $\approx_{E_\oplus}$-soundness.

We now prove the unconditional $\not\vdash_{E_\oplus}$-soundness. Let φ be a frame and T a term, both in normal form, with $\mathrm{names}(\varphi) \cup \mathrm{names}(T) = \{a_1 \ldots a_n\}$. Let α associated to φ as before and $T =_{E_\oplus} \beta_0 \oplus \bigoplus_{j=1}^{n} \beta_j\, a_j$.

Suppose that $\varphi \vdash_{E_\oplus} T$, that is there exists ζ with $\mathrm{names}(\zeta) \cap \{a_1 \ldots a_n\} = \emptyset$, such that $\zeta\varphi =_{E_\oplus} T$ *i.e.* $\zeta\varphi \oplus T =_{E_\oplus} 0$. Assume ζ in AC-normal form. As previously, $\zeta\varphi \oplus T =_{E_\oplus} 0$ implies $\mathrm{names}(\zeta) \subseteq \{a_1 \ldots a_n\}$, hence $\mathrm{names}(\zeta) = \emptyset$. Therefore ζ computes nothing but a linear transformation on the rows of α. We deduce that $\varphi \vdash_{E_\oplus} T$ holds iff $(\beta_0 \ldots \beta_m)$ belongs to the co-image of α, written $\mathrm{coim}(\alpha)$.

Now assume that $\varphi \not\vdash_{E_\oplus} T$ *i.e.* $\beta = (\beta_0 \ldots \beta_n) \notin coim(\alpha)$. Let γ be the $(m+2) \times (n+1)$-matrix obtained by augmenting α with a $(m+1)$-th row equal to β:

$$\gamma = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ \alpha_{1,0} & \alpha_{1,1} & \ldots & \alpha_{1,n} \\ \vdots & & & \vdots \\ \alpha_{m,0} & \alpha_{m,1} & \ldots & \alpha_{m,n} \\ \beta_0 & \beta_1 & \ldots & \beta_n \end{pmatrix}$$

We know that the concrete semantics of φ and T is nothing but the uniform distribution over the image of $A = (1^\eta, 0 \ldots 0) + \{0\} \times (\mathbb{F}_2)^{n\eta}$ by $\widehat{\gamma}$ (defined as $\widehat{\alpha}$ above).

Let us see β a linear function from $(\mathbb{F}_2)^{n+1}$ to $\mathbb{F}_2$ and define $\widehat{\beta}$ as previously. Using the fact that β is independent from the other rows of γ, we prove that the image $\widehat{\gamma}(A)$ is the cartesian product of the two spaces $\widehat{\alpha}(A)$ and $\widehat{\beta}(A)$; the unconditional $\not\vdash_{E_\oplus}$-soundness will then follow.

The inclusion $\widehat{\gamma}(A) \subseteq \widehat{\alpha}(A) \times \widehat{\beta}(A)$ is trivial. As β is independant from the rows of α, there exists a vector $u \in (\mathbb{F}_2)^{n+1}$ such that $\beta(u) = 1$ and $\alpha(u) = 0$. Let $x, y \in A$. We prove that there exists $z \in A$ such that $\widehat{\gamma}(z) = (\widehat{\alpha}(x), \widehat{\beta}(y))$. Indeed, take $z = x + (\widehat{\beta}(y) - \widehat{\beta}(x)) \cdot \widehat{u}$. We have that $\widehat{\alpha}(z) = \widehat{\alpha}(x)$ and $\widehat{\beta}(z) = \widehat{\beta}(y)$. □

F $\approx_E$-soundness for the theory of ciphers and lists

Before detailing the proofs of Section 5.2, notice that each well-sorted term has a unique sort (recall that variables and names have a fixed sort). As the equations themselves are well-sorted (if $t : s$ and $t =_{E_{\mathsf{sym}}} t'$ then $t' : s$), and the indices n of function symbols are redundant with the sorts, we tend to omit the indices in terms. For instance, if $k, k' : Data$, we write: $\mathsf{enc}(\mathsf{cons}(k, \mathsf{nil}), k')$ instead of: $\mathsf{enc}_1(\mathsf{cons}_1(k, \mathsf{nil}), k')$.

F.1 Detailed proof of Proposition 5

Proposition 5. *Let φ be a closed frame. There exists a pattern $\overline{\varphi}$ such that $\varphi \approx_{E_{\mathsf{sym}}} \overline{\varphi}$.*

The proof of Proposition 5 relies on the following Lemma 1, that is used stepwise to rewrite a frame into a pattern. As for the termination of the procedure, assume that φ is not a pattern; define T as the father of the largest non-deducible subterm of φ; it is easy to see that T is necessarily of the form $T = \mathsf{enc}(U, V)$, $T = \mathsf{dec}(U, V)$, $T = \mathsf{head}(V)$ or $T = \mathsf{tail}(V)$ with $\varphi \not\vdash_{E_{\mathsf{sym}}} V$.

Lemma 1. *Let φ be a closed frame in $\mathcal{R}$-normal form. Let T be a subterm of φ of the form $T = \mathsf{enc}(U, V)$, $T = \mathsf{dec}(U, V)$, $T = \mathsf{head}(V)$ or, $T = \mathsf{tail}(V)$ and n a fresh name of the same sort than T. Assume that V is not deducible from φ,* i.e. $\varphi \not\vdash_{E_{\mathsf{sym}}} V$. *Then:*

$$\varphi \approx_{E_{\mathsf{sym}}} \varphi'$$

where $\varphi' = \varphi\{T \mapsto n\}$ is obtained by replacing every *occurrence of T in φ by n.*

We first introduce an handy lemma that gives a characterization of deducible terms.

Lemma 2. *Let φ be a closed frame in $\mathcal{R}$-normal form and T a term in $\mathcal{R}$-normal form. If $\varphi \vdash_{E_{\mathsf{sym}}} T$ then $T = C[T_1, \dots, T_k]$ where the T_i are deducible subterms of φ and C is a context that does not contain names.*

Proof. By definition, $\varphi \vdash_{E_{\mathsf{sym}}} T$ if and only if there exists a term M such that $\mathrm{names}(M) \cap \mathrm{names}(\varphi) = \emptyset$ and $M\varphi =_{E_{\mathsf{sym}}} T$, *i.e.* $M\varphi \rightarrow^*_{\mathcal{R}} T$. We prove Lemma 2 by induction on the size of M. The base case $M = x_i$ is trivial.

If $M = f(M_1, \dots, M_k)$. We only consider the case where $M = \mathsf{dec}(M_1, M_2)$ since the other cases are similar. We have $M_1 \rightarrow^*_{\mathcal{R}} T_1$ and $M_2 \rightarrow^*_{\mathcal{R}} T_2$ Applying the induction hypothesis to M_1 and M_2, we get that $T_1 = C_1[T'_1, \dots, T'_k]$ and

$T_2 = C_2[T'_1, \ldots, T'_k]$ where the T'_i are deducible subterms of φ and C_1, C_2 are contexts that do not contain names. We have $M\varphi \to^*_{\mathcal{R}} \mathsf{dec}(T_1, T_2)$. Either $\mathsf{dec}(T_1, T_2)$ is in $\mathcal{R}$-normal form. In that case and by convergence of $\mathcal{R}$, we have $T = \mathsf{dec}(T_1, T_2)$, hence the result. Or $\mathsf{dec}(T_1, T_2)$ is not in $\mathcal{R}$-normal form. By convergence, we have $\mathsf{dec}(T_1, T_2) \to_{\mathcal{R}} T$. Since T_1 and T_2 are already in normal form, we must have $T_1 = \mathsf{enc}(T'_1, T_2)$ and $T = T'_1$. Either $C_1 = \mathsf{enc}(C'_1, C''_1)$ and we have $T = C'_1[T'_1, \ldots, T'_k]$. Or $C_1 = \_$, which means that T_1 is a deducible subterm of φ. We deduce that T is a deducible subterm of φ, hence the result. □

We can now start the proof of Lemma 1.

Proof. Since $\varphi = \varphi'\{n \mapsto t\}$ and E_{sym} is stable by substitutions of names, we have $\mathrm{eq}_{E_{\mathsf{sym}}}(\varphi') \subseteq \mathrm{eq}_{E_{\mathsf{sym}}}(\varphi)$. To prove $\mathrm{eq}_{E_{\mathsf{sym}}}(\varphi) \subseteq \mathrm{eq}_{E_{\mathsf{sym}}}(\varphi')$, we introduce the following lemma. We set θ to be $\{n \mapsto t\}$. Let $n_1, \ldots, n_p$ be the names occurring in φ'.

Lemma 3. *Let C_1 be a context such that $\varphi' \vdash_{E_{\mathsf{sym}}} C_1[n_1, \ldots, n_p]$ and $C_1[n_1, \ldots, n_p]\theta \to_{\mathcal{R}} T$. Then there exists a public context C_2 such that $C_1 \to_{\mathcal{R}} C_2$ and $T = C_2[n_1, \ldots, n_p]\theta$.*

The lemma is proved by inspection of the rules of $\mathcal{R}$. The reduction occurs at some position p: the reduction $C_1[n_1, \ldots, n_p]|_p\theta \to_{\mathcal{R}} T$ occurs in head. Let $C'_1[n_1, \ldots, n_p] = C_1[n_1, \ldots, n_p]|_p$ If C'_1 is itself an instance of the left-hand-side of a rule of $\mathcal{R}$, than we clearly have that $C'_1 \to_{\mathcal{R}} C'_2$ such that $T = C_2[n_1, \ldots, n_p]\theta$, where C_2 is obtained from C_1 by replacing C'_1 by C'_2 et position p. If C'_1 is not an instance of the left-hand-side of a rule of $\mathcal{R}$ and since t is already in $\mathcal{R}$-normal form, there are only four possibilities for $C'_1[n_1, \ldots, n_p]$.

- $C'_1[n_1, \ldots, n_p] = \mathsf{enc}(n_i, C''_1[n_1, \ldots, n_p])$. It must be the case that $n_i = n$, t is of the form $\mathsf{dec}(u, v)$ and $v = C''_1[n_1, \ldots, n_p]$. ¿From Lemma 2 and since $\varphi' \vdash_{E_{\mathsf{sym}}} C_1[n_1, \ldots, n_p]$, either $C'_1[n_1, \ldots, n_p]$ is subterm of φ' or n_i and $C''_1[n_1, \ldots, n_p]$ are deducible. In both cases, we obtain a contradiction. Indeed, if $C'_1[n_1, \ldots, n_p]$ is subterm of φ' then $C'_1[n_1, \ldots, n_p]\theta = \mathsf{enc}(\mathsf{dec}(u, v), n_j)$ is a subterm of φ, which contradicts that φ is in normal form. If n_i and $C''_1[n_1, \ldots, n_p]$ are deducible then this contradicts $\varphi \not\vdash_{E_{\mathsf{sym}}} v$.
- $C'_1[n_1, \ldots, n_p] = \mathsf{dec}(n_i, n_j)$. This case is very similar to the previous one.
- $C'_1[n_1, \ldots, n_p] = \mathsf{cons}(n_i, C''_1[n_1, \ldots, n_p])$. It must be the case that $n_i = n$, t is of the form $\mathsf{head}(v)$ and $C''_1[n_1, \ldots, n_p] = \mathsf{tail}(v)$. ¿From Lemma 2 and since $\varphi' \vdash_{E_{\mathsf{sym}}} C_1[n_1, \ldots, n_p]$, either $C'_1[n_1, \ldots, n_p]$ is subterm of φ' or n_i and $C''_1[n_1, \ldots, n_p]$ are deducible. Like previously, in both cases, we obtain a contradiction. if $C'_1[n_1, \ldots, n_p]$ is subterm of φ' then $C'_1[n_1, \ldots, n_p]\theta = \mathsf{cons}(\mathsf{head}(v), \mathsf{tail}(v))$ is a subterm of φ, which contradicts that φ is in normal form. If n_i and $C''_1[n_1, \ldots, n_p]$ are deducible then both n and $\mathsf{tail}(v)$ are deducible in $\mathrm{var}\,\phi'$, which means that both $\mathsf{head}(v)$ and $\mathsf{tail}(v)$ are deducible in φ, thus v is deducible in φ, contradiction.
- $C'_1[n_1, \ldots, n_p] = \mathsf{cons}(C''_1[n_1, \ldots, n_p], n_i)$. This case is very similar to the previous one.

Now, let $(M = N) \in \mathrm{eq}_{E_{\mathsf{sym}}}(\varphi)$ and let us show that $(M = N) \in \mathrm{eq}_{E_{\mathsf{sym}}}(\varphi')$. We have $M\varphi =_{E_{\mathsf{sym}}} N\varphi$, *i.e.* $M\varphi'\theta =_{E_{\mathsf{sym}}} N\varphi'\theta$. By convergence of $\mathcal{R}$, we get that there exists a term T such that $M\varphi'\theta \to^*_{\mathcal{R}} T$ and $N\varphi'\theta \to^*_{\mathcal{R}} T$. Applying repeatedly

Lemma 3, we get that $M\varphi' \to_{\mathcal{R}}^* T_1$ such that $T = T_1\theta$ and $N\varphi' \to_{\mathcal{R}}^* T_2$ such that $T = T_2\theta$. Assume that we have proved that $T_1 = T_2$. Then we get $M\varphi' =_{E_{\mathsf{sym}}} N\varphi'$, *i.e.* $(M = N) \in \mathrm{eq}_{E_{\mathsf{sym}}}(\varphi')$, which concludes the proof. It remains us to prove the following lemma.

Lemma 4. *Let T_1 and T_2 two terms such that $\varphi' \vdash_{E_{\mathsf{sym}}} T_i$ ($i = 1, 2$). $T_1\theta = T_2\theta$ implies $T_1 = T_2$.*

The lemma is proved by induction on the sum of the size of T_1 and T_2.

- The base case is trivial.
- If none of T_1 or T_2 is n: $T_1 = f(T_1', \ldots, T_k')$ and $T_2 = f(T_1'', \ldots, T_k'')$. We must have $T_i'\theta T_i''\theta$ for every $1 \leq i \leq k$. Applying the induction hypothesis, we get $T_i' = T_i''$ thus $T_1 = T_2$.
- The most difficult case is when $T_1 = n$ and $T_2 = f(T_1', \ldots, T_k')$. We first notice that since $n\theta = f(T_1', \ldots, T_k')\theta$, n can not occur in T_2, thus $T_2 = T_2\theta = t$. Thanks to Lemma 2 and since $\varphi' \vdash_{E_{\mathsf{sym}}} T_2$, either T_2 is a subterm of φ', which is impossible by construction of φ' or the immediate subterms of T_2 are deducible in φ' (thus in φ), which contradicts the choice of t. □

F.2 A generalized SPRP assumption

In the same way as we generalize the IND-Pi-Cj assumption, we propose here a generalization of SPRP to the case of messages of heterogeneous sizes.

The ω-SPRP assumption consists of the following experience $\mathcal{G}_\eta$ involving an adversary $\mathcal{A}$:

- draw a random bit $b \in \{0, 1\}$;
- if $b = 0$, draw a key k; give $\mathcal{A}$ access to the encryption oracles $\mathcal{E}_{\eta,n}(\cdot, k)$ and decryption oracles $\mathcal{D}_{\eta,n}(\cdot, k)$;
- if $b = 1$, draw a random permutation π_n over $\{0, 1\}^{\eta n}$ for each n; give $\mathcal{A}$ access to the π_n and π_n^{-1} instead of the regular encryption and decryption oracles;
- after interacting with the oracles, $\mathcal{A}$ outputs a bit b'; *$\mathcal{A}$ is successful in $\mathcal{G}_\eta$* iff $b = b'$.

The *advantage* of $\mathcal{A}$ is defined as

$$Adv_{\mathcal{A}}^{\omega\text{-SPRP}}(\eta) = 2 \times \mathbb{P}\left[\mathcal{A} \text{ is successful in } \mathcal{G}_\eta\right] - 1 \tag{1}$$

The ω-SPRP assumption holds for $(\mathcal{E}_{\eta,n}, \mathcal{D}_{\eta,n})$ if the advantage (1) of any probabilistic polynomial-time adversary is a negligible function of η.

It has been shown [16] that the SPRP assumption (on one given cipher) is equivalent to asking that the encryption scheme and its inverse are both IND-P2-C2 or equivalently both IND-P1-C1. We conjecture that this is still the case for strengthened version of these notions—however this falls outside the scope of this paper. Within this (probably simple) conjecture, the $\approx_{E_{\mathsf{sym}}}$-soundness Theorem 4 holds for any ω-SPRP encryption scheme.

F.3 Proof of Theorem 4 ($\approx_{E_{\mathsf{sym}}}$-soundness)

Theorem 4 ($\approx_{E_{\mathsf{sym}}}$-soundness). *Let φ_1 and φ_2 be two well-formed frames of the same domain. Assume that the concrete implementations for the encryption and its inverse satisfy both the ω-IND-P1-C1 assumption. If $\varphi_1 \approx_{E_{\mathsf{sym}}} \varphi_2$ then $([\![\varphi_1]\!]_{A_\eta}) \approx ([\![\varphi_2]\!]_{A_\eta})$.*

We begin by stating a computational counterpart to Lemma 1.

Lemma 5. *Let φ be a closed frame in $\mathcal{R}$-normal form, with only atomic keys and no encryption cycles. Let T be a subterm of φ of the form $T = \mathsf{enc}(U, k)$ (*resp.* $T = \mathsf{dec}(U, k)$), with k name of sort $Data$, and n a fresh name of the same sort as T. Assume that:*

- *the only occurrences of k in φ are in the positions of an encryption or decryption key: $\mathsf{enc}(., k)$ or $\mathsf{dec}(., k)$;*
- *T itself does not appear under an encryption or a decryption with k;*
- *the concrete implementations for the encryption and its inverse satisfy both the ω-IND-P1-C1 assumption.*

Then:

$$([\![\varphi]\!]_{A_\eta}) \approx ([\![\varphi']\!]_{A_\eta})$$

where $\varphi' = \varphi\{T \mapsto n\}$ is obtained by replacing every *occurrence of T in φ by n.*

Notice that the hypothesis of Lemma 5 are stronger than its formal version, Lemma 1. For instance the encryption key k is required to be atomic; the first condition on k trivially implies that k is not deducible from φ. Also nothing is said about head and tail symbols.

Proof (of Lemma 5). Before proving the lemma, let us consider the example of a well-formed frame $\varphi_1 = \{x_1 = \mathsf{enc}(T_1, k),\ x_2 = \mathsf{enc}(T_2, k)\}$, where k does not appear in T_1, T_2, and $T_1 \neq_{E_{\mathsf{sym}}} T_2$. This frame is statically equivalent to $\varphi_2 = \{x_1 = n_1; x_2 = n_2\}$. Our problem here is to prove that $[\![\varphi_1]\!]$ and $[\![\varphi_2]\!]$ are actually indistinguishable. It is not hard to see that this will be the case if and only if the probability that T_1 and T_2 have the same concrete value is negligible. A consequence of this phenomenon is intuitively that we need to prove Lemma 5 and—at least—a limited form of $=_{E_{\mathsf{sym}}}$-faithfulness at the same time.

Formally, let us write $|\varphi|_e$ and $|T|_e$ for the number of distinct subterms with head symbols enc or dec, occurring *resp.* in a frame φ and a term T. Let P_n and Q_n be the two properties:

(P_n) Lemma 5 holds provided that $|\varphi|_e \leq n$.

(Q_n) For all $\mathcal{R}$-normal terms T_1, T_2 of the same sort such that: T_1, T_2 have only atomic keys, the frame $\varphi = \{x = T_1, y = T_2\}$ has no encryption cycles, $T_1 \neq T_2$ and $|\varphi|_e \leq n$, the probability $\mathbb{P}\left[e_1, e_2 \leftarrow [\![T_1, T_2]\!]_{A_\eta}; e_1 = e_2\right]$ is negligible.

We prove P_n and Q_n by mutual induction on n, that is, more precisely we prove the four statements: (1) P_0, (2) $P_{n+1} \Leftarrow Q_n$, (3) Q_0, (4) $Q_{n+1} \Leftarrow (P_{n+1}$ and $Q_n)$.

(1) P_0 is trivially true.

(2) $P_{n+1} \Leftarrow Q_n$. Let $T^0 = \mathsf{enc}_{n_0}(U,k)$ a subterm of φ, k and n two names all satisfying the conditions of Lemma 5. (Of course the case of $T^0 = \mathsf{dec}(U,k)$ is similar.) Let $\varphi = \{x_1 = T_1^0, \ldots, x_n = T_n^0\}$.

Provided an adversary $\mathcal{A}$ able to distinguish $(\llbracket\varphi\rrbracket_{A_\eta})$ and $(\llbracket\varphi'\rrbracket_{A_\eta})$, we build an adversary $\mathcal{B}$ against the ω-IND-P1-C1 assumption on encryption, described as follows:

1. for each name a of sort s appearing in φ, draw a value $\widehat{a} \overset{R}{\longleftarrow} \llbracket s\rrbracket_{A_\eta}$;
2. draw a value $\widehat{a_0} \overset{R}{\longleftarrow} \llbracket s\rrbracket_{A_\eta}$ for some fresh name a_0 of sort $List_{n_0}$;
3. for each x_i ($1 \leq i \leq n$) of sort s_i, compute $\widehat{T_i^0} \in \llbracket s_i\rrbracket_A$ recursively as follows:

   $$\begin{aligned} \widehat{\mathsf{enc}_n(T,k)} &= \mathcal{E}_n(\widehat{T}) \text{ if } T \neq U \\ \widehat{\mathsf{enc}_{n_0}(U,k)} &= \mathcal{E}^*(\widehat{U},\widehat{a_0}) \\ \widehat{\mathsf{dec}_n(T,k)} &= \mathcal{D}_n(\widehat{T}) \\ \widehat{f(T_1,\ldots,T_n)} &= f_{A_\eta}(\widehat{T_1},\ldots,\widehat{T_n}) \text{ if } f(T_1,\ldots,T_n) \notin \{\mathsf{enc}(T',k), \mathsf{dec}(T',k)\} \end{aligned}$$

   where we have written $\mathcal{E}_n(.)$ and $\mathcal{D}_n(.)$ for the encryption and decryption oracles of the ω-IND-P1-C1 game, and $\mathcal{E}^*(\widehat{U},\widehat{a_0})$ for the challenge ciphertext, obtained after submitting the two plaintexts $\widehat{U}$ and $\widehat{a_0}$ (this is done only once, just after $\widehat{U}$ has been computed);
4. submit the concrete frame $\{x_1 = \widehat{T_1}, \ldots, x_n = \widehat{T_n}\}$ to $\mathcal{A}$ and return the same answer.

Note that since T^0 is not a subterm of an encryption or decryption with k, $\mathcal{B}$ is indeed a 1-stage attacker. The distribution computed by $\mathcal{B}$ and submitted to $\mathcal{A}$ equals either $(\llbracket\varphi\rrbracket_{A_\eta})$ or $(\llbracket\varphi'\rrbracket_{A_\eta})$ depending on whichever $\mathcal{E}^*(\widehat{U},\widehat{a_0})$ is the encryption of $\widehat{U}$ or *resp.* that of $\widehat{a_0}$ (in the latter case $\mathcal{E}^*(\widehat{U},\widehat{a_0}) = \mathcal{E}_{n_0}(\widehat{a_0})$ is simply a random number). Thus the probability that $\mathcal{B}$ guesses the right answer is the same as $\mathcal{A}$. Now it may happen that $\mathcal{B}$ does not meet the second requirement for winning the ω-IND-P1-C1 game, that is: (i) there exists a subterm $\mathsf{enc}_{n_0}(T,k)$ such that $T \neq U$ and $\widehat{T} \in \{\widehat{U},\widehat{a_0}\}$ or (ii) there exists a subterm $\mathsf{dec}_{n_0}(T,k)$ such that $\widehat{T} = \mathcal{E}^*(\widehat{U},\widehat{a_0})$.

For (i), the probability that $\widehat{T} = \widehat{a_0}$ is negligible by construction. Moreover, as T and $T^0 = \mathsf{enc}_{n_0}(U,k)$ are two subterms of φ and T^0 is not a subterm of T, the frame $\varphi' = \{x = T, y = U\}$ has no encryption cycles and $|\varphi'|_e < |\varphi|_e = n+1$. The induction hypothesis Q_n implies that the probability for $\widehat{T} = \widehat{U}$ is negligible.

As for (ii), if the challenge ciphertext $\mathcal{E}^*(\widehat{U},\widehat{a_0})$ is the encryption of its second argument, that is $\mathcal{E}_{n_0}(\widehat{a_0})$, then the probability for $\widehat{T} = \mathcal{E}^*(\widehat{U},\widehat{a_0})$ is negligible; otherwise $\mathcal{E}^*(\widehat{U},\widehat{a_0}) = \mathcal{E}_{n_0}(\widehat{U})$. Recall that $T^0 = \mathsf{enc}_{n_0}(U,k)$ is in normal form, thus $U \neq \mathsf{dec}_{n_0}(T,k)$. As T^0 and $\mathsf{dec}_{n_0}(T,k)$ are two subterms of φ and T^0 is not a subterm of $\mathsf{dec}_{n_0}(T,k)$, the frame $\varphi' = \{x = U, y = \mathsf{dec}_{n_0}(T,k)\}$ has no encryption cycles and $|\varphi'|_e < |\varphi|_e = n+1$, hence the induction hypothesis Q_n implies that the probability for $\widehat{T} = \mathcal{E}_{n_0}(\widehat{U})$ is negligible.

To simplify the case analysis of (3) and (4), it is convenient to introduce to following lemma:

Lemma 6. *Let T_1, T_2 be two terms of sort $List_j$. Define for each $0 < i \leq j$, the i-th projection of a term T of sort $List_j$, by:*

$$\pi_i(T) = \mathsf{head}(\underbrace{\mathsf{tail}(\ldots tail}_{i-1 \text{ times}}(T)))$$

Then (i) $T_1 =_{E_{\mathsf{sym}}} T_2$ iff for all $1 \leq i \leq j$, $\pi_i(T_1) =_{E_{\mathsf{sym}}} \pi_i(T_2)$ and moreover (ii) $\mathbb{P}\left[e1, e2 \leftarrow [\![T_1, T_2]\!]_{A_\eta}; e_1 = e_2\right]$ is negligible iff for all $1 \leq i \leq j$,

$$\mathbb{P}\left[e_1^i, e_2^i \leftarrow [\![\pi_i(T_1) \downarrow_{\mathcal{R}}, \pi_i(T_2) \downarrow_{\mathcal{R}}]\!]_{A_\eta}; e_1^i = e_2^i\right]$$

is negligible.

(The notation $T \downarrow_{\mathcal{R}}$ stands for the $\mathcal{R}$-normal form of T.)

Thanks to this lemma, it is sufficient to prove (3) and (4) for T_1 and T_2 of sort $Data$ and in $\mathcal{R}$-normal form. (Indeed notice that if $\varphi = \{x = T_1, y = T_2\}$ has no encryption cycles, then $\varphi' = \{x' = \pi_i(T_1) \downarrow_{\mathcal{R}}, y' = \pi_i(T_2) \downarrow_{\mathcal{R}}\}$ has no encryption cycles and $|\varphi'|_e \leq |\varphi|_e$.)

Given the sorting system and the rewriting rules, a term T of sort $Data$ in $\mathcal{R}$-normal form can only be of the following forms:

1. a constant: 0 or 1,
2. a name of sort $Data$: $T = a$,
3. a projection of name of sort $List_j$: $T = \pi_i(a)$ ($1 \leq i \leq j$),
4. a projection of a encryption/decryption of sort $List_j$: $T = \pi_i(\mathsf{enc}(U, V))$ with $U \notin \{\mathsf{dec}(T', V)\}$ or $T = \pi_i(\mathsf{dec}(U, V))$ with $U \notin \{\mathsf{enc}(T', V)\}$.

(3) Q_0. As T_1 and T_2 contain no encryption/decryption symbol, only the cases $1 - -3$ of the case analysis above can occur; the property follows directly.

(4) $Q_{n+1} \Leftarrow (P_{n+1}$ and $Q_n)$. Let T_1 and T_2 be two distinct closed normal terms and $\varphi = \{x = T_1, y = T_2\}$. Assume that φ has no encryption cycles nor composed keys, and $|\varphi|_e = n + 1$.

1. If one of the two terms—say T_1— is of the form 1 (constant), 2 (name) or 3 (projection of a name). Then T_2 is of the form 4, for instance $T_2 = \pi_i(\mathsf{enc}(U, k))$ with $U \notin \{\mathsf{dec}(T', k)\}$.
   (a) If $T_1 \neq k$, by P_{n+1}, we have $([\![\varphi]\!]_{A_\eta}) \approx ([\![\{x = T_1, y = \pi_i(a)\}]\!]_{A_\eta})$ for some fresh name a. In particular, the probability for the two components x and y to be equal is negligible.
   (b) If $T_1 = k$, assume that T_1 and T_2 yields the same concrete value with significant probability. Let $List_{n_0}$ be the sort of U. We build an adversary $\mathcal{A}$ to the ω-IND-P1-C1 game as follows:
      i. for each name a of sort s appearing in T_2, draw a value $\widehat{a} \xleftarrow{R} [\![s]\!]_{A_\eta}$;
      ii. draw a value $\widehat{a_0} \xleftarrow{R} [\![s]\!]_{A_\eta}$ for some fresh name a_0 of sort $List_{n_0}$;

iii. compute $\widehat{T_2}$ recursively as follows:

$$\begin{aligned}
\widehat{\mathsf{enc}_n(T,k)} &= \mathcal{E}_n(\widehat{T}) \text{ if } T \neq U \\
\widehat{\mathsf{dec}_n(T,k)} &= \mathcal{D}_n(\widehat{T}) \\
\widehat{f(V_1,\dots,V_n)} &= f_{A_\eta}(\widehat{V_1},\dots,\widehat{V_n}) \text{ if } f(V_1,\dots,V_n) \notin \{\mathsf{enc}(T',k),\mathsf{dec}(T',k)\} \\
\widehat{\mathsf{enc}_{n_0}(U,k)} &= \mathcal{E}^*(\widehat{U},\widehat{a_0})
\end{aligned}$$

(same notations as before)

iv. if $\mathcal{E}_{n_0}(\widehat{U},\widehat{T_2}) = \widehat{T_2}$, returns 0, otherwise return 1.

Clearly $\mathcal{A}$ guesses the correct answer with non-negligible probability. As before, we use the property Q_n to conclude that its advantage is non-negligible.

2. Suppose $T_1 = \pi_{i_1}(\mathsf{enc}(u_1,k_1))$ and $T_2 = \pi_{i_2}(\mathsf{enc}(u_2,k_2))$ (the 3 other cases with decryption symbols are similar). As φ has no encryption cycle, we can assume for instance that $k_1 \not<_\varphi k_2$. By P_{n+1}, we have $(\llbracket\varphi\rrbracket_{A_\eta}) \approx (\llbracket\varphi'\rrbracket_{A_\eta})$ where $\varphi' = \varphi\{\mathsf{enc}(u_1,k_1) \mapsto a\} = \{x = T_1', y = T_2'\}$ for some fresh name a. We then apply Q_n to T_1' and T_2'. □

Proof (of Lemma 6). Point (i) is easily shown by induction on i, using the five last equations of E_{sym}. For (ii), notice that:

$$\mathbb{P}\left[e1, e2 \leftarrow \llbracket T_1, T_2\rrbracket_{A_\eta}; e_1 = e_2\right] \leq \sum_{i=1}^{j} \mathbb{P}\left[e_1^i, e_2^i \leftarrow \llbracket \pi_i(T_1), \pi_i(T_2)\rrbracket_{A_\eta}; e_1^i = e_2^i\right]$$

and

$$\forall i, \quad \mathbb{P}\left[e1, e2 \leftarrow \llbracket T_1, T_2\rrbracket_{A_\eta}; e_1 = e_2\right] \geq \mathbb{P}\left[e_1^i, e_2^i \leftarrow \llbracket \pi_i(T_1), \pi_i(T_2)\rrbracket_{A_\eta}; e_1^i = e_2^i\right]$$

Besides it is clear from the unconditional $=_{E_{\mathsf{sym}}}$-soundness, that for any T_1, T_2:

$$\mathbb{P}\left[e1, e2 \leftarrow \llbracket T_1, T_2\rrbracket_{A_\eta}; e_1 = e_2\right] = \mathbb{P}\left[e1, e2 \leftarrow \llbracket T_1 \downarrow_{\mathcal{R}}, T_2 \downarrow_{\mathcal{R}}\rrbracket_{A_\eta}; e_1 = e_2\right] \quad \square$$

Proof (of Theorem 4). Thanks to the (unconditional) $=_{E_{\mathsf{sym}}}$-soundness, it is enough to prove the property on frames in $\mathcal{R}$-normal form.

We begin by proving the following lemma:

Lemma 7. *Assume that the concrete implementations for the encryption and its inverse satisfy both the ω-IND-P1-C1 assumption. For every well-formed $\mathcal{R}$-normalized frame φ, $(\llbracket\varphi\rrbracket_{A_\eta}) \approx (\llbracket\overline{\varphi}\rrbracket_{A_\eta})$ where $\overline{\varphi}$ is the pattern associated to φ following the algorithmic proof of Proposition 5 (this pattern is uniquely defined modulo renaming of names.).*

Now recall that by Proposition 4 and since $\varphi \approx \overline{\varphi}$, we have:

$$\llbracket\overline{\varphi}\rrbracket_{A_\eta} = \llbracket\overline{\varphi}\rrbracket_{A_\eta}^{ideal} = \llbracket\varphi\rrbracket_{A_\eta}^{ideal}$$

Therefore the soundness criterion holds for well-formed $\mathcal{R}$-normalized frames and we conclude by Theorem 1. □

Notice that the use of the ideal semantics could not be (easily) avoided as two statically equivalent patterns may not be equal modulo renaming of bound names.

Proof (of Lemma 7). We prove the property by induction on the number m of encryption and decryption by non-deducible keys in φ.

If $m = 0$, by the well-formedness condition, φ is already a pattern.

Suppose that $m > 0$. As φ has no encryption cycle, we choose a non-deducible (atomic) key k appearing in φ, such that k is maximal for the encryption relation $>_\varphi$.

As k is not deducible, is maximal for $>_\varphi$ and φ contains no head and tail symbols, the only occurrences of k in φ are as encryption or decryption keys. Let T be the largest subterm of φ of the form $T = \mathsf{enc}(U, k)$ or $T = \mathsf{dec}(U, k)$. We apply Lemma 5 on φ and T and conclude by induction hypothesis on the obtained frame φ'. □