Formal Indistinguishability extended to the Random Oracle Model

Cristian Ene, Yassine Lakhnech and Van Chan Ngo *

Université Grenoble 1, CNRS,Verimag

Abstract. Several generic constructions for transforming one-way functions to asymmetric encryption schemes have been proposed. One-way functions only guarantee the weak secrecy of their arguments. That is, given the image by a one-way function of a random value, an adversary has only negligible probability to compute this random value. Encryption schemes must guarantee a stronger secrecy notion. They must be at least resistant against indistinguishability-attacks under chosen plaintext text (IND-CPA). Most practical constructions have been proved in the random oracle model (ROM for short). Such computational proofs turn out to be complex and error prone. Bana et al. have introduced *Formal Indistinguishability Relations (FIR)*, as an abstraction of computational indistinguishability. In this paper, we extend the notion of FIR to cope with the ROM on one hand and adaptive adversaries on the other hand. Indeed, when dealing with hash functions in the ROM and one-way functions, it is important to correctly abstract the notion of weak secrecy. Moreover, one needs to extend frames to include adversaries in order to capture security notions as IND-CPA. To fix these problems, we consider pairs of formal indistinguishability relations and *formal non-derivability relations*. We provide a general framework along with general theorems, that ensure soundness of our approach and then we use our new framework to verify several examples of encryption schemes among which the construction of Bellare Rogaway and Hashed ElGamal.

1 Introduction

Our day-to-day lives increasingly depend upon information and our ability to manipulate it securely. That is, in a way that prevents malicious elements to subvert the available information for their own benefits. This requires solutions based on *provably correct* cryptographic systems (e.g., primitives and protocols). There are two main frameworks for analyzing cryptographic systems; the *symbolic framework*, originating from the work of Dolev and Yao [16], and the *computational approach*, growing out of the work of [18]. A significant amount of effort has been made in order to link both approaches and profit from the advantages of each of them. Indeed, while the symbolic approach is more amenable to automated proof methods, the computation approach can be more realistic.

* Grenoble, email:name@imag.fr This work has been partially supported by the ANR projects SCALP, AVOTE and SFINCS

In their seminal paper [1] Abadi and Rogaway investigate the link between the symbolic model on one hand and the computational model on the other hand. More precisely, they introduce an equivalence relation on terms and prove that equivalent terms correspond to indistinguishable distributions ensembles, when interpreted in the computational model. The work of Abadi and Rogaway has been extended to active adversaries and various cryptographic primitives in e.g. [21, 20, 14, 19]. An other line of work, also considering active adversaries is followed by Backes, Pfitzmann and Waidner using *reactive simulatability* [5, 4] and Canetti [12, 13] using *universal composability*.

Related works. A recently emerging branch of relating symbolic and computational models for passive adversaries is based on *static equivalence* from π-calculus [3], induced by an *equational theory*. Equational theories provide a framework to specify algebraic properties of the underlying signature, and hence, symbolic computations in a similar way as for abstract data types. That is, for a fixed equational theory, a term describes a computation in the symbolic model. Thus, an adversary can distinguish two terms, if he is able to come up with two computations that yield the same result when applied to one term but different results when applied to the other term. Such a pair of terms is called a *test*. This idea can be extended to *frames*, which roughly speaking are tuples of terms. Thus, a *static equivalence* relation is fully determined by the underlying equational theory, as two frames are *statically equivalent*, if there is no test that separates them. In [8] Baudet, Cortier and Kremer study soundness and faithfulness of static equivalence for general equational theories and use their framework to prove soundness of exclusive or as well as certain symmetric encryptions. Abadi et al. [2] use static equivalence to analyze guessing attacks.

Bana, Mohassel and Stegers [7] argue that even though static equivalence works well to obtain soundness results for the equational theories mentioned above, it does not work well in other important cases. Consider for instance the Decisional Diffie Hellman assumption (DDH for short) that states that the tuples (g, g^a, g^b, g^{ab}) and (g, g^a, g^b, g^c), are indistinguishable for randomly sampled a, b, c. It does not seem to be obvious to come up with an equational theory for group exponentiation such that the induced static equivalence includes this pair of tuples without including others whose computational indistinguishability is not proved to be a consequence of the DDH assumption. The static equivalence induced by the equational theory for group exponentiation proposed in [8] includes the pair (g, g^a, g^b, g^{a^2b}) and (g, g^a, g^b, g^c). It is unknown whether the computational indistinguishability of these two distributions can be proved under the DDH assumption. Therefore, Bana et al. propose an alternative approach to build symbolic indistinguishability relations and introduce *formal indistinguishability relations (FIR)*. A FIR is defined as a closure of an initial set of equivalent frames with respect to simple operations which correspond to steps in proofs by reduction. This leads to a flexible symbolic equivalence relation. FIR has nice properties. In order to prove soundness of a FIR it is enough to prove soundness of the initial set of equivalences. Moreover, static equivalence

is one instance of a FIR. Bana et al. show that it is possible to come up with a FIR whose soundness is equivalent to the DDH assumption.

The techniques introduced in this paper, borrow and generalize to arbitrary equational theories some ideas from [15]. In [15] the authors provide a specialized Hoare-like logic to reason about encryption schemes in the random oracle model, and apply their logic to prove IND-CPA of several schemes, including the generic encryption scheme of Bellare and Rogaway [10].

Contributions. In this paper, we extend Bana et al.'s approach by introducing a notion of symbolic equivalence that allows us to prove security of encryption schemes symbolically. More specifically, we would like to be able to treat generic encryption schemes that transform one-way functions to IND-CPA secure encryption schemes. Therefore, three problems need to be solved. First, we need to cope with one-way functions. This is a case where the static equivalence does not seem to be appropriate. Indeed, let f be a one-way function, that is, a function that is easy to compute but difficult to invert. It does not seem easy to come with a set of equations that capture the one-wayness of such a function. Consider the term $f(a|b)$, where $|$ is bit-string concatenation. We know that we cannot easily compute $a|b$ given $f(a|b)$ for uniformly sampled a and b. However, nothing prevents us from being able to compute a for instance. Introducing equations that allow us to compute a from $f(a|b)$, e.g., $g(f(a|b)) = a$, may exclude some one-way functions and does not solve the problem. For instance, nothing prevents us from computing a prefix of b, a prefix of the prefix, etc ... The second problem that needs to be solved is related to the fact that almost all practical provably secure encryption schemes are analyzed in the random oracle model (ROM for short). ROM is an idealized model in which hash functions are randomly sampled functions. In this model, adversaries have oracle access to these functions. An important property is that if an adversary is unable to compute the value of an expression a and if $H(a)$ has not been leaked then $H(a)$ looks like a uniformly sampled value. Thus, we need to be able to symbolically prove that a value of a given expression a cannot be computed by any adversary. This is sometimes called *weak secrecy* in contrast to indistinguishability based secrecy. To cope with this problem, our notion of symbolic indistinguishability comes along with a *non-derivability* symbolic relation. Thus in our approach, we start from an initial pair of a non-derivability relation and a frame equivalence relation. Then, we provide rules that define a closure of this pair of relations in the spirit of Bana et al.'s work. Also in our case, soundness of the obtained relations can be checked by checking soundness of the initial relations. The third problem is related to the fact that security notions for encryption schemes such IND-CPA and real-or-random indistinguishability of cipher-text under chosen plaintext involve *active* adversaries. Indeed, these security definitions correspond to two-phase games, where the adversary first computes a value, then a challenge is produced, then the adversary tries to solve the challenge. Static equivalence and FIR (as defined in [7]) consider only passive adversaries. To solve this problem we consider frames that include variables that correspond to adversaries. As frames are finite terms, we only have finitely many such variables. This is the reason why we only

have a degenerate form of active adversaries which is enough to treat security of encryption schemes and digital signature, for instance. The closure rules we propose in our framework are designed with the objective of minimizing the initial relations which depend on the underlying cryptographic primitives and assumptions. We illustrate the framework by considering security proofs of the construction of Bellare and Rogaway [10] and Hash El Gamal [6].

Outline of the paper. In Section 2, we introduce the symbolic model used for describing generic asymmetric encryption schemes. In Section 3, we describe the computational framework and give definitions that relate the two models. In Section 4, we introduce our definition of formal indistinguishability relation and formal non-derivability relation. We also present our method for proving IND-CPA security. In Section 5, we illustrate our framework: we prove the constructions of Bellare and Rogaway [10], Hash El Gamal [6], and the encryption scheme proposed by Pointcheval in [24]. Finally, in Section 7 we conclude.

2 Symbolic semantics

A *signature* $\Sigma = (\mathcal{S}, \mathcal{F}, \mathcal{H})$ consists of a countable infinite set of *sorts* $\mathcal{S} = \{s, s_1, ...\}$, a finite set of *function symbols*, $\mathcal{F} = \{f, f_1, ...\}$, and a finite set of *oracle symbols*, $\mathcal{H} = \{g, h, h_1, ...\}$ together with arities of the form $ar(f)$ or $ar(h) = s_1 \times ... \times s_k \rightarrow s, k \geq 0$. Symbols in $\mathcal{F}$ that take $k = 0$ as arguments are called *constants*. We suppose that there are three pairwise disjoint countable sets $\mathcal{N}$, $\mathcal{X}$ and $\mathcal{P}$. $\mathcal{N}$ is the set of names, $\mathcal{X}$ is the set of first-order variables, and $\mathcal{P}$ is the set of second order variables. We assume that both names and variables are sorted, that is, to each name or variable u, a sort $\mathbf{s}$ is assigned; we use $\mathbf{s}(u)$ for the sot of u. Variables $p \in \mathcal{P}$ have arities $ar(p) = \mathbf{s}_1 \times ... \times \mathbf{s}_k \rightarrow \mathbf{s}$.

A renaming is a bijection $\tau : \mathcal{N} \rightarrow \mathcal{N}$ such that $\mathbf{s}(a) = \mathbf{s}(\tau(a))$. As usual, we extend the notation $\mathbf{s}(T)$ to denote the sort of a term T. Terms of sort $\mathbf{s}$ are defined by the grammar:

$T ::=$	x	*variable* x *of sort* $\mathbf{s}$
	$\|n$	*name* n *of sort* $\mathbf{s}$
	$\|p(T_1, \ldots, T_k)$	*variable* p *of arity* $\mathbf{s}(T_1) \times ... \times \mathbf{s}(T_k) \rightarrow \mathbf{s}$
	$\|f(T_1, \ldots, T_k)$	*application of* $f \in \mathcal{F}$ *with arity* $\mathbf{s}(T_1) \times ... \times \mathbf{s}(T_k) \rightarrow \mathbf{s}$
	$\|h(T_1, \ldots, T_k)$	*call of* $h \in \mathcal{H}$ *with arity* $\mathbf{s}(T_1) \times ... \times \mathbf{s}(T_k) \rightarrow \mathbf{s}$

We use $fn(T)$, $pvar(T)$ and $var(T)$ for the set of free names, the set of p-variables and the set of variables that occur in the term T, respectively. Meta-variables u, v, w range over names and variables. We use $st(T)$ for the set of subterms of T, defined in the usual way: $st(u) \stackrel{def}{=} \{u\}$ if u is a name or a variable, and $st(l(T_1, \ldots, T_k)) \stackrel{def}{=} \{l(T_1, \ldots, T_k)\} \bigcup_{i \in \{1,...k\}} st(T_i)$, if $l \in \mathcal{F} \cup \mathcal{H} \cup \mathcal{P}$. A term T is closed if it does not have any free variables (but it may contain p-variables), that means $var(T) = \emptyset$. The set of terms is denoted by $\mathbf{T}$.

Symbols in $\mathcal{F}$ are intended to model cryptographic primitives, symbols in $\mathcal{H}$ are intended to model cryptographic oracles (in particular, hash functions in the ROM model), and names in $\mathcal{N}$ are used to model secrets, i.e. concretely random

numbers. Variables $p \in \mathcal{P}$ are intended to model queries and challenges made by adversaries (and can depend on previous queries).

Definition 1 (Substitution). *A substitution $\sigma = \{x_1 = T_1, ..., x_n = T_n\}$ is a mapping from variables to terms whose domain $dom(\sigma) = \{x_1, ..., x_n\}$ is finite and such that $\sigma(x) \neq x$, for each x in the domain.*

A substitution as above is *well-sorted* if x_i and T_i have the same sort for each i, and there is no circular dependence $x_{i_2} \in var(T_{i_1})$, $x_{i_3} \in var(T_{i_2})$, ..., $x_{i_1} \in var(T_{i_k})$. The application of a substitution σ to a term T is written as $\sigma(T) = T\sigma$. This definition is lifted in a standard way to the application of a substitution to set of terms or substitutions. The *normal form* σ^* of a well-sorted substitution σ is the iterative composition of σ with itself until it remains unchanged : $\sigma^* = (\dots((\sigma)\sigma)\dots)\sigma$. For example, if $\sigma = \{x_1 = a, x_2 = f(b, x_1), x_3 = g(x_1, x_2)\}$, then $\sigma^* = \{x_1 = a, x_2 = f(b, a), x_3 = g(a, f(b, a)\}$. A substitution is *closed* if all terms (of its normal form) T_i are closed. We let $var(\sigma) = \cup_i var(T_i)$, $pvar(\sigma) = \cup_i pvar(T_i)$, $n(\sigma) = \cup_i fn(T_i)$, and extend the notations $pvar(.)$, $var(.)$, $n(.)$ and $st(.)$ to tuples and set of terms in the obvious way.

The abstract semantics of symbols is described by an equational theory E, that is an equivalence (denoted as $=_E$) which is stable with respect to application of contexts and well-sorted substitutions of variables.

Definition 2 (Equational Theory.). *An equational theory for a given signature is an equivalence relation $E \subseteq \mathcal{T} \times \mathcal{T}$ (written as $=_E$ in infix notation) on the set of terms such that*
1) $T_1 =_E T_2$ implies $T_1\sigma =_E T_2\sigma$ for every substitution σ;
2) $T_1 =_E T_2$ implies $T\{x = T_1\} =_E T\{x = T_2\}$ for every term T and every variable x;
3) $T_1 =_E T_2$ implies $\tau(T_1) =_E \tau(T_2)$ for every renaming τ.

Frames ([3]) represent sequences of messages observed by an adversary. Formally:

Definition 3 (Frame). *A frame is an expression of the form $\phi = \nu\widetilde{n}.\sigma$ where σ is a well-sorted substitution, and $\widetilde{n}$ is $n(\sigma)$, the set of all names occurring in σ. By abuse of notation we also use $n(\phi)$ for $\widetilde{n}$, the set of names bounded in the frame ϕ. We note $fv(\phi) \stackrel{def}{=} var(\sigma) \setminus dom(\sigma)$ the set of free variables of ϕ.*

The novelty of our definition of frames consists in permitting adversaries to interact with frames using p-variables. This is necessary to be able to cope with adaptive adversaries. We note the set of frames by $\mathbf{F}$.

The normal form ϕ^* of a frame $\phi = \nu\widetilde{n}.\sigma$ is the frame $\phi^* = \nu\widetilde{n}.\sigma^*$. From now on, we tacitly identify substitutions and frames with their normal form. Next, we define composition of frames. Let $\phi = \nu\widetilde{n}.\{x_1 = T_1, ..., x_n = T_n\}$ and $\phi' = \nu\widetilde{n'}.\sigma$ be frames with $\widetilde{n} \cap \widetilde{n'} = \emptyset$. Then, $\phi\phi'$ denotes the frame $\nu(\widetilde{n} \cup \widetilde{n'}).\{x_1 = T_1\sigma, ..., x_n = T_n\sigma\}$.

Definition 4 (Equational equivalence). *Let ϕ and ϕ' be two frames such that $\phi^* = \nu\widetilde{n}.\sigma$ and $\phi'^* = \nu\widetilde{n}.\sigma'$ with $\sigma = \{x_1 = T_1, ..., x_n = T_n\}$ and $\sigma' = \{x_1 = T'_1, ..., x_n = T'_n\}$. Given the equational theory E, we say that ϕ and ϕ' are equationally equivalent written $\phi =_E \phi'$, if and only if $T_i\sigma =_E T'_i\sigma'$ for all i.*

3 Computational Semantics

3.1 Distributions and indistinguishability

Let us note $\eta \in \mathrm{N}$ the security parameter. We are interested in analyzing generic schemes for asymmetric encryption in the *random oracle model* [17, 10]. We write $h \stackrel{r}{\leftarrow} \Omega$ to denote that h is randomly chosen from the set of functions with appropriate domain (depending on η). By abuse of notation, for a list $\boldsymbol{H} = h_1, \cdots, h_m$ of hash functions, we write $\boldsymbol{H} \stackrel{r}{\leftarrow} \Omega$ instead of the sequence $h_1 \stackrel{r}{\leftarrow} \Omega, \ldots, h_m \stackrel{r}{\leftarrow} \Omega$. We fix a finite set $\mathcal{H} = \{h_1, \ldots, h_n\}$ of hash functions. A *distribution ensemble* is a countable sequence of distributions $\{X_\eta\}_{\eta \in \mathrm{N}}$. We only consider distribution ensembles that can be constructed in polynomial time by probabilistic algorithms that have oracle access to $\mathcal{O} = \mathcal{H}$. Given two distribution ensembles $X = \{X_\eta\}_{\eta \in \mathrm{N}}$ and $X' = \{X'_\eta\}_{\eta \in \mathrm{N}}$, an algorithm $\mathcal{A}$ and $\eta \in \mathrm{N}$, the *advantage* of $\mathcal{A}$ in distinguishing X_η and X'_η is defined by:

$$\mathsf{Adv}(\mathcal{A}, \eta, X, X') = \mathsf{Pr}[x \stackrel{r}{\leftarrow} X_\eta : \mathcal{A}^{\mathcal{O}}(\eta, x) = 1] - \mathsf{Pr}[x \stackrel{r}{\leftarrow} X'_\eta : \mathcal{A}^{\mathcal{O}}(\eta, x) = 1].$$

Then, two distribution ensembles X and X' are called *indistinguishable* (denoted by $X \sim X'$) if for any probabilistic polynomial-time algorithm $\mathcal{A}$, the advantage $\mathsf{Adv}(\mathcal{A}, \eta, X, X')$ is negligible as a function of η, that is, for any $n > 0$, it become eventually smaller than η^{-n} as η tends to infinity.

3.2 Frames as distributions

We now give terms and frames a computational semantics parameterized by a computable implementation of the primitives in ROM. Provided a set of sorts $\mathcal{S}$ and a set of symbols $\mathcal{F}$, a *computational algebra* $A = (\mathcal{S}, \mathcal{F})$ consists of

- a sequence of non-empty finite set of bit strings $[\![s]\!]_A = \{[\![s]\!]_{A,\eta}\}_{\eta \in \mathrm{N}}$ with $[\![s]\!]_{A,\eta} \subseteq \{0,1\}^*$ for each sort $s \in S$. For simplicity of the presentation, we assume that all sorts are large domains, whose cardinalities are exponential in the security parameter η;

- a sequence of polynomial time computable functions $[\![f]\!]_A = \{[\![f]\!]_{A,\eta}\}_{\eta \in \mathrm{N}}$ with $[\![f]\!]_{A,\eta} : [\![s_1]\!]_{A,\eta} \times ... \times [\![s_k]\!]_{A,\eta} \rightarrow [\![s]\!]_{A,\eta}$ for each $f \in \mathcal{F}$ with $ar(f) = s_1 \times ... \times s_k \rightarrow s$;

- a polynomial time computable congruence $=_{A,\eta,s}$ for each sort s, in order to check the equality of elements in $[\![s]\!]_{A,\eta}$ (the same element may be represented by different bit strings). By congruence, we mean a reflexive, symmetric, and transitive relation such that $e_1 =_{A,s_1,\eta} e'_1, ..., e_k =_{A,s_k,\eta} e'_k \Rightarrow [\![f]\!]_{A,\eta}(e_1, ..., e_k) =_{A,s,\eta} [\![f]\!]_{A,\eta}(e'_1, ..., e'_k)$ (we usually omit s,η and A and write $=$ for $=_{A,s,\eta}$);

- a polynomial time procedure to draw random elements from $[\![s]\!]_{A,\eta}$; we denote such a drawing by $x \leftarrow^R [\![s]\!]_{A,\eta}$; for simplicity, in this paper we suppose that all these drawing follow a uniform distribution.

From now on we assume a fixed computational algebra $(\mathcal{S}, \mathcal{F})$, and a fixed η, and for simplicity we omit the indices A,s and η. For lack of space, we use *ppt* to stand for probabilistic polynomial-time. Given $\mathcal{H}$ a fixed set of hash functions, and $(\mathcal{A}_i)_{i \in I}$ a fixed set of ppt functions (can be seen as a ppt adversary $\mathcal{A}^{\mathcal{O}}$ taking

an additional input i), we associate to each frame $\phi = \nu\widetilde{n}.\{x_1 = T_1, \ldots, x_k = T_k\}$ a sequence of distributions $[\![\phi]\!]_{\mathcal{H},\mathcal{A}}$ computed as follows:

- for each name n of sort s appearing in $\widetilde{n}$, draw a value $\hat{n} \xleftarrow{r} [\![s]\!]$;
- for each variable $x_i (1 \leq i \leq k)$ of sort s_i, compute $\hat{T}_i \in [\![s_i]\!]$ recursively on the structure of terms: $\hat{x_i} = \hat{T}_i$;
- for each call $h_i(T'_1, \ldots, T'_m)$ compute recursively on the structure of terms: $\widehat{h_i(T'_1, \ldots, T'_m)} = h_i(\hat{T}'_1, \ldots, \hat{T}'_m)$;
- for each call $f(T'_1, \ldots, T'_m)$ compute recursively on the structure of terms: $\widehat{f(T'_1, \ldots, T'_m)} = [\![f]\!](\hat{T}'_1, \ldots, \hat{T}'_m)$;
- for each call $p_i(T'_1, \ldots, T'_m)$ compute recursively on the structure of terms and draw a value $\widehat{p_i(T'_1, \ldots, T'_m)} \xleftarrow{r} \mathcal{A}^{\mathcal{O}}(i, \hat{T}'_1, \ldots, \hat{T}'_m)$;
- return the value $\hat{\phi} = \{x_1 = \hat{T}_1, \ldots, x_k = \hat{T}_k\}$.

Such $\phi = \{x_1 = bse_1, \ldots, x_n = bse_n\}$ with $bse_i \in [\![s_i]\!]$ are called *concrete frames*. We extend the notation $[\![.]\!]$ to (sets of) closed terms in the obvious way.

Now the concrete semantics of a frame ϕ with respect to an adversary $\mathcal{A}$, is given by the following sequence of distributions (one for each implicit η):

$[\![\phi]\!]_{\mathcal{A}} = \left[\mathcal{H} \xleftarrow{r} \Omega; \mathcal{O} = \mathcal{H}; \hat{\phi} \xleftarrow{r} [\![\phi]\!]_{\mathcal{H},\mathcal{A}} : \hat{\phi}\right]$

When $pvar(\phi) = \emptyset$, semantics of ϕ does not depend on the adversary $\mathcal{A}$ and we will use the notation $[\![\phi]\!]$ (or $[\![\phi]\!]_{\mathcal{H}}$) instead of $[\![\phi]\!]_{\mathcal{A}}$ (respectively $[\![\phi]\!]_{\mathcal{H},\mathcal{A}}$).

3.3 Soundness and Completeness

The computational model of a cryptographic scheme is closer to reality than its formal representation by being a more detailed description. Therefore, the accuracy of a formal model can be characterized based on how close it is to the computational model. For this reason, we introduce the notions of soundness and completeness (inspired from [8]) that relate relations in the symbolic model with respect to similar relations in the computational model. Let E be an equivalence theory and let $R_1 \subseteq \mathbf{T} \times \mathbf{T}$, $R_2 \subseteq \mathbf{F} \times \mathbf{T}$, and $R_3 \subseteq \mathbf{F} \times \mathbf{F}$ be relations on closed frames, on closed terms, and relations on closed frames and terms, respectively.

- R_1 is $=$-sound iff for all terms T_1, T_2 of the same sort, $(T_1, T_2) \in R_1$ implies that $\mathsf{Pr}[\hat{e_1}, \hat{e_2} \xleftarrow{r} [\![T_1, T_2]\!]_{\mathcal{A}} : \hat{e}_1 \neq \hat{e}_2))]$ is negligible for any ppt adversary $\mathcal{A}$.
- R_1 is $=$-complete iff for all terms T_1, T_2 of the same sort, $(T_1, T_2) \notin R_1$ implies that $\mathsf{Pr}[\hat{e_1}, \hat{e_2} \xleftarrow{r} [\![T_1, T_2]\!]_{\mathcal{A}} : \hat{e}_1 \neq \hat{e}_2))]$ is non-negligible for some ppt adversary $\mathcal{A}$.
- R_1 is $=$-faithful iff for all terms T_1, T_2 of the same sort, $(T_1, T_2) \notin R_1$ implies that $\mathsf{Pr}[\hat{e_1}, \hat{e_2} \xleftarrow{r} [\![T_1, T_2]\!]_{\mathcal{A}} : \hat{e}_1 = \hat{e}_2))]$ is negligible for any ppt adversary $\mathcal{A}$.
- R_2 is $\not\vdash$-sound iff all frame ϕ and term T, $(\phi, T) \in R_2$ implies that $\mathsf{Pr}[\hat{\phi}, \hat{e} \xleftarrow{r} [\![\phi, T]\!]_{\mathcal{A}} : \mathcal{A}^{\mathcal{O}}(\hat{\phi}) = \hat{e}]$ is negligible for any ppt adversary $\mathcal{A}$.
- R_2 is $\not\vdash$-complete iff for all frame ϕ and term T, $(\phi, T) \notin R_2$ implies that $\mathsf{Pr}[\hat{\phi}, \hat{e} \xleftarrow{r} [\![\phi, T]\!]_{\mathcal{A}} : \mathcal{A}^{\mathcal{O}}(\hat{\phi}) = \hat{e}]$ is non-negligible for some ppt adversary $\mathcal{A}$.
- R_3 is $\approx_E$-sound iff for all frames ϕ_1, ϕ_2 with the same domain, $(\phi_1, \phi_2) \in R_3$ implies that $([\![\phi_1]\!]_{\mathcal{A}}) \sim ([\![\phi_2]\!]_{\mathcal{A}})$ for any ppt adversary $\mathcal{A}$.

- R_3 is $\approx_E$-complete iff for all frames ϕ_1, ϕ_2 with the same domain, $(\phi_1, \phi_2) \notin R_3$ implies that $(\llbracket \phi_1 \rrbracket_{\mathcal{A}}) \not\sim (\llbracket \phi_2 \rrbracket_{\mathcal{A}})$ for some ppt adversary $\mathcal{A}$.

4 Formal relations

One challenge of the paper is to propose appropriate symbolic relations that correctly abstract computational properties as indistinguishability of two distributions or weak secrecy of some random value (the adversary has only negligible probability to compute it). In this section we provide two symbolic relations (called formal indistinguishability relation and formal non-derivability relation) that are sound abstractions for the two above computational properties.

First we define well-formed relations and we recall a simplified definition of a formal indistinguishability relation as proposed in [7].

Definition 5 (Well-formed relations). *A relation $S_d \subseteq \boldsymbol{F} \times \boldsymbol{T}$ is called* **well-formed** *if $fn(M) \subseteq n(\phi)$ for any $(\phi, M) \in S_d$, and a relation $S_i \subseteq \boldsymbol{F} \times \boldsymbol{F}$ is* **well-formed** *if $dom(\phi_1) = dom(\phi_2)$ for any $(\phi_1, \phi_2) \in S_i$.*

Definition 6. *[FIR [7]] A well-formed relation $\cong \subseteq \boldsymbol{F} \times \boldsymbol{F}$ is called a* **formal indistinguishability relation (FIR for short)** *with respect to the equational theory $=_E$, if $\cong$ is closed with respect to the following closure rules:*
(GE1) If $\phi_1 \cong \phi_2$ then $\phi\phi_1 \cong \phi\phi_2$, for any frame ϕ such that $var(\phi) \subseteq dom(\phi_i)$ and $n(\phi) \cap n(\phi_i) = \emptyset$.
(GE2) $\phi \cong \phi'$ for any frame ϕ' such that $\phi' =_E \phi$.
(GE3) $\tau(\phi) \cong \phi$ for any renaming τ.

This definition is a good starting point to capture indistinguishability in the following sense: if we have a correct implementation of the abstract algebra (i.e. $=_E$ is $=$-sound) and we were provided with some initial relation S (reflecting some computational assumption) which is $\approx$-sound , then the closure of S using the above rules produces a larger relation which still remains $\approx$-sound. But in order to use this definition for real cryptographic constructions , we need to enrich it in several aspects. First, most of constructions which are proposed in the literature, ([9], [28], [22], [24], [26], [10]) use bijective functions (XOR-function or permutations) as basic bricks. To deal with these constructions, we add the following closure rule:
(GE4) If M, N are terms such that $N[M/z] =_E y$, $M[N/y] =_E z$, $var(M) = \{y\}$ and $var(N) = \{z\}$, then for any substitution σ such that $r \notin (fn(\sigma) \cup fn(M) \cup fn(N))$ and $x \notin dom(\sigma)$ it holds $\nu\widetilde{n}.r.\{\sigma, x = M[r/y]\}) \cong \nu\widetilde{n}.r.\{\sigma, x = r\}$.

Second, cryptographic constructions use often hash functions. In ideal models, if one applies a hash function (modeled by random functions [10] or pseudo-random permutations [23]) to a argument that is weakly secret, it returns a random value. And they are quite frequent primitives in cryptography that only ensure weak secrecy. One-way functions only guarantee that an adversary that possesses the image by a one-way function of a random value, has only a negligible probability to compute this value. The computational Diffie-Hellman (CDH)

assumption states that if given the tuple g, g^a, g^b for some randomly-chosen generator g and some random values a, b, it is computationally intractable to compute g^{a*b} (equivalently g^{a*b} is a weakly secret value). This motivates us to introduce the **formal non-derivability relation** as an abstraction of weak secrecy. Let us explain the basic closure rules of this relation. Since we assume that all sorts are implemented by large finite sets of bit strings, it is clearly that *(GD1)* $\nu r.\emptyset \not\vdash r$.

Renaming does not change the concrete semantics of terms or frames.
(GD2) If $\phi \not\vdash M$ then $\tau(\phi) \not\vdash \tau(M)$ for any renaming τ.

If the equational theory is preserved in the computational world, then equivalent terms or frames are indistinguishable.
(GD3) If $\phi \not\vdash M$ then $\phi \not\vdash N$ for any term $N =_E M$.
(GD4) If $\phi \not\vdash M$ then $\phi' \not\vdash M$ for any frame $\phi' =_E \phi$.

If some bit string (concrete implementation of term M) is weakly secret, then any polynomially computation (abstracted by the frame ϕ') does not change this.
(GD5) If $\phi \not\vdash M$ then $\phi'\phi \not\vdash M$ for any frame ϕ' such that $n(\phi') \cap n(\phi) = \emptyset$.

Next rule gives a relationship between indistiguishability and secrecy: if two distributions are indistinguishable, then they leak exactly the same information.
(GD6) For all substitutions σ_1, σ_2 such that $x \notin dom(\sigma_i)$, if $\nu\widetilde{n}.\{\sigma_1, x = M\} \cong \nu\widetilde{n}.\{\sigma_2, x = N\}$ and $\nu\widetilde{n}.\sigma_1 \not\vdash M$ then $\nu\widetilde{n}.\sigma_2 \not\vdash N$.

If the concrete implementation of the symbolic contextual term $T(z)$ is a feasible computation, that is, the adversary has all the needed information to compute $T(\cdot)$ ($fn(T) \cap n(\phi) = \emptyset$), then the concrete implementation of $(T\phi)[M/z]$ is weakly secret only because the implementation of M itself is weakly secret.
(GD7) If $\phi \not\vdash (T\phi)[M/z]$ then $\phi \not\vdash M$, where T is such that $fn(T) \cap n(\phi) = \emptyset$.

One can remark now that *(GD6)* may be generalized to the rule below
(GD6g) If T, U are terms such that $(fn(T) \cup fn(U)) \cap \widetilde{n} = \emptyset$, $z \in var(T) \setminus var(U)$ and $U[T/y] =_E z$, then for all substitutions σ_1, σ_2 such that $x \notin dom(\sigma_i)$ and $\nu\widetilde{n}.\{\sigma_1, x = T[M/z]\} \cong \nu\widetilde{n}.\{\sigma_2, x = T[N/z]\}$ and $\nu\widetilde{n}.\sigma_1 \not\vdash M$ then $\nu\widetilde{n}.\sigma_2 \not\vdash N$.

Actually, *(GD6g)* is consequence of rules *(GD3)*, *(GD6)* and *(GD7)*.

Now the rules that capture hash functions in the ROM: the image by a random function of a weakly secret value is a completely random value.
(HD1) If $\nu\widetilde{n}.r.\sigma[r/h(T)] \not\vdash T$ and $r \notin n(\sigma)$, and if $\sigma[r/h(T)]$ does not contain any subterm of the form $h(\bullet)$, then $\nu\widetilde{n}.\sigma \not\vdash T$.
(HE1) If $\nu\widetilde{n}.r.\sigma[r/h(T)] \not\vdash T$ and $r \notin n(\sigma)$, and if $\sigma[r/h(T)]$ does not contain any subterm of the form $h(\bullet)$, then $\nu\widetilde{n}.r.\sigma \cong \nu\widetilde{n}.r.\sigma[r/h(T)]$.

The definition below formalizes the tight connection between FIR and FNDR.

Definition 7 (FNDR and FIR). *A pair of well formed relations $(\not\vdash, \cong)$ is a pair of* **(formal non-derivability relation, formal indistinguishability relation)** *with respect to the equational theory $=_E$, if $(\not\vdash, \cong)$ is closed with respect to the rules (GD1), ..., (GD7),(GE1),...,(GE4), (HD1),(HE1) and $\cong$ is an equivalence.*

The theorem 1 shows that if a pair (FIR,FNDR) was generated by relations S_d and S_i, then it is sufficient to check only soundness of elements in S_d and S_i to

ensure that the closures $\langle S_d\rangle_{\not\vdash}$ and $\langle S_i\rangle_{\cong}$ are sound. We define $(D_1, I_1) \sqsubset (D_2, I_2)$ if and only if $D_1 \subseteq D_2$ and $I_1 \subseteq I_2$. It is easy to see that $\sqsubset$ is an order.

Theorem 1. *Let (S_d, S_i) be a well-formed pair of relations. Then, it exists a unique smallest (with respect to $\sqsubset$) pair denoted $(\langle S_d\rangle_{\not\vdash}, \langle S_i\rangle_{\cong})$ of (FNDR, FIR) such that $\langle S_d\rangle_{\not\vdash} \supseteq S_d$ and $\langle S_i\rangle_{\cong} \supseteq S_i$. In addition, if $=_E$ is $=$-sound, S_d is $\not\vdash$-sound and S_i is $\approx$-sound, then also $\langle S_d\rangle_{\not\vdash}$ is $\not\vdash$-sound and $\langle S_i\rangle_{\cong}$ is $\approx$-sound.*

The reader should notice that rules *(HE1)* and *(HD1)* can be strengthened if $=_E$ is $=$-faithful: "if $\sigma[r/h(T)]$ does not contain any subterm of the form $h(\bullet)$" can be replaced with "$T \neq_E T'$ for any subterm $h(T')$ of $\sigma[r/h(T)]$".

5 Applications

We apply the framework of Section 4 in order to prove IND-CPA security of several generic constructions for asymmetric encryptions. So we will consider pairs of relations $(\not\vdash, \cong) = (\langle S_d\rangle_{\not\vdash}, \langle S_i\rangle_{\cong})$ generated by some initial sets (S_d, S_i), in different equational theories. We assume that all $=_E$, S_d, S_i that are considered in this section satisfy the conditions of Theorem 1. We emphasize the following fact: adding other equations than those considered does not break the computational soundness of results proved in this section, as long as the computational hypothesis encoded by S_d and S_i still hold.
First we introduce a general abstract algebra that we will extend in order to cover different constructions. We consider three sorts $Data$, $Data^1$, $Data^2$, and the symbols $||$: $Data^1 \times Data^2 \rightarrow Data$, $\oplus_S : S \times S \rightarrow S$, $0_S : S$, with $S \in \{Data, Data^1, Data^2\}$ and $\pi_j : Data \rightarrow Data^j$, with $j \in \{1, 2\}$. For simplicity, we omit S when using $\oplus_S$ or 0_S . The equational theory E_g is generated by:
(XEq1) $x \oplus 0 =_{E_g} x$ *(XEq2)* $x \oplus y =_{E_g} y \oplus x$ *(PEq1)* $\pi_1(x||y) =_{E_g} x$
(XEq2) $x \oplus x =_{E_g} 0$ *(XEq4)* $x \oplus (y \oplus z) =_{E_g} (x \oplus y) \oplus z$ *(PEq2)* $\pi_2(x||y) =_{E_g} y$

$||$ is intended to model concatenation, $\oplus$ is the classical XOR and π_j are the projections. Next rules are consequences of the closure rules from Section 4.
(SyE) If $\phi_1 \cong \phi_2$ then $\phi_2 \cong \phi_1$.
(TrE) If $\phi_1 \cong \phi_2$ and $\phi_2 \cong \phi_3$ then $\phi_1 \cong \phi_3$.
(XE1) If $r \notin (fn(\sigma) \cup fn(T))$ then $\nu\widetilde{n}.r.\{\sigma, x = r \oplus T\} \cong \nu\widetilde{n}.r.\{\sigma, x = r\}$.
(CD1) If $(\phi \not\vdash T_1 \vee \phi \not\vdash T_2)$ then $\phi \not\vdash T_1||T_2$.
(XD1) If $\nu\widetilde{n}.\sigma \not\vdash T$ and $r \notin (\widetilde{n} \cup fn(T))$ then $\nu\widetilde{n}.r.\{\sigma, x = r \oplus T\} \not\vdash T$.

5.1 Trapdoor one-way functions in the symbolic model

We extend the above algebra in order to model trapdoor one-way functions. We add a sort $iData$ and new symbols $f : Data \times Data \rightarrow iData$,$f^{-1} : iData \times Data \rightarrow Data$, $pub : Data \rightarrow Data$. f is a trapdoor permutation, with f^{-1} being the inverse function. We extend the equational theory:
(OEq1) $f^{-1}(f(x, pub(y)), y) =_{E_g} x$.

To simplify the notations, we will use $f_k(\bullet)$ instead of $f(\bullet, pub(k))$. Now we want to capture the one wayness of function f. Computationally, a one-way function only ensures the weakly secrecy of a random argument r (as long

as the key k is not disclosed to the adversary). Hence we define $S_i = \emptyset$ and $S_d = \{(\nu k.r.\{x_k = pub(k), x = f_k(r)\}, r)\}$.

The following frame encodes the Bellare-Rogaway encryption scheme ([10]):
$\phi_{br}(m) = \nu k.r.\{x_k = pub(k), x_a = f_k(r), y = g(r) \oplus m, z = h(m||r)\}$
where m is the plaintext to be encrypted, f is a trapdoor one-way function, and g and h are hash functions (hence oracles in the ROM model).

Now we can see the necessity of p-variables in order to encode IND-CPA security of an encryption scheme. It is not enough to prove that for any two messages m_1 and m_2 the following equivalence holds:

$$\nu k.r.\{x_k = pub(k), x_a = f_k(r), y = g(r) \oplus m_1, z = h(m_1||r)\} \cong$$
$$\nu k.r.\{x_k = pub(k), x_a = f_k(r), y = g(r) \oplus m_2, z = h(m_2||r)\}$$

We did not capture that the adversary is adaptive and she can choose her challenges depending on the public key. We must prove a stronger equivalence: for any terms $p(x_k)$ and $p'(x_k)$,

$$\nu k.r.\{x_k = pub(k), x_a = f_k(r), y = g(r) \oplus p(x_k), z = h(p(x_k)||r)\} \cong$$
$$\nu k.r.\{x_k = pub(k), x_a = f_k(r), y = g(r) \oplus p'(x_k), z = h(p'(x_k)||r)\}$$

The reader noticed that for asymmetric encryption, this suffices to ensure IND-CPA: possessing the public key and having access to hash-oracles allow to encrypt any message (having an oracle to encrypt messages becomes superfluous).

Actually, it suffices to prove $\nu k.r.s.t.\{x_k = pub(k), x_a = f_k(r), y = g(r) \oplus p(x_k), z = h(p(x_k)||r)\} \cong \nu k.r.s.t.\{x_k = pub(k), x_a = f_k(r), y = s, z = t\}$. By transitivity, this implies: for any two challenges that adversary chooses for $p(x_k)$, the distributions she gets are indistinguishable.

Next rules are consequences of the definition of S_d and of the closure rules.
(OD1) If f is a one-way function, then $\nu k.r.\{x_k = pub(k), x = f_k(r)\} \not\succ r$.
(ODg1) If f is a one-way function and $\nu\widetilde{n}.\nu k.\{x_k = pub(k), x = T\} \cong \nu r.\nu k.\{x_k = pub(k), x = r\}$, then $\nu\widetilde{n}.\nu k.\{x_k = pub(k), x = f_k(T)\} \not\succ T$.

The proof of IND-CPA security of Bellare-Rogaway scheme is presented in Figure 1. To simplify the notations, implicitly, all names in frames are restricted and we note $\sigma_2 \equiv x_k = pub(k), x_a = f_k(r)$, and $\sigma_3 \equiv \sigma_2, y = g(r) \oplus p(x_k)$.

5.2 Partially one-way functions in the symbolic model

In this subsection, we show how we can deal with trapdoor partially one-way functions ([24]). We demand for function f a stronger property than one-wayness. Let $Data_1$ be a new sort, and let $f : Data_1 \times Data \times Data \to iData$ and $f^{-1} : iData \times Data \to Data_1$ be functions such that
(OEq1) $f(f^{-1}(x, y), z, pub(y)) =_{E_g} x$.

The function f is said *partially one way*, if for any given $f(r, s, pub(k))$, it is impossible to compute in polynomial time a corresponding r without the trapdoor k. In order to deal with fact that f is partially one-way, we define $S_i = \emptyset$ and $S_d = \{(\nu k.r.s.\{x_k = pub(k), x = f_k(r, s)\}, r)\}$.
The frame below encodes the encryption scheme proposed by Pointcheval ([24]).
$\phi_{po}(m) = \nu k.r.s.\{x_k = pub(k), x_a = f_k(r, h(m||s)), y = g(r) \oplus (m||s)\}$
where m is the plaintext to be encrypted, f is a trapdoor partially one-way function, and g and h are hash functions. To prove IND-CPA security of this

$$\dfrac{\dfrac{HE1\ \dfrac{CD1\ \dfrac{GD5\ \dfrac{HD1\ \dfrac{GD5\ \dfrac{OD1\ \dfrac{}{\{\sigma_2\} \not\vdash r}}{\{\sigma_2, y = s'\} \not\vdash r}}{\{\sigma_2, y = g(r)\} \not\vdash r}}{\{\sigma_2, y = g(r) \oplus p(x_k), z = t\} \not\vdash r}}{\{\sigma_2, y = g(r) \oplus p(x_k), z = t\} \not\vdash p(x_k)||r}}{\{\sigma_2, y = g(r) \oplus p(x_k), z = h(p(x_k)||r)\} \cong \{\sigma_2, y = g(r) \oplus p(x_k), z = t\}}}{\{\sigma_2, y = g(r) \oplus p(x_k), z = h(p(x_k)||r)\} \cong \{x_k = pub(k), x_a = f_k(r), y = s, z = t\}}\ (T1)$$

Fig. 1. Proof of IND-CPA security of Bellare-Rogaway scheme.

$$\dfrac{GE1\ \dfrac{TrE\ \dfrac{GE1\ \dfrac{HE1\ \dfrac{GD5\ \dfrac{OD1\ \dfrac{}{\{\sigma_2\} \not\vdash r}}{\{\sigma_2, y = s\} \not\vdash r}}{\{\sigma_2, y = g(r)\} \cong \{\sigma_2, y = s\}}}{\{\sigma_3\} \cong \{\sigma_2, y = s \oplus p(x_k)\}}\ XE1\ \dfrac{}{\{\sigma_2, y = s \oplus p(x_k)\} \cong \{\sigma_2, y = s\}}}{\{\sigma_2, y = g(r) \oplus p(x_k)\} \cong \{\sigma_2, y = s\}}}{\{\sigma_2, y = g(r) \oplus p(x_k), z = t\} \cong \{\sigma_2, y = s, z = t\}}$$

Fig. 2. Tree $(T1)$ from Figure 1.

scheme, we show that $\nu k.r.s.s_1.s_2\{x_k = pub(k), x_a = f_k(r, h(p(x_k)||s)), y = g(r) \oplus (p(x_k)||s)\} \cong \nu k.r.s.s_1.s_2.\{x_k = pub(k), x_a = f_k(r, s_1), y = s_2\}$.

Next rule is a consequence of the definition of S_d.

(ODp1) If f is a one-way function, then $\nu k.r.s.\{x_k = pub(k), x = f_k(r, s)\} \not\vdash r$.

The proof of IND-CPA security of Pointcheval scheme is presented in Figure 3. To simplify notations we suppose that all names in frames are restricted and we note $\sigma_2 \equiv x_k = pub(k), x_a = f_k(r, h(p(x_k)||s))$ and $\sigma_3 \equiv \sigma_2, y = s_2 \oplus (p(x_k)||s)$.

5.3 Computational Diffie Hellman (CDH) Assumption

In this subsection we prove IND-CPA security of a variant of Hash-ElGamal encryption scheme ([27]) in the random oracle model under the CDH assumption. The proof of the original scheme([6]) can be easily obtained from our proof and it can be done entirely in our framework. We will consider two sorts G and A, symbol functions $exp : G \times A \to G$, $* : A \times A \to A$, $0_A : A$, $1_A : A$, $1_G : G$. We write M^N instead of $exp(M, N)$. We extend E_g by the following equations:

(XEqe1) $(x^y)^z =_{E_g} x^{y*z}$. *(XEqe2)* $x^{1_A} =_{E_g} x$. *(XEqe3)* $x^{0_A} =_{E_g} 1_G$.

To capture the CDH Assumption in the symbolic model we define $S_i = \emptyset$ and $S_d = \{(\nu g.r.s.\{x_g = g, x = g^s, y = g^r\}, g^{s*r})\}$. Then we get the next rule:

(CDH) $\nu g.r.s.\{x_g = g, x = g^s, y = g^r\} \not\vdash g^{s*r}$.

The following frame encodes the Hash-ElGamal encryption scheme.

$\phi_{hel}(m) = \nu g.r.s.\{x_g = g, x = g^s, y = g^r, z = h(g^{s*r}) \oplus m\}$

where m is the plaintext to be encrypted, (g, g^s) is the public key and h is a hash function. The proof of IND-CPA security of Hash-ElGamal's scheme is provided in Figure 6. We supposed that all names are restricted and we noted $\sigma_e \equiv x_g = g, x = g^s, y = g^r$, and $\sigma_f \equiv \sigma_e, z = t \oplus p(x, x_g)$.

$$TrE\ \frac{(T2) \qquad (T3)}{\{\sigma_2, y = g(r) \oplus (p(x_k)||s)\} \cong \{x_k = pub(k), x_a = f_k(r, s_1), y = s_2\}}$$

Fig. 3. Proof of IND-CPA security of Pointcheval scheme.

$$HE1\ \frac{GD6\ \frac{SyE\ \frac{XE1\ \frac{}{\{\sigma_3, x = r\} \cong \{\sigma_2, y = s_2, x = r\}}}{\{\sigma_2, y = s_2, x = r\} \cong \{\sigma_3, x = r\}} \quad GD5\ \frac{ODp1\ \frac{}{\{\sigma_2\} \not\vdash r}}{\{\sigma_2, y = s_2\} \not\vdash r}}{\{\sigma_3\} \not\vdash r}}{\{\sigma_2, y = g(r) \oplus (p(x_k)||s)\} \cong \{\sigma_3\}}$$

Fig. 4. Tree $(T2)$ from Figure 3.

$$TrE\ \frac{XE1\ \frac{}{\{\sigma_3\} \cong \{\sigma_2, y = s_2\}} \quad GE1\ \frac{HE1\ \frac{CD1\ \frac{GD5\ \frac{GD1\ \frac{}{\emptyset \not\vdash s}}{\{x_k = pub(k), x_a = f_k(r, s_1)\} \not\vdash s}}{\{x_k = pub(k), x_a = f_k(r, s_1)\} \not\vdash p(x_k)||s}}{\{\sigma_2\} \cong \{x_k = pub(k), x_a = f_k(r, s_1)\}}}{\{\sigma_2, y = s_2\} \cong \{x_k = pub(k), x_a = f_k(r, s_1), y = s_2\}}}{\{\sigma_3\} \cong \{x_k = pub(k), x_a = f_k(r, s_1), y = s_2\}}$$

Fig. 5. Tree $(T3)$ from Figure 3.

$$TrE\ \frac{GE1\ \frac{HE1\ \frac{GD5\ \frac{CDH\ \frac{}{\{\sigma_e\} \not\vdash g^{s*r}}}{\{\sigma_e, z = t\} \not\vdash g^{s*r}}}{\{\sigma_e, z = h(g^{s*r})\} \cong \{\sigma_e, z = t\}}}{\{\sigma_e, z = h(g^{s*r}) \oplus p(x, x_g)\} \cong \{\sigma_f\}} \quad XE1\ \frac{}{\{\sigma_f\} \cong \{\sigma_e, z = t\}}}{\{x_g = g, x = g^s, y = g^r, z = h(g^{s*r}) \oplus p(x, x_g)\} \cong \{x_g = g, x = g^s, y = g^r, z = t\}}$$

Fig. 6. Proof of IND-CPA security of Hash-ElGamal's scheme

6 Static equivalence and FIR

In this section we adapt the definition of deductibility and static equivalence ([8]) to our framework. After, we justify why they are too coarse to be appropriate abstractions for indistinguishability and weak secrecy. Actually, Proposition 1 states that they are coarser approximations of indistinguishability and weak secrecy than FIR and FNDR.

If ϕ is a frame, and M, N are terms, then we use $(M =_E N)\phi$ for $M\phi =_E N\phi$.

Definition 8 (Deductibility). *A (closed) term T is* **deductible** *from a frame ϕ where $(p_i)_{i\in I} = pvar(\phi)$, written $\phi \vdash T$, if and only if there exists a term M and a set of terms $(M_i)_{i\in I}$, such that $var(M) \subseteq dom(\phi)$, $ar(M_i) = ar(p_i)$, $fn(M, M_i) \cap n(\phi) = \emptyset$ and $(M =_E T)(\phi[(M_i(T_{i_1}, \ldots, T_{i_k})/p_i(T_{i_1}, \ldots, T_{i_k}))_{i\in I}])$. We denote by $\nvdash$ the logical negation of $\vdash$.*

For instance, we consider the frame $\phi = \nu k_1.k_2.s_1.s_2.\{x_1 = k_1, x_2 = k_2, x_3 = h((s_1 \oplus k_1) \oplus p(x_1, x_2)), x_4 = h((s_2 \oplus k_2) \oplus p(x_1, x_2))\}$ and the equational theory E_g. Then $h(s_1)\oplus k_2$ is deductible from ϕ since $h(s_1)\oplus k_2 =_{E_g} x_3[x_1/p(x_1, x_2)]\oplus x_2$ but $h(s_1) \oplus h(s_2)$ is not deductible.

If we consider the frame $\phi' = \nu k.r.s.\{x_k = pub(k), x = f_k(r||s)\}$ where f is a trapdoor one-way function, then neither $r||s$, nor r is deductible from ϕ'. The one-wayness of f is modelled by the impossibility of inverting f if k is not disclosed. While this is fair for $r||s$ according to the computational guarantees of f, it seems too strong of assuming that r alone cannot be computed if f is "just" one-way. This raises some doubts about the fairness of $\not\vdash$ as a good abstraction of weak secrecy. We can try to correct this and add an equation of the form $g(f(x||z, pub(y)), y) =_{E_g} x$. And now, what about r_1, if one gives $f((r_1||r_2)||s)$? In the symbolic setting r_1 is not deductible; computationally, we have no guarantee; hence, when one stops to add equations? Moreover, in this way we could exclude "good" one-way functions: computationally, if f is a one-way function, then $f'(x||y) \stackrel{def}{=} x||f(y)$, is another one-way function. The advantage of defining non-deductibility as we did it in the Section 4, is that first, we capture "just" what is supposed to be true in the computational setting, and second, if we add more equations to our abstract algebra (because we discovered that the implementation satisfies more equations) in a coherent manner with respect to the initial computational assumptions, then our proofs still remain computationally sound. This is not true for $\not\vdash$.

Definition 9. *A* **test** *for a frame ϕ is a tuple $\Upsilon = ((M_i)_{i\in I}, M, N)$ such that $ar(M_i) = ar(p_i)$, $var(M, N) \subseteq dom(\phi)$, $fn(M, N, M_i) \cap n(\phi) = \emptyset$. Then ϕ* **passes** *Υ if and only if $(M =_E N)(\phi[(M_i(T_{i_1}, \ldots, T_{i_k})/p_i(T_{i_1}, \ldots, T_{i_k}))_{i\in I}])$.*

Definition 10 (Statically Equivalent). *Two frames ϕ_1 and ϕ_2 are* **statically equivalent***, written as $\phi_1 \approx_E \phi_2$, if and only if*
(i) $dom(\sigma_1) = dom(\sigma_2)$;
(ii) for any test Υ, ϕ_1 passes the test Υ if and only if ϕ_2 passes the test Υ.

For instance, the two frames $\phi_1 = \nu k.s.\{x_1 = k, x_2 = h(s) \oplus (k \oplus p(x_1))\}$ and $\phi_2 = \nu k.s.\{x_1 = k, x_2 = s \oplus (k \oplus p(x_1))\}$ are statically equivalent with respect to E_g. However the two frames $\phi_1' = \nu k.s.\{x_1 = k, x_2 = h(s) \oplus (k \oplus p(x_1)), x_3 = h(s)\}$ and $\phi_2' = \nu k.s.\{x_1 = k, x_2 = s \oplus (k \oplus p(x_1)), x_3 = h(s)\}$ are not. The frame ϕ_2' passes the test $((x_1), x_2, x_3)$, but ϕ_1' does not.

Let us now consider the equational theory from subsection 5.2. Then the following frames $\nu g.a.b.\{x_1 = g, x_2 = g^a, x_3 = g^b, x_4 = g^{a*b}\}$ and $\nu g.a.b.c.\{x_1 = g, x_2 = g^a, x_3 = g^b, x_4 = g^c\}$ are statically equivalent. This seems right, it is the DDH assumption: a computational implementation that satisfies indistinguishability for the interpretations of this two frames will simply satisfy the DDH assumption. But soundness would imply much more. Even $\nu g.a.b.\{x_1 = g, x_2 = g^a, x_3 = g^b, x_4 = g^{a^2*b^2}\}$ and $\nu g.a.b.c.\{x_1 = g, x_2 = g^a, x_3 = g^b, x_4 = g^c\}$ will be statically equivalent. It is unreasonable to assume that this is true for the computational setting. As for non-deductibility, the advantage of considering FIR as the abstraction of indistinguishability, is that if we add equations in a coherent manner with respect to the initial computational assumptions (that is with S_i), then our proofs still remain computationally sound. The proposition below says that if we consider initial reasonable sets S_d and S_i, then we get finer approximations of indistinguishability and weak secrecy than $\not\vdash$ and $\approx_E$.

Proposition 1. *Let (S_d, S_i) be such that $S_d \subseteq \not\vdash$ and $S_i \subseteq \approx_E$. Then $\langle S_d \rangle_{\not\sim} \subseteq \not\vdash$ and $\langle S_i \rangle_{\cong} \subseteq \approx_E$.*

7 Conclusion

In this paper we developed a general framework for relating formal and computational models for generic encryption schemes in the random oracle model. We proposed general definitions of formal indistinguishability relation and formal non-derivability relation, that is symbolic relations that are computationally sound by construction. We extended previous work with respect to several aspects. First, our framework can cope with adaptive adversaries. This is mandatory in order to prove IND-CPA security. Second, many general constructions use one-way functions, and often they are analyzed in the random oracle model: hence the necessity to capture the weak secrecy in the computational world. Third, the closure rules we propose are designed with the objective of minimizing the initial relations which depend of the cryptographic primitives and assumptions. We illustrated our framework on several generic encryption schemes: we proved IND-CPA security of the scheme proposed by Bellare and Rogaway in [10], of Hash El Gamal [6] and of the scheme proposed by Pointcheval in [24].

As future works, we project to study the (relative) completeness of various equational symbolic theories. Other extensions will be to capture fully active adversaries or exact security (as in [11], we could define indistinguishabiliy as up-to some explicit probability p instead of up-to a negligible probability).

References

1. M. Abadi and P. Rogaway. Reconciling two views of cryptography (the computational soundness of formal encryption). In *IFIP International Conference on Theoretical Computer Science (IFIP TCS)*, Sendai, Japan, 2000. Springer-Verlag.
2. Martín Abadi, Mathieu Baudet, and Bogdan Warinschi. Guessing attacks and the computational soundness of static equivalence. In *FoSSaCS*, volume 3921 of *LNCS*, pages 398–412. Springer, 2006.
3. Martín Abadi and Andrew D. Gordon. A bisimulation method for cryptographic protocols. In *ESOP*, volume 1381 of *LNCS*, pages 12–26. Springer, 1998.
4. M. Backes and B. Pfitzmann. Symmetric encryption in a simulatable dolev-yao style cryptographic library. In *CSFW*, pages 204–218. IEEE , 2004.
5. M. Backes, B. Pfitzmann, and M. Waidner. Symmetric authentication within a simulatable cryptographic library. In *ESORICS*, volume 2808 of *LNCS*, pages 271–290. Springer, 2003.
6. Joonsang Baek, Byoungcheon Lee, and Kwangjo Kim. Secure length-saving elgamal encryption under the computational diffie-hellman assumption. In *ACISP*, volume 1841 of *LNCS*, pages 49–58. Springer, 2000.
7. Gergei Bana, Payman Mohassel, and Till Stegers. Computational soundness of formal indistinguishability and static equivalence. In Mitsu Okada and Ichiro Satoh, editors, *ASIAN*, volume 4435 of *LNCS*, pages 182–196. Springer, 2006.

8. Mathieu Baudet, Véronique Cortier, and Steve Kremer. Computationally sound implementations of equational theories against passive adversaries. In *ICALP*, volume 3580 of *LNCS*, pages 652–663. Springer, 2005.
9. M. Bellare and P. Rogaway. Optimal asymmetric encryption. In *EUROCRYPT'04*, volume 950 of *LNCS*, pages 92–111, 1994.
10. Mihir Bellare and Phillip Rogaway. Random oracles are practical: a paradigm for designing efficient protocols. In *CCS'93*, pages 62–73, 1993.
11. Bruno Blanchet and David Pointcheval. Automated security proofs with sequences of games. In *CRYPTO'06*, volume 4117 of *LNCS*, pages 537–554, 2006.
12. Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *FOCS*, pages 136–145, 2001.
13. Ran Canetti and Jonathan Herzog. Universally composable symbolic analysis of mutual authentication and key-exchange protocols. In Shai Halevi and Tal Rabin, editors, *TCC*, volume 3876 of *LNCS*, pages 380–403. Springer, 2006.
14. V. Cortier and B. Warinschi. Computationally sound, automated proofs for security protocols. In Sagiv [25], pages 157–171.
15. Judicaël Courant, Marion Daubignard, Cristian Ene, Pascal Lafourcade, and Yassine Lakhnech. Towards automated proofs for asymmetric encryption schemes in the random oracle model. In *CCS'2008*, pages 371–380. ACM, 2008.
16. D. Dolev and A. C. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–208, 1983.
17. U. Feige, A. Fiat, and A. Shamir. Zero-knowledge proofs of identity. *J. Cryptol.*, 1(2):77–94, 1988.
18. S. Goldwasser and S. Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, April 1984.
19. R. Janvier, Y. Lakhnech, and L. Mazaré. Completing the picture: Soundness of formal encryption in the presence of active adversaries. In Sagiv [25], 172–185.
20. P. Laud. Symmetric encryption in automatic analyses for confidentiality against adaptive adversaries. In *Symposium on Security and Privacy*, pages 71–85, 2004.
21. D. Micciancio and B. Warinschi. Soundness of formal encryption in the presence of active adversaries. *Theory of Cryptography Conference*, 133–151. Springer, 2004.
22. T. Okamoto and D. Pointcheval. React: Rapid enhanced-security asymmetric cryptosystem transform. In *CT-RSA'01*, pages 159–175, 2001.
23. Duong Hieu Phan and David Pointcheval. About the security of ciphers (semantic security and pseudo-random permutations). In *Selected Areas in Cryptography*, volume 3357 of *LNCS*, pages 182–197. Springer, 2004.
24. D. Pointcheval. Chosen-ciphertext security for any one-way cryptosystem. In *PKC'00*, pages 129–146, 2000.
25. Shmuel Sagiv, editor. *Programming Languages and Systems, 14th European Symposium on Programming,ESOP 2005, April 4-8*, volume 3444 of *LNCS*, 2005.
26. V. Shoup. Oaep reconsidered. *J. Cryptology*, 15(4):223–249, 2002.
27. Victor Shoup. Sequences of games: a tool for taming complexity in security proofs. cryptology eprint archive, report 2004/332, 2004.
28. Y. Zheng and J. Seberry. Immunizing public key cryptosystems against chosen ciphertext attacks. *J. on Selected Areas in Communications*, 11(5):715–724, 1993.