Deciding knowledge in security protocols under equational theories

Martín Abadi*
Computer Science Department
University of California
at Santa Cruz
USA

Véronique Cortier †
Loria – CNRS
Nancy
France

October 18, 2005

Abstract

The analysis of security protocols requires precise formulations of the knowledge of protocol participants and attackers. In formal approaches, this knowledge is often treated in terms of message deducibility and indistinguishability relations. In this paper we study the decidability of these two relations. The messages in question may employ functions (encryption, decryption, etc.) axiomatized in an equational theory. One of our main positive results says that deducibility and indistinguishability are both decidable in polynomial time for a large class of equational theories. This class of equational theories is defined syntactically and includes, for example, theories for encryption, decryption, and digital signatures. We also establish general decidability theorems for an even larger class of theories. These theorems require only loose, abstract conditions, and apply to many other useful theories, for example with blind digital signatures, homomorphic encryption, XOR, and other associative-commutative functions.

1 Introduction

Understanding security protocols often requires reasoning about the knowledge of legitimate protocol participants and attackers. As a simple example, let us consider a protocol in which A sends to B a message that consists of a secret s encrypted under a pre-arranged shared key k. One may argue that, after processing this message, B knows s. More interestingly, one may also argue than an attacker with bounded computing power that does not know k but eavesdrops on the communications between A and B and sees the message does not learn s.

*Martín Abadi's work was partly supported by the National Science Foundation under Grants CCR-0204162 and CCR-0208800.

†Véronique Cortier's work was partly supported by the RNTL project PROUVE-03V360 and the ACI Jeunes Chercheurs JC9005. Corresponding author address: Loria, Campus Scientifique, BP 239, 54506 Vandoeuvre-les-Nancy cedex, Nancy, France; Veronique.Cortier@loria.fr.

Accordingly, formal methods for the analysis of security protocols rely on definitions of the knowledge of protocol participants and attackers. In those methods, the knowledge of an attacker is used to determine what messages the attacker can send at each point in time—it can send only messages it knows. Moreover, security guarantees can be phrased in terms of the knowledge of the attacker. For example, a guarantee might be that, at the end of a protocol run, the attacker does not know a particular key, or that the attacker does not know whether a certain ciphertext contains the plaintext "true" or "false". For such applications, although the attacker is typically an active entity that can learn by conducting experiments, the definition of knowledge focuses on a particular point in a protocol execution.

Many formal definitions explain the knowledge of an attacker in terms of message deduction (e.g., [25, 29, 33, 30]). Given a set of messages S and another message M, one asks whether M can be computed from S. The messages are represented by expressions, and correspondingly the computations allowed are symbolic manipulations of those expressions. Intuitively these computations can rely on any step that an eavesdropper who has obtained the messages in S can perform on its own in order to derive M. For example, the eavesdropper can encrypt and decrypt using known keys, and it can extract parts of messages.

Despite its usefulness in proofs about protocol behaviors, the concept of message deduction does not always provide a sufficient account of knowledge, and it is worthwhile to consider alternatives. For instance, suppose that we are interested in a protocol that transmits an encrypted boolean value, possibly a different one in each run. We might like to express that this boolean value remains secret by saying that no attacker can learn it by eavesdropping on the protocol. On the other hand, it is unreasonable to say that an attacker cannot deduce the well-known boolean values "true" and "false". Instead, we may say that the attacker cannot distinguish an instance of the protocol with the value "true" from one with the value "false". More generally, we may say that two systems are equivalent when an attacker cannot distinguish them, and we may then express security guarantees as equivalences. The use of equivalences is common in computational approaches to cryptography (e.g., [24]), and it also figures prominently in several formal methods (e.g., [5, 28, 3]).

Two systems that output messages that an attacker can tell apart are obviously distinguishable. Conversely, in order to establish equivalences between systems, an important subtask is to establish equivalences between the messages that the systems generate (for example, between the encrypted boolean values). These equivalences may be called static equivalences, because they consider only the messages, not the dynamic processes that generate them. Analogously, the deduction relation should perhaps be called static deduction. Despite the static character of these relations, they are useful in analyzing the dynamics of protocols and attacks. In particular, proof methods for properties of protocol behaviors often rely on deduction (e.g., [30]), and process equivalences can be reduced to static equivalences plus fairly standard bisimulation conditions [3] (see also [4, 14]).

In this paper we study the decidability of deduction and static equivalence. We define a relation $\phi \vdash M$ that means that M can be deduced from ϕ, and a relation $\varphi \approx_s \psi$ that means that φ and ψ are statically equivalent; here ϕ, φ, and ψ are all essentially lists of messages, each with a name, represented by formal expressions. For

generating these messages, we allow the application of a wide array of functions—pairing, projections, various flavors of encryption and decryption, digital signatures, one-way hash functions, etc.. Indeed, our results do not make any assumption on any particular cryptographic system beyond fairly general hypotheses on the equational theory that is used for defining the properties of the cryptographic operations.

Our results start with basic observations about the decidability of deduction and static equivalence. Specifically, we demonstrate that, even for decidable equational theories, $\phi \vdash M$ and $\varphi \approx_s \psi$ can be undecidable. Moreover, we establish that deduction can be reduced to static equivalence (not too surprisingly), but that the converse does not hold. Therefore, we investigate hypotheses that would guarantee decidability, allowing for the possibility that the decidability of $\varphi \approx_s \psi$ requires more than the decidability of $\phi \vdash M$.

We identify a simple, syntactically defined class of theories for which $\phi \vdash M$ and $\varphi \approx_s \psi$ are both decidable in polynomial time. These theories, which we call convergent subterm theories, are given by convergent rewriting systems with a finite number of rules of the form $M \to N$ where N is a proper subterm of M or a constant symbol. Convergent subterm theories appear frequently in applications; in particular, standard axiomatizations of encryption, decryption, and digital signatures yield convergent subterm theories.

Going further, we develop decision methods for $\phi \vdash M$ and $\varphi \approx_s \psi$ under an even larger class of equational theories. For this purpose, we assume only loose, abstract conditions, rather than syntactic criteria on the theories. In this respect, we are inspired by Comon-Lundh's current investigations [18] (see Section 6). The general decidability theorems that we obtain subsume the previous ones for convergent subterm theories (with more difficulties and without the same complexity bounds, hence the separate treatment of convergent subterm theories). They also apply to many other useful theories, for example with blind digital signatures, homomorphic encryption, XOR, and other AC functions. Several of the decidability results that we obtain are new.

Checking that a particular theory satisfies our abstract conditions may involve some work, though often less than direct proofs of decidability. In some cases, it may also involve some (fairly elementary and pleasant) mathematics, such as facts on $\mathbb{Z}$-modules. We expect that some of the techniques that we employ in our examples may be reused in the study of other theories.

The problem of deciding knowledge is particularly important in the context of algorithms and tools for automated protocol analysis. Often, special techniques are introduced for particular sets of cryptographic operations of interest, on a case-by-case basis. For example, the classic Dolev-Yao result deals with a fixed, limited suite of public-key operations [23]; more recent decidability results deal with XOR and modular exponentiation (e.g., [16, 17, 19]); many variants and combinations that arise in practice have not yet been explored. On the other hand, other algorithms and tools (e.g., [10, 11, 12]) allow much freedom in the choice of cryptographic operations but their analysis of the knowledge of the attacker is not always guaranteed to terminate. Decidability results under general equational theories have been rare. Comon-Lundh and Treinen have studied the decidability of deduction for a class of equational theories in which, for example, they allow the homomorphism property $\mathsf{enc}(\langle u, v\rangle, k) = \langle \mathsf{enc}(u, k), \mathsf{enc}(v, k)\rangle$ but not the inverse property $I(I(x)) = x$ [20].

These examples illustrate that their class is incomparable with the class of convergent subterm theories; we do not know how their class relates to our results for other theories. Delaune and Jacquemard have shown that deduction is decidable for a subclass of convergent subterm theories, also considering active attacks [21]. These results do not address static equivalence, nor allow associativity and commutativity properties. In fact, even results on specific theories with AC (associative-commutative) functions have been rare. Three important exceptions are decidability results for deduction with XOR [17, 19], in an Abelian group [19], and under certain "AC-like" theories with homomorphisms [27]. We discuss other recent and ongoing related work below.

The next section, Section 2, introduces notations and definitions. Section 3 compares $\vdash$ and $\approx_s$. Section 4 focuses on convergent subterm theories and gives our main decidability results for these theories. In Section 5, we consider the larger class of equational theories. Section 6 concludes. Some proofs appear in the Appendix.

Parts of this paper have been presented, in preliminary form, at ICALP 2004 and CSFW 2005 [1, 2]. This paper represents a synthesis and an extension of the work presented there.

2 Basic definitions

Next we review definitions from previous work. We mostly adopt the definitions of the applied pi calculus [3]. In Section 2.1 we give the syntax of expressions. In Section 2.2 we explain a representation for the information available to an observer who has seen messages exchanged in the course of a protocol execution. In Section 2.3 and 2.4 we present the relations $\vdash$ and $\approx_s$, which (as explained in the introduction) provide two formalizations of the knowledge that the observer has on the basis of that information.

2.1 Syntax

A *signature* Σ consists of a finite set of function symbols, such as enc and pair, each with an arity. We write $\mathsf{arity}(f)$ for the arity of a function symbol f, and let $\mathsf{ar}(\Sigma)$ be the maximal arity of a function symbol in Σ. A function symbol with arity 0 is a constant symbol.

Given a signature Σ, an infinite set of names $\mathcal{N}$, and an infinite set of variables, the set of *terms* is defined by the grammar:

$L, M, N, T, U, V ::=$	terms
$k, \ldots, n, \ldots, s$	name
x, y, z	variable
$f(M_1, \ldots, M_l)$	function application

where f ranges over the function symbols of Σ and l matches the arity of f. Although names, variables, and constant symbols have similarities, we find it clearer to keep them separate. A term is closed when it does not have free variables (but it may contain names and constant symbols). We write $\mathit{fn}(M)$ for the set of names that occur in the term M. We use meta-variables u, v, w to range over names and variables. The *size* $|T|$

of a term T is defined by $|u| = 1$ and $|f(T_1, \ldots, T_l)| = 1 + \sum_{i=1}^{l} |T_i|$. The *DAG-size* $|T|_{\mathsf{DAG}}$ is the number of distinct subterms of T.

We equip the signature Σ with an equational theory E, that is, an equivalence relation on terms that is closed under application of contexts and under substitutions of terms for both names and variables. (While non-standard, the requirement that E be closed under substitutions of terms for names simplifies some technical details and has been harmless in applications.) We write $M =_E N$ when M and N are closed terms and the equation $M = N$ is in E. We use the symbol $==$ to denote syntactic equality of closed terms. As in these definitions, we often focus on closed terms for simplicity.

2.2 Assembling terms into frames

After a protocol execution, an attacker may know a sequence of messages $M_1, \ldots, M_l$. This means that it knows each message but it also knows in which order it received the messages. So it is not enough for us to say that the attacker knows the set of terms $\{M_1, \ldots, M_l\}$. Furthermore, we should distinguish those names that the attacker had before the execution from those that were freshly generated and which may remain secret from the attacker; both kinds of names may appear in the terms.

In the applied pi calculus, such a sequence of messages is organized into a *frame* $\nu\widetilde{n}\sigma$, where $\widetilde{n}$ is a finite set of names (intuitively, the fresh names), and σ is a substitution of the form:

$$\{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}\} \quad \text{with} \quad dom(\sigma) \stackrel{\text{def}}{=} \{x_1, \ldots, x_l\}$$

The variables enable us to refer to each M_i, for example for keeping track of their order of transmission. We always assume that the terms M_i are closed. The size of a frame $\phi = \nu\widetilde{n}\{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}\}$ is $|\phi| \stackrel{\text{def}}{=} \sum_{i=1}^{l} |M_i|$.

2.3 Deduction

Given a frame ϕ that represents the information available to an attacker, we may ask whether a given term closed M may be deduced from ϕ. This relation is written $\phi \vdash M$ (following Schneider [33]). It is axiomatized by the rules:

$$\frac{}{\nu\widetilde{n}\sigma \vdash M} \text{ if } \exists x \in dom(\sigma) \text{ s.t. } x\sigma = M \qquad \frac{}{\nu\widetilde{n}\sigma \vdash s} \; s \notin \widetilde{n}$$

$$\frac{\phi \vdash M_1 \quad \cdots \quad \phi \vdash M_k}{\phi \vdash f(M_1, \ldots, M_k)} \; f \in \Sigma \qquad \frac{\phi \vdash M \quad M =_E M'}{\phi \vdash M'}$$

Since the deducible messages depend on the underlying equational theory, we write $\vdash_E$ when E is not clear from the context. Intuitively, the deducible messages are the messages of ϕ and the names which are not protected in ϕ, closed by equality in E and closed by application of functions. We have the following characterization of deduction:

Proposition 1 *Let M be a closed term and $\nu\widetilde{n}\sigma$ be a frame. Then $\nu\widetilde{n}\sigma \vdash M$ if and only if there exists a term ζ such that $fn(\zeta) \cap \widetilde{n} = \emptyset$ and $\zeta\sigma =_E M$.*

As an example, we consider the equational theory of pairing and symmetric encryption. The signature is $\Sigma_{\mathsf{enc}} = \{\mathsf{pair}, \mathsf{enc}, \mathsf{fst}, \mathsf{snd}, \mathsf{dec}\}$. As usual, we write $\langle x, y\rangle$ instead of $\mathsf{pair}(x, y)$. The theory E_{enc} is defined by the axioms:

$$\mathsf{fst}(\langle x, y\rangle) = x \qquad \mathsf{snd}(\langle x, y\rangle) = y \qquad \mathsf{dec}(\mathsf{enc}(x, y), y) = x$$

Let $\phi \stackrel{\text{def}}{=} \nu k, s\{\mathsf{enc}(s, k)/x, k/y\}$. Then $\phi \vdash k$ and $\phi \vdash s$. Furthermore, we have $k =_{E_{\mathsf{enc}}} y\phi$ and $s =_{E_{\mathsf{enc}}} \mathsf{dec}(x, y)\phi$.

2.4 Static equivalence

Deduction does not always suffice for expressing the knowledge of an attacker, as discussed in the introduction. For example, consider $\phi_1 \stackrel{\text{def}}{=} \nu k\{\mathsf{enc}(0, k)/x, k/y\}$ and $\phi_2 \stackrel{\text{def}}{=} \nu k\{\mathsf{enc}(1, k)/x, k/y\}$, where $0, 1 \in \Sigma$ are constant symbols. The attacker can deduce the same set of terms from these two frames since it knows 0 and 1. But it could tell the difference between these two frames by checking whether the decryption of x with y produces 0 or 1.

We say that two terms M and N are equal in the frame φ for the equational theory E, and write $(M =_E N)\varphi$, if and only if $\varphi = \nu\widetilde{n}.\sigma$, $M\sigma =_E N\sigma$, and $\{\widetilde{n}\} \cap (fn(M) \cup fn(N)) = \emptyset$ for some names $\widetilde{n}$ and substitution σ. Then we say that two frames φ and ψ are *statically equivalent*, and write $\varphi \approx_s \psi$, when $dom(\varphi) = dom(\psi)$ and when, for all terms M and N, we have $(M =_E N)\varphi$ if and only if $(M =_E N)\psi$. We write $\approx_{sE}$ when E is not clear from the context.

In our example, we have $(\mathsf{dec}(x, y) =_{E_{\mathsf{enc}}} 0)\phi_1$ but not $(\mathsf{dec}(x, y) =_{E_{\mathsf{enc}}} 0)\phi_2$. Therefore, $\phi_1 \not\approx_s \phi_2$ although $\nu k\{\mathsf{enc}(0, k)/x\} \approx_s \nu k\{\mathsf{enc}(1, k)/x\}$.

3 Comparison of deduction and static equivalence

We compare equality, deduction, and static equivalence from the point of view of decidability. There is little hope that deduction or static equivalence would be decidable when equality itself is not. (We note however that, for some artificial, especially designed equational theories, deduction may be decidable while equality is undecidable.) Therefore, we focus on equational theories for which equality is at least decidable.

3.1 $\vdash$ may be undecidable

Unfortunately, the decidability of equality is not sufficient for the decidability of deduction and static equivalence. As evidence, let us consider the signature $\Sigma = \{f, \cdot, [\_, \_, \_]\}$ where f is a unary functional symbol, $\cdot$ is a binary functional symbol, and $[\,]$ is a ternary functional symbol, and the equational theory E_{pc} defined by:

$$\begin{aligned} x \cdot (y \cdot z) &= (x \cdot y) \cdot z \\ [x_1, y_1]^z \cdot [x_2, y_2]^z &= [x_1 \cdot x_2, y_1 \cdot y_2]^z \\ f([x, x]^y) &= y \end{aligned}$$

According to these equations, the symbol $\cdot$ is associative and distributes over the symbol $[\,]$, and any term of the form $f([M, M]^k)$ can be collapsed to k. Note that E_{pc} is decidable since orienting the two last equations from left to right leads to a confluent rewriting system. On the other hand, this equational theory enables us to encode the Post Correspondence Problem (PCP) into the deduction problem. The PCP is: given a finite number of pairs of words $(u_i, v_i)_{1\leq i\leq n}$ on the alphabet $A \subset \mathcal{N}$, does there exist a sequence $s_1, \ldots, s_k \in \{1..n\}^*$ such that $u_{s_1} \cdots u_{s_k} = v_{s_1} \cdots v_{s_k}$? We have:

Proposition 2 *Given the PCP instance $(u_i, v_i)_{1\leq i\leq n}$ on the alphabet $A \subset \mathcal{N}$, we define the substitution $\sigma = \{[u_i, v_i]^k/x_i\}$. Then there exists a solution to the PCP instance if and only if $(\nu k)\sigma \vdash_{E_{\mathsf{pc}}} k$.*

It follows:

Proposition 3 *The deduction problem for E_{pc} ($\vdash_{E_{\mathsf{pc}}}$) is undecidable.*

In order to prove Proposition 2, we characterize the terms deducible from $(\nu k)\sigma$. Let Pub be the set of terms built from the names $\mathcal{N} \setminus k$ and the function symbols f, $\cdot$, and $[\,]$ (the *public* terms). Let $\mathcal{L}$ be the set of all terms of the form:

$$[u_{s_1} \cdot \cdots \cdot u_{s_p}, v_{s_1} \cdot \cdots \cdot v_{s_p}]^k$$

where $s_1, \ldots, s_p \in \{1..n\}$. We define the set WF of *well-formed* terms by the grammar:

$$\mathsf{WF} := \mathcal{L} \mid \mathsf{Pub} \mid f(\mathsf{WF}) \mid \mathsf{WF} \cdot \mathsf{WF} \mid [\mathsf{WF}, \mathsf{WF}]^{\mathsf{WF}}$$

Note that if $T \in \mathsf{WF}$ then $T \neq k$ (by induction on the construction of WF).

Lemma 1 *The terms deducible from $(\nu k)\sigma$ are, modulo E_{pc}, in the set WF of well-formed terms.*

This lemma is proved by induction on the construction of deducible terms.

- For every variable x_i, $x_i\sigma$ is well-formed.
- For any name $n \in \mathcal{N}$, if $n \neq k$, then n is well-formed, since $n \in \mathsf{Pub}$.
- If T_1, T_2, and T_3 are well-formed modulo E_{pc}, then $f(T_1)$, $T_1 \cdot T_2$, and $[T_1, T_2]^{T_3}$ are also well-formed modulo E_{pc}.
- If T_1 is well-formed modulo E_{pc} and $T_1 =_{E_{\mathsf{pc}}} T_2$, then T_2 is also well-formed modulo E_{pc}.

We also characterize terms equal to k modulo E_{pc}.

Lemma 2 *Let T be a term. If $k =_{E_{\mathsf{pc}}} T$ and $T \neq k$ then T is of the form:*

$$f([T_1, T_1']^{U_i} \cdots [T_m, T_m']^{U_m})$$

with $U_i =_{E_{\mathsf{pc}}} k$ and $T_1 \cdots T_m =_{E_{\mathsf{pc}}} T_1' \cdots T_m'$.

This lemma is proved by induction on the number of applications of equalities that establish $k =_{E_{\mathsf{pc}}} T$. The only equation that can yield k is $f([x,x]^y) = y$, which leads to a term of the specified form in the base case. In the inductive step, if $T =_{E_{\mathsf{pc}}} T'$ with $T' = f([T_1, T_1']^{U_i} \cdots [T_m, T_m']^{U_m})$, $U_i =_{E_{\mathsf{pc}}} k$, and $T_1 \cdots T_m =_{E_{\mathsf{pc}}} T_1' \cdots T_m'$, and only one equation has been applied to establish $T =_{E_{\mathsf{pc}}} T'$, then

- either the equation has been applied inside one of the terms T_i, T_i' or U_i, and in that case, the property holds immediately,
- or the equation has been applied above the terms T_i, T_i', and U_i; and then either $T = k$ or only the two first equations can have been applied, and in either case the property holds.

Lemma 3 *Let T be a term. If $T =_{E_{\mathsf{pc}}} k$ then T contains k as a subterm.*

This lemma is proved by induction on the size of T. In the base case, $T = k$, and the property holds immediately. In the inductive step, T is of the form $f([T_1, T_1']^{U_i} \cdots [T_m, T_m']^{U_m})$ with $U_i =_{E_{\mathsf{pc}}} k$, by Lemma 2, and by induction hypothesis we obtain that the terms U_i contain k as a subterm, so T contains k as a subterm. An easy consequence of this lemma is that if $T =_{E_{\mathsf{pc}}} k$ then $T \notin \mathsf{Pub}$.

Returning to Proposition 2, let us assume that there exists a solution to a given PCP instance. This assumption means that there exists a sequence $s_1, \ldots, s_p \in \{1..n\}^*$ such that $u_{s_1} \cdots u_{s_p} = v_{s_1} \cdots v_{s_p}$. Then

$$\begin{array}{rcl} f(x_{s_1} \cdots x_{s_k})\sigma & = & f([u_{s_1}, v_{s_1}]^k \cdots [u_{s_p}, v_{s_p}]^k) \\ & =_{E_{\mathsf{pc}}} & f([u_{s_1} \cdots u_{s_p}, v_{s_1} \cdots v_{s_p}]^k) \\ & =_{E_{\mathsf{pc}}} & k \end{array}$$

so k is deducible.

Conversely, assume that k is deducible. By Lemma 1, k must be equal modulo E_{pc} to some term $T \in \mathsf{WF}$. We show by induction on the size of T that there exists a solution to the PCP instance. By Lemma 2 and since $T \neq k$ (since $T \in \mathsf{WF}$), T must be of the form $f([T_1, T_1']^{U_i} \cdots [T_m, T_m']^{U_m})$ with $U_i =_{E_{\mathsf{pc}}} k$ and $T_1 \cdots T_m =_{E_{\mathsf{pc}}} T_1' \cdots T_m'$. Since T cannot be public, $T \in \mathsf{WF}$ implies that the term $T' \stackrel{\text{def}}{=} [T_1, T_1']^{U_i} \cdots [T_m, T_m']^{U_m}$ must be well-formed. If one of the terms U_i is well-formed, we conclude by induction hypothesis, since $U_i =_{E_{\mathsf{pc}}} k$. On the other hand, if none of the terms U_i is well-formed, we proceed as follows. Since $U_i =_{E_{\mathsf{pc}}} k$ and by Lemma 3, all the terms $[T_i, T_i']^{U_i}$ contain k as a subterm, so they are not public. By inspection of the cases in the definition of WF, we deduce that each $[T_i, T_i']^{U_i}$ must be in WF. Since none of the terms U_i is well-formed, we must have that each $[T_i, T_i']^{U_i}$ is in $\mathcal{L}$, so T' is actually equal (syntactically) to

$$[u_{s_1^1} \cdot \cdots \cdot u_{s_{p_1}^1}, v_{s_1^1} \cdot \cdots \cdot v_{s_{p_1}^1}]^k \cdots [u_{s_1^m} \cdot \cdots \cdot u_{s_{p_m}^m}, v_{s_1^m} \cdot \cdots \cdot v_{s_{p_m}^m}]^k$$

with

$$u_{s_1^1} \cdots u_{s_{p_1}^1} \cdots u_{s_1^m} \cdots u_{s_{p_m}^m} = v_{s_1^1} \cdots v_{s_{p_1}^1} \cdots v_{s_1^m} \cdots v_{s_{p_m}^m}$$

for some $s_i^j \in \{1..n\}$. Therefore, there exists a solution to the PCP instance.

3.2 $\vdash$ reduces to $\approx_s$

Next we show that deduction may be reduced to static equivalence by adding only one free unary function symbol (a unary function symbol with no added equations). Thus, the equational theory is basically unchanged in the reduction—it can be given by a fixed set of equational axioms. We leave as an open problem whether the reduction is always possible without even any change to the signature.

Proposition 4 *Let E be an equational theory over some signature Σ. We define $\Sigma' \stackrel{def}{=} \Sigma \uplus \{\mathsf{h}\}$, where h is unary, and let E' be the least equational theory that extends E to terms over Σ'. Let $\phi = \nu\widetilde{n}\{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}\}$ be a frame over Σ, M be a closed term over Σ, and k be a fresh name. Then $\phi \vdash_E M$ if and only if*

$$\nu\widetilde{n}\{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}, {}^{\mathsf{h}(M)}\!/_{x_{l+1}}\} \not\approx_{s\,E'} \nu(\widetilde{n} \cup \{k\})\{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}, {}^{k}\!/_{x_{l+1}}\}$$

We derive that if $\approx_{s\,E'}$ is decidable, then $\vdash_E$ is also decidable (with at most the same complexity).

In order to prove the proposition, we first introduce some notation. We let $\sigma = \{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}\}$, so $\phi = \nu\widetilde{n}\sigma$, and let $\phi_1 = \nu\widetilde{n}\sigma_1$ with $\sigma_1 = \{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}, {}^{\mathsf{h}(M)}\!/_{x_{l+1}}\}$ and $\phi_2 = \nu(\widetilde{n} \cup \{k\})\sigma_2$ with $\sigma_2 = \{{}^{M_1}\!/_{x_1}, \ldots, {}^{M_l}\!/_{x_l}, {}^{k}\!/_{x_{l+1}}\}$.

One direction of Proposition 4 follows easily from Proposition 1. If $\phi \vdash_E M$ then Proposition 1 implies that there exists a term ζ such that $fn(\zeta) \cap \widetilde{n} = \emptyset$ and $\zeta\sigma =_E M$; then $\phi_1 \not\approx_{s\,E'} \phi_2$ because $(\mathsf{h}(\zeta) =_{E'} x_{l+1})\phi_1$ while $(\mathsf{h}(\zeta) \neq_{E'} x_{l+1})\phi_2$.

For the other direction, we use a weak version of a lemma due to Baudet et al. [8]. Given a term $U == \mathsf{h}(U_1)$ and given a name a, the *cutting function* $\mathsf{cut}_{U,a}$ is defined recursively as follows:

$$\begin{array}{rcl} \mathsf{cut}_{U,a}(u) & = & u \quad \text{if } u \text{ is a name or a constant} \\ \mathsf{cut}_{U,a}(g(T_1, \ldots, T_k)) & = & \left\{ \begin{array}{l} a \quad \text{if } g = \mathsf{h},\ k = 1, \text{ and } U_1 =_{E'} T_1 \\ g(\mathsf{cut}_{U,a}(T_1), \ldots, \mathsf{cut}_{U,a}(T_k)) \quad \text{otherwise} \end{array} \right. \end{array}$$

Intuitively, $\mathsf{cut}_{U,a}(T)$ is obtained from T by replacing with a the subterms equal to U modulo E' and whose head symbol is h. The following lemma (adapted from [8]) states that, if an equality holds between terms that mention h, then the equality still holds after cutting subterms whose head symbol is h.

Lemma 4 *Let $U == \mathsf{h}(U_1)$. If $M =_{E'} N$ then $\mathsf{cut}_{U,a}(M) =_{E'} \mathsf{cut}_{U,a}(N)$.*

This lemma relies on the following characterization of E': it is the least transitive relation that contains the equations $L' =_{E'} R'$ for which there exists an equation $L =_E R$, a substitution θ, and a position p such that $L'|_p == L\theta$ and $R' == L'[R\theta]_p$. (As usual, a position is formalized as a sequence of integers that indicates a path in a term; $M|_p$ represents the subterm of M at position p, and $M[R\theta]_p$ is obtained by replacing that subterm with $R\theta$; see Definition 11 in Appendix B.) The lemma is proved by induction on the number of applications of equalities $L =_E R$ required for obtaining $M =_{E'} N$. For the base case, we assume assume that $M =_{E'} N$ and that there exists an equation $L =_E R$, a substitution θ, and a position p such that $M|_p == L\theta$ and $N == M[R\theta]_p$. We consider two cases, distinguished by whether the cutting function $\mathsf{cut}_{U,a}$ cuts a subterm of M above p or not:

1. In the first case, there exists a strict prefix p' of p such that $M|_{p'} == \mathsf{h}(T_1)$ with $U_1 =_{E'} T_1$. We consider the smallest p' that satisfies this property, and let $p = p'.1.p''$, so $N == M[\mathsf{h}(T_1[R\theta]_{p''})]_{p'}$. Since $T_1[R\theta]_{p''} =_{E'} T_1[L\theta]_{p''} == T_1 =_{E'} U_1$, both $\mathsf{h}(T_1)$ and $\mathsf{h}(T_1[R\theta]_{p''})$ are replaced with a by the cutting function, so $\mathsf{cut}_{U,a}(M) == \mathsf{cut}_{U,a}(N)$.

2. In the second case, any p' such that $M|_{p'} = \mathsf{h}(T_1)$ with $U_1 =_{E'} T_1$ is at least as long as p or incomparable. Therefore, $\mathsf{cut}_{U,a}(M[x]_p) == \mathsf{cut}_{U,a}(N[x]_p)$ and $\mathsf{cut}_{U,a}(M) == \mathsf{cut}_{U,a}(M[x]_p)[\mathsf{cut}_{U,a}(L\theta)]_p$, where x is a fresh variable. Moreover, $\mathsf{cut}_{U,a}(L\theta) == L\mathsf{cut}_{U,a}(\theta)$ and $\mathsf{cut}_{U,a}(R\theta) == R\mathsf{cut}_{U,a}(\theta)$ since h does not occur in L nor R. We deduce

$$\begin{array}{rcl} \mathsf{cut}_{U,a}(M) & == & \mathsf{cut}_{U,a}(M[x]_p)[\mathsf{cut}_{U,a}(L\theta)]_p \\ & == & \mathsf{cut}_{U,a}(N[x]_p)[L\mathsf{cut}_{U,a}(\theta)]_p \\ & =_{E'} & \mathsf{cut}_{U,a}(N[x]_p)[R\mathsf{cut}_{U,a}(\theta)]_p \\ & == & \mathsf{cut}_{U,a}(N) \end{array}$$

The inductive step of the proof of Lemma 4 is straightforward.

Lemma 4 yields the following conservativity property, whose converse is evident:

Lemma 5 *If $\phi_1 \vdash_{E'} M$ then $\phi \vdash_E M$.*

By Proposition 1, we establish this conservativity property by assuming that there exists a term ζ' over Σ' such that $fn(\zeta') \cap \widetilde{n} = \emptyset$ and $\zeta'\sigma_1 =_{E'} M$ and proving that there then exists a term ζ over Σ such that $fn(\zeta) \cap \widetilde{n} = \emptyset$ and $\zeta\sigma =_E M$. The symbol h does not appear in M since M is over Σ, but it may appear in ζ'. Intuitively, we obtain ζ from ζ' by cutting subterms where h appears, as follows. Suppose that h appears in $\zeta'\sigma_1$, so there exists a subterm $U == \mathsf{h}(V)$ of $\zeta'\sigma_1$. Let a be a fresh name. We apply the cutting function $\mathsf{cut}_{U,a}$ to the equality $\zeta'\sigma_1 =_{E'} M$, and derive $\mathsf{cut}_{U,a}(\zeta'\sigma_1) =_{E'} \mathsf{cut}_{U,a}(M) == M$ by Lemma 4. Moreover, we can write $\mathsf{cut}_{U,a}(\zeta'\sigma_1)$ in the form $\zeta''\sigma_1$ where ζ'' is a term over Σ' such that $fn(\zeta'') \cap \widetilde{n} = \emptyset$. (We construct ζ'' from ζ' in the following way: for each path p such that $\zeta'\sigma_1|_p == \mathsf{h}(M')$ with $M' =_{E'} V$, p must be a path of ζ' since neither M nor the terms M_i contain h, so we define ζ'' by replacing $\zeta'|_p$ with a at each such position p.) Applying this transformation to all occurrences of h, we eventually obtain ζ'' over Σ and also eliminate any occurrences of x_{l+1}. We thus reduce to the case in which h does not appear in $\zeta'\sigma_1$. In this case, we obtain $\zeta'\sigma_1 == \zeta'\sigma$ (because x_{l+1} cannot occur in ζ' in this case) and $\zeta'\sigma_1 =_E M$ (because E' does not equate any more terms over Σ than E), so $\zeta'\sigma =_E M$.

In order to establish Proposition 4, it remains to prove that if $\phi_1 \not\approx_{s\,E'} \phi_2$ then $\phi \vdash_E M$. For this purpose, we assume that $\phi \not\vdash_E M$ and show that $\phi_1 \approx_{s\,E'} \phi_2$, using Lemma 4 as follows. Let V_1 and V_2 be two terms that do not contain the names $\widetilde{n} \cup \{k\}$.

- Assume that $V_1\sigma_2 =_{E'} V_2\sigma_2$. By substituting k with $\mathsf{h}(M)$ in the equality, we get $V_1\sigma_1 =_{E'} V_2\sigma_1$ since k occurs only in σ_2, and any equation that holds for a fresh name such as k holds for any term.

- Conversely, assume that $V_1\sigma_1 =_{E'} V_2\sigma_1$. Let $U == \mathsf{h}(M)$. We apply the cutting function $\mathsf{cut}_{U,k}$ to the equality, and derive $\mathsf{cut}_{U,k}(V_1\sigma_1) =_{E'} \mathsf{cut}_{U,k}(V_2\sigma_1)$

by Lemma 4. Let us show that $\mathsf{cut}_{U,k}(V_1\sigma_1) == V_1\mathsf{cut}_{U,k}(\sigma_1)$. We argue by contradiction, and assume that $\mathsf{cut}_{U,k}(V_1\sigma_1) == V_1\mathsf{cut}_{U,k}(\sigma_1)$ does not hold. This assumption means that there exists a subterm V_1' of V_1 such that V_1' is not a variable and $V_1'\sigma_1 == \mathsf{h}(T')$ with $T' =_{E'} M$. Since V_1' is not a variable, V_1' must be of the form $\mathsf{h}(V_1'')$ with $V_1''\sigma_1 == T' =_{E'} M$. Since V_1 does not contain the names $\widetilde{n}$, neither do V_1' and V_1'', so $V_1''\sigma_1 =_{E'} M$. Therefore, we have $\phi_1 \vdash_{E'} M$ by Proposition 1, and hence $\phi \vdash_E M$ by Lemma 5, contradicting our assumption that $\phi \nvdash_E M$. We obtain $\mathsf{cut}_{U,k}(V_1\sigma_1) == V_1\mathsf{cut}_{U,k}(\sigma_1)$, and similarly we obtain $\mathsf{cut}_{U,k}(V_2\sigma_1) == V_2\mathsf{cut}_{U,k}(\sigma_1)$, so $V_1\mathsf{cut}_{U,k}(\sigma_1) =_{E'} V_2\mathsf{cut}_{U,k}(\sigma_1)$. Finally, since $\mathsf{cut}_{U,k}(\sigma_1) == \sigma_2$, we deduce that $V_1\sigma_2 =_{E'} V_2\sigma_2$.

We conclude that $\phi_1 \approx_{s\,E'} \phi_2$.

3.3 $\approx_s$ does not reduce to $\vdash$ in general

The converse is not true: $\vdash$ may be decidable while $\approx_s$ is not. Indeed, we can encode an undecidable problem into the static equivalence problem in such a way that the deduction problem remains decidable.

Proposition 5 *There exists an equational theory such that $\approx_s$ is undecidable while $\vdash$ is decidable.*

A preliminary presentation of our work [1] includes a first construction of a suitable equational theory, with only a brief proof sketch. Following our work, Borgström has recently provided an alternative construction, based on context-free grammars, with a complete proof [15]. In what follows we describe our original construction, as it may remain instructive, but refer the reader to Borgström's paper for a rigorous argument.

We consider the following construction: Given two deterministic Turing machines $M_1 = (Q, A, q_0, Q_f, \delta_1)$ and $M_2 = (Q, A, q_0, Q_f, \delta_2)$ with the same control states, where $\delta_1, \delta_2 : Q \times A \to Q \times A \times \{L, R\}$, we construct the machine $\mathcal{M}(M_1, M_2) = (Q, A, q_0, Q_f, \delta)$ where $\delta : \{1, 2\} \times Q \times A \to Q \times A \times \{L, R\}$ such that $\delta(1, q, a) = \delta_1(q, a)$ and $\delta(2, q, a) = \delta_2(q, a)$. At each step, the machine $\mathcal{M}(M_1, M_2)$ plays a transition of either M_1 or M_2. Since the machines M_1 and M_2 are deterministic, a run of the machine $\mathcal{M}(M_1, M_2)$ on a word w may be described by a word s of $\{1, 2\}^*$, which gives the list of choices made by $\mathcal{M}(M_1, M_2)$ at each step. $\mathcal{M}(M_1, M_2), w \xrightarrow{s}$ denotes the machine (with its current tape) after the sequence of choices s on the word w. We assume that the local control state is written on the tape.

Proposition 6 *The following problem is undecidable.*
Input: *Two machines $\mathcal{M}(M_1, M_2)$ and $\mathcal{M}(M_1', M_2')$ and a word w of A^*.*
Output: *Does the following property hold for $\mathcal{M}(M_1, M_2)$ and $\mathcal{M}(M_1', M_2')$: for any sequences $s_1, s_2 \in \{1, 2\}^*$, $\mathcal{M}(M_1, M_2), w \xrightarrow{s_1}$ and $\mathcal{M}(M_1, M_2), w \xrightarrow{s_2}$ have the same tape if and only if $\mathcal{M}(M_1', M_2'), w \xrightarrow{s_1}$ and $\mathcal{M}(M_1', M_2'), w \xrightarrow{s_2}$ have the same tape?*

We reduce this undecidable problem to the $\approx_s$ problem under an equational theory E_{tm} such that $\vdash$ remains decidable. The intuitive idea of our encoding is that a

frame ϕ represents a machine of the form $\mathcal{M}(M_1, M_2)$, a term M represents a sequence of choices such that $M\phi$ represents the tape of the machine (and the number of choices) after this sequence of choices. Then, for two "machines" ϕ and ϕ', it is undecidable whether there exists two sequences of choices M_1, M_2 such that $(M_1 =_{E_{\mathsf{tm}}} M_2)\phi$ and $(M_1 \neq_{E_{\mathsf{tm}}} M_2)\phi'$, that is, whether $\phi \not\approx_s \phi'$.

On the other hand, it is possible to decide whether there exists a sequence of choices M such that $M\phi =_{E_{\mathsf{tm}}} N$, that is, whether $\phi \vdash N$) for a given term N. The term N contains the number of choices, so it is sufficient to test any sequence of choices of length equal to this number of choices.

Appendix A contains a proof of Proposition 6, as well as details on how we use the problem in question.

4 Deciding knowledge under convergent subterm theories

In this section, in order to obtain decidability results for both $\vdash$ and $\approx_s$, we restrict attention to *subterm theories*, defined by a finite set of equations of the form $M = N$ where N is a proper subterm of M or a constant symbol. In Section 4.1, we motivate and introduce a convergence condition on subterm theories. Convergent subterm theories are quite common in applications, as we illustrate with examples in Section 4.2. We present our main decidability results for these theories in Section 4.3.

4.1 Convergence

The definition of subterm theories is almost vacuous on its own. Even equality may be undecidable for subterm theories. Any equational theory defined by a finite set of equations $M = M'$ with variables can be encoded as a subterm theory, with the two equations:

$$\mathsf{Whichever}(M, M') = M \qquad \mathsf{Whichever}(M, M') = M'$$

for each original equation $M = M'$. In light of this encoding, we should add the assumption that, by orienting the equations that define a subterm theory from left to right, we obtain a convergent rewriting system:

Definition 1 *A equational theory E, defined by a finite set of equations $\bigcup_{i=1}^{n}\{M_i = N_i\}$ where $fn(M_i) = fn(N_i) = \emptyset$, is a* convergent subterm theory *if the set of rewriting rules $\mathcal{R} \stackrel{def}{=} \bigcup_{i=1}^{n}\{M_i \to N_i\}$ is convergent and if each N_i is a proper subterm of M_i or a constant. We write $U \to V$ if U and V are closed terms and U may be rewritten to V (in one step) using a rule of $\mathcal{R}$.*

As usual, if $\mathcal{R}$ is convergent then for all terms U and V we have $U =_E V$ if and only if $U{\downarrow} = V{\downarrow}$, where $U{\downarrow}$ and $V{\downarrow}$ are the normal forms of U and V.

We write $\to_E$ instead of $\to$ when the equational theory is not clear from the context.

4.2 Examples

Important destructor-constructor rules like those for pairing, encryption, and signature may be expressed in subterm theories (typically convergent ones):

$$\begin{array}{rcl@{\qquad}rcl} \mathsf{fst}(< x, y >) & = & x & \mathsf{dec}(\mathsf{enc}(x, y), y) & = & x \\ \mathsf{snd}(< x, y >) & = & y & \mathsf{check}(x, \mathsf{sign}(x, \mathsf{sk}(y)), \mathsf{pk}(y)) & = & \mathsf{ok} \end{array}$$

Additional examples can be found in previous work (e.g., [3, 12]). Convergent subterm theories also enable us to capture sophisticated but sensible properties, as in:

$$\begin{array}{ll} E_{\mathsf{inv}}: & \{I(I(x)) = x, I(x) \times x = 1, x \times I(x) = 1\} \\ E_{\mathsf{idem}}: & \{h(h(x)) = h(x)\} \\ E_{\mathsf{sym}}: & \{\mathsf{enc}(\mathsf{enc}(x, y), y) = x\} \end{array}$$

The theory E_{inv} models an inverse function. The theory E_{idem} models a hash function that is idempotent on small inputs (since the hash of a hash gives the same hash). The theory E_{sym} represents an encryption function that also decrypts: the encryption of a plaintext, twice with the same key, returns the plaintext.

A rewriting system is convergent if and only if it is terminating and locally confluent (by Newmann's Lemma [22]). For theories with the subterm property, termination holds immediately, so it suffices to examine critical pairs in order to establish convergence. For example, the theory E_{enc} has no critical pairs, so it is convergent; the theory E_{sym} allows rewriting $\mathsf{enc}(\mathsf{enc}(\mathsf{enc}(x, y), y), y)$ in two different ways, but they both yield $\mathsf{enc}(x, y)$, so E_{sym} is convergent as well; on the other hand, the theory $E_{\mathsf{enc}} \cup E_{\mathsf{sym}}$ is not convergent because of the critical pair that consists of $\mathsf{dec}(\mathsf{enc}(\mathsf{enc}(x, y), y), y) \to \mathsf{enc}(x, y)$ and $\mathsf{dec}(\mathsf{enc}(\mathsf{enc}(x, y), y), y) \to \mathsf{dec}(x, y)$.

4.3 Decidability results

For convergent subterm theories, both $\vdash$ and $\approx_s$ become decidable. Let E be a convergent subterm theory given by $\bigcup_{i=1}^{n} \{M_i = N_i\}$, and $c_E = \max_{1 \leq i \leq n}(|M_i|, \mathsf{ar}(\Sigma)+1)$. By convention, if the equational theory E is empty, we set $c_E = 1$.

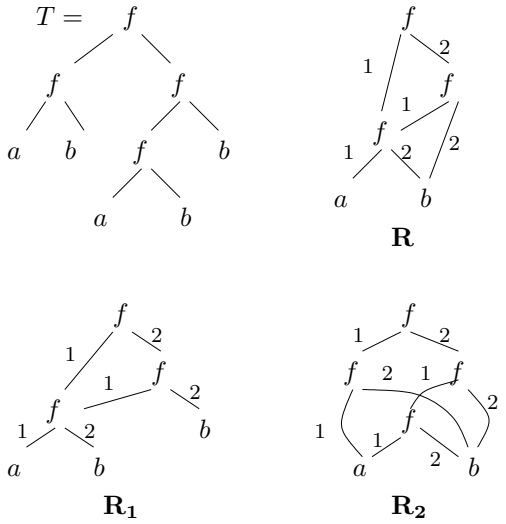
Theorem 1 *For any frames ϕ and ϕ', for any closed term M, we can decide $\phi \vdash M$ and $\phi \approx_s \phi'$ in polynomial time in $|\phi|$, $|\phi'|$, and $|M|$.*

In order to obtain a polynomial bound, we have to consider DAG representations of terms. We define and study them in the next section.

4.3.1 DAG representation for terms

Let us define what is a DAG representation of a term.

Definition 2 (DAG representation) *A DAG representation of a term is a direct acyclic graph (V, l, E, v_0), where V is the set of vertices, $l : V \to \Sigma$ a labelling function, $E \subseteq V \times V \times \{1..\mathsf{ar}(\Sigma)\}$ the set of edges, and $v_0 \in V$ the* root *of the graph. In addition, we assume that for every $v \in V$, for every integer i such that $0 \leq i \leq \mathsf{arity}(l(v))$, there*



The DAG $\mathbf{R}$ is the minimal representation of T but $\mathbf{R_1}$ and $\mathbf{R_2}$ are also DAG representations of T.

Figure 1: Examples of DAG representations.

exists a unique v' (denoted by $E(v,i)$) such that (v,v',i) is in E and that there is no edge of the form (v,v',i) for $i > \mathsf{arity}(l(v))$.

The size of R, written $|R|$, is the number of vertices of R.

The term $t(V,l,E,v_0)$ represented *by a DAG (V,l,E,v_0) is defined recursively by $t(V,l,E,v_0) = l(v_0)(t(V,l,E,e(v_0,1)),\ldots,t(V,l,E,e(v_0,\mathsf{arity}(l(v_0)))))$.*

A DAG representation (V,l,E,v_0) is minimal *if there are no distinct vertices v_1 and v_2 such that $t(V,l,E,v_1) = t(V,l,E,v_2)$.*

Although the memory size needed for representing a DAG R is larger than $|R|$, it is polynomial (actually quadratic) in $|R|$. Thus the measure $|R|$ is sufficient for our purposes. Furthermore, with each term T, we can associate a unique minimal DAG representation of T such that its number of vertices is equal to the number $|T|_{\mathsf{DAG}}$ of subterms of T. See figure 1 for examples.

Proposition 7 *Given a DAG representation R, we can compute the minimal DAG representation of $t(R)$ in polynomial time in $|R|$. Therefore, checking whether $t(R_1) == t(R_2)$ where R_1 and R_2 are two DAG-representations can be done in polynomial time in $|R_1|$ and $|R_2|$.*

Given a DAG representation R, we repeatedly check (at most $|R|$ times) whether there exist two distinct vertices v_1 and v_2 (at most $|R|^2$ possibilities) such that $l(v_1) = l(v_2)$ and for every i such that $0 \leq i \leq \mathsf{arity}(l(v_1))$, $E(v_1,i) = E(v_2,i)$. When such v_1 and v_2 exist, we suppress v_1 in the set of vertices and replace each occurrence of v_1 in E by v_2. We end with the minimal representation of $t(R)$. The total cost of this procedure is at most $\mathcal{O}(|R|^3)$.

Proposition 8 *Given a convergent subterm equational theory and a minimal DAG representation R of a term T, we can compute a (minimal) DAG representation of the normal form $T\downarrow$ of T in polynomial time in $|R|$. Therefore, checking whether $t(R_1) =_E t(R_2)$ where R_1 and R_2 are two minimal DAG-representations can be done in polynomial time in $|R_1|$ and $|R_2|$.*

Let $R = (V, l, E, v_0)$ be a minimal DAG representation of a term T. For every rewriting rule of the form $C[x_1, \ldots, x_n] \to C'[x_1, \ldots, x_n]$ or $C[x_1, \ldots, x_n] \to c$ of the theory, we check (from the root) if the pattern C appears in R (with at most $|C||R|$ tests). If it is the case, that is, there exists some $v \in V$ such that $t(V, l, E, v) == C[x_1, \ldots, x_n]\theta$ for some θ, then we replace the vertex v by one of the vertices that represents $C'[x_1, \ldots, x_n]\theta$ or we add the a vertex that represents c. We minimize the resulting DAG, via Proposition 7, in time $\mathcal{O}(|R|^3)$. At each step (except for a constant number of cases), one of the vertices is suppressed, so this procedure stops after at most $|R|$ steps. We end with a DAG-representation of $T\downarrow$, in time $\mathcal{O}(|R|^4)$.

4.3.2 Proof of Theorem 1

The end of this section is devoted to the proof of the theorem.

Step 1 of the proof: saturating a frame ϕ. We first associate with each frame ϕ the set of subterms of messages in ϕ that may be deduced from ϕ by applying only small contexts. We prove that this set can be computed in polynomial time. In addition, we show that each term in this set has a "recipe" whose DAG-size is polynomial.

Definition 3 *Let $\phi = \nu\widetilde{n}\{M_1/x_1, \ldots, M_l/x_l\}$ be a frame. Let $\mathsf{st}(\phi)$ be the set of subterms of the terms M_i. The saturation $\mathsf{sat}(\phi)$ of ϕ is the minimal set such that:*

1. *for every $1 \leq i \leq l$, $M_i \in \mathsf{sat}(\phi)$,*
2. *if $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and $f(M_1, \ldots, M_k) \in \mathsf{st}(\phi)$, then $f(M_1, \ldots, M_k) \in \mathsf{sat}(\phi)$,*
3. *if $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and $C[M_1, \ldots, M_k] \to M$, where C is a context, $|C| \leq c_E$, $fn(C) \cap \widetilde{n} = \emptyset$, and $M \in \mathsf{st}(\phi)$, then $M \in \mathsf{sat}(\phi)$.*

Proposition 9 *Let ϕ be a frame, $\phi = \nu\widetilde{n}\sigma$.*

1. *The set $\mathsf{sat}(\phi)$ can be computed in time $\mathcal{O}(|\phi|^{\max(\mathsf{ar}(\Sigma), c_E)+2})$.*
2. *For every $M \in \mathsf{sat}(\phi)$, there exists a term ζ_M such that $fn(\zeta_M) \cap \widetilde{n} = \emptyset$, $|\zeta_M|_{\mathsf{DAG}} \leq c_E|\phi|$, and $\zeta_M\sigma =_E M$. The term ζ_M is called a* recipe *of M and is chosen arbitrarily from among the terms that verify these properties.*

The set $\mathsf{sat}(\phi)$ is obtained by saturating the set $\{M_1, \ldots, M_k\}$ by applying the rules 2 and 3 of Definition 3. Since $\mathsf{sat}(\phi) \subseteq \mathsf{st}(\phi)$, this set is saturated in at most $|\phi|$ steps. At each step, we have to compute:

- Every closed term of the form $C[M_1, \ldots, M_k]$ (up to renamings in C), where $|C| \leq c_E$ and the terms M_i are already in the set, and check if it is an instance of some left-hand side of a rule. Thus we need at most $\mathcal{O}(|\phi|^{c_E+1})$ computations.

- Every term $f(M_1, \ldots, M_k)$ that is also in $\mathsf{st}(\phi)$. Thus we have to construct at most $|\Sigma||\phi|^{\mathsf{ar}(\Sigma)}$ terms.

Since each step requires at most $\mathcal{O}(|\phi|^{\max(\mathsf{ar}(\Sigma),c_E+1)})$ computations and since there are at most $|\phi|$ steps, $\mathsf{sat}(\phi)$ may be computed in time $\mathcal{O}(|\phi|^{\max(\mathsf{ar}(\Sigma),c_E)+2})$. For the second part of Proposition 9, we know by Proposition 1 that for each term M of $\mathsf{sat}(\phi)$ there exists ζ_M such that $fn(\zeta_M) \cap \widetilde{n} = \emptyset$ and $\zeta_M \sigma =_E M$. By construction of $\mathsf{sat}(\phi)$, the term ζ_M may be chosen so that:

1. $\zeta_M = x_i$ if $\sigma(x_i) = M$,
2. $\zeta_M = f(\zeta_{M_1}, \ldots, \zeta_{M_k})$ with $M_i \in \mathsf{sat}(\phi)$ if M is obtained by the rule 2,
3. $\zeta_M = C[\zeta_{M_1}, \ldots, \zeta_{M_k}]$ with $M_i \in \mathsf{sat}(\phi)$ if M is obtained by the rule 3.

Assume that we build a graph that contains every DAG that corresponds to the chosen terms ζ_M for $M \in \mathsf{sat}(\phi)$.

1. For every $1 \leq i \leq l$, there is a vertex v_i, labelled by x_i.
2. If $\zeta_M = f(\zeta_{M_1}, \ldots, \zeta_{M_k})$ with $M_i \in \mathsf{sat}(\phi)$, we add a vertex labelled by f and connect this vertex to the vertices that correspond to $\zeta_{M_1}, \ldots, \zeta_{M_k}$.
3. If $\zeta_M = C[\zeta_{M_1}, \ldots, \zeta_{M_k}]$ with $M_i \in \mathsf{sat}(\phi)$, we add a graph that corresponds to $C[\_1, \ldots, \_k]$ (at most $|C| \leq c_E$ vertices) connected to the vertices that correspond to $\zeta_{M_1}, \ldots, \zeta_{M_k}$.

Each step costs one vertex or c_E vertices. Since there are at most $|\mathsf{sat}(\phi)| \leq |\phi|$ steps (one for each term M), the maximal DAG-size of a term ζ_M embedded in this graph is $c_E|\phi|$. Therefore, choosing the recipes from among those terms yields the desired size bound. In what follows, for each ϕ, we assume fixed the set of recipes that corresponds to the terms of $\mathsf{sat}(\phi)$.

Example 1 *We consider again the equational theory E_{enc} defined in Section 2.3. We have $C_{E_{\mathsf{enc}}} = 5$, Let $\phi \stackrel{def}{=} \nu k, s\{\mathsf{enc}(s,k)/x, k/y\}$. By application of rule 1 of Definition 3, we have $\{M_1, M_2\} \subseteq \mathsf{sat}(\phi)$, where $M_1 = \mathsf{enc}(s,k)$ and $M_2 = k$. By application of the rule 3 with the context $C = \mathsf{dec}(\_,\_)$ ($|C| \leq 5$), we have $\mathsf{dec}(M_1, M_2) == \mathsf{dec}(\mathsf{enc}(s,k),k) \rightarrow s$ and $s \in \mathsf{st}(\phi)$. Thus $s \in \mathsf{sat}(\phi)$. Let $M_3 \stackrel{def}{=} s$. Since $\{M_1, M_2, M_3\} \subseteq \mathsf{sat}(\phi) \subseteq \mathsf{st}(\phi) \subseteq \{M_1, M_2, M_3\}$, we deduce that $\mathsf{sat}(\phi) = \{M_1, M_2, M_3\}$.*

The recipes for each term of $\mathsf{sat}(\phi)$ may be chosen in the following way: $\zeta_{M_1} = x$, $\zeta_{M_2} = y$, and $\zeta_{M_3} = \mathsf{dec}(x,y)$.

Step 2 of the proof: Introducing a finite set of equalities to characterize a frame. With each frame ϕ, we associate a set of equalities $\mathsf{Eq}(\phi)$ (finite modulo renaming) such that two frames are equivalent if and only if they satisfy the equalities from each other's set: ϕ' satisfies the equalities $\mathsf{Eq}(\phi)$ and ϕ satisfies the equalities $\mathsf{Eq}(\phi')$.

Definition 4 *Let $\phi = \nu\widetilde{n}\sigma$ be a frame. The set $\mathsf{Eq}(\phi)$ is the set of equalities*

$$C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] = C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}]$$

such that $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\phi$, $|C_1|, |C_2| \leq c_E$, and the terms M_i and M'_i are in $\mathsf{sat}(\phi)$. If ϕ' is a frame such that $(M =_E N)\phi'$ for every $(M = N) \in \mathsf{Eq}(\phi)$, we write $\phi' \models \mathsf{Eq}(\phi)$.

Example 2 *We continue Example 1. Recall that $M_1 = \mathsf{enc}(s, k)$, $M_2 = k$, and $M_3 = s$. We are looking for equalities between small contexts over these terms, modulo the equational theory E_{enc}. By removing trivial or redundant equalities, we obtain that $\mathsf{Eq}(\phi) = \{\mathsf{enc}(\zeta_{M_3}, \zeta_{M_2}) = \zeta_{M_1}\}$, that is, $\mathsf{Eq}(\phi) = \{\mathsf{enc}(\mathsf{dec}(x, y), y) = x\}$. Intuitively, this equality corresponds to the ability of an intruder that can check whether the first message $\mathsf{enc}(s, k)$ is an encrypted message whose encryption key is the second message k, by decrypting and re-encrypting the first message with the second.*

Although $\mathsf{Eq}(\phi)$ may be infinite since the contexts C_1 and C_2 may contain arbitrary names, $\mathsf{Eq}(\phi)$ is finite modulo some renamings that we explain at the end of the section.

Two crucial lemmas show that it is sufficient to consider these equalities:

Lemma 6 *Let $\phi = \nu\widetilde{n}\sigma$ and $\phi' = \nu\widetilde{n'}\sigma'$ be two frames such that $\phi' \models \mathsf{Eq}(\phi)$. For all contexts C_1 and C_2 such that $(fn(C_1) \cup fn(C_2)) \cap \widetilde{n} = \emptyset$, for all terms M_i, $M'_i \in \mathsf{sat}(\phi)$, if $C_1[M_1, \ldots, M_k] == C_2[M'_1, \ldots, M'_l]$, then $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\phi'$.*

Lemma 7 *Let $\phi = \nu\widetilde{n}\sigma$ be a frame. For every context C_1 such that $fn(C_1) \cap \widetilde{n} = \emptyset$, for every $M_i \in \mathsf{sat}(\phi)$, for every term T such that $C_1[M_1, \ldots, M_k] \rightarrow_E^* T$, there exist a context C_2 such that $fn(C_2) \cap \widetilde{n} = \emptyset$, and terms $M'_i \in \mathsf{sat}(\phi)$, such that $T == C_2[M'_1, \ldots, M'_l]$ and for every frame $\phi' \models \mathsf{Eq}(\phi)$, $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\phi'$.*

These two lemmas are proved in a more general setting in Appendix B. How these lemmas are used for proving the decidability of deduction and static equivalence is explained in steps 3 and 4 of the proof, respectively.

Step 3 of the proof: decidability of $\vdash$. Here we show that any message deducible from a frame ϕ is actually a context over terms in $\mathsf{sat}(\phi)$.

Proposition 10 *Let $\phi = \nu\widetilde{n}\sigma$ be a frame, M be a closed term and $M\!\downarrow$ its normal form. Then $\phi \vdash M$ if and only if there exist C and $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ such that $fn(C) \cap \widetilde{n} = \emptyset$ and $M\!\downarrow == C[M_1, \ldots, M_k]$.*

If $M\downarrow == C[M_1, \ldots, M_k]$ with $fn(C) \cap \widetilde{n} = \emptyset$, then $M =_E C[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma$, by construction of the terms ζ_{M_i}. Thus, by Proposition 1, $\phi \vdash M$. Conversely, if $\phi \vdash M$, then by Proposition 1, there exists ζ such that $fn(\zeta) \cap \widetilde{n} = \emptyset$ and $M =_E \zeta\sigma$. Thus $M\downarrow == (\zeta\sigma)\downarrow$. Applying Lemma 7, we obtain that $(\zeta\sigma)\downarrow == C[M_1, \ldots, M_k]$ for some $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and C such that $fn(C) \cap \widetilde{n} = \emptyset$.

We derive that $\phi \vdash M$ can be decided by checking whether $M\downarrow$ is of the form $C[M_1, \ldots, M_k]$ with $M_i \in \mathsf{sat}(\phi)$. Given a term M, $M\downarrow$ can be computed in polynomial time. Once $\mathsf{sat}(\phi)$ is computed (in polynomial time by Proposition 9), checking whether there exist C and $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ such that $fn(C) \cap \widetilde{n} = \emptyset$ and $M\downarrow == C[M_1, \ldots, M_k]$ may be done in time $\mathcal{O}(|M||\phi|^2)$. The procedure is basically as follows:

- Sort $\mathsf{sat}(\phi)$ by the size of the terms (with cost $|\mathsf{sat}(\phi)|^2$).
- For each term T of $\mathsf{sat}(\phi)$ (from terms of maximal size to terms of minimal size), check whether T is equal to a subterm of M. When it is the case, delete this subterm from M. There are $|M|$ subterms in M, the equality test costs $|T| \leq |\phi|$ computations, so this loop can be done in $|M||\phi|^2$.
- Check whether the remaining part of M still contains private names in $\widetilde{n}$. If it is not the case, we have found a context C and $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ such that $fn(C) \cap \widetilde{n} = \emptyset$ and $M\downarrow == C[M_1, \ldots, M_k]$; otherwise such a context does not exist.

This procedure is correct because, when cutting subterms of M equal to terms in $\mathsf{sat}(\phi)$, we start with terms in $\mathsf{sat}(\phi)$ of maximal size. We conclude that $\phi \vdash M$ is decidable in polynomial time.

Step 4 of the proof: decidability of $\approx_s$.

Proposition 11 *For all frames ϕ and ϕ', we have $\phi \approx_s \phi'$ if and only if $\phi \models \mathsf{Eq}(\phi')$ and $\phi' \models \mathsf{Eq}(\phi)$.*

By definition of static equivalence, if $\phi \approx_s \phi'$ then $\phi \models \mathsf{Eq}(\phi')$ and $\phi' \models \mathsf{Eq}(\phi)$. Conversely, assume that $\phi' \models \mathsf{Eq}(\phi)$ and consider M and N such that there exist $\widetilde{n}$ and σ such that $\phi = \nu\widetilde{n}\sigma$, $(fn(M) \cup fn(N)) \cap \widetilde{n} = \emptyset$, and $(M =_E N)\phi$. Then $M\sigma =_E N\sigma$, so $(M\sigma)\downarrow == (N\sigma)\downarrow$. Let $T = (M\sigma)\downarrow$. Applying Lemma 7, we obtain that there exist $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and C_M such that $fn(C_M) \cap \widetilde{n} = \emptyset$ and

$$T == C_M[M_1, \ldots, M_k] \text{ and } M\sigma' =_E C_M[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma'$$

Since $T == (N\sigma)\downarrow$, we obtain similarly that there exist $M'_1, \ldots, M'_l \in \mathsf{sat}(\phi)$ and C_N such that $fn(C_N) \cap \widetilde{n} = \emptyset$ and

$$T == C_N[M'_1, \ldots, M'_l] \text{ and } N\sigma' =_E C_N[\zeta_{M'_1}, \ldots, \zeta_{M'_l}]\sigma'$$

Moreover, since $C_M[M_1, \ldots, M_k] == C_N[M'_1, \ldots, M'_l]$, we derive from Lemma 6 that $C_M[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma' =_E C_N[\zeta_{M'_1}, \ldots, \zeta_{M'_l}]\sigma'$, thus $(M =_E N)\phi'$. Conversely,

when $(M =_E N)\phi'$ and $\phi \models \mathsf{Eq}(\phi')$, we also have that $(M =_E N)\phi$. We conclude that $\phi \approx_s \phi'$.

Therefore, given ϕ and ϕ', in order to decide whether $\phi \approx_s \phi'$ we construct $\mathsf{sat}(\phi)$ and $\mathsf{sat}(\phi')$. This construction can be done in polynomial time by Proposition 9. For each term M of $\mathsf{sat}(\phi)$ or $\mathsf{sat}(\phi')$, the term ζ_M has a polynomial DAG-size.

As noted previously, $\mathsf{Eq}(\phi')$ may be infinite since the equalities may contain arbitrary names. However, each equation of $\mathsf{Eq}(\phi)$ is of the form $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])$ with $|C_1|, |C_2| \leq c_E$, so each equality of $\mathsf{Eq}(\phi)$ contains at most $2c_E$ distinct names besides the names of the recipes. The following lemma, whose proof is easy, says that those $2c_E$ names can be fixed:

Lemma 8 *Let $K = 2c_E$ and $\{n_1, \ldots, n_K\}$ be any set of K distinct names, distinct from the names of the recipes for the terms of $\mathsf{sat}(\phi)$. Let $\mathsf{Eq}'(\phi)$ be the set consisting on the all the equalities*

$$C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] = C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}]$$

such that $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\phi$, $|C_1|, |C_2| \leq c_E$, the terms M_i and M'_i are in $\mathsf{sat}(\phi)$, and $fn(C_1) \cup fn(C_2) \subseteq \{n_1, \ldots, n_K\}$. Then, for any frame ϕ', $\phi \models \mathsf{Eq}(\phi)$ if and only if $\phi \models \mathsf{Eq}'(\phi)$.

Thus, instead of checking whether $\phi \models \mathsf{Eq}(\phi)$, we can check whether $\phi \models \mathsf{Eq}'(\phi)$. More precisely, for all contexts C_1 and C_2 such that $|C_1|, |C_2| \leq c_E$ and $fn(C_1) \cup fn(C_2) \subseteq \{n_1, \ldots, n_K\}$, for all $M_i, M'_i \in \mathsf{sat}(\phi)$, we would check whether $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\phi$ and $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\phi'$.

There are at most $\mathcal{O}((|\phi|^{c_E})^2)$ equalities in $\mathsf{Eq}'(\phi)$. Each term of the form $C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}]\phi$ has a polynomial DAG-size. The equality of two terms represented by DAGs can be checked in polynomial time: we do not need to expand the DAGs in order to test for equality. We conclude that $\phi \approx_s \phi'$ can be decided in polynomial time in $|\phi|$ and $|\phi'|$.

Although this proof is effective, the complexity bounds that we obtain from it appear rather high. For example, for the equational theory E_{enc} of Section 2.3, we can obtain that $\phi \vdash M$ is decidable in time $\mathcal{O}(|M|^3|\phi|^7)$. It should be possible to do much better.

5 Deciding knowledge under more general equational theories

Next, we relax our hypotheses on equational theories. Instead of requiring convergence, we consider equational theories with some associative and commutative symbols that come with a rewriting system $\mathcal{R}$ such that a $\mathcal{R}$ is convergent modulo AC rewriting. Moreover, instead of imposing a syntactic condition (such as a subterm property), we introduce a condition on the set $\mathsf{sat}(\phi)$ associated with each frame ϕ. We present the resulting hypotheses in Section 5.1. We give examples of theories that satisfy the hypotheses in Section 5.2. Finally, we prove general decidability results in Section 5.3.

5.1 The hypotheses

We establish decidability results for equational theories that satisfy three properties. The purpose of this section is to define and start to explain these three properties; Section 5.2 explains them further through examples.

5.1.1 AC-convergence

Our first hypothesis is an adaptation of the standard notion of convergence for theories with AC symbols.

Let E an equational theory, and let $\oplus_1, \ldots, \oplus_k$ be the binary functional symbols such that the equations $x \oplus_i (y \oplus_i z) = (x \oplus_i y) \oplus_i z$ (associativity) and $x \oplus_i y = y \oplus_i x$ (commutativity) are in E.

For two terms U and V, we write $U =_{\mathsf{AC}} V$ if U and V are equal in the theory induced by the equations $x \oplus_i (y \oplus_i z) = (x \oplus_i y) \oplus_i z$ and $x \oplus_i y = y \oplus_i x$ for $1 \leq i \leq k$. When this theory is empty (because we have no AC symbols), $=_{\mathsf{AC}}$ is simply syntactic equality.

When $\mathcal{R}$ is a rewriting system, we write $U \to_{\mathsf{AC}} V$ if there exists W such that $U =_{\mathsf{AC}} W$ and $W \to V$. The relation $\to^*_{\mathsf{AC}}$ denotes the reflexive and transitive closure of $\to_{\mathsf{AC}}$.

Definition 5 (AC-convergent) *An equational theory E is* AC-convergent *if there exists a finite rewriting system $\mathcal{R}$ such that:*

- *$\mathcal{R}$ is* AC-terminating*, that is, for every closed term U, there is no infinite sequence $U \to_{\mathsf{AC}} U_1 \to_{\mathsf{AC}} \cdots U_k \to_{\mathsf{AC}} \cdots$.*

  For every term U, the set of normal forms $U{\downarrow}$ (closed modulo AC) of U is the set of terms V such that $U \to^*_{\mathsf{AC}} V$ and V has no successor for $\to_{\mathsf{AC}}$.

- *$\mathcal{R}$ is* AC-confluent*, that is, for every closed terms U, U_1, and U_2 such that $U \to_{\mathsf{AC}} U_1$ and $U \to_{\mathsf{AC}} U_2$, there exist V_1 and V_2 such that $U_1 \to^*_{\mathsf{AC}} V_1$, $U_2 \to^*_{\mathsf{AC}} V_2$, and $V_1 =_{\mathsf{AC}} V_2$.*

- *For all closed terms U and V, the equality $U =_E V$ holds if and only if there exists a term $T \in (U{\downarrow} \cap V{\downarrow})$.*

By AC-convergence, the set $U{\downarrow}$ is always finite and for all $V, W \in U{\downarrow}$, the equality $V =_{\mathsf{AC}} W$ holds. AC-convergence immediately implies the decidability of equations on closed terms.

In what follows, E is an AC-convergent equational theory and $\mathcal{R}$ is a rewriting system associated with E that satisfies the conditions of Definition 5. If $\mathcal{R}$ consists of a finite set of rules $\bigcup_{i=1}^{k} \{M_i \to N_i\}$, the size c_E of the theory E is defined as $c_E = \max_{1 \leq i \leq k}(|M_i|, |N_i|, \mathsf{ar}(\Sigma) + 1)$. As a special case, $c_E = \mathsf{ar}(\Sigma) + 1$ when $\mathcal{R}$ is empty. As another special case, we obtain the definition of c_E given in Section 4.3 for subterm theories.

Note that E need not have AC symbols. A theory defined by a convergent rewriting system without AC symbol is of course an AC-convergent theory. In that case, we may simply say that the theory is convergent.

Example 3 *As a first example, we consider the theory of an encryption scheme that has an homomorphism property. This property is simply that the encryption of a pair is the pair of the encryptions; the literature (e.g., [31]) suggests other homomorphism properties. This property is modeled by the equation:*

$$\mathsf{enc}(\langle x, y\rangle, z) = \langle \mathsf{enc}(x, z), \mathsf{enc}(y, z)\rangle$$

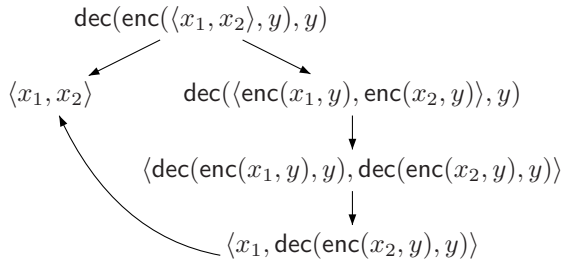
We also assume an analogous equation for decryption:

$$\mathsf{dec}(\langle x, y\rangle, z) = \langle \mathsf{dec}(x, z), \mathsf{dec}(y, z)\rangle$$

As usual, we write $\langle x, y\rangle$ *instead of* $\mathsf{pair}(x, y)$*. The signature* Σ_{homo} *is* $\{\mathsf{pair}, \mathsf{enc}, \mathsf{fst}, \mathsf{snd}, \mathsf{dec}\}$*, and the theory* E_{homo} *is defined by the axioms:*

$$\begin{array}{rcl} \mathsf{enc}(\langle x, y\rangle, z) & = & \langle \mathsf{enc}(x, z), \mathsf{enc}(y, z)\rangle \\ \mathsf{dec}(\langle x, y\rangle, z) & = & \langle \mathsf{dec}(x, z), \mathsf{dec}(y, z)\rangle \\ \mathsf{fst}(\langle x, y\rangle) & = & x \\ \mathsf{snd}(\langle x, y\rangle) & = & y \\ \mathsf{dec}(\mathsf{enc}(x, y), y) & = & x \end{array}$$

We consider the rewriting system $\mathcal{R}_{\mathsf{homo}}$ *obtained from* E_{homo} *by orienting the equations from left to right. With this choice of* $\mathcal{R}_{\mathsf{homo}}$*, the theory* E_{homo} *is convergent: its only critical pair is joinable.*



Example 4 *The theory of XOR is also AC-convergent. The XOR operator is represented by the* $\oplus$ *function symbol, with the following properties:*

$$E_{\mathsf{xor}} = \left\{ \begin{array}{rcl} x \oplus (y \oplus z) & = & (x \oplus y) \oplus z \\ x \oplus y & = & y \oplus x \\ x \oplus x & = & 0 \\ x \oplus 0 & = & x \end{array} \right\}$$

where 0 *is a constant symbol and the signature* Σ_{xor} *is* $\{0, \oplus\}$*. We associate to* E_{xor} *the rewriting system* $\mathcal{R}_{\mathsf{xor}}$*:*

$$\mathcal{R}_{\mathsf{xor}} = \left\{ \begin{array}{rcl} x \oplus x & \to & 0 \\ x \oplus 0 & \to & x \end{array} \right\}$$

Using this choice of $\mathcal{R}_{\mathsf{xor}}$*, it is easy to verify that* E_{xor} *is AC-convergent.*

5.1.2 Local stability

Our second hypothesis roughly says that, for every frame, there is a finite set of terms deducible from the frame that satisfies certain closure conditions. Stating this hypotheses precisely requires a few auxiliary definitions and notations.

Assume that there exists some rule $M_0 \to N_0$ of the rewriting system $\mathcal{R}$ and some substitution θ such that either there exists a term U_1 such that $U =_{\mathsf{AC}} U_1$, $U_1 = M_0\theta$, and $V = N_0\theta$, or there exist terms U_1 and U_2 such that $U =_{\mathsf{AC}} U_1 \oplus U_2$ for some AC symbol $\oplus$, $U_1 = M_0\theta$, and $V =_{\mathsf{AC}} N_0\theta \oplus U_2$. Then we say that the reduction $U \to V$ occurs in head, and we write $U \xrightarrow{h} V$.

We write $\alpha \cdot_\oplus M$ for the term $M \oplus \cdots \oplus M$, α times (for $\alpha \in \mathbb{N}^*$). We simply write αM when the AC symbol is clear from the context. Given a set of terms S and a set of names $\widetilde{n}$, we write $\mathsf{sum}_\oplus(S, \widetilde{n})$ for the set of arbitrary sums of terms of S and other names, closed modulo AC-rewriting:

$$\mathsf{sum}_\oplus(S,\widetilde{n}) \stackrel{\text{def}}{=} \left\{ \begin{array}{l|l} (\alpha_1 \cdot_\oplus T_1) \oplus \cdots \oplus (\alpha_n \cdot_\oplus T_n) & \alpha_i, \beta_i \in \mathbb{N}^*, \\ \qquad \oplus & n_i \notin \widetilde{n}, \\ (\beta_1 \cdot_\oplus n_1) \oplus \cdots \oplus (\beta_k \cdot_\oplus n_k) & T_i \in S \end{array} \right\}$$

Typically, the names in $\widetilde{n}$ will be private, and the others public. Then we define $\mathsf{sum}(S, \widetilde{n})$ as the union of the $\mathsf{sum}_\oplus(S, \widetilde{n})$ for any AC symbol $\oplus$ of the theory.

For convergent subterm theories, the main step of the proof of the decidability of $\vdash$ and $\approx_s$ shows the existence, for each frame ϕ, of a set $\mathsf{sat}(\phi)$ stable by application of "small" contexts. We generalize this condition by requiring that the application of a rewriting rule to a "small" context C applied to arbitrary sums of terms in $\mathsf{sat}(\phi)$ is again a "small" context C' applied to sums of terms in $\mathsf{sat}(\phi)$. The definition of "small" is partly arbitrary; we bound the size of C by c_E and the size of C' by ${c_E}^2$, but other finite size bounds may be suitable.

Definition 6 (locally stable) *An AC-convergent equational theory E is* locally stable *if, for every frame $\phi = \nu\widetilde{n}\{M_1/x_1, \ldots, M_k/x_k\}$, where the terms M_i are closed and in normal form, there exists a finite (computable) set $\mathsf{sat}(\phi)$, closed modulo AC, such that*

1. *for every $1 \le i \le k$, $M_i \in \mathsf{sat}(\phi)$, and for every $n \in fn(\phi)$, $n \in \mathsf{sat}(\phi)$,*
2. *if $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and $f(M_1, \ldots, M_k) \in \mathsf{st}(\mathsf{sat}(\phi))$, then $f(M_1, \ldots, M_k) \in \mathsf{sat}(\phi)$,*
3. *if $C[S_1, \ldots, S_l] \xrightarrow{h} M$, where C is a context such that $|C| \le c_E$ and $fn(C) \cap \widetilde{n} = \emptyset$, and where $S_1, \ldots, S_l \in \mathsf{sum}_\oplus(\mathsf{sat}(\phi), \widetilde{n})$ for some AC symbol $\oplus$ (or $S_i \in \mathsf{sat}(\phi)$ if there is no AC symbol), then there exist a context C', a term M', and $S'_1, \ldots, S'_k \in \mathsf{sum}_\oplus(\mathsf{sat}(\phi), \widetilde{n})$ (or $S'_1, \ldots, S'_k \in \mathsf{sat}(\phi)$ if there is no AC symbol), such that $|C'| \le c_E^2$, $fn(C') \cap \widetilde{n} = \emptyset$, and $M \to^*_{\mathsf{AC}} M' =_{\mathsf{AC}} C'[S'_1, \ldots, S'_k]$,*
4. *if $M \in \mathsf{sat}(\phi)$ then $\phi \vdash M$.*

The set $\mathsf{sat}(\phi)$ need not be unique, nor minimal. Any set that satisfies the four conditions is adequate for our present purposes.

Example 5 *For the equational theory E_{homo} of Example 3, given a frame ϕ in normal form, the set $\mathsf{sat}(\phi)$ is simply obtained by adding subterms of ϕ deducible from ϕ. Suppose for example that the attacker listens to two messages: $\mathsf{enc}(\langle n_1, n_2\rangle, k)$ and $\mathsf{enc}(n_3, \mathsf{enc}(n_1, k))$. Since $\mathsf{enc}(\langle n_1, n_2\rangle, k) =_{E_{\mathsf{homo}}} \langle \mathsf{enc}(n_1, k), \mathsf{enc}(n_2, k)\rangle$, the corresponding frame can be written*

$$\phi_2 = \nu(n_1, n_2, n_3, k)\{\langle \mathsf{enc}(n_1, k), \mathsf{enc}(n_2, k)\rangle / x_1, \mathsf{enc}(n_3, \mathsf{enc}(n_1, k)) / x_2\}$$

Then, the deducible subterms of the frame ϕ_2 are $\mathsf{enc}(n_1, k)$, $\mathsf{enc}(n_2, k)$, and n_3, so $\mathsf{sat}(\phi_2)$ is the set

$$\{\langle \mathsf{enc}(n_1, k), \mathsf{enc}(n_2, k)\rangle, \mathsf{enc}(n_3, \mathsf{enc}(n_1, k)), \mathsf{enc}(n_1, k), \mathsf{enc}(n_2, k), n_3\}$$

In Section 5.2.2 we prove that this construction satisfies the requirements.

In general, establishing that an equational theory is locally stable may be difficult. We give other examples of locally stable theories in Section 5.2.

5.1.3 Local finiteness and local decidability

For our third hypothesis, we consider a certain set of "small" equations that a frame satisfies. One of our results says that this set characterizes the frame. The third hypothesis, which this section presents, pertains to deciding whether another frame satisfies this set. In fact, this section discusses two versions of the third hypothesis, called local finiteness and local decidability. Either is sufficient for our purposes; the former has been more attractive in applications; the latter is more general. As the use of equations may suggest, we rely on the third hypothesis in the study of static equivalence but not deduction.

For each frame $\phi = \nu\widetilde{n}\sigma$, we assume a fixed set of terms $\rho(\phi) = \{\zeta_M \mid M \in \mathsf{sat}(\phi)\}$ such that for each ζ_M, $fn(\zeta_M) \cap \widetilde{n} = \emptyset$ and $\zeta_M\sigma =_E M$. Intuitively, the term ζ_M explains how M may be obtained from the terms of ϕ. Since all the terms of $\mathsf{sat}(\phi)$ are deducible, such a set exists by Proposition 1. For instance, for Example 5, the terms associated with $\mathsf{enc}(n_1, k)$, $\mathsf{enc}(n_2, k)$, and n_3 are respectively $\zeta_{\mathsf{enc}(n_1,k)} = \mathsf{fst}(x_1)$, $\zeta_{\mathsf{enc}(n_2,k)} = \mathsf{snd}(x_1)$, and $\zeta_{n_3} = \mathsf{dec}(x_2, \mathsf{fst}(x_1))$.

Much as in Section 4.3, we associate a set of "small" equations $\mathsf{Eq}(\phi)$ with each frame ϕ, in such a way that two frames are equivalent if and only if they satisfy the equations of each other's set (see Proposition 17).

Definition 7 *Let $\phi = \nu\widetilde{n}\sigma$ be a frame in normal form. The set $\mathsf{Eq}(\phi)$ is the set of equations of the form*

$$C_1[\chi_1, \ldots, \chi_k] = C_2[\chi'_1, \ldots, \chi'_l]$$

where $(C_1[\chi_1, \ldots, \chi_k] =_E C_2[\chi'_1, \ldots, \chi'_l])\phi$, $(fn(C_1) \cup fn(C_2)) \cap \widetilde{n} = \emptyset$, $|C_1| \leq c_E$, $|C_2| \leq c_E^2$, and the terms χ_i and χ'_i are in the set $\mathsf{sum}_{\oplus}(\rho(\phi), \widetilde{n})$ for some AC symbol $\oplus$ (or χ_i and χ'_i are in $\rho(\phi)$ if there is no AC symbol).

When ϕ and ψ are frames and $(M =_E N)\psi$ for every $(M = N) \in \mathsf{Eq}(\phi)$, we say that ψ satisfies the equations of $\mathsf{Eq}(\phi)$, and write $\psi \models \mathsf{Eq}(\phi)$.

Definition 8 (locally decidable) *A locally stable equational theory is* locally decidable *if the question of whether $\psi \models \mathsf{Eq}(\phi)$, for frames ϕ and ψ, is decidable.*

The set $\mathsf{Eq}(\phi)$ may in general be infinite since the terms χ_i may be of arbitrary size. Local finiteness means that the set $\mathsf{Eq}(\phi)$ is always equivalent to a finite set of equations.

Definition 9 (locally finite) *A locally stable equational theory is* locally finite *if, for every frame ϕ, there exists a finite (computable) set of equations $\mathsf{Eq}'(\phi)$ such that, for every frame ψ:*

$$\psi \models \mathsf{Eq}(\phi) \quad \textit{if and only if} \quad \psi \models \mathsf{Eq}'(\phi)$$

This property suffices for local decidability:

Proposition 12 *Every locally finite equational theory is locally decidable.*

Local finiteness is always true when there are no AC symbols since then the set $\mathsf{Eq}(\phi)$ contains only finitely many equations up to renaming:

Proposition 13 *Let E be a locally stable equational theory with no AC symbols. Then, for any frame ϕ, there exists a finite set of equations $\mathsf{Eq}'(\phi)$ such that for every frame ψ, we have $\psi \models \mathsf{Eq}(\phi)$ if and only if $\psi \models \mathsf{Eq}'(\phi)$. In other words, E is locally finite.*

Each equation of $\mathsf{Eq}(\phi)$ is of the form $C_1[\chi_1, \ldots, \chi_k] = C_2[\chi'_1, \ldots, \chi'_l]$ with χ_i, χ'_i in $\rho(\phi)$. Thus it contains a finite number of names (bounded by $c_E + c_E^2$). The set $\mathsf{Eq}'(\phi)$ is obtained from $\mathsf{Eq}(\phi)$ by renaming the names to a fixed set of names.

In Section 5.2 we present some non-trivial examples of locally finite theories with AC symbols. Establishing local finiteness is our preferred way of proving local decidability for such theories. Here we show that at least an (infinite) subset of $\mathsf{Eq}(\phi)$ may always be replaced by a finite number of equations.

Definition 10 *Let $\phi = \nu\widetilde{n}\sigma$ be a frame. Let N be a set of public names (that is, such that $N \cap \widetilde{n} = \emptyset$). We write $\mathsf{Eq}_{AC}(\phi, N)$ for the set of equations of the form $\chi_1 = \chi_2$ such that $\chi_1, \chi_2 \in \mathsf{sum}_{\oplus}(\rho(\phi), \widetilde{n})$, $fn(\chi_1) \cup fn(\chi_2) \subseteq N$, and $(\chi_1 =_E \chi_2)\phi$.*

Note that $\mathsf{Eq}_{AC}(\phi, N)$ is a subset of $\mathsf{Eq}(\phi)$. We show that the set $\mathsf{Eq}_{AC}(\phi, N)$ may always be replaced by a finite number of equations if N is a finite set of public names.

Proposition 14 *Let $\phi = \nu\widetilde{n}\sigma$ be a frame and N a finite set of names such that $N \cap \widetilde{n} = \emptyset$. There exists a finite set $\mathsf{Eq}_{bAC}(\phi, N) \subseteq \mathsf{Eq}_{AC}(\phi, N)$, such that for every frame ψ:*

$$\psi \models \mathsf{Eq}_{AC}(\phi, N) \quad \textit{if and only if} \quad \psi \models \mathsf{Eq}_{bAC}(\phi, N)$$

In addition, the cardinality of $\mathsf{Eq}_{bAC}(\phi)$ is at most the cardinality of $\mathsf{sat}(\phi)$ plus the cardinality of N.

This proposition can be proved using elementary results on $\mathbb{Z}$-modules. (Facts on $\mathbb{Z}$-module may be found in [32], for example.) Assume that $\mathsf{sat}(\phi) = \{M_1, \ldots, M_k\}$, $N = \{n_1, \ldots, n_l\}$, and let $\Gamma \in \mathbb{Z}^{k+l}$. For $1 \leq i \leq k+l$, Γ_i denotes the ith coefficient of Γ, and $\widehat{\Gamma}$ denotes the equation:

$$\bigoplus_{\Gamma_i>0, i\leq k} \Gamma_i \zeta_{M_i} \;\oplus\; \bigoplus_{\Gamma_i>0, i>k} \Gamma_i n_i \;=\; \bigoplus_{\Gamma_i<0, i\leq k} (-\Gamma_i)\zeta_{M_i} \;\oplus\; \bigoplus_{\Gamma_i<0, i>k} (-\Gamma_i) n_i$$

Let $\mathsf{Eq}'_{AC}(\phi, N) = \{\widehat{\Gamma} \mid \Gamma \in \mathbb{Z}^{k+l}, (\widehat{\Gamma})\phi\}$. It is easy to verify that for any frame ψ, $\psi \models \mathsf{Eq}_{AC}(\phi, N)$ if and only if $\psi \models \mathsf{Eq}'_{AC}(\phi, N)$. It is also easy to verify (simplifying the equations) that $\mathsf{Eq}'_{AC}(\phi, N)$ is a $\mathbb{Z}$-submodule of $\mathbb{Z}^{k+l}$ and thus can be generated by a finite number of vectors $V_1, \ldots, V_r$ with $r \leq k+l$. We define $\mathsf{Eq}_{bAC}(\phi, N) = \{\widehat{V_1}, \ldots, \widehat{V_r}\}$. It is then easy to conclude that, for any frame ψ, $\psi \models \mathsf{Eq}_{AC}(\phi, N)$ if and only if $\psi \models \mathsf{Eq}_{bAC}(\phi, N)$.

Example 6 *Consider for example a pure AC theory with only one AC symbol $+$ (and no other function symbol), and the frame*

$$\phi_3 = \nu(n_1, n_2, n_3)\{3n_1 + 2n_2 + 4n_3/x_1, n_2 + 3n_3/x_2, n_1 + 2n_3/x_3, 3n_2 + n_3/x_4\}$$

The set $\mathsf{Eq}(\phi_3)$ consists of the equations of the form $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 + T = \alpha'_1 x_1 + \alpha'_2 x_2 + \alpha'_3 x_3 + \alpha'_4 x_4 + T'$ with $\alpha_i, \alpha'_i \in \mathbb{N}$, and T and T' sums of names distinct from n_1, n_2, and n_3. By convention, if $\alpha_i = 0$ (resp. $\alpha'_i = 0$) then the term $\alpha_i x_i$ (resp. $\alpha'_i x_i$) does not appear in the sum. Since the equation is true for ϕ_3, we must have $T = T'$, thus it is sufficient to consider the equations of the form $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 = \alpha'_1 x_1 + \alpha'_2 x_2 + \alpha'_3 x_3 + \alpha'_4 x_4$ with $\alpha_i, \alpha'_i \in \mathbb{N}$. Adopting the convention that a negative term αx (with $\alpha < 0$) in an equation actually appears on the other side of the equation, it is sufficient to consider the equations of the form $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 = 0$, with $\alpha_i, \alpha'_i \in \mathbb{Z}$. For example, the equation $3x_1 - 2x_2 + x_3 = 0$ stands for the equation $3x_1 + x_3 = 2x_2$. Then, the set of vectors $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ such that the equation $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 = 0$ holds for ϕ_3 is exactly the set of vectors U of $\mathbb{Z}^4$ such that $AU = 0$ with

$$A = \begin{pmatrix} 3 & 0 & 1 & 0 \\ 2 & 1 & 0 & 3 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

By using classical elementary operations on rows and columns, we find that $AU = 0$ if and only if

$$U = \lambda \begin{pmatrix} 1 \\ 1 \\ -3 \\ -1 \end{pmatrix}$$

for $\lambda \in \mathbb{Z}$. We deduce that the set of equations satisfied by ϕ_3 is exactly the set of equations of the form: $\lambda x_1 + \lambda x_2 = 3\lambda x_3 + \lambda x_4$. Thus, in order to decide whether a frame ψ satisfies $\mathsf{Eq}(\phi_3)$, it is sufficient to check whether ψ satisfies the single equation $x_1 + x_2 = 3x_3 + x_4$.

5.2 Examples

In this section, we give examples of locally stable and locally finite equational theories. In Section 5.3, we prove that local stability implies the decidability of deduction, and that local stability and local finiteness imply the decidability of static equivalence.

Several equational theories related to cryptographic operations are locally stable and locally finite. In particular, we prove that convergent subterm theories are locally stable. We show that a theory of homomorphic encryption, a simple theory for addition, and a theory for blind signatures (which are not subterm theories) are also locally stable. These equational theories do not have AC symbols, so local finiteness follows from Proposition 13. As examples of theories with AC symbols, we prove that the pure AC theory and a theory of the XOR operator are locally stable and locally finite. The proofs of these properties require only a few lines, and thus are much simpler than direct proofs of decidability. We have also drafted proofs that the theory of Abelian groups is locally stable and locally finite, but in that case the proofs are quite tedious—probably more than direct proofs of the decidability of deduction and static equivalence.

As the examples may suggest, proving local stability often requires a precise understanding of the cryptographic primitives represented by an equational theory. In particular, removing some equations need not always preserve local stability.

5.2.1 Convergent subterm theories

It is easy to verify that the definition of $\mathsf{sat}(\phi)$ given in Definition 1 fits our requirements for local stability.

Proposition 15 *Every convergent subterm theory is a locally finite theory.*

Consequently, we obtain again that both deducibility and static equivalence are decidable for convergent subterm theories.

5.2.2 Homomorphism

We consider again the equational theory E_{homo} (defined in Example 3), which represents an encryption scheme with a homomorphism property. The size of the theory is 7.

Comon-Lundh and Treinen have investigated a very similar equational theory [20]. They have shown that its deduction relation is decidable in PTIME. Here we show that E_{homo} is locally stable, and it is obviously locally finite (since it has no AC symbol). These properties will imply that both deduction and static equivalence are decidable.

Let $\phi = \nu\widetilde{n}\{M_1/x_1, \ldots, M_k/x_k\}$ be any frame in normal form. We define $\mathsf{sat}(\phi)$ to be the smallest set such that:

1. for every $1 \leq i \leq k$, $M_i \in \mathsf{sat}(\phi)$, and for every $n \in \mathit{fn}(\phi)$, $n \in \mathsf{sat}(\phi)$,
2. if $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and $f(M_1, \ldots, M_k) \in \mathsf{st}(\mathsf{sat}(\phi))$, then $f(M_1, \ldots, M_k) \in \mathsf{sat}(\phi)$,

3. if $M_1, M_2 \in \mathsf{sat}(\phi)$ and $\mathsf{dec}(M_1, M_2) \xrightarrow{h} M$ and the rule $\mathsf{dec}(\mathsf{enc}(x, y), y) \to x$ has been applied, or $\mathsf{fst}(M_1) \xrightarrow{h} M$, or $\mathsf{snd}(M_1) \xrightarrow{h} M$, then $M \in \mathsf{sat}(\phi)$.

The set $\mathsf{sat}(\phi)$ is finite since we add only subterms of terms of ϕ. It trivially satisfies conditions 1, 2, and 4 of Definition 6. Let us show that it satisfies condition 3. Let $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and assume that $C[M_1, \ldots, M_k] \xrightarrow{h} M$ where $|C| \leq 7$. The case where C is a single hole is covered by the fact that the terms are in normal form. The other cases are covered by rule 3 except in the following cases:

- $C = \mathsf{enc}(\_, \_)$, $C = \mathsf{enc}(\_, T)$, or $C = \mathsf{enc}(T, \_)$ where $fn(T) \cap \widetilde{n} = \emptyset$ and $|T| \leq 5$.
  - For $\mathsf{enc}(M_1, M_2) \to M$ with $M_1, M_2 \in \mathsf{sat}(\phi)$: In this case, M_1 must be of the form $M_1 = \langle M_1', M_2' \rangle$ and $M = \langle \mathsf{enc}(M_1', M_2), \mathsf{enc}(M_2', M_2) \rangle$. By rule 3, we know that both M_1' and M_2' are in $\mathsf{sat}(\phi)$ since $\mathsf{fst}(M_1) \to M_1'$ and $\mathsf{snd}(M_1) \to M_2'$. Thus M is a context over terms of $\mathsf{sat}(\phi)$ where the context may be chosen as $C' = \langle \mathsf{enc}(\_, \_), \mathsf{enc}(\_, \_) \rangle$ since $|C'| = 7 \leq 7^2 = 49$.
  - For $\mathsf{enc}(M_1, T) \to M$ with $M_1 \in \mathsf{sat}(\phi)$, $fn(T) \cap \widetilde{n} = \emptyset$, and $|T| \leq 5$: We have similarly that $M = \langle \mathsf{enc}(M_1', T), \mathsf{enc}(M_2', T) \rangle$ with M_1' and M_2' in $\mathsf{sat}(\phi)$. Thus M is a context over terms of $\mathsf{sat}(\phi)$ where the context may be chosen as $C' = \langle \mathsf{enc}(\_, T), \mathsf{enc}(\_, T) \rangle$ since $|C'| \leq 5 + 2|T| \leq 15 \leq 7^2 = 49$.
  - For $\mathsf{enc}(T, M_2) \to M$ with $M_2 \in \mathsf{sat}(\phi)$, $fn(T) \cap \widetilde{n} = \emptyset$, and $|T| \leq 5$: We must have $T = \langle T_1, T_2 \rangle$ with $|T_1| + |T_2| \leq 4$. We obtain $M = \langle \mathsf{enc}(T_1, M_2), \mathsf{enc}(T_2, M_2) \rangle$, so M is a context over terms of $\mathsf{sat}(\phi)$ where the context may be chosen as $C' = \langle \mathsf{enc}(T_1, \_), \mathsf{enc}(T_2, \_) \rangle$ since $|C'| \leq 5 + |T_1| + |T_2| \leq 9 \leq 49$.
- $C = \mathsf{dec}(\_, \_)$, $C = \mathsf{dec}(\_, T)$, or $C = \mathsf{dec}(T, \_)$ where $fn(T) \cap \widetilde{n} = \emptyset$ and $|T| \leq 5$, and the rule $\mathsf{dec}(\langle x, y \rangle, z) \to \langle \mathsf{dec}(x, z), \mathsf{dec}(y, z) \rangle$ has been applied.

  These three cases are very similar to the three cases above.

5.2.3 Addition

We consider a simple theory for addition. Let Σ_{add} be any signature that contains 0, s, pred, and plus, with the equations:

$$E_{\mathsf{add}} = \left\{ \begin{array}{rcl} \mathsf{plus}(x, s(y)) & = & \mathsf{plus}(s(x), y) \\ \mathsf{plus}(x, 0) & = & x \\ \mathsf{pred}(s(x)) & = & x \end{array} \right\}$$

The size $c_{E_{\mathsf{add}}}$ of this theory is at least 4 (and possibly higher if Σ_{add} contains symbols other than 0, s, pred, and plus). We define $\mathcal{R}_{\mathsf{add}}$ by simply orienting the equations from left to right. Using this choice of $\mathcal{R}_{\mathsf{add}}$, it is easy to verify that E_{add} is convergent. (Note that E_{add} has no AC symbol.) For local stability, when $\phi = \nu\widetilde{n}\{M_1/x_1, \ldots, M_k/x_k\}$ is any frame in normal form, we define $\mathsf{sat}(\phi)$ to be the smallest set such that:

1. for every $1 \leq i \leq k$, $M_i \in \mathsf{sat}(\phi)$, and for every $n \in fn(\phi)$, $n \in \mathsf{sat}(\phi)$,
2. if $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and $f(M_1, \ldots, M_k) \in \mathsf{st}(\mathsf{sat}(\phi))$, then $f(M_1, \ldots, M_k) \in \mathsf{sat}(\phi)$,
3. if $\mathsf{pred}(M) \xrightarrow{h} M'$ and $M \in \mathsf{sat}(\phi)$ then $M' \in \mathsf{sat}(\phi)$.

The set $\mathsf{sat}(\phi)$ is finite since we add only subterms of terms of ϕ. The set $\mathsf{sat}(\phi)$ trivially satisfies conditions 1, 2, and 4 of Definition 6. Let us show that it satisfies condition 3. Assume that $C[M_1, \ldots, M_k] \xrightarrow{h} M$ with $M_i \in \mathsf{sat}(\phi)$ and $|C| \leq c_{E_{\mathsf{add}}}$. The only non-trivial case is the one where $\mathsf{plus}(M_1, M_2) \xrightarrow{h} M'$ with $M_1, M_2 \in \mathsf{sat}(\phi)$ and the rule $\mathsf{plus}(x, s(y)) \rightarrow \mathsf{plus}(s(x), y)$ has been applied. We must have that $M_2 = s(M_2')$. Hence $\mathsf{pred}(M_2) \xrightarrow{h} M_2'$, so $M_2' \in \mathsf{sat}(\phi)$. Now, we have $M' = \mathsf{plus}(s(M_1), M_2')$, with $M_1, M_2' \in \mathsf{sat}(\phi)$ and $|\mathsf{plus}(s(\_), \_)| = 4 \leq 4^2$, so condition 3 is satisfied.

Note that, were we to omit the equation $\mathsf{pred}(s(x)) = x$ in our equational theory, the proof of local stability would no longer be valid.

5.2.4 Blind signatures

We consider a theory recently introduced by Kremer and Ryan in order to model blind signatures and related constructs in their analysis of a protocol for electronic voting [26]. This theory treats signatures much like that of Section 4, with four differences: the checking construct is called checksign (rather than check); checking does not require plaintext; there is no separate signature-key computation (no function sk); and, most importantly, this theory also describes signature blinding and unblinding functions. Let Σ_{blind} be any signature that contains open, commit, getpk, host, checksign, sign, unblind, and blind, with the equations:

$$E_{\mathsf{blind}} = \left\{ \begin{array}{rcl} \mathsf{open}(\mathsf{commit}(x,y),y) & = & x \\ \mathsf{getpk}(\mathsf{host}(x)) & = & x \\ \mathsf{checksign}(\mathsf{sign}(x,y),\mathsf{pk}(y)) & = & x \\ \mathsf{unblind}(\mathsf{blind}(x,y),y) & = & x \\ \mathsf{unblind}(\mathsf{sign}(\mathsf{blind}(x,y),z),y) & = & \mathsf{sign}(x,z) \end{array} \right\}$$

The size $c_{E_{\mathsf{blind}}}$ of the theory is at least 7 (and possibly higher if Σ_{blind} contains additional symbols). We define $\mathcal{R}_{\mathsf{blind}}$ by simply orienting the equations from left to right. The theory E_{blind} is clearly convergent. To prove that E_{blind} is locally stable, we extend the definition of subterms by requiring that $\mathsf{sign}(M_1, M_3)$ is a subterm of $\mathsf{sign}(\mathsf{blind}(M_1, M_2), M_3)$. More formally, we define:

$$\begin{array}{l} \mathsf{st}_{\mathsf{ext}}(u) = u \\ \mathsf{st}_{\mathsf{ext}}(\mathsf{sign}(\mathsf{blind}(M_1, M_2), M_3)) = \\ \quad \{\mathsf{sign}(M_1, M_3)\} \cup \{\mathsf{sign}(\mathsf{blind}(M_1, M_2), M_3)\} \\ \quad \cup\, \mathsf{st}_{\mathsf{ext}}((\mathsf{blind}(M_1, M_2)) \cup \mathsf{st}_{\mathsf{ext}}(M_3) \\ \mathsf{st}_{\mathsf{ext}}(f(M_1, \ldots, M_k)) = \\ \quad \{f(M_1, \ldots, M_k)\} \cup \bigcup_{i=1}^{k} \mathsf{st}_{\mathsf{ext}}(M_i) \\ \qquad \text{otherwise (that is, for other terms)} \end{array}$$

When $\phi = \nu\widetilde{n}\{M_1/x_1, \ldots, M_k/x_k\}$ is any frame in normal form, we define $\mathsf{sat}(\phi)$ to be the smallest set such that:

1. for every $1 \leq i \leq k$, $M_i \in \mathsf{sat}(\phi)$, and for every $n \in fn(\phi)$, $n \in \mathsf{sat}(\phi)$,
2. if $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and $f(M_1, \ldots, M_k) \in \mathsf{st}(\mathsf{sat}(\phi))$, then $f(M_1, \ldots, M_k) \in \mathsf{sat}(\phi)$,
3. if $C[M_1, \ldots, M_k] \xrightarrow{h} M$, $M_i \in \mathsf{sat}(\phi)$ and $M \in \mathsf{st}_{\mathsf{ext}}(\mathsf{sat}(\phi))$ then $M \in \mathsf{sat}(\phi)$.

The set $\mathsf{sat}(\phi)$ is finite since we add only extended subterms of terms of ϕ. The set $\mathsf{sat}(\phi)$ trivially satisfies conditions 1, 2, and 4 of Definition 6. Let us show that it satisfies condition 3. Assume that $C[M_1, \ldots, M_k] \xrightarrow{h} M$ with $M_i \in \mathsf{sat}(\phi)$ and $|C| \leq c_{E_{\mathsf{blind}}}$. If one of the four first rules of $\mathcal{R}_{\mathsf{blind}}$ has been applied, then M is a subterm of $C[M_1, \ldots, M_k]$. Thus either $M = C'[M_1, \ldots, M_k]$ for some context C' and condition 3 is satisfied or M is a subterm of one of the terms M_i, thus $M \in \mathsf{sat}(\phi)$ and condition 3 is satisfied. If the fifth rule of $\mathcal{R}_{\mathsf{blind}}$ has been applied, then three (non-trivial) cases may arise.

- If $M_2 \xrightarrow{h} M$ then M is an extended subterm of M_2, so $M \in \mathsf{sat}(\phi)$ and condition 3 is satisfied.
- Similarly, if $\mathsf{unblind}(M_1, M_2) \xrightarrow{h} M$ then M is an extended subterm of M_1, so $M \in \mathsf{sat}(\phi)$ and condition 3 is satisfied.
- Finally, suppose that $\mathsf{unblind}(\mathsf{sign}(M_1, M_2), M_3) \xrightarrow{h} M$. It must be the case that $M_1 = \mathsf{blind}(M_1', M_3)$. Since $\mathsf{unblind}(M_1, M_3) \xrightarrow{h} M_1'$ and M_1' is a subterm of M_1, we have $M_1' \in \mathsf{sat}(\phi)$. Now, since $M = \mathsf{sign}(M_1', M_2)$ and $|\mathsf{sign}(\_, \_)| = 3 \leq 7^2$, condition 3 is satisfied.

5.2.5 Pure AC theory

We consider the case where the signature contains only constant symbols and AC symbols $\oplus_1, \ldots, \oplus_k$ and the equational theory E_{ac} contains only the AC equations for each symbol:

$$E_{\mathsf{ac}} = \bigcup_{i=1}^{k} \left\{ \begin{array}{rcl} (x \oplus_i y) \oplus_i z & = & x \oplus_i (y \oplus_i z) \\ x \oplus_i y & = & y \oplus_i x \end{array} \right\}$$

With the empty rewriting system $\mathcal{R}_{\mathsf{ac}} = \emptyset$, E_{ac} is an AC-convergent theory. When $\phi = \nu\widetilde{n}\{M_1/x_1, \ldots, M_k/x_k\}$ is any frame, we define $\mathsf{sat}(\phi)$ to be the smallest set such that:

1. for every $1 \leq i \leq k$, $M_i \in \mathsf{sat}(\phi)$, and for every $n \in fn(\phi)$, $n \in \mathsf{sat}(\phi)$,
2. if $M_1, M_2 \in \mathsf{sat}(\phi)$ and $M_1 \oplus_i M_2 \in \mathsf{st}(\mathsf{sat}(\phi))$, then $M_1 \oplus_i M_2 \in \mathsf{sat}(\phi)$,
3. if $M_1 =_{\mathsf{AC}} M_2$ and $M_1 \in \mathsf{sat}(\phi)$ then $M_2 \in \mathsf{sat}(\phi)$.

The set $\mathsf{sat}(\phi)$ is finite since we add only terms smaller or equal than the maximal size of the terms of ϕ. The set $\mathsf{sat}(\phi)$ trivially satisfies conditions 1, 2, and 4 of Definition 6. It also satisfies condition 3 since the rewriting system $\mathcal{R}_{\mathsf{ac}}$ is empty. Thus E_{ac} is locally stable.

Now, for any frame $\phi = \nu\widetilde{n}\sigma$, the set of equations $\mathsf{Eq}(\phi)$ simply consists of $\mathsf{Eq}_{AC}(\phi, \mathcal{N} - \widetilde{n})$. Since names that do not appear in ϕ need not be considered, $\mathsf{Eq}_{AC}(\phi, \mathcal{N} - \widetilde{n})$ is equivalent to $\mathsf{Eq}_{AC}(\phi, N)$ where N is the set of free names of ϕ, in the sense that for any frame ψ, $\psi \models \mathsf{Eq}_{AC}(\phi, \mathcal{N} - \widetilde{n})$ if and only if $\psi \models \mathsf{Eq}_{AC}(\phi, N)$. By Proposition 14, we conclude that the equational theory E_{ac} is locally finite.

5.2.6 XOR

We consider the theory E_{xor} of the XOR operator (defined in Example 3).

We have seen that E_{xor} is AC-convergent. We wish to verify that E_{xor} is locally stable. When $\phi = \nu\widetilde{n}\{M_1/x_1, \ldots, M_k/x_k\}$ is any frame in normal form, we define $\mathsf{sat}(\phi)$ to be the smallest set, closed under AC, such that:

1. for every $1 \leq i \leq k$, $M_i \in \mathsf{sat}(\phi)$, and for every $n \in \mathit{fn}(\phi)$, $n \in \mathsf{sat}(\phi)$, and $0 \in \mathsf{sat}(\phi)$,
2. if $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and $f(M_1, \ldots, M_k) \in \mathsf{st}(\mathsf{sat}(\phi))$, then $f(M_1, \ldots, M_k) \in \mathsf{sat}(\phi)$,
3. if $M_1, M_2 \in \mathsf{sat}(\phi)$, then $(M_1 \oplus M_2)\!\downarrow \subseteq \mathsf{sat}(\phi)$,
4. if a is a name not in $\widetilde{n}$ and if $M \oplus a \rightarrow_{\mathsf{AC}} M'$ with $M' \in \mathsf{st}(\mathsf{sat}(\phi))$, then $M' \in \mathsf{sat}(\phi)$.

Let us first show that $\mathsf{sat}(\phi)$ is finite. Let the set $\mathsf{sst}(\phi)$ of *simple subterms* of ϕ be the set of subterms of ϕ whose head symbol is not $\oplus$. Let $S = \{T_1 \oplus \cdots \oplus T_n \mid T_i \in \mathsf{sst}(\phi), T_i \neq 0, T_i = T_j \Rightarrow i = j\}$ be the set of sums of distinct terms of $\mathsf{sst}(\phi)$. The set S is finite and $\mathsf{sat}(\phi) \subseteq S$. Indeed, it is easy to show that S satisfies the four conditions above, using that $\mathsf{st}(S) = S$.

The set $\mathsf{sat}(\phi)$ trivially satisfies conditions 1, 2, and 4 of Definition 6. Let us show that it satisfies condition 3. Let $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and C be a context such that $\mathit{fn}(C) \cap \widetilde{n} = \emptyset$ and assume that $C[M_1, \ldots, M_k] \xrightarrow{h} M$. We have that $C[M_1, \ldots, M_k] =_{\mathsf{AC}} \bigoplus_{i=1}^{k} M_i \oplus \bigoplus_{i=1}^{n} a_i$, where each a_i is a name not in $\widetilde{n}$ or the constant 0. Let us show that one of the normal forms of $C[M_1, \ldots, M_k]$ is a context of terms in $\mathsf{sat}(\phi)$. Applying recursively rule 3, we obtain that $(\bigoplus_{i=1}^{k} M_i)\!\downarrow \subseteq \mathsf{sat}(\phi)$. Now, applying recursively rule 4, we obtain that $C[M_1, \ldots, M_k]\!\downarrow =_{\mathsf{AC}} M' \oplus \bigoplus_{j=1}^{r} a_{i_j}$, with $M' \in \mathsf{sat}(\phi)$. By AC-convergence, we know that $M \rightarrow_{\mathsf{AC}}^{*} =_{\mathsf{AC}} M' \oplus \bigoplus_{j=1}^{r} a_{i_j}$ with $M' \oplus \bigoplus_{j=1}^{r} a_{i_j} \in \mathsf{sum}_{\oplus}(\mathsf{sat}(\phi), \widetilde{n})$, since no a_{i_j} is 0 (for otherwise the term would not be in normal form), so the context C' that simply consists of a hole satisfies the required conditions.

Like in the pure AC case, for any frame ϕ, the set of equation $\mathsf{Eq}(\phi)$ simply consists of $\mathsf{Eq}_{AC}(\phi, \mathcal{N} - \widetilde{n})$ since the only constant is 0 and 0 is itself in $\mathsf{sat}(\phi)$. Since names that do not appear in ϕ do not need to be considered, $\mathsf{Eq}_{AC}(\phi, \mathcal{N} - \widetilde{n})$ is equivalent to

$\mathsf{Eq}_{AC}(\phi, N)$ where N is the set of free names of ϕ, in the sense that for any frame ψ, $\psi \models \mathsf{Eq}_{AC}(\phi, \mathcal{N} - \widetilde{n})$ if and only if $\psi \models \mathsf{Eq}_{AC}(\phi, N)$. Thus, by Proposition 14, the equational theory E_{xor} is locally finite.

Note that, in this example, we can also conclude without using Proposition 14. Indeed, we can consider the set $\mathsf{Eq}'(\phi)$ that consists of the equations

$$\bigoplus_{j=1}^{k_1} \zeta_{M_{i_j}} \oplus \bigoplus_{j=1}^{k_2} n_{i_j} = \bigoplus_{j=k_1+1}^{l_1} \zeta_{M_{i_j}} \oplus \bigoplus_{j=k_2+1}^{l_2} n_{i_j}$$

such that

$$\left(\bigoplus_{j=1}^{k_1} \zeta_{M_{i_j}} \oplus \bigoplus_{j=1}^{k_2} n_{i_j} =_E \bigoplus_{j=k_1+1}^{l_1} \zeta_{M_{i_j}} \oplus \bigoplus_{j=k_2+1}^{l_2} n_{i_j} \right) \phi$$

$n_{i_j} \in fn(\phi)$, and $l \neq j \implies M_{i_l} \neq M_{i_j}, n_{i_l} \neq n_{i_j}$. Clearly, $\mathsf{Eq}'(\phi)$ is finite and it is easy to verify that, for any frame ψ, $\psi \models \mathsf{Eq}_{AC}(\phi, \widetilde{n})$ if and only if $\psi \models \mathsf{Eq}'(\phi)$.

5.3 Decidability results

In this section, we state and prove our decidability results for deduction and static equivalence.

5.3.1 Decidability of deduction

Theorem 2 *For locally stable equational theories, deduction is decidable. More precisely, given a frame ϕ and a term M, once $M{\downarrow}$ and $\mathsf{sat}(\phi)$ are computed, $\phi \vdash M$ can be decided in polynomial time in $|M{\downarrow}|$ and $|\mathsf{sat}(\phi)|$.*

The proof is based on the following lemma.

Lemma 9 *Let E be a locally stable theory. Let $\phi = \nu\widetilde{n}\sigma$ be a frame. For every context C_1 such that $fn(C_1) \cap \widetilde{n} = \emptyset$, for every $M_i \in \mathsf{sat}(\phi)$, for every term T such that $C_1[M_1, \ldots, M_k] \rightarrow_{\mathsf{AC}} T$, there exist a context C_2 such that $fn(C_2) \cap \widetilde{n} = \emptyset$, and terms $M'_i \in \mathsf{sat}(\phi)$, such that $T \rightarrow^*_{\mathsf{AC}} C_2[M'_1, \ldots, M'_l]$.*

This lemma is a weak version of Lemma 11 presented in Section 5.3.2. Applying repeatedly this lemma leads to the following corollary.

Corollary 1 *Let E be a locally stable theory. Let $\phi = \nu\widetilde{n}\sigma$ be a frame. For every context C_1 such that $fn(C_1) \cap \widetilde{n} = \emptyset$, for every $M_i \in \mathsf{sat}(\phi)$, for every term T in normal form such that $C_1[M_1, \ldots, M_k] \rightarrow^*_{\mathsf{AC}} T$, there exist a context C_2 such that $fn(C_2) \cap \widetilde{n} = \emptyset$, and terms $M'_i \in \mathsf{sat}(\phi)$, such that $T =_{\mathsf{AC}} C_2[M'_1, \ldots, M'_l]$.*

Assuming Lemma 9, let $\phi = \nu\widetilde{n}\sigma$ be a frame, C_1 be a context such that $fn(C_1) \cap \widetilde{n} = \emptyset$, $M_i \in \mathsf{sat}(\phi)$, and T a term in normal form such that $C_1[M_1, \ldots, M_k] \rightarrow^*_{\mathsf{AC}} T$. Either $C_1[M_1, \ldots, M_k] =_{\mathsf{AC}} T$ and we are done or we have $C_1[M_1, \ldots, M_k] \rightarrow_{\mathsf{AC}} T' \rightarrow^*_{\mathsf{AC}} T$. By Lemma 9, there exist a context C_2 such that $fn(C_2) \cap \widetilde{n} = \emptyset$, and terms $M'_i \in \mathsf{sat}(\phi)$, such that $T' \rightarrow^*_{\mathsf{AC}} C_2[M'_1, \ldots, M'_l]$. By AC-confluence of the

equational theory and since T is in normal form, $C_2[M'_1, \ldots, M'_l] \rightarrow^*_{\mathsf{AC}} T$. Since the equational theory is AC-terminating, we repeat this transformation until we obtain that $T =_{\mathsf{AC}} C_3[M''_1, \ldots, M''_l]$ for some terms $M''_i \in \mathsf{sat}(\phi)$ and some context C_3.

We show that for any term deducible from a frame ϕ, one of its normal forms is a context over terms in $\mathsf{sat}(\phi)$.

Proposition 16 *Let $\phi = \nu\widetilde{n}\sigma$ be a frame, M be a closed term, and $M{\downarrow}$ its set of normal forms. Then $\phi \vdash M$ if and only if there exist a term $T \in M{\downarrow}$, a context C, and terms $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ such that $fn(C) \cap \widetilde{n} = \emptyset$ and $T == C[M_1, \ldots, M_k]$.*

If there exists $T \in M{\downarrow}$ such that $T == C[M_1, \ldots, M_k]$ with $fn(C) \cap \widetilde{n} = \emptyset$, then $T =_E C[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma$, by construction of ζ_{M_1}, ..., ζ_{M_k}. Therefore, by Proposition 1, $\phi \vdash T$, so $\phi \vdash M$.

Conversely, if $\phi \vdash M$, then by Proposition 1, there exists ζ such that $fn(\zeta) \cap \widetilde{n} = \emptyset$ and $M =_E \zeta\sigma$. Thus there exists $T' \in (M{\downarrow} \cap (\zeta\sigma){\downarrow})$. Since $\zeta\sigma \rightarrow^*_{\mathsf{AC}} T'$, applying Corollary 1, we obtain that $T' =_{\mathsf{AC}} C[M_1, \ldots, M_k]$ for some $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and C such that $fn(C) \cap \widetilde{n} = \emptyset$. Thus we end the proof by choosing $T == C[M_1, \ldots, M_k]$.

We derive that $\phi \vdash M$ can be decided by checking whether one of the terms in $M{\downarrow}$ is of the form $C[M_1, \ldots, M_k]$ with $M_i \in \mathsf{sat}(\phi)$. Regarding the complexity, once $M{\downarrow}$ and $\mathsf{sat}(\phi)$ are computed, $\phi \vdash M$ can be decided in polynomial time in $|M{\downarrow}|$ and $|\mathsf{sat}(\phi)|$ using the same procedure as for Theorem 1.

5.3.2 Decidability of static equivalence

Theorem 3 *For locally decidable equational theories, static equivalence is decidable. A fortiori, for locally finite equational theories, static equivalence is decidable.*

The complexity of the resulting decision procedure closely depends on the complexity of the procedure that ensures local decidability. For locally decidable equational theories, this complexity is simply the complexity of checking whether $\phi \models \mathsf{Eq}(\psi)$ given the frames ϕ and ψ. For locally finite equational theories, it depends polynomially on the time needed to compute $\mathsf{Eq}'(\phi)$ and the time needed to check whether ψ satisfies each equation of $\mathsf{Eq}'(\phi)$.

Our result relies on three hypotheses, namely AC-convergence, locally stability, and local decidability. We leave as an open problem whether the third hypothesis is essential. As far as we know, it might be that AC-convergence and local stability imply local decidability. However, our experience with proofs of local decidability suggests that this implication does not hold, at least not trivially.

The proof is based on two main lemmas that we prove in Appendix B.

Lemma 10 *Let E be a locally stable theory. Let $\phi = \nu\widetilde{n}\sigma$ and $\psi = \nu\widetilde{n'}\sigma'$ be two frames such that $\psi \models \mathsf{Eq}(\phi)$. For all contexts C_1 and C_2 such that $(fn(C_1) \cup fn(C_2)) \cap \widetilde{n} = \emptyset$, for all terms $M_i, M'_i \in \mathsf{sat}(\phi)$, if $C_1[M_1, \ldots, M_k] =_{\mathsf{AC}} C_2[M'_1, \ldots, M'_l]$, then $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\psi$.*

Lemma 11 *Let E be a locally stable theory. Let $\phi = \nu\widetilde{n}\sigma$ be a frame. For every context C_1 such that $fn(C_1) \cap \widetilde{n} = \emptyset$, for every $M_i \in \mathsf{sat}(\phi)$, for every term T such that*

$C_1[M_1, \ldots, M_k] \rightarrow_{\mathsf{AC}} T$, *there exist a context* C_2 *such that* $fn(C_2) \cap \widetilde{n} = \emptyset$, *and terms* $M'_i \in \mathsf{sat}(\phi)$, *such that* $T \rightarrow^*_{\mathsf{AC}} C_2[M'_1, \ldots, M'_l]$ *and for every frame* $\psi \models \mathsf{Eq}(\phi)$, $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\psi$.

As for Corollary 1, applying repeatedly Lemma 11 leads to the following corollary.

Corollary 2 *Let* E *be a locally stable theory. Let* $\phi = \nu\widetilde{n}\sigma$ *be a frame. For every context* C_1 *such that* $fn(C_1) \cap \widetilde{n} = \emptyset$, *for every* $M_i \in \mathsf{sat}(\phi)$, *for every term* T *in normal form such that* $C_1[M_1, \ldots, M_k] \rightarrow^*_{\mathsf{AC}} T$, *there exist a context* C_2 *such that* $fn(C_2) \cap \widetilde{n} = \emptyset$, *and terms* $M'_i \in \mathsf{sat}(\phi)$, *such that* $T =_{\mathsf{AC}} C_2[M'_1, \ldots, M'_l]$ *and for every frame* $\psi \models \mathsf{Eq}(\phi)$, $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\psi$.

In order to check whether two frames satisfy the same equations, we show (using these two lemmas) that it is sufficient to check whether they satisfy the same "small" equations.

Proposition 17 *Let* E *be a locally stable theory. For all frames* ϕ *and* ψ, *we have* $\phi \approx_s \psi$ *if and only if* $\phi \models \mathsf{Eq}(\psi)$ *and* $\psi \models \mathsf{Eq}(\phi)$.

By definition of static equivalence, if $\phi \approx_s \psi$ then $\phi \models \mathsf{Eq}(\psi)$ and $\psi \models \mathsf{Eq}(\phi)$.

Conversely, assume now that $\psi \models \mathsf{Eq}(\phi)$ and consider M and N such that there exist $\widetilde{n}$ and σ such that $\phi = \nu\widetilde{n}\sigma$, $(fn(M) \cup fn(N)) \cap \widetilde{n} = \emptyset$, and $(M =_E N)\phi$. Then $M\sigma =_E N\sigma$, so $((M\sigma)\downarrow \cap (N\sigma)\downarrow) \neq \emptyset$. Let $T \in ((M\sigma)\downarrow \cap (N\sigma)\downarrow)$. Since $M\sigma \rightarrow^*_{\mathsf{AC}} T$, by applying Corollary 2 we obtain that there exist C_M and $M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ such that $fn(C_M) \cap \widetilde{n} = \emptyset$, $T =_{\mathsf{AC}} C_M[M_1, \ldots, M_k]$, and $(M =_E C_M[\zeta_{M_1}, \ldots, \zeta_{M_k}])\psi$. Since $N\sigma \rightarrow^*_{\mathsf{AC}} T$, we obtain similarly that there exist C_N and $M'_1, \ldots, M'_l \in \mathsf{sat}(\phi)$ such that $fn(C_N) \cap \widetilde{n} = \emptyset$, $T =_{\mathsf{AC}} C_N[M'_1, \ldots, M'_l]$, and $(N =_E C_N[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\psi$. Moreover, since $C_M[M_1, \ldots, M_k] =_{\mathsf{AC}} C_N[M'_1, \ldots, M'_l]$, we derive from Lemma 10 that $(C_M[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_N[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\psi$, so $(M =_E N)\psi$. Symmetrically, if $(M =_E N)\psi$ and $\phi \models \mathsf{Eq}(\psi)$, then $(M =_E N)\phi$. We conclude that $\phi \approx_s \psi$.

Therefore, given ϕ and ψ, we may consider $\mathsf{Eq}(\phi)$ and $\mathsf{Eq}(\psi)$ in order to decide whether $\phi \approx_s \psi$. By local decidability of the theory, we can decide whether $\phi \models \mathsf{Eq}(\psi)$ and $\psi \models \mathsf{Eq}(\phi)$.

6 Conclusion

This paper investigates decidability questions for message deducibility and static equivalence, two formal representations for knowledge in the analysis of security protocols. This investigation yields a few somewhat negative results, for example that static equivalence cannot always be reduced to message deducibility. On the other hand, the main results are strong, positive ones: message deducibility and static equivalence are decidable under a wide class of equational theories. This class includes, in particular, standard theories for basic cryptographic primitives. It also includes some less standard, more advanced examples: theories of XOR, homomorphic encryption, blind signatures, addition, and pure AC theories. We succeed in giving a unified treatment for this disparate collection of theories, with a body of techniques that apply to all of them plus

special techniques for verifying that particular theories belong in the class. In addition, for a simple, syntactically defined subclass of theories, we prove that deducibility and static equivalence are actually decidable in polynomial time.

The performances of the corresponding decision procedures obviously depend on the choice of equational theory, and we do not expect them to be very good in many cases. Nevertheless, for many theories of interest, deciding deducibility and static equivalence may well be practical. Baudet has recently implemented a variant of our procedures [6]. The tool ProVerif supports another approach for establishing static equivalences [13].

As indicated in the introduction, deduction and static equivalence are static notions, but they play an important role in analyses with respect to active attacks. Nevertheless, it remains challenging to obtain decidability results with respect to active attacks. This problem is addressed in recent and ongoing work. That work is still largely under way, so detailed descriptions may be premature, but we briefly mention some interesting developments. Going beyond the work of Delaune and Jacquemard [21] (described in the introduction), Baudet has proved that both deduction and static equivalence are decidable under convergent subterm theories [7]. Comon-Lundh is studying the decidability of deduction under general equational theories, including associativity and commutativity properties [18]. Overall, this field appears as a lively one, with increasingly sophisticated techniques and powerful theorems. We may therefore look forward to much progress in algorithmic reasoning about the knowledge of active attackers in security protocols.

Acknowledgments

We are grateful to Michael Rusinowitch and Mathieu Baudet for very helpful discussions.

References

[1] Martín Abadi and Véronique Cortier. Deciding knowledge in security protocols under equational theories. In *Proc. 31st Int. Coll. Automata, Languages, and Programming (ICALP'2004)*, volume 3142 of *Lecture Notes in Computer Science*, pages 46–58. Springer, July 2004.

[2] Martín Abadi and Véronique Cortier. Deciding knowledge in security protocols under (many more) equational theories. In *Proc. 18th IEEE Computer Security Foundations Workshop (CSFW'05)*, pages 62–76. IEEE Comp. Soc. Press, June 2005.

[3] Martín Abadi and Cédric Fournet. Mobile values, new names, and secure communication. In *Proceedings of the 28th ACM Symposium on Principles of Programming Languages (POPL'01)*, pages 104–115, January 2001.

[4] Martín Abadi and Andrew D. Gordon. A bisimulation method for cryptographic protocols. *Nordic Journal of Computing*, 5(4):267–303, Winter 1998.

[5] Martín Abadi and Andrew D. Gordon. A calculus for cryptographic protocols: The spi calculus. *Information and Computation*, 148(1):1–70, January 1999.

[6] Mathieu Baudet. Private communication, 2005.

[7] Mathieu Baudet. Deciding security of protocols against off-line guessing attacks. In *Proceedings of the 12th ACM Conference on Computer and Communications Security (CCS'05)*. ACM Press, November 2005. To appear.

[8] Mathieu Baudet, Véronique Cortier, and Steve Kremer. Computationally sound implementations of equational theories against passive adversaries. Manuscript, extended version of [9], 2005.

[9] Mathieu Baudet, Véronique Cortier, and Steve Kremer. Computationally sound implementations of equational theories against passive adversaries. In *Proc. 32nd Int. Coll. Automata, Languages and Programming (ICALP'2005)*, volume 3580 of *Lecture Notes in Computer Science*, pages 652–663. Springer, July 2005.

[10] Bruno Blanchet. An efficient cryptographic protocol verifier based on Prolog rules. In *14th IEEE Computer Security Foundations Workshop (CSFW-14)*, pages 82–96, June 2001.

[11] Bruno Blanchet. From secrecy to authenticity in security protocols. In Manuel Hermenegildo and Germán Puebla, editors, *9th Int. Static Analysis Symposium (SAS'02)*, volume 2477 of *LNCS*, pages 342–359. Springer Verlag, September 2002.

[12] Bruno Blanchet. Automatic proof of strong secrecy for security protocols. In *IEEE Symposium on Security and Privacy*, pages 86–100, May 2004.

[13] Bruno Blanchet, Martín Abadi, and Cédric Fournet. Automated verification of selected equivalences for security protocols. In *20th IEEE Symposium on Logic in Computer Science (LICS 2005)*, pages 331–340. IEEE Computer Society, June 2005.

[14] Michele Boreale, Rocco De Nicola, and Rosario Pugliese. Proof techniques for cryptographic processes. In *Proceedings of the Fourteenth Annual IEEE Symposium on Logic in Computer Science*, pages 157–166, July 1999.

[15] Johannes Borgström. Static equivalence *is* harder than knowledge. In Jos Baeten and Iain Phillips, editors, *Proceedings of the 12th International Workshop on Expressiveness in Concurrency (EXPRESS'05)*, Electronic Notes in Theoretical Computer Science, pages 44–55. Elsevier Science Publishers, August 2005.

[16] Yannick Chevalier, Ralf Kuester, Michael Rusinowitch, and Mathieu Turani. Deciding the security of protocols with Diffie-Hellman exponentiation and products in exponents. In P. K. Pandya and J. Radhakrishnan, editors, *FST TCS 2003: Foundations of Software Technology and Theoretical Computer Science, 23rd Conference*, volume 2914 of *LNCS*, pages 124–135. Springer Verlag, 2003.

[17] Yannick Chevalier, Ralf Kuester, Michael Rusinowitch, and Mathieu Turani. An NP decision procedure for protocol insecurity with xor. In *Proceedings of the 18th Annual IEEE Symposium on Logic In Computer Science (LICS'03)*, pages 261–270, 2003.

[18] Hubert Comon-Lundh. Intruder theories (ongoing work). In *Foundations of Software Science and Computation Structures (FoSSaCS'04)*, volume 2987 of *LNCS*, pages 1–4. Springer, 2004.

[19] Hubert Comon-Lundh and Vitaly Shmatikov. Intruder deductions, constraint solving and insecurity decision in presence of exclusive or. In *Proceedings of the 18th Annual IEEE Symposium on Logic In Computer Science (LICS'03)*, pages 271–280, 2003.

[20] Hubert Comon-Lundh and Ralf Treinen. Easy intruder deductions. Technical Report LSV-03-8, Laboratoire Spécification et Vérification, ENS de Cachan, France, 2003.

[21] Stéphanie Delaune and Florent Jacquemard. Narrowing-based constraint solving for the verification of security protocols. Technical Report LSV-04-8, Laboratoire Spécification et Vérification, ENS de Cachan, France, April 2004.

[22] Nachum Dershowitz and Jean-Pierre Jouannaud. *Handbook of theoretical computer science*, volume B: formal models and semantics, chapter Rewrite systems, pages 243–320. MIT Press, 1991.

[23] Danny Dolev and Andrew C. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, IT-29(12):198–208, March 1983.

[24] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28:270–299, April 1984.

[25] Richard A. Kemmerer, Catherine Meadows, and Jonathan K. Millen. Three systems for cryptographic protocol analysis. *Journal of Cryptology*, 7(2):79–130, Spring 1994.

[26] Steve Kremer and Mark Ryan. Analysis of an electronic voting protocol in the applied pi calculus. submitted, 2004.

[27] Pascal Lafourcade, Denis Lugiez, and Ralf Treinen. Intruder deduction for ac-like equational theories with homomorphisms. In *Proceedings of the 16th International Conference on Rewriting Techniques and Applications (RTA'05)*, volume 3467 of *LNCS*, pages 308–322. Springer, April 2005.

[28] Patrick Lincoln, John Mitchell, Mark Mitchell, and Andre Scedrov. A probabilistic poly-time framework for protocol analysis. In *Proceedings of the Fifth ACM Conference on Computer and Communications Security*, pages 112–121, 1998.

[29] Gavin Lowe. Breaking and fixing the Needham-Schroeder public-key protocol using FDR. In *Tools and Algorithms for the Construction and Analysis of Systems*, volume 1055 of *LNCS*, pages 147–166. Springer Verlag, 1996.

[30] Lawrence C. Paulson. The inductive approach to verifying cryptographic protocols. *Journal of Computer Security*, 6(1–2):85–128, 1998.

[31] Ronald L. Rivest, Leonard Adleman, and Michael L. Dertouzos. On data banks and privacy homomorphisms. In R. DeMillo, D. Dobkin, A. Jones, and R. Lipton, editors, *Foundations of Secure Computation*, pages 169–180. Academic Press, 1978.

[32] Pierre Samuel. *Algebraic Theory of Numbers*. Hermann, 1970.

[33] Steve Schneider. Security properties and CSP. In *IEEE Symposium on Security and Privacy*, pages 174–187, 1996.

Appendix

A Proof of Proposition 6 and additional material on Proposition 5

Proposition 6. *The following problem is undecidable.*
Input: *Two machines $\mathcal{M}(M_1, M_2)$ and $\mathcal{M}(M_1', M_2')$ and a word w of A^*.*
Output: *Does the following property* **(P)** *hold for $\mathcal{M}(M_1, M_2)$ and $\mathcal{M}(M_1', M_2')$: for any sequences $s_1, s_2 \in \{1,2\}^*$, $\mathcal{M}(M_1, M_2), w \xrightarrow{s_1}$ and $\mathcal{M}(M_1, M_2), w \xrightarrow{s_2}$ have the same tape if and only if $\mathcal{M}(M_1', M_2'), w \xrightarrow{s_1}$ and $\mathcal{M}(M_1', M_2'), w \xrightarrow{s_2}$ have the same tape?*

The halting problem for a deterministic Turing machine can be reduced to this problem. Given any deterministic Turing machine $M = (Q, A, q_0, Q_f, \delta)$, we construct the deterministic Turing machine $\mathcal{T}(M) = (Q, A \uplus \{c_o\}, q_0, Q_f, \delta')$, where we modify the transitions for the final states:

$$\begin{cases} \delta'(q,a) = \delta(q,a) & \forall a \in A, q \notin Q_f \\ \delta'(q,a) = (q, c_0, L) & \forall a \in A, q \in Q_f. \end{cases}$$

Then $\mathcal{M}(M, \mathcal{T}(M)), w \xrightarrow{s_1}$ and $\mathcal{M}(M, \mathcal{T}(M)), w \xrightarrow{s_2}$ have the same tape for any sequences $s_1, s_2 \in \{1,2\}^*$ if and only if M does not reach its final state on w.

Now, let M_0 be any fixed deterministic Turing machine. For any sequences $s_1, s_2 \in \{1,2\}^*$, $\mathcal{M}(M_0, M_0), w \xrightarrow{s_1}$ and $\mathcal{M}(M_0, M_0), w \xrightarrow{s_2}$ have the same tape. We deduce that M does not reach its final state on w if and only $\mathcal{M}(M, \mathcal{T}(M))$ and $\mathcal{M}(M_0, M_0)$ satisfy the property (P). This ends the proof of proposition 6.

In order to reduce this undecidable problem to $\approx_s$, we consider the equational theory E_{tm} displayed in figure 2. By orienting the equations from left to right, we obtain convergent rewriting rules such that $M =_{E_{\mathsf{tm}}} M'$ if and only if $M{\downarrow} == M'{\downarrow}$ where $M{\downarrow}$ is the normal form of M for these rewriting rules. Intuitively, we consider terms of the form $h(w_1, q, w_2, s^n(0))$, where w_1 represents the tape before the machine's head, w_2 represents the tape after the machine's head, q is the control state, and $s^n(0)$ is a counter that represents the number of rules that have been applied. A term

$$\begin{aligned}
&\mathsf{Apply}(1,[(x_q, x_1 \to x_{q'}, x_2, R), y], h(z_1, x_q, x_1 \cdot z_2, x')) \\
&\quad = h(z_1 \cdot x_2, x_{q'}, z_2, s(x')) \\
&\mathsf{Apply}(1,[(x_q, x_1 \to x_{q'}, x_2, R), y], h(z_1, x_q, x_1, x')) \\
&\quad = h(z_1 \cdot x_2, x_{q'}, \#, s(x')) \\
&\mathsf{Apply}(1,[(x_q, x_1 \to x_{q'}, x_2, L), y], h(z_1 \cdot x_3, x_q, x_1 \cdot z_2, x')) \\
&\quad = h(z_1, x_{q'}, x_3 \cdot (x_2 \cdot z_2), s(x')) \\
&\mathsf{Apply}(2,[y, (x_q, x_1 \to x_{q'}, x_2, R)], h(z_1, x_q, x_1 \cdot z_2, x')) \\
&\quad = h(z_1 \cdot x_2, x_{q'}, z_2, s(x')) \\
&\mathsf{Apply}(2,[y, (x_q, x_1 \to x_{q'}, x_2, R)], h(z_1, x_q, x_1, x')) \\
&\quad = h(z_1 \cdot x_2, x_{q'}, \#, s(x')) \\
&\mathsf{Apply}(2,[y, (x_q, x_1 \to x_{q'}, x_2, L)], h(z_1 \cdot x_3, x_q, x_1 \cdot z_2, x')) \\
&\quad = h(z_1, x_{q'}, x_3 \cdot (x_2 \cdot z_2), s(x'))
\end{aligned}$$

Figure 2: The equational theory E_{tm}.

$[(q, a \to q_1, a_1, D_1), (q, a \to q_2, a_2, D_2)]$ represents a couple of rules of two Turing machine. Then the term

$$\mathsf{Apply}(i, [(q, a \to q_1, a_1, D_1), (q, a \to q_2, a_2, D_2)], h(w_1, q, w_2, s^n(0))),$$

where $i \in \{1, 2\}, D_1, D_2 \in \{L, R\}$, represents the application of the rule number 1 or 2 (depending on i) on the tape $h(w_1, q, w_2, s^n(0))$. The result of this application is given by the equational theory E_{tm}.

Now, to each machine $\mathcal{M}(M_1, M_2)$, we associate the frame $\phi_{\mathcal{M}(M_1,M_2)}$:

$$\nu(A \cup Q)[h(\#, q_0, \#, 0)/x_0] \cup \bigcup_{a \in A, q \in Q} [[(q, a \to \delta_1(q, a)), (q, a \to \delta_2(q, a))]/x_{a,q}]$$

Then we can verify that two machines $\mathcal{M}(M_1, M_2)$ and $\mathcal{M}(M_1', M_2')$ verify the property (P) of proposition 6 if and only if $\phi_{\mathcal{M}(M_1,M_2)} \approx_s \phi_{\mathcal{M}(M_1',M_2')}$. We deduce that $\approx_s$ is undecidable for the equational theory E_{tm}.

At the same time, $\vdash$ remains decidable: in order to decide whether $\phi \vdash M$, where $\phi = \nu\widetilde{n}\sigma$, it is sufficient to decide if there exists ζ such that $fn(\zeta) \cap \widetilde{n} = \emptyset$ and $\zeta\sigma =_{E_{\mathsf{tm}}} M$, that is, $\zeta\sigma\downarrow == M\downarrow$. Intuitively, for ϕ of the form $\phi_{\mathcal{M}(M_1,M_2)}$ and for M of the form $h(w_1, q, w_2, s^n(0))$, we are looking for some sequences of choices (represented by ζ) such that the tape of the machine $\mathcal{M}(M_1, M_2)$ after this sequence of choices is equal to M. Since the term M contains the number of rules that have been applied, it is sufficient to test any sequence of choices of length equal to this number of rules, so there is a finite number of sequences to check. This idea can be generalized to any ϕ and M, establishing that $\vdash$ is decidable. (We do not give the proof of this generalization, in light of Borgström's alternative proof of Proposition 5.)

B Proofs of Lemmas 10 and 11

Definition 11 *The set $\mathcal{P}(M)$ of* paths *of a term M is defined inductively by:*

$$\begin{aligned}\mathcal{P}(u) &= \epsilon \\ \mathcal{P}(f(M_1,\ldots,M_n)) &= \epsilon \cup \bigcup_{i=1}^{n} i\cdot\mathcal{P}(M_i) \quad \textit{for } i \le n\end{aligned}$$

The subterm of M at position $p \in \mathcal{P}(M)$, written $M|_p$, is defined inductively by:

$$\begin{aligned}M|_\epsilon &= M \\ f(M_1,\ldots,M_n)|_{i\cdot p} &= M_i|_p \quad \textit{for } i \le n\end{aligned}$$

Lemma 10. *Let E be a locally stable theory. Let $\phi = \nu\widetilde{n}\sigma$ and $\psi = \nu\widetilde{n'}\sigma'$ be two frames such that $\psi \models \mathsf{Eq}(\phi)$. For all contexts C_1 and C_2 such that $(fn(C_1)\cup fn(C_2))\cap \widetilde{n} = \emptyset$, for all terms $M_i, M'_i \in \mathsf{sat}(\phi)$, if $C_1[M_1,\ldots,M_k] =_{\mathsf{AC}} C_2[M'_1,\ldots,M'_l]$, then $(C_1[\zeta_{M_1},\ldots,\zeta_{M_k}] =_E C_2[\zeta_{M'_1},\ldots,\zeta_{M'_l}])\psi$.*

This lemma is proved by induction on the sum of the sizes of C_1 and C_2.

Base case: If $|C_1|, |C_2| \le c_E$, then the equation

$$(C_1[\zeta_{M_1},\ldots,\zeta_{M_k}] = C_2[\zeta_{M'_1},\ldots,\zeta_{M'_l}])$$

is in $\mathsf{Eq}(\phi)$ since $|C_1| \le c_E$ and $|C_2| \le c_E \le c_E^2$, so $\psi \models \mathsf{Eq}(\phi)$ implies $(C_1[\zeta_{M_1},\ldots,\zeta_{M_k}] =_E C_2[\zeta_{M'_1},\ldots,\zeta_{M'_l}])\psi$.

Inductive step: If neither C_1 nor C_2 is a hole, then $C_1 == f(C_1^1,\ldots,C_1^r)$ and $C_2 == f(C_2^1,\ldots,C_2^r)$. There are two cases.

- f is not an AC symbol. Then, for every $1 \le i \le r$, $C_1^i[M_1,\ldots,M_k] =_{\mathsf{AC}} C_2^i[M'_1,\ldots,M'_l]$. By applying the induction hypothesis, we obtain

  $$(C_1^i[\zeta_{M_1},\ldots,\zeta_{M_k}] =_E C_2^i[\zeta_{M'_1},\ldots,\zeta_{M'_l}])\psi$$

  so

  $$(C_1[\zeta_{M_1},\ldots,\zeta_{M_k}] =_E C_2[\zeta_{M'_1},\ldots,\zeta_{M'_l}])\psi$$

- f is an AC symbol $\oplus$. We write $C_1 = C_1^1 \oplus \cdots \oplus C_1^r \oplus x_1 \oplus \cdots \oplus x_p$ and $C_2 = C_2^1 \oplus \cdots \oplus C_2^{r'} \oplus y_1 \oplus \cdots \oplus y_{p'}$ in such a way that the head symbol of the contexts C_1^i and C_2^j is not $\oplus$, C_1^i and C_2^j are not holes, and the variables x_i and y_j refer to the holes of C_1 and C_2. If the equation can be split, with $C_1 =_{\mathsf{AC}} C'_1 \oplus C''_1$ and $C_2 =_{\mathsf{AC}} C'_2 \oplus C''_2$ such that $(C'_1[\zeta_{M_1},\ldots,\zeta_{M_k}] =_E C'_2[\zeta_{M'_1},\ldots,\zeta_{M'_l}])\phi$ and $(C''_1[\zeta_{M_1},\ldots,\zeta_{M_k}] =_E C''_2[\zeta_{M'_1},\ldots,\zeta_{M'_l}])\phi$, then we conclude as above, applying the induction hypothesis. On the other hand, if the equation cannot be split, for every $1 \le i \le r$, $N_i \stackrel{\text{def}}{=} C_1^i[M_1,\ldots,M_k]$ is not equal to some $C_2^j[M'_1,\ldots,M'_l]$ so it must be a subterm of some M'_j. Since each M'_j is in $\mathsf{sat}(\phi)$ and by

applying recursively rule 2 of Definition 6, we get that N_i is in $\mathsf{sat}(\phi)$, thus there exists $\zeta_{N_i} \in \rho(\phi)$ such that $\zeta_{N_i}\sigma =_E N_i$. Symmetrically, for every $1 \leq j \leq r$, $N'_j \stackrel{\text{def}}{=} C'^j_1[M'_1, \ldots, M'_k]$ is not equal to some $C^i_1[M_1, \ldots, M_l]$, so $N'_j \in \mathsf{sat}(\phi)$ and there exists $\zeta_{N'_j} \in \rho(\phi)$ such that $\zeta_{N'_j}\sigma =_E N'_j$.

- From $N_i == C^i_1[M_1, \ldots, M_k]$ and applying the induction hypothesis, we get $\zeta_{N_i}\sigma' =_E C^i_1[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma'$ and similarly, $\zeta_{N'_j}\sigma' =_E C'^j_2[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma'$.
- Renaming $C^i_1[M_1, \ldots, M_k]$ by N_i in our initial equation, we get $N_1 \oplus \cdots \oplus N_r \oplus M_1 \oplus \cdots \oplus M_p = N'_1 \oplus \cdots \oplus N'_{r'} \oplus M'_1 \oplus \cdots \oplus M'_{p'}$. Applying the base case, we get $(\zeta_{N_1} \oplus \cdots \oplus \zeta_{N_r} \oplus \zeta_{M_1} \oplus \cdots \oplus \zeta_{M_p} =_E \zeta_{N'_1} \oplus \cdots \oplus \zeta_{N'_{r'}} \oplus \zeta_{M'_1} \oplus \cdots \oplus \zeta_{M'_{p'}})\sigma$. Since this equation is in $\mathsf{Eq}(\phi)$, we deduce $(\zeta_{N_1} \oplus \cdots \oplus \zeta_{N_r} \oplus \zeta_{M_1} \oplus \cdots \oplus \zeta_{M_p} =_E \zeta_{N'_1} \oplus \cdots \oplus \zeta_{N'_{r'}} \oplus \zeta_{M'_1} \oplus \cdots \oplus \zeta_{M'_{p'}})\sigma'$.

Combining these equations, we get

$$(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\psi$$

If C_1 or C_2 is a hole, then let us say $C_1 == f(C^1_1, \ldots, C^r_1)$ and $C_2 == \_$. Let $M, M_1, \ldots, M_k \in \mathsf{sat}(\phi)$ and assume $C_1[M_1, \ldots, M_k] =_{\mathsf{AC}} M$. Again we consider two cases.

- f is not an AC symbol. Then we have

$$f(C^1_1[M_1, \ldots, M_k], \ldots, C^r_1[M_1, \ldots, M_k]) =_{\mathsf{AC}} M$$

For every $1 \leq i \leq r$, let $N_i \stackrel{\text{def}}{=} C^i_1[M_1, \ldots, M_k]$. Thus, each N_i is a subterm of M, so it is in $\mathsf{st}(\mathsf{sat}(\phi))$. Since each M_j is in $\mathsf{sat}(\phi)$ and by applying repeatedly rule 2 of Definition 6, we get that N_i is in $\mathsf{sat}(\phi)$. Thus there exists $\zeta_{N_i} \in \rho(\phi)$ such that $\zeta_{N_i}\sigma =_E N_i$.

  - From $N_i == C^i_1[M_1, \ldots, M_k]$ and applying the induction hypothesis, we get $\zeta_{N_i}\sigma' =_E C^i_1[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma'$.
  - From $M =_{\mathsf{AC}} f(N_1, \ldots, N_r)$ and applying the base case, we get $\zeta_M\sigma' =_E f(\zeta_{N_1}, \ldots, \zeta_{N_r})\sigma'$.

  Combining these equations, we get

$$(\zeta_M =_E C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}])\psi$$

- f is an AC symbol $\oplus$. We write $C_1 = C^1_1 \oplus \cdots \oplus C^r_1 \oplus x_1 \oplus \cdots \oplus x_p$ and $C_2 = x$, and we have $C^1_1[M_1, \ldots, M_k] \oplus \ldots \oplus C^r_1[M_1, \ldots, M_k] \oplus M'_1 \oplus \ldots \oplus M'_p =_{\mathsf{AC}} M$. Each $N_i \stackrel{\text{def}}{=} C^i_1[M_1, \ldots, M_k]$ is a subterm of $M \in \mathsf{sat}(\phi)$ thus is in $\mathsf{sat}(\phi)$. Again, there exists $\zeta_{N_i} \in \rho(\phi)$ such that $\zeta_{N_i}\sigma =_E N_i$.

  - From $N_i == C^i_1[M_1, \ldots, M_k]$ and applying the induction hypothesis, we get $\zeta_{N_i}\sigma' =_E C^i_1[\zeta_{M_1}, \ldots, \zeta_{M_k}]\sigma'$.

– From $N_1 \oplus \ldots \oplus N_r \oplus M'_1 \oplus \ldots \oplus M'_p =_{\mathsf{AC}} M$ and by the equation $\zeta_{N_1} \oplus \cdots \oplus \zeta_{N_r} \oplus \zeta_{M'_1} \oplus \cdots \oplus \zeta_{M'_p} =_E \zeta_M$ is in $\mathsf{Eq}(\phi)$, we get $(\zeta_{N_1} \oplus \cdots \oplus \zeta_{N_r} \oplus \zeta_{M'_1} \oplus \cdots \oplus \zeta_{M'_p} =_E \zeta_M)\sigma'$.

Combining these equations, we get

$$(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E \zeta_M)\psi$$

Lemma 11. *Let E be a locally stable theory. Let $\phi = \nu\widetilde{n}\sigma$ be a frame. For every context C_1 such that $fn(C_1) \cap \widetilde{n} = \emptyset$, for every $M_i \in \mathsf{sat}(\phi)$, for every term T such that $C_1[M_1, \ldots, M_k] \to_{\mathsf{AC}} T$, there exist a context C_2 such that $fn(C_2) \cap \widetilde{n} = \emptyset$, and terms $M'_i \in \mathsf{sat}(\phi)$, such that $T \to^*_{\mathsf{AC}} C_2[M'_1, \ldots, M'_l]$ and for every frame $\psi \models \mathsf{Eq}(\phi)$, $(C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] =_E C_2[\zeta_{M'_1}, \ldots, \zeta_{M'_l}])\psi$.*

An easy case is when the reduction occurs inside one of the terms M_i: $M_i \to_{\mathsf{AC}} M'_i$. By definition of $\mathsf{sat}(\phi)$ (since E is locally stable), we know that there exists C such that $|C| \leq c_E^2$, $fn(C) \cap \widetilde{n} = \emptyset$, and $M'_i \to^*_{\mathsf{AC}} C[M''_1, \ldots, M''_l]$ where $M''_i \in \mathsf{sat}(\phi)$. In addition, the equation $\zeta_{M_i} = C[\zeta_{M''_1}, \ldots, \zeta_{M''_l}]$ is in $\mathsf{Eq}(\phi)$ (since $|C| \leq c_E^2$), thus $(\zeta_{M_i} =_E C[\zeta_{M''_1}, \ldots, \zeta_{M''_l}])\psi$. We obtain that

$$\begin{aligned} T == C_1[M_1, \ldots, M_{i-1}, M'_i, M_{i+1}, \ldots, M_k] \\ \to^*_{\mathsf{AC}} C_1[M_1, \ldots, C[M''_1, \ldots, M''_l], \ldots, M_k] \end{aligned}$$

and

$$\left(\begin{array}{l} (C_1[\zeta_{M_1}, \ldots, \zeta_{M_k}] \\ \quad =_E \\ C_1[\zeta_{M_1}, \ldots, C[\zeta_{M''_1}, \ldots, \zeta_{M''_l}], \ldots, \zeta_{M_k}] \end{array} \right) \psi$$

We now consider the case where the reduction does not occur inside the terms M_i. We can assume that

> *for every path p of C_1,*
> *if $C_1|_p[M_1, \ldots, M_k]$ is in $\mathsf{sat}(\phi)$,* (*)
> *then $C_1|_p$ is the single hole context.*

Indeed, if there exists a path p of C_1 such that $T_1 \stackrel{\text{def}}{=} C_1|_p[M_1, \ldots, M_k] \in \mathsf{sat}(\phi)$ and $C_1|_p$ is not a hole then $C_1[M_1, \ldots, M_k] == C'_1[T_1, M_1, \ldots, M_k]$ where $T_1, M_i \in \mathsf{sat}(\phi)$ and C'_1 is a context strictly smaller than C_1. In that case, we consider $C'_1[T_1, M_1, \ldots, M_k]$ instead of $C_1[M_1, \ldots, M_k]$ and we apply the transformation again until property (*) holds.

We have

$$C_1[M_1, \ldots, M_k] \quad == \quad C_3[M'' \oplus M' \oplus \bigoplus_{i=1}^{r} C'_i[M_1, \ldots, M_k], M_1, \ldots, M_k]$$

where $M' = M'_1 \oplus \ldots \oplus M'_l$, $M'' = M''_1 \oplus \ldots \oplus M''_l$ with $M'_i \oplus M''_i \in \mathsf{sat}(\phi)$, the head symbol of the context C'_i is not $\oplus$, C'_i is not a single hole, and $T_1 \stackrel{\text{def}}{=} M' \oplus \bigoplus_{i=1}^{r} C'_i[M_1, \ldots, M_k]$ is an instance $M_0\theta$ (modulo AC) of the left-hand side of some rule $M_0 \to N_0$ of the rewriting system associated with E.

For each variable x of M_0, we consider the occurrences of $x\theta$ in T_1.

1. Either $x\theta$ occurs as a subterm of one of the terms M_i or M'_i;

2. or there exists a subterm of T_1 of the form $N_1 \oplus \ldots \oplus N_p$ with $N_i =_{\mathsf{AC}} N'_i \oplus N''_i \in \mathsf{sat}(\phi)$ for some N''_i such that $x\theta =_{\mathsf{AC}} N'_1 \oplus \ldots \oplus N'_p$;

3. or there exists a subterm of T_1 of the form

$$N_1 \oplus \ldots \oplus N_p \oplus \bigoplus_{i=1}^{r'} C''_i[M_1, \ldots, M_k]$$

(modulo AC) where the head symbols of the contexts C''_i are not $\oplus$ and the contexts C''_i are not a hole, and

$$x\theta =_{\mathsf{AC}} N'_1 \oplus \ldots \oplus N'_p \oplus \bigoplus_{i=1}^{r'} C''_i[M_1, \ldots, M_k]$$

with $N_i =_{\mathsf{AC}} N'_i \oplus N''_i \in \mathsf{sat}(\phi)$ for some N''_i, thus the terms N'_i are subterms of terms of $\mathsf{sat}(\phi)$.

Note that case 3 cannot occur simultaneously with case 1 or case 2 for the same variable x. If case 3 were to occur simultaneously with case 1 or case 2, we would have that some $C''_i[M_1, \ldots, M_k]$ is a subterm of some M_i or M'_i, thus applying recursively rule 2 of Definition 6, we would get that $C''_i[M_1, \ldots, M_k] \in \mathsf{sat}(\phi)$, which contradicts property (*) (since C''_i is not a hole).

Without loss of generality, we assume that the variables of M_0 are $x_1, \ldots, x_{k_1}, y_1, \ldots, y_{k_2}$ where the variables x_i are in case 1 or case 2 and the variables y_j are in case 3. For each variable y_j, we consider the l occurrences of y_j in T_1.

$$\begin{array}{rcl} y_j\theta & =_{\mathsf{AC}} & N^1_1 \oplus \ldots \oplus N^1_{k_1} \oplus \bigoplus_{i=1}^{r_1} C^1_i[M_1, \ldots, M_k] \\ & & \vdots \\ & =_{\mathsf{AC}} & N^l_1 \oplus \ldots \oplus N^l_{k_l} \oplus \bigoplus_{i=1}^{r_l} C^l_i[M_1, \ldots, M_k] \end{array}$$

where the terms N^j_i are subterms of terms in $\mathsf{sat}(\phi)$ and the head symbols of the contexts C^j_i are not $\oplus$.

We write $cl(C^j_i[M_1, \ldots, M_k])$ for the class of $C^j_i[M_1, \ldots, M_k]$ modulo AC, and we associate a fresh name symbol $a_{cl(C^j_i[M_1,\ldots,M_k])}$ with each $cl(C^j_i[M_1, \ldots, M_k])$. Therefore, $a_{cl(C^{j_1}_{i_1}[M_1,\ldots,M_k])}$ and $a_{cl(C^{j_2}_{i_2}[M_1,\ldots,M_k])}$ are the same symbol whenever $C^{j_1}_{i_1}[M_1, \ldots, M_k] =_{\mathsf{AC}} C^{j_2}_{i_2}[M_1, \ldots, M_k]$. In each equation

$$N^{j_1}_1 \oplus \ldots \oplus N^{j_1}_{k_{j_1}} \oplus \bigoplus_{i=1}^{r_{j_1}} C^{j_1}_i[M_1, \ldots, M_k] \\ =_{\mathsf{AC}} N^{j_2}_1 \oplus \ldots \oplus N^{j_2}_{k_{j_2}} \oplus \bigoplus_{i=1}^{r_{j_2}} C^{j_2}_i[M_1, \ldots, M_k]$$

every $C_i^{j_1}[M_1, \ldots, M_k]$ must be equal modulo AC to one of the terms $C_i^{j_2}[M_1, \ldots, M_k]$. If $C_i^{j_1}[M_1, \ldots, M_k]$ were equal to some subterm of the terms $N_i^{j_2}$, $C_i^{j_1}[M_1, \ldots, M_k]$ would be a term of $\mathsf{sat}(\phi)$, contradicting property (*). Thus, we obtain that

$$\begin{array}{rl} & N_1^1 \oplus \ldots \oplus N_{k_1}^1 \oplus \bigoplus_{i=1}^{r_1} a_{C_i^1[M_1,\ldots,M_k]} \\ & \vdots \\ =_{\mathsf{AC}} & N_1^l \oplus \ldots \oplus N_{k_l}^l \oplus \bigoplus_{i=1}^{r_l} a_{C_i^l[M_1,\ldots,M_k]} \stackrel{\text{def}}{=} T_{y_j} \end{array}$$

We consider the substitution θ' such that $x_i\theta' = x_i\theta$ and $y_j\theta' = T_{y_j}$. We define $\theta''(a_{cl(C_i^j[M_1,\ldots,M_k])}) = C_i^j[M_1, \ldots, M_k]$.

We also consider the term T_2 that is obtained from $\bigoplus_{i=1}^{r} C_i'[M_1, \ldots, M_k]$ by replacing each $C_i^j[M_1, \ldots, M_k]$ with $a_{cl(C_i^j[M_1,\ldots,M_k])}$.

We have $T_2 == C_2[S_1, \ldots, S_k]$ for some context C_2 such that $|\_ \oplus C_2| \leq |M_0| \leq c_E$ and $S_i \in \mathsf{sum}_\oplus(\mathsf{sat}(\phi), \widetilde{n})$. Since $M'' \oplus T_2$ is an instance $M_0\theta'$ of M_0 we have $M' \oplus M'' \oplus T_2 \to_{\mathsf{AC}} M' \oplus N_0\theta'$. Applying condition 3 of Definition 6, there exist $S_i' \in \mathsf{sum}_\oplus(\mathsf{sat}(\phi), \widetilde{n})$, there exists a context C', such that $|C'| \leq c_E^2$, $fn(C') \cap \widetilde{n} = \emptyset$, and $M' \oplus N_0\theta' \to_{\mathsf{AC}}^* C'[S_1', \ldots, S_l']$. Applying the substitution θ'', we deduce that $M' \oplus N_0\theta =_{\mathsf{AC}} M' \oplus N_0\theta'\theta'' \to_{\mathsf{AC}}^* C'[S_1', \ldots, S_l']\theta''$. Note that $C'[S_1', \ldots, S_l']\theta''$ is a context of terms of $\mathsf{sat}(\phi)$:

$$C'[S_1', \ldots, S_l']\theta'' = C''[M_1, \ldots, M_k, S_1', \ldots, S_l']$$

To each sum $S = \alpha_1 M_1 \oplus \cdots \oplus \alpha_n M_n \oplus \beta_1 n_1 \oplus \cdots \oplus \beta_k n_k$ in $\mathsf{sum}_\oplus(\mathsf{sat}(\phi), \widetilde{n})$, we associate the term $\zeta_S = \alpha_1 \cdot_\oplus \zeta_{M_1} \oplus \cdots \oplus \alpha_n \cdot_\oplus \zeta_{M_n} \oplus \beta_1 \cdot_\oplus n_1 \oplus \cdots \oplus \beta_k \cdot_\oplus n_k$.

Now, since the equation $\zeta_{M' \oplus M''} \oplus C_2[\zeta_{S_1}, \ldots, \zeta_{S_k}] = C'[\zeta_{S_1'}, \ldots, \zeta_{S_l'}]$ is in $\mathsf{Eq}(\phi)$, we deduce

$$(\zeta_{M' \oplus M''} \oplus C_2[\zeta_{S_1}, \ldots, \zeta_{S_k}] = C'[\zeta_{S_1'}, \ldots, \zeta_{S_l'}])\psi$$

If $a_{cl(C_{i_1}^{j_1}[M_1,\ldots,M_k])} = a_{cl(C_{i_2}^{j_2}[M_1,\ldots,M_k])}$, we have

$$C_{i_1}^{j_1}[M_1, \ldots, M_k] =_{\mathsf{AC}} C_{i_2}^{j_2}[M_1, \ldots, M_k]$$

thus (by Lemma 10) we have

$$(C_{i_1}^{j_1}[\zeta_{M_1}, \ldots, \zeta_{M_k}] = C_{i_2}^{j_2}[\zeta_{M_1}, \ldots, \zeta_{M_k}])\psi$$

So we can reconstruct $M'' \oplus T_1$ and obtain

$$\zeta_{M' \oplus M''} \oplus \bigoplus_{i=1}^{r} C_i'[\zeta_{M_1}, \ldots, \zeta_{M_k}] = C''[\zeta_{M_1}, \ldots, \zeta_{M_k}, \zeta_{S_1'}, \ldots, \zeta_{S_l'}])\psi$$

which allows us to conclude the proof of Lemma 11.