Proof System for Applied Pi Calculus

Jia Liu, Huimin Lin

Proof System for Applied Pi Calculus

Jia Liu[1,2] ⋆ and Huimin Lin[1]

[1] State Key Laboratory of Computer Science
Institute of Software, Chinese Academy of Sciences
[2] Graduate University, Chinese Academy of Sciences
{jliu, lhm}@ios.ac.cn

Abstract. A symbolic-style proof system is presented to reason about observational equivalence for applied pi-calculus. The proofs of the soundness and completeness of the system rely on a recently developed theory of symbolic bisimulation for applied pi-calculus. The completeness result of the proof system is restricted to the finite fragment of applied pi-calculus which admits finite partition, and it is demonstrated that this fragment covers an important subset of applied pi-calculus which is practically useful for analyzing security protocols.

1 Introduction

The applied pi-calculus is a descendant of the pi-calculus designed for cryptographic applications. It extends pi-calculus with value-passing, primitive function symbols and equational theory. To capture the knowledge exposed by processes to the environment, *active substitutions* are employed. For example, let $A = \nu k.(a(x).\ \mathit{if}\ dec(x,k) = m\ \mathit{then}\ \overline{a}\ \mathit{else}\ \overline{c} \mid \{enc(m,k)/y\})$. Process A contains an active substitution $\{enc(m,k)/y\}$, where $enc(m,k)$ denotes a ciphertext obtained by encrypting the plaintext m by the secret key k and y can be regarded as an alias of the ciphertext. The secret key k in process A is restricted since we do not wish k to be visible to the environment, while the ciphertext can be accessed through the alias y. To model the shared-key cryptography, we use the equation $dec(enc(w_1,w_2),w_2) = w_1$ to decrypt the ciphertext. Thus the equality test $dec(x,k) = m$ can be satisfied when x takes the value represented by y, leading to the following transitions in concrete semantics:

$$\begin{aligned} A &\xrightarrow{a(y)} \nu k.(\ \mathit{if}\ dec(y,k) = m\ \mathit{then}\ \overline{a}\ \mathit{else}\ \overline{c} \mid \{enc(m,k)/y\}) \\ &\equiv \nu k.(\ \mathit{if}\ dec(enc(m,k),k) = m\ \mathit{then}\ \overline{a}\ \mathit{else}\ \overline{c} \mid \{enc(m,k)/y\}) \\ &\xrightarrow{\tau} \nu k.(\overline{a} \mid \{enc(m,k)/y\}) \end{aligned}$$

Security protocols are modeled as processes in the applied pi calculus and security properties such as anonymity, privacy and strong secrecy can be expressed as indistinguishability properties from the view of attackers, formalized

⋆ This work is supported by the National Natural Science Foundation of China (Grants No.60721061 and No.60833001).

by the notion of *observational equivalence*. Two processes are observationally equivalent if they cannot be distinguished in any context. A context models an active attacker which can intercept and forge messages. The universal quantification over contexts makes observational equivalence difficult to check, hence an alternative notion of *labeled bisimilarity* is introduced in [1] which relies on direct comparison of labeled transitions rather than contexts. However, in labeled transitional semantics, an input prefix may give rise to infinitely many branches, as in $a(x).P \xrightarrow{a(M)} P\{M/x\}$, for *every* term M, which hinders computer-assisted verification. To hurdle this problem, *symbolic bisimulations* have recently been advocated for the applied pi-calculus [9] and [15], and the later is shown exactly captures observational equivalence. The aim of this paper is to formulate a proof system to reason about observational equivalence, based on the symbolic bisimulation theory of [15].

The statements of our proof system are of the form $(\mathcal{D}, \Phi) \rhd A = B$ where $(\mathcal{D}, \Phi)$ is a *constraint* consisting of a trail $\mathcal{D}$ and a formula Φ. The proof system consists of *axioms* and *inference rules*. Different from the previous works [10, 16, 3, 12, 14], the basic entities of the proof system are *agents* of the form $\nu\widetilde{n}.(P \mid \sigma)$, where σ is a collection of active substitutions, rather than process P. The reasoning crosses through the frame and directly applies to the process part, as in the rule

$$\textbf{Tau} \quad \frac{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(P \mid \sigma) = \nu\widetilde{m}.(Q \mid \theta)}{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(\tau.P \mid \sigma) = \nu\widetilde{m}.(\tau.Q \mid \theta)}.$$

This is because the equality tests in P should be evaluated with the knowledge represented by the "frame" $\nu\widetilde{n}.\sigma$. For example, we can derive $(\emptyset, \mathit{true}) \rhd \nu s.(a(x).[x = s]\overline{b}\langle c\rangle \mid \{s/y\}) = \nu k.(a(x).[\mathit{dec}(x,k) = m]\overline{b}\langle c\rangle \mid \{\mathit{enc}(m,k)/y\})$; However, we cannot derive $(\emptyset, \mathit{true}) \rhd a(x).[x = s]\overline{b}\langle c\rangle = a(x).[\mathit{dec}(x,k) = m]\overline{b}\langle c\rangle$, because the equality tests $[x = s]$ and $[\mathit{dec}(x,k) = m]$ cannot be satisfied at the same time without the knowledge exposed by the frames $\nu s.\{s/y\}$ and $\nu k.\{\mathit{enc}(m,k)/y\}$.

The proof system is for agent equivalence and has to inevitably rely on some form of reasoning about the underlying equational theories on terms (which are parameters to applied pi-calculus). We have decided to factor out reasoning on terms from the proof system, using "semantical judgments" of the form $\Phi \models_{\mathcal{D}} \Psi$, as can be seen in the following rule:

$$\textbf{Partition} \quad \frac{(\mathcal{D}, \Phi_i) \rhd A = B,\ i = 1, 2,\ \Phi \models_{\mathcal{D}} \Phi_1 \vee \Phi_2}{(\mathcal{D}, \Phi) \rhd A = B}.$$

The rule states that, if we can infer $(\mathcal{D}, \Phi_1) \rhd A = B$ and $(\mathcal{D}, \Phi_2) \rhd A = B$ in the proof system, and we know, by some means, that Φ semantically implies $\Phi_1 \vee \Phi_2$ under $\mathcal{D}$, then we can derive $(\mathcal{D}, \Phi) \rhd A = B$. One may think of such semantical judgments as questions about the term domain, to be answered by an "oracle". In practice they can be resolved by invoking some decision procedures, like the one in [2] for instance, or appealing to a separate proof system specially designed for the underlying equational theories.

Our proof system is sound in general while complete on a class of finite processes on which finite partition on constraint systems always suffices. We will show that this class of processes covers an important fragment of the applied pi-calculus termed *simple processes*, which has been used for describing and analyzing security protocols.

Due to space limitation proofs are sketched. For a complete and rigorous treatment please refer to the full version of this paper, available at `http://lcs.ios.ac.cn/~jliu`.

2 Applied Pi Calculus

Applied pi-calculus [1] is an extension of pi-calculus with value-passing, primitive functions and equational theory. We assume two disjoint, infinite sets $\mathcal{N}$ and $\mathcal{V}$ of names and variables, respectively. An implicit sort system, including a *base sort* and a *channel sort*, splits $\mathcal{N}$ (resp. $\mathcal{V}$) into base sort $\mathcal{N}_b$ (resp. $\mathcal{V}_b$) and channel sort $\mathcal{N}_{ch}$ (resp. $\mathcal{V}_{ch}$). Unless otherwise stated, we will use a, b, c to range over channel names, s, k over base names, and m, n over names of either sort; we will also use x, y, z to range over variables, and u, v, w over either names or variables. Function symbols, such as f, enc, dec etc., are required to take arguments and produce results of base sort only. Terms, ranged over by M, N, are builded up from names and variables by function applications. We shall write $var(M)$ and $name(M)$ for variables and names respectively in M. *Extended processes* are created by extending *plain processes* with *active substitutions* of the form $\{M/x\}$ which is required to be defined on base sort only.

$P_r, Q_r, R_r ::=$ plain processes	$A_r, B_r, C_r ::=$ extended processes
0	P_r
$P_r \mid Q_r$	$A_r \mid B_r$
$!P_r$	$\nu n.A_r$
$\nu n.P_r$	$\nu x.A_r$
if $M = N$ then P_r then Q_r	$\{M/x\}$
$u(x).P_r$	
$\overline{u}\langle N\rangle.P_r$	

In an extended process, there is at most one substitution for each variable and exactly one when the variable is restricted. Substitutions are sort-respecting partial mappings of finite domains. Substitutions of terms for variables, ranged over by σ, θ, are always required to be cycle-free. The domain and range of σ are denoted $dom(\sigma)$ and $ran(\sigma)$, respectively. $Z\sigma$ is the result of applying σ to Z. The null process 0 is identified with the empty substitution. A substitution $\theta = \{M_1/x_1, \cdots, M_n/x_n\}$ will be identified with the parallel composition $\{M_1/x_1\} \mid \cdots \mid \{M_n/x_n\}$, and $\theta\sigma$ is defined as $\{M_1\sigma/x_1, \cdots, M_n\sigma/x_n\}$. A substitution σ is *idempotent* if $dom(\sigma) \cap var(ran(\sigma)) = \emptyset$. We shall write σ^* for the result of iterating σ until reaching idempotence, and use ϱ to denote one-to-one renaming of names and variables. To avoid confusion, we write $\varrho(Z)$ for the application of ϱ to Z, and $\varrho(\theta)$ means $\{\varrho(M_1)/\varrho(x_1), \cdots, \varrho(M_n)/\varrho(x_n)\}$.

We shall write $fn(A_r)$, $bn(A_r)$, $fv(A_r)$ and $bv(A_r)$ for the sets of free and bound names, free and bound variables, respectively, of A_r. A_r is *closed* if every variable in A_r is either bound or defined by an active substitution.

Terms are equipped with an equational theory $=_E$ that is an equivalence relation closed under substitutions of terms for variables, one-to-one renamings, and term contexts.

Observational equivalence $\approx$ [1] is a contextual equivalence relation on closed extended processes such that $A_r \approx B_r$ implies $C[A_r] \approx C[B_r]$ for any context C. Contexts model active attackers who can intercept and forge messages. Thus observational equivalence captures security properties in the presence of attackers, such as anonymity and privacy. Since the universal quantification over contexts makes $\approx$ difficult to verify, an alternative characterisation, namely *labeled bisimilarity*, is introduced in [1] which relies on direct comparison of labeled transitions rather than contexts. To overcome the problem of infinite branching caused by input transitions in labeled bisimulation, *symbolic bisimulations* are proposed in [9] and [15], and the notion of symbolic bisimulation presented in the later has been shown to be sound and complete w.r.t. $\approx$. We shall briefly review the symbolic semantics of [15] in next section.

3 Symbolic Semantics

Language For technical reasons, symbolic semantics [15] is built up on top of "intermediate processes", originally proposed in [9], which is a sufficient subset of extended processes. For the purpose of axiomatisation we extend the language of [15] with summation.

S, T	$::=$	$\mathit{true} \mid M = N \mid \neg S \mid S \wedge T$	
π	$::=$	$\tau \mid u(x) \mid \overline{u}\langle M\rangle$	prefix
P, Q, R	$::=$	$0 \mid S\pi.P \mid P + Q \mid P \mid Q \mid !P_r$	plain agents
F, G, H	$::=$	$P \mid \{M/x\} \mid F \mid G$	framed agents
A, B, C	$::=$	$F \mid \nu n.A \mid A + B$	agents

Here $S\pi.P$ is one-armed conditional, and the two-armed conditional operator "*if* $M = N$ *then* P *else* Q" of [1] and [15] can be defined as "$(M = N)\tau.P + \neg(M = N)\tau.Q$". We abbreviate $\mathit{true}\,\pi.P$ to $\pi.P$ and $\neg(M = N)$ to $M \neq N$. The domain of a framed agent F, denoted by $dom(F)$, is the set of variables x for which F contains a substitution $\{M/x\}$. Each framed agent F is required to be *applied*, that is, every variable in $dom(F)$ occurs only once in F. For example, $\overline{a}\langle k\rangle \mid \{k/x\}$ is applied but $\overline{a}\langle x\rangle \mid \{k/x\}$ is not.

The choice operator + does not appear in the original applied pi-calculus. We introduce it here in order to axiomatize parallel composition, as in the case of CCS and pi-calculus. Thus the operator merely serves as a vehicle to achieve a complete axiomatization, not intended to be used by the users. Since + is only used when a parallel composition is expanded, it is reasonable to required $dom(A) = dom(B)$ in a summation $A+B$, and $dom(A+B)$ is defined as $dom(A)$.

For an agent A, we define the *frame* $\varphi(A)$ *of* A as follows:

$$\varphi(0) = \varphi(S\pi.P) = 0 \quad \varphi(\{M/x\}) = \{M/x\} \quad \varphi(A \mid B) = \varphi(A) \cup \varphi(B)$$
$$\varphi(A+B) = \begin{cases} 0 & \text{if } \varphi(A) = 0 \text{ or } \varphi(B) = 0 \\ \varphi(A) + \varphi(B) & \text{otherwise} \end{cases}$$
$$\varphi(\nu n.A) = \textstyle\sum_{i\in I} \nu n.\nu\widetilde{m}_i.\sigma_i \text{ where } \varphi(A) = \sum_{i\in I} \nu\widetilde{m}_i.\sigma_i$$

Constraints A *constraint* $(\mathcal{D}, \Phi)$ is a pair where $\mathcal{D}$ is a *trail* and Φ a *formula*. A trail abstractly represents the ability of the attackers to deduce messages from a given set of messages. We shall use $\mathcal{D}, \mathcal{E}, \mathcal{F}$ to range over trails.

Formally, a *trail* is a set of the form $\{x_1 : U_1, \cdots, x_\ell : U_\ell\}$ where x_i are variables and U_i are finite sets of channel names and base variables, satisfying:
1. $x_1, \cdots, x_\ell$ are pairwise-distinct and do not appear in any U_j, $1 \le j \le \ell$;
2. for each $1 \le i < \ell$, $name(U_i) \supseteq name(U_{i+1})$ and $var(U_i) \subseteq var(U_{i+1})$.

For a trail $\mathcal{D} = \{x_i : U_i\}_{i\in I}$, let $dom(\mathcal{D}) = \{x_i\}_{i\in I}$ and $fnv(\mathcal{D}) = dom(\mathcal{D}) \cup \bigcup_{i\in I} U_i$. Let A be an agent with $\varphi(A) = \sum_{j\in J} \nu\widetilde{n}_j.\sigma_j$. $\mathcal{D}$ is *compatible with* A if the following conditions are satisfied:
1. $dom(\mathcal{D}) \cap dom(A) = \emptyset$,
2. $var(\bigcup_{i\in I} U_i) \subseteq dom(A)$, $fv(A) \subseteq dom(A) \cup dom(\mathcal{D})$, and
3. for any $x_i : U_i$ and $y \in U_i$ with $i \in I$, $x_i \notin var(y\sigma_j)$ for every $j \in J$.

Intuitively, the variables x_i in $\mathcal{D}$ are input variables. The corresponding U_i records all the variables that *can* be used by x_i and the names that *cannot* be used by x_i, at the moment when the input action of x_i fires.

A substitution θ *respects* $\mathcal{D}$ if
1. $dom(\theta) = dom(\mathcal{D})$,
2. for any $i \in I$, $var(x_i\theta) \subseteq U_i$ and $name(x_i\theta) \cap U_i = \emptyset$.

Example 1. Let $A = \overline{c}\langle x\rangle \mid \{h(y)/w_1, g(y,z)/w_2\}$ and $\mathcal{D} = \{x : \{c\}, y : \emptyset, z : \{w_1\}\}$. Then $\mathcal{D}$ is a trail which is compatible with A. The substitution $\{a/x, h(k)/y, f(k,w_1)/z\}$ respects $\mathcal{D}$, while $\{c/x, h(k)/y, f(w_2)/z\}$ does not (because c cannot be used by x, and w_2 cannot be used by z).

Formulas are specified by the following grammar:

$$\Phi,\ \Psi \ ::= S \ \mid\ \sigma \blacktriangleright \Phi \ \mid\ \Phi \wedge \Psi \ \mid\ \neg\Phi \ \mid\ \mathrm{H}n.\Phi$$

S is a formula as defined in the previous page. In $\sigma \blacktriangleright \Phi$, σ is an idempotent substitution that represents the environmental knowledge accumulated so far to define some variables occurring in Φ. $\mathrm{H}n.\Phi$ hides n in Φ and n is binding. We shall identify α-convertible formulas. We write *false* for $\neg$*true*, $\Phi \vee \Psi$ for $\neg(\neg\Phi \wedge \neg\Psi)$, $\Phi \Rightarrow \Psi$ for $\neg\Phi \vee \Psi$, and $\Phi \Leftrightarrow \Psi$ for $(\Phi \Rightarrow \Psi) \wedge (\Psi \Rightarrow \Phi)$.

The satisfiability relation $\models$ is defined between idempotent substitutions and formulas as follows, where the standard clauses for negation and conjunction are omitted:

$$\theta \models M = N \ \ \textit{if} \ \ M\theta =_E N\theta$$
$$\theta \models \sigma \blacktriangleright \Phi \ \ \textit{if} \ \ \theta\sigma \text{ is cycle-free and } (\theta\sigma)^* \models \Phi$$
$$\theta \models \mathrm{H}n.\Phi \ \ \textit{if} \ \ \exists m \notin fn(\mathrm{H}n.\,\Phi) \cup name(\theta) \text{ such that } \theta \models \{m/n\}\Phi$$

We write $\Phi \models_{\mathcal{D}} \Psi$ to mean: $\theta \models \Phi$ implies $\theta \models \Psi$ for any θ respecting $\mathcal{D}$.

Definition 1 (Partition). *A collection of formulas Σ is a* partition of Φ under $\mathcal{D}$ *if for any θ respecting $\mathcal{D}$ it holds that $\theta \models \Phi$ implies $\theta \models \Psi$ for some $\Psi \in \Sigma$.*

Example 2. Let $\mathcal{D} = \{x : \{y\}\}$, $\Phi = \mathrm{H}s.(\{enc(m,s)/y\} \blacktriangleright dec(x,s) = m)$ and $\Psi = \mathrm{H}k.(\{k/y\} \blacktriangleright x = k)$, with the equation $dec(enc(w_1,w_2),w_2) =_E w_1$. Then we have $\{y/x\} \models \Phi$ because $\{enc(m,s)/x\} \models dec(x,s) = m$. Similarly $\{y/x\} \models \Psi$. Moreover we can deduce that $(x = y) \models_{\mathcal{D}} \Phi \wedge \Psi$.

Symbolic Semantics Symbolic semantics will be defined modulo *symbolic structural equivalence* $\equiv_s$, which is defined by the AC properties of $|$ with neutral 0 and the AC properties of $+$, such as $(\overline{a}\langle b\rangle \mid 0) + \overline{c}\langle k\rangle \equiv_s \overline{c}\langle k\rangle + \overline{a}\langle b\rangle$.

Symbolic actions are of the form $\tau, u(x), \overline{u}\langle v\rangle$ or $\nu w.\overline{u}\langle w\rangle$, where $u, v \in \mathcal{N}_{ch} \cup \mathcal{V}_{ch}$ and $w \in \mathcal{N}_{ch} \cup \mathcal{V}_b$. For two symbolic actions α and β with the same bound objects, we use $[\alpha = \beta]$ to denote the formula obtained by comparison of their subjects and free objects; for instance $[\overline{u}\langle w\rangle = \overline{v}\langle w'\rangle]$ denotes $(u = v) \wedge (w = w')$ and $[u(x) = v(x)]$ denotes $u = v$.

Symbolic transition relations, $\{ \xrightarrow{\Phi,\alpha} \mid \Phi \text{ a formula, } \alpha \text{ a symbolic action} \}$, are defined on agents by the following typical rules:

$$S\,u(x).P \xrightarrow{S,\ u(x)} P \qquad S\,\overline{u}\langle v\rangle.P \xrightarrow{S,\overline{u}\langle v\rangle} P$$

$$S\overline{u}\langle M\rangle.P \xrightarrow[x \in \mathcal{V}_b,\ x \notin fv(S\overline{u}\langle M\rangle.P)]{S,\nu x.\overline{u}\langle x\rangle} P \mid \{M/x\} \qquad !P_r \xrightarrow[\Gamma(P_r) = \nu\widetilde{m}.P]{true,\tau} \nu\widetilde{m}.(P \mid !P_r)$$

$$\frac{A \xrightarrow{\Phi,\alpha} A' \quad n \notin name(\alpha)}{\nu n.A \xrightarrow{\mathrm{H}n.\Phi,\alpha} \nu n.A'} \qquad \frac{A \xrightarrow{\Phi,\overline{u}\langle c\rangle} A' \quad u \neq c}{\nu c.A \xrightarrow{\mathrm{Hc}.\Phi,\nu c.\overline{u}\langle c\rangle} A'} \qquad \frac{A \xrightarrow{\Phi,\alpha} A'}{A + B \xrightarrow{\Phi,\alpha} A'}$$

$$\frac{A \xrightarrow{\Psi,\alpha} \nu\widetilde{n}.F \quad bv(\alpha) \cap fv(B) = \{\widetilde{n}\} \cap fn(B) = \emptyset}{A \mid B \xrightarrow{\Phi,\alpha} \nu\widetilde{n}.(F \mid B)}$$

$\Phi = (\sigma \cup \varphi(B)) \blacktriangleright S$ if $\Psi = \sigma \blacktriangleright S,\ dom(B) \cap dom(\sigma) = \emptyset$

Example 3. Let $P = [x = s]\overline{b}\langle c\rangle$. Then $\nu c\, s.(a(x).P \mid \{s/y\}) \xrightarrow{\mathrm{Hc},\ s.(\{s/y\} \blacktriangleright true),a(x)} \nu c\ s.(P \mid \{s/y\})$ and $\nu c\ s.(P \mid \{s/y\}) \xrightarrow{\mathrm{Hc}\ s.(\{s/y\} \blacktriangleright x=s),\nu c.\overline{b}\langle c\rangle} \nu s.\{s/y\}$.

After each symbolic transition, we need to update the relevant trail. Let $\mathcal{D} = \{x_i : U_i\}_{i \in I}$ be compatible with A, and $A \xrightarrow{\Phi,\alpha} A'$ with $bnv(\alpha) \cap fnv(\mathcal{D}) = \emptyset$, then the result of $\mathcal{D}$ updated by this transition is defined thus:

$$\mathcal{X}(\alpha, dom(A), \mathcal{D}) \triangleq \begin{cases} \mathcal{D} \cup \{x : dom(A)\} & \alpha \text{ is } u(x) \\ \{x_i : (U_i \cup \{c\})\}_{i \in I} & \alpha \text{ is } \nu c.\,\overline{u}\langle c\rangle \\ \mathcal{D} & \text{otherwise} \end{cases}$$

It can be shown that $\mathcal{X}(\alpha, dom(A), \mathcal{D})$ is also a trail and compatible with A' [15]. Intuitively, it records the current abstract knowledge (i.e. $dom(A)$) on input and prevents the prior input variables from using the fresh name (i.e. c) yielded by the opening of channel name.

Example 4. For the symbolic transitions in Example 3, we have $\mathcal{X}(a(x), \{y\}, \emptyset) = \{x : \{y\}\}$ and $\mathcal{X}(\nu c.\overline{b}\langle c\rangle, \{y\}, \{x : \{y\}\}) = \{x : \{y, c\}\}$.

$\Gamma(0) = 0 \qquad \Gamma(\{M/x\}) = \{M/x\} \quad \Gamma(u(x).P_r) = \nu\widetilde{n}.u(x).P$, where $\Gamma(P_r) = \nu\widetilde{n}.P$
$\Gamma(!P_r) = !P_r \quad \Gamma(\nu n.A_r) = \nu n.\Gamma(A_r) \quad \Gamma(\overline{u}\langle N\rangle.P_r) = \nu\widetilde{n}.\overline{u}\langle N\rangle.P$, where $\Gamma(P_r) = \nu\widetilde{n}.P$
$\Gamma(\nu x.A_r) = \Gamma(A_r)_{\setminus x} \quad \Gamma(\mathit{if}\ M = N\ \mathit{then}\ P_r\ \mathit{else}\ Q_r) = \nu\widetilde{n}.\nu\widetilde{m}.\mathit{if}\ M = N\ \mathit{then}\ P\ \mathit{else}\ Q$
where $\Gamma(P_r) = \nu\widetilde{n}.P, \Gamma(Q_r) = \nu\widetilde{m}.Q$
$\Gamma(A_r \mid B_r) = \nu\widetilde{n}.\nu\widetilde{m}.(F \mid G)(\varphi(F) \cup \varphi(G))^*$, where $\Gamma(A_r) = \nu\widetilde{n}.F, \Gamma(B_r) = \nu\widetilde{m}.G$
where $\Gamma(A_r)_{\setminus x}$ is obtained by replacing $\{M/x\}$ in $\Gamma(A_r)$ to 0

Fig. 1. Transformation Γ

Weak symbolic transitions $\stackrel{\Phi,\gamma}{\Longrightarrow}$ (γ is α or ϵ) are generated by absorbing τ transitions as usual. We write $\stackrel{\Phi,\widehat{\alpha}}{\Longrightarrow}$ to mean $\stackrel{\Phi,\alpha}{\Longrightarrow}$ if α is not τ and $\stackrel{\Phi,\epsilon}{\Longrightarrow}$ otherwise.

To capture observational equivalence in applied pi-calculus we also need a means to compare the environmental knowledge exposed by agents:

Definition 2 (Symbolic Static Equivalence). *Let A, B be agents with $\varphi(A) = \sum_{i\in I} \nu\widetilde{n}_i.\sigma_i$ and $\varphi(B) = \sum_{j\in J} \nu\widetilde{m}_j.\theta_j$. We write $A \sim^{(\mathcal{D},\Phi)} B$ if*

1. *$\mathcal{D}$ is compatible with A and B*
2. *$dom(A) = dom(B)$*
3. *for some fresh $x, y \in \mathcal{V}_b$, it holds that $\Phi \models_\varepsilon (\bigvee_{i\in I} \Phi_i) \Leftrightarrow (\bigvee_{j\in J} \Psi_j)$, where $\Phi_i = \mathrm{H}\widetilde{n}_i.(\sigma_i \blacktriangleright x = y), \Psi_j = \mathrm{H}\widetilde{m}_j.(\theta_j \blacktriangleright x = y), i \in I, j \in J$ and $\mathcal{E} = \mathcal{D} \cup \{x : dom(A)\} \cup \{y : dom(B)\}$.*

Definition 3 (Symbolic Bisimilarity). *$\{\approx^{(\mathcal{D},\Phi)} \mid (\mathcal{D}, \Phi)$ a constraint$\}$ is the largest family of symmetric relations on agents such that whenever $A \approx^{(\mathcal{D},\Phi)} B$ then*

1. *$A \sim^{(\mathcal{D},\Phi)} B$*
2. *if $A \xrightarrow{\Phi_1,\alpha} A'$ with $bnv(\alpha) \cap fnv(A, B, \Phi, \mathcal{D}) = \emptyset$, let $\mathcal{F} = \mathcal{X}(\alpha, dom(A), \mathcal{D})$, then there is a partition Σ of $\Phi \wedge \Phi_1$ under $\mathcal{F}$, such that for any $\Psi \in \Sigma$ there are Φ_2, β, B_1 satisfying $B \stackrel{\Phi_2,\widehat{\beta}}{\Longrightarrow} B_1$, $\Psi \models_\varepsilon [\alpha = \beta] \wedge \Phi_2$ and $A_1 \approx^{(\mathcal{E},\Psi)} B_1$.*

To relate symbolic bisimulation to observational equivalence, which is defined on extended processes in the previous section, we employ the function Γ, as defined in Fig. 1, to turn extended processes into an agent, by pulling name binders to the top level, applying active substitutions and eliminating variable restrictions. For example, $\Gamma(\nu x.(\overline{a}\langle x\rangle.\nu n.\overline{a}\langle n\rangle \mid \nu k.\{k/x\})) = \nu n.\nu k.(\overline{a}\langle k\rangle.\overline{a}\langle n\rangle \mid 0)$. The soundness and completeness of symbolic bisimulation w.r.t. observational equivalence was shown in [15]:

Theorem 1. *Let A_r, B_r be closed extended processes. Then $A_r \approx B_r$ iff $\Gamma(A_r) \approx^{(\emptyset,\, true)} \Gamma(B_r)$.*

This result was shown in [15] for the applied pi calculus without choice operator +. As explained before, the choice operator is used in the current work only for the sake of axiomatization. When starting from a +-free agent, the semantic constructions defined so far do not introduce this operator. Hence the theorem also holds here.

4 Proof System

This section is devoted to presenting a proof system for symbolic bisimulation and proving its soundness and completeness. The following discussion is confined to the finite fragment of the calculi, namely the fragment which does not contain replications. Our proof system can be viewed as a general extension of the previous works [10, 16, 3, 12, 14]

The statements of the proof system are of the form $(\mathcal{D}, \Phi) \rhd A = B$. The proof system consists of axioms and inference rules. The axioms are shown in Fig. 2. Apart from those familiar axioms from CCS and pi-calculus, we have **Es** to distribute active substitutions over summation.

The inference rules are presented in Fig. 3. Different from the proof systems for value-passing CCS [12] or pi-calculus [3, 14], the basic entities are of the form $\nu\widetilde{n}.(P \mid \sigma)$, where P is a plain process, rather than just P. The main reason is that the evaluation of the equality tests occurred in P may depend on the knowledge exposed by frame $\nu\widetilde{n}.\sigma$. This will be further explained later. In **Par**, the side conditions ensure that the trail $\mathcal{E}$ is compatible with the agents in the derived equation $(\mathcal{E}, \Phi) \rhd A \mid C = B \mid C$. Rule **Frame** relates frames which are symbolically static equivalent, namely they expose the same knowledge to the environment. The equation $x = y$ in formula $\Phi_i = \mathrm{H}\widetilde{n}.(\sigma_i \rhd x = y)$ abstractly represents the set of tests $\{\, M = N \mid var(M, N) \subseteq dom(\sigma_i) \,\}$. Φ_i holds means these tests can be satisfied under the knowledge exposed by $\nu\widetilde{n}_i.\sigma_i$. When $\bigvee \Phi_i$ is equavelant to $\bigvee \Psi_j$, the frames $\sum_i \nu\widetilde{n}_i.\sigma_i$ and $\sum_j \nu\widetilde{m}_j.\theta_j$ expose the same knowledge. In **Outt**, the active substitutions $\{M/y\}$ and $\{N/x\}$ in the premise are eliminated when output prefixes are introduced. This reflects the fact that active substitutions are generated by output transitions: $S\overline{u}\langle M\rangle.P \xrightarrow{S,\, \nu x.\overline{u}\langle x\rangle} P \mid \{M/x\}$. The rule **Partition** permits a case analysis on formula Φ.

The proof system is designed to reason about agent equivalence and has to inevitably rely on some form of reasoning on the underlying equational theories on terms, which are taken as parameters to the applied pi-calculus. We have decided to factor out reasoning on terms and substitutions from the proof system, using "semantical judgments" of the form $\Phi \models_{\mathcal{D}} \Psi$, as can be seen in **Frame**, **Input**, **Outt**, **Outch**, and **Partition**. One may think of these as questions about the term domain, to be answered by an "oracle". In practice they can be resolved by invoking some decision procedures, as the one in [2] for instance, or appealing to a separate proof system specially designed for the underlying equational theories. The following lemma is easy to prove (using **Guard**):

Lemma 1. *Assume $\mathcal{D}$ is compatible with $\nu\widetilde{n}.(S\pi.P \mid \sigma)$.*

1. *If $\Phi \models_{\mathcal{D}} \mathrm{H}\widetilde{n}.(\sigma \blacktriangleright S)$ then $\vdash (\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(S\pi.P \mid \sigma) = \nu\widetilde{n}.(\pi.P \mid \sigma)$.*
2. *If $\Phi \wedge \mathrm{H}\widetilde{n}.(\sigma \blacktriangleright S) \models_{\mathcal{D}} false$ then $\vdash (\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(S\pi.P \mid \sigma) = \nu\widetilde{n}.\sigma$.*

Example 5. Assuming $dec(enc(w_1, w_2), w_2) =_E w_1$, let us prove:

$$(\emptyset, true) \rhd \nu s.\overline{a}\langle s\rangle.a(x).[x = s]\overline{b}\langle c\rangle = \nu k.\overline{a}\langle enc(m, k)\rangle.a(x).[dec(x, k) = m]\overline{b}\langle c\rangle.$$

By OUTT, it suffices to derive

$$(\emptyset, true) \rhd \nu s.(a(x).[x = s]\overline{b}\langle c\rangle \mid \{s/y\}) = \nu k.(a(x).[dec(x,k) = m]\overline{b}\langle c\rangle \mid \{enc(m,k)/y\})$$

Invoking INPUT and PARTITION leads to the following two statements:
$(\mathcal{D}, x = y) \rhd \nu s.([x = s]\overline{b}\langle c\rangle \mid \{s/y\}) = \nu k.([dec(x,k) = m]\overline{b}\langle c\rangle \mid \{enc(m,k)/y\})$
$(\mathcal{D}, x \neq y) \rhd \nu s.([x = s]\overline{b}\langle c\rangle \mid \{s/y\}) = \nu k.([dec(x,k) = m]\overline{b}\langle c\rangle \mid \{enc(m,k)/y\})$
where $\mathcal{D} = \{x : \{y\}\}$.

We continue with the first one and the other is similar. From Example 2, we know that $(x = y) \models_{\mathcal{D}} \mathrm{H}s.(\{s/y\} \blacktriangleright x = s) \wedge \mathrm{H}k.(\{enc(m,k)/y\} \blacktriangleright dec(x,k) = m)$, hence by Lemma 1 this statement can be reduced to

$$(\mathcal{D}, x = y) \rhd \nu s.(\overline{b}\langle c\rangle \mid \{s/y\}) = \nu k.(\overline{b}\langle c\rangle \mid \{enc(m,k)/y\}).$$

Applying OUTCH, we are left to show that

$$(\mathcal{D}, x = y) \rhd \nu s.\{s/y\} = \nu k.\{enc(m,k)/y\}.$$

By FRAME, this leads to the "semantical judgement" $(x = y) \models_{\mathcal{D}} \mathrm{H}s.(\{s/y\} \blacktriangleright z_1 = z_2) \Leftrightarrow \mathrm{H}k.(\{enc(m,k)/y\} \blacktriangleright z_1 = z_2)$, which can be easily verified by the algorithm developed in [2], for instance.

As shown in this example, we can derive $(\emptyset, true) \rhd \nu s.(a(x).[x = s]\overline{b}\langle c\rangle \mid \{s/y\}) = \nu k.(a(x).[dec(x,k) = m]\overline{b}\langle c\rangle \mid \{enc(m,k)/y\})$; However, we cannot derive $(\emptyset, true) \rhd a(x).[x = s]\overline{b}\langle c\rangle = a(x).[dec(x,k) = m]\overline{b}\langle c\rangle$, because the equality tests $[x = s]$ and $[dec(x,k) = m]$ cannot be satisfied at the same time without the knowledge exposed by the frames $\nu s.\{s/y\}$ and $\nu k.\{enc(m,k)/y\}$. This explains why the basic entities of the proof system are agents of the form $\nu\widetilde{n}.(P \mid \sigma)$, which are plain processes equipped with frames, not just plain processes.

Since weak bisimilarity is not preserved by summation, we need to introduce a refined equivalence which takes care of initial τ moves. The equivalence is defined on top of weak bisimilarity as follows:

Definition 4 (Symbolic Congruence). $\{\cong^{(\mathcal{D},\Phi)} \mid (\mathcal{D}, \Phi)$ *a constraint* $\}$ *is the largest family of symmetric relations between agents and whenever* $A \cong^{(\mathcal{D},\Phi)} B$,

1. $A \sim^{(\mathcal{D},\Phi)} B$
2. *if* $A \xrightarrow{\Phi_1, \alpha} A'$ *with* $bnv(\alpha) \cap fnv(A, B, \Phi, \mathcal{D}) = \emptyset$, *let* $\mathcal{E} = \mathcal{X}(\alpha, dom(A), \mathcal{D})$, *then there is a partition* Σ *of* $\Phi \wedge \Phi_1$ *under* $\mathcal{E}$, *such that for any* $\Psi \in \Sigma$ *there are* Φ_2, β, B_1 *satisfying* $B \overset{\Phi_2, \beta}{\Longrightarrow} B_1$, $\Psi \models_{\mathcal{E}} [\alpha = \beta] \wedge \Phi_2$ *and* $A_1 \approx^{(\mathcal{E},\Psi)} B_1$.

Theorem 2 (Soundness). *If* $\vdash (\mathcal{D}, \Phi) \rhd A = B$ *then* $A \cong^{(\mathcal{D},\Phi)} B$.

Soundness ensures correctness of the proof system. It is easy to see that $\cong^{(\mathcal{D},\Phi)} \subseteq \approx^{(\mathcal{D},\Phi)}$. Combining with Theorem 1, we know that the proof system is sound w.r.t observational equivalence.

Now we turn to completeness. Since the rule **Partition** can only be used finitely many times in a proof, to capture $A \cong^{(\mathcal{D},\Phi)} B$ by purely syntactical inferencing requires the partitions in Def. 3 and Def. 4 must be finite. It has been shown that in the case of value-passing CCS and pi-calculus, such finite partitions always exist for processes whose symbolic transition graphs are finite

P1	$A = A \mid 0$	**S1**	$A + 0 = A$
P2	$A \mid B = B \mid A$	**S2**	$A + A = A$
P3	$(A \mid B) \mid C = A \mid (B \mid C)$	**S3**	$A + B = B + A$
		S4	$(A + B) + C = A + (B + C)$
R1	$\nu n.A = A$ if $n \notin fn(A)$	**T1**	$\pi.\tau.P = \pi.P$
R2	$\nu n.\nu m.A = \nu m.\nu n.A$	**T2**	$P + \tau.P = \tau.P$
R3	$\nu n.(S\pi.P \mid \sigma) = \nu n.\sigma$ if $n \in sub(\pi)$	**T3**	$\pi.(P + \tau.Q) + \pi.Q = \pi.(P + \tau.Q)$
Er	$\nu n.(A + B) = \nu n.A + \nu n.B$	**Es**	$(A + B) \mid \sigma = (A \mid \sigma) + (B \mid \sigma)$

Ep Let $P = \sum_{i\in I} S_i\pi_i.P_i$ and $Q = \sum_{j\in J} T_j\pi'_j.Q_j$
with $bnv(\pi_i) \cap fnv(Q) = bnv(\pi'_j) \cap fnv(P) = \emptyset$.

$$P \mid Q = \sum_{i\in I} S_i\pi_i.(P_i \mid Q) + \sum_{j\in J} T_j\pi'_j.(P \mid Q_j) + \sum_{\pi_i \ opp \ \pi'_j} S_i \wedge T_j \wedge (u_i = v_j)\tau.R_{ij}$$

where π_i *opp* π'_j and R_{ij} are defined as follows
1. $\pi_i = u_i(x)$, $\pi'_j = \overline{v}_j\langle M\rangle$ with M, x the same sort; then $R_{ij} = P_i\{M/x\} \mid Q_j$
2. The converse of the above clause;

where subject of prefix π is $sub(\tau) = \emptyset$ and $sub(u(x)) = sub(\overline{u}\langle M\rangle) = \{u\}$.

Fig. 2. The Axioms

[11, 14]. However, the situation is less clear in the applied pi-calculus, since here we have to consider not only dynamic behaviors of processes but also static equivalence of knowledge, i.e. $\sim^{(\mathcal{D},\Phi)}$, which depends on the expressiveness of the constraint systems [11, 4, 13]. Let us say that a class of agents *admit finite partition* if symbolic equivalences on them can be established when the phrase "there exists a partition Σ" is replaced by "there exists a finite partition Σ" in Def. 3 and Def. 4. The completeness of the proof system holds on agents that admit finite partition. In next section we will demonstrate that a widely used fragment of applied pi-calculus admits finite partition, hence this restriction is acceptable in practical applications. In what follows all agents are assumed to admit finite partition.

The following lemma "lifts" $\approx^{(\mathcal{D},\Phi)}$ to $\cong^{(\mathcal{D},\Phi)}$:

Lemma 2 (Lifting). *$\nu\widetilde{n}.(P \mid \sigma) \approx^{(\mathcal{D},\Phi)} \nu\widetilde{m}.(Q \mid \theta)$ iff there exists a finite partition Σ of Φ under $\mathcal{D}$ such that for any $\Psi \in \Sigma$ we have either $\nu\widetilde{n}.(P \mid \sigma) \cong^{(\mathcal{D},\Psi)} \nu\widetilde{m}.(\tau.Q \mid \theta)$, or $\nu\widetilde{n}.(\tau.P \mid \sigma) \cong^{(\mathcal{D},\Psi)} \nu\widetilde{m}.(Q \mid \theta)$, or $\nu\widetilde{n}.(P \mid \sigma) \cong^{(\mathcal{D},\Psi)} \nu\widetilde{m}.(Q \mid \theta)$.*

Definition 5 (Normal Forms).

- *Agent A is in* head normal form *if $A = \sum_i \nu\widetilde{n}_i.(P_i \mid \sigma_i)$ with each $P_i = S_i\pi_i.Q_i$ or 0 and $sub(\pi_i) \cap \{\widetilde{n}_i\} = \emptyset$.*
- *A head normal form A is a* full normal form *if $A \stackrel{\Phi_1,\epsilon}{\Longrightarrow} \xrightarrow{\Phi_2,\alpha} A'$ implies $A \xrightarrow{\Phi,\alpha} A'$ with $(\Phi_1 \wedge \Phi_2) \Leftrightarrow \Phi$.*

The *height* of an agent A, $|\,A\,|$, is defined inductively thus: $|\,0\,| = |\,\{M/x\}\,| = 0$, $|\,S\pi.P\,| = |\,P\,| + 1$, $|\,A \mid B\,| = |\,A\,| + |\,B\,|$, $|\,A + B\,| = \max(|\,A\,|, |\,B\,|)$ and

Axiom $\dfrac{}{(\mathcal{D}, true) \rhd A = B}$ $A = B$ is an axiom and $\mathcal{D}$ is compatible with A, B

Equiv $\dfrac{(\mathcal{D}, \Phi) \rhd A = B}{(\mathcal{D}, \Phi) \rhd B = A} \qquad \dfrac{(\mathcal{D}, \Phi) \rhd A = B,\ B = C}{(\mathcal{D}, \Phi) \rhd A = C}$

Par $\dfrac{(\mathcal{D}, true) \rhd A = B}{(\mathcal{E}, true) \rhd A \mid C = B \mid C}$
$\mathcal{E} = \{x_i : U_i\}_i$ is compatible with $A \mid C$ and $B \mid C$
$\mathcal{D} = \{x_i : (U_i - dom(C))\}_i$

Frame $\dfrac{\Phi \models_{\mathcal{E}} (\bigvee_i \Phi_i) \Leftrightarrow (\bigvee_j \Psi_j)}{(\mathcal{D}, \Phi) \rhd \sum_i \nu\widetilde{n}_i.\sigma_i = \sum_j \nu\widetilde{m}_j.\theta_j}$
$\mathcal{D}$ is compatible with each σ_i and θ_j
$dom(\sigma_i) = dom(\theta_j)$, $\Phi_i = \mathrm{H}\widetilde{n}_i.(\sigma_i \blacktriangleright (x = y))$, $\Psi_j = \mathrm{H}\widetilde{m}_j.(\theta_j \blacktriangleright (x = y))$
$\mathcal{E} = \mathcal{D} \cup \{\, x : dom(\sigma_i),\ y : dom(\sigma_i) \,\}$ for some fresh $x, y \in \mathcal{V}_b$.

Tau $\dfrac{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(P \mid \sigma) = \nu\widetilde{m}.(Q \mid \theta)}{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(\tau.P \mid \sigma) = \nu\widetilde{m}.(\tau.Q \mid \theta)}$

Input $\dfrac{(\mathcal{E}, \Phi) \rhd \sum_i \nu\widetilde{n}_i.(\tau.P_i \mid \sigma_i) = \sum_j \nu\widetilde{m}_j.(\tau.Q_j \mid \theta_j),\ \Phi \models_{\mathcal{D}} \bigwedge_{i,j} u_i = v_j}{(\mathcal{D}, \Phi) \rhd \sum_i \nu\widetilde{n}_i.(u_i(x).P_i \mid \sigma_i) = \sum_j \nu\widetilde{m}_j.(v_j(x).Q_j \mid \theta_j)}$
$\mathcal{E} = \mathcal{D} \uplus \{x : dom(\sigma_i)\}$, $x \notin fv(\mathcal{D}, \Phi, \{\sigma_i, \theta_j\}_{i,j})$
$var(\{u_i, v_j\}_{i,j}) \subseteq dom(\mathcal{D})$, $name(\{u_i, v_j\}_{i,j}) \cap \{\widetilde{n}_i, \widetilde{m}_j\}_{i,j} = \emptyset$

Guard $\dfrac{(\mathcal{D}, \Phi \wedge \mathrm{H}\widetilde{n}.(\sigma \blacktriangleright S)) \rhd \nu\widetilde{n}.(\pi.P \mid \sigma) = A,\ (\mathcal{D}, \Phi \wedge \mathrm{H}\widetilde{n}.(\sigma \blacktriangleright \neg S)) \rhd \nu\widetilde{n}.\sigma = A}{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(S\pi.P \mid \sigma) = A}$
$var(S) \subseteq dom(\mathcal{D})$

Outt $\dfrac{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(P \mid \sigma \mid \{M/y\}) = \nu\widetilde{m}.(Q \mid \theta \mid \{N/y\}),\ \Phi \models_{\mathcal{D}} u = v}{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(\overline{u}\langle M\rangle.P \mid \sigma) = \nu\widetilde{m}.(\overline{v}\langle N\rangle.Q \mid \theta)}$
$var(u, v) \subseteq dom(\mathcal{D})$, $name(u, v) \cap \{\widetilde{n}, \widetilde{m}\} = \emptyset$ and $y \notin fv(\mathcal{D}, \Phi)$

Outch $\dfrac{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(P \mid \sigma) = \nu\widetilde{m}.(Q \mid \theta),\ \Phi \models_{\mathcal{D}} [\overline{u}\langle w\rangle = \overline{v}\langle w'\rangle]}{(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}.(\overline{u}\langle w\rangle.P \mid \sigma) = \nu\widetilde{m}.(\overline{v}\langle w'\rangle.Q \mid \theta)}$
$var(\overline{u}\langle w\rangle, \overline{v}\langle w'\rangle) \subseteq dom(\mathcal{D})$, $name(\overline{u}\langle w\rangle, \overline{v}\langle w'\rangle) \cap \{\widetilde{n}, \widetilde{m}\} = \emptyset$

Sum $\dfrac{(\mathcal{D}, \Phi) \rhd A_i = B_i,\ i = 1, 2}{(\mathcal{D}, \Phi) \rhd A_1 + A_2 = B_1 + B_2}$

Res $\dfrac{(\mathcal{D}, \Phi) \rhd A = B}{(\mathcal{E}, \mathrm{H}n.\Phi) \rhd \nu n.A = \nu n.B}$
$n \notin fn(\Phi)$ if $n \notin name(\mathcal{D})$
$\mathcal{E} = \{x_i : (U_i - \{n\})\}_i$ if $\mathcal{D} = \{x_i : U_i\}_i$

Partition $\dfrac{(\mathcal{D}, \Phi_i) \rhd A = B,\ i = 1, 2,\ \Phi \models_{\mathcal{D}} \Phi_1 \vee \Phi_2}{(\mathcal{D}, \Phi) \rhd A = B}$

Absurd $\dfrac{}{(\mathcal{D}, false) \rhd A = B}$ $\mathcal{D}$ is compatible with A, B

Fig. 3. The Inference Rules

$|\,\nu n.A\,| = |\,A\,|$. Every agent can be rewritten to head normal form and full normal form without increasing its height, as stated in the lemma below.

Lemma 3. *For any agent A, there is a head/full normal form B such that, for any $\mathcal{D}$ compatible with A,*

1. $\vdash (\mathcal{D}, true) \rhd A = B$, $|\,B\,| \leq |\,A\,|$, $fnv(B) \subseteq fnv(A)$ *and* $dom(B) = dom(A)$
2. *if* $A = \nu\widetilde{n}.F$ *then* B *has the form* $\sum_{i\in I} \nu\widetilde{n}.(S_i\pi_i.P_i \mid \varphi(F))$.

For example, let $A = \nu s.[((\overline{a}\langle s\rangle \mid a(x)) + \tau.\overline{a}\langle c\rangle) \mid \{s/y\}]$. Then we can deduce a full normal form for A by the following sequence:

$$\begin{aligned}
(\emptyset, true) \rhd A &= \nu s.((\overline{a}\langle s\rangle.a(x) + a(x).\overline{a}\langle s\rangle + \tau + \tau.\overline{a}\langle c\rangle) \mid \{s/y\}) && \textbf{Ep}\\
&= \nu s.((\overline{a}\langle s\rangle.a(x) + a(x).\overline{a}\langle s\rangle + \tau + \tau.\overline{a}\langle c\rangle + \overline{a}\langle c\rangle) \mid \{s/y\}) && \textbf{T2}\\
&= \nu s.(\overline{a}\langle s\rangle.a(x) \mid \{s/y\}) + \nu s.(a(x).\overline{a}\langle s\rangle \mid \{s/y\})\\
&\quad + \nu s.(\tau \mid \{s/y\}) + \nu s.(\tau.\overline{a}\langle c\rangle \mid \{s/y\}) + \nu s.(\overline{a}\langle c\rangle \mid \{s/y\}) && \textbf{Er}, \textbf{Es}
\end{aligned}$$

Theorem 3 (Completeness). *If $A \cong^{(\mathcal{D},\Phi)} B$ then $\vdash (\mathcal{D}, \Phi) \rhd A = B$.*

Proof. The proof proceeds by induction on the joint height $|\,A\,| + |\,B\,|$. By Lemma 3, we rewrite A and B to full normal form $\sum_{i\in I} \nu\widetilde{n}_i.(\widehat{P}_i \mid \sigma_i)$ and $\sum_{j\in J} \nu\widetilde{m}_j.(\widehat{Q}_i \mid \theta_j)$ respectively, where $\widehat{P}_i = S_i\pi_i.P_i$ or 0 and $\widehat{Q}_j = T_j\pi'_j.Q_j$ or 0. We group the summands of A according to type γ of π_i and write A_γ for the result. It suffices to show that $(\mathcal{D}, \Phi) \rhd A_\gamma + B = B$ and $(\mathcal{D}, \Phi) \rhd B_\gamma + A = A$ for each γ.

We only sketch the proof for the case $\gamma = \tau$. Assume $A_\tau = \sum_i \nu\widetilde{n}_i.(S_i\tau.P_i \mid \sigma_i)$ and $B_\tau = \sum_j \nu\widetilde{m}_j.(T_j\tau.Q_j \mid \theta_j)$. Then $A \xrightarrow{\Phi_i,\tau} A_i \equiv_s \nu\widetilde{n}_i.(P_i \mid \sigma_i)$ where $\Phi_i = \mathrm{H}\widetilde{n}_i.(\sigma_i \blacktriangleright S_i)$. Since $A \cong^{(\mathcal{D},\Phi)} B$, there is a finite partition Σ of $\Phi \wedge \Phi_i$ under $\mathcal{D}$, and for each $\Psi \in \Sigma$, there exist $B \xrightarrow{\Psi_j,\tau} B_j \equiv_s \nu\widetilde{m}_j.(Q_j \mid \theta_j)$, such that $\Psi \models_{\mathcal{D}} \Psi_j$ and $A_i \approx^{(\mathcal{D},\Psi)} B_j$. By Theorem 2, induction hypothesis, **T2** and **Partition**, we can derive $(\mathcal{D}, \Psi) \rhd \nu\widetilde{n}_i.(\tau.P_i \mid \sigma_i) = \nu\widetilde{m}_j.(\tau.Q_j \mid \theta_j)$. By Lemma 1.*1*, we have $(\mathcal{D}, \Psi) \rhd \nu\widetilde{n}_i.(S_i\tau.P_i \mid \sigma_i) = \nu\widetilde{m}_j.(T_j\tau.Q_j \mid \theta_j)$, and hence $(\mathcal{D}, \Psi) \rhd \nu\widetilde{n}_i.(S_i\tau.P_i \mid \sigma_i) + B = \nu\widetilde{m}_j.(T_j\tau.Q_j \mid \theta_j) + B = B$. By **Partition**, we obtain $(\mathcal{D}, \Phi \wedge \Phi_i) \rhd \nu\widetilde{n}_i.(S_i\tau.P_i \mid \sigma_i) + B = B$. By Lemma 1.*2*, we have $(\mathcal{D}, \Phi \wedge \neg\Phi_i) \rhd \nu\widetilde{n}_i.(S_i\tau.P_i \mid \sigma_i) = \nu\widetilde{n}_i.\sigma_i$ since we can deduce that $\neg\Phi_i \wedge \mathrm{H}\widetilde{n}.(\sigma_i \blacktriangleright S_i) \models_{\mathcal{D}} \mathit{false}$. Adding B to both sides we have $(\mathcal{D}, \Phi \wedge \neg\Phi_i) \rhd \nu\widetilde{n}_i.(S_i\tau.P_i \mid \sigma_i) + B = B$. By **Partition** again, $(\mathcal{D}, \Phi) \rhd \nu\widetilde{n}_i.(S_i\tau.P_i \mid \sigma_i) + B = B$. Finally, by **Sum** we obtain $(\mathcal{D}, \Phi) \rhd A_\tau + B = B$. Similarly we can derive $(\mathcal{D}, \Phi) \rhd B_\tau + A = A$.

For any extended processes A_r, it is easy to see that $\Gamma(A_r) \equiv_s \nu\widetilde{n}.(P \mid \sigma)$ for some $\widetilde{n}, P, \sigma$. The following theorem is a direct corollary of Theorem 1, 2 and 3, using Lemma 2 and axiom **T1**.

Theorem 4. *Let A_r, B_r be closed extended processes and $\Gamma(A_r) \equiv_s \nu\widetilde{n}.(P \mid \sigma)$ and $\Gamma(B_r) \equiv_s \nu\widetilde{m}.(Q \mid \theta)$. Then $A_r \approx B_r$ iff $\vdash (\emptyset, true) \rhd \nu\widetilde{n}.(\tau.P \mid \sigma) = \nu\widetilde{m}.(\tau.Q \mid \theta)$.*

Thus our proof system is sound and complete w.r.t. observational equivalence for finite extended processes which admit finite partition.

5 Finiteness of Partition

In practice we do not always need full applied pi-calculus for describing and analyzing security protocols. For example, as argued in [7, 8], it is generally assumed that all communications are controlled by the attacker thus private channels between processes are not accurate. We shall show in this section that a useful fragment of the applied pi-calculus, called "simple processes" [7, 8], admit finite partition. Simple processes are built up from "basic processes". A basic process represents a session of protocol role which knows exactly what to do next. Simple processes are used to analyze security protocols whose roles have a deterministic behavior, such as the protocols in [6]. For simple processes without ELSE branch nor replications, it is shown in [8] that symbolic trace equivalence coincides with observational equivalence. In comparison, we use symbolic bisimilarity to fully capture observational equivalence, and we will show that finite partitions are sufficient for simple processes, even in the presence of ELSE branch and replications.

The sets of *basic processes* $\mathcal{B}(c, U)$ with $c \in \mathcal{N}_{ch}$ and finite $U \subset \mathcal{V}_b$ are the least sets of processes such that

1. $0 \in \mathcal{B}(c, U)$
2. if $B \in \mathcal{B}(c, U)$, $var(M) \subseteq U$ and $name(M) \subset \mathcal{N}_b$ then $\overline{c}\langle M\rangle.B \in \mathcal{B}(c, U)$
3. if $B_1, B_2 \in \mathcal{B}(c, U)$, $var(M, N) \subseteq U$ and $name(M, N) \subset \mathcal{N}_b$, then *if* $M = N$ *then* B_1 *else* $B_2 \in \mathcal{B}(c, U)$
4. if $B \in \mathcal{B}(c, U \uplus \{x\})$ and $x \in \mathcal{V}_b$, then $c(x).B \in \mathcal{B}(c, U)$.

Let us abbreviate $A_1 \mid A_2 \mid \cdots \mid A_m$ to $\prod_{i\in I} A_i$. Then *simple processes* are those of the form

$$\nu\widetilde{n}.(\prod\nolimits_{i\in I} \nu\widetilde{n}_i.(B_i \mid \sigma_i) \mid \prod\nolimits_{j\in J} !(\nu c_j, \widetilde{m}_j).\overline{b}_j\langle c_j\rangle.B'_j)$$

where $B_i \in \mathcal{B}(a_i, \emptyset)$, $B'_j \in \mathcal{B}(c_j, \emptyset)$; a_i, b_j with $i \in I, j \in J$ are pairwise-distinct channel names. As argued in [8], the pairwise-distinct channel names for each basic process correspond to the fact that the attacker is able to schedule the messages and know which process the message comes from (*e.g.* via IP addresses).

To cater simple processes, in symbolic semantics it is adequate to consider *simple agents* of the form

$$A \equiv_s \nu\widetilde{n}.(\prod\nolimits_{i\in I} B_i \mid \prod\nolimits_{j\in J} !(\nu c_j, \widetilde{m}_j).\overline{b}_j\langle c_j\rangle.B'_j \mid \sigma)$$

where $B_i \in \mathcal{B}(a_i, U_i), B'_j \in \mathcal{B}(c_j, \emptyset)$; U_i with $i \in I$ are pairwise-distinct and σ is idempotent with $dom(\sigma) \cap \bigcup_{i\in I} U_i = \emptyset$; a_i, b_j with $i \in I, j \in J$ are pairwise-distinct channel names. In what follows we shall use A to range over simple agents.

For simple agents we do not have to use the rule $!P_r \xrightarrow{true,\tau} \nu\widetilde{m}.(P \mid !P_r)$, where $\Gamma(P_r) = \nu\widetilde{m}.P$, to expand replications since replications in simple agents

are always guarded by bound output. Instead, we can use the following simpler rule (conflicts on $\widetilde{m}$ can be avoided by α-conversion):

$$!(\nu c, \widetilde{m}).\overline{b}\langle c\rangle.B \xrightarrow{true,\nu c.\overline{b}\langle c\rangle} \nu\widetilde{m}.(B \mid !(\nu c, \widetilde{m}).\overline{b}\langle c\rangle.B).$$

We still use $\approx^{(\mathcal{D},\Phi)}$ to denote the symbolic bisimilarity in the resulting symbolic semantics. Note that the definition of simple agents is closed under $\xrightarrow{\Phi,\alpha}$.

Theorem 5. *For simple agents A and B, $A \approx^{(\mathcal{D},\Phi)} B$ can be established when the phrase "there exists a partition Σ" is replaced by "there exists a finite partition Σ" in Definition 3.*

Proof. To show finite partition always suffices, assume $A \approx^{(\mathcal{D},\Phi)} B$ and $A \xrightarrow{\Phi_1,\alpha} A_1$. Let $\Sigma_\alpha = \{\, \Phi \wedge \Phi_1 \wedge \Psi_1 \mid B \overset{\Psi_1,\widehat{\alpha}}{\Longrightarrow} B_1 \not\xrightarrow{\tau} \,\}$, where $B_1 \not\xrightarrow{\tau}$ denotes that there is no Ψ', B_1' such that $B_1 \xrightarrow{\Psi',\tau} B_1'$. We can verify that Σ_α is a finite partition of $\Phi \wedge \Phi_1$ under $\mathcal{E}$, and $A_1 \approx^{(\mathcal{E},\Phi\wedge\Phi_1\wedge\Psi_1)} B_1$.

Thus, by Theorem 4, our proof system is sound and complete for observational equivalence on finite fragment of simple processes.

6 Conclusions

We have presented a proof system for observational equivalence in the applied pi-calculus, and shown its soundness and completeness. The completeness result is obtained via a recently developed theory of symbolic bisimulation which exactly captures observational equivalence. This is the first inference system for the applied pi calculus which makes it possible to reason on security properties by syntactic manipulations.

As the applied pi-calculus is parameterized on equational theories for cryptographic operations while our proof system mainly concerns with behavioural properties of processes, "static" reasoning about cryptographic operations has been factored out from the proof system, as "semantic judgements" of the form $\Phi \models_{\mathcal{D}} \Psi$. The verification of $\Phi \models_{\mathcal{D}} \Psi$ is a second order E-unification problem. The reasoning about some special class of the problem is discussed in [2], where sound and complete transformation rules are proposed to handle the constraint systems without negation for convergent equational theories, and a decision procedure for convergent subterm theories. The ongoing work of [5] mainly dedicates to finding a simpler decision algorithm than [2] for a larger class of equational theories in the presence of negation.

Our completeness result is confined to finite processes which admit finite partition. This contrasts to the proof systems for value-passing CCS and pi-calculus, where finite partitions are sufficient for finite processes. The expressiveness of formulas is highly relevant in this regard. The formula language in this paper includes two operators $\sigma \blacktriangleright \Phi$ and $\mathrm{H}n.\Phi$, which are mainly needed for symbolic output transitions: $\nu k.(\mathit{if}\ x = k\ \mathit{then}\ P\ \mathit{else}\ Q \mid \{k/y\}) \xrightarrow{\mathrm{H}k.(\{k/y\}\blacktriangleright x=k),true} P$.

When an agent tries to match a symbolic transition from the other, the choices on the branches are closely dependent on symbolic static equivalence. It is still unclear whether the expressiveness of the formulas is sufficient to guarantee finite partitions for symbolic static equivalence, or how to extend the formula language if not.

References

1. M. Abadi and C. Fournet. Mobile values, new names, and secure communication. In *POPL*, pages 104–115, 2001.
2. M. Baudet. Deciding security of protocols against off-line guessing attacks. In *CCS '05: Proceedings of the 12th ACM conference on Computer and communications security*, pages 16–25, New York, NY, USA, 2005. ACM.
3. M Boreale and De Nicola, R. A symbolic semantics for the pi-calculus (extended abstract). In *CONCUR '94: Proceedings of the Concurrency Theory*, pages 299–314, London, UK, 1994. Springer-Verlag.
4. J. Borgström. A complete symbolic bisimilarity for an extended spi calculus. *Electron. Notes Theor. Comput. Sci.*, 242(3), 2009.
5. V. Cheval, H. Comon-Lundh, and S. Delaune. A decision procedure for proving observational equivalence. In Michele Boreale and Steve Kremer, editors, *Preliminary Proceedings of the 7th International Workshop on Security Issues in Coordination Models, Languages and Systems (SecCo'09)*, Bologna, Italy, October 2009.
6. J. Clark and J. Jacob. A survey of authentication protocol literature, 1997. Available at `http://www.cs.york.ac.uk/~jac/papers/drareviewps.ps`.
7. H. Comon-Lundh and V. Cortier. Computational soundness of observational equivalence. In *CCS '08: Proceedings of the 15th ACM conference on Computer and communications security*, pages 109–118, New York, NY, USA, 2008. ACM.
8. V. Cortier and S. Delaune. A method for proving observational equivalence. In *Proceedings of the 22nd IEEE Computer Security Foundations Symposium (CSF'09)*, pages 266–276, Port Jefferson, NY, USA, July 2009. IEEE Computer Society Press.
9. S. Delaune, S. Kremer, and M.D. Ryan. Symbolic bisimulation for the applied pi calculus. In *FSTTCS*, pages 133–145, 2007.
10. M Hennessy. A proof system for communicating processes with value-passing (extended abstract). In *Proceedings of the Ninth Conference on Foundations of Software Technology and Theoretical Computer Science*, pages 325–339, London, UK, 1989. Springer-Verlag.
11. M. Hennessy and H. Lin. Symbolic bisimulations. *Theor. Comput. Sci.*, 138(2):353–389, 1995.
12. M. Hennessy and H. Lin. Proof systems for message-passing process algebras. *Formal Asp. Comput.*, 8(4):379–407, 1996.
13. M. Johansson, B. Victor, and J. Parrow. A fully abstract symbolic semantics for psi-calculi., 2009. Accepted for SOS'09.
14. H. Lin. Complete inference systems for weak bisimulation equivalences in the pi-calculus. *Inf. Comput.*, 180(1):1–29, 2003.
15. J Liu and H Lin. A complete symbolic bisimulation for full applied pi calculus. In *SOFSEM '10: Proceedings of the 36th Conference on Current Trends in Theory and Practice of Computer Science*, pages 552–563, Berlin, Heidelberg, 2010. Springer-Verlag.
16. J. Parrow and D. Sangiorgi. Algebraic theories for name-passing calculi. *Information and Computation*, 120:174–197, 1994.