A Quantitative Approach to Noninterference for Probabilistic Systems

Alessandro Aldini [2,1]

Istituto STI, Università Carlo Bo, Urbino, Italy

Alessandra Di Pierro [3,1]

Dipartimento di Informatica, Università di Pisa, Italy

Abstract

We present a technique for measuring the security of a system which relies on a probabilistic process algebraic formalisation of noninterference. We define a mathematical model for this technique which consists of a linear space of processes and linear transformations on them. In this model the measured quantity corresponds to the norm of a suitably defined linear operator associated to the system. The probabilistic model we adopt is reactive in the sense that processes can react to the environment with a probabilistic choice on a set of inputs; it is also generative in the sense that outputs autonomously chosen by the system are governed by a probability distribution. In this setting, noninterference is formulated in terms of a probabilistic notion of weak bisimulation. We show how the probabilistic information in this notion can be used to estimate the maximal information leakage, i.e. the security degree of a system against a most powerful attacker.

Key words: Probabilistic Noninterference, Process Algebra, Similarity Relation, Weak Bisimulation

1 Introduction

Several characterisations of the security of a system against illegal information flows from a high-security level enclave to a low-security level enclave have been proposed in the literature, which employ qualitative methods based, e.g., on a logical interpretation of the noninterference idea by Goguen and Meseguer [GM82]. This idea ultimately depends on some notion of indistinguishability of process behaviours (see, e.g., [FG95,RS01]). However, in many practical situations the complete absence of any information flow is difficult to be guaranteed and it is more realistic to assume that some amount of information leakage can be tolerated [Gra90,R$^+$01]. In

[1] This work has been partially funded by Progetto MEFISTO (Metodi Formali per la Sicurezza e il Tempo)

[2] Email: aldini@sti.uniurb.it

[3] Email: dipierro@di.unipi.it

This is a preliminary version. The final version will be published in Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

such a scenario security is not an absolute requirement and approximated versions of security properties would be more appropriate. The latter would also allow us to estimate the difference between the real system of interest and an idealised perfectly secure system by means of a measure of the approximation.

On the basis of such considerations, in [A$^+$03] it is shown how a security property termed Probabilistic Noninterference (*PNI*) can be used for establishing the security of a system in an approximated way. *PNI* extends with probabilities a nondeterministic notion of noninterference [FG95], which checks whether, from the viewpoint of a low-level observer, the behaviour of the system is invariant with respect to the behaviour of high-level users interacting with the system. The particular class of high-level users considered by *PNI*, denoted $\mathcal{A}_{PNI}$, can be viewed as a family of probabilistic adversaries that try to set up a covert communication channel from high level to low level by affecting the behaviour of the system. Such adversaries are memoryless in the sense that their probabilistic behaviour is fixed at the beginning and does not change during the system execution. Formally, for each $H \in \mathcal{A}_{PNI}$, an equivalence check is performed between the process modelling the behaviour of the system without high-level interferences and the process modelling the behaviour of the system interacting with H. If such low-level views of the system are equivalent, then a low-level observer cannot infer anything about the activity of H or, in other words, H cannot affect the low-level behaviour of the system (see [A$^+$03] and the references therein). The main idea behind an approximated notion of probabilistic noninterference is to replace equivalence by similarity. This allows for some tolerance in the comparison of the two low-level views of the system rather than giving a binary answer to the potential equality between them. In this paper we concentrate on the problem of estimating such a tolerance factor in terms of "how much" similar two probabilistic processes are. The relaxed notion of indistinguishability we propose implies the one described in [A$^+$03] and allows us to apply a methodology to detect the most powerful adversary; this corresponds to the high-level user of $\mathcal{A}_{PNI}$ that maximises the difference between the two low-level views considered by *PNI*.

The criterion we adopt to establish the indistinguishability of two probabilistic processes refers to a probabilistic notion of weak bisimulation [BH97]. An approximated version of noninterference can be formalised through the definition of a corresponding similarity relation. We define such a relation in terms of the observable difference between two processes, where observability is based on the weak probabilistic bisimulation semantics. The maximal observable difference between processes will give us a measure of their indistinguishability, that is in our formalisation of noninterferference a measure of the information leakage.

A formal justification of such quantity is given in a mathematical framework in which the operational semantics of the probabilistic calculus is defined in terms of linear operators on a vector space representing the state space. In this setting we show that the metric induced by a particular operator norm on the process terms corresponds to a notion of distance which coincides with the measure above. This result exploits techniques introduced in [DHW02,DHW03].

The probabilistic model we adopt was introduced in [BB00] and used in [A$^+$03,ABG03] to formalise the *PNI* property. This model includes both probabilities and nondeterminism. On the one hand, a probability distribution is used to govern the choice among the various output actions that can be autonomously performed by the system. On the other hand, the choice among different types of input actions that can be accepted by the system depends on the environment behaviour. Therefore, such choices are nondeterministic. More precisely, once the environment has decided which type of input will be activated, the system reacts according to a specified probability distribution associated with the input actions of that type. Thus, the model is a mixed generative-reactive model, where output operations follow the generative model of probabilities and input operations follow the reactive model of probabilities [GSS95].

The presence of nondeterminism in the model is in a sense not compatible with our quantitative approach which requires that all transition probabilities are specified in order to calculate the behavioural difference between processes. In fact, in our security analysis nondeterminism is resolved by the probabilistic adversaries in $\mathcal{A}_{PNI}$, which play the role of probabilistic schedulers. Such schedulers are responsible of the information that may flow from the high to the low level, and therefore represent the "attackers" in our model of security.

In the rest of the paper, we first describe the formal model surveyed above (Section 2) and then present the *PNI* property and its application to an exact verification of the system security (Section 3), as done in [A$^+$03,ABG03]. In Section 4 we introduce the main contribution of this paper by presenting our quantitative approach towards measuring noninterference. In Section 5 we show that the quantity introduced as a measure for noninterference can be formalised in terms of the metric on probabilistic processes induced by a particular operator norm. Finally, in Section 6 we conclude by discussing some related work.

2 The Probabilistic Model

In this section we present the probabilistic framework in which we formalise a quantitative approach to noninterference. In particular, we consider a probabilistic calculus that was introduced in [BA03] and used in [ABG03] to define a probabilistic extension of the nondeterministic approach to noninterference of [FG95]. Such a calculus derives from a simple nondeterministic process algebra where actions are syntactically divided into input actions and output actions. Formally, for each visible action of type a, we distinguish the output action a and the input action a_*. Process terms synchronously communicate with the environment through their inputs and outputs, and perform internal computations through unobservable actions, termed τ actions.

Probabilities are introduced by adding probabilistic information to the algebraic operators. As an example, the classical CCS choice operator $P + Q$ is replaced by a probabilistic choice operator $P +^p Q$, where p is the parameter of a probability distribution that guides the choice between P (chosen with probability p) and Q (chosen with probability $1 - p$).

As far as the model of probabilities is concerned, we adopt a mixture of the generative and reactive approaches of [GSS95]. We assume that output actions behave as *generative* actions, i.e. the system autonomously decides, on the basis of a probability distribution, which output action will be offered to the environment and how to behave after such an event. In particular, τ is a generative action, since it expresses an autonomous internal move that does not react to external stimuli. For example, the process $a.P +^p \tau.Q$ will either execute a with probability p and then proceed as P, or execute τ with probability $1 - p$ and then proceed as Q.

Input actions are modelled as *reactive* actions, i.e. the system internally reacts to the choice of an action type performed by the environment. Once the action type has been (nondeterministically) chosen, a particular reactive action of that type is executed on the basis of a probability distribution. Thus, we can see the input actions as underspecified, since their execution is guided by the environment behaviour. For example, the process $b_*.P +^p (a_*.Q +^q a_*.R)$ will react to one of two action types, a or b, which can be selected by the environment. Note that such a choice is purely nondeterministic and does not depend on the probability distribution specified by parameter p. If the chosen event is a, the system performs one of the possible reactive input actions a_* according to the probability distribution specified by parameter q. Instead, if the chosen event is b, the system reacts by executing the unique reactive input action b_* it can perform with probability 1.

In the following, we present the syntax and the semantics of the probabilistic calculus. The interested reader is referred to [A$^+$03,ABG03,BA03,BB00] for more details. Then, we describe a notion of process equivalence [BH97] based on a probabilistic extension of the weak bisimulation of [Mil89].

2.1 Syntax

The syntax of the probabilistic process calculus is as follows:

$$P ::= \underline{0} \mid \pi.P \mid P +^p P \mid P \parallel^p_S P \mid P\backslash L \mid P/^p_a \mid A.$$

We use $\underline{0}$ to represent the terminated process (we usually omit it).

Action π is drawn from the set of actions Act and can be an internal action τ, an output action a, or an input action a_*, where a belongs to the set of visible action types $AType$. $\pi.P$ performs the action π with probability 1 and then behaves like P.

The alternative choice operator $P +^p Q$, where $p \in (0,1)$, performs a mixed probabilistic/nondeterministic choice among the actions of P and Q. More precisely, $P +^p Q$ executes a generative (reactive of type a) action of P with probability p and a generative (reactive of type a) action of Q with probability $1-p$. If one process P or Q cannot execute generative (reactive of type a) actions, $P +^p Q$ chooses a generative (reactive of type a) action of the other process with probability 1. The choice among generative and reactive actions and among reactive actions of different types is purely nondeterministic. Hence, the parameter that probabilistically guides the choices comes into play if and only if a probabilistic choice is actually performed.

The parallel composition operator $P \parallel^p_S Q$, where $p \in (0,1)$ and $S \subseteq AType$, asynchronously performs all the actions of P and Q that do not belong to the synchronisation set S and imposes synchronisation for all the actions belonging to S. Two actions can synchronise if they are of the same type a and either they are both input actions (and the result is an input action of type a), or one of them is an output action and the other one is an input action (and the result is an output action of type a). The probabilistic choice mechanism among the actions of P and Q is the same as described in the case of the choice operator. Because of the synchronisation policy, the execution of some actions of P may be prevented in $P \parallel^p_S Q$. Thus, we normalise the probabilities of the generative actions of P executable by $P \parallel^p_S Q$ in order to obtain a probability distribution. A symmetric argument holds for Q.

The restriction operator $P\backslash L$ prevents the execution of the actions of type in $L \subseteq AType$. The semantics of this operator can be expressed in terms of the parallel operator. In fact, we have that $P\backslash L$ corresponds to $P \parallel^p_L \underline{0}$, for any choice of parameter p.

The hiding operator $P/^p_a$ turns visible actions of type a into internal actions τ. In particular, when hiding an action a_*, we must pay attention to the side effect caused by the fact that a reactive action becomes a generative action. To this purpose, we use parameter p to express the probability that generative actions τ obtained by hiding reactive actions a_* of P are executed with respect to the generative actions previously enabled by P. Obviously, p is not used when hiding generative actions, because in such a particular case no nondeterminism must be resolved.

Example 2.1 Consider the process P defined by $a +^q b$, where the probabilistic choice is governed by parameter q. The semantics of $P/^p_a$ is given by $\tau +^q b$, i.e. it is again a probabilistic choice governed by parameter q. In such a case, parameter p of the hiding operator is not used since no nondeterministic choice must be resolved.

Now, consider process P given by $a_* +^q b$, where the choice is purely nondeterministic (parameter q is not considered). The semantics of $P/^p_a$ is the probabilistic choice $\tau +^p b$, governed by parameter p, between τ (obtained by hiding a_*) and b.

By turning reactive actions into generative invisible actions, the hiding operator allows us to obtain closed (fully generative) systems from open systems (i.e., systems enabling reactive choices). In order to obtain a closed system, the nondeterministic choices due to possible interactions with the environment have to be resolved, and parameter p turns such choices into probabilistic choices. In this sense, the hiding operator allows us to obtain a more concrete (fully specified) system which is an essential requirement for a quantitative reasoning about the system.

Constants A are used to specify recursive systems. In general, when defining a process term, we assume a set of constants defining equations of the form $A \stackrel{\Delta}{=} P$ (with P a guarded term [Mil89]) to be given.

In the rest of the paper, we denote by $\mathcal{G}$ the set of finite state, guarded, and closed terms [Mil89], called processes, generated by the syntax above. Moreover, we assume $p = \frac{1}{2}$ in the case parameter p of a probabilistic operator is omitted.

2.2 Operational Semantics

The semantics of the probabilistic process calculus is expressed in terms of mixed generative/reactive transition systems, which we now introduce. To this aim, we assume the following notation. Sets *RAct* and *GAct* denote the sets of input actions and of output and internal actions, respectively. We use the abbreviation $P \xrightarrow{\pi}$ for $P \xrightarrow{\pi, p} P'$, denoting that P can execute action π with probability p and then behave as P', for some $p \in]0, 1]$ and some process P'. We also use $P \xrightarrow{G}$, with $G \subseteq GAct$, to indicate $P \xrightarrow{a}$, for some $a \in G$, denoting that P can execute a generative action belonging to the set G.

The operational semantics of the probabilistic process algebra is given by the labeled transition system $(\mathcal{G}, Act, T)$, called generative/reactive transition system, where states are process terms and the transition relation T is the least multiset satisfying the operational rules reported in Table 1 and in Table 2. As far as the rules for $P +^p Q$ and $P \,\|^p_S\, Q$ are concerned, in addition to the reported rules, which refer to the local moves of the left-hand process P, we also consider the symmetric rules taking into account the local moves of the right-hand process Q. Such symmetric rules are obtained by exchanging the roles of terms P and Q in the premises and by replacing p with $1-p$ in the label of the derived transitions.

The semantics rules reflect the informal presentation of the syntax of the operators. Here, we go through some details concerning the parallel operator. If both P and Q can execute some synchronising actions a_* in $P \,\|_S^p\, Q$, then the composed system can execute some actions a_*: the probability of each action a_* executable by $P \,\|_S^p\, Q$ is the product of the probabilities of the two actions a_* (one of P and one of Q) that are involved in the synchronisation. The probabilities of the generative actions of P (Q) that are executable by $P \,\|_S^p\, Q$ are normalised in order to obtain a probability distribution [GSS95]. To this purpose, we employ some additional notation.

- The set $G_{S,Q} = \{a \in \mathit{AType} \cup \{\tau\} \mid a \notin S \vee (a \in S \wedge Q \xrightarrow{a_*})\}$ contains the action types not belonging to set S and the action types belonging to S for which an input action of Q can be performed. Intuitively, $G_{S,Q}$ contains all the types of the actions that any process P can execute within $P \,\|_S^p\, Q$.
- The function $\nu_P(G_{S,Q}) : \mathcal{P}(\mathit{AType} \cup \{\tau\}) \longrightarrow]0,1]$ computes the sum of the probabilities of the generative actions of P with type in $G_{S,Q}$. The value $\nu_P(G_{S,Q})$ is used to normalise the probabilities of the generative actions of P executable by $P \,\|_S^p\, Q$.

2.3 Weak Probabilistic Bisimulation

The security analysis we conduct is based on the semantics of processes (i.e., the security check considers the program behaviour), so that we need an equivalence relation allowing for a comparison among the observable behaviours of different systems. As argued in [Smith03,ABG03] a natural notion of observational equivalence on which to base notions of confinement is weak bisimulation: a semantics based on weak bisimulation allows us to neglect details about the internal behaviour of a system which are not important for a security analysis (such as the running time of a concurrent thread) and to concentrate only on those behaviours which are interesting for the analysis (e.g. behaviours which are observable from an external or low-level viewpoint). We consider here a probabilistic variant of the weak bisimulation which was introduced in [BH97]. Such a relation, denoted by $\approx_{\mathrm{PB}}$, is a probabilistic extension of the weak bisimulation ($\approx_{\mathrm{B}}$) of [Mil89]. In essence, $\approx_{\mathrm{PB}}$ replaces the classical weak transitions of $\approx_{\mathrm{B}}$ by the probability of reaching classes of equivalent states. More precisely, we use a function *Prob* such that $\mathit{Prob}(P, \pi, C)$ denotes the aggregate probability of going from P to a term in the class (of equivalent terms) C by executing an action π, and $\mathit{Prob}(P, \tau^* a, C)$ expresses the aggregate probability of going from P to a term in the equivalence class C via sequences of any number of τ actions followed by an action a, possibly equal to τ.

Lemma 2.2 *The value of $\mathit{Prob}(P, \tau^* a, C)$ is the minimal non-negative solution to the equation system:*

$$\begin{cases} 1 & \text{if } a = \tau \wedge P \in C \\ \sum_{Q \in \mathcal{G}} \mathit{Prob}(P, \tau, Q) \cdot \mathit{Prob}(Q, \tau^*, C) & \text{if } a = \tau \wedge P \notin C \\ \sum_{Q \in \mathcal{G}} \mathit{Prob}(P, \tau, Q) \cdot \mathit{Prob}(Q, \tau^* a, C) + \mathit{Prob}(P, a, C) & \text{if } a \neq \tau \end{cases}$$

As shown in [ABG03], this system has a least solution. We are now ready to define the

$$\pi.P \xrightarrow{\pi,1} P$$

$$\frac{P \xrightarrow{a_*,q} P' \quad Q \xrightarrow{a_*}}{P +^p Q \xrightarrow{a_*,p\cdot q} P'} \qquad \frac{P \xrightarrow{a_*,q} P' \quad Q \not\xrightarrow{a_*}}{P +^p Q \xrightarrow{a_*,q} P'}$$

$$\frac{P \xrightarrow{a,q} P' \quad Q \xrightarrow{GAct}}{P +^p Q \xrightarrow{a,p\cdot q} P'} \qquad \frac{P \xrightarrow{a,q} P' \quad Q \not\xrightarrow{GAct}}{P +^p Q \xrightarrow{a,q} P'}$$

$$\frac{P \xrightarrow{a_*,q} P' \quad P \xrightarrow{GAct}}{P/^p_a \xrightarrow{\tau,p\cdot q} P'/^p_a} \qquad \frac{P \xrightarrow{a_*,q} P' \quad P \not\xrightarrow{GAct}}{P/^p_a \xrightarrow{\tau,q} P'/^p_a}$$

$$\frac{P \xrightarrow{b_*,q} P'}{P/^p_a \xrightarrow{b_*,q} P'/^p_a} \; a \neq b$$

$$\frac{P \xrightarrow{b,q} P' \quad P \xrightarrow{a_*}}{P/^p_a \xrightarrow{b,(1-p)\cdot q} P'/^p_a} \; a \neq b \qquad \frac{P \xrightarrow{a,q} P' \quad P \xrightarrow{a_*}}{P/^p_a \xrightarrow{\tau,(1-p)\cdot q} P'/^p_a}$$

$$\frac{P \xrightarrow{b,q} P' \quad P \not\xrightarrow{a_*}}{P/^p_a \xrightarrow{b,q} P'/^p_a} \; a \neq b \qquad \frac{P \xrightarrow{a,q} P' \quad P \not\xrightarrow{a_*}}{P/^p_a \xrightarrow{\tau,q} P'/^p_a}$$

$$\frac{P \xrightarrow{\pi,q} P'}{A \xrightarrow{\pi,q} P'} \; \text{if } A \triangleq P$$

Table 1
Operational semantics (part I)

weak probabilistic bisimulation equivalence.

Definition 2.3 An equivalence relation $R \subseteq \mathcal{G} \times \mathcal{G}$ is a weak probabilistic bisimulation if and only if, whenever $(P, Q) \in R$, then for all $C \in \mathcal{G}/R$:

- $Prob(P, \tau^* a, C) = Prob(Q, \tau^* a, C)$ $\forall a \in GAct$
- $Prob(P, a_*, C) = Prob(Q, a_*, C)$ $\forall a_* \in RAct$.

Two terms $P, Q \in \mathcal{G}$ are weakly probabilistically bisimulation equivalent, denoted $P \approx_{\mathrm{PB}} Q$, if

$$\frac{P \xrightarrow{a_*,q} P' \quad Q \xrightarrow{a_*}}{P \,\|^p_S\, Q \xrightarrow{a_*,p\cdot q} P' \,\|^p_S\, Q}\; a \notin S \qquad \frac{P \xrightarrow{a_*,q} P' \quad Q \not\xrightarrow{a_*}}{P \,\|^p_S\, Q \xrightarrow{a_*,q} P' \,\|^p_S\, Q}\; a \notin S$$

$$\frac{P \xrightarrow{a_*,q} P' \quad Q \xrightarrow{a_*,q'} Q'}{P \,\|^p_S\, Q \xrightarrow{a_*,q\cdot q'} P' \,\|^p_S\, Q'}\; a \in S$$

$$\frac{P \xrightarrow{a,q} P' \quad Q \xrightarrow{G_{S,P}}}{P \,\|^p_S\, Q \xrightarrow{a,p\cdot q/\nu_P(G_{S,Q})} P' \,\|^p_S\, Q}\; a \notin S$$

$$\frac{P \xrightarrow{a,q} P' \quad Q \not\xrightarrow{G_{S,P}}}{P \,\|^p_S\, Q \xrightarrow{a,q/\nu_P(G_{S,Q})} P' \,\|^p_S\, Q}\; a \notin S$$

$$\frac{P \xrightarrow{a,q} P' \quad Q \xrightarrow{a_*,q'} Q' \quad Q \xrightarrow{G_{S,P}}}{P \,\|^p_S\, Q \xrightarrow{a,p\cdot q'\cdot q/\nu_P(G_{S,Q})} P' \,\|^p_S\, Q'}\; a \in S$$

$$\frac{P \xrightarrow{a,q} P' \quad Q \xrightarrow{a_*,q'} Q' \quad Q \not\xrightarrow{G_{S,P}}}{P \,\|^p_S\, Q \xrightarrow{a,q'\cdot q/\nu_P(G_{S,Q})} P' \,\|^p_S\, Q'}\; a \in S$$

Table 2
Operational semantics (part II)

there exists a weak probabilistic bisimulation R including the pair (P, Q).

Note that such a definition requires two equivalent terms to be strongly equivalent in the case of reactive actions and weakly equivalent in the case of generative actions. This is because τ is a generative action, therefore computing the probability associated with a mixed trace of generative/reactive actions (like, e.g., $\tau^* a_*$) does not actually make sense. Note also that, as shown in [BH97], the first equation in Definition 2.3 can be equivalently written as $Prob(P, \tau^* a \tau^*, C) = Prob(Q, \tau^* a \tau^*, C)$.

Example 2.4 Processes $P \triangleq a +^{\frac{1}{2}} b$ and $Q \triangleq \tau.Q +^{\frac{1}{3}} (a +^{\frac{1}{2}} b)$ behave the same from the viewpoint of an external observer, who can see either an output action a or an output action b with equal probabilities. Formally, P and Q are weakly probabilistically bisimulation equivalent and the relation that satisfies Definition 2.3 is $R = \{C, [\underline{0}]\}$, with $C = \{P, Q\}$ and $[\underline{0}] = \{\underline{0}\}$. The only interesting case to be verified is related to the execution of a visible action (possibly preceded by a sequence of internal actions) starting from the initial state and reaching the null term. As it is easy to see, we have $Prob(P, \tau^*\pi, [\underline{0}]) = \frac{1}{2}$, with $\pi \in \{a, b\}$. As far as $Prob(Q, \tau^*\pi, [\underline{0}])$ is concerned, we observe that Q can execute an arbitrary number of times the action τ before reaching state $\underline{0}$ via an action a (b). Hence, the probability $\frac{1}{3}$ associated with the outgoing internal transition of Q is distributed among the other outgoing transitions of Q. Formally, by resolving the equation system of Lemma 2.2, we have $Prob(Q, \tau^* a, [\underline{0}]) = \frac{1}{3} \cdot Prob(Q, \tau^* a, [\underline{0}]) + \frac{1}{3}$, from which we derive $Prob(Q, \tau^* a, [\underline{0}]) = \frac{1}{2}$ (similarly for b). Therefore, R is a weak probabilistic bisimulation and $P \approx_{\mathrm{PB}} Q$.

3 Probabilistic Noninterference

According to the standard definition of noninterference given by Goguen and Meseguer [GM82], a high-level user (High, for short) is said to interfere with a low-level user (Low, for short) if what High can do is reflected on what Low can observe. In this setting, High can perform high-level activities only and observe all the interactions between the system and the environment. Instead, Low can perform low-level activities only and is not allowed to directly observe the occurrence of high-level events. In spite of this, Low may succeed in detecting the High behaviour by simply interacting with the low-level interface of the system. In other words, even if there does not exist a direct communication channel from High to Low, High may have the possibility of indirectly passing information to Low through interactions with the system. Noninterference analysis mainly aims at checking the presence of indirect information flows, called covert channels, from High to Low. In our probabilistic framework, what Low can see in order to infer the High behaviour is not only the logical low-level interface of the system interacting with the environment, but also the probability distribution of the events representing such interactions. In this section we describe a formalisation of the noninterference approach in the probabilistic process calculus surveyed in the previous section [A$^+$03,ABG03].

The noninterference-based security analysis roughly consists of deriving two models from the system specification at hand, corresponding to two different low-level views of the system, and then checking the semantic equivalence between such derived models. The semantic equivalence between processes is based on the weak probabilistic bisimulation $\approx_{\mathrm{PB}}$, introduced in Section 2.3, while the choice of the low-level models to be compared depends on the definition of the security property. The property we consider here was introduced in [ABG03] as a probabilistic extension of the Strong Nondeterministic Noninterference property of [FG95]. Such a property, which we call Probabilistic Noninterference (*PNI*), compares the low-level view of the system in the absence of high-level interactions and the low-level view of the system in the presence of high-level interactions.

Formally, we divide actions into high-level actions and low-level actions, denoted *High* and *Low*, respectively, depending on the nature of the activities they represent. *High* and *Low* are two disjoint sets that form a covering of *AType*. Given a process P, we denote with $\bar{h}^P = h_1^P, \ldots, h_n^P$ the sequence (in alphabetic order) of types of the high-level input actions that syntactically occur in the action prefix operators within P. Then, the application of the security check to P is as follows. The low-level view of P in the absence of high-level operations is obtained by preventing P from executing its high-level actions. This is carried out by applying the restriction operator to P, i.e. $P \backslash High$. The low-level view of P in the presence of high-level interactions is obtained by turning all the high-level actions of P into invisible actions, since Low is not expected to observe them. This is carried out by applying the hiding operator to P, thus obtaining a family of processes $P/^{p_1}_{h_1^P} \ldots /^{p_n}_{h_n^P}$, with $p_1, \ldots, p_n \in (0, 1)$, where each possible sequence $\bar{p} = p_1, \ldots, p_n$ expresses the probability distribution (chosen by High) of the hidden high-level input actions executable by P. In the following, we sometimes use the abbreviation $P/^{\bar{p}}_{\bar{h}^P}$ to stand for $P/^{p_1}_{h_1^P} \ldots /^{p_n}_{h_n^P}$. Finally, for each possible $\bar{p}$, we compare $P\backslash High$ and $P/^{\bar{p}}_{\bar{h}^P}$ to check whether they are weak probabilistic bisimilar. If such a condition holds we say that P satisfies the *PNI* property, or $P \in PNI$.

Definition 3.1 $P \in PNI \Leftrightarrow P\backslash High \approx_{PB} P/^{p_1}_{h_1^P} \ldots /^{p_n}_{h_n^P} \; \forall p_1, \ldots, p_n \in (0,1)$.

Note that the sequence $\bar{p} = p_1, \ldots, p_n$ represents the particular probabilistic behaviour (i.e. the strategy) followed by High. Hence, the universal quantification over all possible sequences imposes that the equivalence check must hold for each High strategy. In particular, we can interpret each $\bar{p}$ as representing an adversary whose probabilistic behaviour may be responsible for setting up a covert channel from High to Low. The *PNI* definition takes into account a family $\mathcal{A}_{PNI}$ of adversaries (against which *PNI* checks the presence of information flows from High to Low) that turn out to be:

- *active*: an adversary can alter the probabilistic low-level behaviour of the system, since the application of the hiding operator affects the probability distribution of the generative low-level actions.
- *memoryless*: an adversary cannot alter its strategy step by step, since the probability distribution of the hidden high-level inputs, expressed by parameters $p_1, \ldots, p_n$ and chosen by the adversary, does not change during the system execution.

If *PNI* holds, Low cannot infer the behaviour of any adversary in $\mathcal{A}_{PNI}$, that means the system does not leak information from High to Low.

As shown in [A$^+$03,ABG03], probabilistic noninterference reveals covert channels that are not observable in a purely nondeterministic setting, and offers the means for measuring the information leakage in terms of probability of observing the related covert channel. We now provide some examples showing the expressive power of *PNI*.

Example 3.2 Consider process $P \triangleq h_*.(l_*.\underline{0} + l'_*.\underline{0}) + (l_*.\underline{0} + l'_*.\underline{0})$, which may accept a high-level input of type h before interacting with Low through one of the low-level actions, l_* or l'_*. From the viewpoint of Low, the observable interface of the system cannot be altered by the strategy followed by High. This is reflected by the security check, which states that P satisfies the *PNI* property, since $(l_*.\underline{0} + l'_*.\underline{0}) \approx_{PB} \tau.(l_*.\underline{0} + l'_*.\underline{0}) +^p (l_*.\underline{0} + l'_*.\underline{0})$ for any choice of parameter p.

An example of an information flow from High to Low is given by the following process $P \triangleq l.(h_*.\underline{0} + l'.\underline{0}) + l'.l.\underline{0}$. If P produces the sequence of low-level outputs $l'.l$, then High cannot interact with the system and no information can flow from High to Low. Instead, if the system first chooses l, then High can be responsible for deciding whether l' will be executed. Formally, $l.l'.\underline{0} + l'.l.\underline{0}$, which expresses the semantics of $P\backslash\{h\}$, is not weak probabilistic bisimilar to $l.(\tau.\underline{0} +^p l'.\underline{0}) + l'.l.\underline{0}$, which expresses the semantics of $P/^p_h$. In particular, they cannot be equivalent for any choice of parameter p of the hiding operator. The same example revisited in a nondeterministic scenario reveals the same covert channel described above, which turns out to be a purely possibilistic information flow.

An example of a probabilistic information leakage is given by the process $P \triangleq l.h.l'.\underline{0} +$

$(l.l'.\underline{0}+l.\underline{0})$. The behaviour of High does not affect the set of possible results, $l.l'$ or l, observable by Low. However, High can alter the probability distribution of such results. Formally, a probabilistic covert channel is captured by the *PNI* property, since $l.\underline{0} + (l.l'.\underline{0} + l.\underline{0})$, which is the semantics of $P\backslash\{h\}$, is equivalent to $l.l'.\underline{0} +^{1/4} l.\underline{0}$, and $P \triangleq l.\tau.l'.\underline{0} + (l.l'.\underline{0} + l.\underline{0})$, which is the semantics of $P/_h$, is equivalent to $l.l'.\underline{0} +^{3/4} l.\underline{0}$. Hence, $P\backslash\{h\}$ and $P/_h$ are not weak probabilistic bisimilar. From a statistical viewpoint, if High interferes and Low observes repeated executions of P, then, on average, the result of $\frac{3}{4} \cdot n$ experiments over n will be $l.l'$. On the other hand, the same observation occurs $\frac{1}{4} \cdot n$ times over n (on average) in the case High does not interact with P. From the viewpoint of Low, a small number of experiments is sufficient to guess the behaviour of High (see [A+03] for a mathematical justification of this statistical interpretation).

The following example shows an application of the noninterference general idea to study interferences between honest users and malicious parties (see, e.g., [FGM00] for an application of the noninterference approach to the analysis of cryptographic protocols).

Example 3.3 Let us assume that Low represents a user that interacts with the system in order to obtain a service, while High is a potential adversary interacting with the system with some malicious intentions. In particular, we consider an abstraction of a low-level shared resource with password-based access. Low can access and consume the resource, while High is not allowed to. In spite of this, High can try to guess the access password in order to consume the resource in place of Low. The overall system is given by the parallel composition of two processes, $LU \parallel_{Act} Resource$, where LU expresses the behaviour of the low-level user:

$$\begin{aligned} LU \triangleq\ & low\_request\,.\,low\_insert\_password\,.\,low\_consume\_resource\,.\,\underline{0} + \\ & low\_not\_available_*\,.\,\underline{0} \end{aligned}$$

and *Resource* models the shared resource that reacts to external requests, which may arrive either from LU or from the high-level adversary:

$$\begin{aligned} Resource \triangleq\ & low\_request_*\,.\,low\_insert\_password_*\,.\,low\_consume\_resource_*\,.\,\underline{0} + \\ & high\_request_*\,.\,(high\_try\_password_*\,.\,high\_consume\_resource_*\,. \\ & \qquad low\_not\_available\,.\,\underline{0} +^q \\ & \qquad high\_try\_password_*\,.\,Resource) \end{aligned}$$

In the absence of the adversary, Low can normally access and consume the resource. Formally, $(LU \parallel_{Act} Resource)\backslash High$ does not enable the action *low_not_available*. On the other hand, if High tries to interfere, then Low may not be able to access the resource. In particular, we have that $(LU \parallel_{Act} Resource)/^{p}_{high\_request}/_{high\_try\_password}/_{high\_consume\_resource}$ reaches the null term by performing sequences of the form $\tau^* low\_not\_available$ with probability $p \cdot q \cdot \sum_{i=0}^{\infty}(p \cdot (1-q))^i = \frac{p \cdot q}{1-p+p\cdot q}$. Note that when hiding the action of type *high_request*, parameter p of the hiding operator is used to resolve the choice between the resulting invisible action and the synchronising action *low_request*. On the other hand, when hiding the action of type *high_try_password* (resp. *high_consume_resource*) the parameter of the hiding operator is not used, since the system performs an action of this type with probability 1.

4 Measuring Noninterference

In this section we show how to exploit the probabilistic information associated with the behaviour of a system in order to give a quantitative estimate of possible information leakages. This gives us a means to evaluate the effectiveness of a covert channel that is responsible for an illegal information flow. As shown in [A+03], this is related to the number of tests (system executions) needed to an external observer for detecting such an information flow.

The technique we are going to introduce aims at quantifying the information leakage of a system by calculating the maximal difference between the transition probabilities observed by a low-level user when the system is interacting with High and when it is not, respectively. As a consequence a basic requirement for this technique is that the systems we analyse are *fully specified* from the viewpoint of a low-level observer. This means that the nondeterminism due to possible interactions with the environment has to be resolved. In fact, this is the effect of the hiding operator when turning reactive actions into generative internal actions. Since in our modelling of the *PNI* property hiding is applied only to high-level actions, we only need to assure that these are the only reactive transitions enabled by the system. We think that this assumption is not restrictive, since a quantitative estimate of a covert channel observed by Low really makes sense if (i) the behaviour of Low is fully specified and (ii) we use a technique that takes into account all the possible ways in which the behaviour of High may influence the probability distribution of the low-level events. Our approach would be applicable also to the general case of systems including low-level reactive actions, provided that we complicated the model in order to consider all the possible associated interactions.

The probability of observing an information flow from high level to low level can be estimated by relaxing the behavioural equivalence relation expressed by the weak probabilistic bisimulation defined in Section 2.3. The intuitive idea, inspired by [A+03,ABG03], is as follows. According to the *PNI* property introduced in Section 3, a process P is not secure if the low-level models corresponding to the behaviours of P with and without High interferences are not equivalent in the probabilistic weak bisimulation semantics. Therefore, an information leakage is detected when, for a given sequence $\bar{p}$ chosen by High, for each equivalence relation $R \subseteq \mathcal{G} \times \mathcal{G}$ including the pair $(P \backslash High, P/^{\bar{p}}_{\bar{h}^P})$, there exist $C \in \mathcal{G}/R$, $a \in GAct$, and a pair $(P', P'') \in R$, such that

$$Prob(P', \tau^* a, C) \neq Prob(P'', \tau^* a, C).$$

The difference between these two probabilities can be used to give an estimate of the amount of information leakage. More precisely, for every equivalence relation R including the pair $(P \backslash High, P/^{\bar{p}}_{\bar{h}^P})$, we consider the pair of states (of a class in $\mathcal{G}/R$) where the weak transition probabilities are maximally different and calculate the difference. We can then define a measure of the security of P as the minimal of these differences over all equivalence relations.

More formally, we define the quantity $\delta^R_{\bar{p}}(P)$ (or simply $\delta^R_{\bar{p}}$ when process P is clear from the context), which expresses the behavioural distance between the low-level models $P \backslash High$ and $P/^{\bar{p}}_{\bar{h}^P}$ of a system P with respect to a given relation $R \subseteq \mathcal{G} \times \mathcal{G}$ including the pair $(P \backslash High, P/^{\bar{p}}_{\bar{h}^P})$ and a particular choice of the sequence of parameters $\bar{p} = p_1, \ldots, p_n$ governing the interaction of each high-level input action of P with High.

Definition 4.1 Let P be a process, $R \subseteq \mathcal{G} \times \mathcal{G}$ an equivalence relation including the pair

$(P\backslash High, P/^{\bar{p}}_{\bar{h}P})$, and $\bar{p} = p_1, \ldots, p_n$ a sequence of parameters such that $p_i \in (0,1), 1 \leq i \leq n$. We define

$$\delta^R_{\bar{p}} = \sup_{\substack{(P', P'') \in R \\ a \in GAct \\ C \in \mathcal{G}/R}} \mid Prob(P', \tau^* a, C) - Prob(P'', \tau^* a, C) \mid .$$

By using this quantity we can then define a measure for the security degree of a system P as follows.

Definition 4.2 Let P be a process and let R, $\bar{p}$ and $\delta^R_{\bar{p}}$ as in Definition 4.1. Then we define

$$\varepsilon_{\bar{p}} = inf_R \; \delta^R_{\bar{p}}.$$

The quantity $\varepsilon_{\bar{p}}$ expresses the maximal distance between the process without High interferences and the process interacting with the high-level user modelled by sequence $\bar{p}$, obtained for the particular relation that is the best approximation of a weak probabilistic bisimulation. We point out that this quantity depends on parameters $p_1, \ldots, p_n$ forming the sequence $\bar{p}$. These parameters represent an hypothetical high-level adversary that can pass to Low an amount of information according to the quantity $\varepsilon_{\bar{p}}$. The measure $\varepsilon_{\bar{p}}$ can also be interpreted as the "effectiveness" of the adversary strategy corresponding to the sequence $\bar{p}$. In fact, it determines *how easy* it is for a low-level user to obtain some confidential information, in terms of the number of tests (system executions) the low-level user needs to perform in order to distinguish the behaviours with and without the interference of such an adversary. This number of tests can be analysed by using various standard statistical methods, such as the so-called *hypothesis testing* method [Shao99]. This method provides a simple way to estimate how many tests are needed to distinguish two processes and the confidence that the tests outcome is correct. The application of this method for a statistical interpretation of the approximation of confinement properties was first described in [DHW02]; a detailed description of this statistical interpretation in our process algebraic setting can be found in [A$^+$03].

In [A$^+$03,ABG03] an approximated notion of noninterference is proposed by employing a relaxed version of the weak probabilistic bisimulation $\approx_{\mathrm{PB}}$, termed weak probabilistic bisimulation with ε-precision ($\approx_{PB_\varepsilon}$), which is a non-transitive relation formally defined as follows.

Definition 4.3 A relation $R \subseteq \mathcal{G} \times \mathcal{G}$ is a weak probabilistic bisimulation with ε-precision, where $\varepsilon \in (0,1)$, if and only if, whenever $(P, Q) \in R$, then for all $C \in \mathcal{G}/R$:

- $\mid Prob(P, \tau^* a, C) - Prob(Q, \tau^* a, C) \mid \leq \varepsilon \;\; \forall a \in GAct$
- $\mid Prob(P, a_*, C) - Prob(Q, a_*, C) \mid \leq \varepsilon \;\; \forall a_* \in RAct.$

If $P \approx_{PB_\varepsilon} Q$ then there exists a weak probabilistic bisimulation with ε-precision including the pair (P, Q).

By replacing $\approx_{\mathrm{PB}}$ by $\approx_{PB_\varepsilon}$ in the *PNI* definition, we obtain a relaxed property that checks whether a system P is an approximated version of a secure system up to a tolerance ε. We recall that in this section we apply the security check to fully generative processes, since the systems we consider do not enable reactive low-level transitions. Hence, the verification of the condition of Definition 4.3 is applied to the generative actions $a \in GAct$ only.

With respect to the *PNI* security check based on Definition 4.3, in this section we have proposed a specific notion of measure $\varepsilon_{\bar{p}}$ of the security of a system P against the adversary corresponding to sequence $\bar{p}$. An important consequence of Definitions 4.1 and 4.2 is that P turns out to be approximately secure (up to ε) against the adversary modelled by $\bar{p}$ – namely there exists a relation including the pair $(P\backslash High, P/^{\bar{p}}_{\bar{h}P})$ that is a weak probabilistic bisimulation with ε-precision – if and only if $\varepsilon_{\bar{p}}$ is less than ε.

Proposition 4.4 *Let P be a process and $\varepsilon_{\bar{p}}$ as in Definition 4.2. Then*

(i) $P\backslash High \approx_{PB_{\varepsilon_{\bar{p}}}} P/^{\bar{p}}_{\bar{h}P}$.

(ii) $P\backslash High \approx_{PB_{\varepsilon}} P/^{\bar{p}}_{\bar{h}P}$ *for each* $\varepsilon > \varepsilon_{\bar{p}}$.

(iii) *There does not exist* $\varepsilon < \varepsilon_{\bar{p}}$ *such that* $P\backslash High \approx_{PB_{\varepsilon}} P/^{\bar{p}}_{\bar{h}P}$.

Proof. We immediately derive (i) from Definitions 4.2, 4.1 and from the definition of weak probabilistic bisimulation with ε-precision. Condition (ii) simply holds if we take the relation R such that $\varepsilon_{\bar{p}} = \delta^R_{\bar{p}}$. Finally, to derive (iii), it suffices to observe that if such an ε exists, then there also exists a relation R' including the pair $(P\backslash High, P/^{\bar{p}}_{\bar{h}P})$ such that, whenever $(P', P'') \in R'$, for each $C \in \mathcal{G}/R'$ it holds that $|\,Prob(P', \tau^*a, C) - Prob(P'', \tau^*a, C)\,| \leq \varepsilon \ \ \forall a \in GAct$. That means $\delta^{R'}_{\bar{p}} \leq \varepsilon$, thus violating the hypothesis that $\varepsilon_{\bar{p}}$ is the minimum, over all relations R, of the family of values $\delta^R_{\bar{p}}$. □

As we have seen, *PNI* checks whether the system is secure against a class $\mathcal{A}_{PNI}$ of adversaries, among which we are interested in evaluating the effectiveness of the adversary that maximises the information leakage, i.e. the most powerful adversary. Formally, if $\bar{q} = q_1, \ldots, q_n$ is the sequence representing such an adversary, we have that $\varepsilon_{\bar{q}} = sup_{\bar{p}}\ \varepsilon_{\bar{p}}$. The problem of estimating the most effective adversary corresponds to the problem of finding the least upper bound of the function $\varepsilon_{\bar{p}}$.

4.1 Examples

We now provide some intuitive examples that explain the role of our approximation in estimating the security degree of systems and determining the worst case, i.e. the maximal information leakage that the class of adversaries defined by *PNI* may set up from high level to low level.

Example 4.5 Process $P \stackrel{\Delta}{=} h_*.l + \tau.(h_*.l'.\underline{0} + l.\underline{0})$ performs either a high-level input operation of type h or an internal move. Such a choice is nondeterministic, as it depends on the behaviour of High. Then, if the action τ is executed, another nondeterministic choice is to be performed between the high-level input operation of type h and a low-level output l. Note that an information flow from High to Low occurs if the action l' is observed by Low, as l' is executed only if High interacts with the system. On the other hand, the execution of l does not reveal anything about the strategy followed by High. In Figure 1 we show the labeled transition systems modelling the low-level views of P to be compared through equivalence checking. The left-hand system represents the behaviour of $P\backslash High$. The right-hand system expresses what Low can observe in the case High decides to interact with P, i.e. $P/^p_h$. In such a case, each nondeterministic choice is probabilistically resolved by High according to parameter p. In other words, each $p \in (0, 1)$ expresses a high-level strategy followed by the adversary.

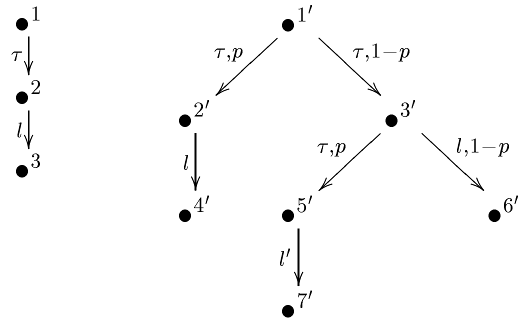

Fig. 1. Process $h_*.l + \tau.(h_*.l'.\underline{0} + l.\underline{0})$: what Low can observe with or without High interactions. Transition probabilities are omitted when equal to 1.

Let us consider the most relevant equivalence relations.
For $R_1 = \{\{1, 1'\}, \{2, 2', 3'\}, C = \{5'\}, \{3, 4', 6', 7'\}\}$ we have

$$\delta_p^{R_1} = |\, Prob(2, \tau^*, C) - Prob(3', \tau^*, C)\,| = |\, 0 - p\,| = p,$$

and for $R_2 = \{\{1, 1'\}, C = \{2, 2'\}, \{3'\}, \{5'\}, \{3, 4', 6', 7'\}\}$ we have

$$\delta_p^{R_2} = |\, Prob(1, \tau^*, C) - Prob(1', \tau^*, C)\,| = 1 - p$$

from which we derive $\varepsilon_p = min\{p, 1 - p\}$. That means the probabilistic behaviour of High directly affects the capability of distinguishing between the two low-level views of the system. In order to evaluate the worst case, we maximise function ε_p, thus obtaining $p = \frac{1}{2}$, for which we have $\varepsilon_p = \frac{1}{2}$. In fact, it is easy to see that $\forall p \neq \frac{1}{2}$, $\varepsilon_p < \frac{1}{2}$. We can then conclude that the most powerful adversary is the high-level user expressed by parameter $p = \frac{1}{2}$. In the statistical interpretation mentioned above, where ε_p is inversely proportional to the number of tests the adversary has to perform to breach the security of the system, the most powerful adversary is the attacker which needs the minimal number of tests. In particular, assume that the most powerful adversary, modelled by parameter $p = \frac{1}{2}$, is interacting with the system. By applying the technique explained in [A$^+$03], we obtain that the probability for a low-level observer of guessing the activity of such a high-level user is about 96.5% after 10 tests. As another example, consider the high-level user modelled by parameter $p = 0.01$. Such an adversary interferes with the system in a way that is rarely revealed by the low-level observer. Indeed, in order to correctly guess the behaviour of such an adversary with the same probability of success as above, the observer needs about 1312 tests, while after 10 tests such a probability is about 56%, which is a value very close to the success probability of a blind guess.

Example 4.6 Consider the second process of Example 3.2, $P \triangleq l.(h_*.\underline{0} + l'.\underline{0}) + l'.l.\underline{0}$. The low-level models to be compared are $l.l'.\underline{0} + l'.l.\underline{0}$ and $l.(\tau.\underline{0} +^p l'.\underline{0}) + l'.l.\underline{0}$, which, as we have seen, cannot be weakly bisimulation equivalent (see Fig. 2). In particular, the distinguishing behaviour arises if we execute l with probability $\frac{1}{2}$. In fact, after such an event the former process executes the action l' with probability 1, while the latter process executes either an internal move with probability p or the action l' with probability $1 - p$.
Formally, for $R_1 = \{\{1, 1'\}, \{2, 2'\}, \{3, 3'\}, C = \{4, 5, 4', 5', 6'\}\}$ we have

$$\delta_p^{R_1} = |\, Prob(2, \tau^* l', C) - Prob(2', \tau^* l', C)\,| = |\, 1 - (1 - p)\,| = p,$$

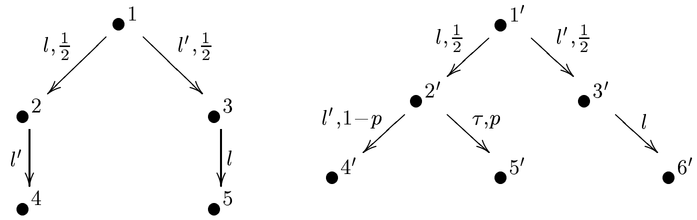

Fig. 2. Two low-level views of process $l.(h_*.\underline{0} + l'.\underline{0}) + l'.l.\underline{0}$

and for $R_2 = \{\{1, 1'\}, C = \{2\}, \{2'\}, \{3, 3'\}, \{4, 5, 4', 5', 6'\}\}$ we have

$$\delta_p^{R_2} = |\, Prob(1, \tau^* l, C) - Prob(1', \tau^* l, C)\,| = \frac{1}{2}.$$

Since for any relation R different from R_1 and R_2 we have $\delta_p^R \geq \frac{1}{2}$, it follows $\varepsilon_p = min\{\frac{1}{2}, p\}$. Note that the information leakage is negligible as p tends to zero, because in such a case also ε_p tends to zero. That means High is not interfering (rarely interferes) with the system. Clearly, the closer p is to 1, the easier it is for Low to reveal the behaviour of High. However, even for the worst case (corresponding to the limiting scenario $p = 1$), we have $\varepsilon_p = \frac{1}{2}$, i.e. the maximal probability of observing an information leakage depends on the probability of reaching the state where High can actually interfere.

Example 4.7 Consider the shared low-level resource described in Example 3.3. The low-level views to be compared through equivalence checking are depicted in Fig. 3.

If, for instance, we take the relation $R_1 = \{\{1, 1'\}, \{2, 2'\}, \{3, 3'\}, \{4, 4'\}, C = \{5'\}, \{6', 7'\}\}$, we obtain

$$\delta_p^{R_1} = |\, Prob(1, \tau^*, C) - Prob(1', \tau^*, C)\,| = p.$$

Instead, for $R_2 = \{\{1, 1', 5'\}, \{2, 2'\}, \{3, 3'\}, \{4, 4'\}, C = \{6', 7'\}\}$, from Example 3.3 we derive

$$\delta_p^{R_2} = |\, Prob(1, \tau^*, C) - Prob(5', \tau^*, C)\,| = q \cdot \sum_{i=0}^{\infty} (p \cdot (1-q))^i = \frac{q}{1 - p + p \cdot q}.$$

Moreover, it can be verified that considering other relations is not meaningful if we want to estimate ε_p. Hence, we have

$$\varepsilon_p = min\{p, \frac{q}{1 - p + p \cdot q}\}.$$

We recall that q is the probability of guessing the password and p is the parameter of the hiding operator, which guides the choice between low-level access requests and high-level access requests. Therefore, it is easy to verify that if High lets parameter p tend to 1 then ε_p tends to 1. That means if the low-level access request is always preempted by the high-level access request, High will always succeed in guessing the password and consuming the resource.

5 The Measure $\varepsilon_{\bar{p}}$ and Operator Norms

The quantity $\varepsilon_{\bar{p}}$ introduced in the previous section is based on a "behavioural distance" between processes defined by considering all possible equivalence relations. We now give a more formal justification of such a distance in terms of an appropriate metric on the space of the

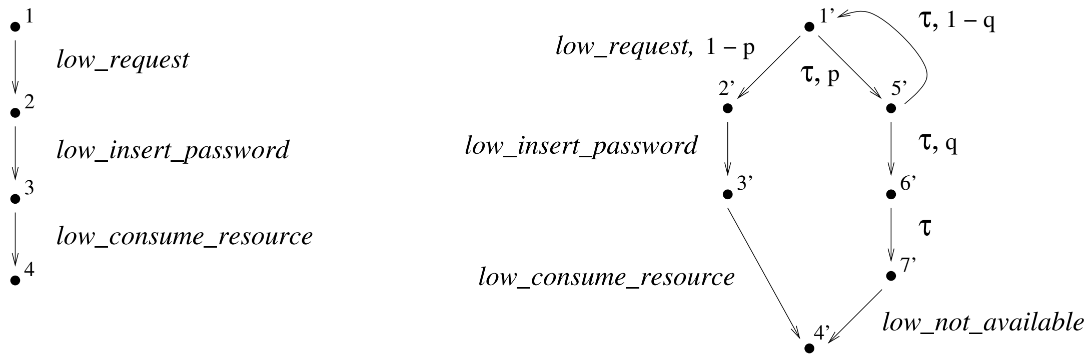

Fig. 3. Low-level views for the shared resource example.

processes in the calculus defined in Section 2.1. We use to this purpose the linear operators framework introduced in [DHW03a,DHW03] for defining approximate process equivalences. In this framework, the operational semantics of a probabilistic process is described by a linear operator representing its transition graph, and the distance between processes is defined via the notion of *operator norm*. In the following, we first re-cast our security framework based on weak probabilistic bisimulation in this linear operators setting, and then we show that the quantity $\varepsilon_{\bar{p}}$ corresponds to a particular operator norm which captures the idea of behavioural distance introduced in Section 4.

5.1 Weak Probabilistic Bisimulation via Linear Operators

We will base our treatment on the basic assumption introduced in Section 4 that systems are fully specified from the viewpoint of a low-level observer or, in other words, all choices are probabilistic after the potential interaction of the system with High has been resolved. As a consequence, we can actually consider as a reference model a restriction of the model introduced in Section 2 where only generative transitions are executable. More precisely, we can refer to a restricted version of the labelled transition system $(\mathcal{G}, Act, T)$ that considers fully generative processes only.

Based on such a restriction, as described in [DHW03], where fully probabilistic processes are considered, we can associate to the probabilistic relation T the following linear operator:

$$\mathbf{M}(T) = \bigoplus_{a \in GAct} \mathbf{M}_a(T),$$

where for all $a \in GAct$ and $P, Q \in \mathcal{G}$, the matrix defined by:

$$(\mathbf{M}_a(T))_{PQ} = \sum \{q \mid \text{there exists } P \xrightarrow{a,\, q} Q \text{ and } q \neq 0\}$$

represents a one step transition on action $a \in GAct$ in the transition system $(\mathcal{G}, GAct, T)$. Note that if there are no transitions from P to Q then $(\mathbf{M}_a(T))_{PQ} = 0$, as the sum over an empty set is 0. The symbol $\oplus$ represents the direct sum operation defined for a given set $\{\mathbf{M}_i\}_{i=1}^k$ of

$n_i \times m_i$ matrices by the $(\sum_{i=1}^{k} n_i) \times (\sum_{i=1}^{k} m_i)$ matrix:

$$\bigoplus_i \mathbf{M}_i = \begin{pmatrix} \mathbf{M}_1 & 0 & 0 \dots & 0 \\ 0 & \mathbf{M}_2 & 0 \dots & 0 \\ \vdots & \vdots & \vdots \ddots & \vdots \\ 0 & 0 & 0 \dots & \mathbf{M}_k \end{pmatrix}.$$

For a given process P we denote by $\mathbf{M}(P)$ the associated transition matrix. This is a $n \times n$ matrix, where n is the cardinality of the set $S_P \subseteq \mathcal{G}$ of the processes which can be reached by a computation starting from P.

These linear operators are defined on the vector space $\mathcal{V}(\mathcal{G})$ containing all distributions over the set of states (processes) $\mathcal{G}$, which is the space of all formal linear combinations of elements in $\mathcal{G}$ with coefficients in $\mathbb{R}$. For the sake of simplicity we assume that the set $\mathcal{G}$ is finite so that we can restrict ourself, for the time being, to consider only finite-dimensional vector spaces and finite-dimensional linear operators. The results we present can nevertheless be extended to the general infinite case along the lines of [DHW03].

Our aim is to re-formulate the weak probabilistic bisimulation semantics introduced in Section 2.3 in terms of the above defined linear operators $\mathbf{M}(P)$ representing the operational semantics of a process P. In order to take into account transitions involving sequences of τ actions (the Milner "double arrow relation") we extend the single step operator $\mathbf{M} = \bigoplus_{a \in GAct} \mathbf{M}_a$ to encode transitions on strings $\sigma \in GAct^*$ of finite length as follows:

$$\mathbf{S}_\sigma = \mathbf{M}_{a_1}\mathbf{M}_{a_2} \dots \mathbf{M}_{a_n}, \text{ and } \mathbf{S}_\varepsilon = \mathbf{I},$$

where ε denotes the empty sequence in $GAct^*$.

In order to express the condition for weak probabilistic bisimulation in Definition 2.3 (first equation), we now need to define an operator which encodes probabilities of the form $Prob(P, \tau^*a, C)$, with C an equivalence class in a given relation $R \subseteq \mathcal{G} \times \mathcal{G}$. As shown in [DHW03], equivalence relations are in a one-to-one correspondence with a particular class of linear operators called *classification operators*: if $\approx$ is an equivalence relation on a set X, then there is a classification operator $\mathbf{K} : \mathcal{V}(X) \to \mathcal{V}(X/_{\approx})$, and vice versa. Note that $\mathbf{K}$ is a $n \times m$ matrix, where n is the cardinality of the set X (i.e. the dimension of the space $\mathcal{V}(X)$) and m is the number of $\approx$-equivalence classes in the partition of X (i.e. the dimension of the space $\mathcal{V}(X/_{\approx})$).

By using classification matrices we can now express the probability $Prob(P, \sigma, C)$, where $\sigma \in GAct^*$ and $C \in \mathcal{G}/R$ for some equivalence relation represented by $\mathbf{K}$, by the operator $\mathbf{S}_\sigma\mathbf{K}$. In particular, we have that $Prob(P, \sigma, C) = (\mathbf{S}_\sigma\mathbf{K})_{P,C}$.

We can then re-phrase Lemma 2.2 by stating that probabilities $Prob(P, \tau^*a, C)$, with $a \in GAct$, resolve the equation system:

$$(\mathbf{S}_{\tau^*a}\mathbf{K})_{P,C} = \begin{cases} 1 & \text{if } a = \tau \wedge P \in C \\ (\mathbf{M}_\tau \cdot \mathbf{S}_{\tau^*}\mathbf{K})_{P,C} & \text{if } a = \tau \wedge P \notin C \\ (\mathbf{M}_\tau \cdot \mathbf{S}_{\tau^*a}\mathbf{K})_{P,C} + (\mathbf{M}_a)_{P,C} & \text{if } a \neq \tau \end{cases}$$

For $\sigma = a_1a_2 \dots a_n \in GAct^n$, we denote by $\mathbf{P}_\sigma(P, \mathbf{K}_P)$ the operator $\mathbf{S}(P)_\sigma\mathbf{K}_P$, where

$\mathbf{S}(P)_\sigma = \mathbf{M}(P)_{a_1}\mathbf{M}(P)_{a_2}\ldots\mathbf{M}(P)_{a_n}$ is the restriction of $\mathbf{S}$ to the states reachable by P. The matrix $\mathbf{K}_P$ is a restricted matrix $\mathbf{K}$, where all the rows corresponding to states which are not reachable by P have been eliminated. So $\mathbf{K}_P$ is a $n_P \times m$ matrix, where n_P is the number of states reachable by P. We will also denote by $\mathbf{P}_{\tau^*a}(P, \mathbf{K}_P)$ the operator $\mathbf{S}_{\tau^*a}\mathbf{K}$ restricted to P.

Given a $n \times m$ matrix $\mathbf{K}$ and $n' \geq n$ we define the *completion* to n', $\bar{\mathbf{K}}$, of $\mathbf{K}$ as the $n' \times m$ matrix:

$$\bar{\mathbf{K}} = \mathbf{K} \oplus \mathbf{O}_{n'-n},$$

where $\mathbf{O}_k$ indicates the k-dimensional null matrix, that is the $k \times k$ matrix with only zero entries. We will use this operation to define a process which operates on the same number of abstract states (or classes, represented by the columns) as $\mathbf{K}$ but without any transitions between the "extra" $n' - n$ states.

The weak probabilistic bisimulation relation introduced in [BH97] can now be formulated in a linear operator setting as follows:

Definition 5.1 Let $P, Q \in \mathcal{G}$ be two processes and let n_P and n_Q be the number of states reachable by P and Q respectively. Then P and Q are weak probabilistic bisimilar iff there exists an $n \times m$ classification matrix $\mathbf{K}$ with $n = n_P + n_Q$ such that

$$\bar{\mathbf{P}}_{\tau^*a}(P, \mathbf{K}_P) = \bar{\mathbf{P}}_{\tau^*a}(Q, \mathbf{K}_Q) \text{ for all } a \in GAct,$$

where $\bar{\mathbf{P}}_{\tau^*a}(P, \mathbf{K}_P)$ and $\bar{\mathbf{P}}_{\tau^*a}(Q, \mathbf{K}_Q)$ are the completions to n of $\mathbf{P}_{\tau^*a}(P, \mathbf{K}_P)$ and $\mathbf{P}_{\tau^*a}(Q, \mathbf{K}_Q)$ respectively.

This definition corresponds to (the first equation of) Definition 2.3 in the restricted case of fully specified processes.

5.2 A Metric for Weak Probabilistic Bisimulation

In this section we show that the quantity $\varepsilon_{\bar{p}}$ introduced in Section 4 as a measure for the confinement of a given system, corresponds to the notion of distance induced by a particular *operator norm*. In general, the norm of an operator describes the maximal "stretching factor" of normalised vectors. We formally define it after recalling the basic definition of a *vector norm*:

Definition 5.2 A *norm* on a vector space $\mathcal{V}$ is a map $||.|| : \mathcal{V} \mapsto \mathbb{R}$ such that:

(i) $||\vec{x}|| \geq 0$,

(ii) $||\vec{x}|| = 0 \Leftrightarrow \vec{x} = \vec{o}$,

(iii) $||\alpha\vec{x}|| = |\alpha|||\vec{x}||$,

(iv) $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$,

with $\vec{o} \in \mathcal{V}$ the null vector.

Given a normed vector space $\mathcal{V}$ the *operator norm* for linear operators on $\mathcal{V}$ is defined by:

$$||\mathbf{M}|| = \sup_{\vec{x} \in \mathcal{V}} \frac{||\mathbf{M}(\vec{x})||}{||\vec{x}||} = \sup_{||\vec{x}||=1} ||\mathbf{M}(\vec{x})||.$$

The exact numerical value of an operator norm depends, of course, on the particular vector norm used. We will consider the *supremum norm* defined by

$$\|\vec{x}\|_\infty = \|(x_i)_i\|_\infty = \sup_i |x_i|,$$

and use the corresponding operator norm to define a metric on the set of linear operators representing the semantics of our probabilistic processes.

Definition 5.3 Let $\mathcal{S}$ be the set

$$\mathcal{S} = \{\mathbf{L} : \mathcal{V}(\mathcal{G}) \to \mathcal{V}(\mathcal{G}/R) \mid R \text{ is an equivalence relation}\}.$$

We define the metric d on $\mathcal{S}$ as the metric induced by the supremum operator norm:

$$d(\mathbf{L}_1, \mathbf{L}_2) = \|\mathbf{L}_1 - \mathbf{L}_2\|_\infty.$$

This metric is particularly suited for expressing the notion of behavioural distance introduced in Section 4. In particular, it can be used to calculate the values of the quantity $\delta_{\bar{p}}^R$ of Definition 4.1, as shown in the following.

We recall that in order to verify the *PNI* property for a given a system P, we check whether the two models $P_1 \equiv P \backslash High$ and $P_2 \equiv P/^{p_1}_{h_1^P} \dots /^{p_n}_{h_n^P}$ are weak probabilistic bisimilar for all probabilities $\bar{p} = p_1, \dots, p_n$. This equivalence checking can be performed by comparing for each equivalence relation R on the set $S = S_{P_1} \cup S_{P_2}$ of the states reached by both P_1 and P_2 (or, equivalently, for each classification matrix $\mathbf{K}$ on the vector space $\mathcal{V}(S)$) the operators $\bar{\mathbf{P}}_{\tau^* a}(P_1, \mathbf{K}_{P_1})$ and $\bar{\mathbf{P}}_{\tau^* a}(P_2, \mathbf{K}_{P_2})$, cf. Definition 5.1. If we use the distance d of Definition 5.3 then the outcome of this comparison is a numerical quantity which coincides with the measure $\varepsilon_{\bar{p}}$.

Proposition 5.4 *Let $P \in \mathcal{G}$, $P_1 \equiv P \backslash High$ and $P_2 \equiv P/^{p_1}_{h_1^P} \dots /^{p_n}_{h_n^P}$, with $\bar{p} = p_1, \dots, p_n$ a sequence of parameters such that $p_i \in (0,1), 1 \le i \le n$. Let $R \subseteq \mathcal{G} \times \mathcal{G}$ be an equivalence relation on the set $S = S_{P_1} \cup S_{P_2} \subseteq \mathcal{G}$, and let $\mathbf{K}$ be the corresponding classification matrix. Then*

$$\delta_{\bar{p}}^R = \max_{a \in GAct} \; d(\bar{\mathbf{P}}_{\tau^* a}(P_1, \mathbf{K}_{P_1}), \bar{\mathbf{P}}_{\tau^* a}(P_2, \mathbf{K}_{P_2})),$$

where $\mathbf{K}_{P_1}$ and $\mathbf{K}_{P_2}$ are the restrictions of $\mathbf{K}$ to the states in S_{P_1} and S_{P_2} respectively.

Proof. Let $\vec{x}$ be a normalised vector in $\mathcal{V}(S)$. Then the vector

$$\bar{\mathbf{P}}_{\tau^* a}(P_1, \mathbf{K}_{P_1})\vec{x} - \bar{\mathbf{P}}_{\tau^* a}(P_2, \mathbf{K}_{P_2})\vec{x}$$

is the vector in $\mathcal{V}(S/R)$ whose components are the difference of the probabilities of going from each state in S to each equivalence class via the sequence $\tau^* a$, for a given $a \in GAct$. Now observe that the choice of vector $\vec{x}$ corresponds to the choice of the pair (P', P'') in Definition 4.1. Moreover, since $GAct$ is finite, the least upper bound is actually the maximal element of the set. Then the thesis follows from the definition of the supremum operator norm. □

If for some $\mathbf{K}$ this distance is zero for all sequences $\bar{p}$, then we can conclude that P is secure (according to Definition 5.1 and the definition of PNI property). Otherwise, we can consider the relation $\mathbf{K}$ which minimises the distance $d(\bar{p})$. This corresponds to the quantity $\varepsilon_{\bar{p}}$ in Definition 4.2:

Corollary 5.5 *Let P be a process and let R, $\bar{p}$ as in Proposition 5.4. Then*
$$\varepsilon_{\bar{p}} = \inf_{\mathbf{K}} d(\bar{\mathbf{P}}_{\tau^* a}(P_1, \mathbf{K}_{P_1}), \bar{\mathbf{P}}_{\tau^* a}(P_2, \mathbf{K}_{P_2})).$$

6 Conclusion and Related Work

We have introduced a formal definition of the amount of information flowing in a system from High to Low, based on a notion of process similarity corresponding to an approximate probabilistic version of the weak bisimulation of [Mil89]. Our approach is able to detect and measure probabilistic covert channels from High to Low, by comparing the effect on the low-level view of the absence/presence of the high-level user.

A different approach aiming to the same objective of quantifying information flow has been proposed in [Lowe02], where the "quantity" is defined in terms of the behaviours of the high-level user that are distinguishable from a low-level user point of view. This approach does not consider probabilistic behaviours; instead it relies on a worst case analysis based on all possible ways to resolve nondeterminism.

Another related approach is the one presented in [CHM01], where the amount of confidential information which may be leaked by programs written in a simple imperative language is analysed by using Shannon's information theory.

Desharnais et al. [D^+02] propose an extension of [D^+00] by defining a fixed-point characterisation of a pseudometric for approximating weak bisimulation, where zero distance corresponds to weak bisimilarity. A quantitative meaning of such a metric is that nearby processes have nearby propensities to leak information. The metric is defined in the context of the alternating model for labelled concurrent Markov chains (LCMCs). In a LCMC states are either probabilistic or nondeterministic. Transitions from probabilistic states are not labelled and are associated with a probability distribution on the set of reachable nondeterministic states. Transitions from nondeterministic states are labelled and finitely branching and lead to probabilistic states. The probability of reaching a state through a weak transition is computed by taking the supremum over all possible computations, where a computation is a purely probabilistic transition system obtained by resolving the nondeterministic choices as follows: for each nondeterministic state at most one outgoing transition is picked up. With respect to such a framework, in our model states can have both probabilistic and nondeterministic choices, depending on the nature of the interactions with the environment. Then, before computing the distance between different processes, nondeterminism is resolved through a probabilistic scheduler rather than in a deterministic way. The notion of pseudometric of [D^+02] is strictly related to the notion of channel capacity from information theory [CT91]. On the other hand, our notion of confinement between processes provides a natural statistical interpretation, in terms of number of experiments that are needed on average to distinguish confined processes [A^+03]. Moreover, we have shown that our measure of the security degree of systems has a formal interpretation as the metric induced by the norm of a linear operator associated to our probabilistic processes.

Other works deal with approximate reasoning in order to obtain a relaxed notion of truth - the goal is moving from a qualitative scenario where $\{0, 1\}$ are the unique truth values to a quantitative scenario where the interval $[0, 1]$ is given as the collection of the truth values. Along this line, several approaches are investigated in [Koz81,JS90,BW01], which are not related to information flow capacity issues. Finally, in [L+98] an asymptotic notion of probabilistic equivalence is defined to estimate a secrecy property. In particular, in the setting of a variant of spi-calculus, observational equivalence is expressed in terms of indistinguishability by polynomial time statistical tests. Then, secrecy is checked by verifying whether the protocol under analysis is observationally equivalent to an idealized protocol. With respect to our framework, probabilities are intended as a means for a polynomial time treatment of cryptographic primitives and do not come into play in the modelling of the protocol behaviour.

References

[A+03] A. Aldini, M. Bravetti, A. Di Pierro, R. Gorrieri, C. Hankin, and H. Wiklicky. Two Formal Approaches for Approximating Noninterference Properties. *Foundations of Security Analysis and Design II – Tutorial Lectures*, Springer LNCS 2946:1–43, 2004.

[ABG03] A. Aldini, M. Bravetti, and R. Gorrieri. A Process-algebraic Approach for the Analysis of Probabilistic Noninterference. *Journal of Computer Security* 12(2), 2004.

[BH97] C. Baier and H. Hermanns. Weak Bisimulation for Fully Probabilistic Processes. In Proc. of *9th Int. Conf. on Computer Aided Verification (CAV'97)*, Springer LNCS 1254:119–130, 1997.

[BA03] M. Bravetti and A. Aldini. Discrete Time Generative-reactive Probabilistic Processes with Different Advancing Speeds. *Theoretical Computer Science* 290(1):355–406, 2003.

[BB00] M. Bravetti and M. Bernardo. Compositional Asymmetric Cooperations for Process Algebras with Probabilities, Priorities, and Time. In Proc. of *1st Workshop on Models for Time-Critical Systems (MTCS'00)*, ENTCS 39(3), 2000.

[BW01] F. van Breugel and J. Worrell. Towards Quantitative Verification of Probabilistic Systems (extended abstract). In Proc. of *28th International Colloquium on Automata, Languages and Programming (ICALP'01)*, Springer LNCS 2076:421–432, 2001.

[CHM01] D. Clark, S. Hunt, and P. Malacaria. Quantitative Analysis of the Leakage of Confidential Data. In ENTCS 59(3), (A. Di Pierro and H. Wiklicky, Eds.), Elsevier Science Publishers, 2002.

[CT91] T. Cover and J. Thomas. *Elements of Information Theory.* John Wiley, New York, 1991.

[D+00] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximation of Labeled Markov Processes. In Proc. of *15th Symposium on Logic in Computer Science (LICS'00)*, pp. 95–106, IEEE CS Press, 2000.

[D+02] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. The Metric Analogue of Weak Bisimulation for Probabilistic Processes. In Proc. of *17th Symposium on Logic in Computer Science (LICS'02)*, pp. 413–422, IEEE CS Press, 2002.

[DHW03] A. Di Pierro, C. Hankin, and H. Wiklicky. Quantitative Relations and Approximate Process Equivalences. In Proc. of *14th Int. Conf. on Concurrency Theory (CONCUR'03)*, Springer LNCS 2761:508–522, 2003.

[DHW03a] A. Di Pierro, C. Hankin, and H. Wiklicky. Measuring the Confinement of Concurrent Probabilistic Systems. In Proc. of *WITS'03 – 2003 IFIP WG 1.7, ACM SIGPLAN and GI FoMSESS Workshop on Issues in the Theory of Security*, 2003.

[DHW02] A. Di Pierro, C. Hankin, and H. Wiklicky. Approximate Noninterference. In Proc. of *15th Computer Security Foundations Workshop (CSFW'02)*, pp. 3–17, IEEE CS Press, 2002.

[FG95] R. Focardi and R. Gorrieri. A Classification of Security Properties. *Journal of Computer Security* 3(1):5–33, 1995.

[FGM00] R. Focardi, R. Gorrieri, and F. Martinelli. Non Interference for the Analysis of Cryptographic Protocols. In Proc. of *27th Int. Colloquium on Automata, Languages and Programming (ICALP'00)*, Springer LNCS 1853:354–372, 2000.

[GSS95] R. J. van Glabbeek, S. A. Smolka, and B. Steffen. Reactive, Generative and Stratified Models of Probabilistic Processes. *Information and Computation* 121:59–80, 1995.

[GM82] J.A. Goguen and J. Meseguer. Security Policy and Security Models. In Proc. of *Symposium on Security and Privacy (SSP'82)*, pp. 11–20, IEEE CS Press, 1982.

[Gra90] J. W. Gray III. Probabilistic Interference. In Proc. of *Symposium on Security and Privacy (SSP'90)*, pp. 170–179, IEEE CS Press, 1990.

[JS90] C.-C. Jou and S. A. Smolka. Equivalences, Congruences, and Complete Axiomatizations for Probabilistic Processes. In Proc. of *1st Int. Conf. on Concurrency Theory (CONCUR'90)*, Springer LNCS 458:367–383, 1990.

[Koz81] D. Kozen. Semantics of Probabilistic Programs. *Journal of Computer and Systems Sciences* 22(3):328–350, 1981.

[L^+98] P. Lincoln, J. C. Mitchell, M. Mitchell, and A. Scedrov. A Probabilistic Poly-Time Framework for Protocol Analysis. In Proc. of *5th Conf. on Computer and Communications Security*, pp. 112–121, ACM Press, 1998.

[Lowe02] G. Lowe. Quantifying Information Flow. In Proc. of *15th Computer Security Foundation Workshop (CSFW'02)*, pp. 18–31, IEEE CS Press, 2002.

[Mil89] R. Milner. *Communication and Concurrency*, Prentice Hall, 1989.

[R^+01] P.Y.A. Ryan, J. McLean, J. Millen, and V. Gligor. Non-interference: Who needs It? In Proc. of *14th Computer Security Foundations Workshop (CSFW'01)*, pp. 237–238, IEEE CS Press, 2001.

[RS01] P.Y.A. Ryan, S. Schneider. Process Algebra and Non-interference. *Journal of Computer Security* 9:75–103, 2001.

[Shao99] J. Shao. *Mathematical Statistics.* Springer Texts in Statistics, Springer Verlag, New York – Berlin – Heidelberg, 1999.

[Smith03] G. Smith. Probabilistic Noninterference through Weak Probabilistic Bisimulation. In Proc. of *16th Computer Security Foundations Workshop (CSFW'03)*, pp. 3–13, IEEE CS Press, 2003.