Metrics for Differential Privacy in Concurrent Systems*

Lili Xu[1,3,4,5], Konstantinos Chatzikokolakis[2,3], and Huimin Lin[4]

[1] INRIA, Paris, France
[2] CNRS, Paris, France
[3] Ecole Polytechnique, Paris, France
[4] Institute of Software, Chinese Academy of Sciences, Beijing, China
[5] Graduate University, Chinese Academy of Sciences, Beijing, China

Abstract. Originally proposed for privacy protection in the context of statistical databases, differential privacy is now widely adopted in various models of computation. In this paper we investigate techniques for proving differential privacy in the context of concurrent systems. Our motivation stems from the work of Tschantz et al., who proposed a verification method based on proving the existence of a stratified family between states, that can track the privacy leakage, ensuring that it does not exceed a given leakage budget. We improve this technique by investigating a state property which is more permissive and still implies differential privacy. We consider two pseudometrics on probabilistic automata: The first one is essentially a reformulation of the notion proposed by Tschantz et al. The second one is a more liberal variant, relaxing the relation between them by integrating the notion of amortisation, which results into a more parsimonious use of the privacy budget. We show that the metrical closeness of automata guarantees the preservation of differential privacy, which makes the two metrics suitable for verification. Moreover we show that process combinators are non-expansive in this pseudometric framework. We apply the pseudometric framework to reason about the degree of differential privacy of protocols by the example of the Dining Cryptographers Protocol with biased coins.

1 Introduction

Differential privacy [14] was originally proposed for privacy protection in the context of statistical databases, but nowadays it is becoming increasingly popular in many other fields, ranging from programming languages [24] to social networks [23] and geolocation [20]. One of the reasons of its success is its independence from side knowledge, which makes it robust to attacks based on combining various sources of information.

In the original definition, a query mechanism $\mathcal{A}$ is ϵ-differentially private if for any two databases u_1 and u_2 which differ only for one individual (one row), and any property Z, the probability distributions of $\mathcal{A}(u_1), \mathcal{A}(u_2)$ differ on Z at most by e^ϵ, namely, $\Pr[\mathcal{A}(u_1) \in Z] \leq e^\epsilon \cdot \Pr[\mathcal{A}(u_2) \in Z]$. This means that the presence (or the data) of an individual cannot be revealed by querying the database. In [7], the principle of

* This work has been partially supported by the project ANR-12-IS02-001 PACE, by the project ANR-11-IS02-0002 LOCALI, by the INRIA Large Scale Initiative CAPPRIS and by the National Natural Science Foundation of China (Grant No.60833001).

E. Ábrahám and C. Palamidessi (Eds.), FORTE 2014, LNCS 8461, pp. 199–215, 2014.

differential privacy has been formally extended to measure the degree of protection of secrets in more general settings.

In this paper we deal with the problem of verifying differential privacy properties for concurrent systems, modeled as probabilistic automata admitting both nondeterministic and probabilistic behavior. In such systems, reasoning about the probabilities requires *solving* the nondeterminism first, and to such purpose the usual technique is to consider functions, called *schedulers*, which select the next step based on the history of the computation. However, in our context, as well as in security in general, we need to restrict the power of the schedulers and make them unable to distinguish between secrets in the histories, or otherwise they would plainly reveal them by their choice of the step. See for instance [6,8,2] for a discussion on this issue. Thus we consider a restricted class of schedulers, called *admissible schedulers*, following the definition of [2]. Admissibility is introduced to deal with bisimulation-like notions in security contexts: Two bisimilar processes are typically considered to be indistinguishable, yet an unrestricted scheduler could trivially separate them.

The property of differential privacy requires that the observations generated by two different secret values be probabilistically similar. In standard concurrent systems the notion of similarity is usually formalized as an equivalence, preferably preserved under composition, i.e., a congruence. We mention in particular trace equivalence and bisimulation. The first is often used for its simplicity, but in general is not compositional [17]. The second one is a congruence and it is appealing for its proof technique. Process equivalences have been extensively used to formalize security properties like secrecy [1] and noninterference [15,25,26].

In this paper we focus on metrics suitable for verifying differential privacy. Namely, metrics for which the distance between two processes determines an upper bound on the ratio of the probabilities of the respective observables. We start by considering the framework proposed by Tschantz et al. [27], which was explicitly designed for the purpose of verifying differential privacy. Their verification technique is based on proving the existence of an indexed family of bijections between states. The parameter of the starting states, representing the privacy budget, determines the level of differential privacy of the system, which decreases over time by subtracting the absolute difference of probabilities in each step during mutual simulation. Once the balance reaches zero, processes must behave exactly the same. We reformulate this notion in the form of a pseudometric, showing some novel properties as a distance relation.

The above technique is sound, but has a rather rigid budget management. The main goal of this paper is to make the technique more permissive by identifying a pseudometric that is more relaxed and still implies an upper bound on the privacy leakage.

In particular, the pseudometric we propose is based on a thriftier use of the privacy budget, which is inspired by the notion of *amortisation* used in some quantitative bisimulations [18,10]. The idea is that, when constructing the bijections between states, the differences among the probabilities of related states are kept with their sign, and added with their sign through each step. In this way, successive differences can compensate (amortise) each other, and rather than always being consumed, the privacy budget may also be refurbished. In [18] the idea of amortisation is applied on a set of cost-based

actions. The quantitative feature considered here is discrete probability distributions over states, which is shown to benefit from the theory of amortisation as well.

Furthermore, there is a soundness criterion on the distance notion for probabilistic concurrent systems defined in [13]. It says that 0-distance in a pseudometric is expected to fully characterise bisimilarity. We show that 0-distance in the two pseudometrics implies bisimilarity while the converse does not hold. Although the pseudometrics do not thoroughly satisfy the criterion, we prove that several process combinators including parallel composition are non-expansive in the pseudometrics. Non-expansiveness gives a desirable property that when close processes are placed in the same context, the resulting processes are still close in the distance. This can be viewed as an analogue of the congruence properties of bisimulation. Finally, we illustrate the verification technique of differential privacy using the example of the Dining Cryptographers Problem(DCP) with biased coins.

More related Work. Verification of differential privacy has become an active area of research. Among the approaches based on formal methods, we mention those based on type-systems [24,16] and logical formulations [3].

In a previous paper [28], one of the authors has developed a compositional method for proving differential privacy in a probabilistic process calculus. The technique there is rather different from the ones presented in paper: the idea is based on decomposing a process in simpler processes, computing the level of privacy of these, and combining them to obtain the level of privacy of the original program.

A line of one very interesting approach related to ours in spirit - considering pseudometrics on probabilistic automata - includes the work by Desharnais et al. [13] and Deng et al. [11]. They both use the metric à la Kantorovich proposed in [13], which represents a cornerstone in the area of bisimulation metrics. It would be attractive to see how the Kantorovich metric can be adapted to reason about differential privacy.

Finally, among several formalizations of the notion of information protection based on probability theory, we mention some rather popular approaches, mainly based on information theory, in particular, to consider different notions of entropy depending on the kind of adversary, and to express the leakage of information in terms of the notion of mutual information. We name a few works also discussed in the models of probabilistic automata and process algebra: Boreale [4] establishs a framework for quantifying information leakage using absolute leakage, and introduces a notion of rate of leakage. Deng et al. [12] use the notion of relative entropy to measure the degree of anonymity. Compositional methods based on Bayes risk method are discussed by Braun et al. [5]. A metric for probabilistic processes based on the Jensen-Shannon divergence is proposed in [22] for measuring information flow in reactive processes. Unlike the information-theoretical approach, differential privacy provides strong privacy guarantees independently from side knowledge. However, progress for differential privacy has been relatively new and going slowly. It would be interesting to see how the issues stressed and the reasoning techniques developed there can be adapted for differential privacy.

Contribution. The main contributions of this paper can be summarized as follows:

- We reformulate the notion of approximate similarity proposed in [27] in terms of a pseudometric and we study the properties of the distance relation (in Section 3).

- We propose the second pseudometric which is more liberal than the former one, in the sense that the total differences of probabilities get amortised during the mutual simulation. We show that the level of differential privacy is continuous with respect to the metric, which says that if every two processes running on two adjacent secrets of a system are close in the metric then the system is differentially private, making the metric suitable for verification (in Section 4).
- We show that 0-distance in the pseudometrics implies bisimilarity (in Section 5).
- We present the non-expansiveness property in the pseudometrics for CCS_p operators in a probabilistic variant of Milner's CCS [21] (in Section 6).
- We use the pseudometric framework to show that the Dining Cryptographers protocol with probability-p biased coins is $|\ln \frac{p}{1-p}|$-differentially private. (in Section 7).

The rest of the Paper. In the next section we recall some preliminary notions about probabilistic automata, differential privacy and pseudometrics. Section 8 concludes. Long proofs can be found in the appendix.

2 Preliminaries

2.1 Probabilistic Automata

Given a set X, we denote by $Disc(X)$ the set of discrete sub-probability measures over X; the support of a measure μ is defined as $supp(\mu) = \{x \in X | \mu(x) > 0\}$. A *probabilistic automaton* (henceforth PA) $\mathcal{A}$ is a tuple $(S, \overline{s}, A, D)$ where S is a finite set of *states*, $\overline{s} \in S$ is the *start state*, A is a finite set of action *labels*, and $D \subseteq S \times A \times Disc(S)$ is a *transition relation*. We write $s \xrightarrow{a} \mu$ for $(s, a, \mu) \in D$, and we denote by $act(d)$ the action of the transition $d \in D$. A PA $\mathcal{A}$ is *fully probabilistic* if from each state of $\mathcal{A}$ there is at most one transition available.

An *execution* α of a PA is a (possibly infinite) sequence $s_0 a_1 s_1 a_2 s_2 \ldots$ of alternating states and labels, such that for each $i : s_i \xrightarrow{a_{i+1}} \mu_{i+1}$ and $\mu_{i+1}(s_{i+1}) > 0$. We use $lstate(\alpha)$ to denote the last state of a finite execution α. We use $Exec^*(\mathcal{A})$ and $Exec(\mathcal{A})$ to represent the set of finite and all executions of $\mathcal{A}$, respectively. A *scheduler* of a PA $\mathcal{A} = (S, \overline{s}, A, D)$ is a function $\zeta : Exec^*(\mathcal{A}) \mapsto D$ such that $\zeta(\alpha) = s \xrightarrow{a} \mu \in D$ implies that $s = lstate(\alpha)$. The idea is that a scheduler selects a transition among the ones available in D, basing its decision on the history of the execution. The *execution tree* of $\mathcal{A}$ with respect to the scheduler ζ, denoted by $\mathcal{A}_\zeta$, is a fully probabilistic automaton $(S', \overline{s}', A', D')$ such that $S' \subseteq Exec^*(\mathcal{A})$, $\overline{s}' = \overline{s}$, $A' = A$, and $\alpha \xrightarrow{a} \nu \in D'$ if and only if $\zeta(\alpha) = lstate(\alpha) \xrightarrow{a} \mu$ for some μ and $\nu(\alpha a s) = \mu(s)$. Intuitively, $\mathcal{A}_\zeta$ is produced by unfolding the executions of $\mathcal{A}$ and resolving all non-deterministic choices using ζ. Note that $\mathcal{A}_\zeta$ is a simple and fully probabilistic automaton. We use α with primes and indices to range over states in an execution tree.

A *trace* is a sequence of labels in $A^* \cup A^\omega$ obtained from executions by removing states. We use $[\,]$ to represent the empty trace, and $\frown$ to concatenate two traces. A state α of $\mathcal{A}_\zeta$ induces a probability measure over traces as follows. The basic measurable events are the cones of finite traces, where the cone of a finite trace $\boldsymbol{t}$, denoted by $C_{\boldsymbol{t}}$, is the set $\{\boldsymbol{t}' \in A^* \cup A^\omega | \boldsymbol{t} \leq \boldsymbol{t}'\}$, where $\leq$ is the standard prefix preorder on sequences.

The probability of a cone $C_{\boldsymbol{t}}$ induced by state α, denoted by $\Pr_\zeta[\alpha \rhd \boldsymbol{t}]$, is defined recursively as follows.

$$\Pr_\zeta[\alpha \rhd \boldsymbol{t}] = \begin{cases} 1 & \text{if } \boldsymbol{t} = [\,], \\ 0 & \text{if } \boldsymbol{t} = a^\frown \boldsymbol{t}' \text{ and } act(\zeta(\alpha)) \neq a, \\ \sum_{s_i \in supp(\mu)} \mu(s_i) \Pr_\zeta[\alpha a s_i \rhd \boldsymbol{t}'] & \\ \qquad \text{if } \boldsymbol{t} = a^\frown \boldsymbol{t}' \text{ and } \zeta(\alpha) = s \xrightarrow{a} \mu. \end{cases} \tag{1}$$

Admissible schedulers. We consider a restricted class of schedulers, called *admissible schedulers*, following the definition of [2]. Essentially this definition requires that whenever given two *adjacent* states s, s', namely, differing only for the choice for some secret value, then the choice made by the scheduler on s and s' should be consistent, i.e. the scheduler should not be able to make a different choice on the basis of the secret. Note that in [27] scheduling is not an issue since non-determinism is not allowed.

More precisely, in [2] admissibility is achieved by introducing tags for transitions. Admissible schedulers are viewed as entities that have access to a system through a screen with buttons, where each button represents one (current) available option, i.e. an enabled tag. A scheduler ζ is admissible if for all finite executions having the same sequence of screens, ζ decides the same tagged transition for them.

Pseudometrics on states. A pseudometric[1] on S is a function $m : S^2 \to \mathbb{R}$ satisfying the following properties: $m(s,s) = 0$ (reflexivity), $m(s,t) = m(t,s)$ (symmetry) and $m(s,t) \leq m(s,u) + m(u,t)$ (triangle inequality). We define $m_1 \preceq m_2$ iff $\forall s,t : m_1(s,t) \geq m_2(s,t)$ (note that the order is reversed).

2.2 Differential Privacy

Differential privacy [14] was originally defined in the context of statistical databases, by requiring that a mechanism (i.e. a probabilistic query) gives similar answers on *adjacent* databases, that is those differing on a single row. More precisely, a mechanism $\mathcal{K}$ satisfies *ϵ-differential privacy* iff for all adjacent databases x, x': $\Pr[\mathcal{K}(x) \in Z] \leq e^\epsilon \cdot \Pr[\mathcal{K}(x') \in Z]$ for all $Z \subseteq range(\mathcal{K})$. Differential privacy imposes looser restrictions on non-adjacent secrets, which is considered as another merit of it.

In this paper, we study concurrent systems taking a secret as input and producing an observable trace as output. Let U be a set of secrets and $\sim$ an adjacency relation on U, where $u \sim u'$ denotes the fact that two close secrets u, u' should not be easily distinguished by the adversary after seeing observable traces. A *concurrent system* $\mathcal{A}$ is a mapping of secrets to probabilistic automata, where $\mathcal{A}(u), u \in U$ is the automaton modelling the behaviour of the system when running on u. Differential privacy can be directly adapted to this context:

Definition 1 (Differential Privacy). *A concurrent system $\mathcal{A}$ satisfies* ϵ-differential privacy *(DP) iff for any $u \sim u'$, any finite trace $\boldsymbol{t}$ and any admissible scheduler ζ:*

$$\Pr_\zeta[\mathcal{A}(u) \rhd \boldsymbol{t}] \leq e^\epsilon \cdot \Pr_\zeta[\mathcal{A}(u') \rhd \boldsymbol{t}]$$

[1] Unlike a metric, points in a pseudometric need not be distinguishable; that is, one may have $m(s,t) = 0$ for distinct values $s \neq t$.

3 The Accumulative Pseudometric

In this section, we present the first pseudometric based on a reformulation of the relation family proposed in [27]. We reformulate their notion in the form of an approximate bisimulation relation, named *accumulative bisimulation*, and then use it to construct a pseudometric on the state space.

We start by defining an approximate lifting operation that lifts a relation over states to a relation over distributions. Intuitively, we use a parameter ϵ to represent the total privacy leakage budget. A parameter c ranging over $[0, \epsilon]$, starting from 0, records the current amount of leakage and increasing over time by adding the maximum absolute difference of probabilities, denoted by σ, in each step during mutual simulation. Once c reaches the budget bound ϵ, processes must behave exactly the same. Since the total bound is ϵ, only a total of ϵ privacy can be leaked, a fact that will be used later to verify differential privacy. We use D to simply differentiate notions of this section from the following sections.

Definition 2. *Let $\epsilon \geqslant 0$, $c \in [0, \epsilon]$, $\mathcal{R} \subseteq S \times S \times [0, \epsilon]$. The* D-approximate lifting *of $\mathcal{R}$ up to c, denoted by $\mathcal{L}^D(\mathcal{R}, c)$, is the relation on $Disc(S)$ defined as:*

$$\mu \mathcal{L}^D(\mathcal{R}, c)\mu' \quad \textit{iff} \quad \exists \textit{ bijection } \beta : supp(\mu) \to supp(\mu') \textit{ such that}$$
$$\forall s \in supp(\mu) : (s, \beta(s), c + \sigma) \in \mathcal{R} \quad \textit{where} \quad \sigma = \max_{s \in supp(\mu)} |\ln \frac{\mu(s)}{\mu'(\beta(s))}|$$

This lifting allows us to define an approximate bisimulation relation:

Definition 3 (Accumulative bisimulation). *A relation $\mathcal{R} \subseteq S \times S \times [0, \epsilon]$ is a* ϵ-accumulative bisimulation *iff for all $(s, t, c) \in \mathcal{R}$:*

1. *$s \xrightarrow{a} \mu$ implies $t \xrightarrow{a} \mu'$ with $\mu \mathcal{L}^D(\mathcal{R}, c)\mu'$*
2. *$t \xrightarrow{a} \mu'$ implies $s \xrightarrow{a} \mu$ with $\mu \mathcal{L}^D(\mathcal{R}, c)\mu'$*

We can now define a pseudometric based on accumulative bisimulation as:

$$m^D(s, t) = \min\{\epsilon \mid (s, t, 0) \in \mathcal{R} \text{ for some } \epsilon\text{-accumulative bisimulation } \mathcal{R}\}$$

Proposition 1. *m^D is a pseudometric, that is:*

1. *(reflexivity) $m^D(s, s) = 0$*
2. *(symmetry) $m^D(s_1, s_2) = m^D(s_2, s_1)$*
3. *(triangle inequality) $m^D(s_1, s_3) \leq m^D(s_1, s_2) + m^D(s_2, s_3)$*

Proof Sketch. The proof proceeds by showing for each clause respectively that: 1. $Id_S = \{(s, s, 0) | s \in S\}$ is a 0-accumulative bisimulation; 2. Assume that $(s_1, s_2, 0)$ is in a ϵ-accumulative bisimulation $\mathcal{R}$, then $\mathcal{R}' = \{(s_2', s_1', c) | (s_1', s_2', c) \in \mathcal{R}\}$ is a ϵ-accumulative bisimulation; 3. Assume that $(s_1, s_2, 0)$ is in the ϵ_1-accumulative bisimulation $\mathcal{R}_1 \subseteq S \times S \times [0, \epsilon_1]$, $(s_2, s_3, 0)$ is in the ϵ_2-accumulative bisimulation $\mathcal{R}_2 \subseteq S \times S \times [0, \epsilon_2]$. Their relational composition $\mathcal{R}_1 \mathcal{R}_2 \subseteq S \times S \times [0, \epsilon_1 + \epsilon_2]$:

$$\{(s_1', s_3', c) | \exists s_2', c_1, c_2 . (s_1', s_2', c_1) \in \mathcal{R}_1 \wedge (s_2', s_3', c_2) \in \mathcal{R}_2 \wedge c \leq c_1 + c_2\}$$

is a $\epsilon_1 + \epsilon_2$-accumulative bisimulation. □

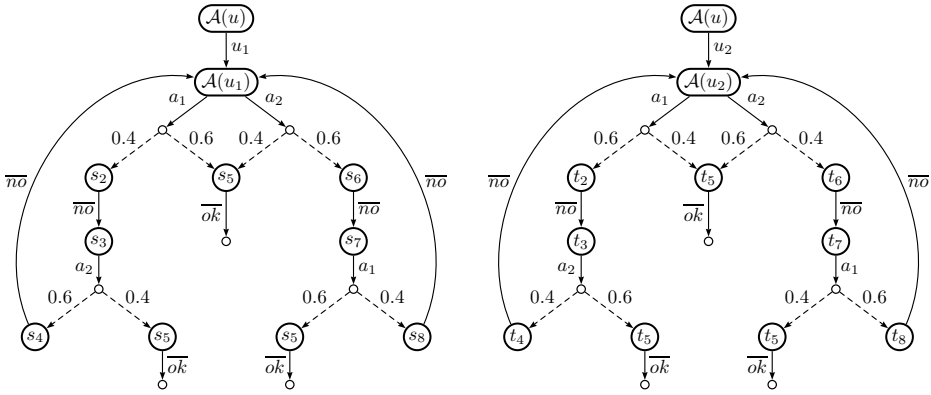


Fig. 1. A PIN-checking system: $m^D(\mathcal{A}(u_1), \mathcal{A}(u_2)) = \infty$, $m^A(\mathcal{A}(u_1), \mathcal{A}(u_2)) = \ln \frac{9}{4}$

Verification of differential privacy using m^D. As already shown in [27], the closeness of processes in the relation family implies a level of differential privacy. We here restate this result in terms of the metric m^D.

Lemma 1. *Given a PA $\mathcal{A}$, let $\mathcal{R}$ be a ϵ-accumulative bisimulation, $c \in [0, \epsilon]$, let ζ be an admissible scheduler, $\boldsymbol{t}$ be a finite trace, α_1, α_2 two finite executions of $\mathcal{A}$. If $(lstate(\alpha_1), lstate(\alpha_2), c) \in \mathcal{R}$, then*

$$\frac{1}{e^{\epsilon - c}} \leq \frac{\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}]}{\Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}]} \leq e^{\epsilon - c}$$

The above lemma shows that in a ϵ-accumulative bisimulation, two states related by a current leakage amount c, produce distributions over the same trace that only deviate by a factor $(\epsilon - c)$ representing the remaining amount of leakage. Then it is easy to get that the level of differential privacy is continuous with respect to m^D.

Theorem 1. *A concurrent system $\mathcal{A}$ is ϵ-differentially private if $m^D(\mathcal{A}(u), \mathcal{A}(u')) \leq \epsilon$ for all $u \sim u'$.*

4 The Amortised Pseudometric

As shown in the previous section, m^D is useful for verifying differential privacy. However, a drawback of this metric is that the definition of accumulative bisimulation is too restrictive: first, the amount of leakage is only accumulated, independently from whether the difference in probabilities is negative or positive. Moreover, the accumulation is the same for all branches, and equal to the worst branch, although the actual difference on some branch might be small. As a consequence, m^D is inapplicable in several systems, as shown by the following toy example.

Example 1. Consider a PIN-checking system $\mathcal{A}(u)$ in which the PIN variable u is designated from two secret codes u_1 and u_2. In order to protect the secrecy of the two PINs, rather than announcing to a user deterministically that whether the password he enters is correct or wrong, the system makes a response probabilistically. The idea is to give a positive answer with a higher probability when the password and the PIN match, and to give a negative answer with a higher probability otherwise.

The PIN-checking system could be defined as the PA shown in Fig. 1.We use label a_i to model the behavior that the password entered by a user is u_i, where $i \in \{1, 2\}$. We use label $\overline{ok}$ and $\overline{no}$ to represent a positive and a negative answer, respectively.

Consider an admissible scheduler always choosing for $\mathcal{A}(u_1)$ the a_1-branch (the case for the a_2-branch is similar), thus scheduling for $\mathcal{A}(u_2)$ also the a_1-branch. It is easy to see that the ratio of probabilities for $\mathcal{A}(u_1)$ and $\mathcal{A}(u_2)$ producing the same finite sequences $(a_1\overline{no}\,a_2\,\overline{no})^*$ is $(\frac{0.4\times0.6}{0.6\times0.4})^* = 1$. For the rest sequences $(a_1\overline{no}\,a_2\,\overline{no})^*a_1\overline{ok}$ and $(a_1\overline{no}\,a_2\,\overline{no})^*a_1\overline{no}\,a_2\,\overline{ok}$, we can check that the ratios are bounded by $\frac{9}{4}$. Thus, $\mathcal{A}$ satisfies $\ln\frac{9}{4}$-differential privacy. However, we can not find an accumulative bisimulation with a bounded ϵ between $\mathcal{A}(u_1)$ and $\mathcal{A}(u_2)$. The problem lies in that the leakage amount is always accumulated by adding the absolute differences during cyclic simulations, resulting in a convergence to ∞.

In order to obtain a more relaxed metric, we employ the *amortised bisimulation* relation of [18,10]. The main intuition behind this notion is that the privacy leakage amount in each simulation step may be either reduced due to a negative difference of probabilities, or increased due to a positive difference. Hence, the long-term budget gets amortised, in contrast to the accumulative bisimulation in which the budget is always consumed. We start by defining the corresponding lifting, using A to represent amortised bisimulation-based notions. Note that the current leakage c ranges over $[-\epsilon, \epsilon]$.

Definition 4. *Let $\epsilon \geqslant 0$, $c \in [-\epsilon, \epsilon]$, $\mathcal{R} \subseteq S \times S \times [-\epsilon, \epsilon]$. The* A-approximate lifting *of $\mathcal{R}$ up to c, denoted by $\mathcal{L}^A(\mathcal{R}, c)$, is a relation on $Disc(S)$ defined as:*

$$\mu\mathcal{L}^A(\mathcal{R}, c)\mu' \quad \textit{iff} \quad \exists \textit{ bijection } \beta : supp(\mu) \to supp(\mu') \textit{ such that}$$
$$\forall s \in supp(\mu) : (s, \beta(s), c + \ln\frac{\mu(s)}{\mu'(\beta(s))}) \in \mathcal{R}$$

Note that if $\ln\frac{\mu(s)}{\mu'(\beta(s))}$ is positive, then after this mutual step, the current leakage for s and $\beta(s)$ gets increased, otherwise decreased. We are now ready to define amortised bisimulation.

Definition 5 (Amortised bisimulation). *A relation $\mathcal{R} \subseteq S \times S \times [-\epsilon, \epsilon]$ is a* ϵ-amortised bisimulation *iff for all $(s, t, c) \in \mathcal{R}$:*

1. *$s \xrightarrow{a} \mu$ implies $t \xrightarrow{a} \mu'$ with $\mu\mathcal{L}^A(\mathcal{R}, c)\mu'$*
2. *$t \xrightarrow{a} \mu'$ implies $s \xrightarrow{a} \mu$ with $\mu\mathcal{L}^A(\mathcal{R}, c)\mu'$*

Similarly to the previous section, we can finally define a pseudometric on states as:

$$m^A(s, t) = \min\{\epsilon \mid (s, t, 0) \in \mathcal{R} \text{ for some } \epsilon\text{-amortised bisimulation } \mathcal{R}\}$$

Proposition 2. *m^A is a pseudometric.*

Proof Sketch. The proof proceeds by showing that: 1. $Id_S = \{(s,s,0)|s \in S\}$ is a 0-amortised bisimulation; 2. Assume that $(s_1,s_2,0)$ is in a ϵ-amortised bisimulation $\mathcal{R}$, then $\mathcal{R}' = \{(s_2',s_1',c)\ |(s_1',s_2',-c) \in \mathcal{R}\}$ is a ϵ-amortised bisimulation; 3. Let $(s_1,s_2,0)$ be in the ϵ_1-amortised bisimulation $\mathcal{R}_1 \subseteq S \times S \times [-\epsilon_1,\epsilon_1]$, $(s_2,s_3,0)$ be in the ϵ_2-amortised bisimulation $\mathcal{R}_2 \subseteq S \times S \times [-\epsilon_2,\epsilon_2]$. Their relational composition $\mathcal{R}_1\mathcal{R}_2 \subseteq S \times S \times [-\epsilon_1-\epsilon_2, \epsilon_1+\epsilon_2]$:

$$\{(s_1',s_3',c)|\exists s_2', c_1, c_2.(s_1',s_2',c_1) \in \mathcal{R}_1 \wedge (s_2',s_3',c_2) \in \mathcal{R}_2 \wedge c_1+c_2=c\}$$

is a $\epsilon_1+\epsilon_2$-amortised bisimulation. □

Verification of differential privacy using m^A. We now show that m^A can be used to verify differential privacy.

Lemma 2. *Given a PA $\mathcal{A}$, let $\mathcal{R}$ be a ϵ-amortised bisimulation, $c \in [-\epsilon,\epsilon]$, let ζ be an admissible scheduler, $\boldsymbol{t}$ be a finite trace, α_1,α_2 two finite executions of $\mathcal{A}$. If $(lstate(\alpha_1), lstate(\alpha_2), c) \in \mathcal{R}$, then*

$$\frac{1}{e^{\epsilon+c}} \leq \frac{\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}]}{\Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}]} \leq e^{\epsilon-c}$$

Note that there is a subtle difference between Lemmas 1 and 2, in that the denominator in the left-hand bound is $e^{\epsilon+c}$ instead of $e^{\epsilon-c}$. This comes from the amortised nature of $\mathcal{R}$. We can now show that differential privacy is continuous with respect to m^A as well.

Theorem 2. *A concurrent system $\mathcal{A}$ is ϵ-differentially private if $m^A(\mathcal{A}(u),\mathcal{A}(u')) \leq \epsilon$ for all $u \sim u'$.*

Proof Sketch. Since $m^A(\mathcal{A}(u),\mathcal{A}(u')) \leq \epsilon$ for all $u \sim u'$, by the definition of m^A, there exists a ϵ-amortised bisimulation $\mathcal{R}$ such that $(\mathcal{A}(u),\mathcal{A}(u'),0) \in \mathcal{R}$. By Lemma 2, for any admissible scheduler ζ, any finite trace $\boldsymbol{t}$:

$$\frac{1}{e^{\epsilon}} \leq \frac{\Pr_\zeta[\mathcal{A}(u) \rhd \boldsymbol{t}]}{\Pr_\zeta[\mathcal{A}(u') \rhd \boldsymbol{t}]} \leq e^{\epsilon}$$

Thus, we obtain that $\mathcal{A}$ is ϵ-differentially private. □

Example 2 (Example 1 revisited). Consider again the concurrent system shown in Fig. 1. Let S and T denote the state space of $\mathcal{A}(u_1)$ and $\mathcal{A}(u_2)$, respectively. Let $\mathcal{R} \subseteq S \times T \times [\ln\frac{4}{9}, \ln\frac{9}{4}]$. It is straightforward to check according to Def. 5 that the following relation is an amortised bisimulation between $\mathcal{A}(u_1)$ and $\mathcal{A}(u_2)$.

$$\begin{aligned}\mathcal{R} = \{\ & (\mathcal{A}(u_1),\mathcal{A}(u_2),0),\\ & (s_2,t_2,\ln\tfrac{2}{3}),(s_5,t_5,\ln\tfrac{3}{2}),(s_3,t_3,\ln\tfrac{2}{3}),(s_4,t_4,0),(s_5,t_5,\ln\tfrac{4}{9}),\\ & (s_6,t_6,\ln\tfrac{3}{2}),(s_5,t_5,\ln\tfrac{2}{3}),(s_7,t_7,\ln\tfrac{3}{2}),(s_8,t_8,0),(s_5,t_5,\ln\tfrac{9}{4})\ \}\end{aligned}$$

Thus $m^A(\mathcal{A}(u_1),\mathcal{A}(u_2)) \leq \ln\frac{9}{4}$. By Theorem 2, $\mathcal{A}$ is $\ln\frac{9}{4}$-differentially private.

5 Comparing the Two Pseudometrics

In this section, we formally compare the two metrics, showing that our pseudometric is indeed more liberal than the first one. Moreover,we investigate whether they can fully characterise bisimilarity. We show that m^D and m^A only imply bisimilarity, while the converse direction does not hold because of the strong requirement of the bijections in their definitions.

We show that m^A is bounded by m^D. Note the converse does not hold, since Examples 1 and 2 already show the cases in which m^D is infinite while m^A is finite.

Lemma 3. $m^D \preceq m^A$.

Proof Sketch. Let $\mathcal{R}^D \subseteq S \times S \times [0, \epsilon]$ be the ϵ-accumulative bisimulation such that $(s, t, 0) \in \mathcal{R}^D$. It is sufficient to show that the relation $\mathcal{R}^A \subseteq S \times S \times [-\epsilon, \epsilon]$ defined on the basis of $\mathcal{R}^D$ as follows is a ϵ-amortised bisimulation.

$$(s', t', c^A) \in \mathcal{R}^A \text{ iff } \exists c^D.(s', t', c^D) \in \mathcal{R}^D \wedge |c^A| \leq c^D$$

Relations with probabilistic bisimilarity. We adopt the notion of probabilistic bisimilarity which was first defined in [19]. An equivalence relation over S can be lifted to a relation over distributions over S by stating that two distributions are equivalent if they assign the same probability to the same equivalence class.

Formally, let $\mathcal{R} \subseteq S \times S$ be an equivalence relation. Two probability distributions μ_1 and μ_2 are $\mathcal{R}$-equivalent, written $\mu_1 \mathcal{L}(\mathcal{R}) \mu_2$, iff for every equivalence class $E \in S/\mathcal{R}$ we have $\mu_1(E) = \mu_2(E)$, in which $\mu_i(E) = \sum_{s \in E} \mu_i(s)$, $i = 1, 2$.

Definition 6. *An equivalence relation $\mathcal{R} \subseteq S \times S$ is a* strong bisimulation *iff for all $(s, t) \in \mathcal{R}$, $s \xrightarrow{a} \mu$ implies $t \xrightarrow{a} \mu'$ with $\mu \mathcal{L}(\mathcal{R}) \mu'$[2]. We write $s \sim t$ whenever there is a strong bisimulation that relates them. $\sim$ is the maximum strong bisimulation, namely* strong bisimilarity.

Proposition 3. *The following hold:*

- $m^D(s, t) = 0 \Rightarrow s \sim t$
- $m^A(s, t) = 0 \Rightarrow s \sim t$

The proofs are achieved by showing that the relation $\mathcal{R}$ induced by 0 distance in m^A (or m^D), namely, $(s, t) \in \mathcal{R}$ iff $m^A(s, t) = 0$, is a strong bisimulation.

6 Process Algebra

Process algebras provide the link to the desired compositional reasoning about approximate equality in such a pseudometric framework. We would like process operators to be *non-expansive* in the pseudometrics, which allows us to estimate the degree of differential privacy of a complex system from its components. In this section we consider a simple process calculus whose semantics is given by probabilistic automata. We define

[2] The converse is implied by the symmetry of the equivalence relation $\mathcal{R}$.

$$\text{ACT}\ \frac{}{\alpha.P \xrightarrow{\alpha} \delta(P)} \qquad \text{PROB}\ \frac{}{\bigoplus_{i\in I} p_i\, P_i \xrightarrow{\tau} \sum_i p_i\, P_i}$$

$$\text{SUM1}\ \frac{P \xrightarrow{\alpha} \mu}{P+Q \xrightarrow{\alpha} \mu} \qquad \text{PAR1}\ \frac{P \xrightarrow{\alpha} \mu}{P \,|\, Q \xrightarrow{\alpha} \mu \,|\, Q}$$

$$\text{COM}\ \frac{P \xrightarrow{a} \delta(P') \quad Q \xrightarrow{\overline{a}} \delta(Q')}{P \,|\, Q \xrightarrow{\tau} \delta(P' \,|\, Q')} \qquad \text{RES}\ \frac{P \xrightarrow{\alpha} \mu \qquad \alpha \neq a, \overline{a}}{(\nu a)P \xrightarrow{\alpha} (\nu a)\mu}$$

Fig. 2. The semantics of CCS_p. SUM1 and PAR1 have corresponding right rules SUM2 and PAR2, omitted for simplicity.

prefixing, non-deterministic choice, probabilistic choice, restriction and parallel composition constructors for the process calculus, and show that they are non-expansive in the sense that when neighboring processes are placed in the same context, the resulting processes are still neighboring.

The syntax of CCS_p is:

$$\begin{array}{rcll} \alpha & ::= & a \mid \overline{a} \mid \tau & \text{prefixes} \\ P, Q & ::= & \alpha.P \mid P \,|\, Q \mid P+Q \mid \bigoplus_{i\in 1..n} p_i P_i \mid (\nu a)P \mid \mathbf{0} & \text{processes} \end{array}$$

Here $\bigoplus_{i\in 1..n} p_i P_i$ stands for a probabilistic choice constructor, where the p_i's represent positive probabilities, i.e., they satisfy $p_i \in (0,1]$ and $\sum_{i\in 1..n} p_i = 1$. It may be occasionally written as $p_1 P_1 \oplus \cdots \oplus p_n P_n$. The rest constructors are the standard ones in Milner's CCS [21].

The semantics of a CCS_p term is a probabilistic automaton defined according to the rules in Fig. 2. We write $s \xrightarrow{a} \mu$ when (s, a, μ) is a transition of the probabilistic automaton. We also denote by $\mu|Q$ the measure μ' such that $\mu'(P|Q) = \mu(P)$ for all processes P and $\mu'(R) = 0$ if R is not of the form $P|Q$. Similarly $(\nu a)\mu = \mu'$ such that $\mu'((\nu a)P) = \mu(P)$. A transition of the form $P \xrightarrow{a} \delta(P')$, i.e. a transition having for target a Dirac measure, corresponds to a transition of a non-probabilistic automaton.

Proposition 4. *If $m(P,Q) \leq \epsilon$, where $m \in \{m^D, m^A\}$, then*

1. $m(a.P, a.Q) \leq \epsilon$
2. $m(pR \oplus (1-p)P, pR \oplus (1-p)Q) \leq \epsilon$
3. $m(R+P, R+Q) \leq \epsilon$
4. $m((\nu a)P, (\nu a)Q) \leq \epsilon$
5. $m(R \,|\, P, R \,|\, Q) \leq \epsilon$.

Proof sketch. The proof proceeds by finding a ϵ-accumulative (resp. amortised) bisimulation relation witnessing their distance in m not greater than ϵ. Let $\mathcal{R}$ be a ϵ-accumulative (resp. amortised) bisimulation relation witnessing $m(P,Q) \leq \epsilon$. Define

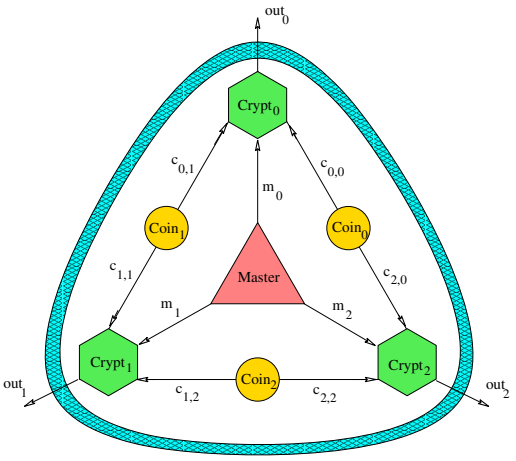


Fig. 3. Chaum's system for the Dining Cryptographers

the relation $Id_S = \{(s, s, 0)|s \in S\}$. We construct for each clause a relation $\mathcal{R}'$ as follows and show that it is a ϵ-accumulative (resp. amortised) bisimulation relation.

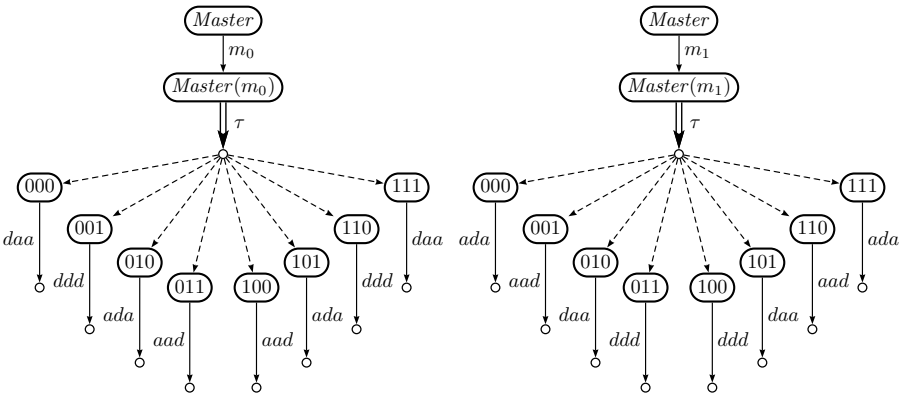
1. $\mathcal{R}' = \{ (a.P, a.Q, 0) \} \cup \mathcal{R}$,
2. $\mathcal{R}' = \{ (pR \oplus (1-p)P, pR \oplus (1-p)Q, 0) \} \cup \mathcal{R} \cup Id_R$,
3. $\mathcal{R}' = \{ (R + P, R + Q, 0) \} \cup \mathcal{R} \cup Id_R$,
4. $\mathcal{R}' = \{ ((\nu a)P', (\nu a)Q', c) \mid (P', Q', c) \in \mathcal{R} \}$,
5. $\mathcal{R}' = \{ (R' \mid P', R' \mid Q', c) \mid (P', Q', c) \in \mathcal{R} \} \cup Id_R$.

7 An Application to the Dining Cryptographers Protocol

In this section we use the pseudometric method to reason about the degree of differential privacy of the Dining Cryptographers Protocol [9] with biased coins. In particular, we show that with probability-p biased coins, the degree of differential privacy in the case of three cryptographers is $|\ln \frac{p}{1-p}|$. This result can also be generalized to the case of n cryptographers.

The problem of the Dining Cryptographers is the following: Three cryptographers dine together. After the dinner, the bill has to be paid by either one of them or by another agent called the master. The master decides who will pay and then informs each of them separately whether he has to pay or not. The cryptographers would like to find out whether the payer is the master or one of them. However, in the latter case, they wish to keep the payer anonymous.

The Dining Cryptographers Protocol (DCP) solves the above problem as follows: each cryptographer tosses a fair coin which is visible to himself and his neighbor to the left. Each cryptographer checks his own coin and the one to his right and, if he is not paying, announces "agree" if the two coins are the same and "disagree" otherwise. However, the paying cryptographer says the opposite. It can be proved that the master is paying if and only if the number of disagrees is even [9].



(a) $Master(m_0)$ (b) $Master(m_1)$

Fig. 4. The probabilistic automata of the Dining cryptographers

The graph shown in Fig. 3 illustrates the dinner-table and the allocation of the coins between the three cryptographers. We consider the coins which are probability-p biased, i.e., producing 0 (for "head") with probability p, and 1 (for "tail") with $1 - p$. We consider the final announcement in the order of $out_0\, out_1\, out_2$, with $out_i \in \{a, d\}$ (a for "agree" and d for "disagree", $i \in \{0, 1, 2\}$) announced by $Crytp_i$. For example, if $Crytp_0$ is designated to pay, $Coin_0\, Coin_1\, Coin_2 = 010$, then $out_0\, out_1\, out_2 = ada$.

We are interested in the case when one of the cryptographers is paying, since that is the case in which they want to keep the payer anonymous. We use $Master(m_i)$ to denote the system in which $Crytp_i$ is designated to pay. To show that the DCP is differentially private, both pseudometrics introduced before can be used. In this problem, it suffices to find between $Master(m_i)$'s bounded distances in the accumulative pseudometric m^D, more precisely, bounded accumulative bisimulation relations.

Proposition 5. *A DCP with three cryptographers and with probability-p biased coins is $|\ln \frac{p}{1-p}|$-differentially private.*

Proof. Fig. 4 shows two probabilistic automata $Master(m_0)$ and $Master(m_1)$ when $Crytp_0$ and $Crytp_1$ are paying respectively. Basically they are probabilistic distributions over all possible outcomes $Coin_0\, Coin_1\, Coin_2$ (i.e. inner states) produced by the three-coins toss, followed by an announcement determined by each outcome. For simplicity initial τ transitions are merged harmlessly. Let $b_0b_1b_2$ and $c_0c_1c_2$ represent two inner states of $Master(m_0)$ and $Master(m_1)$ respectively. There exists a bijection function f between them:

$$c_0c_1c_2 = f(b_0b_1b_2) = b_0(b_1 \oplus 1)b_2$$

where $\oplus$ represents the addition modulo 2 (xor), such that the announcement of $b_0b_1b_2$ can be shown equal to the one of $c_0c_1c_2$.

Note that, the probability of reaching an inner state $b_0b_1b_2$ from $Master(m_0)$ is $p^i(1-p)^{(3-i)}$, where $i \in \{0, 1, 2, 3\}$ is the number of 0 in $\{b_0, b_1, b_2\}$. Because

$c_0 = b_0, c_1 = b_1 \oplus 1, c_2 = b_2$, the ratio between the probabilities of reaching $b_0b_1b_2$ from $Master(m_0)$ and $c_0c_1c_2$ from $Master(m_1)$ differs at most by $|\ln \frac{p}{1-p}|$. It is easy to see that $\{(Master(m_0), Master(m_1), 0)\} \cup \{\, (b_0b_1b_2, f(b_0b_1b_2), |\ln \frac{p}{1-p}|) \mid b_0, b_1, b_2 \in \{0,1\}\,\}$ forms a $|\ln \frac{p}{1-p}|$-accumulative bisimulation relation. Thus $m^D(Master(m_0), Master(m_1)) \leq |\ln \frac{p}{1-p}|$.

Similarly, we consider the probabilistic automata $Master(m_2)$ when $Crytp_2$ is paying (though omitted in Fig. 4). Let $e_0e_1e_2$ represent one of its inner states. We can also find a bijection f' between $c_0c_1c_2$ and $e_0e_1e_2$: $e_0e_1e_2 = f'(c_0c_1c_2) = c_0c_1(c_2 \oplus 1)$, and a bijection f'' between $b_0b_1b_2$ and $e_0e_1e_2$: $e_0e_1e_2 = f''(b_0b_1b_2) = (b_0 \oplus 1)b_1b_2$ such that they output same announcements, The rest proceeds as above. By Theorem 1, the DCP is $|\ln \frac{p}{1-p}|$-differentially private. □

The above proposition can be extended to the case of n dining cryptographers where $n \geq 3$. We assume that the n cryptographers are fully connected, i.e., that a coin exists between every pair of cryptographers. Let c_{kl} $(k, l \in Z, k, l \in [0, n-1], k < l)$ be the coin linking two cryptographers $Crytp_k$ and $Crytp_l$. In this case the output of $Crytp_i$ would be $out_i = c_{0i} \oplus c_{1i} \oplus \cdots c_{i(n-1)} \oplus pay(i)$, where $pay(i) = 1$ if $Crytp_i$ pays and 0 otherwise.

Proposition 6. *A DCP with n fully connected cryptographers and with probability-p biased coins is $|\ln \frac{p}{1-p}|$-differentially private.*

We can see that the more the coins are biased, the worse the privacy gets. If the coins are fair, namely, $p = 1 - p = \frac{1}{2}$, then the DCP is 0-differentially private, in which case the privacy is well protected. With the help of the pseudometric method, we get a general proposition about the degree of differential privacy of DCP. Moreover, it is obtained through some local information, rather than by computing globally the summations of probabilities for each trace.

8 Conclusion and Future Work

We have investigated two pseudometrics on probabilistic automata: the first one is a reformulation of the notion proposed in [27], the second one is designed in the sense that the total privacy leakage bound gets amortised. Each of them establishs a framework for the formal verification of differential privacy for concurrent systems. Namely, the closer processes are in the pseudometrics, the higher level of differential privacy they can preserve. We have showed that our pseudometric is more liberal than the former one. They both impliy strong bisimilarity, and the typical process algebra operators are non-expansive with respect to the distance in the pseudometrics. We have used the pseudometric verification method to learn that: A Dining Cryptographers protocol with probability-p biased coins is $|\ln \frac{p}{1-p}|$-differentially private.

In this paper we have mainly focused on developing a basic framework for the formal verification of differential privacy for concurrent systems. In the future we plan to develop more realistic case-studies and applications. Another interesting direction, which is also our ongoing work, is to investigate a new pseudometric, adapted from the

metric à la Kantorovich proposed in [13], see whether it can fully characterise bisimilarity, and moreover, release the bijection requirement in the definition of the quantitative bisimulations considered in this paper.

Acknowledgements. The authors wish to thank Catuscia Palamidessi, Frank D. Valencia and the anonymous reviewers for providing constructive comments and recommendations on an earlier version of this paper.

References

1. Abadi, M., Gordon, A.D.: A calculus for cryptographic protocols: The spi calculus. Inf. and Comp. 148(1), 1–70 (1999)
2. Andrés, M.E., Palamidessi, C., Sokolova, A., Rossum, P.V.: Information Hiding in Probabilistic Concurrent Systems. TCS 412(28), 3072–3089 (2011)
3. Barthe, G., Köpf, B., Olmedo, F., Béguelin, S.Z.: Probabilistic relational reasoning for differential privacy. In: Proc. of POPL. ACM (2012)
4. Boreale, M.: Quantifying information leakage in process calculi. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds.) ICALP 2006. LNCS, vol. 4052, pp. 119–131. Springer, Heidelberg (2006)
5. Braun, C., Chatzikokolakis, K., Palamidessi, C.: Compositional methods for information-hiding. In: Amadio, R.M. (ed.) FOSSACS 2008. LNCS, vol. 4962, pp. 443–457. Springer, Heidelberg (2008)
6. Canetti, R., Cheung, L., Kaynar, D., Liskov, M., Lynch, N., Pereira, O., Segala, R.: Task-structured probabilistic i/o automata. In: Proc. of WODES (2006)
7. Chatzikokolakis, K., Andrés, M.E., Bordenabe, N.E., Palamidessi, C.: Broadening the scope of differential privacy using metrics. In: De Cristofaro, E., Wright, M. (eds.) PETS 2013. LNCS, vol. 7981, pp. 82–102. Springer, Heidelberg (2013)
8. Chatzikokolakis, K., Palamidessi, C.: Making random choices invisible to the scheduler. In: Caires, L., Vasconcelos, V.T. (eds.) CONCUR 2007. LNCS, vol. 4703, pp. 42–58. Springer, Heidelberg (2007)
9. Chaum, D.: The dining cryptographers problem: Unconditional sender and recipient untraceability. Journal of Cryptology 1, 65–75 (1988)
10. de Frutos-Escrig, D., Rosa-Velardo, F., Gregorio-Rodríguez, C.: New bisimulation semantics for distributed systems. In: Derrick, J., Vain, J. (eds.) FORTE 2007. LNCS, vol. 4574, pp. 143–159. Springer, Heidelberg (2007)
11. Deng, Y., Chothia, T., Palamidessi, C., Pang, J.: Metrics for action-labelled quantitative transition systems. In: Proc. of QAPL. ENTCS, vol. 153, pp. 79–96. Elsevier (2006)
12. Deng, Y., Pang, J., Wu, P.: Measuring anonymity with relative entropy. In: Dimitrakos, T., Martinelli, F., Ryan, P.Y.A., Schneider, S. (eds.) FAST 2006. LNCS, vol. 4691, pp. 65–79. Springer, Heidelberg (2007)
13. Desharnais, J., Jagadeesan, R., Gupta, V., Panangaden, P.: The metric analogue of weak bisimulation for probabilistic processes. In: Proc. of LICS, pp. 413–422. IEEE (2002)
14. Dwork, C.: Differential privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds.) ICALP 2006. LNCS, vol. 4052, pp. 1–12. Springer, Heidelberg (2006)
15. Focardi, R., Gorrieri, R.: Classification of security properties (part i: Information flow). In: Focardi, R., Gorrieri, R. (eds.) FOSAD 2000. LNCS, vol. 2171, pp. 331–396. Springer, Heidelberg (2001)
16. Gaboardi, M., Haeberlen, A., Hsu, J., Narayan, A., Pierce, B.C.: Linear dependent types for differential privacy. In: POPL, pp. 357–370 (2013)

17. Jou, C.-C., Smolka, S.: Equivalences, congruences, and complete axiomatizations for probabilistic processes. In: Baeten, J.C.M., Klop, J.W. (eds.) CONCUR 1990. LNCS, vol. 458, pp. 367–383. Springer, Heidelberg (1990)
18. Kiehn, A., Arun-Kumar, S.: Amortised bisimulations. In: Wang, F. (ed.) FORTE 2005. LNCS, vol. 3731, pp. 320–334. Springer, Heidelberg (2005)
19. Larsen, K.G., Skou, A.: Bisimulation through probabilistic testing. Inf. and Comp. 94(1), 1–28 (1991)
20. Machanavajjhala, A., Kifer, D., Abowd, J.M., Gehrke, J., Vilhuber, L.: Privacy: Theory meets practice on the map. In: Proc. of ICDE, pp. 277–286. IEEE (2008)
21. Milner, R.: Communication and Concurrency. Series in Comp. Sci. Prentice Hall (1989)
22. Mu, C.: Measuring information flow in reactive processes. In: Qing, S., Mitchell, C.J., Wang, G. (eds.) ICICS 2009. LNCS, vol. 5927, pp. 211–225. Springer, Heidelberg (2009)
23. Narayanan, A., Shmatikov, V.: De-anonymizing social networks. In: Proc. of S&P, pp. 173–187. IEEE (2009)
24. Reed, J., Pierce, B.C.: Distance makes the types grow stronger: a calculus for differential privacy. In: Proc. of ICFP, pp. 157–168. ACM (2010)
25. Ryan, P.Y.A., Schneider, S.A.: Process algebra and non-interference. Journal of Computer Security 9(1/2), 75–103 (2001)
26. Smith, G.: Probabilistic noninterference through weak probabilistic bisimulation. In: CSFW, pp. 3–13 (2003)
27. Tschantz, M.C., Kaynar, D., Datta, A.: Formal verification of differential privacy for interactive systems (extended abstract). ENTCS 276, 61–79 (2011)
28. Xu, L.: Modular reasoning about differential privacy in a probabilistic process calculus. In: Palamidessi, C., Ryan, M.D. (eds.) TGC 2012. LNCS, vol. 8191, pp. 198–212. Springer, Heidelberg (2013)

A Appendix

A.1 Proof of Lemma 2

Proof. We prove by induction on the length of trace $\boldsymbol{t}$: $|\boldsymbol{t}|$.

1. $|\boldsymbol{t}| = 0$: According to equation (1), for any ζ, $\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}] = \Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}] = 1$.
2. IH: For any two executions α_1 and α_2 of $\mathcal{A}$, let $s_1 = lstate(\alpha_1)$ and $s_2 = lstate(\alpha_2)$. $(s_1, s_2, c) \in \mathcal{R}$ implies that for any admissible scheduler ζ, trace $\boldsymbol{t}'$ where $|\boldsymbol{t}'| \leq L$: $\frac{1}{e^{\epsilon+c}} \leq \frac{\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}']}{\Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}']} \leq e^{\epsilon-c}$.
3. We have to show that for any admissible scheduler ζ, trace $\boldsymbol{t}$ with $|\boldsymbol{t}| = L + 1$, $(s_1, s_2, c) \in \mathcal{R}$ implies $\frac{1}{e^{\epsilon+c}} \leq \frac{\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}]}{\Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}]} \leq e^{\epsilon-c}$.
   Assume that $\boldsymbol{t} = a^\frown \boldsymbol{t}'$. We prove first the right-hand part $\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}] \leq e^{\epsilon-c} * \Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}]$. According to equation (1), two cases must be considered:
   - Case $act(\zeta(\alpha_1)) \neq a$. Then $\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}] = 0$. Since ζ is admissible, it schedules for α_2 a transition consistent with α_1, namely, not a transition labeled by a either. Thus $\Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}] = 0$, the inequality is satisfied.
   - Case $\zeta(\alpha_1) = s_1 \xrightarrow{a} \mu_1$. So, $\Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}] = \sum_{s_i \in supp(\mu_1)} \mu_1(s_i) \Pr_\zeta[\alpha_1 a s_i \rhd \boldsymbol{t}']$. Since $(s_1, s_2, c) \in \mathcal{R}$, there must be also a transition from s_2 such that $s_2 \xrightarrow{a} \mu_2$ and $\mu_1 \mathcal{L}^A(\mathcal{R}, c) \mu_2$. Since ζ is admissible, $\zeta(\alpha_2) = s_2 \xrightarrow{a} \mu_2$. We use t_i to range over elements in $supp(\mu_2)$. Thus, $\Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}] = \sum_{t_i \in supp(\mu_2)} \mu_2(t_i) \cdot$

$\Pr_\zeta[\alpha_2 a t_i \rhd \boldsymbol{t}']$. Since $\mu_1 \mathcal{L}^A(\mathcal{R}, c)\mu_2$, there is a bijection $\beta : supp(\mu_1) \longrightarrow supp(\mu_2)$, s.t. for any $s_i \in supp(\mu_1)$, there is a state $t_i \in supp(\mu_2)$, $t_i = \beta(s_i)$ and $(s_i, t_i, c + \ln \mu_1(s_i) - \ln \mu_2(t_i)) \in \mathcal{R}$. Apply the inductive hypothesis to $\alpha_1 a s_i, \alpha_2 a t_i$ and $\boldsymbol{t}'$, we get that:

$$\Pr_\zeta[\alpha_1 a s_i \rhd \boldsymbol{t}'] \leq e^{\epsilon - (c + \ln \mu_1(s_i) - \ln \mu_2(t_i))} * \Pr_\zeta[\alpha_2 a t_i \rhd \boldsymbol{t}'] \tag{2}$$

Thus,

$$\begin{aligned}
& \Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}] && (3)\\
= & \sum_{s_i \in supp(\mu_1)} \mu_1(s_i) \Pr_\zeta[\alpha_1 a s_i \rhd \boldsymbol{t}'] && (4)\\
\leq & \sum_{s_i \in supp(\mu_1)} \mu_1(s_i) e^{\epsilon - (c + \ln \mu_1(s_i) - \ln \mu_2(\beta(s_i)))} \Pr_\zeta[\alpha_2 a \beta(s_i) \rhd \boldsymbol{t}'] && (5)\\
= & \sum_{s_i \in supp(\mu_1)} \mu_1(s_i) * \frac{\mu_2(\beta(s_i))}{\mu_1(s_i)} * e^{\epsilon - c} * \Pr_\zeta[\alpha_2 a \beta(s_i) \rhd \boldsymbol{t}'] && (6)\\
= & \sum_{t_i \in supp(\mu_2)} \mu_2(t_i) * e^{\epsilon - c} * \Pr_\zeta[\alpha_2 a t_i \rhd \boldsymbol{t}'] && (7)\\
= & e^{\epsilon - c} \sum_{t_i \in supp(\mu_2)} \mu_2(t_i) \Pr_\zeta[\alpha_2 a t_i \rhd \boldsymbol{t}'] && (8)\\
= & e^{\epsilon - c} * \Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}] && (9)
\end{aligned}$$

which completes the proof of right-hand part. Lines (4) and (9) follow from the equation (1). Line (5) follow from the inductive hypothesis, i.e. Line (2).
For the left-hand part $\Pr_\zeta[\alpha_2 \rhd \boldsymbol{t}] \leq e^{\epsilon + c} * \Pr_\zeta[\alpha_1 \rhd \boldsymbol{t}]$, exchange the roles of s_1 and s_2, use β^{-1} instead of β, and all the rest is analogous. □

A.2 Proof of Proposition 6

Proof sketch. The proof proceeds analogously to the case of three cryptographers. To find an accumulative bisimulation relation between every two instances of the DCP $Master(m_i)$ and $Master(m_j)$, $(i, j \in Z, i, j \in [0, n-1], i < j)$, we point out here mainly the bijection function between their inner states. Let $b_{12} b_{13} \cdots b_{(n-1)n}$ and $c_{12} c_{13} \cdots c_{(n-1)n}$ represent the inner states of $Master(m_i)$ and $Master(m_j)$ respectively, where the subscript (kl), $(k, l \in Z, k, l \in [0, n-1], k < l)$, indicates the coin linking two cryptographers $Crytp_k$ and $Crytp_l$. There exists a bijection function f between them defined as: $c_{12} c_{13} \cdots c_{(n-1)n} = f(b_{12} b_{13} \cdots b_{(n-1)n})$, precisely,

$$c_{kl} = \begin{cases} b_{kl} \oplus 1 & \text{if } kl = ij, \\ b_{kl} & \text{otherwise.} \end{cases}$$

We can check that the bijective states defined in this way produce the same announcement in $Master(m_i)$ and $Master(m_j)$. Moreover, only the coin (ij) is different, the ratio between the probability mass of every bijective states is at most $|\ln \frac{p}{1-p}|$. □