

Exploring probabilistic bisimulations, part I

Matthew Hennessy

Department of Computer Science, Trinity College, Dublin 1, Ireland

Abstract. We take a fresh look at strong probabilistic bisimulations for processes which exhibit both non-deterministic and probabilistic behaviour. We suggest that it is natural to interpret such processes as distributions over states in a probabilistic labelled transition system, a pLTS; this enables us to adapt the standard notion of contextual equivalence to this setting. We then prove that a novel form of bisimulation equivalence between distributions are both sound and complete with respect to this contextual equivalence. We also show that a very simple extension to HML, Hennessy–Milner Logic, provides finite explanations for inequivalences between distributions. Finally we show that our bisimulations between distributions in a pLTS are simply an alternative characterisation of a standard notion of probabilistic bisimulation equivalence, defined between states in a pLTS.

Keywords: Probabilistic processes, Contextual equivalence, Bisimulation equivalence, Logical characterisation, Equational theory

1. Introduction

Bisimulations [Mil89] provide a well-established and elegant theory of the behaviour of non-deterministic processes. Let us review the framework.

- (1) *Labelled transitions systems, LTSs:* These provide an intensional semantics for processes, describing their computations or more generally the interactions between processes and their environment.
- (2) *Process calculi:* Formal description languages for describing processes and their specifications. These usually consist of a small number of *combinators* with which processes can be described by their structure.
- (3) *Behavioural equivalence:* This determines which process descriptions are extensionally equivalent; that is which processes can not be distinguished by their users or more generally their environments. Perhaps the most uncontroversial is the so-called *contextual equivalence*, \sim_{ext} [MS92, SW01, HY95, Hen07], defined in terms of simple properties one would expect of a behavioural equivalence.
- (4) *Bisimulations:* These are relations between processes which satisfy simple properties expressed in terms of the intensional semantics. They provide an elegant proof methodology for demonstrating process equivalence; to show two processes behaviourally equivalent it is sufficient to exhibit a witness bisimulation containing them. In many settings this proof method is not only sound with respect to contextual equivalence but also complete.

- (5) *Property logics*: These describe extensional properties of processes, but are also used as a proof methodology for demonstrating inequivalence between processes. For example in many settings, [Mil89, Cle90] if two processes are not equivalent there is a property expressed in the modal logic HML which one satisfies and the other does not. Thus these properties provide an finite explanation of why the two processes are inequivalent.
- (6) *Equational theories*: these are equations, in terms of the combinators of a process calculus, which can be used to manipulate process descriptions while preserving behavioural equivalence. Many process calculi have *complete* equational theories, in the sense that if two finite process descriptions are behaviourally equivalent they can be proved equivalent using the equations.

If we now turn our attention to probabilistic processes, or more correctly processes which exhibit both non-deterministic and probabilistic behaviour, the situation is not as clear, despite the intense research into their semantics, [Seg95, SL95, PLS00, BEMC00, BS01, LS89, MM05, Sto02]. The purpose of this paper is to show that a similar framework can be identified, with relatively little effort. We confine our attention to the *strong case*, in which the internal behaviour of processes is not taken into account; treating the *weak case*, which abstracts from internal actions, would require considerable more effort.

Let us take the individual ingredients above in turn.

- (1) In the purely non-deterministic setting a process corresponds to a state in an LTS, and the intensional behaviour is determined by relations of the form $s \xrightarrow{\mu} t$. Here the action label μ represents an interaction between the process and its environment. In [Seg95] probabilistic automata were suggested as a convenient intensional model for probabilistic processes, and these, or minor variations, are now widely used for this purpose. Here we use the popular *probabilistic labelled transition systems* or *pLTSs*, in which intensional behaviour is determined by relations of the form $s \xrightarrow{\mu} \Delta$, where s is a state and Δ is a distribution. However we view a process not as a state in a pLTS but as a distribution over states. This view is supported by the well-known natural generalisation of the relations $s \xrightarrow{\mu} \Delta$, over states to distributions, to relations between distributions, $\Delta \xrightarrow{\mu} \Theta$.
- (2) Standard process calculus, such as ACP, CCS or CSP, have been extended to describe probabilistic behaviour, often by adding one extra combinator for probabilistic choice, [And02, Tof94, MMSS95]. Thus $Q_p \oplus R$ is a process which behaves like Q with probability p and like R with probability $(1 - p)$. In this paper, as an example, we use a very simple CCS-based calculus which we call bpCCS, for basic probabilistic CCS. We will interpret every term in bpCCS as a distribution over a set of states in a pLTS.
- (3) Contextual equivalences have long been used to define the behavioural semantics of sequential and concurrent languages. Following [DH11] we adapt this approach to probabilistic systems. Our definition is based on *reduction barbed congruence* [HY95, MS92] which requires some notion of *observation* or *barb*; see page 116 of [SW01] or page 41 of [Hen07] for more discussion on this point. In a probabilistic setting there is a certain amount of choice here, but we show that the resulting behavioural equivalence, which we denote by \sim_{rbc} , is independent of the natural choices.
- (4) There is a natural, and well-known, generalisation of (strong) bisimulation equivalence in LTSs to probabilistic versions in pLTSs, [SL95, PS07, BS01], defined as a relation between states. In this paper we show that this can be recast as a relation between processes, that is distributions. Moreover one of the main results of the paper is that the resulting bisimulation equivalence coincides with our contextual equivalence \sim_{rbc} . This result is subject to minor conditions, similar to those required for the corresponding result to hold in the purely non-deterministic setting.
- (5) In the purely non-deterministic case inequivalences are explained using a simple property logic HML, obtained by adding to Boolean logic one modal operator $\langle \mu \rangle \varphi$, [Mil89]. Various extensions to this logic have been proposed in various probabilistic settings, in both the strong and the weak case, in order to explain inequivalences between probabilistic systems; see for example [Han91, PLS03, HPS⁺11, PS07]. In the second main

result of the paper we show that it is sufficient to add one particularly simple further operator, $\varphi_1 \oplus_p \varphi_2$ parameterised on the probability p . Intuitively a probabilistic process, that is a distribution in a pLTS, satisfies this property if it satisfies property φ_1 with probability p and property φ_2 with probability $(1 - p)$.

- (6) Finally (strong) bisimulation equivalence has a well-known equational axiomatisation for finite CCS expressions, [Mil89]. We show that this can be extended to the probabilistic case by adding natural equations which capture the behaviour of probabilistic choice. This result is a simple variation on similar results in [And02, BS01], which consider slight variations on bpCCS.

The remainder of the paper is organised as follows. After recalling some mathematical notation, in Sect. 3 we review the various definitions of bisimulations in a probabilistic setting. The main result here, Theorem 3.17, is that our version of bisimulation equivalence between distributions in a pLTS, $\Delta \sim_{sd} \Theta$, is determined by the standard notion of (strong) probabilistic bisimulation equivalence [SL95], a relation between the states of a pLTS. In Sect. 4 we explain our probabilistic extension of HML and how probabilistic processes satisfies these properties. Here the main result, Theorem 4.3, states that two probabilistic processes are not contextually equivalent if and only if there is some property in pHML which distinguishes them. Note that formulae in pHML are by definition finite and therefore they offer finite explanations for inequivalences.

In Sect. 5 we define our contextual equivalence between processes in a pLTS, which we call \sim_{rc} ; this is a natural extension of the corresponding definition in the purely non-deterministic case. We then prove, Theorem 5.6, that subject to some expressivity constraints, this behavioural equivalence coincides with our bisimulation equivalence from Sect. 3; thus these bisimulations also provide a sound and complete proof methodology for contextual equivalence. Finally in Sect. 6 we outline the axiomatisation of contextual equivalence over finite terms in bpCCS. The paper ends with a brief look at related work.

2. Mathematical preliminaries

We begin with some notation. A (discrete) probability distribution over a set S is a function $\Delta : S \rightarrow [0, 1]$ with $\sum_{s \in S} \Delta(s) = 1$; the support of such an Δ is the set $\text{supp}(\Delta) = \{s \in S \mid \Delta(s) > 0\}$. The point distribution \bar{s} assigns probability 1 to s and 0 to all other elements of S , so that $\text{supp}(\bar{s}) = s$. We use $\mathcal{D}(S)$ to denote the set of distributions, and we extend elements Δ of $\mathcal{D}(S)$ to functions over the powerset of S in the natural way, by letting $\Delta(T) = \sum_{s \in T} \Delta(s)$, for any subset T of S .

Let $\{p_i \mid 1 \leq i \leq n\}$ be a finite collection of real numbers satisfying $\sum_{1 \leq i \leq n} p_i = 1$ and let Δ_i , $1 \leq i \leq n$, be a collection of distributions. Then we use $\sum_{1 \leq i \leq n} p_i \cdot \Delta_i$ to denote the distribution determined by $(\sum_{1 \leq i \leq n} p_i \cdot \Delta_i)(s) = \sum_{1 \leq i \leq n} p_i \cdot \Delta_i(s)$. We sometimes abbreviate $p \cdot \Delta_1 + (1 - p) \cdot \Delta_2$ to $\Delta_1 \oplus_p \Delta_2$. For any subset D of $\mathcal{D}(S)$ we use $cc(D)$ to denote the convex-closure of D , that is the least subset of $\mathcal{D}(S)$ which contains D and is closed under the operations $-\oplus_p-$, where p ranges over $[0, 1]$.

Definition 2.1 Let $\mathcal{R} \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$. It is said to be linear if $\Delta_i \mathcal{R} \Theta_i$, $i = 1, 2$, implies $(\Delta_1 \oplus_p \Delta_2) \mathcal{R} (\Theta_1 \oplus_p \Theta_2)$ for any $0 \leq p \leq 1$. It is said to be left-decomposable if $(\Delta_1 \oplus_p \Delta_2) \mathcal{R} \Theta$ implies $\Theta = (\Theta_1 \oplus_p \Theta_2)$, where $\Delta_i \mathcal{R} \Theta_i$, $i = 1, 2$. The concept of right-decomposable is defined analogously, and \mathcal{R} is said to be decomposable if it is both left- and right-decomposable. Note that if \mathcal{R} is symmetric then left-decomposable implies decomposable. ■

Definition 2.2 (Lifting) Let $\mathcal{R} \subseteq S \times \mathcal{D}(S)$ be a relation from S to distributions over S . Then $lift(\mathcal{R}) \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$ is the smallest linear relation in $\mathcal{D}(S) \times \mathcal{D}(S)$ that satisfies $s \mathcal{R} \Theta$ implies $\bar{s} lift(\mathcal{R}) \Theta$. ■

There are numerous alternative formulations of this idea. For example *weight* functions are used in [BEMC00, Sto02]; for a comparison of the various approaches see [DD11]. The following is taken from [DvG10, DH11].

Lemma 2.3 For $\mathcal{R} \subseteq S \times S$, $\Delta \text{lift}(\mathcal{R}) \Theta$ if and only if there is a finite index set I such that

- (i) $\Delta = \sum_{i \in I} p_i \cdot \bar{s}_i$
- (ii) $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$
- (iii) $s_i \mathcal{R} \Theta_i$ for each $i \in I$

Proof See Lemma 2.5 of [DH11]. □

Lemma 2.4 For any $\mathcal{R} \subseteq S \times \mathcal{D}(S)$, the relation $\text{lift}(S)$ is left-decomposable.

Proof See Proposition 2.7 of [DH11]. □

It is more usual to consider liftings of relations from $S \times S$, rather than $S \times \mathcal{D}(S)$, [Seg95, SL94, PS07, BEMC00]. We can achieve this indirectly via Definition 2.2. For $\mathcal{R} \subseteq S \times S$ let \mathcal{R}^e be the obvious extension to $S \times \mathcal{D}(S)$ given by $s \mathcal{R}^e \Theta$ if and only if $\Theta = \bar{t}$ for some $t \in S$ such that $s \mathcal{R} t$. Then we use $\text{slift}(\mathcal{R})$ to denote its lifting, $\text{lift}(\mathcal{R}^e)$. The properties of the operation $\text{lift}(-)$ given in Lemma 2.4 and Lemma 2.3 are easily generalised to $\text{slift}(-)$; we leave the proofs to the reader:

Lemma 2.5 Suppose $\mathcal{R} \subseteq S \times S$. Then

- (a) $\text{slift}(\mathcal{R})$ is both left- and right-decomposable
- (b) $\Delta \text{slift}(\mathcal{R}) \Theta$ if and only if there is a finite index set I such that

- (i) $\Delta = \sum_{i \in I} p_i \cdot \bar{s}_i$
 - (ii) $\Theta = \sum_{i \in I} p_i \cdot \bar{t}_i$
 - (iii) $s_i \mathcal{R} t_i$ for each $i \in I$.
-

Lemma 2.6 For any $\mathcal{R}_1, \mathcal{R}_2 \subseteq (S \times S)$, $\text{slift}(\mathcal{R}_1) \circ \text{slift}(\mathcal{R}_2) \subseteq \text{slift}(\mathcal{R}_1 \circ \mathcal{R}_2)$.

Proof Suppose $\Delta \text{slift}(\mathcal{R}_1) \circ \text{slift}(\mathcal{R}_2) \Theta$. This means there is some distribution Γ such that $\Gamma \text{slift}(\mathcal{R}_2) \Theta$ and

$$\Delta = \sum_{i \in I} p_i \cdot s_i \quad \Gamma = \sum_{i \in I} p_i \cdot r_i \quad \text{where } s_i \mathcal{R}_1 r_i \text{ for each } i \in I$$

Left-decomposability of $\text{slift}(\mathcal{R}_2)$ gives $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ where $\bar{r}_i \text{slift}(\mathcal{R}_2) \Theta_i$ for each $i \in I$. Right-decomposability means in turn that $r_i \mathcal{R}_2 t$ for every $t \in [\Theta_i]$. But this means that for each such t , $s_i \mathcal{R}_1 \circ \mathcal{R}_2 t$ and therefore $\bar{s}_i \text{slift}(\mathcal{R}_1 \circ \mathcal{R}_2) \Theta_i$. The required $\Delta \text{slift}(\mathcal{R}_1 \circ \mathcal{R}_2) \Theta$ now follows by the linearity of $\text{slift}(-)$. □

There is another useful characterisation of $\text{slift}(\mathcal{R})$, when \mathcal{R} is an equivalence relation, [BEMC00, DD11].

Proposition 2.7 Suppose $\mathcal{R} \subseteq S \times S$ is an equivalence relation. Then $\Delta \text{slift}(\mathcal{R}) \Theta$ if and only if for every equivalence class E of \mathcal{R} , $\Delta(E) = \Theta(E)$.

Proof (Outline) First suppose $\Delta \text{slift}(\mathcal{R}) \Theta$. Employing part (b) of Lemma 2.5 we know that for some finite index set I

$$\Delta = \sum_{i \in I} p_i \cdot \bar{s}_i \quad \Theta = \sum_{i \in I} p_i \cdot \bar{t}_i \tag{1}$$

where $s_i \mathcal{R} t_i$ for each $i \in I$. Then for any subset E of S simple calculations give

$$\Delta(E) = \sum \{p_i \mid i \in I, s_i \in E\} \quad \Theta(E) = \sum \{p_i \mid i \in I, t_i \in E\}$$

Now suppose E is actually an equivalence class over S . Then it follows that $\Delta(E) = \Theta(E)$ as in this case $t_i \in E$ if and only if $s_i \in E$ for each $i \in I$.

Conversely, suppose $\Delta(E) = \Theta(E)$ for every equivalence class E of some equivalence relation $\mathcal{R} \subseteq S \times S$. We have to come up with a decomposition of Δ and Θ as in (1) above. To this end let I be the finite indexing set $\{(s, t) \mid s \mathcal{R} t\}$, and for each $(s, t) \in I$ let $l_{(s,t)}, r_{(s,t)}$ denote s, t respectively.

We now have to describe the appropriate probabilities p_i for each $i \in I$. If we use $[s]$ to denote the equivalence class $\{t \mid s \mathcal{R} t\}$ then we can let $p_{(s,t)}$ be $\frac{\Delta(s) \cdot \Theta(t)}{\Delta([s])}$; note that in the denominator we could equally well have used $\Theta([t])$. Then calculations show

$$\Delta = \sum_{(s,t) \in I} p_{(s,t)} \cdot \overline{l_{(s,t)}} \quad \Theta = \sum_{(s,t) \in I} p_{(s,t)} \cdot \overline{r_{(s,t)}}$$

and the result now follows since $l_{(s,t)} \mathcal{R} r_{(s,t)}$ for each (s, t) in I . \square

3. Transition systems and bisimulations

Definition 3.1 A labelled transition system (LTS) is a triple $\langle S, \text{Act}, \rightarrow \rangle$ where

- (a) S is a set of states
- (b) Act_τ is a set of transition labels, with distinguished element τ
- (c) the relation \rightarrow is a subset of $S \times \text{Act}_\tau \times S$.

Instances of the relation are usually written $s \xrightarrow{\mu} t$. The special action τ will be used in Sect. 5 to represent synchronisation. \blacksquare

Recall, [Mil89], that (strong) bisimulation equivalence \sim is the largest relation over the states of an LTS such that if $s \sim t$ then for every $\mu \in \text{Act}_\tau$

- (a) $s \xrightarrow{\mu} s'$ implies $t \xrightarrow{\mu} t'$ for some t' such that $s' \sim t'$
- (b) conversely, $t \xrightarrow{\mu} t'$ implies $s \xrightarrow{\mu} s'$ for some s' such that $s' \sim t'$.

However there is another less well-known formulation which relies on the fact that \sim is an equivalence relation.

Definition 3.2 An equivalence relation \mathcal{R} over the states of an LTS is called an *equiv-bisimulation* if whenever $s \mathcal{R} t$ then for every $\mu \in \text{Act}_\tau$ and equivalence class E of \mathcal{R} ,

- (a) if $s \xrightarrow{\mu} s'$ for some $s' \in E$ then $t \xrightarrow{\mu} t'$ for some $t' \in E$.

Note that the symmetric condition on t follows automatically from the fact that \mathcal{R} is an equivalence class. \blacksquare

Proposition 3.3 In an arbitrary LTS bisimulation equivalence \sim is the largest equiv-bisimulation.

Proof Straightforward calculations. \square

One of the first generalisations of LTSs to include probabilistic behaviour is the following, taken from [LS89].

Definition 3.4 A labelled Markov chain (LMC) is a four-tuple $\langle S, \text{Act}_\tau, C, P \rangle$, where

- (a) S is a set of states
- (b) Act_τ is a set of transition labels, with distinguished action τ
- (c) for each $\mu \in \text{Act}_\tau$, $C(\mu)$ is a subset of S
- (d) for each $\mu \in \text{Act}_\tau$ and each $s \in C(\mu)$, $P(\mu)(s)$ is a probabilistic distribution over S , that is an element of $\mathcal{D}(S)$.

Intuitively $s \in C(\mu)$ means that in state s the action μ can happen and $P(\mu)(s)(t)$ gives the probability that the next state is t when the action μ is fired in state s . \blacksquare

Intuitively LMCs are obtained from LTSs by replacing non-determinism with probabilistic choice. In view of Proposition 3.3 the following is a reasonable formulation of (strong) bisimulation for these kinds of probabilistic processes.

$$\begin{array}{c}
\text{(ACTION)} \\
\mu.P \xrightarrow{\mu} [P] \\
\\
\frac{\text{(EXT.L)} \quad P \xrightarrow{\mu} \Delta,}{P + Q \xrightarrow{\mu} \Delta} \qquad \qquad \qquad \frac{\text{(EXT.R)} \quad Q \xrightarrow{\mu} \Delta,}{P + Q \xrightarrow{\mu} \Delta}
\end{array}$$

Fig. 1. Operational semantics of bpCCS

Definition 3.5 (Larsen–Skou bisimulations [LS89]) For an equivalence relation $\mathcal{R} \subseteq S \times S$, where S is the set of states in some LMC, $\mathcal{B}_{ls}(\mathcal{R})$ is the relation over $S \times S$ determined by letting $s \mathcal{B}_{ls}(\mathcal{R}) t$ whenever, for every $\mu \in \text{Act}_\tau$,

- (a) $s \in C(\mu)$ if and only if $t \in C(\mu)$
- (b) and for every equivalence class E of \mathcal{R} , $P(a)(s)(E) = P(a)(t)(E)$.

Recall that $P(\mu)(s)$ is a distribution, and so $P(\mu)(s)(E)$ is the accumulated probability of arriving in E from the state s , after performing μ ; $P(\mu)(s)(E) = \sum_{t \in E} P(\mu)(s)(t)$. We use $\sim_{ls} \subseteq S \times S$ to denote the largest equivalence relation such that $\sim_{ls} \subseteq \mathcal{B}_{ls}(\sim_{ls})$. ■

The (simple) probabilistic automata (PA) from [SL94] can be viewed as another generalisation of LTSs to include probabilistic behaviour; however in PAs both non-deterministic and probabilistic behaviour are present. Here we use a minor variation of PAs [BEMC00, DvGHM09].

Definition 3.6 An *probabilistic labelled transition system* (pLTS) is a triple $\langle S, \text{Act}_\tau, \rightarrow \rangle$, where

- (a) S is a set of states
- (b) Act_τ is a set of transition labels, with distinguished element τ
- (c) the relation \rightarrow is a subset of $S \times \text{Act} \times \mathcal{D}(S)$.

As with LTSs, instances of the relation are normally written $s \xrightarrow{\mu} \Delta$. We also use $(s \text{ after } \mu)$ to denote the set $\{\Delta \mid s \xrightarrow{\mu} \Delta\}$. A pLTS is called *deterministic* if for all $\mu \in \text{Act}_\tau$, the set $(s \text{ after } \mu)$ contains at most one distribution. It is *finitary* if all its components are finite sets. ■

pLTSs will be the main focus of this paper and so it will be convenient to introduce a syntax for describing them. The process calculus CCS can be viewed as a syntax for describing LTSs and so we devise a simple extension for pLTSs, called bpCCS meaning *basic probabilistic CCS*:

$$P \in \text{bpCCS} ::= \mathbf{0} \mid \mu.P, \mu \in \text{Act}_\tau \mid P + P \mid P_p \oplus P, p \in (0, 1)$$

We have deliberately kept this language basic; more complicated behaviour can be obtained by adding recursive definitions, and the language will be further extended in Sect. 5. Following [DvGH⁺07] we interpret each process term as a distribution over a pLTS, whose states are given by a subset of terms, and whose next-state relations are defined by a collection of inductive rules. The states in the pLTS are given by:

$$s \in \text{bpCCS}_s ::= \mathbf{0} \mid \mu.P, \mu \in \text{Act}_\tau \mid s + s \tag{2}$$

Intuitively these are process terms which have no top-level probabilistic behaviour. The interpretation of arbitrary terms $[P]$ as distributions over these states is given by

- (a) $[\mathbf{0}] = \bar{\mathbf{0}}$ and $[a.P] = \overline{a.P}$
- (b) $[P_1_p \oplus P_2] = p \cdot [P_1] + (1 - p) \cdot [P_2]$
- (c) $[P_1 + P_2]$ is determined by $[P_1 + P_2](s) = \begin{cases} [P_1](s_1) \cdot [P_2](s_2), & \text{if } s = s_1 + s_2, \text{ where } s_i \in [P_i] \\ 0, & \text{otherwise} \end{cases}$

Note that (c) effectively makes external choice $+$ into an operator on states, which is lifted in the standard manner to distributions. In other words if we were to assume that $+$ is a syntactic operator on state terms only then we could extend it to distributions by defining

$$(\Delta + \Theta)(s) = \begin{cases} \Delta(s_1) \cdot \Theta(s_2), & \text{if } s = s_1 + s_2 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

The next-state relations, over states in bpCCS_s , is determined by the three simple rules in Fig. 1. In the remainder of the paper when giving examples of terms in bpCCS we use standard abbreviations; for example we abbreviate $a. \mathbf{0}$ to a .

It is not too difficult to see that LMCs are essentially deterministic pLTSs. First let M denote the LMC $\langle S, \text{Act}_\tau, C, P \rangle$. The corresponding pLTS, denoted $\text{plts}(M)$, is given by $\langle S, \text{Act}_\tau, \rightarrow \rangle$ where $s \xrightarrow{\mu} \Delta$ whenever

- (a) $s \in C(\mu)$
- (b) Δ is the distribution $P(\mu)(s)$.

Note that $\text{plts}(M)$ is always deterministic by construction.

Conversely, consider a deterministic pLTS $M = \langle S, \text{Act}_\tau, \rightarrow \rangle$. The LMC $\text{lmc}(M)$ is determined by $\langle S, \text{Act}, C, P \rangle$, where

- (a) $s \in C(\mu)$ if $s \xrightarrow{\mu} \Delta$ for some Δ
- (b) for each $s \in C(\mu)$, $P(a)(s)$ is the unique distribution Δ such that $s \xrightarrow{\mu} \Delta$.

Note that these two operations are inverses, $\text{plts}(\text{lmc}(M)) = M$ and $\text{lmc}(\text{plts}(M)) = M$.

Definition 3.5 can easily be adapted to pLTSs.

Definition 3.7 For an equivalence relation $\mathcal{R} \subseteq (S \times S)$, where S is the set of states in some pLTS, $\mathcal{B}_{ls}(\mathcal{R})$ is the relation over $S \times S$ determined by letting $s \mathcal{B}_{ls}(\mathcal{R}) t$ whenever for every $\mu \in \text{Act}_\tau$,

- (a) $s \xrightarrow{\mu} \Delta$ implies $t \xrightarrow{\mu} \Theta$ for some Θ such that $\Delta \text{slift}(\mathcal{R}) \Theta$
- (b) conversely, $t \xrightarrow{\mu} \Theta$ implies $s \xrightarrow{\mu} \Delta$ for some Δ such that $\Delta \text{slift}(\mathcal{R}) \Theta$.

We use the same notation, $\sim_{ls} \subseteq S \times S$, to denote the largest equivalence relation such that $\sim_{ls} \subseteq \mathcal{B}_{ls}(\sim_{ls})$. This is referred to as (strong) *bisimulation equivalence* in papers such as [LS89, SL95, BEMC00, PS07, And02], although typically the formulation is slightly different, using the alternative characterisation of $\text{slift}(\mathcal{R})$ in Proposition 2.7. ■

Theorem 3.8

1. If $s \sim_{ls} t$ in an LMC M then $s \sim_{ls} t$ in the pLTS $\text{plts}(M)$.
2. Conversely, if $s \sim_{ls} t$ in a deterministic pLTS M then $s \sim_{ls} t$ in the LMC $\text{lmc}(M)$.

Proof Straightforward using Proposition 2.7. □

However it is felt, see for example page 17 of [SL95], that Definition 3.7 is not an appropriate notion of behavioural equivalence for arbitrary pLTSs, as the following example underlines.

Example 3.9 Let p denote the process $a.b.\mathbf{0} + a.c.\mathbf{0}$ and q the process $p + a.(b.\mathbf{0} \frac{1}{2} \oplus c.\mathbf{0})$ respectively. Then $p \not\sim_{ls} q$.

Nevertheless it is not unreasonable to expect that p and q should be behaviourally equivalent. This intuition comes from an interpretation of non-deterministic choice based on *schedulers*. So consider the reason why $p \not\sim_{ls} q$; the move $q \xrightarrow{a} (b.\mathbf{0} \frac{1}{2} \oplus c.\mathbf{0})$ can not be matched by a corresponding move from p . But consider a scheduler for p which when deciding which a action to perform, half the time chooses the first possibility and half the time the second. The effect of this scheduler on p could be formalised by allowing the move $p \xrightarrow{a} (b.\mathbf{0} \frac{1}{2} \oplus c.\mathbf{0})$, thereby allowing the missing corresponding move. ■

Definition 3.10 (Combined moves [Seg95, PS07]) In a pLTS let $s \xrightarrow{\mu}_{cc} \Delta$ whenever $\Delta \in cc(s \text{ after } \mu)$. Note that if $s \xrightarrow{\mu} \Delta$ for no Δ then also $s \xrightarrow{\mu}_{cc} \Delta$ for no Δ . ■

The intuition here is that we should allow the moves $s \xrightarrow{\mu}_{cc} \Delta$ when playing bisimulation games because there is some (probabilistic) scheduler which can choose between the set of possible residuals (s after μ) to obtain any distribution in $cc(s \text{ after } \mu)$.

This leads to a much more mainstream definition of probabilistic bisimulations.

Definition 3.11 For an equivalence relation $\mathcal{R} \subseteq (S \times S)$, where S is the set of states in some pLTS, $\mathcal{B}_{cls}(\mathcal{R})$ be the relation over $S \times S$ determined by letting $s \mathcal{B}_{cls}(\mathcal{R}) t$ whenever for every $\mu \in \text{Act}$,

- (a) $s \xrightarrow{\mu} \Delta$ implies $t \xrightarrow{\mu}_{cc} \Theta$ for some Θ such that $\Delta \text{ slift}(\mathcal{R}) \Theta$
- (b) conversely, $t \xrightarrow{\mu} \Theta$ implies $s \xrightarrow{\mu}_{cc} \Delta$ for some Δ such that $\Delta \text{ slift}(\mathcal{R}) \Theta$.

We use the notation $\sim_{cls} \subseteq S \times S$, to denote the largest equivalence relation such that $\sim_{cls} \subseteq \mathcal{B}_{cls}(\sim_{cls})$. Slight variations on this definition is called *strong probabilistic bisimulation equivalence* in [Seg95, SL95, PS07, DvG10]. ■

There is one further critique we can make of this formulation of bisimulation equivalence. The defining functional $\mathcal{B}_{cls}(-)$ is only applied to equivalence relations. This means that by definition \sim_{cls} is of course an equivalence relation over states. But another consequence is that witness bisimulations used to prove process equivalences must in turn be equivalence relations over the state space. This is very restricting. And indeed this restriction to equivalence relations in the definition of probabilistic bisimulations is unnecessary. The following definition is taken from [DvG10].

Definition 3.12 (*Strong s-bisimulations*) For any $\mathcal{R} \subseteq (S \times S)$, where S is the set of states in some pLTS, $\mathcal{B}_{ss}(\mathcal{R})$ be the relation over $S \times S$ determined by letting $s \mathcal{B}_{ss}(\mathcal{R}) t$ whenever for every $\mu \in \text{Act}$,

- (a) $s \xrightarrow{\mu} \Delta$ implies $t \xrightarrow{\mu}_{cc} \Theta$ for some Θ such that $\Delta \text{ slift}(\mathcal{R}) \Theta$
- (b) conversely, $t \xrightarrow{\mu} \Theta$ implies $s \xrightarrow{\mu}_{cc} \Delta$ for some Δ such that $\Delta \text{ slift}(\mathcal{R}) \Theta$.

We say \mathcal{R} is a *strong s-bisimulation* if $\mathcal{R} \subseteq \mathcal{B}_{ss}(\mathcal{R})$ and we use $\sim_{ss} \subseteq S \times S$ to denote the largest such relation. ■

The advantage of this formulation is that witness bisimulations, used to establish an equivalence between processes, can be chosen arbitrarily. But a potential disadvantage is that it is not obvious that \sim_{ss} is an equivalence relation.

Proposition 3.13 *In an arbitrary pLTS,*

- (a) \sim_{ss} is an equivalence relation
- (b) $s \sim_{ss} t$ if and only if $s \sim_{cls} t$.

Proof Let \mathcal{R} denote the relational composition $\sim_{ss} \circ \sim_{ss}$. Part (a) follows if we can show $\mathcal{R} \subseteq \mathcal{B}_{ss}(\mathcal{R})$. Suppose $s \mathcal{R} t$ and $s \xrightarrow{\mu} \Delta$. By definition there exists some state r such that $s \sim_{ss} r$ and $r \sim_{ss} t$. Therefore using the transfer properties of \sim_{ss} there exist moves $r \xrightarrow{\mu} \Gamma$ and $t \xrightarrow{\mu} \Theta$ such that $\Delta \text{ slift}(\sim_{ss}) \Gamma$ and $\Gamma \text{ slift}(\sim_{ss}) \Theta$. It follows from Lemma 2.6 that $\Delta \text{ slift}(\mathcal{R}) \Theta$, and therefore we have matched the move from s by a move from t . A similar argument will show that moves from t can also be matched appropriately by s .

Part (b) now follows immediately; $\sim_{cls} \subseteq \sim_{ss}$ because by definition $\sim_{cls} \subseteq \mathcal{B}_{ss}(\sim_{cls})$, while because \sim_{ss} is an equivalence relation we also have $\sim_{ss} \subseteq \mathcal{B}_{cls}(\sim_{ss})$ from which $\sim_{ss} \subseteq \sim_{cls}$ follows. □

Example 3.14 Consider the pLTS given in Fig. 2, where we use an obvious notation for depicting distributions. Let \mathcal{R} be the relation given by

$$\mathcal{R} = \{(s_0, t_0), (s_1, t_1), (s_1, t_2), (s_2, t_3), (s_3, t_3)\} \cup \{(\mathbf{0}, \mathbf{0})\}$$

It will follow that $s_0 \sim_{ss} t_0$ if we can show that \mathcal{R} is a strong bisimulation, as defined in Definition 3.12.

Let Δ, Θ denote the distributions $\frac{1}{2} \cdot \bar{s}_1 + \frac{1}{4} \cdot \bar{s}_2 + \frac{1}{4} \cdot \bar{s}_3$ and $\frac{1}{4} \cdot \bar{t}_1 + \frac{1}{4} \cdot \bar{t}_2 + \frac{1}{4} \cdot \bar{t}_3$ respectively; note that $\Delta \text{ slift}(\mathcal{R}) \Theta$. Since $s_0 \xrightarrow{a} \Delta$ and $t_0 \xrightarrow{a} \Theta$ it follows that the moves from s_0 and t_0 can be appropriately matched. All the other requirements, on the other pairs in \mathcal{R} , are trivially fulfilled.

Note incidently that \mathcal{R} is obviously not an equivalence relation. ■



Fig. 2. An example

In view of Proposition 3.13 it is reasonable to focus on Definition 3.12 as an appropriate definition of (strong) bisimulation for pLTSs. However in a pLTS it is more natural to consider behaviour in terms of distributions instead of states, for at least two reasons. The natural interpretation of process terms, such as those in bpCCS, are interpreted as distributions over the states of the underlying pLTS; further, as computations of interactions evolve distributions rather than states are obtained as residuals.

To change the focus from states to distributions we use the *lifting* of the actions $s \xrightarrow{\mu} \Delta'$, relations in $S \times \mathcal{D}(S)$ to $\Delta \text{ lift}(\xrightarrow{\mu}) \Delta'$, relations in $\mathcal{D}(S) \times \mathcal{D}(S)$, as given in Definition 2.2; to render the notation less forbidding we abbreviate $\Delta \text{ lift}(\xrightarrow{\mu}) \Delta'$ to $\Delta \xrightarrow{\mu} \Delta'$. Note that this endows the set of distributions over the state space of a pLTS with the structure of a standard LTS.

Lemma 3.15 *In a pLTS, $\bar{s} \xrightarrow{\mu} \Delta$ if and only if $s \xrightarrow{\mu}_{cc} \Delta$.*

Proof Straightforward, using Lemma 2.3. □

Now we can reformulate the bisimulation equivalence in Definition 3.12 purely in terms of distributions.

Definition 3.16 (*Strong d-bisimulations*) For $\mathcal{R} \subseteq (\mathcal{D}(S) \times \mathcal{D}(S))$, where S is the set of states in some pLTS, $\mathcal{B}_{sd}(\mathcal{R})$ is the relation over $\mathcal{D}(S) \times \mathcal{D}(S)$ determined by letting $\Delta \mathcal{B}_{sd}(\mathcal{R}) \Theta$ whenever

- (a) for every $\mu \in \text{Act}_\tau$, $\Delta \xrightarrow{\mu} \Delta'$ implies $\Theta \xrightarrow{\mu} \Theta'$ for some Θ' such that $\Delta' \mathcal{R} \Theta'$
- (b) conversely, for every $\mu \in \text{Act}$, $\Theta \xrightarrow{\mu} \Theta'$ implies $\Delta \xrightarrow{\mu} \Delta'$ for some Δ' such that $\Delta' \mathcal{R} \Theta'$
- (c) $\Delta = \Delta_1 \oplus \Delta_2$ implies $\Theta = \Theta_1 \oplus \Theta_2$ such that $\Delta_i \mathcal{R} \Theta_i$ for $i = 1, 2$
- (d) conversely, $\Theta = \Theta_1 \oplus \Theta_2$ implies $\Delta = \Delta_1 \oplus \Delta_2$ such that $\Delta_i \mathcal{R} \Theta_i$ for $i = 1, 2$.

We use $\sim_{sd} \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$ to denote the largest relation such that $\sim_{sd} \subseteq \mathcal{B}_{sd}(\sim_{sd})$. ■

Note that by definition \sim_{sd} is both left and right decomposable; moreover a standard augment shows that it is an equivalence relation. However because of the complexity of the underlying LTS of distributions it is not straightforward to exhibit strong d-bisimulations.

Theorem 3.17 *In an arbitrary pLTS $\Delta \sim_{sd} \Theta$ if and only if $\Delta \text{ slift}(\sim_{ss}) \Theta$.*

Proof First suppose $\Delta \sim_{sd} \Theta$; the proof that $\Delta \text{ slift}(\sim_{ss}) \Theta$ follows in three steps. Let $\mathcal{R} \subseteq S \times S$ be defined by letting $s \mathcal{R} t$ whenever $\bar{s} \sim_{sd} \bar{t}$.

- (a) Suppose $\bar{s} \sim_{sd} \Theta$. Using the fact that \sim_{sd} is right decomposable it follows that $s \mathcal{R} t$ for every $t \in [\Theta]$, and therefore $\bar{s} \text{ slift}(\mathcal{R}) \Theta$.
- (b) More generally suppose $\Delta \sim_{sd} \Theta$. Writing Δ as $\sum_{s \in [\Delta]} \Delta(s) \cdot \bar{s}$ and using the left-decomposability of \sim_{sd} we see that Θ can be written as $\sum_{s \in [\Delta]} \Delta(s) \cdot \Theta_s$, where $\bar{s} \sim_{sd} \Theta_s$, or using (a), where $\bar{s} \text{ slift}(\mathcal{R}) \Theta_s$. It follows by Linearity of $\text{slift}(-)$ that $\Delta \text{ slift}(\mathcal{R}) \Theta$.
- (c) Using (b) we now show that $\mathcal{R} \subseteq \mathcal{B}_{ss}(\mathcal{R})$. Suppose $s \mathcal{R} t$ and $s \xrightarrow{\mu} \Delta$. Since by definition $\bar{s} \sim_{sd} \bar{t}$ it follows that $\bar{t} \xrightarrow{\mu} \Theta$ such that $\Delta \sim_{sd} \Theta$. From (b) we know $\Delta \text{ slift}(\mathcal{R}) \Theta$ and therefore by Lemma 3.15 we have the required matching move, $p \xrightarrow{\mu}_{cc} \Theta$. The proof that moves from t can be matched by s is identical.

From (c) we know that $\mathcal{R} \subseteq \sim_{ss}$ and it now follows from (b) that $\Delta \sim_{sd} \Theta$ implies $\Delta \text{slift}(\sim_{ss}) \Theta$.

To prove the converse it is sufficient to show that $\text{slift}(\sim_{ss})$ is a strong d-bisimulation, that is $\text{slift}(\sim_{ss}) \subseteq \mathcal{B}_{sd}(\sim_{ss})$. Conditions (c) and (d) of Definition 3.16 follow from the fact that $\text{slift}(\mathcal{R})$, for any relation \mathcal{R} , is both left- and right-decomposable. We show how condition (a) is satisfied; the proof for (b) is symmetric.

Suppose $\Delta \text{slift}(\sim_{ss}) \Theta$ and $\Delta \xrightarrow{\mu} \Delta'$. By the second part of Lemma 2.5 we know

$$\Delta = \sum_{i \in I} p_i \cdot \bar{s}_i, \quad \Theta = \sum_{i \in I} p_i \cdot \bar{t}_i$$

where for each $i \in I$, $s_i \sim_{ss} t_i$.

The arrows $\xrightarrow{\mu}$ are lifted relations, and therefore left-decomposable. So we know Δ' can be written as $\sum_{i \in I} p_i \cdot \Delta'_i$ where $s_i \xrightarrow{\mu} \Delta'_i$ for each $i \in I$.

Using the fact that $s_i \sim_{ss} t_i$ each of these moves can be matched by a move $t_i \xrightarrow{\mu} \Theta'_i$, or via Lemma 3.15 $\bar{t}_i \xrightarrow{\mu} \Theta'_i$, such that $\Delta'_i \text{slift}(\sim_{ss}) \Theta'_i$.

Now letting Θ' denote $\sum_{i \in I} p_i \cdot \Theta'_i$ we have by the linearity of lifted relations that $\Theta \xrightarrow{\mu} \Theta'$, which is the required matching move, since linearity again gives $\Delta' \text{slift}(\sim_{ss}) \Theta'$. \square

The proof of this theorem made extensive use of the fact that \sim_{sd} is decomposable, and indeed conditions (c) and (d) in Definition 3.16 are essential. Without them the trivial process $\mathbf{0}$ would be identified with $(a \frac{1}{2} \oplus b)$, since the latter distribution can not perform any action. This same example demonstrates that using the standard definition of bisimulation equivalence on the derived LTS of distributions also does not capture \sim_{ss} .

4. A finitary distinguishing logic

We use a slight extension of the familiar HML logic [Mil89], denoted pHML, whose formulae are defined as follows:

$$\varphi ::= \text{tt} \mid \varphi_1 \vee \varphi_2 \mid \langle \mu \rangle \varphi, \mu \in \text{Act}_\tau \mid \neg \varphi \mid \varphi_{1_p} \oplus \varphi_2, p \in [0, 1]$$

The satisfaction relation $\Delta \models \varphi$, where Δ is a distribution over the states of a pLTS, is defined by induction on φ :

- $\Delta \models \text{tt}$ for all Δ
- $\Delta \models \varphi_1 \vee \varphi_2$ if $\Delta \models \varphi_1$ or $\Delta \models \varphi_2$
- $\Delta \models \langle \mu \rangle \varphi$ if $\Delta \xrightarrow{\mu} \Delta'$ such that $\Delta' \models \varphi$
- $\Delta \models \varphi_{1_p} \oplus \varphi_2$ if $\Delta = \Delta_{1_p} \oplus \Delta_2$ for some Δ_i such that $\Delta_i \models \varphi_i$
- $\Delta \models \neg \varphi$ whenever $\Delta \not\models \varphi$

Let $\text{pHML}(\Delta) = \{ \varphi \mid \Delta \models \varphi \}$. We have the usual derived operators:

- **ff** for $\neg \text{tt}$
- $\varphi_1 \wedge \varphi_2$ for $\neg (\neg \varphi_1 \vee \neg \varphi_2)$
- $[\mu]\varphi$ for $\neg \langle \mu \rangle \neg \varphi$

So for example $\Delta \models [\mu]\varphi$ if $\Delta \xrightarrow{\mu} \Delta'$ implies $\Delta' \models \varphi$.

Proposition 4.1 $\Delta \sim_{sd} \Theta$ implies $\text{pHML}(\Delta) = \text{pHML}(\Theta)$.

Proof Suppose $\Delta \sim_{sd} \Theta$ and $\Delta \models \varphi$. Then a straightforward argument by induction on φ will show that $\Theta \models \varphi$. We give one example.

Suppose $\varphi = \varphi_{1_p} \oplus \varphi_2$. $\Delta \models \varphi$ implies $\Delta = \Delta_{1_p} \oplus \Delta_2$ such that $\Delta_i \models \varphi_i$, $i = 1, 2$. Since \sim_{sd} is decomposable we must have $\Theta = \Theta_{1_p} \oplus \Theta_2$ such that $\Delta_i \sim_{sd} \Theta_i$. By induction we have $\Theta_i \models \varphi_i$ and therefore we can conclude that $\Theta \models \varphi$. \square

Example 4.2 Consider the two process terms $P_1 = a_{\frac{1}{2}} \oplus (b + c)$ and $Q_1 = (a_{\frac{1}{2}} \oplus b) + (a_{\frac{1}{2}} \oplus c)$. We have

$$\llbracket P_1 \rrbracket \models [a]\text{ff } \frac{1}{2} \oplus \text{tt} \quad \text{but} \quad \llbracket Q_1 \rrbracket \not\models [a]\text{ff } \frac{1}{2} \oplus \text{tt}$$

Consequently $\llbracket P_1 \rrbracket \not\sim_{sd} \llbracket Q_1 \rrbracket$.

Similarly we can show $\llbracket P_2 \rrbracket \not\sim_{sd} \llbracket Q_2 \rrbracket$, where $P_2 = b.c + b.d$, $Q_2 = b.c_{\frac{1}{2}} \oplus b.d$, by providing a distinguishing formula. For example the formula

$$\langle b \rangle (\langle c \rangle \text{tt } \frac{1}{2} \oplus \text{tt}) \wedge (\langle d \rangle \text{tt } \frac{1}{2} \oplus \text{tt})$$

is satisfied by Q_2 but not P_2 . ■

In fact Proposition 4.1 can be strengthened to

Theorem 4.3 *In a finitary pLTS, $\Delta \sim_{sd} \Theta$ if and only if $\text{pHML}(\Delta) = \text{pHML}(\Theta)$.*

The proof uses characteristic formulae for distributions, not relative to \sim_{sd} but rather their inductive approximations, \sim_{ss}^k for all $k \geq 0$, to the state based equivalence \sim_{ss} . These are defined as follows:

- $s \sim_{ss}^0 t$ for all $s, t \in S$
- $s \sim_{ss}^{(k+1)} t$ whenever $s \mathcal{B}_{ss}(\sim_{ss}^k) t$.

To define the characteristic formulae we need to extend the modal formulae slightly, by allowing into the grammar the extra clause

$$\varphi ::= \dots \mid \varphi_1 \oplus \varphi_2 \mid \dots$$

where $\Delta \models \varphi_1 \oplus \varphi_2$ is taken to hold whenever there exists some $0 < p < 1$ such that $\Delta = \Delta_{1_p} \oplus \Delta_2$ and $\Delta_i \models \varphi_i$, $i = 1, 2$. We denote the extended logic by pHML^+ and define $\text{pHML}^+(\Delta)$ in the obvious manner.

Lemma 4.4 $\text{pHML}(\Delta) = \text{pHML}(\Theta)$ if and only if $\text{pHML}^+(\Delta) = \text{pHML}^+(\Theta)$.

Proof It suffices to show that if $\text{pHML}(\Delta) = \text{pHML}(\Theta)$ then

$$\text{for every } \varphi \in \text{pHML}^+, \Delta \models \varphi \text{ if and only if } \Theta \models \varphi \tag{4}$$

Let $\varphi_1 \prec \varphi_2$ if

- $|\varphi_1| < |\varphi_2|$; here we are referring to the number of symbols in φ_i
- or φ_2 has the form $\varphi^1 \oplus \varphi^2$ and $\varphi_1 = \varphi^1_p \oplus \varphi^2$ for some $0 < p < 1$.

It is easy to check that \prec is well-founded; so we prove (4) above by induction on this ordering. We give two examples.

- Suppose φ is $\varphi_1 \oplus \varphi_2$. If $\Delta \models \varphi$ then there exists some $0 < p < 1$ such that $\Delta = \Delta_{1_p} \oplus \Delta_2$ where $\Delta_i \models \varphi_i$. This means that $\Delta \models \varphi_{1_p} \oplus \varphi_2$ and by induction, since $\varphi_{1_p} \oplus \varphi_2 \prec \varphi_1 \oplus \varphi_2$, we know $\Theta \models \varphi_{1_p} \oplus \varphi_2$, and therefore that $\Theta \models \varphi$.
If $\Theta \models \varphi$ then the argument is the same.
- Suppose φ is $\neg \varphi_1$. Here if $\Delta \models \varphi$ then $\Delta \not\models \varphi_1$. So by induction $\Theta \not\models \varphi_1$, which means $\Theta \models \varphi$.
The argument when $\Theta \models \varphi$ is symmetrical. □

The characteristic formulae use some derived operators, in addition to those mentioned above. For any finite non-empty set of probabilities $\{p_i \mid 1 \leq i \leq n\}$ summing up to 1, we use $\sum\{p_i \cdot \varphi_i \mid 1 \leq i \leq n\}$ to denote the obvious generalisation of the binary (definite) probabilistic choice $\varphi_{1_p} \oplus \varphi_2$; all instances can easily be rendered in the base language pHML^+ . On the other hand for an arbitrary finite index set I , possibly empty, we use $\bigoplus_{i \in I} \varphi_i$ to denote the generalisation of the binary (indefinite) choice operator, $\varphi_1 \oplus \varphi_2$. If I is empty this denotes ff .

Otherwise it has the obvious interpretation; for example if $I = \{1, 2, 3\}$ then it denotes $\varphi_1 \oplus (\varphi_2 \oplus \varphi_3)$. This notation is simplified to $\oplus \mathcal{F}$ where \mathcal{F} is a finite set of formulae.

We use \oplus to define an even more complicated derived formula. Let \mathcal{F} be an arbitrary finite set of formulae. We define an operator $\text{pch}(-)$ such that $\Delta \models \text{pch}(\mathcal{F})$ if and only if there exists a finite subset $\{\varphi_1, \dots, \varphi_n\}$ of \mathcal{F} such that $\Delta \models \varphi_1 \oplus \dots \oplus \varphi_n$. This is achieved by letting

$$\text{pch}(\mathcal{F}) = \bigvee \{ \oplus S \mid S \subseteq \mathcal{F} \}$$

Note that if the set \mathcal{F} is empty then this disjunction is the formula ff .

Finally we use $\text{sub}(\varphi)$ as an abbreviation for $\neg(\neg\varphi \oplus \text{tt})$. Note that $\Delta \models \text{sub}(\varphi)$ means that for any $0 < p < 1$ and for any Δ' such that $\Delta = \Delta'_p \oplus \Delta''$ for some distribution Δ'' , $\Delta' \models \varphi$. So for example $\Delta \models \text{sub}(\varphi)$ ensures that $\Delta \models \varphi$; it also ensures that $\bar{s} \models \varphi$ for every $s \in [\Delta]$, reminiscent of the primitive operator \downarrow in the logic of [DvG10]; here however it is derived. Also in pHML the fact that $\bar{s} \models \varphi$ for every s in the support of Δ does not ensure that $\Delta \models \text{sub}(\varphi)$, as the following example shows.¹

Example 4.5 Let φ be the formula $\neg((a)\text{tt} \frac{1}{2} \oplus (b)\text{tt})$ and Δ the distribution $(\bar{a} \frac{1}{2} \oplus \bar{b})$. Then $\bar{a} \models \varphi$ and $\bar{b} \models \varphi$ but $\Delta \not\models \varphi$. This in turn means that $\Delta \not\models \text{sub}(\varphi)$. ■

We now define the characteristic formulae $\varphi^k(s)$ and $\Phi^k(\Delta)$ by induction on k :

$$k = 0: \varphi^0(s) = \Phi^0(\Delta) = \text{tt}$$

$$k + 1: \varphi^{(k+1)}(s) = \text{Can}(k + 1, s) \wedge \text{Must}(k + 1, s) \text{ and } \Phi^{(k+1)}(\Delta) = \bigoplus \{ \Delta(s) \cdot \text{sub}(\varphi^{(k+1)}(s)) \mid s \in [\Delta] \}$$

where

$$\text{Can}(k, s) = \bigwedge_{s \xrightarrow{\mu} \Delta} \langle \mu \rangle \Phi^k(\Delta) \quad \text{Must}(k, s) = \bigwedge_{\mu \in \text{Act}_\tau} [\mu] \text{pch}(\{ \Phi^k(\Delta) \mid s \xrightarrow{\mu} \Delta \})$$

Note that the construction of these (finite) formulae rely on the fact that underlying pLTS is finitary.

Proposition 4.6 For all $k \geq 0$

- (a) $\bar{s} \models \varphi^k(s)$ and $\Delta \models \Phi^k(\Delta)$
- (b) $\bar{t} \models \varphi^k(s)$ implies $s \sim_{ss}^k t$
- (c) $\Theta \models \Phi^k(\Delta)$ implies $\Delta \text{slift}(\sim_{ss}^k) \Theta$.

Proof By induction on k , with the case $k = 0$ being trivial. So suppose all three statements are true for k we show that this implies they are also true for $(k + 1)$.

(a) A straightforward inductive argument. Notice in particular that if $s \xrightarrow{\mu} \Theta$ then the formula $\Phi^k(\Theta)$ is a disjunct of $\text{pch}(\{ \Phi^k(\Delta) \mid s \xrightarrow{\mu} \Delta \})$; this makes it easy to see that $\bar{s} \models \text{Must}(k + 1, s)$.

(b) Suppose $\bar{t} \models \varphi^{(k+1)}(s)$; we have to show $s \sim_{ss}^{(k+1)} t$. First suppose $s \xrightarrow{\mu} \Delta$. Because $\bar{t} \models \text{Can}(k + 1, s)$ we have $\bar{t} \xrightarrow{\mu} \Theta$, and therefore by Lemma 3.15, $t \xrightarrow{\mu}_{cc} \Theta$, such that $\Theta \models \Phi^k(\Delta)$; induction on (c) gives the required $\Delta \text{slift}(\sim_{ss}^k) \Theta$.

Conversely suppose $t \xrightarrow{\mu}_{cc} \Theta$, which implies $\bar{t} \xrightarrow{\mu} \Theta$. Because $\bar{t} \models \text{Must}(k + 1, s)$ we have that $\Theta \models \text{pch}(\{ \Phi^k(\Delta) \mid s \xrightarrow{\mu} \Delta \})$.

This in turn means that there exists $\Delta_1, \dots, \Delta_n$, $n \geq 1$, such that $s \xrightarrow{\mu} \Delta_i$ and p_1, \dots, p_n , each greater than 0, such that $\Theta \models \bigoplus \{ p_i \cdot \Phi^k(\Delta_i) \mid 1 \leq i \leq n \}$. Let $\Delta = \sum_{1 \leq i \leq n} p_i \cdot \Delta_i$; note that by definition $s \xrightarrow{\mu}_{cc} \Delta$ since $\Delta \in cc(\{ \Gamma \mid s \xrightarrow{\mu} \Gamma \})$.

By definition of the derived operator we therefore have $\Theta = \sum_{1 \leq i \leq n} p_i \cdot \Theta_i$ such that $\Theta_i \models \Phi^k(\Delta_i)$. Induction now yields $\Delta_i \text{slift}(\sim_{ss}^k) \Theta_i$ and the required $\Delta \text{slift}(\sim_{ss}^k) \Theta$ follows by the linearity of lifting.

¹ Notice that in the definition of $\Delta \models \varphi_1 \oplus \varphi_2$ if we had allowed the probability p to range from $0 \leq p \leq 1$, rather than $0 < p < 1$, then $\text{sub}(\varphi)$ would be logically equivalent to tt .

$$\begin{array}{c}
\text{(PAR.L)} \\
\frac{s \xrightarrow{\mu} \Delta}{s \mid t \xrightarrow{\mu} \Delta \mid \bar{t}} \quad \mu \in \text{Act}_\tau \\
\\
\text{(PAR.I)} \\
\frac{s \xrightarrow{a} \Delta, t \xrightarrow{\bar{a}} \Theta}{s \mid t \xrightarrow{\tau} \Delta \mid \Theta} \quad a \in \text{Act}
\end{array}
\qquad
\begin{array}{c}
\text{(PAR.R)} \\
\frac{t \xrightarrow{\mu} \Theta}{s \mid t \xrightarrow{\mu} \bar{s} \mid \Theta} \quad \mu \in \text{Act}_\tau
\end{array}$$

Fig. 3. Parallel composition of pLTSs

(c) Suppose $\Theta \models \Phi^{(k+1)}(\Delta)$, which means that $\Theta = \sum_{s \in [\Delta]} \Delta(s) \cdot \Theta_s$ such that $\Theta_s \models \text{sub}(\varphi^{(k+1)}(s))$. So for any state $t \in [\Theta_s]$ we know $\bar{t} \models \varphi^{(k+1)}(s)$ and therefore by (b), $s \sim_{ss}^{(k+1)} \bar{t}$. Linearity of lifting now gives $\bar{s} \text{slift}(\sim_{ss}^{(k+1)}) \Theta_s$, and applying linearity once more we obtain the required $\Delta \text{slift}(\sim_{ss}^{(k+1)}) \Theta$. \square

Proof of Theorem 4.3: Because of Proposition 4.1 and Lemma 4.4 it is sufficient to show $\text{pHML}^+(s) = \text{pHML}^+(t)$ implies $s \sim_{ss} t$. However in a finitary pLTS the set $S \times S$ is finite, and since $\sim_{ss}^{k+1} \subseteq \sim_{ss}^k$ there must be some $n \geq 0$ such that $s \sim_{ss}^n t$ implies $s \sim_{ss}^m t$ for all $m \geq n$. This in turn means that $s \sim_{ss}^n t$ implies $s \sim_{ss} t$, which further implies that $\Delta \text{slift}(\sim_{ss}^n) \Theta$ implies $\Delta \text{slift}(\sim_{ss}) \Theta$. Note that the particular n depends on the pLTS in hand.

Now suppose $\text{pHML}^+(\Delta) = \text{pHML}^+(\Theta)$. Proposition 4.6 (c) immediately gives $\Delta \text{slift}(\sim_{ss}^n) \Theta$, and therefore $\Delta \text{slift}(\sim_{ss}) \Theta$.

5. Behavioural justification

Here we develop the contextual equivalence for probabilistic processes alluded to in point (3) on page 749 of Sect. 1. Let us assume that the set of actions Act has an inverse operation, with \bar{a} giving the inverse of a ; this should satisfy the property $\bar{\bar{a}} = a$. Intuitively the simultaneous occurrence of the action a and its inverse \bar{a} constitutes a synchronisation between two independent components. Now given two pLTSs $\mathcal{P}_1, \mathcal{P}_2$ we can define a new pLTS, which we refer to as $\mathcal{P}_1 \mid \mathcal{P}_2$, whose states are all of the form $s_1 \mid s_2$ where s_i are states in the pLTS \mathcal{P}_i and whose next-state relations are determined by the rules in Fig. 3; note that the rule (PAR.I) captures the intuitive idea of synchronisation. All rules require the parallel operator to be lifted to distributions. This is done in the standard manner, similar to the lifting of $+$ in (3) on page 755:

$$(\Delta \mid \Theta)(s) = \begin{cases} \Delta(s_1) \cdot \Theta(s_2), & \text{if } s = s_1 \mid s_2 \\ 0, & \text{otherwise} \end{cases}$$

Definition 5.1 We say a pLTS \mathcal{P} is *par-closed* if $\mathcal{P} \mid \mathcal{P}$ is already a sub-pLTS of \mathcal{P} . We say it is *sufficiently expressive* if it contains a denotation for every term in bpCCS. \blacksquare

To construct an example of such a pLTS let us extend the set of state terms in (2) by

$$s \in \text{fpCCS}_s ::= \dots \mid s \mid s \mid \dots$$

thereby obtaining the language of finite probabilistic CCS, namely fpCCS. The rules governing the behaviour of the new states $s \mid t$ are those already given in Fig. 3. The resulting pLTS has an interpretation for every process term in fpCCS as a distribution; moreover it is par-closed and sufficiently expressive.

Theorem 5.2 (Compositionality of \sim_{sd}) *In an arbitrary par-closed pLTS, $\Delta_1 \sim_{sd} \Delta_2$ implies $\Delta_1 \mid \Theta \sim_{sd} \Delta_2 \mid \Theta$.*

Proof It is difficult to give a direct proof in terms of \sim_{sd} ; instead we use the state-based formulation of bisimulation and proceed in three steps.

(a) Let \mathcal{R} be the relation over $S \times S$ defined by

$$\{(s_1 \mid t), (s_2 \mid t) \mid s_1 \sim_{ss} s_2\}$$

A straightforward argument will show that $\mathcal{R} \subseteq \mathcal{B}_{ss}(\mathcal{R})$. As an example of the transfer properties required we consider one case. Suppose $s_1 \mid t \mathcal{R} s_2 \mid t$ and $s_1 \mid t \xrightarrow{\tau} \Delta_1 \mid \Theta$ because $s_1 \xrightarrow{a} \Delta_1$ and $t \xrightarrow{a} \Theta$. We have to find a matching move from $s_2 \mid t$.

Since $s_1 \sim_{ss} s_2$ we know that $s_2 \xrightarrow{a} \Delta_2$ such that $\Delta_1 \text{slift}(\mathcal{R}) \Delta_2$. This move can be combined with that from t to obtain $s_2 \mid t \xrightarrow{\tau} \Delta_2 \mid \Theta$. So it remains to see that

$$(\Delta_1 \mid \Theta) \mathcal{R} (\Delta_2 \mid \Theta) \tag{5}$$

First notice that $\Delta_1 \text{slift}(\mathcal{R}) \Delta_2$ implies $(\Delta_1 \mid \bar{r}) \text{slift}(\mathcal{R}) (\Delta_2 \mid \bar{r})$ for any state r . Since $(\Delta_i \mid \Theta)$ can be written as $\sum_{r \in [\Theta]} \Theta(r) \cdot (\Delta_i \mid \bar{r})$, (5) now follows from the linearity of lifting.

So we have established $s_1 \sim_{ss} s_2$ implies $s_1 \mid t \sim_{ss} s_2 \mid t$.

(b) Now suppose $\Delta_1 \text{slift}(\sim_{ss}) \Delta_2$. We show that for any state t this implies $\Delta_1 \mid \bar{t} \text{slift}(\sim_{ss}) \Delta_2 \mid \bar{t}$. We have

$$\Delta_1 = \sum_{i \in I} p_i \cdot \bar{s}_{1i} \quad \Delta_2 = \sum_{i \in I} p_i \cdot \bar{s}_{2i} \quad \text{where } s_{1i} \sim_{ss} s_{2i} \text{ for each } i \in I$$

Part (a) gives $s_{1i} \mid t \sim_{ss} s_{2i} \mid t$ and so we also have the decompositions

$$\Delta_1 \mid \bar{t} = \sum_{i \in I} p_i \cdot \overline{s_{1i} \mid t} \quad \Delta_2 \mid \bar{t} = \sum_{i \in I} p_i \cdot \overline{s_{2i} \mid t} \quad \text{where } s_{1i} \mid t \sim_{ss} s_{2i} \mid t \text{ for each } i \in I.$$

This means $\Delta_1 \mid \bar{t} \text{slift}(\sim_{ss}) \Delta_2 \mid \bar{t}$.

(c) For any Θ the distribution $\Delta \mid \Theta$ can be written as $\sum_{t \in \Theta} \Delta \mid \bar{t}$. Therefore using linearity and part (b) we can prove that $\Delta_1 \sim_{ss} \Delta_2$ implies $\Delta_1 \mid \Theta \text{slift}(\sim_{ss}) \Delta_2 \mid \Theta$.

The result now follows from Theorem 3.17 □

We now turn our attention to the precise formulation of contextual equivalence alluded to in the Introduction. Following the well-established approach of [HY95, FG05, SKS07] we define a touchstone equivalence in terms of desirable properties which we would expect of a reasonable behavioural equivalence. The first is *compositionality* which we have already formulated, in Theorem 5.2. The second is a formalisation of the intuitive idea that processes should have similar non-deterministic choice structures.

Definition 5.3 A relation \mathcal{R} over distributions in a pLTS is *s-reduction-closed* if $\Delta \mathcal{R} \Theta$ implies:

- if $\Delta \xrightarrow{\tau} \Delta'$ then $\Theta \xrightarrow{\tau} \Theta'$ such that $\Delta' \mathcal{R} \Theta'$
- conversely, if $\Theta \xrightarrow{\tau} \Theta'$ then $\Delta \xrightarrow{\tau} \Delta'$ such that $\Delta' \mathcal{R} \Theta'$. ■

The final property used is that processes should react in the same way to a collection of primitive experiments or observations, often referred to as *barbs*. Formally a barb b is a predicate on processes, which in pLTSs are distributions. There are a number of possibilities for barbs in pLTSs but for the moment let us just assume some arbitrary collection B of barbs. We say the relation $\mathcal{R} \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$ preserves the barbs in B if whenever $\Delta \mathcal{R} \Theta$ then for every barb $b \in B$, Δ satisfies b if and only if Θ satisfies b .

Definition 5.4 [Strong reduction barbed closure] In an arbitrary pLTS let \sim_{rbc}^B be the largest relation over distributions which is compositional, s-reduction closed, and preserves the barbs in B . ■

One natural collection of barbs can be defined as follows: for every $a \in \text{Act}$ and $p \in [0, 1]$ let us write $\Delta \downarrow_{\geq p}^a$ if $\sum\{\Delta(s) \mid s \xrightarrow{a}\} \geq p$. Let sG denote this set of barbs, $\{\downarrow_{\geq p}^a \mid a \in \text{Act}\}$.

There are other variations on this theme. For example $\Delta \downarrow_{\leq p}^a$ if $\sum\{\Delta(s) \mid s \xrightarrow{a}\} \leq p$, and $\Delta \downarrow_{=p}^a$ if $\sum\{\Delta(s) \mid s \xrightarrow{a}\} = p$. We use sL and sE to denote the sets $\{\downarrow_{< p}^a \mid a \in \text{Act}\}$ and $\{\downarrow_{=p}^a \mid a \in \text{Act}\}$ respectively.

Proposition 5.5 *In an arbitrary $p\text{LTS}$, the three relations \sim_{rbc}^{sG} , \sim_{rbc}^{sL} and \sim_{rbc}^{sE} coincide.*

Proof It is a question of checking that the sets of barbs can simulate each other. For example \sim_{rbc}^{sG} and \sim_{rbc}^{sL} coincide by virtue of the fact that $\Delta \downarrow_{\leq p}^a$ if and only not $\Delta \downarrow_{< p}^a$.

Also $\Delta \sim_{rbc}^{\text{sG}} \Theta$ implies $\Delta \sim_{rbc}^{\text{sE}} \Theta$ because $\Gamma \downarrow_{=p}^a$ if and only if $\Gamma \downarrow_{\geq p}^a$ and not $\Gamma \downarrow_{< p}^a$.

Finally suppose $\Delta \sim_{rbc}^{\text{sE}} \Theta$. To prove that this implies $\Delta \sim_{rbc}^{\text{sG}} \Theta$ suppose $\Delta \downarrow_{\geq p}^a$. Let q denote the probability $\sum\{\Delta(s) \mid s \in [\Delta], s \xrightarrow{a}\}$. Then by definition $\Delta \downarrow_{=q}^a$ and since $\Delta \sim_{rbc}^{\text{sE}} \Theta$ this gives $\Theta \downarrow_{=q}^a$. But since $p \leq q$ we therefore have the required $\Theta \downarrow_{\geq p}^a$. \square

Thus it does not really matter which of these sets of barbs are used. Consequently we often use the abbreviation \sim_{rbc} and employ whatever barbs are appropriate.

Proposition 5.6 (*Soundness of d -bisimulations*) *In an arbitrary $p\text{LTS}$, $\Delta \sim_{sd} \Theta$ implies $\Delta \sim_{rbc} \Theta$.*

Proof It is simply a question of showing that \sim_{sd} satisfies the defining properties of \sim_{rbc}^{sG} . Theorem 5.2 gives compositionality, while s -reduction-closure follows from the definition of \sim_{sd} . So let us look at barb preservation.

Suppose $\Delta \sim_{sd} \Theta$ and $\Delta \downarrow_{\geq p}^a$. Let q denote the probability $\sum\{\Delta(s) \mid s \xrightarrow{a}\}$; by definition $q \geq p$. Now consider the distribution Δ_a determined by

$$\Delta_a(s) = \begin{cases} \frac{\Delta(s)}{q}, & \text{if } s \xrightarrow{a} \\ 0, & \text{otherwise} \end{cases}$$

Then Δ can be written as $\Delta_a \oplus \Delta'$ for some distribution Δ' .

Since \sim_{sd} is decomposable we must have $\Theta = \Theta_a \oplus \Theta'$ such that $\Delta_a \sim_{sd} \Theta_a$. Now $\Delta_a \xrightarrow{a}$ from which it follows that $\Theta_a \xrightarrow{a}$ and so we can conclude $\Theta \downarrow_{\geq p}^a$. \square

The behavioural justification for this form of probabilistic bisimulation, \sim_{sd} , lies in the fact that the converse to this proposition is also true. The main technical tool underlying this proof is the following:

Lemma 5.7 *In a sufficiently expressive $p\text{LTS}$, suppose $(\Delta_1 \mid \text{succ}_1)_p \oplus (\Delta_2 \mid \text{succ}_2) \sim_{rbc} (\Theta_1 \mid \text{succ}_1)_p \oplus (\Theta_2 \mid \text{succ}_2)$ where $p > 0$ and succ_i are two fresh actions. Then $\Delta_i \sim_{rbc} \Theta_i$, $i = 1, 2$.*

Proof Let \mathcal{R} be the relation defined by

$$\{(\Delta_1, \Theta_1) \mid (\Delta_1 \mid \text{succ}_1)_p \oplus (\Delta_2 \mid \text{succ}_2) \sim_{rbc} (\Theta_1 \mid \text{succ}_1)_p \oplus (\Theta_2 \mid \text{succ}_2) \\ \text{for some } \Delta_2, \Theta_2, p > 0 \text{ and fresh } \text{succ}_i, i = 1, 2\}$$

The result follows by symmetry if we show that \mathcal{R} is compositional, s -reduction-closed and preserves the barbs in G . The first is straightforward because of the compositionality of \sim_{rbc} ; let us look at the other two. But first note that if $\Delta_1 \mathcal{R} \Theta_1$ then we know that

$$\Gamma \sim_{rbc} \Lambda \text{ where } \Gamma = (\Delta_1 \mid \text{succ}_1)_p \oplus (\Delta_2 \mid \text{succ}_2) \text{ and } \Lambda = (\Theta_1 \mid \text{succ}_1)_p \oplus (\Theta_2 \mid \text{succ}_2) \quad (6)$$

for some Δ_2, Θ_2 and fresh actions succ_i .

s-reduction-closed: Suppose $\Delta_1 \mathcal{R} \Theta_1$ and $\Delta_1 \xrightarrow{\tau} \Delta'_1$. Let Γ, Λ be determined as in (6) above and consider the testing process

$$T = \text{stay} + \overline{\text{succ}_2}.go$$

where stay and go are two new fresh actions.

Then consider the action² $\Gamma \mid T \xrightarrow{\tau} \Gamma'$ where Γ' is the distribution

$$(\Delta'_1 \mid succ_1 \mid T)_p \oplus (\Delta_2 \mid go)$$

Γ' has the barbs $\Gamma' \downarrow_{=p}^{stay}$, $\Gamma' \downarrow_{=(1-p)}^{go}$ and $\Gamma' \downarrow_{=0}^{succ_2}$.

By the compositionality of \sim_{rbc} we have $\Gamma \mid T \sim_{rbc} \Lambda \mid T$ and therefore $\Lambda \mid T$ must have a matching move, that is $\Lambda \mid T \xrightarrow{\tau} \Lambda'$ such that Λ' has these same barbs. By examining the possibilities it becomes clear that Λ' can only take the form

$$(\Theta'_1 \mid succ_1 \mid T)_p \oplus (\Theta_2 \mid go)$$

for some Θ'_1 such that $\Theta_1 \xrightarrow{\tau} \Theta'_1$. We now have to show that $\Delta'_1 \mathcal{R} \Theta'_1$.

To do so consider the new testing process $\overline{stay} + \overline{go}.succ_3$ which we denote by S ; here $succ_3$ is again assumed to be fresh. We know $(\Gamma' \mid S) \sim_{rbc} (\Lambda' \mid S)$ and so consider the move $(\Gamma' \mid S) \xrightarrow{\tau} \Gamma''$ where $\Gamma'' = (\Delta'_1 \mid succ_1)_p \oplus (\Delta_2 \mid succ_3)$. This must have a matching move from $(\Lambda' \mid S)$ whose residual is s-reduction barbed congruent to Γ'' . Again considering the barbs of Γ'' the only possibility is for the residual to be $(\Theta'_1 \mid succ_1)_p \oplus (\Theta_2 \mid succ_3)$. We have just shown the required $\Delta'_1 \mathcal{R} \Theta'_1$.

barb-preserving: Here the style of argument is similar, but the tests used are different. Suppose $\Delta_1 \mathcal{R} \Theta_1$ and $\Delta_1 \downarrow_a^{=q}$ for some action a and probability q . Let Γ and Λ be as in (6) above.

Now consider the test $T = \overline{a}.succ_2.succ$ where as usual $succ$ is fresh. Then $(\Gamma \mid T) \downarrow_{succ}^{=pq}$ and by contextuality we must also have $(\Lambda \mid T) \downarrow_{succ}^{=pq}$. In view of the fact that $p > 0$ this can only be the case if $\Theta_1 \downarrow_a^{=q}$. \square

Proposition 5.8 *In a sufficiently expressive pLTS, \sim_{rbc} is decomposable.*

Proof By symmetry, it is sufficient to prove that \sim_{rbc} is left-decomposable. Suppose $\Delta_1 \oplus \Delta_2 \sim_{rbc} \Theta$. We can assume that $p > 0$ for otherwise the required decomposition of Θ is trivial.

Here we use the test $T = stay + \tau.succ_1 + \tau.succ_2$. Consider $(\Delta_1 \oplus \Delta_2) \mid T$. This can perform a τ action to $\Gamma = (\Delta_1 \mid succ_1)_p \oplus (\Delta_2 \mid succ_2)$, and note the barbs $\Gamma \downarrow_{succ_1}^{=p}$, $\Gamma \downarrow_{succ_2}^{=(1-p)}$ and $\Gamma \downarrow_{stay}^0$.

For the usual reasons we must have a move $(\Theta \mid T) \xrightarrow{\tau} \Lambda$ where Λ also has these barbs. This forces Λ to be of the form $(\Theta_1 \mid succ_1)_p \oplus (\Theta_2 \mid succ_2)$ with $\Theta = \Theta_1 \oplus \Theta_2$. Moreover by s-reduction-closure we must have $\Gamma \sim_{rbc} \Lambda$ and an application of the previous technical lemma now gives the required $\Delta_i \sim_{rbc} \Theta_i$, $i = 1, 2$. \square

Theorem 5.9 (Full-abstraction for \sim_{sd}) *In a sufficiently expressive pLTS, $\Delta \sim_{sd} \Theta$ if and only if $\Delta \sim_{rbc} \Theta$.*

Proof Because of Proposition 5.6 it is sufficient to show that $\sim_{rbc} \subseteq \mathcal{B}_{sd}(\sim_{rbc})$. Moreover the previous result shows that \sim_{rbc} is decomposable, and therefore by symmetry all we need to show is that if $\Delta \sim_{rbc} \Theta$ and $\Delta \xrightarrow{\mu} \Delta'$ then $\Theta \xrightarrow{\mu} \Theta'$ such that $\Delta' \sim_{rbc} \Theta'$; if μ is τ then this follows directly from the fact that \sim_{rbc} is s-reduction closed. So we consider the case when μ is $a \in \text{Act}$. Let T be the testing process $fail + \overline{a}.succ$ where $fail$ and $succ$ are fresh actions. Then $\Delta \mid \overline{T} \xrightarrow{\tau} \Delta' \mid succ$. Since $\Delta \sim_{rbc} \Theta$ we know $\Delta \mid \overline{T} \sim_{rbc} \Theta \mid \overline{T}$. Since \sim_{rbc} is s-reduction-closed, there is some Γ such that $\Theta \mid \overline{T} \xrightarrow{\tau} \Gamma$ and $\Delta' \mid succ \sim_{rbc} \Gamma$. Now \sim_{rbc} preserves barbs from sE and since $(\Delta' \mid succ) \downarrow_{=1}^{succ}$ and $(\Delta' \mid succ) \not\downarrow_{=0}^{fail}$, the same barbs should also be true for Γ . By the construction of the test T it follows that Γ must have the form $\Theta' \mid succ$ for some Θ' with $\Theta \xrightarrow{a} \Theta'$.

We have therefore established that for some Θ' such that $\Theta \xrightarrow{a} \Theta'$, $\Delta' \mid succ \sim_{rbc} \Theta' \mid succ$, where $succ$ is fresh. From this it follows easily that $\Delta' \sim_{rbc} \Theta'$, as required. To justify this last step it is sufficient to show that the relation

$$\mathcal{R} = \{(\Delta, \Theta) \mid \Delta \mid succ \sim_{rbc} \Theta \mid succ \text{ for some fresh actions } succ\}$$

satisfies the defining properties of \sim_{rbc} ; all three are perfectly straightforward. \square

² Here we are abbreviating the point distribution \overline{T} by T .

$$\begin{array}{ll}
(A1) & x + y = y + x \\
(A2) & \mu.x + \mu.x = \mu.x \\
(A3) & (x + y) + z = x + (y + z) \\
(A4) & x + \mathbf{0} = x \\
(P1) & x_p \oplus y = y_p \oplus x \\
(P2) & x_p \oplus x = x \\
(P3) & (x_p \oplus y)_q \oplus z = x_{p \cdot q} \oplus (y_{\frac{(1-p) \cdot q}{1-p \cdot q}} \oplus z) \\
(CC) & \mu.x + \mu.y = \mu.x + \mu.y + \mu.(x_p \oplus y) \\
(PD) & x + (y_p \oplus z) = (x + y)_p \oplus (x + z)
\end{array}$$

Fig. 4. The equational theory

6. Axiomatisation

In this section we discuss briefly the equational theory of \sim_{sd} . To keep things straightforward we confine our attention to the simple language bpCCS.

The axioms we use are given in Fig. 4. The first set, (A1)–(A4), are known to be sound and complete for strong bisimulation equivalence for finite CCS. But note we have replaced the more standard idempotency equation

$$x + x = x$$

with the weaker (A2). This is because the idempotency law is not sound for bpCCS.

Example 6.1 Let P denote the term $a_{\frac{1}{2}} \oplus b$. Then $\llbracket P + P \rrbracket$ is the distribution

$$\frac{1}{4} \cdot \overline{a+a} + \frac{1}{4} \cdot \overline{a+b} + \frac{1}{4} \cdot \overline{b+a} + \frac{1}{4} \cdot \overline{b+b}$$

So the distribution $\llbracket P + P \rrbracket$ has the barb $\downarrow_{\geq \frac{3}{4}}^a$ and obviously $\llbracket P \rrbracket$ does not.

Consequently $\llbracket P \rrbracket \not\sim_{rbc} \llbracket P + P \rrbracket$ and therefore by Theorem 5.9 $\llbracket P \rrbracket \not\sim_{sd} \llbracket P + P \rrbracket$. ■

The axioms (P1)–(P3) are the obvious axioms required for re-arranging distributions syntactically. However there are two new axioms. The first, (CC) which is parametric in the probability p , is a reflexion of the fact that in the state-based definition of strong bisimulation, Definition 3.12, combined moves are allowed, while (PD) results from the fact effectively $+$ is an operator on states rather than distributions.

For terms in bpCCS we write $P =_{SA} Q$ if P and Q can be proved equal using the axioms in Fig. 4.

Proposition 6.2 (Soundness) *In bpCCS, $P =_{SA} Q$ implies $\llbracket P \rrbracket \sim_{sd} \llbracket Q \rrbracket$.*

Proof It is straightforward to show

- each of the axioms are sound
- $lift(\sim_{ss})$ is preserved by the operators $\mu \cdot -$, $- + -$ and $-_p \oplus -$ acting on distributions.

The result then follows. □

The converse to this result, completeness, rests on the following, a generalisation of the well-known property of the axioms for finite CCS, Lemma 16 page 137 of [Mil89]:

Lemma 6.3 (Derivative lemma) *If S is a state term then $\llbracket S \rrbracket \xrightarrow{\mu} \Delta$ implies $S =_{SA} S + \mu \cdot P$, where $\llbracket P \rrbracket = \Delta$.*

Proof There are two steps.

- (a) Suppose $S \xrightarrow{\mu} \Delta$. Then by induction on the derivation of this judgement using the rules in Fig. 1 one can show that $S =_{SA} S + \mu \cdot P$ for some term P such that $P = \llbracket \Delta \rrbracket$. This proof uses all of the axioms (A1)–(A4).
- (b) Using (a) we can then show, using the axiom (CC), that $S \xrightarrow{\mu}_{cc} \Delta$ implies $S =_{SA} S + \mu \cdot P$ for some term P satisfying $P = \llbracket \Delta \rrbracket$.

The result now follows since $\llbracket S \rrbracket$ is \bar{S} and therefore by Lemma 3.15 $\llbracket S \rrbracket \xrightarrow{\mu} \Delta$ means $S \xrightarrow{\mu}_{cc} \Delta$. □

A derivative lemma can not be derived for arbitrary terms; indeed it would be unsound, as the following example shows.

Example 6.4 Let P denote the term $a.b \frac{1}{2} \oplus a.c$. Then $\llbracket P \rrbracket \xrightarrow{a} \bar{b} \frac{1}{2} \oplus \bar{c}$ but $\llbracket P \rrbracket \not\sim_{rc} \llbracket P + a.(b \frac{1}{2} \oplus c) \rrbracket$, and therefore $\llbracket P \rrbracket \not\sim_{sd} \llbracket P + a.(b \frac{1}{2} \oplus c) \rrbracket$.

To see this let \bar{Q} denote $P + a.(b \frac{1}{2} \oplus c)$ and consider the test $(- \mid \bar{a})$. Then $\llbracket \bar{Q} \mid \bar{a} \rrbracket \xrightarrow{\tau} \llbracket b \frac{1}{2} \oplus (b \frac{1}{2} \oplus c) \rrbracket$ and this residual has the barb $\downarrow_{\frac{3}{4}}^b$.

However there is a unique Δ such that $\llbracket P \mid \bar{a} \rrbracket \xrightarrow{\tau} \Delta$, which does not have this barb. \blacksquare

Theorem 6.5 In bpCCS, $P =_{SA} Q$ if and only if $\llbracket P \rrbracket \sim_{sd} \llbracket Q \rrbracket$.

Proof We already have soundness, Proposition 6.2, and therefore we only have to show completeness, namely $\llbracket P \rrbracket \sim_{sd} \llbracket Q \rrbracket$ implies $P =_{SA} Q$.

Let us use $\sum\{p_i \cdot P_i \mid 1 \leq i \leq n\}$, where it is assumed that each $p_i > 0$ and $\sum_{1 \leq i \leq n} p_i = 1$, to denote the obvious generalisation of the probabilistic choice $P_1 \oplus P_2$ such that $\llbracket \sum\{p_i \cdot P_i \mid 1 \leq i \leq n\} \rrbracket$ is the distribution

$$p_1 \cdot \llbracket P_1 \rrbracket + \dots + p_n \cdot \llbracket P_n \rrbracket$$

If n is 1, in which case the probability p_i must also be 1, this term denotes P_1 ; otherwise it represents some term in bpCCS constructed from the P_i using $-_{r_i} \oplus -$ for appropriate probabilities r_i .

A term $\sum\{p_i \cdot S_i \mid 1 \leq i \leq n\}$ in bpCCS is said to be in *standard form* if

- (i) each S_i is a state term
- (ii) if $\mu \cdot R$ summand of any S_i then R is in standard form.

A systematic use of the axiom (PD) will convert an arbitrary term into *standard form*. Therefore we only have to show Completeness for *standard forms*, which we do by induction on the combined size of the terms involved. There are three cases.

- (i) Suppose first that P and Q are state terms. Here the proof proceeds as in the completeness proof for CCS, Proposition 15 in Chapter 7 of [Mil89]. By symmetry we only have to show $P =_{SA} P + Q$. If Q is $\mathbf{0}$ this follows by the axiom (A4). Otherwise we show $P =_{SA} P + \mu \cdot R$ for every summand $\mu \cdot R$ of Q . Now $\llbracket P \rrbracket \sim_{sd} \llbracket Q \rrbracket$ implies $P \sim_{ss} Q$, and because $Q \xrightarrow{\mu} \llbracket R \rrbracket$, we know $P \xrightarrow{\mu}_{cc} \Delta$ such that $\Delta \text{ lift}(\sim_{ss}) \llbracket R \rrbracket$. The required $P =_{SA} P + \mu \cdot R$ now follows by the Derivative lemma, Lemma 6.3 and induction, since $\text{lift}(\sim_{ss})$ coincides with \sim_{sd} .
- (ii) Suppose only one of terms, say P is a state term; so Q has the form $\sum\{p_i \cdot S_i \mid 1 < i \leq n\}$. Here since \sim_{sd} is decomposable it follows that $\llbracket P \rrbracket \sim_{sd} \llbracket S_i \rrbracket$ for each i . By part (i) we know $P =_{SA} S_i$ and therefore $P =_{SA} Q$ follows by using the axiom (P2).
- (iii) Finally suppose P and Q are both non-trivial probabilistic sums. In particular P must have the form $S_p \oplus P'$ for some state term S . Again using the decomposability of \sim_{sd} we know that $Q = Q_1 \oplus Q_2$ such that $S \sim_{sd} Q_1$ and $P' \sim_{sd} Q_2$. Induction gives $P' =_{SA} Q_2$ and part (ii) that $S =_{SA} Q_1$, from which $P =_{SA} Q$ follows. \square

7. Conclusions

In this paper we have re-examined (strong) probabilistic bisimulation equivalence for processes which exhibit both non-deterministic and probabilistic behaviour, and have demonstrated that it enjoys all of the properties commonly associated with bisimulation equivalence for purely non-deterministic systems. The intensional behavioural of such processes are captured via pLTSs, essentially the simple Probabilistic Automata of [SL95]. However instead of focusing on the states of these pLTSs, we have switched attention to distributions of states, which more properly correspond to processes. This enabled us to apply a standard notion of contextual equivalence, [HY95, SW01, Hen07], to obtain a touchstone behavioural equivalence for such processes, Definition 5.4. Then we have shown that, subject to expressivity constraints on the underlying pLTS, a version of (strong)

probabilistic bisimulations over processes, Definition 3.16, provide a sound and complete proof methodology for the contextual equivalence, Theorem 5.9. Moreover we have shown, Theorem 3.17, that these bisimulations over processes (i.e. distributions) are determined by the standard notion of strong probabilistic bisimulation from the literature, [SL95]. We have also shown that a modest extension of the modal logic HML is sufficient to distinguish inequivalent processes, provided the underlying pLTS is finitary, Theorem 4.3. Finally we have shown how the complete axiomatisation of strong bisimulation equivalence over finite CCS terms can be extended to a probabilistic version of CCS, Theorem 6.5.

The paper builds on the extensive literature on probabilistic bisimulations, which we have tried to reference in situ. However there are also novel results. For example the version of probabilistic bisimulations which apply directly to distributions is new; it is a simplification of version of (weak) bisimulations given in Definition 2.12 of [DH11]. The characterisation of the contextual equivalence in terms of these bisimulations is also new, Theorem 5.9.

In [PS07] there is an extension to HML different from our pHML which also characterise strong probabilistic bisimulation equivalence for finitary pLTSs; infinitary variations to their logic are used in [HPS⁺11] for characterising the behaviour in non-finitary pLTSs. That extension to HML uses two non-standard logical constructs; their satisfaction by a distribution depends on the behaviour of the point distributions of every state in its support. This is discussed in more detail in [DvG10], page 15. We feel that the addition of our single extra logical construct $\varphi_1 \oplus_p \varphi_2$ is more natural. However our use of modal formulae is simply to distinguish between inequivalence processes, which for finite branching systems can be achieved using a finitary logic. The more complicated question of characterising the entire behaviour of (regular) processes via a single formula is studied in [DvG10]; in general such logics require infinitary constructs.

Axiomatisations of various forms of probabilistic bisimulation equivalence are given in [BS01]. This includes strong bisimulation equivalence and their set of axioms are naturally very similar to ours. However formally they are incomparable as the language they use, Probabilistic Process Algebra, is different from our bpCCS; specifically they have a two-level syntax, for states and distributions, and this distinction plays a role in their axioms. A similar range of axiomatisations for various probabilistic extensions of ACP can be found in [And02].

Finally, our results only pertain to **strong** probabilistic bisimulations, where processes do not have any internal behaviour. It would be very interesting to re-do this exercise for the weak case, in which τ actions are unobservable.

Acknowledgments

The author would like to thank the anonymous referees for their useful comments.

References

- [And02] Andova S (2002) Probabilistic process algebra. PhD thesis, Eindhoven University of Technology
- [BEMC00] Baier C, Engelen B, Majster-Cederbaum ME (2000) Deciding bisimilarity and similarity for probabilistic processes. *J Comput Syst Sci* 60(1):187–231
- [BS01] Bandini E, Segala R (2001) Axiomatizations for probabilistic bisimulation. In: Orejas F, Spirakis PG, van Leeuwen J (eds) ICALP, volume 2076 of Lecture Notes in Computer Science, Springer, Berlin, pp 370–381
- [Cle90] Cleaveland R (1990) On automatically explaining bisimulation inequivalence. In: Clarke EM, Kurshan RP (eds) CAV, volume 531 of Lecture Notes in Computer Science. Springer, New York, pp 364–372
- [DD11] Deng Y, Du W (2011) Logical, metric, and algorithmic characterisations of probabilistic bisimulation. Technical Report CMU-CS-11-110, Carnegie Mellon University
- [DH11] Deng Y, Hennessy M (2011) On the semantics of Markov automata. In: Proceedings of the 38th international colloquium on automata, languages and programming (ICALP'11). Volume 6756 of Lecture Notes in Computer Science. Springer, New York, pp 307–318 (full version available as technical report TCS-2011-06, Trinity College).
- [DvG10] Deng Y, van Glabbeek RJ (2010) Characterising probabilistic processes logically (extended abstract). In: Fermüller CG, Voronkov A (eds) LPAR (Yogyakarta). Volume 6397 of Lecture Notes in Computer Science. Springer, New York, pp 278–293
- [DvGH⁺07] Deng Y, van Glabbeek R, Hennessy M, Morgan C, Zhang C (2007) Remarks on testing probabilistic processes. *Electron Notes Theor Comput Sci* 72:359–397
- [DvGHM09] Deng Y, van Glabbeek R, Hennessy M, Morgan C (2009) Testing finitary probabilistic processes (extended abstract). In: Proceedings of the 20th international conference on concurrency theory. Volume 5710 of Lecture Notes in Computer Science. Springer, New York, pp 274–288 (full version available at <http://basics.sjtu.edu.cn/~yuxin/publications/finitary.pdf>)
- [FG05] Fournet C, Gonthier G (2005) A hierarchy of equivalences for asynchronous calculi. *J Logic Algebr Programm* 63(1):131–173
- [Han91] Hansson HA (1991) Time and probability in formal design of distributed systems. SICS Dissertation Series. Uppsala University
- [Hen07] Hennessy M (2007) A distributed Pi-calculus. Cambridge University Press, New York

- [HPS⁺11] Hermanns H, Parma A, Segala R, Wachter B, Zhang L (2011) Probabilistic logical characterization. *Inf Comput* 209:154–172
- [HY95] Honda K, Yoshida N (1995) On reduction-based process semantics. *Theor Comput Sci* 152:437–486
- [LS89] Larsen KG, Skou A (1989) Bisimulation through probabilistic testing (preliminary report). In: Proceedings of the 16th ACM SIGPLAN-SIGACT symposium on Principles of programming languages, New York pp 344–352
- [Mil89] Milner R (1989) Communication and concurrency. Prentice-Hall, Englewood Cliffs
- [MM05] McIver A, Morgan C (2005) Abstraction, refinement and proof for probabilistic systems. Springer, Berlin
- [MMSS95] Morgan C, McIver A, Seidel K, Sanders JW (1995) Refinement-oriented probability for CSP. *Formal Aspects Comput* 8:8–16
- [MS92] Milner R, Sangiorgi D (1992) Barbed bisimulation. In: Kuich W (ed) ICALP. Volume 623 of Lecture Notes in Computer Science. Springer, Berlin, pp 685–695
- [PLS00] Philippou A, Lee I, Sokolsky O (2000) Weak bisimulation for probabilistic systems. In: Proceedings of the 11th international conference on concurrency theory (CONCUR). Volume 1877 of Lecture Notes in Computer Science. Springer, New York, pp 334–349
- [PLS03] Philippou A, Lee I, Sokolsky O (2003) Weak bisimulation for probabilistic systems. Technical Report FLAGS-TR-6, IST Programme. Foundational aspects of global computing. <http://ru1.cti.gr/flags/online-reports.jsp>
- [PS07] Parma A, Segala R (2007) Logical characterizations of bisimulations for discrete probabilistic systems. In: Seidl H (ed) FoSSaCS. Volume 4423 of Lecture Notes in Computer Science. Springer, Berlin pp 287–301
- [Seg95] Segala R (1995) Modeling and verification of randomized distributed real-time systems. Technical Report MIT/LCS/TR-676, PhD thesis, MIT, Dept. of EECS
- [SKS07] Sangiorgi D, Kobayashi N, Sumii E (2007) Environmental bisimulations for higher-order languages. In: Proceedings of the 22nd IEEE symposium on logic in computer science. IEEE Computer Society, pp 293–302
- [SL94] Segala R, Lynch NA (1994) Probabilistic simulations for probabilistic processes. In: Jonsson B, Parrow J (eds) CONCUR. Volume 836 of Lecture Notes in Computer Science. Springer, Berlin, pp 481–496
- [SL95] Segala R, Lynch NA (1995) Probabilistic simulations for probabilistic processes. *Nordic J Comput* 2(2):250–273
- [Sto02] Stoelinga M (2002) Alea jacta est: verification of probabilistic, real-time and parametric systems. PhD thesis, University of Nijmegen, The Netherlands. <http://www.soe.ucsc.edu/~marielle>
- [SW01] Sangiorgi D, Walker D (2001) The π -calculus: a theory of mobile processes. Cambridge University Press, Cambridge
- [Tof94] Tofts CMN (1994) Processes with probabilities, priority and time. *Formal Aspects Comput* 6(5):536–564

Received 21 December 2011

Accepted in revised form 24 April 2012 by Peter Höfner, Robert van Glabbeek and Ian Hayes

Published online 2 July 2012