On the Semantics of Markov Automata

Yuxin Deng[1,3,4] * Matthew Hennessy[2†]
[1] Shanghai Jiao Tong University
[2] Trinity College Dublin
[3] Carnegie Mellon University
[4] Chinese Academy of Sciences

October 16, 2011

Abstract

Markov automata describe systems in terms of events which may be nondeterministic, may occur probabilistically, or may be subject to time delays. We define a novel notion of weak bisimulation for such systems and prove that this provides both a sound and complete proof methodology for a natural extensional behavioural equivalence between such systems, a generalisation of *reduction barbed congruence*, the well-known touchstone equivalence for a large variety of process description languages.

Keywords: Markov automata; Bisimulation; Soundness; Completeness

1 Introduction

Markov Automata (MA), as defined in [9], describe system behaviour in terms of nondeterministic, probabilistic and timed events. The first two kinds of events are well-known from Probabilistic Automata (PA) [25, 26] and Probabilistic Labelled Transition Systems (pLTSs) [6], while the third are taken to be random delays, governed by

*Partially supported by the Natural Science Foundation of China (61173033, 61033002), the Qatar National Research Fund under grant NPRP 09-1107-1-168, and the Opening Fund of Top Key Discipline of Computer Software and Theory in Zhejiang Provincial Colleges at Zhejiang Normal University.

†Supported financially by SFI project no. SFI 06 IN.1 1898.

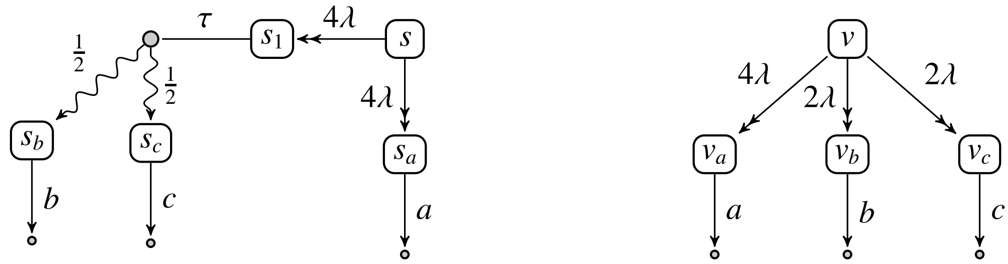

Figure 1: Timed transitions and distributions

negative exponential distributions parametrised by some delay $\lambda \in \mathbb{R}^+$. As explained in [11] these timed events can be given a straightforward operational semantics in terms only of their parametric delays.

For example, consider the MAs in Figure 1, taken from [9]. In such diagrams we use double headed arrows between states to denote time delays. From the initial state of the first automaton, s, there is a race between two possible timed events, each governed by the same rate, 4λ, for some arbitrary $\lambda \in \mathbb{R}^+$. If the right hand event wins, the state of the automaton changes to s_a, from which some external action a can happen. If the other timed event wins, the change of state is to s_1, from which an internal unobservable action, denoted by τ, can occur. Moreover, the effect of this internal action is probabilistic; fifty percent of the time the state change will be to s_b, where the external action b can occur, while with the same probability the change will be to s_c, where c can occur. Formally, this probabilistic behaviour is represented as an action from a state, such as s_1, to a distribution over states, represented diagrammatically as a darkened circle connected to states in the support of the distribution, labelled with their probabilities.

On the other hand, the second automaton is much more straightforward. From its initial state there is a race between three timed events, two running at the same rate and one at double the rate. Then one of the (external) actions a, b, c occurs depending on which event wins the race.

Providing a satisfactory behavioural model of MAs is necessarily a complicated undertaking. But as pointed out in [9], because of the nature of their underlying distributions, the timed events can be satisfactorily explained in terms of simple probabilistic distributions determined by their rates. They propose a translation of MAs into PAs, which we will explain in Section 2. Since behavioural theories have already been

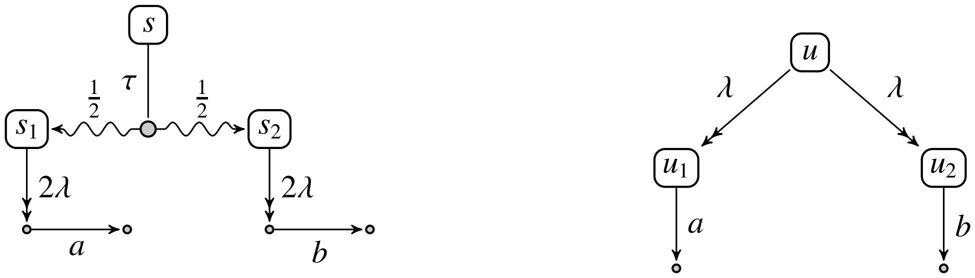

Figure 2: Timed transitions and distributions, again

developed for PAs [27, 8, 20, 4], we therefore automatically obtain such theories for MAs, via their induced PAs.

However, if one uses a standard behavioural theory for PAs, such as weak bisimulation equivalence as defined in [16, 27, 20, 8] then the two MAs in Figure 1 are distinguished. Instead the authors of [9] propose a new version of bisimulation equivalence between PAs, which enjoys desired standard properties such as compositionality, and which identifies these two MAs. But as the authors point out their equivalence still distinguishes between the MAs in Figure 2. The question naturally arises: which MAs should be distinguished behaviourally, and which be deemed equivalent. This is the topic of the current paper.

We approach the question indirectly, by giving criteria for reasonable behavioural equivalences between MAs; this induces a *touchstone* extensional equivalence between systems, namely the largest equivalence, $\approx_{\text{behav}}$, which satisfies these criteria. Thus two MAs should only be distinguished on the basis of the chosen criteria.

Having an independent notion of which systems should, and which should not, be distinguished, one can then justify a particular notion of bisimulation by showing that it captures precisely the touchstone equivalence, $\approx_{\text{behav}}$. In other words, a particular definition of bisimulation is appropriate because $\approx_{bis}$, the associated bisimulation equivalence,

(i) is *sound* with respect to the touchstone equivalence, that is $s_1 \approx_{bis} s_2$ implies $s_1 \approx_{\text{behav}} s_2$

(ii) provides a *complete* proof methodology for the touchstone equivalence, that is $s_1 \approx_{\text{behav}} s_2$ implies $s_1 \approx_{bis} s_2$.

This approach originated in [14] but has now been widely used for different process description languages; for example see [15, 23] for its application to higher-order process languages, [21] for mobile ambients and [10] for asynchronous languages. Moreover in each case the distinguishing criteria are more or less the same. The touchstone equivalence should

(i) be *compositional*; that is preserved by some natural operators for constructing systems

(ii) *preserve barbs*; barbs are simple experiments which observers may perform on systems [22]

(iii) be *reduction-closed*; this is a natural condition on the reduction semantics of systems which ensures that nondeterministic choices are in some sense preserved.

We adapt this approach to MAs. Using natural versions of these criteria for MAs we obtain an appropriate touchstone equivalence, which we call *reduction barbed congruence*, $\approx_{rbc}$. We then develop a new theory of bisimulations which is both sound and complete for $\approx_{rbc}$.

The remainder of the paper is organised as follows. In the next section we give our definition of Markov automata, a slight generalisation of that in [9]; in addition to the timed events parametrised on specific delays, we have special timed events which have indefinite, or imprecise delay times associated with them. In order to model the delay operators probabilistically, we then show how to translate a MA into a PA, as suggested in [9]. For this purpose we use a slight variation, called MLTSs, in which there are distinguished actions labelled by weights. We then develop our new definition of bisimulation equivalence for MLTSs, thereby inducing bisimulation equivalence between MAs; this construction is illustrated via examples. In Section 3 we show how MAs can be composed, using a parallel operator based on CCS [18]. In fact this is extended to an interpretation of a Markovian extension of CCS, mCCS, as a MA. We then show that bisimulation equivalence is preserved by this form of composition.

Section 4 contains the main theoretical results of the paper. We give a formal definition of the touchstone equivalence $\approx_{rbc}$, and detail the proof that this is captured precisely by our new notion of bisimulation. The paper ends with a brief comparison with related work in Section 5; in particular we resume our discussion of [9], which originally stimulated our interest in bisimulations for Markovian processes.

2 Markov automata

We have divided this material into three sections. In the first we describe the two kinds of automata of interest, and the relation between them. We then have a section devoted to properties and extensions to the action relations which underlie these automata. In the third section we describe our notion of bisimulation equivalence.

2.1 Automata

We begin with some notation. A (discrete) probability subdistribution over a set S is a function $\Delta : S \rightarrow [0,1]$ with $\sum_{s\in S} \Delta(s) \leq 1$; the support of such an Δ is the set $\lceil\Delta\rceil = \{\, s \in S \mid \Delta(s) > 0 \,\}$. The mass of a distribution Δ, denoted by $|\Delta|$, is defined to be the sum $\sum_{s\in\lceil\Delta\rceil} \Delta(s)$. A subdistribution is a (total, or full) distribution if its mass is 1. The point distribution $\overline{s}$ assigns probability 1 to s and 0 to all other elements of S, so that $\lceil\overline{s}\rceil = s$. We use $\mathcal{D}_{sub}(S)$ to denote the set of subdistributions over S, and $\mathcal{D}(S)$ its subset of full distributions.

We write $\mathbb{R}^+$ for the set of all positive real numbers. Let $\{\Delta_k \mid k \in K\}$ be a set of subdistributions, possibly infinite. Then $\sum_{k\in K} \Delta_k$ is the partial real-valued function in $S \rightarrow \mathbb{R}^+$ defined by $(\sum_{k\in K} \Delta_k)(s) := \sum_{k\in K} \Delta_k(s)$. This is a partial operation on subdistributions because for some state s the sum of $\Delta_k(s)$ might not have an upper bound. If the index set is finite, say $\{1..n\}$, we often write $\Delta_1 + \ldots + \Delta_n$. For p a real number from $[0,1]$ we use $p \cdot \Delta$ to denote the subdistribution given by $(p \cdot \Delta)(s) := p \cdot \Delta(s)$. Note that if $\sum_{k\in K} p_k = 1$ for some collection of $p_k \geq 0$, and the Δ_k are distributions, then so is $\sum_{k\in K} p_k \cdot \Delta_k$. We sometimes abbreviate $p \cdot \Delta_1 + (1-p) \cdot \Delta_2$ into $\Delta_1\, {}_p\!\oplus\, \Delta_2$.

Definition 2.1. A *Markov automaton* (MA), is a quadruple $\langle S, \mathsf{Act}_\tau, \rightarrow, \mapsto\rangle$, where

(i) S is a set of states
(ii) Act_τ is a set of transition labels, with distinguished element τ
(iii) the relation $\rightarrow$ is a subset of $S \times \mathsf{Act}_\tau \times \mathcal{D}(S)$
(iv) the relation $\mapsto$ is a subset of $S \times (\mathbb{R}^+ \cup \{\delta\}) \times S$

satisfying

(a) $s \overset{\mathbf{d}}{\mapsto} t$ implies $s \not\xrightarrow{\tau}$, where $\mathbf{d} = \delta$ or $\mathbf{d} = \lambda \in \mathbb{R}^+$
(b) $s \overset{\delta}{\mapsto} t_1$ and $s \overset{\delta}{\mapsto} t_2$ implies $t_1 = t_2$. □

In (a) and (b) we use the standard notation for actions, for example, representing $(s,\lambda,t) \in\ \mapsto$ as $s \overset{\lambda}{\mapsto} t$; this notation is used throughout the paper. We will also use notation such as Act_τ, rather than $\mathsf{Act} \cup \{\tau\}$, to emphasise that τ is a special element not in Act. We have seen in the introduction how we represent these automata graphically. However, to make these diagrams simpler we will sometimes represent a point distribution simply as a state rather than, more correctly, using a darkened circle.

Our definition of a MA is a mild generalisation of that in [9]; for example *maximal progress*, assumption (a), is built in to the definition. But the major extension is the introduction of the indefinite delay actions denoted by the special action δ, $s \overset{\delta}{\mapsto} t$; this can be viewed as a timed action whose underlying rate is unknown. Such indefinite actions, often called *passive* when they are external, are widely used in the literature [3, 13], although their precise properties vary between publications; see [11], page 66 for a discussion. The role of indefinite delay actions will become clear in Section 3.1 when we define the parallel composition of two MAs.

Following [9], we study MAs indirectly, by considering derived structures, which we call *MLTSs*.

Definition 2.2. A *Markov labelled transition system* (MLTS) is a triple $\langle S, \mathsf{Act}_\tau, \rightarrow\rangle$, where

(i) S is a set of states

(ii) Act_τ is a set of transition labels, with distinguished element τ

(iii) the relation $\rightarrow$ is a subset of $S \times (\mathsf{Act}_{\tau,\delta} \cup \mathbb{R}^+) \times \mathcal{D}(S)$

satisfying

(a) $s \xrightarrow{\mathbf{d}} \Delta$ implies $s \not\xrightarrow{\tau}$, where $\mathbf{d} = \delta$ or $\mathbf{d} = \lambda \in \mathbb{R}^+$
(b) $s \xrightarrow{\delta} \Delta_1$ and $s \xrightarrow{\delta} \Delta_2$ implies $\Delta_1 = \Delta_2$
(c) $s \xrightarrow{\lambda_1} \Delta_1$ and $s \xrightarrow{\lambda_2} \Delta_2$ implies $\lambda_1 = \lambda_2$ and $\Delta_1 = \Delta_2$.

The first two constraints are inherited directly from MAs while (c) means that actions labelled by λs, in this context refered to as *weights*, are deterministic. □

A (non-probabilistic) labelled transition system (LTS) may be viewed as a degenerate MLTS, one in which only point distributions are used, and the special actions labelled by δ and $\lambda \in \mathbb{R}^+$ are vacuous. An MLTS is *finitary* if the state set S is finite and for each $s \in S$ the set $\{(\mu, \Delta) \mid s \xrightarrow{\mu} \Delta\}$ is finite; in this paper we are primarily concerned with finitary MLTSs.

Admittedly, MAs and MLTSs are very similar; the difference lies in the intent. In the former, timed events are represented explicitly as occurrences of actions $s \overset{\lambda_i}{\mapsto} t_i$, with race conditions represented by multiple timed actions with the same source s. In the latter, MLTSs, these races will be represented implicitly as actions $s \xrightarrow{\lambda} \Delta$ where Δ is a probability distribution representing the probability of the various target states t_i by the race; the label λ, the weight, will be required for compositional reasoning. Thus in MLTSs the passage of time is modelled probabilistically. We are primarily interested in MAs; however it is difficult to apply certain concepts to them, such as bisimulations. MLTSs are more amenable. We therefore study the semantic theory of MAs in terms of their derived MLTSs.

The intuitive ideas outlined above underlie the formal interpretation of MAs in MLTSs. The essential ingredient in the interpretation is the function on the states of a MA, defined by $\mathsf{Rate}(s) = \sum\{\, \lambda_i \mid s \overset{\lambda_i}{\mapsto} t_i \,\}$.

Given a MA M as in Definition 2.1 the MLTS $\mathsf{mlts}(M)$ is given by $\langle S, \mathsf{Act}_\tau, \rightarrow\rangle$ where:

(a) for $\mu \in \mathsf{Act}_\tau$ the actions $s \xrightarrow{\mu} \Delta$ are inherited from M

(b) $s \xrightarrow{\delta} \overline{t}$ whenever $s \overset{\delta}{\mapsto} t$ in M

(c) for $\lambda \in \mathbb{R}^+$, $s \xrightarrow{\lambda} \Delta$ if $\mathsf{Rate}(s) = \lambda > 0$ and $\Delta = \sum\{\, p_i \cdot \overline{t_i} \mid s \overset{\lambda_i}{\mapsto} t_i \,\}$ where $p_i = \frac{\lambda_i}{\mathsf{Rate}(s)}$

Example 2.3. The derived MLTSs of the two MAs in Figure 1 are given in Figure 3. Note that the time dependent race between the evolution of s to s_a or s_1 in Figure 1 is represented in Figure 3 by a single arrow labelled by the total rate of s to a distribution representing the chances of s_1 and s_2 winning the race. Similarly, in the second MA the race from v to v_a, v_b, v_c is now represented by a single weighted arrow to a similar distribution. The weights on these arrows will be used for compositional reasoning.
$\square$

2.2 Actions over distributions

In a MLTS actions are only performed by states, in that actions are given by relations from states to distributions. But in general we allow distributions over states to perform an action. For this purpose, we *lift* these relations so that they also apply to subdistributions [6].

Definition 2.4. [Lifting] Let $\mathcal{R} \subseteq S \times \mathcal{D}_{sub}(S)$ be a relation from states to subdistributions in a MLTS. Then $\mathit{lift}(\mathcal{R}) \subseteq \mathcal{D}_{sub}(S) \times \mathcal{D}_{sub}(S)$ is the smallest relation that satisfies

(i) $s \mathrel{\mathcal{R}} \Theta$ implies $\overline{s} \mathrel{\mathit{lift}(\mathcal{R})} \Theta$, and

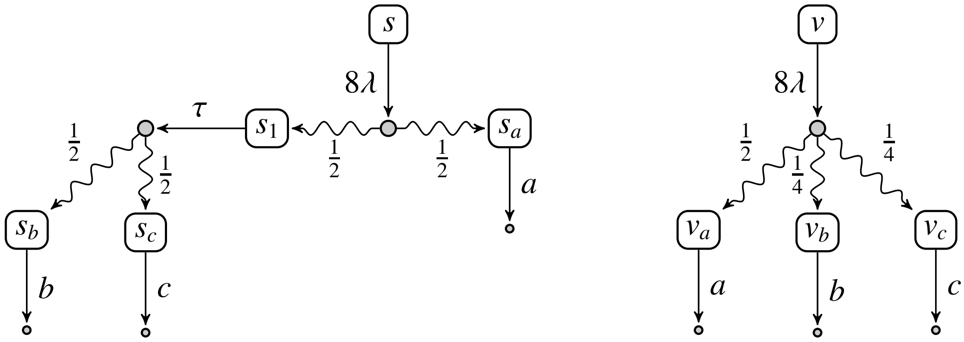

Figure 3: Derived MLTSs of MAs in Figure 1

(ii) (Linearity) Δ_i *lift*($\mathcal{R}$) Θ_i for $i \in I$ implies $(\sum_{i\in I} p_i \cdot \Delta_i)$ *lift*($\mathcal{R}$) $(\sum_{i\in I} p_i \cdot \Theta_i)$ for any $p_i \in [0, 1]$ with $\sum_{i\in I} p_i = 1$, where I is a finite index set. □

Note that the definition of Linearity uses only a finite index set I; this is sufficient for our purposes as our primary focus are on finite state systems. Indeed in the remainder of the paper all index sets can be taken to be finite, unless indicated otherwise.

There are numerous ways of formulating this concept. The following is particularly useful.

Lemma 2.5. Δ *lift*($\mathcal{R}$) Θ if and only if there is a finite index set I such that

(i) $\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}$

(ii) $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$

(iii) $s_i \; \mathcal{R} \; \Theta_i$ for each $i \in I$.

Proof. ($\Leftarrow$) Suppose there is an index set I such that (i) $\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}$, (ii) $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$, and (iii) $s_i \; \mathcal{R} \; \Theta_i$ for each $i \in I$. By (iii) and the first rule in Definition 2.4, we have $\overline{s_i}$ *lift*($\mathcal{R}$) Θ_i for each $i \in I$. By the second rule in Definition 2.4 we obtain that $(\sum_{i\in I} p_i \cdot \overline{s_i})$ *lift*($\mathcal{R}$) $(\sum_{i\in I} p_i \cdot \Theta_i)$, that is Δ *lift*($\mathcal{R}$) Θ.

($\Rightarrow$) We proceed by rule induction.

- If Δ *lift*($\mathcal{R}$) Θ because of $\Delta = \overline{s}$ and $s \; \mathcal{R} \; \Theta$, then we can simply take I to be the singleton set $\{i\}$ with $p_i = 1$ and $\Theta_i = \Theta$.

- If Δ *lift*($\mathcal{R}$) Θ because of the conditions $\Delta = \sum_{i\in I} p_i \cdot \Delta_i$, $\Theta_i = \sum_{i\in I} p_i \cdot \Theta_i$ for some index set I, and Δ_i *lift*($\mathcal{R}$) Θ_i for each $i \in I$, then by induction hypothesis there are index sets J_i such that $\Delta_i = \sum_{j\in J_i} p_{ij} \cdot \overline{s_{ij}}$, $\Theta_i = \sum_{j\in J_i} p_{ij} \cdot \Theta_{ij}$, and $s_{ij}\ \mathcal{R}\ \Theta_{ij}$ for each $i \in I$ and $j \in J_i$. It follows that $\Delta = \sum_{i\in I}\sum_{j\in J_i} p_i p_{ij} \cdot \overline{s_{ij}}$, $\Theta = \sum_{i\in I}\sum_{j\in J_i} p_i p_{ij} \cdot \Theta_{ij}$, and $s_{ij}\ \mathcal{R}\ \Theta_{ij}$ for each $i \in I$ and $j \in J_i$. Therefore, it suffices to take $\{ij \mid i \in I, j \in J_i\}$ to be the index set and $\{p_i p_{ij} \mid i \in I, j \in J_i\}$ be the collection of probabilities.

□

We apply this operation to the relations $\xrightarrow{\mu}$ in the MLTS for $\mu \in \mathsf{Act}_{\tau,\delta} \cup \mathbb{R}^+$, where we also write $\xrightarrow{\mu}$ for *lift*($\xrightarrow{\mu}$). Thus as source of a relation $\xrightarrow{\mu}$ we now also allow distributions, and even subdistributions. But note that $\overline{s} \xrightarrow{\mu} \Delta$ is more general than $s \xrightarrow{\mu} \Delta$. In papers such as [27, 5] the former is refered to as a *combined transition* because if $\overline{s} \xrightarrow{\mu} \Delta$ then there is a collection of distributions Δ_i and probabilities p_i such that $s \xrightarrow{\mu} \Delta_i$ for each $i \in I$ and $\Delta = \sum_{i\in I} p_i \cdot \Delta_i$ with $\sum_{i\in I} p_i = 1$.

Relations over subdistributions obtained by *lifting* enjoy some very useful properties, which we encapsulate in the following definition.

Definition 2.6. [Left-decomposable] A binary relation over subdistributions, $\mathcal{R} \subseteq \mathcal{D}_{sub}(S) \times \mathcal{D}_{sub}(S)$, is called *left-decomposable* if $(\sum_{i\in I} p_i \cdot \Delta_i)\ \mathcal{R}\ \Theta$, where I is a finite index set, implies that Θ can be written as $(\sum_{i\in I} p_i \cdot \Theta_i)$ such that $\Delta_i\ \mathcal{R}\ \Theta_i$ for every $i \in I$. □

Proposition 2.7. For any $\mathcal{R} \subseteq S \times \mathcal{D}_{sub}(S)$ the relation *lift*($\mathcal{R}$) over subdistributions is left-decomposable.

Proof. Suppose $\Delta = (\sum_{i\in I} p_i \cdot \Delta_i)$ and Δ *lift*($\mathcal{R}$) Θ. We have to find a family of Θ_i such that

(i) Δ_i *lift*($\mathcal{R}$) Θ_i for each $i \in I$

(ii) $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$.

From the alternative characterisation of lifting, Lemma 2.5, we know that

$$\Delta = \sum_{j\in J} q_j \cdot \overline{s_j} \qquad\qquad s_j\ \mathcal{R}\ \Theta^j \qquad\qquad \Theta = \sum_{j\in J} q_j \cdot \Theta^j$$

Define Θ_i to be

$$\sum_{s\in\lceil\Delta_i\rceil} \Delta_i(s) \cdot \Big(\sum_{\{\, j\in J \mid s=s_j \,\}} \frac{q_j}{\Delta(s)} \cdot \Theta^j\Big)$$

Note that $\Delta(s)$ can be written as $\sum_{\{ j \in J \mid s = s_j \}} q_j$ and therefore

$$\Delta_i = \sum_{s \in \lceil \Delta_i \rceil} \Delta_i(s) \cdot (\sum_{\{ j \in J \mid s = s_j \}} \frac{q_j}{\Delta(s)} \cdot \overline{s_j})$$

Since $s_j \mathcal{R} \Theta_j$ this establishes (i) above.

To establish (ii) above let us first abbreviate the sum $\sum_{\{ j \in J \mid s = s_j \}} \frac{q_j}{\Delta(s)} \cdot \Theta^j$ to $X(s)$. Then $\sum_{i \in I} p_i \cdot \Theta_i$ can be written as

$$\begin{aligned}
& \sum_{s \in \lceil \Delta \rceil} \sum_{i \in I} p_i \cdot \Delta_i(s) \cdot X(s) \\
= & \sum_{s \in \lceil \Delta \rceil} (\sum_{i \in I} p_i \cdot \Delta_i(s)) \cdot X(s) \\
= & \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot X(s)
\end{aligned}$$

The last equation is justified by the fact that $\Delta(s) = \sum_{i \in I} p_i \cdot \Delta_i(s)$.

Now $\Delta(s) \cdot X(s) = \sum_{\{ j \in J \mid s = s_j \}} q_j \cdot \Theta^j$ and therefore we have

$$\begin{aligned}
\sum_{i \in I} p_i \cdot \Theta_i &= \sum_{s \in \lceil \Delta \rceil} \sum_{\{ j \in J \mid s = s_j \}} q_j \cdot \Theta^j \\
&= \sum_{j \in J} q_j \cdot \Theta^j \\
&= \Theta
\end{aligned}$$

□

As a consequence we can now assume that the action relations $\Delta \xrightarrow{\mu} \Theta$ over distributions are both linear and left-decomposable.

As remarked in [6], in MLTSs it is necessary to have an infinitary version of the standard weak internal action $\xrightarrow{\tau}{}^*$ used in LTSs.

Definition 2.8. [Hyper-derivations] In a MLTS a hyper-derivation consists of a collec-

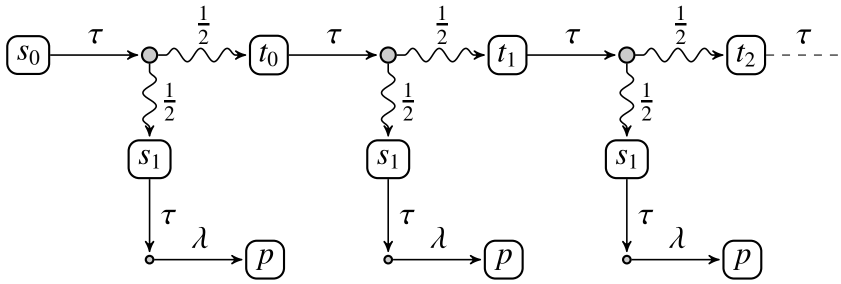

Figure 4: Limiting internal moves

tion of subdistributions $\Delta, \Delta_k^{\rightarrow}, \Delta_k^{\times}$, for $k \geq 0$, with the following properties:

$$\begin{array}{rcl} \Delta & = & \Delta_0^{\rightarrow} + \Delta_0^{\times} \\ \Delta_0^{\rightarrow} & \xrightarrow{\tau} & \Delta_1^{\rightarrow} + \Delta_1^{\times} \\ & \vdots & \\ \Delta_k^{\rightarrow} & \xrightarrow{\tau} & \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times} \\ & \vdots & \\ \Delta' & = & \sum_{k=0}^{\infty} \Delta_k^{\times} \end{array} \tag{1}$$

We call Δ' a *hyper-derivative* of Δ, and write $\Delta \Longrightarrow \Delta'$. □

Example 2.9. Consider the MLTS in Figure 4, where for graphical convenience we have multiple occurrences of the same state s_1. Starting from the initial state s_0 an ever increasing number of internal τ moves are performed before the eventual timed λ action, but with ever decreasing probability. This is captured formally in the following hyper-derivation:

$$\begin{array}{lclcl} \overline{s_0} & = & \overline{s_0} & + & \varepsilon \\ \overline{s_0} & \xrightarrow{\tau} & \frac{1}{2}\cdot\overline{t_0} & + & \frac{1}{2}\cdot\overline{s_1} \\ \frac{1}{2}\cdot\overline{t_0} & \xrightarrow{\tau} & \frac{1}{2^2}\cdot\overline{t_1} & + & \frac{1}{2^2}\cdot\overline{s_1} \\ & \vdots & & & \\ \frac{1}{2^{(k+1)}}\cdot\overline{t_k} & \xrightarrow{\tau} & \frac{1}{2^{(k+2)}}\cdot\overline{t_{(k+1)}} & + & \frac{1}{2^{(k+2)}}\cdot\overline{s_1} \\ & \vdots & & & \end{array}$$

where ε represents the empty subdistribution. Therefore,

$$\overline{s_0} \Longrightarrow \sum_{k>0} \frac{1}{2^k} \cdot \overline{s_1}$$

that is $s_0 \Longrightarrow \overline{s_1}$, because $\overline{s_1} = \sum_{k>0} \frac{1}{2^k} \cdot \overline{s_1}$. □

Note that in general hyper-derivations are defined over subdistributions. But as our example shows they can lead to hyper-derivations between (full) distributions; indeed in the paper we will only use such instances of hyper-derivations. We refer to [6] for more comprehensive discussion on hyper-derivations where they have been studied in detail. Here we will only summarise the properties we require for the present paper.

Theorem 2.10. In an arbitrary MLTS, the relation $\Longrightarrow$ over distributions is

(i) linear

(ii) left-decomposable

(iii) reflexive and transitive.

Proof. See Appendix A. □

With these concepts we can now define the appropriate notion of weak moves in a MLTS, which we may then use to define our concept of bisimulations. We write $\Delta \stackrel{\tau}{\Longrightarrow} \Delta'$ to mean $\Delta \Longrightarrow \Delta'$ and $\Delta \stackrel{\alpha}{\Longrightarrow} \Delta'$, for $\alpha \in \mathsf{Act}_\delta \cup \mathbb{R}^+$, to mean $\Delta \Longrightarrow\stackrel{\alpha}{\longrightarrow}\Longrightarrow \Delta'$.
As a side remark we have:

Corollary 2.11. In an arbitrary MLTS, the action relations $\stackrel{\mu}{\Longrightarrow}$ are both linear and left-decomposable.

Proof. It is easy to check that both properties are preserved by composition; that is if $\mathcal{R}_i, i = 1, 2$, are linear, left-decomposable respectively, then so is $\mathcal{R}_1 \cdot \mathcal{R}_2$. The result now follows since $\stackrel{\mu}{\Longrightarrow}$ is formed by composition from two relations which we know are both linear and left-decomposable. □

2.3 Markov bisimulations

Definition 2.12. [Markov bisimulations] For $\mathcal{R} \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$, where S is the set of states in a MLTS, let $\mathcal{B}(\mathcal{R})$ be the relation over $\mathcal{D}(S) \times \mathcal{D}(S)$ determined by letting $\Delta\ \mathcal{B}(\mathcal{R})\ \Theta$ if, for each $\mu \in \mathsf{Act}_{\tau,\delta} \cup \mathbb{R}^+$ and all finite sets of probabilities $\{ p_i \mid i \in I \}$ satisfying $\sum_{i \in I} p_i = 1$,

(i) whenever $\Delta \overset{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta_i$, for any distributions Δ_i, there are some distributions Θ_i with $\Theta \overset{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta_i$, such that $\Delta_i \mathcal{R} \Theta_i$ for each $i \in I$

(ii) symmetrically, whenever $\Theta \overset{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta_i$, for any distributions Θ_i, there are some distributions Δ_i with $\Delta \overset{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta_i$, such that $\Delta_i \mathcal{R} \Theta_i$ for each $i \in I$.

A relation $\mathcal{R}$ is called a *Markov bisimulation* if $\mathcal{R} \subseteq \mathcal{B}(\mathcal{R})$. The largest Markov bisimulation, which is guaranteed to exist using standard arguments, is denoted by $\approx_{bis}$. For most of the paper *Markov bisimulation* will be abbreviated to simply *bisimulation*.
□

Proposition 2.13. $\approx_{bis}$ is an equivalence relation.

Proof. Straightforward because of the form of the functional $\mathcal{B}$. □

However, due to the use of weak arrows and the quantification over sets of probabilities, it is not easy to exhibit witness bisimulations. We therefore give an alternative characterisation of $\approx_{bis}$ in terms of a relation between *states* and distributions.

Definition 2.14. [Simple bisimulations] For $\mathcal{R} \subseteq S \times \mathcal{D}(S)$, where again S is the set of states in a MLTS, let $\mathcal{SB}(\mathcal{R})$ be the relation over $S \times \mathcal{D}(S)$ defined by letting $s\ \mathcal{SB}(\mathcal{R})\ \Theta$ if, for each $\mu \in \mathsf{Act}_{\tau,\delta} \cup \mathbb{R}^+$,

(i) whenever $s \overset{\mu}{\longrightarrow} \Delta'$, there is some $\Theta \overset{\mu}{\Longrightarrow} \Theta'$, such that $\Delta'\ \mathit{lift}(\mathcal{R})\ \Theta'$

(ii) there exists some $\Delta \in \mathcal{D}(S)$ such that $\overline{s} \overset{\tau}{\Longrightarrow} \Delta$ and $\Theta\ \mathit{lift}(\mathcal{R})\ \Delta$.

We use $\approx_{sbis}$ to denote the largest solution to $\mathcal{R} = \mathcal{SB}(\mathcal{R})$. □

Note that both forms of bisimulation equivalence are defined for MLTSs. But in the paper we will apply them to the states and distributions of MAs. For example, we write $s \approx_{sbis} \Delta$, where s is a state in a MA M and Δ a distribution, to mean $s \approx_{sbis} \Delta$ in the derived $\mathsf{mlts}(M)$.

Example 2.15. Consider again the MLTSs in Figure 3, derived from the MAs in Figure 1. Here $s \approx_{sbis} \overline{v}$ because the following relation

$$\{\langle s, \overline{v}\rangle,\ \langle s_1, \frac{1}{2}\cdot\overline{v_b} + \frac{1}{2}\cdot\overline{v_c}\rangle,\ \langle s_a, \overline{v_a}\rangle,\ \langle s_b, \overline{v_b}\rangle,\ \langle s_c, \overline{v_c}\rangle,\ \langle v, \overline{s}\rangle,\ \langle v_a, \overline{s_a}\rangle,\ \langle v_b, \overline{s_b}\rangle,\ \langle v_c, \overline{s_c}\rangle\}$$

is a simple bisimulation.

Now consider the MLTS in Figure 4. We have already seen in Example 2.9 that $\overline{s_0} \Longrightarrow \overline{s_1}$, and therefore $\overline{s_0} \overset{\lambda}{\Longrightarrow} \overline{p}$; with similar reasoning we can show that $t_i \overset{\lambda}{\Longrightarrow} \overline{p}$ for every $i \geq 0$. It follows that the relation

$$\{\,\langle s_i, \overline{\lambda.p}\rangle,\ \langle \lambda.p, \overline{s_i}\rangle\ \mid\ i = 0,1\,\} \cup \{\,\langle t_i, \overline{\lambda.p}\rangle, \langle \lambda.p, \overline{t_i}\rangle\ \mid\ i \geq 0\,\}\ \cup\ \{\langle \lambda.p, \overline{\lambda.p}\rangle, \langle p, \overline{p}\rangle\}$$

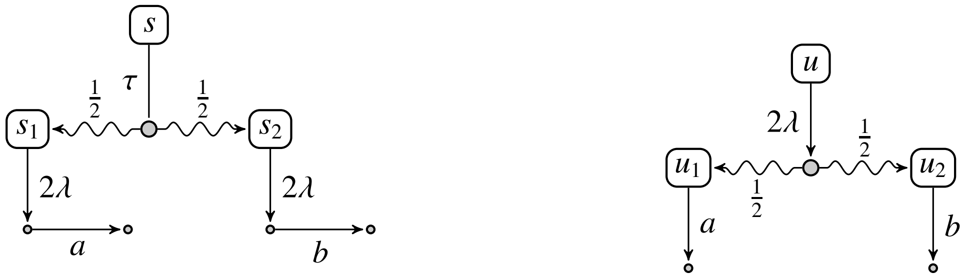

Figure 5: Derived MLTSs of MAs in Figure 2

is a simple bisimulation, and therefore $s_0 \approx_{sbis} \overline{\lambda.p}$, where, as we will see $\lambda.p$ describes in an obvious manner the MA which does the timed action at rate λ and evolves to the state p.

Now consider the MA in Figure 2. We describe their MLTSs in Figure 5 but note that the structure of the first automata does not change. Let us examine the big MLTS whose state space and transition relation are the unions of those in the two sub-systems. Here $s \not\approx_{sbis} \overline{u}$ because the transition $s \xrightarrow{\tau} \frac{1}{2} \cdot \overline{s_1} + \frac{1}{2} \cdot \overline{s_2}$ cannot be matched by any transition from u. The state u cannot enable internal actions, so the only weak internal transition from $\overline{u}$ is $\overline{u} \overset{\tau}{\Longrightarrow} \overline{u}$. However, the derivative $\overline{u}$ is not able to simulate $\frac{1}{2} \cdot \overline{s_1} + \frac{1}{2} \cdot \overline{s_2}$ according to the lifted relation $lift(\approx_{sbis})$. Suppose for a contradiction that $(\frac{1}{2} \cdot \overline{s_1} + \frac{1}{2} \cdot \overline{s_2})\ lift(\approx_{sbis})\ \overline{u}$. Then we must have $s_1 \approx_{sbis} \overline{u}$ and $s_2 \approx_{sbis} \overline{u}$; obviously neither of these holds. □

The precise relationship between the two forms of bisimulations are given by:

Theorem 2.16. Let Δ and Θ be two distributions in a finitary MLTS.

(i) If $\Delta \approx_{bis} \Theta$ then there is some Θ' with $\Theta \overset{\tau}{\Longrightarrow} \Theta'$ and $\Delta\ lift(\approx_{sbis})\ \Theta'$

(ii) If $\Delta\ lift(\approx_{sbis})\ \Theta$ then $\Delta \approx_{bis} \Theta$.

The remainder of this section is devoted to the proof of this theorem; it involves first developing a number of subsidiary results.

Proposition 2.17. Suppose $\Delta\ lift(\approx_{sbis})\ \Theta$ and $\Delta \xrightarrow{\mu} \Delta'$, in an arbitrary MLTS. Then there exists some distribution Θ' such that $\Theta \overset{\mu}{\Longrightarrow} \Theta'$ and $\Delta'\ lift(\approx_{sbis})\ \Theta'$.

Proof. Suppose $\Delta\ lift(\approx_{sbis})\ \Theta$ and $\Delta \xrightarrow{\mu} \Delta'$. By Lemma 2.5 there is a finite index set I such that (i) $\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}$, (ii) $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$, and (iii) $s_i \approx_{sbis} \Theta_i$ for each $i \in I$. By the condition $\Delta \xrightarrow{\mu} \Delta'$, (i) and Proposition 2.7, we can decompose Δ' into $\sum_{i\in I} p_i \cdot \Delta'_i$ for some Δ'_i such that $\overline{s_i} \xrightarrow{\mu} \Delta'_i$. By Lemma 2.5 again, for each $i \in I$, there is an index set J_i such that $\Delta'_i = \sum_{j\in J_i} q_{ij} \cdot \Delta'_{ij}$ and $s_i \xrightarrow{\mu} \Delta'_{ij}$ for each $j \in J_i$ and $\sum_{j\in J_i} q_{ij} = 1$. By (iii) there is some Θ'_{ij} such that $\Theta_i \stackrel{\mu}{\Longrightarrow} \Theta'_{ij}$ and $\Delta'_{ij}\ lift(\approx_{sbis})\ \Theta'_{ij}$. Let $\Theta' = \sum_{i\in I, j\in J_i} p_i q_{ij} \cdot \Theta'_{ij}$. Since $\stackrel{\mu}{\Longrightarrow}$ is linear by Corollary 2.11, we know that $\Theta = \sum_{i\in I} p_i \sum_{j\in J_i} q_{ij}\Theta_i \stackrel{\mu}{\Longrightarrow} \Theta'$. By the linearity of $lift(\approx_{sbis})$, we notice that $\Delta' = (\sum_{i\in I} p_i \sum_{j\in J_i} q_{ij} \cdot \Delta'_{ij})\ lift(\approx_{sbis})\ \Theta'$. □

Theorem 2.18. In a finitary MLTS, if $s \approx_{sbis} \Theta$ and $\overline{s} \stackrel{\tau}{\Longrightarrow} \Delta'$ then there is some Θ' with $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$ and $\Delta'\ lift(\approx_{sbis})\ \Theta'$.

Proof. See Appendix A. The proof depends crucially on the restriction to finitary MLTSs. □

Corollary 2.19. In a finitary MLTS, suppose $\Delta\ lift(\approx_{sbis})\ \Theta$ and $\Delta \stackrel{\mu}{\Longrightarrow} \Delta'$. Then there is some Θ' with $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$ and $\Delta'\ lift(\approx_{sbis})\ \Theta'$.

Proof. Given the two previous results this is fairly straightforward. Suppose $\Delta \stackrel{\mu}{\Longrightarrow} \Delta'$ and $\Delta\ \ lift(\approx_{sbis})\ \ \Theta$. If μ is τ then the required Θ' follows by an application of the theorem, since the relation $\stackrel{\tau}{\Longrightarrow}$ is actually defined to be $\Longrightarrow$.

Otherwise, by definition we know $\Delta \Longrightarrow \Delta_1$, $\Delta_1 \xrightarrow{\mu} \Delta_2$ and $\Delta_2 \Longrightarrow \Delta'$. An application of the theorem gives a Θ_1 such that $\Theta \Longrightarrow \Theta_1$ and $\Delta_1\ lift(\approx_{sbis})\ \Theta_1$. An application of the proposition gives a Θ_2 such that $\Theta_1 \stackrel{\mu}{\Longrightarrow} \Theta_2$ and $\Delta_2\ lift(\approx_{sbis})\ \Theta_2$. Finally another application of the theorem gives $\Theta_2 \Longrightarrow \Theta'$ such that $\Delta'\ lift(\approx_{sbis})\ \Theta'$.

The result now follows since the transitivity of hyper-derivations, Theorem 2.10, gives $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$. □

Theorem 2.20. In a finitary MLTS, $\Delta\ lift(\approx_{sbis})\ \Theta$ implies $\Delta \approx_{bis} \Theta$.

Proof. Let $\mathcal{R}$ denote the relation $lift(\approx_{sbis}) \cup (lift(\approx_{sbis}))^{-1}$. We show that $\mathcal{R}$ is a bisimulation relation, that is $\mathcal{R} \subseteq \mathcal{B}(\mathcal{R})$, from which the result follows.

Suppose that $\Delta\ \mathcal{R}\ \Theta$. There are two possibilities:

(a) $\Delta\ lift(\approx_{sbis})\ \Theta$.

To show $\Delta\ \mathcal{B}(\mathcal{R})\ \Theta$ first suppose $\Delta \stackrel{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta'_i$. By Corollary 2.19 there is some distribution Θ' with $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$ and $(\sum_{i\in I} p_i \cdot \Delta'_i)\ lift(\approx_{sbis})\ \Theta'$. But by Proposition 2.7 we know that the relation $lift(\approx_{sbis})$ is left-decomposable. This means that $\Theta' = \sum_{i\in I} p_i \cdot \Theta'_i$ for some distributions Θ'_i such that $\Delta'_i\ lift(\approx_{sbis})\ \Theta'_i$ for each $i \in I$. We hence have the required matching move from Θ.

For the converse suppose $\Theta \stackrel{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta'_i$. We have to find a matching move, $\Delta \stackrel{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta'_i$, such that $\Delta'_i \mathrel{\mathcal{R}} \Theta'_i$. In fact it is sufficient to find a move $\Delta \stackrel{\mu}{\Longrightarrow} \Delta'$ such that $\sum_{i\in I} p_i \cdot \Theta'_i \; \mathit{lift}(\approx_{sbis}) \; \Delta'$, since $(\mathit{lift}(\approx_{sbis}))^{-1} \subseteq \mathcal{R}$ and the deconstruction of Δ' into the required sum $\sum_{i\in I} p_i \cdot \Delta'_i$ will again follow from the fact that $\mathit{lift}(\approx_{sbis})$ is left-decomposable. To this end let us abbreviate $\sum_{i\in I} p_i \cdot \Theta'_i$ to simply Θ'.

We know from $\Delta \; \mathit{lift}(\approx_{sbis}) \; \Theta$, using left-decomposability, that $\Theta = \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot \Theta_s$ for some Θ_s with $s \approx_{sbis} \Theta_s$. Then by the definition of $\approx_{sbis}$, $\overline{s} \stackrel{\tau}{\Longrightarrow} \Delta_s$ for some Δ_s such that $\Theta_s \; \mathit{lift}(\approx_{sbis}) \; \Delta_s$. Now using the left-decomposability of weak actions, from $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$ we have $\Theta' = \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot \Theta'_s$ such that $\Theta_s \stackrel{\mu}{\Longrightarrow} \Theta'_s$, for each s in the support of Δ.

Applying Corollary 2.19 to $\Theta_s \; \mathit{lift}(\approx_{sbis}) \; \Delta_s$ we have, again for each s in the support of Δ, a matching move $\Delta_s \stackrel{\mu}{\Longrightarrow} \Delta'_s$ such that $\Theta'_s \; \mathit{lift}(\approx_{sbis}) \; \Delta'_s$.

But, since $\overline{s} \stackrel{\tau}{\Longrightarrow} \Delta_s$, this gives $\overline{s} \stackrel{\mu}{\Longrightarrow} \Delta'_s$ for each $s \in \lceil\Delta\rceil$; using the linearity of weak moves, these moves from the states s in the support of Δ can be combined to obtain the action $\Delta \stackrel{\mu}{\Longrightarrow} \sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\Delta'_s$. The required Δ' is this sum, $\sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\Delta'_s$, since linearity of $\mathit{lift}(\approx_{sbis})$ gives $\Delta' (\mathit{lift}(\approx_{sbis}))^{-1} \Theta'$.

(b) The second possibility is that $\Delta \; (\mathit{lift}(\approx_{sbis}))^{-1} \; \Theta$, that is $\Theta \; \mathit{lift}(\approx_{sbis}) \; \Delta$. But in this case the proof that the relevant moves from Θ and Δ can be properly matched is exactly the same as in case (a).

□

We also have a partial converse to Theorem 2.20:

Proposition 2.21. In a finitary MLTS, $\overline{s} \approx_{bis} \Theta$ implies $s \approx_{sbis} \Theta$.

Proof. Let $\approx^s_{bis}$ be the restriction of $\approx_{bis}$ to $S \times \mathcal{D}(S)$, in the sense that $s \approx^s_{bis} \Theta$ whenever $\overline{s} \approx_{bis} \Theta$.

We show that $\approx^s_{bis} \subseteq \mathcal{SB}(\approx^s_{bis})$. Suppose $s \approx^s_{bis} \Theta$.

(i) First suppose $s \stackrel{\mu}{\longrightarrow} \Delta'$. Then since $\overline{s} \approx_{bis} \Theta$ there must exist some $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$ such that $\Delta' \approx_{bis} \Theta'$. Now consider the degenerate action $\Delta' \stackrel{\tau}{\Longrightarrow} \sum_{t\in\lceil\Delta'\rceil} \Delta'(t) \cdot \overline{t}$. There must be a matching move from Θ', $\Theta' \stackrel{\tau}{\Longrightarrow} \Theta'' = \sum_{t\in\lceil\Delta'\rceil} \Delta'(t) \cdot \Theta'_t$ such that $\overline{t} \approx_{bis} \Theta'_t$, that is $t \approx^s_{bis} \Theta'_t$ for each $t \in \lceil\Delta'\rceil$.

By linearity, this means $\Delta' \; \mathit{lift}(\approx^s_{bis}) \; \Theta''$ and by the transitivity of $\Longrightarrow$ we have the required matching move $\Theta \stackrel{\mu}{\Longrightarrow} \Theta''$.

(ii) To establish the second requirement, consider the trivial move $\Theta \stackrel{\tau}{\Longrightarrow} \Theta$. Since $\overline{s} \approx_{bis} \Theta$ there must exist a corresponding move $\overline{s} \stackrel{\tau}{\Longrightarrow} \Delta$ such that $\Delta \approx_{bis} \Theta$. By Proposition 2.13, we also have $\Theta \approx_{bis} \Delta$. Now by an argument symmetric to that used in part (i) we can show that this implies the existence of some Δ' such that $\Delta \stackrel{\tau}{\Longrightarrow} \Delta'$, that is $\overline{s} \stackrel{\tau}{\Longrightarrow} \Delta'$ and $\Theta\ lift(\approx^{s}_{bis})\ \Delta'$.

$\square$

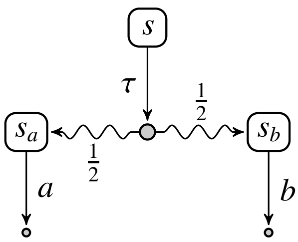

Figure 6: An MLTS

But in general the relations $\approx_{bis}$ and $lift(\approx_{sbis})$ do not coincide for arbitrary distributions. Consider the MLTS in Figure 6 and let Δ denote the distribution $\frac{1}{2} \cdot \overline{s_a} + \frac{1}{2} \cdot \overline{s_b}$. Then it is easy to see that $\Delta \approx_{bis} \overline{s}$ but not $\Delta\ lift(\approx_{sbis})\ \overline{s}$; the latter follows because the point distribution $\overline{s}$ can not be decomposed as $\frac{1}{2} \cdot \Theta_a + \frac{1}{2} \cdot \Theta_b$ so that $\Theta_a \approx_{sbis} \overline{s_a}$ and $\Theta_b \approx_{sbis} \overline{s_b}$.
The nearest to a general converse to Theorem 2.20 is the following:

Proposition 2.22. Suppose $\Delta \approx_{bis} \Theta$ in a finitary MLTS. Then there is some Θ' with $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$ and $\Delta\ lift(\approx_{sbis})\ \Theta'$.

Proof. Now suppose $\Delta \approx_{bis} \Theta$. We can rewrite Δ as $\sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \overline{s}$, and trivially $\Delta \stackrel{\tau}{\Longrightarrow} \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \overline{s}$. Since $\approx_{bis}$ is a bisimulation this move can be matched by some $\Theta \stackrel{\tau}{\Longrightarrow} \Theta' = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Theta_s$ such that $\overline{s} \approx_{bis} \Theta_s$. But we have just shown in the previous proposition that this means $s \approx_{sbis} \Theta_s$.

By Definition 2.4, $\Delta\ lift(\approx_{sbis})\ \Theta'$ and therefore $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$ is the required move. $\square$

Bisimulation equivalence, $\approx_{bis}$ from Definition 2.12, is our primary behavioural equivalence but we will often develop properties of it via the connection we have just established with $\approx_{sbis}$ from Definition 2.14; the latter is more amenable as it only required strong moves to be matched. However we can also prove properties of $\approx_{sbis}$ by using this connection to bisimulation equivalence; a simple example is the following:

Corollary 2.23. In a finitary MLTS suppose $s \approx_{sbis} \Theta$ where $s \not\xrightarrow{\tau}$. Then whenever $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$ it follows that $s \approx_{sbis} \Theta'$.

Proof. Suppose $s \approx_{sbis} \Theta$, which means $\overline{s}\ lift(\approx_{sbis})\ \Theta$ and therefore by Theorem 2.20 $\overline{s} \approx_{bis} \Theta$. The move $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$ must be matched by a corresponding move from $\overline{s}$. However since $s \not\xrightarrow{\tau}$ the only possibility is the empty move, giving $\overline{s} \approx_{bis} \Theta'$. Now by Proposition 2.21 we have the required $s \approx_{sbis} \Theta'$. $\square$

Corollary 2.24. In any finitary MLTS, the relation $\approx_{bis}$ is linear.

Proof. Consider any collection of probabilities p_i with $\sum_{i\in I} p_i = 1$, where I is a finite index set. Suppose further that $\Delta_i \approx_{bis} \Theta_i$ for each $i \in I$. We need to show that $\Delta \approx_{bis} \Theta$, where $\Delta = \sum_{i\in I} p_i \cdot \Delta_i$ and $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$.

By Proposition 2.22, there is some Θ'_i with $\Theta_i \overset{\tau}{\Longrightarrow} \Theta'_i$ and $\Delta_i\ \mathit{lift}(\approx_{sbis})\ \Theta'_i$. By Theorem 2.10 (i) and Definition 2.4, both $\overset{\tau}{\Longrightarrow}$ and $\mathit{lift}(\approx_{sbis})$ are linear. Therefore, we have $\Theta \overset{\tau}{\Longrightarrow} \Theta'$ and $\Delta\ \mathit{lift}(\approx_{sbis})\ \Theta'$, where $\Theta' = \sum_{i\in I} p_i \cdot \Theta'_i$. It follows from Theorem 2.20 that $\Delta \approx_{bis} \Theta'$.

Now for any transition $\Delta \overset{\mu}{\Longrightarrow} (\sum_{j\in J} q_j \cdot \Delta_j)$, where J is finite, there is a matching transition $\Theta' \overset{\mu}{\Longrightarrow} (\sum_{j\in J} q_j \cdot \Theta_j)$ such that $\Delta_j \approx_{bis} \Theta_j$ for each $j \in J$. Note that we also have the transition $\Theta \overset{\mu}{\Longrightarrow} (\sum_{j\in J} q_j \cdot \Theta_j)$ according to the transitivity of $\overset{\tau}{\Longrightarrow}$. By symmetrical arguments, any transition $\Theta \overset{\mu}{\Longrightarrow} (\sum_{j\in J} q_j \cdot \Theta_j)$ can be matched by some transition $\Delta \overset{\mu}{\Longrightarrow} (\sum_{j\in J} q_j \cdot \Delta_j)$ such that $\Delta_j \approx_{bis} \Theta_j$ for each $j \in J$. □

3 Composing Markov automata

In this section, we first introduce and language mCCS and show how to interpret it as a Markov automaton. Then we prove that bisimulation equivalence is preserved by most of the operators in the language.

3.1 mCCS

Here we assume that the set of actions Act is equipped with a complementation function $\bar{\cdot} : \mathsf{Act} \to \mathsf{Act}$ satisfying $\bar{\bar{a}} = a$; we say $\bar{a}$ is the complement of a. Then given two MAs, $M_i = \langle S_1, \mathsf{Act}_\tau, \to, \mapsto, \rangle$ for $i = 1, 2$, their composition $(M_1 \mid M_2)$ is given by $\langle S_1 \mid S_2, \mathsf{Act}_\tau, \to, \mapsto, \rangle$ where the set of states $S_1 \mid S_2 = \{\, s_1 \mid s_2 \mid s_i \in S_i, i = 1, 2 \,\}$ and the relations are determined by the rules in Figure 7. The rules use the obvious extension of the function | on pairs of states to pairs of distributions. To be precise $\Delta \mid \Theta$ is the distribution defined by:

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

This construction can also be explained as follows:

Lemma 3.1.

(i) $\Delta \mid \bar{t} = \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot (\bar{s} \mid \bar{t})$

$$\text{(PAR.A)}\quad \frac{s \xrightarrow{\mu} \Delta}{s \mid t \xrightarrow{\mu} \Delta \mid \overline{t}}\ \mu \in \mathsf{Act}_\tau \qquad \text{(PAR.R)}\quad \frac{t \xrightarrow{\mu} \Theta}{s \mid t \xrightarrow{\mu} \overline{s} \mid \Theta}\ \mu \in \mathsf{Act}_\tau$$

$$\text{(PAR.I)}\quad \frac{s \xrightarrow{a} \Delta,\ t \xrightarrow{\overline{a}} \Theta}{s \mid t \xrightarrow{\tau} \Delta \mid \Theta}$$

$$\text{(PAR.L.T)}\quad \frac{s \overset{\mathbf{d}}{\mapsto} s',\ t \overset{\delta}{\mapsto} t',\ s \mid t \not\xrightarrow{\tau}}{s \mid t \overset{\mathbf{d}}{\mapsto} s' \mid t'} \qquad \text{(PAR.R.T)}\quad \frac{s \overset{\delta}{\mapsto} s', t \overset{\mathbf{d}}{\mapsto} t',\ s \mid t \not\xrightarrow{\tau}}{s \mid t \overset{\mathbf{d}}{\mapsto} s' \mid t'}\ \mathbf{d} = \delta, \lambda$$

Figure 7: Composing Markov automata

(ii) $\Delta \mid \Theta = \sum_{t \in \lceil \Theta \rceil} \Theta(t) \cdot (\Delta \mid \overline{t})$.

Proof. Straightforward calculation. □

Lemma 3.2. If M_1 and M_2 are Markov automata, then so is $(M_1 \mid M_2)$. □

Proof. Straightforward. It is simply a question of checking that the resulting automata satisfy conditions (a) and (b) of Definition 2.1. □

We can internalise this composition relation by considering MAs which are *par-closed*:

Definition 3.3. A Markov automaton M is *par-closed* if $(M \mid M)$ is already a sub-MA of M. □

The simplest way of constructing a par-closed MA is by interpreting a process algebra as a universal Markov automaton. To this end we introduce the language mCCS whose terms are given by:

$$\begin{aligned} P, Q \ &::= \ \mathbf{0} \ \mid \ \delta.P \ \mid \ \lambda.P,\ \lambda \in \mathbb{R}^+ \ \mid \ \mu{:}D,\ \mu \in \mathsf{Act}_\tau \ \mid \ P + Q \ \mid \ P \mid Q \ \mid \ A \\ D \ &::= \ (\oplus_{i \in I} p_i \cdot P_i),\ \text{where } \textstyle\sum_{i \in I} p_i = 1 \end{aligned}$$

where A ranges over a set of process constants, with each of which is associated a definition, $A \Leftarrow \mathrm{Def}(A)$, where $\mathrm{Def}(A)$ is some term in the language. mCCS is interpreted

$$\text{(ACTION)} \quad \mu{:}D \xrightarrow{\mu} [\![D]\!]$$

$$\text{(RECURSION)} \quad \frac{\mathrm{Def}(A) \xrightarrow{\alpha} \Delta}{A \xrightarrow{\alpha} \Delta} \quad \alpha = \mu, \lambda, \delta$$

$$\text{(EXT.L)} \quad \frac{P \xrightarrow{\mu} \Delta,}{P + Q \xrightarrow{\mu} \Delta}$$

$$\text{(EXT.L.L)} \quad \frac{P \overset{\lambda}{\mapsto} P',\ Q \not\xrightarrow{\tau}}{P + Q \overset{\lambda}{\mapsto} P'}$$

$$\text{(DELAY)} \quad \lambda.P \overset{\lambda}{\mapsto} P,$$

$$\text{(DELAY.}\delta\text{)} \quad \lambda.P \overset{\delta}{\mapsto} \lambda.P$$

$$(\delta\text{.E}) \quad \frac{P \xrightarrow{\mu} \Delta}{\delta.P \xrightarrow{\mu} \Delta}$$

$$(\delta\text{.D}) \quad \frac{P \not\xrightarrow{\tau}}{\delta.P \overset{\delta}{\mapsto} P}$$

$$\text{(EXT)} \quad \frac{P \overset{\delta}{\mapsto} P',\ Q \overset{\delta}{\mapsto} Q'}{P + Q \overset{\delta}{\mapsto} P' + Q'}$$

$$\text{(EXT.D.L)} \quad \frac{P \overset{\delta}{\mapsto} P',\ Q \overset{\delta}{\not\mapsto},\ Q \not\xrightarrow{\tau}}{P + Q \overset{\delta}{\mapsto} P'}$$

Figure 8: Operational semantics of mCCS

as a Markov automaton whose states are all the terms in the language, and whose arrows are determined by the rules in Figure 8, together with those in Figure 7; we have omitted the obvious symmetric counterparts to the rules (EXT.L), (EXT.L.L) and (EXT.D.L). Other operations, such as the standard hiding $Q\backslash a, a \in Act$, can also be easily given an interpretation. We say a process P from mCCS is *finitary* if the sub-MA consisting of all states reachable from P is finitary, and we use *finitary* mCCS to refer to the MA consisting of all such finitary P.

The rule (ACTION) uses the notation $[\![D]\!]$, where D has the form $(\oplus_{i\in I} p_i \cdot P_i)$, to denote the obvious distribution over process terms, whose support consists of the terms $P_1, \ldots P_n$, each with weight p_i respectively. Most of the other rules should be self-explanatory, although the justification for the rules for λ transitions depends on non-trivial properties of exponential distributions; these are explained in detail in [11]. Nevertheless this interpretation of mCCS is quite different from that of other Markovian process calculi, such as those in [11, 3]. First the actions $\mu{:}D$ are *insistent* rather than *lazy*; they do not allow time to pass. For example, the process $(\lambda.Q \mid a{:}P)$ is stuck with respect to time; it cannot let time pass. This is because the parallel operator requires each component to perform a timed transition which $a{:}P$ can not do, before time can pass. To obtain lazy actions one can define $a.P$ by the declaration $A \Leftarrow a{:}P + \delta.A$. Then we have the transition

$$\lambda.Q \mid a.P \overset{\lambda}{\mapsto} Q \mid a.P$$

by an application of the rule (PAR.L.T) to the transitions $\lambda.Q \overset{\lambda}{\mapsto} Q$ and $a.P \overset{\delta}{\mapsto} a.P$.

The parallel operator is even more constraining in that at most one of its components can perform a definite delay. Again, this is reminiscent of many existing Markovian process algebras [2, 3], although these tend to have delays associated with external actions. But in the setting of mCCS the net effect is an operational semantics very similar to that in [9]. For example, consider the process $Q = (\lambda_1.P_1 \mid \lambda_2.P_2)$. This has three timed transitions

(i) $Q \overset{\lambda_1}{\mapsto} (P_1 \mid \lambda_2.P_2)$ via an application of the rule (PAR.L.T) to the transitions $\lambda_1.P_1 \overset{\lambda_1}{\mapsto} P_1$ and $\lambda_2.P_2 \overset{\delta}{\mapsto} \lambda_2.P_2$

(ii) $Q \overset{\lambda_2}{\mapsto} (\lambda_1 P_2 \mid P_2)$ via an application of (PAR.R.T) to the transitions $\lambda_1.P_1 \overset{\delta}{\mapsto} \lambda_1.P$ and $\lambda_2.P_2 \overset{\lambda_2}{\mapsto} P_2$

(iii) $Q \overset{\delta}{\mapsto} Q$ via an application of either of (PAR.L.T) or (PAR.R.T) to the transitions $\lambda_1.P_1 \overset{\delta}{\mapsto} \lambda_1.P_1$ and $\lambda_1.P_1 \overset{\delta}{\mapsto} \lambda_1.P_1$.

Proposition 3.4. mCCS, endowed with the actions from Figures 7 and 8, is a Markov automaton.

Proof. It is just a matter of checking that the rules enforce the properties (a) and (b) from Definition 2.1. □

The language CCS is a sublangauge of our mCCS. Let $\approx$ be the standard definition of observational equivalence for CCS, for example as defined in Definition 6 on page 109 of [18]. Our behavioural theory is a conservative extension of this standard theory:

Proposition 3.5. For all terms in CCS $P \approx Q$ if and only if $\overline{P} \approx_{bis} \overline{Q}$.

Proof. Let $P \overset{\mu}{\mapsto} P'$, $\mu \in \mathsf{Act}_\tau$, be the standard operational semantics for CCS terms, as given for example in Chapter 2.5 of [18]. Then it is straightforward to prove:

- $P \overset{\mu}{\mapsto} P'$ implies $P \xrightarrow{\mu} \overline{P'}$
- $P \xrightarrow{\mu} \Delta$ implies $\Delta = \overline{P'}$ for some P' such that $P \overset{\mu}{\mapsto} P'$

From these two points it is straightforward to prove the result. □

3.2 Compositionality

The main operator of interest in mCCS is the parallel composition; we show that it preserves bisimulation equivalence. This requires some preliminary results, particularly on composing actions from the components of a parallel composition.

Lemma 3.6. In a par-closed MLTS,

(i) $\Delta \xrightarrow{\mu} \Delta'$ implies $\Delta \mid \Theta \xrightarrow{\mu} \Delta' \mid \Theta$, for $\mu \in \mathsf{Act}_\tau$

(ii) $\Delta_1 \xrightarrow{a} \Delta_1'$ and $\Delta_2 \xrightarrow{\overline{a}} \Delta_2'$ implies $\Delta_1 \mid \Delta_2 \xrightarrow{\tau} \Delta_1' \mid \Delta_2'$

(iii) $\Delta_1 \xrightarrow{\mathbf{d}} \Delta_1'$ and $\Delta_2 \xrightarrow{\delta} \Delta_2'$ implies $\Delta_1 \mid \Delta_2 \xrightarrow{\mathbf{d}} \Delta_1' \mid \Delta_2'$, for $\mathbf{d} = \delta, \lambda$.

Proof. Each case follows by straightforward linearity arguments. As an example we outline the proof of (i). $\Delta \xrightarrow{\mu} \Delta'$ means that

$$\Delta = \sum_{i \in I} p_i \cdot \overline{s_i} \qquad\qquad s_i \xrightarrow{\mu} \Delta_i \qquad\qquad \Delta' = \sum_{i \in I} p_i \cdot \Delta_i$$

For any state t, $s_i \mid t \xrightarrow{\mu} \Delta_i \mid \overline{t}$ using the rule (PAR.A) in Figure 7. By linearity we have $\sum_{i \in I} p_i \cdot (\overline{s_i} \mid \overline{t}) \xrightarrow{\mu} \sum_{i \in I} p_i \cdot (\Delta_i \mid \overline{t})$ and this may be rendered as

$$\Delta \mid \overline{t} \xrightarrow{\mu} \Delta' \mid \overline{t}$$

By the second part of Lemma 3.1 $(\Delta \mid \Theta)$ may be written as $\sum_{t \in \lceil \Theta \rceil} \Theta(t) \cdot (\Delta \mid t)$ and therefore another application of linearity gives $\Delta \mid \Theta \xrightarrow{\mu} \sum_{t \in \lceil \Theta \rceil} \Theta(t) \cdot (\Delta' \mid \overline{t})$ and by the same result this residual coincides with $(\Delta' \mid \Theta)$. □

Lemma 3.7. In a par-closed MLTS,

(i) $\Delta \Longrightarrow \Delta'$ implies $\Delta \mid \Theta \Longrightarrow \Delta' \mid \Theta$

(ii) $\Delta \overset{\mu}{\Longrightarrow} \Delta'$ implies $\Delta \mid \Theta \overset{\mu}{\Longrightarrow} \Delta' \mid \Theta$, for $\mu \in \mathsf{Act}_\mu$

(iii) $\Delta_1 \overset{a}{\Longrightarrow} \Delta_1'$ and $\Delta_2 \overset{\overline{a}}{\Longrightarrow} \Delta_2'$ implies $\Delta_1 \mid \Delta_2 \overset{\tau}{\Longrightarrow} \Delta_1' \mid \Delta_2'$

(iv) $\Delta_1 \overset{\mathbf{d}}{\Longrightarrow} \Delta_1'$ and $\Delta_2 \overset{\delta}{\Longrightarrow} \Delta_2'$ implies $\Delta_1 \mid \Delta_2 \overset{\mathbf{d}}{\Longrightarrow} \Delta_1' \mid \Delta_2'$, for $\mathbf{d} = \delta, \lambda$

Proof. Parts (ii), (iii) and (iv) follow from (i) and the corresponding result in the previous lemma.

For (i) suppose $\Delta \Longrightarrow \Delta'$. First note that a hyper-derivation from Δ to $\sum_{k=0}^{\infty} \Delta_k^{\times} = \Delta'$, as in Definition 2.8, can easily be transformed into a hyper-derivation from $(\Delta \mid \overline{t})$ to $\sum_{k=0}^{\infty}(\Delta_k^{\times} \mid \overline{t})$. This means that for any state t we have a $(\Delta \mid \overline{t}) \Longrightarrow (\Delta' \mid \overline{t})$.

By the second part of Lemma 3.1 $(\Delta \mid \Theta)$ can be written as $\sum_{t \in \lceil\Theta\rceil} \Theta(t) \cdot (\Delta \mid \overline{t})$, and since $\Longrightarrow$ is linear, Theorem 2.10, this means $(\Delta \mid \Theta) \Longrightarrow \sum_{t \in \lceil\Theta\rceil} \Theta(t) \cdot (\Delta' \mid \overline{t})$ and again Lemma 3.1 renders this residual to be $(\Delta' \mid \Theta)$. □

Composing weighted actions in a MLTS is more complicated although we are helped by the fact that both weighted actions and δ actions are unique, if they exist. For example, if $\Delta \xrightarrow{\delta} \Delta'$ then we know exactly the structure Δ' must take. For every $s \in \lceil\Delta\rceil$ there is a unique distribution Δ_δ^s such that $s \xrightarrow{\delta} \Delta_\delta^s$ and Δ' must coincide with $\Delta_\delta = \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \Delta_\delta^s$. Similarly, if Δ has any weighted action it must take the form $\Delta \xrightarrow{\lambda} \Delta_w$ for some $\lambda \in \mathbb{R}^+$ where $\Delta_w = \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \Delta_w^s$ and each Δ_w^s is the unique distribution, guaranteed to exist, such that $s \xrightarrow{\lambda} \Delta_w^s$.

Theorem 3.8. Suppose

- $\Delta \xrightarrow{\beta} \Delta_\beta$ and $\Delta \xrightarrow{\delta} \Delta_\delta$
- $\Theta \xrightarrow{\gamma} \Theta_\gamma$ and $\Theta \xrightarrow{\delta} \Theta_\delta$

Then $\Delta \mid \Theta \xrightarrow{\beta+\gamma} \frac{\beta}{\beta+\gamma} \cdot (\Delta_\beta \mid \Theta_\delta) + \frac{\gamma}{\beta+\gamma} \cdot (\Delta_\delta \mid \Theta_\gamma)$

Proof. We look at three cases:

(i) First suppose Δ, Θ are single point distributions $\overline{s}, \overline{t}$ respectively.

In this case $\overline{s} \xrightarrow{\beta} \Delta_\beta$ means that in the underlying MA

- $s \overset{\beta_i}{\longmapsto} s_i$ for $i \in I$, where I is a non-empty family
- $\beta = \sum_{i \in I} \beta_i$
- $\Delta_\beta = \sum_{i \in I} \frac{\beta_i}{\beta} \cdot \overline{s_i}$.

Note that in the notation developed above, Δ_β coincides with Δ_w^s. Moreover Δ_δ coincides with $\overline{s_\delta}$.

Similarly Θ_δ is $\overline{t_\delta}$ and we have

- $t \overset{\gamma_j}{\mapsto} t_j$ for j ranging over some non-empty family J
- $\gamma = \sum_{j \in J} \gamma_j$
- $\Theta_\gamma = \Delta^t_w = \sum_{j \in J} \frac{\gamma_j}{\gamma} \cdot \overline{t_j}$

Note that, still in the MA, this means

$$s \mid t \overset{\beta_i}{\mapsto} s_i \mid t_\delta \qquad\qquad s \mid t \overset{\gamma_j}{\mapsto} s_\delta \mid t_j \qquad \text{for all } i \in I,\ j \in J$$

Now in the derived MLTS these give rise to the weighted action

$$\overline{s} \mid \overline{t} \xrightarrow{(\beta+\gamma)} \sum_{i \in I} \frac{\beta_i}{(\beta+\gamma)} \cdot \overline{s_i \mid t_\delta} \ + \ \sum_{j \in J} \frac{\gamma_j}{(\beta+\gamma)} \cdot \overline{s_\delta \mid t_j}$$

But the sum $\sum_{i \in I} \frac{\beta_i}{(\beta+\gamma)} \cdot \overline{s_i \mid t_\delta}$ can be rewritten as $(\sum_{i \in I} \frac{\beta_i}{(\beta+\gamma)} \cdot \overline{s_i}) \mid \overline{t_\delta}$, which coincides with $\frac{\beta}{(\beta+\gamma)} \cdot (\Delta^s_w \mid \overline{t_\delta})$.

Similarly the sum $\sum_{j \in J} \frac{\gamma_j}{(\beta+\gamma)} \cdot \overline{s_\delta \mid t_j}$ can be rewritten as $\frac{\gamma}{(\beta+\gamma)} \cdot (\overline{s_\delta} \mid \Delta^t_w)$ and therefore we get the required move

$$\overline{s} \mid \overline{t} \xrightarrow{(\beta+\gamma)} \frac{\beta}{(\beta+\gamma)} \cdot (\Delta^s_w \mid \overline{t_\delta}) \ + \ \frac{\gamma}{(\beta+\gamma)} \cdot (\overline{s_\delta} \mid \Delta^t_w) \tag{2}$$

(ii) Let us generalise this to the case of an arbitrary Δ, but where Θ is still the one point distribution $\overline{t}$.

From Lemma 3.1 $\Delta \mid \overline{t} = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot (s \mid t)$. Applying (2) above, by linearity we get the move

$$\Delta \mid \overline{t} \xrightarrow{(\beta+\gamma)} \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \frac{\beta}{(\beta+\gamma)} \cdot (\Delta^s_w \mid \overline{t_\delta}) \ + \ \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \frac{\gamma}{(\beta+\gamma)} \cdot (\overline{s_\delta} \mid \Delta^t_w)$$

In this case note that, since weighted moves are deterministic, $\Delta_\beta = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Delta^s_w$ and $\Delta_\delta = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \overline{s_\delta}$. Therefore the first sum in this residual can be rewritten as $\frac{\beta}{(\beta+\gamma)} \cdot (\Delta_\beta \mid \overline{t_\delta})$ and the second as $\frac{\gamma}{(\beta+\gamma)} \cdot (\Delta_\delta \mid \Delta^t_w)$ and so we have the required move

$$\Delta \mid \overline{t} \xrightarrow{(\beta+\gamma)} \frac{\beta}{(\beta+\gamma)} \cdot (\Delta_\beta \mid \overline{t_\delta}) \ + \ \frac{\gamma}{(\beta+\gamma)} \cdot (\Delta_\delta \mid \Delta^t_w) \tag{3}$$

(iii) Let us finally consider an arbitrary Δ and Θ. Again we use Lemma 3.1; this time part (ii) gives $\Delta \mid \Theta = \sum_{t\in\lceil\Theta\rceil} \Theta(t)\cdot(\Delta \mid \bar{t})$. Again linearity and (3) above gives us the move

$$\Delta \mid \Theta \xrightarrow{(\beta+\gamma)} \sum_{t\in\lceil\Theta\rceil} \Theta(t)\cdot\frac{\beta}{(\beta+\gamma)}\cdot(\Delta_\beta \mid \overline{t_\delta}) \;+\; \sum_{t\in\lceil\Theta\rceil} \Theta(t)\cdot\frac{\gamma}{(\beta+\gamma)}\cdot(\Delta_\delta \mid \Delta^t_w)$$

Now using the fact that $\Theta_\delta = \sum_{t\in\lceil\Theta\rceil} \Theta(t)\cdot\overline{t_\delta}$ and Θ_γ must be $\sum_{t\in\lceil\Theta\rceil} \Theta(t)\cdot\Delta^t_w$ this can be rewritten into the required move:

$$\Delta \mid \Theta \xrightarrow{(\beta+\gamma)} \frac{\beta}{\beta+\gamma}\cdot(\Delta_\beta \mid \Theta_\delta) + \frac{\gamma}{\beta+\gamma}\cdot(\Delta_\delta \mid \Theta_\gamma)$$

□

Theorem 3.9. [Compositionality of $\approx_{sbis}$] Let s,t be states and Θ a distribution in an arbitrary MA, if $s \approx_{sbis} \Theta$ then $s \mid t \approx_{sbis} \Theta \mid \bar{t}$.

Proof. We construct the following relation

$$\mathcal{R} = \{(s \mid t, \Theta \mid \bar{t}) \mid s \approx_{sbis} \Theta\}$$

and check that $\mathcal{R} \subseteq \mathcal{SB}(\approx_{sbis})$ in the associated MLTS. This will imply that $\mathcal{R} \subseteq \approx_{sbis}$, from which the result follows. Note that by construction we have that

(a) Δ_1 *lift*$(\approx_{sbis})$ Δ_2 implies $(\Delta_1 \mid \Theta)$ *lift*$(\mathcal{R})$ $(\Delta_2 \mid \Theta)$ for any distribution Θ

We use this property throughout the proof.

Let $(s \mid t,\ \Theta \mid \bar{t}) \in \mathcal{R}$. We first prove property (ii) in Definition 2.14, which turns out to be straightforward. Since $s \approx_{sbis} \Theta$, there is some Δ such that $\bar{s} \overset{\tau}{\Longrightarrow} \Delta$ and Θ *lift*$(\approx_{sbis})$ Δ. An application of Lemma 3.7(ii) gives $\bar{s} \mid \bar{t} \overset{\tau}{\Longrightarrow} \Delta \mid \bar{t}$ and property (a) that $(\Theta \mid \bar{t})$ *lift*$(\mathcal{R})$ $(\Delta \mid \bar{t})$.

Let us concentrate on property (i): we must prove that every move from $s \mid t$ in the derived MLTS has a matching move from $\Theta \mid \bar{t}$. The first possibility is that $s \mid t \xrightarrow{\mu} \Gamma$ with $\mu \in \mathsf{Act}_\tau$; the matching move from $\Theta \mid t$ depends on the derivation of the move $s \mid t \xrightarrow{\mu} \Gamma$ from the rules in Figure 7.

- Suppose Γ is $\Delta' \mid \bar{t}$, where $s \xrightarrow{\mu} \Delta'$. Here we have $\Theta \overset{\mu}{\Longrightarrow} \Theta'$ such that Δ' *lift*$(\approx_{sbis})$ Θ', since $s \approx_{sbis} \Theta$. Moreover by Lemma 3.7(ii), we can deduce $\Theta \mid \bar{t} \overset{\mu}{\Longrightarrow} \Theta' \mid \bar{t}$. Again by (a) we have $(\Delta' \mid \bar{t},\ \Theta' \mid \bar{t}) \in$ *lift*$(\mathcal{R})$, and therefore a matching move.

- Suppose Γ is $\overline{s} \mid \Delta'$ where $t \xrightarrow{\mu} \Delta'$. Here a symmetric version of Lemma 3.6(i) gives $\Theta \mid \overline{t} \xrightarrow{\mu} \Theta \mid \Delta'$. This is the required matching move since we can use (a) above to deduce $(\overline{s} \mid \Delta', \Theta \mid \Delta') \in \mathit{lift}(\mathcal{R})$.

- The final possibility for μ is τ and Γ is $(\Delta_1 \mid \Delta_2)$ where $s \xrightarrow{a} \Delta_1$ and $t \xrightarrow{\overline{a}} \Delta_2$ for some $a \in \mathsf{Act}$. Here, since $s \approx_{sbis} \Theta$, we have a move $\Theta \stackrel{a}{\Longrightarrow} \Theta'$ such that $\Delta_1 \ \mathit{lift}(\approx_{sbis}) \ \Theta'$. By combining these moves using part (iii) of Lemma 3.7 we obtain $\Theta \mid \overline{t} \stackrel{\tau}{\Longrightarrow} \Theta' \mid \Delta_2$. Again this is the required matching move since an application of (a) above gives $(\Delta_1 \mid \Delta_2, \Theta' \mid \Delta_2) \in \mathit{lift}(\mathcal{R})$.

Now suppose $s \mid t \xrightarrow{\delta} \Gamma$ in the MLTS. In this case it must be that $\Gamma = \Delta_\delta \mid \Theta_\delta$ where these two distributions are the unique ones such that $s \xrightarrow{\delta} \Delta_s$ and $t \xrightarrow{\delta} \Theta_\delta$. This case is very similar to the previous one, but using part (iv) of Lemma 3.7 rather than (iii).

The final possible move, the most complicated case, is $s \mid r \xrightarrow{\lambda} \Gamma$ for some $\lambda \in \mathbb{R}^+$. This move in the derived MLTS must be because in the underlying MA there are a non-empty set of timed transitions from $s \mid t$ whose residuals combine in the derived MLTS to form Γ. These individual timed transitions come in two forms, depending on whether they are inferred by the rule (PAR.L.T) or (PAR.R.T) from Figure 7. Hence, there are two disjoint index sets I_1, I_2 such that

(i) For each $i \in I_1$ we have $s \stackrel{\beta_i}{\mapsto} s_i$, and $t \stackrel{\delta}{\mapsto} t_\delta$

(ii) For each $i \in I_2$ we have $t \stackrel{\gamma_i}{\mapsto} t_i$, and $s \stackrel{\delta}{\mapsto} s_\delta$

(iii) $\Gamma = \sum_{i \in I_1} \frac{\beta_i}{\beta+\gamma} \cdot (\overline{s_i} \mid \overline{t_\delta}) + \sum_{i \in I_2} \frac{\gamma_i}{\beta+\gamma} \cdot (\overline{s_\delta} \mid \overline{t_i})$ where $\beta = \sum_{i \in I_1} \beta_i$, $\gamma = \sum_{i \in I_2} \gamma_i$ and $\lambda = \beta + \gamma$.

In the following argument we assume that both index sets I_1 and I_2 are non-empty; when either are empty the reasoning is simpler and is omitted.

Before proceeding let us first reorganise Γ so that it is expressed in terms of the derived weighted actions from s and t. First note that $s \xrightarrow{\beta} \Delta_w$, where Δ_w denotes $\sum_{i \in I_1} \frac{\beta_i}{\beta} \cdot \overline{s_i}$ and $t \xrightarrow{\gamma} \Gamma_w$ where Γ_w is $\sum_{i \in I_2} \frac{\gamma_i}{\gamma} \cdot \overline{t_i}$. Then with some simple reorganisation we can see that

$$\Gamma = \frac{\beta}{\beta+\gamma} \cdot (\Delta_w \mid \overline{t_\delta}) + \frac{\gamma}{\beta+\gamma} \cdot (\overline{s_\delta} \mid \Gamma_w)$$

We have to find a weak move $\Theta \mid \overline{t} \stackrel{(\beta+\gamma)}{\Longrightarrow} \Gamma'$ such that $\Gamma \ \mathit{lift}(\mathcal{R}) \ \Gamma'$.

First consider the move $s \xrightarrow{\beta} \Delta_w$. Since $s \approx_{sbis} \Theta$ there is a matching move of the form $\Theta \Longrightarrow \Theta_d \xrightarrow{\beta} \Theta_\beta \Longrightarrow \Theta_w$ such that $\Delta_w \ \mathit{lift}(\approx_{sbis}) \ \Theta_w$.

Now, since $s \not\xrightarrow{\tau}$, an application of Corollary 2.23 gives us that $s \approx_{sbis} \Theta_d$. Hence, we also have a matching move from Θ_d for the move $s \xrightarrow{\delta} \overline{s_\delta}$. Moreover, since $\Theta_d \xrightarrow{\beta}$ we know by maximal progress that $\Theta_d \not\xrightarrow{\tau}$. Therefore, the matching move must take the form $\Theta_d \xrightarrow{\delta} \Theta_\delta \Longrightarrow \Theta'$ such that $\overline{s_\delta}$ *lift*$(\approx_{sbis})$ Θ'.

We are now ready to construct the required matching move:

$$\begin{array}{rcl} \Theta \mid \bar{t} & \Longrightarrow & \Theta_d \mid \bar{t} \\ & \xrightarrow{(\beta+\gamma)} & \dfrac{\beta}{\beta+\gamma} \cdot (\Theta_\beta \mid \overline{t_\delta}) + \dfrac{\gamma}{\beta+\gamma} \cdot (\Theta_\delta \mid \Gamma_w) \\ & \Longrightarrow & \dfrac{\beta}{\beta+\gamma} \cdot (\Theta_w \mid \overline{t_\delta}) + \dfrac{\gamma}{\beta+\gamma} \cdot (\Theta' \mid \Gamma_w) \end{array}$$

Here the second move is an application of Theorem 3.8 and the third an application of Lemma 3.7 (i) and the linearity of $\Longrightarrow$.

Letting Γ' denote this residual, it follows from property (a) above that Γ *lift*$(\mathcal{R})$ Γ', since Δ_w *lift*$(\approx_{sbis})$ Θ_w and $\overline{s_\delta}$ *lift*$(\approx_{sbis})$ Θ'. □

Corollary 3.10. In an arbitrary MA, Δ *lift*$(\approx_{sbis})$ Θ implies $(\Delta \mid \Gamma)$ *lift*$(\approx_{sbis})$ $(\Theta \mid \Gamma)$

Proof. A simple consequence of the previous compositionality result, using a straightforward linearity argument. □

Theorem 3.11. [Compositionality of $\approx_{bis}$] Let Δ, Θ and Γ be any distributions in a finitary par-closed MA. If $\Delta \approx_{bis} \Theta$ then $\Delta \mid \Gamma \approx_{bis} \Theta \mid \Gamma$.

Proof. We show that the relation

$$\mathcal{R} = \{(\Delta \mid \Gamma, \Theta \mid \Gamma) \mid \Delta \approx_{bis} \Theta\} \cup \approx_{bis}$$

is a bisimulation, from which the result follows.

Suppose $(\Delta \mid \Gamma, \Theta \mid \Gamma) \in \mathcal{R}$. Since $\Delta \approx_{bis} \Theta$, we know from Theorem 2.16 that some Θ' exists such that $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$ and Δ *lift*$(\approx_{sbis})$ Θ' and the previous corollary implies that $(\Delta \mid \Gamma)$ *lift*$(\approx_{sbis})$ $(\Theta' \mid \Gamma)$; by Theorem 2.16 this gives $(\Delta \mid \Gamma) \approx_{bis} (\Theta' \mid \Gamma)$.

We now show that $\mathcal{R} \subseteq \mathcal{B}(\mathcal{R})$. Consider the actions from $(\Delta \mid \Gamma)$ and $(\Theta \mid \Gamma)$; by symmetry it is sufficient to show that the actions of the former can be matched by the latter. Suppose that $(\Delta \mid \Gamma) \stackrel{\mu}{\Longrightarrow} (\sum_i p_i \cdot \Delta'_i)$. Then $(\Theta' \mid \Gamma) \stackrel{\mu}{\Longrightarrow} (\sum_i p_i \cdot \Theta'_i)$ with $\Delta'_i \approx_{bis} \Theta'_i$ for each i. But by part (i) of Lemma 3.7 $(\Theta \mid \Gamma) \stackrel{\tau}{\Longrightarrow} (\Theta' \mid \Gamma)$ and therefore we have the required matching move $(\Theta \mid \Gamma) \stackrel{\mu}{\Longrightarrow} (\sum_i p_i \cdot \Theta'_i)$. □

A particular application of this compositionality result is that bisimulation equivalence is preserved by the parallel operator | in the language mCCS. As expected it is not preserved by the choice operator; the standard example from CCS applies: $\overline{\tau.a.\mathbf{0}} \approx_{bis} \overline{a.\mathbf{0}}$ but $\overline{b.\mathbf{0}+\tau.a.\mathbf{0}} \not\approx_{bis} \overline{b.\mathbf{0}+a.\mathbf{0}}$. Recall that terms are interpreted as states in the MA for mCCS and therefore these results are expressed using point distributions.

However, bisimulation is preserved by all the other operators.

Proposition 3.12.

(i) $\overline{P} \approx_{bis} \overline{Q}$ implies $\overline{\lambda.P} \approx_{bis} \overline{\lambda.Q}$ and $\delta.P \approx_{bis} \overline{\delta.Q}$

(ii) $[\![D]\!] \approx_{bis} [\![E]\!]$ implies $\overline{\mu{:}D} \approx_{bis} \overline{\mu{:}E}$.

Proof. Straightforward, by first proving the corresponding results for $\approx_{sbis}$. We outline one example.

Let $\mathcal{R}= \{\, (\lambda.P, \overline{\lambda.Q}) \mid \overline{P} \approx_{bis} \overline{Q} \,\}$. Recall from Proposition 2.22 that if $\overline{P} \approx_{bis} \overline{Q}$ then there is some Θ' such that $\overline{Q} \stackrel{\tau}{\Longrightarrow} \Theta'$ and $\overline{P}$ *lift*$(\approx_{sbis})$ Θ'. With this remark we can show that $\mathcal{R} \cup \approx_{sbis}$ is a simple bisimulation.

- Consider the strong move from $\lambda.P$ in the derived MLTS, $\lambda.P \xrightarrow{\lambda} \overline{P}$; this can be matched by $\overline{\lambda.Q} \stackrel{\lambda}{\Longrightarrow} \Theta'$.
- The move $\lambda.P \xrightarrow{\delta} \overline{\lambda.P}$ is matched by $\lambda.Q \xrightarrow{\delta} \overline{\lambda.Q}$, as $\overline{\lambda.P}$ *lift*$(\mathcal{R})$ $\overline{\lambda.Q}$ holds by definition.
- Also $\lambda.P \stackrel{\tau}{\Longrightarrow} \overline{\lambda.P}$ via the empty move and by definition $\overline{\lambda.Q}$ *lift*$(\mathcal{R})$ $\overline{\lambda.P}$

It follows that if $\overline{P} \approx_{bis} \overline{Q}$ then $\lambda.P \approx_{sbis} \overline{\lambda.Q}$. But now an application of Theorem 2.20 gives the required $\overline{\lambda.P} \approx_{bis} \overline{\lambda.Q}$. □

4 Soundness and completeness

Consider an arbitrary par-closed MA $M = \langle S, \mathsf{Act}_\tau, \rightarrow, \mapsto \rangle$. Experimenting on processes in M consists in observing what communications a process can perform, as it evolves by both internal moves and the passage of time. We make this precise in the following definition:

Definition 4.1. [Evolution] Let $\Delta \Longrightarrow \Delta'$ be the least reflexive relation satisfying:
(a) $\Delta \Longrightarrow \Delta_1$ and $\Delta_1 \stackrel{\tau}{\Longrightarrow} \Delta'$ implies $\Delta \Longrightarrow \Delta'$
(b) $\Delta \Longrightarrow \Delta_1$ and $\Delta_1 \xrightarrow{\lambda} \Delta'$ implies $\Delta \Longrightarrow \Delta'$, where $\lambda \in \mathbb{R}^+$

(c) $\Delta_i \Longrightarrow \Delta_i'$ for each $i \in I$, where I is a finite index set, implies $(\sum_{i\in I} p_i \cdot \Delta_i) \Longrightarrow (\sum_{i\in I} p_i \cdot \Delta_i')$ for any $\sum_{i\in I} p_i = 1$. □

Thus $\Delta \Longrightarrow \Delta'$ is a relation between distributions in a MA which allows reduction either by internal actions τ or definite delay actions λ; with the latter the reductions are to distributions determined by the rates of the states in the support of Δ. But it is important to note that the passage of time is not recorded directly, and indeed passes at independent speeds in the individual states in the support of Δ. For example, suppose Δ is $\frac{1}{2} \cdot \Delta_1 + \frac{1}{2} \cdot \Delta_2$ where Δ_1, Δ_2 are $2\lambda.(\frac{1}{4} \cdot \overline{b} + \frac{3}{4} \cdot \overline{a})$ and $4\lambda.(\frac{1}{2} \cdot \overline{b} + \frac{1}{2} \cdot \overline{c})$. Then $\Delta \Longrightarrow \frac{3}{8} \cdot \overline{b} + \frac{3}{8} \cdot \overline{a} + \frac{1}{4} \cdot \overline{c}$, despite the fact that in Δ_2 time passes at twice the rate as in Δ_1.

Definition 4.2. [Barbs] For $\Delta \in \mathcal{D}(S)$ and $a \in \mathsf{Act}$ let $\mathcal{V}_a(\Delta) = \sum\{\, \Delta(s) \mid s \xrightarrow{a} \}$. We write $\Delta \Downarrow_a^{\geq p}$ whenever $\Delta \Longrightarrow \Delta'$, where $\mathcal{V}_a(\Delta') \geq p$. We also we use the notation $P \not\Downarrow_a^{>0}$ to mean that $P \Downarrow_a^{\geq p}$ does not hold for any $p > 0$. □

Then we say a relation $\mathcal{R}$ is *barb-preserving* if $\Delta \Downarrow_a^{\geq p}$ iff $\Theta \Downarrow_a^{\geq p}$ whenever $\Delta \mathrel{\mathcal{R}} \Theta$. It is *reduction-closed* if $\Delta \mathrel{\mathcal{R}} \Theta$ implies

(i) whenever $\Delta \Longrightarrow \Delta'$, there is a $\Theta \Longrightarrow \Theta'$ such that $\Delta' \mathrel{\mathcal{R}} \Theta'$
(ii) whenever $\Theta \Longrightarrow \Theta'$, there is a $\Delta \Longrightarrow \Delta'$ such that $\Delta' \mathrel{\mathcal{R}} \Theta'$.

Finally, we say that in a par-closed MA $\mathcal{R}$ is *compositional* if $\Delta_1 \mathrel{\mathcal{R}} \Delta_2$ implies $(\Delta_1 \mid \Theta) \mathrel{\mathcal{R}} (\Delta_2 \mid \Theta)$ for every distribution Θ.

Definition 4.3. In a par-closed MA, let $\approx_{rbc}$ be the largest relation over the states which is barb-preserving, reduction-closed and compositional. □

Example 4.4. Consider the two processes $P_1 = \lambda_1.Q_1$ and $P_2 = \lambda_2.Q_2$ where $\lambda_1 < \lambda_2$ and Q_i are two arbitrary processes. We can show that $P_1 \not\approx_{rbc} P_2$ by exhibiting a *testing* process T such that the barbs of $(P_1 \mid T)$ and $(P_2 \mid T)$ are different. For example, let $T = \delta.\tau.\mathbf{0} + \lambda_1.succ$. In $(P_i \mid T)$ there is a race between two timed events; in $(P_2 \mid T)$ their rates are λ_1 versus λ_2 while in $(P_1 \mid T)$ both events have the same rate. If the timed event in the test wins out, the action *succ* will occur. Consequently $(P_1 \mid T) \Downarrow_{succ}^{\geq\frac{1}{2}}$. However, $(P_2 \mid T)$ does not have this barb; instead $(P_2 \mid T) \Downarrow_{succ}^{\geq q}$, where $q = \frac{\lambda_1}{\lambda_1+\lambda_2}$; q is strictly smaller than $\frac{1}{2}$ since $\lambda_1 < \lambda_2$.

It follows, by a suitable instantiation of Q_1, Q_2, that $\lambda_1.\lambda_2.P \not\approx_{rbc} \lambda_2.\lambda_1.P$ when λ_1 and λ_2 are different. □

Example 4.5. Consider the processes $P_1 = a{:}Q$, $P_2 = a.Q$, and $P_3 = \lambda.P_2$, where Q is an arbitrary process, and we have seen that $a.Q$ is shorthand for a recursively defined process $A \Leftarrow a{:}Q + \delta.A$.

Note that according to our semantics P_1 does not let time pass. Let T be the testing process $\lambda.(\bar{a}.succ + \tau.\mathbf{0})$. The process $P_1 \mid T$ cannot evolve, thus $(P_1 \mid T) \not\Downarrow^{>0}_{succ}$. However, we have $P_2 \mid T \overset{\lambda}{\mapsto} P_2 \mid (\bar{a}.succ + \tau.\mathbf{0}) \xrightarrow{\tau} Q \mid succ$, thus $(P_2 \mid T) \Downarrow^{\geq 1}_{succ}$. The only comparable barb for P_3 is $(P_3 \mid T) \Downarrow^{\geq \frac{1}{2}}_{succ}$, because if the timed event in the test takes place, then by maximal progress the τ action must happen before the timed event in the process. It follows that the three processes P_1, P_2 and P_3 can be distinguished. □

Example 4.6. Consider the two MAs s and u from Figure 2, discussed in the Introduction. Let T be the process $\tau.\delta.\bar{a}.succ + \tau.\delta.\bar{b}.succ$ and Δ denote the point distribution $\overline{\mathbf{0} \mid succ}$. Since $s \mid T \Longrightarrow\!\!\blacktriangleright \Delta$, we have $(s \mid T) \Downarrow^{\geq 1}_{succ}$.

However, the weak derivatives of $u \mid T$ under the evolution relation are very few, and one can easily check that none of them will have exactly the barbs of Δ because if $(u \mid T) \Downarrow^{\geq p}_{succ}$ then p is at most $\frac{1}{2}$. It follows that $s \not\approx_{rbc} u$, i.e. s and u are indeed behaviourally different. □

Lemma 4.7. The relation $\Longrightarrow\!\!\blacktriangleright$ over distributions is linear and left-decomposable.

Proof. The relation $\Longrightarrow\!\!\blacktriangleright$ is linear by definition. For left-decomposability, we proceed by rule induction. The relation $\Longrightarrow\!\!\blacktriangleright$ contains the identity relation and is closed under the three rules (a), (b) and (c) in Definition 4.1. As an example, we consider rule (a).

Suppose $(\sum_{i\in I} p_i \cdot \Delta_i) \Longrightarrow\!\!\blacktriangleright \Delta' \overset{\tau}{\Longrightarrow} \Delta''$. By induction, there are distributions Δ'_i such that $\Delta' = \sum_{i\in I} p_i \cdot \Delta'_i$ and $\Delta_i \Longrightarrow\!\!\blacktriangleright \Delta'_i$ for each $i \in I$. By Theorem 2.10 the relation $\overset{\tau}{\Longrightarrow}$ is left-decomposable. Hence, there exist distributions Δ''_i such that $\Delta'' = \sum_{i\in I} p_i \cdot \Delta''_i$ and $\Delta'_i \overset{\tau}{\Longrightarrow} \Delta''_i$ for each $i \in I$. It follows that $\Delta_i \Longrightarrow\!\!\blacktriangleright \Delta''_i$ for each $i \in I$, by using rule (a) in Definition 4.1. □

Proposition 4.8. In an arbitrary MA, $\approx_{bis}$ is reduction-closed.

Proof. Suppose $\Delta \approx_{bis} \Theta$ and $\Delta \Longrightarrow\!\!\blacktriangleright \Delta'$. We have to show that $\Theta \Longrightarrow\!\!\blacktriangleright \Delta'$ such that $\Delta' \approx_{bis} \Theta'$.

The proof is by rule induction on how $\Delta \Longrightarrow\!\!\blacktriangleright \Delta'$ is derived using the rules (a), (b) and (c) in Definition 4.1. The base case is, by reflexivity, when Δ' is Δ, and is trivial. Hence, there are three cases:

(a) Suppose $\Delta \Longrightarrow\!\!\blacktriangleright \Delta'' \overset{\tau}{\Longrightarrow} \Delta'$ for some distribution Δ''. By induction, there is some Θ'' such that $\Theta \Longrightarrow\!\!\blacktriangleright \Theta''$ and $\Delta'' \approx_{bis} \Theta''$. The latter implies the existence of some Θ' such that $\Theta'' \overset{\tau}{\Longrightarrow} \Theta'$, which yields $\Theta \Longrightarrow\!\!\blacktriangleright \Theta'$, and $\Delta' \approx_{bis} \Theta'$.

(b) Suppose $\Delta \Longrightarrow\!\!\blacktriangleright \Delta'' \xrightarrow{\lambda} \Delta'$ for some distribution Δ''. As in the last case, it can be shown that there exists some Θ' such that $\Theta \Longrightarrow\!\!\blacktriangleright \Theta'$ and $\Delta' \approx_{bis} \Theta'$.

(c) Suppose $\Delta = \sum_{i\in I} p_i \cdot \Delta_i$, $\Delta' = \sum_{i\in I} p_i \cdot \Delta'_i$, and $\Delta_i \Longrightarrow\!\!\!\blacktriangleright \Delta'_i$ for each i in a finite index set I, where $\sum_{i\in I} p_i = 1$. It follows from $\Delta \approx_{bis} \Theta$ and $\Delta \stackrel{\tau}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta_i$ that $\Theta \stackrel{\tau}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta_i$ for some Θ_i with $\Delta_i \approx_{bis} \Theta_i$. By induction, there exists Θ'_i with $\Theta_i \Longrightarrow\!\!\!\blacktriangleright \Theta'_i$ and $\Delta'_i \approx_{bis} \Theta'_i$ for each $i \in I$. Let $\Theta' = \sum_{i\in I} p_i \cdot \Theta'_i$. By Lemma 4.7, the relation $\Longrightarrow\!\!\!\blacktriangleright$ is linear. Then we have $(\sum_{i\in I} p_i \cdot \Theta_i) \Longrightarrow\!\!\!\blacktriangleright \Theta'$, thus $\Theta \Longrightarrow\!\!\!\blacktriangleright \Theta'$. By the linearity of $\approx_{bis}$, Corollary 2.24, we also have $\Delta' \approx_{bis} \Theta'$.

$\square$

Theorem 4.9. [Soundness] In a finitary par-closed MA, if $\Delta \approx_{bis} \Theta$ then $\Delta \approx_{rbc} \Theta$.

Proof. Because of Theorem 3.11 and the previous proposition it is sufficient to prove that $\approx_{bis}$ is barb-preserving.

Suppose $\Delta \approx_{bis} \Theta$ and $\Delta \Downarrow_a^{\geq p}$, for any action a and a probability p; we need to show that $\Theta \Downarrow_a^{\geq p}$. We see from $\Delta \Downarrow_a^{\geq p}$ that $\Delta \Longrightarrow\!\!\!\blacktriangleright \Delta'$ for some Δ' with $\mathcal{V}_a(\Delta') \geq p$. By Proposition 4.8, the relation $\approx_{bis}$ is reduction-closed. Hence, there exists Θ' such that $\Theta \Longrightarrow\!\!\!\blacktriangleright \Theta'$ and $\Delta' \approx_{bis} \Theta'$. The degenerate weak transition $\Delta' \stackrel{\tau}{\Longrightarrow} \sum_{s\in\lceil\Delta'\rceil} \Delta'(s) \cdot \overline{s}$ must be matched by some transition

$$\Theta' \stackrel{\tau}{\Longrightarrow} \sum_{s\in\lceil\Delta'\rceil} \Delta'(s) \cdot \Theta'_s \tag{4}$$

such that $\overline{s} \approx_{bis} \Theta'_s$. By Proposition 2.21 we know that $s \approx_{sbis} \Theta'_s$ for each $s \in \lceil\Delta'\rceil$. Now if $s \stackrel{a}{\longrightarrow}$, then $\Theta'_s \stackrel{a}{\Longrightarrow}$, that is $\Theta'_s \stackrel{\tau}{\Longrightarrow} \Theta''_s \stackrel{a}{\longrightarrow}$ for some distribution Θ''_s. Let S_a be the set of states $\{s \in \lceil\Delta'\rceil \mid s \stackrel{a}{\longrightarrow}\}$, and Θ'' be the distribution

$$(\sum_{s\in S_a} \Delta'(s) \cdot \Theta''_s) + (\sum_{s\in\lceil\Delta'\rceil\setminus S_a} \Delta'(s) \cdot \Theta'_s).$$

By the linearity and reflexivity of $\stackrel{\tau}{\Longrightarrow}$, Theorem 2.10, we have

$$(\sum_{s\in\lceil\Delta'\rceil} \Delta'(s) \cdot \Theta'_s) \stackrel{\tau}{\Longrightarrow} \Theta'' \tag{5}$$

By (4), (5) and the transitivity of $\stackrel{\tau}{\Longrightarrow}$, we obtain $\Theta' \stackrel{\tau}{\Longrightarrow} \Theta''$, thus $\Theta \Longrightarrow\!\!\!\blacktriangleright \Theta''$. It remains to show that $\mathcal{V}_a(\Theta'') \geq p$.

Note that for each $s \in S_a$ we have $\Theta''_s \stackrel{a}{\longrightarrow}$, which means that $\mathcal{V}_a(\Theta''_s) = 1$. It follows that

$$\begin{array}{rcl}
\mathcal{V}_a(\Theta'') & = & \sum_{s\in S_a} \Delta'(s) \cdot \mathcal{V}_a(\Theta''_s) + \sum_{s\in\lceil\Delta'\rceil\setminus S_a} \Delta'(s) \cdot \mathcal{V}_a(\Theta'_s) \\
& \geq & \sum_{s\in S_a} \Delta'(s) \cdot \mathcal{V}_a(\Theta''_s) \\
& = & \sum_{s\in S_a} \Delta'(s) \\
& = & \mathcal{V}_a(\Delta') \\
& \geq & p
\end{array}$$

$\square$

In order to establish a converse to Theorem 4.9, *completeness*, we need to work in a MA which is expressive enough to provide appropriate contexts and barbs in order to distinguish processes which are not bisimilar. For this purpose we use the MA determined by the language mCCS in the previous section. For the remainder of this section we focus on this particular MA.

Lemma 4.10. In mCCS, if $s \xrightarrow{\lambda}$ for any weight $\lambda \in \mathbb{R}^+$ then $s \xrightarrow{\delta}$.

Proof. A straightforward induction on the derivation of $s \xrightarrow{\lambda} \Delta$ from the rules in Figures 7 and 8. □

We will eventually establish the completeness by showing that $\approx_{rbc}$ is a bisimulation, but this requires that we first develop a series of auxiliary properties of $\approx_{rbc}$ in this setting. The technique used normally involves examining the barbs of processes in certain contexts; the following lemma gives extra power to this technique. Here, as in the remainder of the paper, we abbreviate the process $c{:}\mathbf{0}$ to c, for any action name c.

Lemma 4.11. In mCCS suppose $(\Delta \mid succ)\,{}_p\!\oplus \Delta' \approx_{rbc} (\Theta \mid succ)\,{}_p\!\oplus \Theta'$ where $p > 0$ and *succ* is fresh name. Then $\Delta \approx_{rbc} \Theta$.

Proof. Consider the relation

$$\mathcal{R} = \{(\Delta, \Theta) \mid (\Delta \mid c)\,{}_p\!\oplus \Delta' \approx_{rbc} (\Theta \mid c)\,{}_p\!\oplus \Theta' \text{ for some } \Delta', \Theta' \text{ and fresh } c\}$$

We show that $\mathcal{R} \subseteq \approx_{rbc}$, by showing that $\mathcal{R}$ satisfies the three defining properties of $\approx_{rbc}$.

(1) $\mathcal{R}$ is compositional. Suppose $\Delta \mathrel{\mathcal{R}} \Theta$; we have to show that $\Delta \mid \Phi \mathrel{\mathcal{R}} \Theta \mid \Phi$, for any distribution Φ. Since $\Delta \mathrel{\mathcal{R}} \Theta$ there are some Δ', Θ' and fresh c such that

$$\Lambda \approx_{rbc} \Gamma \text{ where } \Lambda = (\Delta \mid c)\,{}_p\!\oplus \Delta',\ \Gamma = (\Theta \mid c)\,{}_p\!\oplus \Theta' \qquad (6)$$

Since since $\approx_{rbc}$ is compositional, we have $\Lambda \mid \Phi \approx_{rbc} \Gamma \mid \Phi$. Therefore, $(\Delta|\Phi \mid c)\,{}_p\!\oplus (\Delta' \mid \Phi) \approx_{rbc} (\Theta' \mid \Phi \mid c)\,{}_p\!\oplus (\Theta' \mid \Phi)$, which means, by definition, that $(\Delta \mid \Phi) \mathrel{\mathcal{R}} (\Theta \mid \Phi)$.

(2) $\mathcal{R}$ is barb-preserving. Suppose $\Delta \Downarrow_a^{\geq q}$ for some action a and probability q, where $\Delta \mathrel{\mathcal{R}} \Theta$. Again we may assume (6) above. Consider the testing process $\bar{a}.\bar{c}.b$, where b is fresh. Since $\approx_{rbc}$ is compositional, we have $(\Lambda \mid \bar{a}.\bar{c}.b) \approx_{rbc} (\Gamma \mid \bar{a}.\bar{c}.b)$. Note that $(\Lambda \mid \bar{a}.\bar{c}.b) \Downarrow_b^{\geq pq}$, which implies $(\Gamma \mid \bar{a}.\bar{c}.b) \Downarrow_b^{\geq pq}$. Since c is fresh for Θ', the latter has no potential to enable the action c, and thus $\Theta' \mid \bar{a}.\bar{c}.b$ is not able to fire the action b. Therefore, it must be the case that $(\Theta \mid c \mid \bar{a}.\bar{c}.b) \Downarrow_b^{\geq q}$, which implies $\Theta \Downarrow_a^{\geq q}$.

(3) $\mathcal{R}$ is reduction-closed. There are three steps to this proof.

(i) We first show that $\mathcal{R}$ is closed with respect to $\stackrel{\tau}{\Longrightarrow}$. Suppose $\Delta\ \mathcal{R}\ \Theta$ and $\Delta \stackrel{\tau}{\Longrightarrow} \Delta''$ for some distribution Δ''. Let Γ and Λ be determined as in (6) above. Then $\Gamma \stackrel{\tau}{\Longrightarrow} (\Delta'' \mid c)\,{}_p{\oplus}\,\Delta'$. Since $\Lambda \approx_{rbc} \Gamma$, there is some Γ' such that $\Gamma \Longrightarrow \Gamma'$ and $(\Delta'' \mid c)\,{}_p{\oplus}\,\Delta' \approx_{rbc} \Gamma'$. Since the component $\Theta \mid c$ cannot enable any weighted action, neither can Γ. Then it must be the case that $\Gamma \stackrel{\tau}{\Longrightarrow} \Gamma'$. In other words, there are some Θ'', Θ''' such that $\Gamma' \equiv (\Theta'' \mid c)\,{}_p{\oplus}\,\Theta'''$ with $\Theta \stackrel{\tau}{\Longrightarrow} \Theta''$ and $\Theta' \stackrel{\tau}{\Longrightarrow} \Theta'''$. Thus $(\Delta'' \mid c)\,{}_p{\oplus}\,\Delta' \approx_{rbc} (\Theta'' \mid c)\,{}_p{\oplus}\,\Theta'''$. Thus by definition $\Delta''\ \mathcal{R}\ \Theta''$.

(ii) Next, we show that $\mathcal{R}$ is closed with respect to weighted actions, in the following sense. Suppose $\Delta\ \mathcal{R}\ \Theta$ and $\Delta \stackrel{\lambda}{\longrightarrow} \Delta''$ for some rate λ and distribution Δ''. We prove that $\Theta \stackrel{\lambda'}{\Longrightarrow} \Theta''$ for some *arbitrary* weight λ' and distribution Θ'' such that $\Delta''\ \mathcal{R}\ \Theta''$.

Again we use the notation from (6) above. Since $\Lambda \approx_{rbc} \Gamma$ and $\approx_{rbc}$ is compositional, we have $\Lambda \mid T \ \approx_{rbc}\ \Gamma \mid T$ where T is the testing process $\bar{c}.(\mathit{fail} + \delta.c')$ for some fresh actions fail, c'. The transition

$$\Lambda \mid T \Longrightarrow \Delta'' \mid c'\,{}_p{\oplus}\,\Delta' \mid T$$

must be matched by some transition $(\Gamma \mid T) \Longrightarrow \Phi$ with $(\Delta'' \mid c'\,{}_p{\oplus}\,\Delta' \mid T) \approx_{rbc} \Phi$. Since $\approx_{rbc}$ is barb-preserving, we have $\Phi \not\Downarrow_{c}^{>0}$, $\Phi \not\Downarrow_{\mathit{fail}}^{>0}$ and $\Phi \Downarrow_{c'}^{\geq p}$. This can happen only if $\Theta \stackrel{\lambda'}{\Longrightarrow} \Theta''$ for some Θ'' and arbitrary weight λ', as Θ' cannot enable action c so as to fire c'. Thus Φ is in the form $\Theta'' \mid c'\,{}_p{\oplus}\,\Theta' \mid T$, and therefore by definition $\Delta''\ \mathcal{R}\ \Theta''$.

(iii) Finally, using (i) and (ii), we can now show that $\mathcal{R}$ is reduction closed. Suppose $\Delta\ \mathcal{R}\ \Theta$ and $\Delta \Longrightarrow \Delta'$. We can use induction on the proof of this derivation from the rules in Definition 4.1 that this can be matched by a derivation $\Theta \Longrightarrow \Theta'$ such that $\Delta'\ \mathcal{R}\ \Theta'$.

$\square$

Proposition 4.12.

(i) In an arbitrary MLTS the relation $\approx_{rbc}$ is linear.

(ii) (Weak-left-decomposable) In mCCS, if $(\sum_{i\in I} p_i \cdot \Delta_i) \approx_{rbc} \Theta$, where I is a finite index set, then there are some Θ_i such that $\Theta \stackrel{\tau}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta_i$ and $\Delta_i \approx_{rbc} \Theta_i$ for each $i \in I$.

Proof. (i) Let $\mathcal{R}$ be the relation

$$\{((\sum_{i\in I} p_i \cdot \Delta_i), (\sum_{i\in I} p_i \cdot \Theta_i)) \mid \Delta_i \approx_{rbc} \Theta_i \text{ for all } i \in I\}$$

where as usual we assume all index sets I to be finite. We will show that show that it is reduction-closed, barb-preserving and compositional. From this we will have that $\mathcal{R} \subseteq \approx_{rbc}$, and then linearity follows.

- $\mathcal{R}$ is reduction closed. Suppose $\Delta \mathrel{\mathcal{R}} \Theta$ and $\Delta \Longrightarrow \Delta'$. By the construction of $\mathcal{R}$ we know that

$$\Delta = \sum_{i\in I} p_i \cdot \Delta_i, \qquad \Theta = \sum_{i\in I} p_i \cdot \Theta_i, \text{ with } \Delta_i \approx_{rbc} \Theta_i \text{ for each } i \in I \tag{7}$$

  By Lemma 4.7, the relation $\Longrightarrow$ is left-decomposable. That is, there are distributions Δ'_i such that $\Delta' = \sum_{i\in I} p_i \cdot \Delta'_i$ and $\Delta_i \Longrightarrow \Delta'_i$ for each $i \in I$. Since $\Delta_i \approx_{rbc} \Theta_i$, there exists Θ'_i such that $\Theta_i \Longrightarrow \Theta'_i$ and $\Delta'_i \approx_{rbc} \Theta'_i$. Let $\Theta' = \sum_{i\in I} p_i \cdot \Theta'_i$. By Lemma 4.7, the relation $\Longrightarrow$ is linear, which implies $\Theta \Longrightarrow \Theta'$. Moreover, note that $(\Delta', \Theta') \in \mathcal{R}$.

  Evolutions from Θ can be matched by Δ in the same manner.

- $\mathcal{R}$ preserves barbs. Suppose $\Delta \mathrel{\mathcal{R}} \Theta$ and $\Delta \Downarrow_a^{\geq p}$; that is $\Delta \Longrightarrow \Delta'$ for some Δ' with $\mathcal{V}_a(\Delta') \geq p$. We have just shown that $\mathcal{R}$ is reduction closed. Therefore, there is some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta' \mathrel{\mathcal{R}} \Theta'$. The latter means that there are distributions Δ_j, Θ_j and probabilities q_j such that $\Delta' = \sum_{j\in J} q_j \cdot \Delta_j$, $\Theta' = \sum_{j\in J} q_j \cdot \Theta_j$, and $\Delta_j \approx_{rbc} \Theta_j$ for each $j \in J$.

  Clearly, $\Delta_j \Downarrow_a^{\geq \mathcal{V}_a(\Delta_j)}$ holds. Then $\Theta_j \Longrightarrow \Theta'_j$ for some Θ'_j with $\mathcal{V}_a(\Theta'_j) \geq \mathcal{V}_a(\Delta_j)$. It follows that $\Theta' \Longrightarrow \sum_{j\in J} q_j \cdot \Theta'_j$ and

$$\mathcal{V}_a(\sum_{j\in J} q_j \cdot \Theta'_j) \;=\; \sum_{j\in J} q_j \mathcal{V}_a(\Theta'_j) \;\geq\; \sum_{j\in J} q_j \mathcal{V}_a(\Delta_j) \;=\; \mathcal{V}_a(\Delta') \;\geq\; p.$$

  Note that we also have $\Theta \Longrightarrow \sum_{j\in J} q_j \cdot \Theta'_j$ by the transitivity of $\Longrightarrow$ and consequently we obtain the required barb, $\Theta \Downarrow_a^{\geq p}$.

- $\mathcal{R}$ is compositional. Suppose $\Delta \mathrel{\mathcal{R}} \Theta$. We have to show $(\Delta \mid \Gamma) \mathrel{\mathcal{R}} (\Delta \mid \Gamma$ for an arbitrary distribution Γ. Using the notation from (7) above, the compositionality of $\approx_{rbc}$ gives $(\Delta_i \mid \Gamma) \approx_{rbc} (\Theta_i \mid \Gamma)$ for each $i \in I$. Since $\Delta \mid \Gamma = \sum_{i\in I} p_i \cdot (\Delta_i \mid \Gamma)$ and $\Theta \mid \Gamma = \sum_{i\in I} p_i \cdot (\Theta_i \mid \Gamma)$ we therefore have the required $(\Delta \mid \Gamma, \Theta \mid \Gamma) \in \mathcal{R}$.

(ii) Without loss of generality, we assume that $p_i \neq 0$ for all $i \in I$. Suppose that $(\sum_{i \in I} p_i \cdot \Delta_i) \approx_{rbc} \Theta$. Consider the testing process $T = b + \sum_{i \in I} \tau{:}a_i$, where a_i and b are fresh actions. By the compositionality of $\approx_{rbc}$, we have $(\sum_{i \in I} p_i \cdot \Delta_i) \mid T \approx_{rbc} \Theta \mid T$. Now $(\sum_{i \in I} p_i \cdot \Delta_i) \mid T \stackrel{\tau}{\Longrightarrow} \sum_{i \in I} p_i \cdot (\Delta_i \mid a_i)$. Since $\approx_{rbc}$ is reduction-closed, there is some Γ such that $\Theta \mid T \Longrightarrow \Gamma$ and $\sum_{i \in I} p_i \cdot (\Delta_i \mid a_i) \approx_{rbc} \Gamma$. Note that $\Theta \mid T$ cannot enable weighted actions, so we actually have $\Theta \mid T \stackrel{\tau}{\Longrightarrow} \Gamma$.

The barbs of $\sum_{i \in I} p_i \cdot (\Delta_i \mid a_i)$ constrain severely the possible structure of Γ. For example, since $\Gamma \not\Downarrow_b^{>0}$, we have $\Gamma \equiv \sum_{k \in K} q_k \cdot (\Theta_k \mid a_{k_i})$ for some index set K, where $\Theta \stackrel{\tau}{\Longrightarrow} \sum_k q_k \cdot \Theta_k$ and $k_i \in I$. For any indices k_1 and k_2, if $a_{k_1} = a_{k_2}$, we can combine the two components $q_{k_1} \cdot \Theta_{k_1} + q_{k_2} \cdot \Theta_{k_2}$ into one component $(q_{k_1} + q_{k_2}) \cdot \Theta_{k_{12}}$ where $\Theta_{k_{12}} = (\frac{q_{k_1}}{q_{k_1} + q_{k_2}} \cdot \Theta_{k_1} + \frac{q_{k_2}}{q_{k_1} + q_{k_2}} \cdot \Theta_{k_2})$. In this way, we see that Γ can be written as $\sum_{i \in I} q_i \cdot (\Theta_i \mid a_i)$. Since $\Gamma \Downarrow_{a_i}^{\geq p_i}$, $q_i \geq p_i$ and $\sum_{i \in I} p_i = \sum_{i \in I} q_i = 1$, we have $p_i = q_i$ for each $i \in I$.

Therefore the required matching move is $\Theta \stackrel{\tau}{\Longrightarrow} \sum_{i \in I} p_i \cdot \Theta_i$. This follows because $\sum_{i \in I} p_i \cdot (\Delta_i \mid a_i) \approx_{rbc} \sum_{i \in I} p_i \cdot (\Theta_i \mid a_i)$, from which Lemma 4.11 implies the required $\Delta_i \approx_{rbc} \Theta_i$ for each $i \in I$.

□

Although by definition $\approx_{rbc}$ is closed with respect to the evolution relation $\Longrightarrow$, we can prove that it is also closed with respect to the individual components, and indeed the definite delay operator, in mCCS. This is proved in the following three propositions. First a straightforward case.

Proposition 4.13. Suppose $\Delta \approx_{rbc} \Theta$ in mCCS. If $\Delta \stackrel{\mu}{\Longrightarrow} \Delta'$ with $\mu \in \mathsf{Act}_\tau$ then $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$ such that $\Delta' \approx_{rbc} \Theta'$.

Proof. We can distinguish two cases.

(1) μ is τ. Then for some fresh action *succ* we have $\Delta \mid succ \approx_{rbc} \Theta \mid succ$ and $\Delta \mid succ \Longrightarrow \Delta' \mid succ$. Since $\approx_{rbc}$ is reduction closed, there is some Γ such that $\Theta \mid succ \Longrightarrow \Gamma$ with $\Delta' \mid succ \approx_{rbc} \Gamma$.

Note that because of the operational rules for parallel composition, in Figure 7, and the fact that actions are insistent, the process $\Theta'' \mid succ$ can not perform any weighted actions, definite or otherwise, for any distribution Θ''. Hence, it must be the case that Γ has the form $\Theta' \mid succ$ with $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$. By Lemma 4.11 and $\Delta' \mid succ \approx_{rbc} \Theta' \mid succ$, it follows that $\Delta' \approx_{rbc} \Theta'$.

(2) μ is a, for some $a \in \mathsf{Act}$. Let T be the process $fail + \bar{a}{:}succ$ where *fail* and *succ* are fresh actions. Then $\Delta \mid T \Longrightarrow \Delta' \mid succ$. Since $\Delta \approx_{rbc} \Theta$ we know $\Delta \mid T \approx_{rbc}$

$\Theta \mid T$. Since $\approx_{rbc}$ is reduction-closed, there is some Γ such that $\Theta \mid T \Longrightarrow \Gamma$ and $\Delta' \mid succ \approx_{rbc} \Gamma$. But again because of the use of insistent actions in the test T we actually must have $\Theta \mid T \stackrel{\tau}{\Longrightarrow} \Gamma$.

Since $\approx_{rbc}$ is barb-preserving we have $\Gamma \not\Downarrow^{>0}_{fail}$ and $\Gamma \Downarrow^{\geq 1}_{succ}$. By the construction of the test T it must be the case that Γ has the form $\Theta' \mid succ$ for some Θ' with $\Theta \stackrel{a}{\Longrightarrow} \Theta'$. By Lemma 4.11 and $\Delta' \mid succ \approx_{rbc} \Theta' \mid succ$, it follows that $\Delta' \approx_{rbc} \Theta'$.

□

To prove that $\approx_{rbc}$ is closed under weighted actions is considerably more difficult. In order to organise the proof it is convenient to introduce some notation. Let us say a state s is *stable* if $s \stackrel{\tau}{\not\longrightarrow}$ and a distribution Δ is stable if s is stable for every $s \in \lceil \Delta \rceil$. Note that if $\Delta \stackrel{\lambda}{\longrightarrow}$ then Δ is stable.

Lemma 4.14. In mCCS, suppose $s \stackrel{\lambda}{\longrightarrow} \Delta_w$ and $\overline{s} \approx_{rbc} \overline{t}$, where t is stable. Then there exists some $\overline{t} \stackrel{\lambda}{\Longrightarrow} \Theta'$ such that $\Delta_w \approx_{rbc} \Theta'$.

Proof. First consider the test process $\delta.succ$ where $succ$ is a fresh action. Then $\overline{s} \mid \delta.succ \approx_{rbc} \overline{t} \mid \delta.succ$ and since $s \stackrel{\lambda}{\longrightarrow} \Delta_w$ we know $(s \mid \delta.succ) \Downarrow^{\geq 1}_{succ}$. Since $\approx_{rbc}$ is barb-preserving this means that $(\overline{t} \mid \delta.succ) \Downarrow^{\geq 1}_{succ}$, which can only be the case if there is some weighted derivative $\overline{t} \stackrel{\lambda_t}{\longrightarrow} \Theta_w$. Thus we have two goals; namely show that $\lambda_t = \lambda$ and $\Theta_w \stackrel{\tau}{\Longrightarrow} \Theta'$ for some distribution such that $\Delta_w \approx_{rbc} \Theta'$.

By Lemma 4.10 we know that both s and t also have δ derivatives; in fact because they are states in mCCS they happen to be point distributions, of the form $s \stackrel{\delta}{\longrightarrow} \overline{s_\delta}$ and $t \stackrel{\delta}{\longrightarrow} \overline{t_\delta}$ respectively.

Let T be the test process

$$fail_1 + \delta.(\tau.succ_1 + fail_2) + \lambda.succ_2$$

where $succ_1, succ_2$ and $fail_1, fail_2$ are all fresh actions. Then $\overline{s} \mid T \Longrightarrow \frac{1}{2} \cdot (\Delta_w \mid succ_1) + \frac{1}{2} \cdot (\overline{s_\delta} \mid succ_2)$. Since $\overline{s} \mid T \approx_{rbc} \overline{t} \mid T$ and $\approx_{rbc}$ is reduction-closed, there is some Γ_t such that $\overline{t} \mid T \Longrightarrow \Gamma_t$ and

$$(\frac{1}{2} \cdot (\Delta_w \mid succ_1) + \frac{1}{2} \cdot (\overline{s_\delta} \mid succ_2)) \approx_{rbc} \Gamma_t \tag{8}$$

By Theorem 3.8 we have the weighted move

$$\overline{t} \mid T \stackrel{\lambda_t + \lambda}{\longrightarrow} \frac{\lambda_t}{\lambda_t + \lambda} \cdot (\Theta_w | (\tau.succ_1 + fail_2 + \lambda.succ_2)) + \frac{\lambda}{\lambda_t + \lambda} \cdot (\overline{t_\delta} \mid succ_2). \tag{9}$$

Since $\approx_{rbc}$ is barb-preserving and $\Gamma_t \not\Downarrow^{>0}_{fail_1}$, the state t must contribute some action in the derivation $\bar{t} \mid T \Longrightarrow \Gamma_t$. The contributed action is not a τ action since $t \mid T \not\xrightarrow{\tau}$, so the weighted action in (9) must happen. Thus,

$$\frac{\lambda_t}{\lambda_t + \lambda} \cdot (\Theta_w \mid (\tau.succ_1 + fail_2 + \lambda.succ_2)) + \frac{\lambda}{\lambda_t + \lambda} \cdot (\overline{t_\delta} \mid succ_2) \quad \Longrightarrow \Gamma_t$$

Since $\Gamma_t \not\Downarrow^{>0}_{fail_2}$, we have

$$\begin{array}{ll} & \frac{\lambda_t}{\lambda_t+\lambda} \cdot (\Theta_w \mid (\tau.succ_1 + fail_2 + \lambda.succ_2)) + \frac{\lambda}{\lambda_t+\lambda} \cdot (\overline{t_\delta} \mid succ_2) \\ \stackrel{\tau}{\Longrightarrow} & \frac{\lambda_t}{\lambda_t+\lambda} \cdot (\Theta^2_t \mid succ_1) + \frac{\lambda}{\lambda_t+\lambda} \cdot (\Theta^3_t \mid succ_2) \\ \Longrightarrow & \Gamma_t \end{array}$$

for some Θ^2_t, Θ^3_t with $\Theta_w \stackrel{\tau}{\Longrightarrow} \Theta^2_t$ and $\overline{t_\delta} \stackrel{\tau}{\Longrightarrow} \Theta^3_t$. Since neither $\Theta^2_t \mid succ_1$ nor $\Theta^3_t \mid succ_2$ can perform any weighted action, in the last step of the above reasoning we can replace $\Longrightarrow$ by $\stackrel{\tau}{\Longrightarrow}$. Therefore, there are some Θ^4_t, Θ^5_t such that

$$\Gamma_t \;=\; \frac{\lambda_t}{\lambda_t + \lambda} \cdot (\Theta^4_t \mid succ_1) + \frac{\lambda}{\lambda_t + \lambda} \cdot (\Theta^5_t \mid succ_2) \tag{10}$$

with $\Theta^2_t \stackrel{\tau}{\Longrightarrow} \Theta^4_t$ and $\Theta^3_t \stackrel{\tau}{\Longrightarrow} \Theta^5_t$. Since $\Gamma \Downarrow^{\geq \frac{1}{2}}_{succ_1}$ and $\Gamma \Downarrow^{\geq \frac{1}{2}}_{succ_2}$, we have

$$\frac{\lambda_t}{\lambda_t + \lambda} \geq \frac{1}{2} \qquad \text{and} \qquad \frac{\lambda}{\lambda_t + \lambda} \geq \frac{1}{2} \tag{11}$$

From the two inequalities in (11) we obtain one of our goals, namely $\lambda_t = \lambda$.

Thus (10) can be simplified to

$$\Gamma_t = \frac{1}{2} \cdot (\Theta^4_t \mid succ_1) + \frac{1}{2} \cdot (\Theta^5_t \mid succ_2)$$

Then an application of Lemma 4.11 to (8) above gives that $\Delta_w \approx_{rbc} \Theta^4_t$. This establishes our second goal because $\Theta_w \stackrel{\tau}{\Longrightarrow} \Theta^4_t$. □

Lemma 4.15. In mCCS, suppose $\Delta \approx_{rbc} \Theta$ and $\Delta \xrightarrow{\lambda}$. Then there is a stable Θ_e such that $\Theta \stackrel{\tau}{\Longrightarrow} \Theta_e$ and $\Delta \approx_{rbc} \Theta_e$.

Proof. First consider any hyper-derivative of Θ, $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$. By Proposition 4.13 there is some Δ' with $\Delta \stackrel{\tau}{\Longrightarrow} \Delta'$ and $\Delta' \approx_{rbc} \Theta'$. But since $\Delta \xrightarrow{\lambda}$ we know that Δ is stable. Therefore this Δ' must be Δ itself. In other words every hyper-derivative of Θ is reduction barbed congruent to Δ. In order to prove the lemma we need only find some stable hyper-derivative of Θ.

Let *succ* be a fresh action. Since $\Delta \approx_{rbc} \Theta$ and $(\Delta \mid \delta.succ) \Downarrow^{\geq 1}_{succ}$ it follows that $(\Theta \mid \delta.succ) \Downarrow^{\geq 1}_{succ}$. That is, there is some derivation $\Theta \mid \delta.succ \Longrightarrow \Lambda$, such that $\mathcal{V}_{succ}(\Lambda) \geq 1$. We prove by induction on this derivation that there is some stable Θ_e such that $\Theta \stackrel{\tau}{\Longrightarrow} \Theta_e$.

To this end, it is convenient to rearrange the inductive definition of the evolution relation in Definition 4.1 slightly. Clearly the derivation $\Theta \mid \delta.succ \Longrightarrow \Lambda$ can not follow from reflexivity. Therefore, there are three possibilities.

(i) $\Theta \mid \delta.succ \xrightarrow{\beta} \Lambda'$, where $\Lambda' \Longrightarrow \Lambda$. Here we must have $t \xrightarrow{\beta} \Theta^t_w$ for each $t \in \lceil\Theta\rceil$, and maximal progress also ensures that $t \not\xrightarrow{\tau}$. In this case we can take the required Θ_e to be Θ itself.

(ii) $\Theta \mid \delta.succ \stackrel{\tau}{\Longrightarrow} \Lambda'$, where $\Lambda' \Longrightarrow \Lambda$. Here it is straightforward to show Λ' must have the form $\Theta' \mid succ$ where $\Theta \stackrel{\tau}{\Longrightarrow} \Theta'$ and the result follows in a straightforward manner by induction.

(iii) The final possibility is that $\Theta \mid \delta.succ = \sum_{i \in I} p_i \cdot \Lambda_i$ where $\Lambda = \sum_{i \in I} p_i \cdot \Lambda'_i$, and $\Lambda_i \Longrightarrow \Lambda'_i$ for each $i \in I$.

Then each Λ_i must be in the form $\Theta_i \mid \delta.succ$ and $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$. By induction each Θ_i has some weak stable derivative Θ^e_i. Let $\Theta^e = \sum_{i \in I} p_i \cdot \Theta^e_i$. Clearly, Θ^e is a derivative of Θ and moreover is stable.

□

Proposition 4.16. Suppose $\Delta \approx_{rbc} \Theta$ in finitary mCCS. If $\Delta \xrightarrow{\lambda} \Delta_w$ with $\lambda \in \mathbb{R}^+$ then $\Theta \stackrel{\lambda}{\Longrightarrow} \Theta'$ such that $\Delta_w \approx_{rbc} \Theta'$.

Proof. Because of the previous lemma we may assume that Θ is stable. Clearly, Δ can be written as $\sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \overline{s}$, and so by Weak-left-decomposability, Proposition 4.12, and the fact that Θ is stable, we know that $\Theta = \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \Theta_s$, where for each s, $\overline{s} \approx_{rbc} \Theta_s$. Now, because $\approx_{rbc}$ is symmetric, by another application of the same result to $\overline{s} \approx_{rbc} \Theta_s$ we also know that for each such s, $\overline{s} \approx_{rbc} \overline{t}$ for each $t \in \lceil\Theta_s\rceil$. These are all stable and so we can use Lemma 4.14 to find a $\overline{t} \stackrel{\lambda}{\Longrightarrow} \Theta_t$ such that $\Delta^s_w \approx_{rbc} \Theta_t$, where Δ^s_w is the unique distribution such that $s \xrightarrow{\lambda} \Delta^s_w$.

We now combine, for a particular Θ_s all these weak λ-actions, using Linearity from Corollary 2.11, to obtain $\Theta_s \stackrel{\lambda}{\Longrightarrow} \sum_{t \in \lceil\Theta_s\rceil} \Theta_s(t) \cdot \Theta_t$. For convenience let us use Θ'_s to denote this residual, and note that by Linearity $\Delta^s_w \approx_{rbc} \Theta'_s$.

Because of the determinacy of weighted actions in a MLTS, we know that Δ_w in the statement of the Proposition can be written as $\sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \Delta^s_w$. Moreover, another application of Linearity for weak λ-actions gives that $\Theta \stackrel{\lambda}{\Longrightarrow} \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \Theta'_s$. It therefore follows by the Linearity of $\approx_{rbc}$, Proposition 4.12, that $\Delta_w \approx_{rbc} \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \Theta'_s$ □

Proposition 4.17. Suppose $\Delta \approx_{rbc} \Theta$ in mCCS. If $\Delta \xrightarrow{\delta} \Delta'$ then $\Theta \xRightarrow{\delta} \Theta'$ such that $\Delta' \approx_{rbc} \Theta'$.

Proof. We distinguish two cases.

(i) Suppose $\Delta \not\xrightarrow{\lambda}$ for any weight λ. Here let T be the process $\mathit{fail} + \lambda_0.\mathit{succ}$ where succ and fail are fresh actions and λ_0 is an arbitrary positive rate. Then we have $\Delta \mid T \Longrightarrow \Delta' \mid \mathit{succ}$. Since $\Delta \approx_{rbc} \Theta$, there is some Γ such that $\Theta \mid T \Longrightarrow \Gamma$ and $\Delta' \mid T \approx_{rbc} \Gamma$. The latter implies $\Gamma \not\Downarrow^{>0}_{\mathit{fail}}$ and $\Gamma \Downarrow^{\geq 1}_{\mathit{succ}}$ and so it must be the case that Γ has the form $\Theta' \mid \mathit{succ}$ with $\Theta \xRightarrow{\delta} \Theta'$. Now it follows from Lemma 4.11 that $\Delta' \approx_{rbc} \Theta'$.

(ii) If $\Delta \xrightarrow{\lambda} \Delta_\lambda$ for some weight λ. By Proposition 4.16 there is a matching transition $\Theta \xRightarrow{\tau} \Theta_d \xrightarrow{\lambda} \Theta'_\lambda \xRightarrow{\tau} \Theta_\lambda$. Note that the move $\Theta \xRightarrow{\tau} \Theta_d$ can only be matched by the degenerate transition $\Delta \xRightarrow{\tau} \Delta$ because by maximal progress $\Delta \not\xrightarrow{\tau}$. It therefore follows that $\Delta \approx_{rbc} \Theta_d$.

Let T be the process

$$\mathit{fail}_1 + \delta.(\tau.\mathbf{0} + \mathit{fail}_2) + \lambda.\mathit{succ}$$

where $\mathit{fail}_1, \mathit{fail}_2$ and succ are fresh actions. Since $\Delta \xrightarrow{\lambda} \Delta_\lambda$, Lemma 4.10 ensures that Δ also has a δ (unique) derivative, $\Delta \xrightarrow{\delta} \Delta_\delta$. Then $\Delta \mid T \Longrightarrow \frac{1}{2} \cdot (\Delta_\lambda \mid \mathbf{0}) + \frac{1}{2} \cdot (\Delta_\delta \mid \mathit{succ})$, using Theorem 3.8. Since $\approx_{rbc}$ is compositional we have $\Delta \mid T \approx_{rbc} \Theta_d \mid T$. Then there is some Γ such that $\Theta_d \mid T \Longrightarrow \Gamma$ and

$$(\frac{1}{2} \cdot (\Delta_\lambda \mid \mathbf{0}) + \frac{1}{2} \cdot (\Delta_\delta \mid \mathit{succ})) \approx_{rbc} \Gamma. \tag{12}$$

This means $\Gamma \not\Downarrow^{>0}_{\mathit{fail}_1}$ and therefore Θ_d must have contributed some action in the derivation $\Theta_d \mid T \Longrightarrow \Gamma$. The contributed action is not a τ action since $\Theta_d \mid T \not\xrightarrow{\tau}$, so a weighted action must happen as follows, using Theorem 3.8 again.

$$\Theta_d \mid T \xrightarrow{2\lambda} (\frac{1}{2} \cdot (\Theta'_\lambda \mid (\tau.\mathbf{0} + \mathit{fail}_2 + \lambda.\mathit{succ})) + \frac{1}{2} \cdot (\Theta_\delta \mid \mathit{succ})) \Longrightarrow \Gamma$$

Again the existence of Θ_δ with $\Theta_d \xrightarrow{\delta} \Theta_\delta$ is assured by Lemma 4.10, because we are using mCCS processes.

Since $\Gamma \not\Downarrow^{>0}_{\mathit{fail}_2}$, we have

$$\begin{array}{ll} & \frac{1}{2} \cdot (\Theta'_\lambda \mid (\tau.\mathbf{0} + \mathit{fail}_2 + \lambda.\mathit{succ})) + \frac{1}{2} \cdot (\Theta_\delta \mid \mathit{succ}) \\ \xRightarrow{\tau} & \frac{1}{2} \cdot (\Theta_1 \mid \mathbf{0}) + \frac{1}{2} \cdot (\Theta_2 \mid \mathit{succ}) \\ \Longrightarrow & \Gamma \end{array}$$

for some Θ_1, Θ_2 with $\Theta'_\lambda \stackrel{\tau}{\Longrightarrow} \Theta_1$ and $\Theta_\delta \stackrel{\tau}{\Longrightarrow} \Theta_2$. Since neither $\Theta_1 \mid \mathbf{0}$ nor $\Theta_2 \mid succ$ can enable any weighted action, in the last step we can replace $\Longrightarrow$ by $\stackrel{\tau}{\Longrightarrow}$. Then there are some Θ_3, Θ_4 such that

$$\Gamma = \frac{1}{2} \cdot (\Theta_3 \mid \mathbf{0}) + \frac{1}{2} \cdot (\Theta' \mid succ) \tag{13}$$

with $\Theta_1 \stackrel{\tau}{\Longrightarrow} \Theta_3$ and $\Theta_2 \stackrel{\tau}{\Longrightarrow} \Theta'$. By (12), (13) and Lemma 4.11, we have $\Delta' \approx_{rbc} \Theta'$.

On the other hand, we have $\Theta \stackrel{\tau}{\Longrightarrow} \Theta_d \stackrel{\delta}{\longrightarrow} \Theta_\delta \stackrel{\tau}{\Longrightarrow} \Theta_2 \stackrel{\tau}{\Longrightarrow} \Theta'$ and therefore $\Theta \stackrel{\delta}{\Longrightarrow} \Theta'$ is the required matching move.

□

Theorem 4.18. [Completeness] In finitary mCCS, $\Delta \approx_{rbc} \Theta$ implies $\Delta \approx_{bis} \Theta$.

Proof. We show that $\approx_{rbc}$ is a bisimulation, that is $\approx_{rbc} \subseteq \mathcal{B}(\approx_{rbc})$, where $\mathcal{B}$ is the functional given in Definition 2.12. Because of symmetry it is sufficient to show that if $\Delta \stackrel{\mu}{\Longrightarrow} \sum_{i \in I} p_i \cdot \Delta_i$ with $\sum_{i \in I} p_i = 1$, where $\mu \in \mathsf{Act}_{\tau,\delta} \cup \mathbb{R}^+$ and I is a finite index set, there is a matching move $\Theta \stackrel{\mu}{\Longrightarrow} \sum_{i \in I} p_i \cdot \Theta_i$ for some Θ_i such that $\Delta_i \approx_{rbc} \Theta_i$.

In fact because of Proposition 4.12 (Weak-left-decomposable) it is sufficient to match a simple move $\Delta \stackrel{\mu}{\Longrightarrow} \Delta'$ with a simple move $\Theta \stackrel{\mu}{\Longrightarrow} \Theta'$ such that $\Delta' \approx_{rbc} \Theta'$. But this can easily be established using Propositions 4.13, 4.16 and 4.17. □

5 Conclusion and related work

The thesis underlying this paper is that bisimulations should be considered as a proof methodology for demonstrating behavioural equivalence between systems, rather than providing the definition of the extensional behavioural equivalence itself. We have adapted the well-known *reduction barbed congruence* used for a variety of process calculi [14, 21, 10], to obtain a touchstone extensional behavioural equivalence for a minor variation of the Markov automata, MAs, originally defined in [9]. Incidently there are also minor variations on the formulation of reduction barbed congruence, often called *contextual equivalence* or *barbed congruence*, in the literature. See [10, 24] for a discussion of the differences.

Then we have defined a novel notion of (weak) bisimulations, called Markov bisimulations, which provide both a sound and complete coinductive proof methodology for establishing the equivalence between such automata. These results were achieved within the context of a rich language, mCCS, for defining MAs. Of particular significance is the presence of indefinite delay actions, insistent actions, and a compositional operator which is sensitive to the passage of time; this combination is reminiscent of synchronous CCS [19], although similar compositional operators have already been used for certain varieties of Markov processes [3]. We should point out that our interpretation of mCCS is somewhat simplistic, in that unlike IMC in [11] it does not take into account the multiplicities of action occurrences. However, our interpretation is sufficient for the purposes of this paper. If we were interested in, for example, developing an algebraic theory for mCCS then a more refined interpretation would be required; this could easily be adapted from [11].

There are already quite a few variations on the theme of bisimulations for PAs which can be used to establish behavioural equivalences between MAs [27, 20, 16, 8, 12]. A characteristic of our formulation is that it allows bisimulations to relate states to distributions rather than simply states, thus differentiating it from most of these. The one exception is [9], where properties of subdistributions are also used in defining their bisimulations. However, our Markov bisimulation equivalence $\approx_{bis}$ is different from the bisimulation equivalence of [9], denoted by $\approx_{MA}$ here, because the former is defined for full distributions while the latter is for subdistributions. Even if we restrict $\approx_{MA}$ to full distributions, they are still different. For example, we have $A \approx_{bis} \mathbf{0}$ but $A \not\approx_{MA} \mathbf{0}$, where $A \Leftarrow \tau : A$. We conjecture that in general $\approx_{bis}$ is strictly coarser than $\approx_{MA}$ (restricted to full distributions), but they coincide for non-divergent systems [9]. We discuss the relationship between our $\approx_{bis}$ and $\approx_{MA}$ in more detail in Appendix B.

Our approach to Markov processes is based directly on that of [9, 11], in which external actions are considered instantaneous, and time can only pass when no more internal activity can be performed. Moreover, it is only timed actions which are subject to Markovian behaviour. However, there is a large literature on a more general framework in which Markovian behaviour applies to all actions. See [13] or Chapter 3 of [1] for a representative exposition. It would be interesting to see if our notion of bisimulation could be adapted to such a framework.

Our notion of reduction relation, $\Longrightarrow$, is closely related to the concept of scheduler, sometimes called policy, adversary, resolution etc., occurring in the literature. One may consider different classes of schedulers (e.g. deterministic schedulers and probabilistic schedulers). In each class only a restricted form of reduction is allowed, which yields a restricted form of reduction barbed congruence. It would be interesting to compare these different variants and characterise them by co-inductively defined relations.

Acknowledgement

We thank Christian Eisentraut for the interesting discussion on clarifying the relationship between $\approx_{bis}$ and $\approx_{MA}$.

A Some properties of hyper-derivations

These results on hyper-derivations were originally proven in the full-version of [6]. We include them here for the sake of completeness.

Lemma A.1. In an arbitrary MLTS the relation $\Longrightarrow$ is linear.

Proof. Let us first introduce some notation for hyper-derivations. Referring to Definition 2.8 let us abbreviate the hyper-derivation (1) from Δ by $\langle \Delta_k^{\rightarrow}, \Delta_k^{\times}, k \geq 0 \rangle$. Recall that this is a hyper-derivation from Δ to $\sum_{k\geq 0} \Delta_k^{\times}$, that is $\Delta \Longrightarrow \sum_{k\geq 0} \Delta_k^{\times}$.

Now suppose we have $\Delta^j \Longrightarrow \Theta^j$ for every $j \in J$. We have to construct a hyper-derivation from $\sum_{j\in J} p_j \cdot \Delta^j$ to $\sum_{j\in J} p_j \cdot \Theta^j$.

From the hypothesis we have for each j a hyper-derivation $\langle \Delta_k^{\rightarrow j}, \Delta_k^{\times j}, k \geq 0 \rangle$, where Θ_j is $\sum_{k\geq 0} \Delta_k^{\times j}$. By construction the single arrow $\xrightarrow{\tau}$ is linear, since it is defined as a lifted relation. Therefore applying linearity for each k we get a hyper-derivation of the form $\langle \sum_{j\in J} p_j \cdot \Delta_k^{\rightarrow j}, \sum_{j\in J} p_j \cdot \Delta_k^{\times j}, k \geq 0 \rangle$.

This is easily seen to be the required hyper-derivation. □

Lemma A.2. For any subdistributions $\Delta, \Theta, \Gamma, \Lambda, \Pi$ we have

(i) If $\Delta \Longrightarrow \Theta$ then $|\Delta| \geq |\Theta|$.

(ii) If $\Delta \Longrightarrow \Theta$ and $p \in \mathbb{R}^+$ such that $|p \cdot \Delta| \leq 1$, then $p \cdot \Delta \Longrightarrow p \cdot \Theta$.

(iii) If $\Gamma + \Lambda \Longrightarrow \Pi$ then $\Pi = \Pi^\Gamma + \Pi^\Lambda$ with $\Gamma \Longrightarrow \Pi^\Gamma$ and $\Lambda \Longrightarrow \Pi^\Lambda$.

Proof. By definition $\Delta \Longrightarrow \Theta$ means that some $\Delta_k, \Delta_k^{\times}, \Delta_k^{\rightarrow}$ exist for all $k \geq 0$ such that

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^{\times} + \Delta_k^{\rightarrow}, \qquad \Delta_k^{\rightarrow} \xrightarrow{\tau} \Delta_{k+1}, \qquad \Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}.$$

(i) A simple inductive proof shows that

$$|\Delta| \quad = \quad |\Delta_i^{\rightarrow}| + \sum_{k\leq i} |\Delta_k^{\times}| \text{ for any } i \geq 0. \tag{14}$$

The sequence $\{\sum_{k\leq i} |\Delta_k^{\times}|\}_{i=0}^{\infty}$ is nondecreasing and by (14) each element of the sequence is not greater than $|\Delta|$. Therefore, the limit of this sequence is bounded by $|\Delta|$. That is,

$$|\Delta| \quad \geq \quad \lim_{i\to\infty} \sum_{k\leq i} |\Delta_k^{\times}| \quad = \quad |\Theta|$$

(ii) Now suppose $p \in \mathbb{R}^+$ such that $|p \cdot \Delta| \leq 1$. From Definition 2.4 it follows that

$$p{\cdot}\Delta = p{\cdot}\Delta_0, \qquad p{\cdot}\Delta_k = p{\cdot}\Delta_k^{\rightarrow} + p{\cdot}\Delta_k^{\times}, \qquad p{\cdot}\Delta_k^{\rightarrow} \xrightarrow{\tau} p{\cdot}\Delta_{k+1}, \qquad p{\cdot}\Theta = \sum_k p{\cdot}\Delta_k^{\times}.$$

Hence Definition 2.8 yields $p \cdot \Delta \Longrightarrow p \cdot \Theta$.

(iii) Next suppose $\Gamma + \Lambda \Longrightarrow \Pi$. By Definition 2.8 there are subdistributions $\Pi_k, \Pi_k^{\rightarrow}, \Pi_k^{\times}$ for $k \in \mathbb{N}$ such that

$$\Gamma + \Lambda = \Pi_0, \qquad \Pi_k = \Pi_k^{\rightarrow} + \Pi_k^{\times}, \qquad \Pi_k^{\rightarrow} \xrightarrow{\tau} \Pi_{k+1}, \qquad \Pi = \sum_k \Pi_k^{\times}.$$

For any $s \in S$, define

$$\begin{array}{rcl} \Gamma_0^{\rightarrow}(s) & := & min(\Gamma(s), \Pi_0^{\rightarrow}(s)) \\ \Gamma_0^{\times}(s) & := & \Gamma(s) - \Gamma_0^{\rightarrow}(s) \\ \Lambda_0^{\times}(s) & := & min(\Lambda(s), \Pi_0^{\times}(s)) \\ \Lambda_0^{\rightarrow}(s) & := & \Lambda(s) - \Lambda_0^{\times}(s)\ , \end{array} \tag{15}$$

and check that $\Gamma_0^{\rightarrow} + \Gamma_0^{\times} = \Gamma$ and $\Lambda_0^{\rightarrow} + \Lambda_0^{\times} = \Lambda$. To show that $\Lambda_0^{\rightarrow} + \Gamma_0^{\rightarrow} = \Pi_0^{\rightarrow}$ and $\Lambda_0^{\times} + \Gamma_0^{\times} = \Pi_0^{\times}$ we fix a state s and distinguish two cases: either (a) $\Pi_0^{\rightarrow}(s) \geq \Gamma(s)$ or (b) $\Pi_0^{\rightarrow}(s) < \Gamma(s)$. In Case (a) we have $\Pi_0^{\times}(s) \leq \Lambda(s)$ and the definitions (15) simplify to $\Gamma_0^{\rightarrow}(s) = \Gamma(s)$, $\Gamma_0^{\times}(s) = 0$, $\Lambda_0^{\times}(s) = \Pi_0^{\times}(s)$ and $\Lambda_0^{\rightarrow}(s) = \Lambda(s) - \Pi_0^{\times}(s)$, whence immediately $\Gamma_0^{\rightarrow}(s) + \Lambda_0^{\rightarrow}(s) = \Pi_0^{\rightarrow}(s)$ and $\Gamma_0^{\times}(s) + \Lambda_0^{\times}(s) = \Pi_0^{\times}(s)$. Case (b) is similar.

Since $\Lambda_0^{\rightarrow} + \Gamma_0^{\rightarrow} \xrightarrow{\tau} \Pi_1$, by Proposition 2.7 we find Γ_1, Λ_1 with $\Gamma_0^{\rightarrow} \xrightarrow{\tau} \Gamma_1$ and $\Lambda_0^{\rightarrow} \xrightarrow{\tau} \Lambda_1$ and $\Pi_1 = \Gamma_1 + \Lambda_1$. Being now in the same position with Π_1 as we were with Π_0, we can continue this procedure to find Λ_k, Γ_k, $\Lambda_k^{\rightarrow}$, $\Gamma_k^{\rightarrow}$, $\Lambda_k^{\times}$ and $\Gamma_k^{\times}$ with

$$\begin{array}{lll} \Gamma = \Gamma_0, & \Gamma_k = \Gamma_k^{\rightarrow} + \Gamma_k^{\times}, & \Gamma_k^{\rightarrow} \xrightarrow{\tau} \Gamma_{k+1}, \\ \Lambda = \Lambda_0, & \Lambda_k = \Lambda_k^{\rightarrow} + \Lambda_k^{\times}, & \Lambda_k^{\rightarrow} \xrightarrow{\tau} \Lambda_{k+1}, \\ \Gamma_k + \Lambda_k = \Pi_k, & \Gamma_k^{\rightarrow} + \Lambda_k^{\rightarrow} = \Pi_k^{\rightarrow}, & \Gamma_k^{\times} + \Lambda_k^{\times} = \Pi_k^{\times}. \end{array}$$

Let $\Pi^{\Gamma} := \sum_k \Gamma_k^{\times}$ and $\Pi^{\Lambda} := \sum_k \Lambda_k^{\times}$. Then $\Pi = \Pi^{\Gamma} + \Pi^{\Lambda}$ and Definition 2.8 yields $\Gamma \Longrightarrow \Pi^{\Gamma}$ and $\Lambda \Longrightarrow \Pi^{\Lambda}$.

□

Together, Lemma A.2(ii) and (iii) imply that $\Longrightarrow$ is left-decomposable, as in Definition 2.6, for finite index sets I.

We now generalise this result to infinite, but still countable, index sets.

Theorem A.3. [Infinite left-decomposition] Let $p_i \in [0, 1]$ for $i \in I$ with $\sum_{i\in I} p_i \leq 1$, and I a countable index set. If $\sum_{i\in I} p_i \cdot \Delta_i \Longrightarrow \Theta$ then $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$ for subdistributions Θ_i such that $\Delta_i \Longrightarrow \Theta_i$ for all $i \in I$.

Proof. In the light of Lemma A.2(ii) it suffices to show that if $\sum_{i=0}^{\infty} \Delta_i \Longrightarrow \Theta$ then $\Theta = \sum_{i=0}^{\infty} \Theta_i$ for subdistributions Θ_i such that $\Delta_i \Longrightarrow \Theta_i$ for all $i \geq 0$.

Since $\sum_{i=0}^{\infty} \Delta_i = \Delta_0 + \sum_{i\geq 1} \Delta_i$ and $\sum_{i=0}^{\infty} \Delta_i \Longrightarrow \Theta$, by Lemma A.2(iii) there are $\Theta_0, \Theta_1^{\geq}$ such that

$$\Delta_0 \Longrightarrow \Theta_0, \qquad \sum_{i\geq 1} \Delta_i \Longrightarrow \Theta_1^{\geq}, \qquad \Theta = \Theta_0 + \Theta_1^{\geq}\ .$$

Using Lemma A.2(iii) once more, we have $\Theta_1, \Theta_2^{\geq}$ such that

$$\Delta_1 \Longrightarrow \Theta_1, \qquad \sum_{i\geq 2} \Delta_i \Longrightarrow \Theta_2^{\geq}, \qquad \Theta_1^{\geq} = \Theta_1 + \Theta_2^{\geq},$$

thus in combination $\Theta = \Theta_0 + \Theta_1 + \Theta_2^{\geq}$. Continuing this process we have that

$$\Delta_i \Longrightarrow \Theta_i, \qquad \sum_{j\geq i+1} \Delta_j \Longrightarrow \Theta_{i+1}^{\geq}, \qquad \Theta = \sum_{j=0}^{i} \Theta_j + \Theta_{i+1}^{\geq}$$

for all $i \geq 0$. Lemma A.2(i) ensures that $|\sum_{j\geq i+1} \Delta_j| \geq |\Theta_{i+1}^{\geq}|$ for all $i \geq 0$. But since $\sum_{i=0}^{\infty} \Delta_i$ is a subdistribution, we know that the tail sum $\sum_{j\geq i+1} \Delta_j$ converges to ε when i approaches ∞, and therefore that $\lim_{i\to\infty} \Theta_i^{\geq} = \varepsilon$. Thus by taking that limit we conclude that $\Theta = \sum_{i=0}^{\infty} \Theta_i$. $\square$

We proceed with the important properties of reflexivity and transitivity of weak derivations. First note that reflexivity is straightforward; in Definition 2.8 it suffices to take $\Delta_0^{\rightarrow}$ to be the empty distribution ε.

Theorem A.4. [Transitivity of $\Longrightarrow$] If $\Delta \Longrightarrow \Theta$ and $\Theta \Longrightarrow \Lambda$ then $\Delta \Longrightarrow \Lambda$.

Proof. By definition $\Delta \Longrightarrow \Theta$ means that some $\Delta_k, \Delta_k^{\times}, \Delta_k^{\rightarrow}$ exist for all $k \geq 0$ such that

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^{\times} + \Delta_k^{\rightarrow} \qquad \Delta_k^{\rightarrow} \xrightarrow{\tau} \Delta_{k+1} \qquad \Theta = \sum_{k=0}^{\infty} \Delta_k^{\times} \tag{16}$$

Since $\Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}$ and $\Theta \Longrightarrow \Lambda$, by Theorem A.3 there are Λ_k for $k \geq 0$ such that $\Lambda = \sum_{k=0}^{\infty} \Lambda_k$ and $\Delta_k^{\times} \Longrightarrow \Lambda_k$ for all $k \geq 0$. For each $k \geq 0$, we know that $\Delta_k^{\times} \Longrightarrow \Lambda_k$ gives us some $\Delta_{kl}, \Delta_{kl}^{\times}, \Delta_{kl}^{\rightarrow}$ for $l \geq 0$ such that

$$\Delta_k^{\times} = \Delta_{k0}, \qquad \Delta_{kl} = \Delta_{kl}^{\times} + \Delta_{kl}^{\rightarrow}, \qquad \Delta_{kl}^{\rightarrow} \xrightarrow{\tau} \Delta_{k,l+1} \qquad \Lambda_k = \sum_{l\geq 0} \Delta_{kl}^{\times}. \tag{17}$$

Therefore we can put all this together with

$$\Lambda \quad = \quad \sum_{k=0}^{\infty} \Lambda_k \quad = \quad \sum_{k,l \geq 0} \Delta^{\times}_{kl} \quad = \quad \sum_{i \geq 0} \left(\sum_{k,l|k+l=i} \Delta^{\times}_{kl} \right) \tag{18}$$

where the last step is a straightforward diagonalisation.

Now from the decompositions above we re-compose an alternative trajectory of Δ'_i's to take Δ via $\Longrightarrow$ directly to Λ. Define

$$\Delta'_i = \Delta'^{\times}_i + \Delta'^{\rightarrow}_i, \qquad \Delta'^{\times}_i = \sum_{k,l|k+l=i} \Delta^{\times}_{kl}, \qquad \Delta'^{\rightarrow}_i = (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_i \ , \tag{19}$$

so that from (18) we have immediately that

$$\Lambda \quad = \quad \sum_{i \geq 0} \Delta'^{\times}_i \ . \tag{20}$$

We now show that

(i) $\Delta = \Delta'_0$

(ii) $\Delta'^{\rightarrow}_i \xrightarrow{\tau} \Delta'_{i+1}$

from which, with (20), we will have $\Delta \Longrightarrow \Lambda$ as required. For (i) we observe that

	Δ	
$=$	Δ_0	(16)
$=$	$\Delta^{\times}_0 + \Delta^{\rightarrow}_0$	(16)
$=$	$\Delta_{00} + \Delta^{\rightarrow}_0$	(17)
$=$	$\Delta^{\times}_{00} + \Delta^{\rightarrow}_{00} + \Delta^{\rightarrow}_0$	(17)
$=$	$(\sum_{k,l\|k+l=0} \Delta^{\times}_{kl}) + (\sum_{k,l\|k+l=0} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_0$	index arithmetic
$=$	$\Delta'^{\times}_0 + \Delta'^{\rightarrow}_0$	(19)
$=$	Δ'_0 .	(19)

For (ii) we observe that

	$\Delta'^{\rightarrow}_i$	
$=$	$(\sum_{k,l\|k+l=i} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_i$	(19)
$\xrightarrow{\tau}$	$(\sum_{k,l\|k+l=i} \Delta_{k,l+1}) + \Delta_{i+1}$	(16), (17), linearity of $\xrightarrow{\tau}$
$=$	$(\sum_{k,l\|k+l=i}(\Delta^{\times}_{k,l+1} + \Delta^{\rightarrow}_{k,l+1})) + \Delta^{\times}_{i+1} + \Delta^{\rightarrow}_{i+1}$	(16), (17)
$=$	$(\sum_{k,l\|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1} + (\sum_{k,l\|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1}$	rearrange
$=$	$(\sum_{k,l\|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta_{i+1,0} + (\sum_{k,l\|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1}$	(17)

$$
\begin{array}{lll}
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1,0} + \Delta^{\rightarrow}_{i+1,0} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & (17) \\
= & (\sum_{k,l|k+l=i+1} \Delta^{\times}_{kl}) + (\sum_{k,l|k+l=i+1} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_{i+1} & \text{index arithmetic} \\
= & \Delta'^{\times}_{i+1} + \Delta'^{\rightarrow}_{i+1} & (19) \\
= & \Delta'_{i+1}\ , & (19)
\end{array}
$$

which concludes the proof. □

We now turn our attention to the proof of Theorem 2.18. Unfortunately this relies on two complex auxiliary results, the first of which is much too long to include here.

Theorem A.5. For any distribution in a finitary MLTS, the set $\{\, \Delta' \mid \Delta \Longrightarrow \Delta' \,\}$ is Cauchy closed in the standard Euclidean space.

Proof. Established as Theorem 1 of [6]. The proof is relatively complex, as it relies on developing hyper-derivations relative to static policies. The restriction to finitary MLTSs is necessary because in such a MLTS the set of static polices is finite. □

Theorem A.6. [Infinite linearity] Suppose $\mathcal{R}$ is a relation over $S \times \mathcal{D}_{sub}(S)$, where S is finite, and $\sum_{i\geq 0} p_i = 1$. Then Δ_i $lift(\mathcal{R})$ Θ_i implies $(\sum_{i\geq 0} p_i \cdot \Delta_i)$ $lift(\mathcal{R})$ $(\sum_{i\geq 0} p_i \cdot \Theta_i)$.

Proof. See Appendix C □

Proof of Theorem 2.18: Suppose s is a state and Θ a distribution in a finitary MLTS such that $s \approx_{sbis} \Theta$ and $\overline{s} \stackrel{\tau}{\Longrightarrow} \Delta'$. Referring to Definition 2.8, there must be Δ_k, $\Delta^{\rightarrow}_k$ and $\Delta^{\times}_k$ for $k \geq 0$ such that $\overline{s} = \Delta_0$, $\Delta_k = \Delta^{\rightarrow}_k + \Delta^{\times}_k$, $\Delta^{\rightarrow}_k \stackrel{\tau}{\longrightarrow} \Delta_{k+1}$ and $\Delta' = \sum_{k=1}^{\infty} \Delta^{\times}_k$. Since $\Delta^{\times}_0 + \Delta^{\rightarrow}_0 = \overline{s}$ $lift(\approx_{sbis})$ Θ, using Proposition 2.7 we can define $\Theta := \Theta^{\times}_0 + \Theta^{\rightarrow}_0$ so that $\Delta^{\times}_0$ $lift(\approx_{sbis})$ $\Theta^{\times}_0$ and $\Delta^{\rightarrow}_0$ $lift(\approx_{sbis})$ $\Theta^{\rightarrow}_0$. Since $\Delta^{\rightarrow}_0 \stackrel{\tau}{\longrightarrow} \Delta_1$ and $\Delta^{\rightarrow}_0$ $lift(\approx_{sbis})$ $\Theta^{\rightarrow}_0$, by Proposition 2.17 we have $\Theta^{\rightarrow}_0 \Longrightarrow \Theta_1$ with Δ_1 $lift(\approx_{sbis})$ Θ_1.

Repeating the above procedure gives us inductively a series $\Theta_k, \Theta^{\rightarrow}_k, \Theta^{\times}_k$ of subdistributions, for $k \geq 0$, such that $\Theta_0 = \Theta$, Δ_k $lift(\approx_{sbis})$ Θ_k, $\Theta_k = \Theta^{\rightarrow}_k + \Theta^{\times}_k$, $\Delta^{\times}_k$ $lift(\approx_{sbis})$ $\Theta^{\times}_k$, $\Delta^{\rightarrow}_k$ $lift(\approx_{sbis})$ $\Theta^{\rightarrow}_k$ and $\Theta^{\rightarrow}_k \stackrel{\tau}{\Longrightarrow} \Theta_k$. We define $\Theta' := \sum_i \Theta^{\times}_i$. By the infinite linearity of lifting operation, Theorem A.6, we have Δ' $lift(\approx_{sbis})$ Θ'. It remains to be shown that $\Theta \Longrightarrow \Theta'$.

For that final step, since the set $\{\Theta'' \mid \Theta \Longrightarrow \Theta''\}$ is closed according to Theorem A.5, we can establish $\Theta \Longrightarrow \Theta'$ by exhibiting a sequence Θ'_i with $\Theta \Longrightarrow \Theta'_i$ for each i and with the Θ'_i's being arbitrarily close to Θ'. Induction establishes for each i that $\Theta \Longrightarrow \Theta'_i := (\Theta^{\rightarrow}_i + \sum_{k\leq i} \Theta^{\times}_k)$. Since Δ' is a full distribution, whose mass is 1, i.e. $|\Delta'| = 1$, we must have $\lim_{i\to\infty} |\Delta^{\rightarrow}_i| = 0$. It is easy to see that for any two subdistributions Γ_1, Γ_2 if Γ_1 $lift(\approx_{sbis})$ Γ_2 then they have the same mass. Therefore, it follows from the condition $\Delta^{\rightarrow}_i$ $lift(\approx_{sbis})$ $\Theta^{\rightarrow}_i$ that $\lim_{i\to\infty} |\Theta^{\rightarrow}_i| = 0$. Thus these Θ'_i's form the sequence we needed. □

B An alternative presentation of bisimulation

We have said in Section 5 that the bisimulation equivalence $\approx_{MA}$ given in [9] is different from our Markov bisimulation equivalence $\approx_{bis}$. Certainly the formulation is quite different in style to our definition. But here we suggest that a good approximation to $\approx_{MA}$ can be formulated in our framework, and we show that the resulting equivalence actually coincides with our Markov bisimulation equivalence, $\approx_{bis}$. This serves to emphasise that, modulo minor differences, the equivalences $\approx_{bis}$ and $\approx_{MA}$ are *essentially the same*.

Definition B.1. A binary relation $\mathcal{R} \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$, where S is the set of states in a MLTS, is called an *ehz-bisimulation*, whenever $\Delta \mathrel{\mathcal{R}} \Theta$ implies

(i) if $\Delta \stackrel{\tau}{\Longrightarrow} (\Delta_1 \,{}_p\!\oplus \Delta_2)$, for any $0 \leq p \leq 1$, then

  (a) $\Theta \stackrel{\tau}{\Longrightarrow} (\Theta_1 \,{}_p\!\oplus \Theta_2)$ with $\Delta_1 \mathrel{\mathcal{R}} \Theta_1$ and $\Delta_2 \mathrel{\mathcal{R}} \Theta_2$

  (b) and for each $\mu \in \mathsf{Act}_\delta \cup \mathbb{R}^+$, $\Delta_1 \stackrel{\mu}{\Longrightarrow} \Delta_1'$ implies $\Theta_1 \stackrel{\mu}{\Longrightarrow} \Theta_1'$ and $\Delta_1' \mathrel{\mathcal{R}} \Theta_1'$

(ii) and symmetrically for Θ.

We write $\approx_{ehz}$ for the largest ehz-bisimulation. □

Although this is similar to the formulation of $\approx_{MA}$ in [9] we should point out the differences.

(i) Our definition of Markov automata, Definition 2.1, and MLTSs, Definition 2.2, assume maximal progress, whereas the automata of [9] do not; instead in [9] maximal progress is implemented in their definition of bisimulation.

(ii) Our definition of $\approx_{ehz}$ uses as weak internal actions, the hyper-derivatives $\Longrightarrow$ from Definition 2.8 whereas that of $\approx_{MA}$ in [9] uses a definition of weak internal moves based on finite branching labelled trees.

(iii) Our definition of $\approx_{ehz}$ restricts attention to (full) distributions whereas in [9] the definition of $\approx_{MA}$ more generally applies to subdistributions. Of course, this difference is insignificant because we can easily extend $\approx_{ehz}$ to a relation between subdistributions by replacing $\mathcal{D}(S)$ with $\mathcal{D}_{sub}(S)$ in Definition B.1 because $\stackrel{\mu}{\Longrightarrow}$ can be a relation between subdistributions in general [6].

Nevertheless we feel that Definition B.1 captures the essence of $\approx_{MA}$ from [9]. Note that the resulting relation $\approx_{ehz}$ is still different from $\approx_{MA}$ because the latter, but not the former, has a stability requirement, which intuitively means that if a state is stable then any state bisimilar to it must be able to perform some internal transitions to reach a stable state.

Lemma B.2. Suppose $\Delta \approx_{ehz} \Theta$ and $\Delta \stackrel{\tau}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta_i$ for some finite index set I. There exists distribution Θ_i for each $i \in I$ such that $\Delta_i \approx_{ehz} \Theta_i$ and $\Theta \stackrel{\tau}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta_i$.

Proof. Since I is a finite set, we can assume without loss of generality that $I = \{1..n\}$ for some $n \geq 1$. We proceed by induction on n.

- $n = 1$. By the first clause of Definition B.1, there exists Θ_1 such that $\Theta \Longrightarrow \Theta_1$ and $\Delta_1 \approx_{ehz} \Theta_1$.
- Suppose the result holds for m where $m \geq 1$. We consider the case that $n = m+1$. From Δ there is a transition

$$\Delta \stackrel{\tau}{\Longrightarrow} (\Delta_n\ {}_{p_n}\!\oplus (\sum_{i=1..m} \frac{p_i}{1-p_n} \cdot \Delta_i)).$$

Since $\Delta \approx_{ehz} \Theta$, we know from Definition B.1 that there are distributions Θ_n and Θ' such that

$$\Theta \stackrel{\tau}{\Longrightarrow} (\Theta_n\ {}_{p_n}\!\oplus \Theta') \tag{21}$$

with $\Delta_n \approx_{ehz} \Theta_n$ and $(\sum_{i=1..m} \frac{p_i}{1-p_n} \cdot \Delta_i) \approx_{ehz} \Theta'$. By the induction hypothesis and the degenerate transition $(\sum_{i=1..m} \frac{p_i}{1-p_n} \cdot \Delta_i) \stackrel{\tau}{\Longrightarrow} (\sum_{i=1..m} \frac{p_i}{1-p_n} \cdot \Delta_i)$, there exists Θ_i for each $i = 1..m$ such that $\Delta_i \approx_{ehz} \Theta_i$ and

$$\Theta' \stackrel{\tau}{\Longrightarrow} \sum_{i=1..m} \frac{p_i}{1-p_n} \cdot \Theta_i \tag{22}$$

Using the reflexivity, linearity and transitivity of $\stackrel{\tau}{\Longrightarrow}$, we know from (21) and (22) that

$$\Theta \stackrel{\tau}{\Longrightarrow} \sum_{i=1..n} p_i \cdot \Theta_i$$

as required.

□

Theorem B.3. In an arbitrary MLTS, $\Delta \approx_{ehz} \Theta$ if and only if $\Delta \approx_{bis} \Theta$.

Proof. (⇐) It is straightforward to show that $\approx_{bis}$ is an ehz-bisimulation.

(⇒) We show that $\approx_{ehz}$ is a Markov bisimulation. Suppose $\Delta \approx_{ehz} \Theta$ and $\Delta \stackrel{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta_i$. If $\mu = \tau$ then we directly appeal to Lemma B.2. Now assume that $\mu \neq \tau$. In the first clause of Definition B.1, by setting $p = 1$ and $\Delta_1 = \Delta$, we see that there exist Θ_1, Θ'_1 such that $\Theta \stackrel{\tau}{\Longrightarrow} \Theta_1$ with $\Delta_1 \approx_{ehz} \Theta_1$ and $\Theta_1 \stackrel{\mu}{\Longrightarrow} \Theta'_1$ with $(\sum_{i\in I} p_i \cdot \Delta_i) \approx_{ehz} \Theta'_1$. By Lemma B.2 there are distributions Θ_i such that $\Theta'_1 \stackrel{\tau}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta_i$ and $\Delta_i \approx_{ehz} \Theta_i$ for each $i \in I$. Note that we also have the transition $\Theta \stackrel{\mu}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Theta_i$. □

C Infinite linearity of lifting operation

For any subset X of $\mathcal{D}_{sub}(S)$ let $cc(X)$, the convex closure of X, be the least set satisfying:

(i) $X \subseteq cc(X)$

(ii) $\Delta, \Theta \in cc(X)$ implies $\Delta\,{}_p\!\oplus\,\Theta \in cc(X)$, for every $0 \leq p \leq 1$.

Equivalently we can say $\Delta \in cc(X)$ if and only if $\Delta = \sum_{i\in I} p_i \cdot \Delta_i$, where $\Delta_i \in X$, for some finite index set I such that $\sum_{i\in I} p_i = 1$. Analogously let $cc_\omega(X)$ be the set of subdistributions of the form $\sum_{i\geq 0} p_i \cdot \Delta_i$, where $\Delta_i \in X$ and $\sum_{i\geq 0} p_i = 1$. Both of these are closure operators in the standard sense; for example $cc(-)$ satisfies:

- $X \subseteq cc(X)$
- $X \subseteq Y$ implies $cc(X) \subseteq cc(Y)$
- $cc(cc(X)) = cc(X)$.

Lemma C.1. If the set S is finite then $cc(X) = cc_\omega(X)$ for any subset X of $\mathcal{D}_{sub}(S)$.

Proof. It is clear that $cc(X) \subseteq cc_\omega(X)$, so we prove the inverse inclusion, $cc_\omega(X) \subseteq cc(X)$. As in [7] we view a subdistribution over S as a point in Euclidean space of dimension $|S|$ and give a geometric proof, by induction on the size of S. More specifically we prove, by induction on k, that if X is a subset in a space of dimension k, that $cc(X) = cc_\omega(X)$. The base case, when $|S| = 1$ is trivial. Let us we consider the inductive case, where the dimension is $(k + 1)$.

Suppose there is a point $x \in cc_\omega(X)$ but $x \notin cc(X)$. We show that this contradicts the inductive hypothesis.

From the Hyperplane separation theorem, Theorem 1.2.4 in [17], there exists a hyperplane H that separates x from $cc(X)$. If h is the normal of H we can assume without loss of generality that there is a constant c satisfying

$$h \cdot x \geq c \text{ and } h \cdot x' \leq c \text{ for all } x' \in X$$

where with a slight abuse of notation we write $\cdot$ for dot product of two vectors of dimension $(k + 1)$.

Since $x \in cc_\omega(X)$, there is a sequence of probabilities p_i with $\sum_{i\geq 0} p_i = 1$ and a sequence of points $x_i \in X$ such that $x = \sum_{i\geq 0} p_i \cdot x_i$. We then have

(i) $c \geq h \cdot x = \sum_{i\geq 0} p_i \cdot (h \cdot x_i)$

(ii) $h \cdot x_i \leq c$ for all $i \geq 0$.

It follows from (i) and (ii) that actually $h \cdot x_i = c$ for all $i \geq 0$. In other words, it must be the case that $h \cdot x_i = c$ for all i, which means that all the points x_i lies in H; in other words the separation of x from $cc(X)$ can not be strict. Therefore, we have that $x \in cc_\omega((X \cap H))$ since $cc_\omega(\{x_i \mid i \geq 0\}) \subseteq cc_\omega((X \cap H))$.

On the other hand, since $x \notin cc(X)$ we have $x \notin cc((X \cap H))$. However $X \cap H$ can be described as a subset in a space of one dimension lower than X, that is of dimension k. We have now contradicted the induction hypothesis. $\square$

In order to use this result to prove Theorem A.6 we need to rephrase the definition of *lifting*, Definition 2.4, in terms of the closure operator $cc(-)$. To this end let us use $\mathcal{R}\,(s)$, for any $\mathcal{R} \subseteq S \times \mathcal{D}_{sub}(S)$, to denote the set $\{\, \Delta \mid a \mathrel{\mathcal{R}} \Delta \,\}$.

Proposition C.2. For subdistributions over a finite set S, Δ *lift*$(\mathcal{R})$ Θ if and only if Θ can be written in the form $\sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot \Theta_s$ where each $\Theta_s \in cc(\mathcal{R}\,(s))$.

Proof. Suppose $\Theta = \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot \Theta_s$ with $\Theta_s \in cc(\mathcal{R}\,(s))$. To show that Δ *lift*$(\mathcal{R})$ Θ, it suffices to prove that $\overline{s}$ *lift*$(\mathcal{R})$ Θ_s for each $s \in \lceil\Delta\rceil$, as *lift*$(\mathcal{R})$ is linear. Since $\Theta_s \in cc(\mathcal{R}\,(s))$, we can rewrite Θ_s as $\Theta_s = \sum_{i\in I} p_i \cdot \Theta_{i_s}$ where $\Theta_{i_s} \in \mathcal{R}\,(s)$ for some finite index set I. The fact that $\overline{s} = \sum_{i\in I} p_i \cdot \overline{s}$ and $s \mathrel{\mathcal{R}} \Theta_{i_s}$ yields that $\overline{s}$ *lift*$(\mathcal{R})$ Θ_s.

Conversely suppose Δ *lift*$(\mathcal{R})$ Θ. By Lemma 2.5 we have that

$$\Delta = \sum_{i\in I} p_i \cdot \overline{s_i} \qquad s_i \mathrel{\mathcal{R}} \Theta_i \qquad \Theta = \sum_{i\in I} p_i \cdot \Theta_i. \tag{23}$$

For each $s \in \lceil\Delta\rceil$, let $I_s = \{i \in I \mid s_i = s\}$. Note that $\Delta(s) = \sum_{i\in I_s} p_i$. Hence, we can rewrite Θ as follows:

$$\begin{array}{rcl} \Theta & = & \sum_{s\in\lceil\Delta\rceil} \sum_{i\in I_s} p_i \cdot \Theta_i \\ & = & \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot (\sum_{i\in I_s} \frac{p_i}{\Delta(s)} \cdot \Theta_i) \end{array}$$

Since the subdistribution $\sum_{i\in I_s} \frac{p_i}{\Delta(s)} \cdot \Theta_i$ is a convex combination of $\{\Theta_i \mid i \in I_s\}$, it must be in $cc(\mathcal{R}\,(s))$ due to (23), and the result follows. $\square$

Proof of Theorem A.6: Suppose $\sum_{i\geq 0} p_i = 1$ and Δ_i *lift*($\mathcal{R}$) Θ_i for each $i \geq 0$. Let Δ, Θ denote $\sum_{i\geq 0} p_i \cdot \Delta_i$ and $\sum_{i\geq 0} p_i \cdot \Theta_i$ respectively. We have to show $\Delta\ \mathcal{R}\ \Theta$. By Proposition C.2 it is sufficient to show

$$\Theta = \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot \Gamma_s \tag{24}$$

where $\Gamma_s \in cc(\mathcal{R}\ (s))$ for each $s \in \lceil\Delta\rceil$.

By the same proposition we know that for each $i \geq 0$, since Δ_i *lift*($\mathcal{R}$) Θ_i,

$$\Theta_i = \sum_{s\in\lceil\Delta_i\rceil} \Delta_i(s) \cdot \Theta_{i_s} \text{ with } \Theta_{i_s} \in cc(\mathcal{R}\ (s)). \tag{25}$$

Therefore,

$$\begin{array}{rcl} \Theta & = & \sum_{i\geq 0} p_i \cdot (\sum_{s\in\lceil\Delta_i\rceil} \Delta_i(s) \cdot \Theta_{i_s}) \\ & = & \sum_{s\in\lceil\Delta\rceil} \sum_{i\geq 0} (p_i \cdot \Delta_i(s)) \cdot \Theta_{i_s} \end{array}$$

Let w_i^s denote $p_i \cdot \Delta_i(s)$ and note that $\Delta(s)$ is the infinite sum $\sum_{i\geq 0} w_i^s$. Therefore we can continue:

$$\begin{array}{rcl} \Theta & = & \sum_{s\in\lceil\Delta\rceil} \sum_{i\geq 0} w_i^s \cdot \Theta_{i_s} \\ & = & \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot (\sum_{i\geq 0} \frac{w_i^s}{\Delta(s)} \cdot \Theta_{i_s}) \end{array}$$

The required (24) above will follow if we can show $(\sum_{i\geq 0} \frac{w_i^s}{\Delta(s)} \cdot \Theta_{i_s}) \in cc(\mathcal{R}\ (s))$ for each $s \in \lceil\Delta\rceil$.

From (25) we know $\Theta_{i_s} \in cc(\mathcal{R}\ (s))$, and therefore by construction we have that $(\sum_{i\geq 0} \frac{w_i^s}{\Delta(s)} \cdot \Theta_{i_s}) \in cc_\omega(cc(\mathcal{R}\ (s)))$. But now an application of Lemma C.1 gives $cc_\omega(cc(\mathcal{R}\ (s))) = cc(cc(\mathcal{R}\ (s)))$, and since $cc(-)$ is a closure operator this coincides with $cc(\mathcal{R}\ (s))$. □

References

[1] Alessandro Aldini, Marco Bernardo, and Flavio Corradini. *A Process Algebraic Approach to Software Architecture Design*. Springer, 2010.

[2] Marco Bernardo and Roberto Gorrieri. A tutorial on EMPA: A theory of concurrent processes with nondeterminism, priorities, probabilities and time. *Theoretical Computer Science*, 202(1-2):1–54, 1998.

[3] Mario Bravetti and Marco Bernardo. Compositional asymmetric cooperations for process algebras with probabilities, priorities, and time. *Electronic Notes in Theoretical Computer Science*, 39(3), 2000.

[4] Yuxin Deng and Wenjie Du. Probabilistic barbed congruence. *Electronic Notes in Theoretical Computer Science*, 190(3):185–203, 2007.

[5] Yuxin Deng and Catuscia Palamidessi. Axiomatizations for probabilistic finite-state behaviors. *Theoretical Computer Science*, 373(1-2):92–114, 2007.

[6] Yuxin Deng, Rob van Glabbeek, Matthew Hennessy, and Carroll Morgan. Testing finitary probabilistic processes (extended abstract). In *Proceedings of the 20th International Conference on Concurrency Theory*, volume 5710 of *Lecture Notes in Computer Science*, pages 274–288. Springer, 2009. Full version available at `http://basics.sjtu.edu.cn/~yuxin/publications/finitary.pdf`.

[7] Yuxin Deng, Rob J. van Glabbeek, Carroll Morgan, and Chenyi Zhang. Scalar outcomes suffice for finitary probabilistic testing. In Rocco De Nicola, editor, *ESOP*, volume 4421 of *Lecture Notes in Computer Science*, pages 363–378. Springer, 2007.

[8] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Weak bisimulation is sound and complete for pCTL*. *Information and Computation*, 208(2):203–219, 2010.

[9] Christian Eisentraut, Holger Hermanns, and Lijun Zhang. On probabilistic automata in continuous time. In *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science*, pages 342–351. IEEE Computer Society, 2010.

[10] Cédric Fournet and Georges Gonthier. A hierarchy of equivalences for asynchronous calculi. *Journal of Logic and Algebraic Programming*, 63(1):131–173, 2005.

[11] Holger Hermanns. *Interactive Markov Chains: The Quest for Quantified Quality*, volume 2428 of *Lecture Notes in Computer Science*. Springer, 2002.

[12] Holger Hermanns, Augusto Parma, Roberto Segala, Björn Wachter, and Lijun Zhang. Probabilistic logical characterization. *Information and Computation*, 209:154–172, February 2011.

[13] Jane Hillston. *A Compositional Approach to Performance Modelling (Distinguished Dissertations in Computer Science)*. Cambridge University Press, New York, NY, USA, 2005.

[14] Kohei Honda and Mario Tokoro. On asynchronous communication semantics. In P. Wegner M. Tokoro, O. Nierstrasz, editor, *Proceedings of the ECOOP '91 Workshop on Object-Based Concurrent Computing*, volume 612 of *LNCS 612*. Springer-Verlag, 1992.

[15] A. Jeffrey and J. Rathke. Contextual equivalence for higher-order pi-calculus revisited. *Logical Methods in Computer Science*, 1(1:4), 2005.

[16] K. G. Larsen and A. Skou. Bisimulation through probabilistic testing (preliminary report). In *Proceedings of the 16th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*, pages 344–352, New York, NY, USA, 1989. ACM.

[17] Jiri Matousek. *Lectures on Discrete Geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer, 2002.

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

[19] Robin Milner. Calculi for synchrony and asynchrony. *Theoretical Computer Science*, 25:267–310, 1983.

[20] Anna Philippou, Insup Lee, and Oleg Sokolsky. Weak bisimulation for probabilistic systems. In *Proceedings of the 11th International Conference on Concurrency Theory*, volume 1877 of *Lecture Notes in Computer Science*, pages 334–349. Springer, 2000.

[21] Julian Rathke and Pawel Sobocinski. Deriving structural labelled transitions for mobile ambients. In *Proceedings of the 19th International Conference on Concurrency Theory*, volume 5201 of *Lecture Notes in Computer Science*, pages 462–476. Springer, 2008.

[22] Julian Rathke and Pawel Sobocinski. Making the unobservable, unobservable. *Electronic Notes in Computer Science*, 229(3):131–144, 2009.

[23] D. Sangiorgi, N. Kobayashi, and E. Sumii. Environmental bisimulations for higher-order languages. In *Proceedings of the 22nd IEEE Symposium on Logic in Computer Science*, pages 293–302. IEEE Computer Society, 2007.

[24] D. Sangiorgi and D. Walker. *The π-calculus: a Theory of Mobile Processes*. Cambridge University Press, 2001.

[25] Roberto Segala. Modeling and verification of randomized distributed real-time systems. Technical Report MIT/LCS/TR-676, PhD thesis, MIT, Dept. of EECS, 1995.

[26] Roberto Segala. Testing probabilistic automata. In *Proceedings of the 7th International Conference on Concurrency Theory*, volume 1119 of *Lecture Notes in Computer Science*, pages 299–314. Springer, 1996.

[27] Roberto Segala and Nancy A. Lynch. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2(2):250–273, 1995.