Raiders of the Lost Equivalence: Probabilistic Branching Bisimilarity

Valentina Castiglioni[a,1], Simone Tini[b,1]

[a] *Reykjavik University, Reykjavik, Iceland*
[b] *University of Insubria, Como, Italy*

Abstract

Probabilistic branching bisimilarity allows us to compare process behavior with respect to their branching structure and probabilistic features, while abstracting away from those computation steps that can be marked as irrelevant to the analysis at hand. To the best of our knowledge, in the context of nondeterministic probabilistic processes, no proof of probabilistic branching bisimilarity being an equivalence has been provided so far. Since, as happened in the fully nondeterministic case, some researchers are using such a result by taking it for granted, we decided to dedicate this note to a formal proof of it. More precisely, we extend and adapt the proof strategy used by Basten in the fully nondeterministic case. Thus, we introduce the probabilistic analogue to the notion of semi-branching bisimilarity and to van Glabbeek and Weijland's Stuttering Lemma to prove that probabilistic branching bisimilarity is indeed an equivalence relation.

Keywords: Nondeterministic probabilistic processes, Probabilistic branching bisimilarity, Semi-branching bisimilarity, Equivalence relations

1. Introduction

Branching bisimulation [19] is a *behavioural equivalence* that generalizes bisimulation [31, 32] to abstract away from (non observable, irrelevant, hidden) *silent* computation steps of processes while preserving their *branching structure*. This key feature is mainly due to the *stuttering* nature of branching bisimulation, which guarantees that the potential of a process is preserved in the execution of a sequence of silent steps [19, 20]. It is therefore not surprising that branching bisimulation has become one of the most studied behavioural equivalences (see among others [5–8, 10, 11, 15–17, 21, 22, 39]).

Following the ever-increasing interest in *probabilistic systems*, several notions of *probabilistic branching bisimulation* have been proposed (see, e.g., [2, 4, 9, 29, 35, 37]). The reason behind such a wealth of definitions is two-fold:

- There are several different probabilistic models, like the alternating [24, 33], fully probabilistic [27], generative, and reactive models [18], labeled Markov chains [23], and nondeterministic probabilistic labeled transition systems (PTSs) [34], each one with distinctive features leading to different formalization of branching bisimulation.
- There are several possible interpretations of the interplay of nondeterminism, probability and weak semantics. This prevents us from identifying a "*unique, correct*" extension of branching bisimulation to the probabilistic setting.

All such notions suffer the same *drawback*, which is actually shared with the fully nondeterministic branching bisimulation: it is not true, in general, that the composition of two (probabilistic) branching bisimulations is in turn a (probabilistic) branching bisimulation [5]. It is therefore not trivial to prove that the relation obtained as the union of all (probabilistic) branching bisimulations, the so called (*probabilistic*) *branching bisimilarity*, is an equivalence relation.

In the fully nondeterministic setting, in [5] Twan Basten proved such a result by exploiting the notion of *semi-branching bisimulation* [19] and the stuttering property of branching bisimulations.

In this note we extend his technique and provide a detailed proof of probabilistic branching bisimilarity being an equivalence relation over processes in the PTS model, namely processes whose next computation step is chosen nondeterministically and each transition takes a process to a (discrete) probability distribution over processes. The result is given for PTSs that are *divergence-free*, namely PTSs where all sequences of silent steps are finite. Interestingly, we provide an example showing that the interplay of nondeterminism and probability breaks the stuttering property of both branching bisimilarity and semi-branching bisimilarity for those PTSs that admit divergence. This suggests us that these behavioural properties are interesting only in the divergence-free case, since stuttering is one of their main desired property. Divergence-freedom is a property of PTSs induced by several process algebras, like, e.g. the family *timed process algebras* originally introduced in [26], where in between two *tick* actions modelling the discrete passage of time only a limited num-

[1] Corresponding authors. E-mail: `vale.castiglioni@gmail.com` (Valentina Castiglioni), `simone.tini@uninsubria.it` (Simone Tini)

ber of silent actions is admitted, in order to prevent Zeno behaviours. Roughly,

1. In the context of PTSs, we introduce the notion of *probabilistic semi-branching bisimulation.* Technically, probabilistic semi-branching bisimulation relaxes probabilistic branching bisimulation by giving to pairs of related processes more freedom in mimicking their initial silent moves.

2. We show that the relation composition of two probabilistic semi-branching bisimulations is again a semi-branching bisimulation (Proposition 3) even if we lift them to relations over distributions (Lemma 5). This allows us to conclude that probabilistic semi-branching bisimilarity, namely the union of all semi-branching bisimulations, is transitive. Since it is also reflexive and symmetric, we infer that it is an equivalence relation (Theorem 1). This result holds also for PTSs with divergence.

3. We introduce the *probabilistic* equivalent to the *stuttering Lemma* from [20]. This states that, in the divergence-free case, *probabilistic semi-branching bisimilarity* satisfies the (probabilistic) stuttering property (Proposition 5).

4. By observing that each probabilistic semi-branching bisimulation with the stuttering property is also a probabilistic branching bisimulation, we show (Corollary 1) that probabilistic semi-branching bisimilarity *coincides with probabilistic branching bisimilarity.* Hence, we conclude that *probabilistic branching bisimilarity is an equivalence relation* (Theorem 2). Clearly, this holds in the divergence-free case.

5. We provide an example showing that both branching bisimilarity and semi-branching bisimilarity are not stuttering for PTSs admitting divergences.

Interestingly, [1] presents a proof of probabilistic branching bisimilarity being an equivalence on a different semantic model, namely that of non-strictly alternating processes [33]. In that model, there is a partition between nondeterministic processes and probabilistic ones, and processes of each type only perform transitions of the same type. Also due to the differences between the considered semantic models, the proof technique of [1] is quite different from ours in that they define each probabilistic branching bisimulation directly as an equivalence relation and they prove that the transitive closure of the union of all probabilistic branching bisimulations, i.e., probabilistic branching bisimilarity, which is an equivalence relation by construction, is again a probabilistic branching bisimulation. The same proof technique is used in [3] to prove that probabilistic branching bisimilarity is an equivalence over fully probabilistic processes [27], namely processes with a deterministic transition to a distribution.

2. Background

In this section we introduce the notion of probabilistic branching bisimulation on the PTS model.

2.1. The PTS model.

Given a countable set X, a *discrete probability distribution* over X is a mapping $\pi\colon X \to [0,1]$ such that $\sum_{x\in X}\pi(x) = 1$. The *support* of π is the set $\mathsf{supp}(\pi) = \{x \in X \mid \pi(x) > 0\}$. By $\Delta(X)$ we denote the set of all *finitely supported* distributions over X, ranged over by $\pi, \pi', \dots$. Given an element $x \in X$, we let $\delta(x)$ denote the *Dirac's distribution on* x, defined by $\delta(x)(x) = 1$ and $\delta(x)(y) = 0$ for all $y \neq x$. For a finite set of indexes I, weights $p_i \in (0,1]$ with $\sum_{i\in I} p_i = 1$ and distributions $\pi_i \in \Delta(X)$ with $i \in I$, the distribution $\sum_{i\in I} p_i\pi_i$ is defined by $(\sum_{i\in I} p_i\pi_i)(x) = \sum_{i\in I} p_i \cdot \pi_i(x)$, for all $x \in X$.

PTSs [34] combine LTSs [28] and discrete time Markov chains [23], to model, at the same time, reactive behavior, nondeterminism and probability. In a PTS, the state space is given by a set $\mathbf{S}$ of *processes*, ranged over by $s, t, \dots$. The transition steps take processes to probability distributions over processes in $\Delta(\mathbf{S})$ and are labeled by *actions* describing the observable activity of the process. We assume a set of actions $\mathcal{A}_\tau = \mathcal{A} \cup \{\tau\}$, where the *silent* action τ models an internal computation step that cannot be observed by the external environment. We let $\alpha, \beta, \dots$ range over $\mathcal{A}_\tau$ and $a, b, \dots$ range over $\mathcal{A}$.

Definition 1 (PTS, [34])**.** A *nondeterministic probabilistic labelled transition system (PTS)* is a triple $(\mathbf{S}, \mathcal{A}_\tau, \rightarrow)$, where: 1. $\mathbf{S}$ is a countable set of processes, 2. $\mathcal{A}_\tau = \mathcal{A} \cup \{\tau\}$ is a countable set of actions, and, 3. $\rightarrow \subseteq \mathbf{S} \times \mathcal{A}_\tau \times \Delta(\mathbf{S})$ is a transition relation.

A *transition* $(s, \alpha, \pi) \in \rightarrow$ is usually denoted by $s \xrightarrow{\alpha} \pi$.

In a PTS, given an infinite set I, a *divergence* is a sequence of processes $\{s_i\}_{i\in I}$ and distributions $\{\pi_i\}_{i\in I}$ with $s_i \xrightarrow{\tau} \pi_i$ and $s_{i+1} \in \mathsf{supp}(\pi_i)$, for all $i \in I$. The proof that probabilistic branching bisimulation is an equivalence is given for divergence-free PTSs. We refer the interested reader to Section 5 for further details on divergence.

2.2. Relation lifting

A *probabilistic branching bisimulation* is a relation over $\mathbf{S}$ that relates two processes if they can mimic each others *observable* transitions and evolve to distributions related, in turn, by the same relation. To formalize this intuition, we need to *lift* relations over processes to distributions. This is obtained via the notion of *matching*, also known as *coupling* or *weight function*, for distributions.

Definition 2 (Matching)**.** Let X, Y be two countable sets and consider two distributions $\pi_X \in \Delta(X)$ and $\pi_Y \in \Delta(Y)$. A *matching* for π_X and π_Y is a distribution over the product space $\mathfrak{w} \in \Delta(X \times Y)$ having π_X and π_Y as left and right marginals, respectively, namely $\sum_{y\in Y} \mathfrak{w}(x,y) =$

$\pi_X(x)$, for all $x \in X$, and $\sum_{x \in X} \mathfrak{w}(x,y) = \pi_Y(y)$, for all $y \in Y$. We denote by $\mathfrak{W}(\pi_X, \pi_Y)$ the set of all matchings for π_X and π_Y.

Definition 3 (Relation lifting, [34])**.** The *lifting* of a relation $\mathcal{R} \subseteq X \times Y$ is the relation $\mathcal{R}^\ell \subseteq \Delta(X) \times \Delta(Y)$ with $\pi_X \mathcal{R}^\ell \pi_Y$ if and only if there is a matching $\mathfrak{w} \in \mathfrak{W}(\pi_X, \pi_Y)$ such that $x \mathcal{R} y$ whenever $\mathfrak{w}(x,y) > 0$.

We recall here some definitions equivalent to Definition 3 which will be useful in our proofs.

Proposition 1 ([12, Def. 2.1 and Thm. 2.4])**.** *Consider a relation $\mathcal{R} \subseteq X \times Y$. Then $\mathcal{R}^\ell \subseteq \Delta(X) \times \Delta(Y)$ is the smallest relation satisfying*

1. *$x \mathcal{R} y$ implies $\delta(x) \mathcal{R}^\ell \delta(y)$;*
2. *$\pi_i \mathcal{R}^\ell \pi'_i$ implies $(\sum_{i \in I} p_i \pi_i) \mathcal{R}^\ell (\sum_{i \in I} p_i \pi'_i)$, for any finite index set I with $p_i \in (0,1]$ and $\sum_{i \in I} p_i = 1$.*

Proposition 2 ([13, Prop. 1])**.** *Consider two sets X, Y. Let $\pi_X \in \Delta(X)$, $\pi_Y \in \Delta(Y)$ and $\mathcal{R} \subseteq X \times Y$. Then $\pi_X \mathcal{R}^\ell \pi_Y$ if and only if there are a finite set of indexes I and of weights $p_i \in (0,1]$ with $\sum_{i \in I} p_i = 1$, such that*

- *$\pi_X = \sum_{i \in I} p_i \delta(x_i)$,*
- *$\pi_Y = \sum_{i \in I} p_i \delta(y_i)$,*
- *$x_i \; \mathcal{R} \; y_i$ for all $i \in I$.*

2.3. Weak transitions

In order to define *weak transitions*, which allow us to abstract away from silents steps, we need to lift the notion of a transition to a relation between probability distributions. This is called a *hyper-transition* in [29, 30] and stems from [14, 30]. Here, we adapt it to the PTS model. The idea is that a transition $\pi \xrightarrow{\alpha} \pi'$ between two distributions $\pi, \pi' \in \Delta(\mathbf{S})$ is enabled if and only if *all processes* in the support of π can perform an α-move and π' is the convex combination of the distributions reached through these transitions, namely for all $s \in \mathsf{supp}(\pi)$ we have $s \xrightarrow{\alpha} \pi_s$, for some $\pi_s \in \Delta(\mathbf{S})$, and $\pi' = \sum_{s \in \mathsf{supp}(\pi)} \pi(s)\pi_s$. However, this way of lifting transitions is too strict to deal with weak semantics like probabilistic branching bisimulation. In fact, as in the nonprobabilistic setting, a process can simulate a τ-move of another process by not performing any move, we need to allow (some of) the processes in the support of distributions to do the same. In order to formalise that some processes in a distribution have the freedom to simulate a τ by idling, we consider the set of labels $\hat{\mathcal{A}}_\tau = \{\hat{\alpha} \mid \alpha \in \mathcal{A}_\tau\}$, with $\hat{\alpha} \notin \mathcal{A}_\tau$, and use $\mathcal{A}_\tau \cup \hat{\mathcal{A}}_\tau$ to label lifted transitions. Then, $\pi \xrightarrow{\tau}_\ell \pi'$ denotes that *all* processes in the support of π make the silent transitions, whereas $\pi \xrightarrow{\hat{\tau}}_\ell \pi'$ denotes that *some* processes in the support of π make the silent transitions.

Definition 4 (Lifted transition)**.** For a PTS $(\mathbf{S}, \mathcal{A}_\tau, \to)$, we define the relation $\to_\ell \colon (\mathbf{S} \cup \Delta(\mathbf{S})) \times (\mathcal{A}_\tau \cup \hat{\mathcal{A}}_\tau) \times \Delta(\mathbf{S})$ from $\to$ as follows:

- $s \xrightarrow{\alpha}_\ell \pi$ if and only if $s \xrightarrow{\alpha} \pi$;
- $s \xrightarrow{\hat{\alpha}}_\ell \pi$ if and only if $\begin{cases} \text{either } s \xrightarrow{\alpha} \pi \\ \text{or } \alpha = \tau \text{ and } \pi = \delta(s); \end{cases}$
- for $(\alpha) \in \{\alpha, \hat{\alpha}\}$, $\pi \xrightarrow{(\alpha)}_\ell \pi'$ if and only if:
  - $s \xrightarrow{(\alpha)}_\ell \pi_s$ for all $s \in \mathsf{supp}(\pi)$, and
  - $\pi' = \sum_{s \in \mathsf{supp}(\pi)} \pi(s)\pi_s$.

Clearly, if $\alpha \neq \tau$ then relations $\xrightarrow{\hat{\alpha}}_\ell$ and $\xrightarrow{\alpha}_\ell$ coincide. The reason for having both notations is that, sometimes, it will be convenient to write $\xrightarrow{\hat{\alpha}}_\ell$ or $\xrightarrow{\alpha}_\ell$ by letting α range to the whole $\mathcal{A}_\tau$.

As usual we write $\xrightarrow{\hat{\varepsilon}}_\ell$ for the reflexive-transitive closure of the relation $\xrightarrow{\hat{\tau}}_\ell$. We remark that in the divergence-free case a lifted transition of the form $s \xrightarrow{\hat{\tau}}_\ell \delta(s)$ can never be inferred from any transition $s \xrightarrow{\tau} \delta(s)$. Therefore, in that case, $s \xrightarrow{\hat{\tau}}_\ell \delta(s)$ is syntactic sugar to denote that no silent-move took place.

2.4. Probabilistic branching bisimilarity

Our definition of probabilistic branching bisimulation is equivalent to the *scheduler-free* version defined in [1, 29].

Definition 5 (Probabilistic branching bisimulation)**.** A binary relation $\mathcal{B} \subseteq \mathbf{S} \times \mathbf{S}$ is a *probabilistic branching bisimulation* if whenever $s \, \mathcal{B} \, t$ then:

1. whenever $s \xrightarrow{\alpha} \pi_s$ then
   (a) either $\alpha = \tau$ and $\pi_s \, \mathcal{B}^\ell \, \delta(t)$;
   (b) or $t \xrightarrow{\hat{\varepsilon}}_\ell \pi \xrightarrow{\alpha}_\ell \pi_t$ with $\delta(s) \, \mathcal{B}^\ell \, \pi$ and $\pi_s \, \mathcal{B}^\ell \, \pi_t$;
2. whenever $t \xrightarrow{\alpha} \pi_t$ then
   (a) either $\alpha = \tau$ and $\delta(s) \, \mathcal{B}^\ell \, \pi_t$;
   (b) or $s \xrightarrow{\hat{\varepsilon}}_\ell \pi \xrightarrow{\alpha}_\ell \pi_s$ with $\pi \, \mathcal{B}^\ell \, \delta(t)$ and $\pi_s \, \mathcal{B}^\ell \, \pi_t$.

Then, we say that s and t are *probabilistic branching bisimilar*, notation $s \approx_{\mathrm{b}} t$, if there exists a probabilistic branching bisimulation relating them.

Notice that an arbitrary union of probabilistic branching bisimulations is in turn a probabilistic branching bisimulation. Therefore the union of all probabilistic branching bisimulations, namely $\approx_{\mathrm{b}}$, is the largest probabilistic branching bisimulation, which will be called *probabilistic branching bisimilarity*. We remark also that Definition 5 automatically guarantees that the inverse of a probabilistic branching bisimulation is a probabilistic branching bisimulation.

2.5. Related work in brief

We remark that we do not require branching bisimulations to be symmetric, although it is easy to prove that our Definition 5 would be equivalent to its formulation in terms of symmetric relations. We opted for the present formulation because, as already pointed out in [5] for the fully nondeterministic case, it is not true in general that the relation composition of two symmetric relations is a symmetric relation, thus implying that the closure of probabilistic branching bisimulations under union would not be immediate. Even further, differently from e.g., [1, 3, 35, 36], we do not require a probabilistic branching bisimulation $\mathcal{B}$ to be an equivalence relation (cf., e.g., [1, Definition 5]). As argued in [25], requiring each relation $\mathcal{B}$ to be an equivalence is very restrictive, although unnecessary, as to verify equivalence we would need to use witness relations that are in turn equivalences. We stress also that our approach is *state-based* as opposite to the *distribution-based* one of, e.g., [14]. Briefly, the main difference in the two approaches is that in the latter, each nondeterministic choice of a process becomes a probabilistic choice. As a consequence, it is always possible to reason on subdistributions over processes, instead of distributions, generating thus a coarser behavioral relation on them. Notice that, from this point of view, the use of randomized schedulers, or combined transitions, in [35] (and thus in all the subsequent ones) is equivalent to the distribution based approach. It is however our opinion that, as perfectly embodied by the PTS model and the chosen state-based approach, nondeterminism and probability are two distinct concepts obeying their own principles and laws and thus we need to model both of them, as well as their interplay, without substituting one with the other.

3. Probabilistic semi-branching bisimilarity

To prove that probabilistic branching bisimilarity is an equivalence, we extend and improve the proof technique used in [5], for the fully nondeterministic case, to deal with the probabilistic behavior of processes. We start, in this section, by relaxing the notion of probabilistic branching bisimulation to that of *probabilistic semi-branching bisimulation*, and we show that the largest probabilistic semi-branching bisimulation is an equivalence relation (Theorem 1). This result is proved also for processes admitting divergence. Informally, the only difference between these two behavioural relations is in the treatment of τ as initial step: for a semi-branching bisimulation $\mathcal{S}$ and processes $s \ \mathcal{S}\ t$, if s performs a τ-transition $s \xrightarrow{\tau} \pi_s$ as in case 1 in Definition 5, then the first possible answer by t formalized in case 1a and consisting in making no move provided that $\delta(t)$ is related with π_s, is relaxed by admitting that there is an arbitrary sequence of silent transitions taking t to a distribution π related with π_s. On the contrary, the second possible answer by t formalised in case 1b is left unchanged. Technically, instead of rewriting the relaxed version of case 1a and the original case 1b, in the following formal definition we compact the two cases by simply asking that t performs a (possibly empty) sequence of τ-transitions $t \xrightarrow{\hat{\varepsilon}}_\ell \pi$ with $\delta(s)\ \mathcal{S}^\ell\ \pi$ and, then, π performs a lifted τ-move $\pi \xrightarrow{\hat{\tau}}_\ell \pi_t$ with $\pi_s\ \mathcal{S}^\ell\ \pi_t$, thus allowing each process in $\mathsf{supp}(\pi)$ to mimic or not the τ-move by s.

Definition 6 (Probabilistic semi-branching bisimulation)**.** A binary relation $\mathcal{S} \subseteq \mathbf{S}\times\mathbf{S}$ is called a *probabilistic semi-branching bisimulation* if whenever $s\ \mathcal{S}\ t$ then:

- whenever $s \xrightarrow{\alpha} \pi_s$ then $t \xrightarrow{\hat{\varepsilon}}_\ell \pi \xrightarrow{\hat{\alpha}}_\ell \pi_t$ with $\delta(s)\ \mathcal{S}^\ell\ \pi$ and $\pi_s\ \mathcal{S}^\ell\ \pi_t$;
- whenever $t \xrightarrow{\alpha} \pi_t$ then $s \xrightarrow{\hat{\varepsilon}}_\ell \pi \xrightarrow{\hat{\alpha}}_\ell \pi_s$ with $\pi\ \mathcal{S}^\ell\ \delta(t)$ and $\pi_s\ \mathcal{S}^\ell\ \pi_t$.

Two processes s, t are called *probabilistic semi-branching bisimilar*, denoted by $s \approx_{\mathrm{sb}} t$, if there exists a probabilistic semi-branching bisimulation relating them.

Notice that an arbitrary union of probabilistic semi-branching bisimulations is in turn a probabilistic semi-branching bisimulation. Therefore, the union of all probabilistic semi-branching bisimulations, namely $\approx_{\mathrm{sb}}$, is the largest probabilistic semi-branching bisimulation, which will be called *probabilistic semi-branching bisimilarity.*

We also remark that, by Definition 4, in case $\alpha = \tau$ the lifted transition $\pi \xrightarrow{\hat{\tau}}_\ell \pi_t$ in the first item of Definition 6 (and, analogously, the lifted transition $\pi \xrightarrow{\hat{\tau}}_\ell \pi_s$ in the second item) does not imply that a τ-step was actually performed, that is the processes in the support of π are not obliged to mimic the τ-move by s. This additional level of freedom will allow us to obtain that the relation composition of two probabilistic semi-branching bisimulations is in turn a probabilistic semi-branching bisimulation. This result, formally stated in Proposition 3, ensures that probabilistic semi-branching bisimilarity is transitive. Being it reflexive and symmetric by definition, we will conclude that it is an equivalence relation (Theorem 1).

In order to show that the probabilistic semi-branching bisimulations over the PTS model are closed under relation composition, we need several preliminary results. Essentially, we need to show that: (i) probabilistic semi-branching bisimulations are preserved by relation lifting, formally stated in Lemma 3, (ii) processes related by a probabilistic semi-branching bisimulation are able to simulate arbitrary sequences of τ-transitions, formally stated in Lemma 4, and (iii) relation lifting distributes over relation composition, formally stated in Lemma 5

3.1. Relation lifting preserves probabilistic semi-branching bisimulations

In order to present Lemma 3, we need two preliminary results. The first states that the lifted transitions performed by a convex combination of distributions $\sum_{i\in I} p_i\pi_i$ are in bijection with the lifted transitions performed by those distributions π_i.

Lemma 1. *Assume that* $\pi = \sum_{i\in I} p_i\pi_i$ *and* $(\alpha) \in \{\alpha, \hat{\alpha}\}$. *Then,*

$$\pi \xrightarrow{(\alpha)}_\ell \pi' \iff \left[\begin{array}{l} \forall i \in I.\, \pi_i \xrightarrow{(\alpha)}_\ell \pi'_i \\ \pi' = \sum_{i\in I} p_i\pi'_i. \end{array}\right.$$

Proof. For each $i \in I$, π_i has the form $\pi_i = \sum_{j\in J_i} q_{i,j}\delta(s_{i,j})$ for a suitable set of indexes J_i and processes $s_{i,j}$. Therefore, $\pi = \sum_{i\in I} p_i \sum_{j\in J_i} q_{i,j}\delta(s_{i,j})$, which can be rewritten as $\pi = \sum_{i\in I, j\in J_i} p_i q_{i,j}\delta(s_{i,j})$. By Definition 4, we have that $\pi \xrightarrow{(\alpha)}_\ell \pi'$ if and only if $s_{i,j} \xrightarrow{(\alpha)}_\ell \pi_{i,j}$ for all $i \in I$ and $j \in J_i$, and $\pi' = \sum_{i\in I, j\in J_i} p_i q_{i,j}\pi_{i,j}$. By Definition 4, for each $i \in I$ we have that $s_{i,j} \xrightarrow{(\alpha)}_\ell \pi_{i,j}$ for all $j \in J_i$ is equivalent to having $\pi_i \xrightarrow{(\alpha)}_\ell \sum_{j\in J_i} q_{i,j}\pi_{i,j}$. Summarizing, we have shown that $\pi \xrightarrow{(\alpha)}_\ell \pi'$ if and only if $\pi_i \xrightarrow{(\alpha)}_\ell \sum_{j\in J_i} q_{i,j}\pi_{i,j}$ for all $i \in I$. If we name the distribution $\sum_{j\in J_i} q_{i,j}\pi_{i,j}$ as π'_i, then it remains to prove that $\pi' = \sum_{i\in I} p_i\pi'_i$, i.e., $\sum_{i\in I, j\in J_i} p_i q_{i,j}\pi_{i,j} = \sum_{i\in I} p_i \sum_{j\in J_i} q_{i,j}\pi_{i,j}$, which clearly holds. □

By exploiting Lemma 1, we can show that any sequence of lifted transitions $\xrightarrow{\hat{\varepsilon}}_\ell\xrightarrow{\hat{\alpha}}_\ell$ performed by a distribution π is in bijection with the same sequence performed by the processes in its support.

Lemma 2. *Assume a distribution* $\pi \in \Delta(\mathbf{S})$. *We have*

$$\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi' \xrightarrow{\hat{\alpha}}_\ell \pi'' \iff \left[\begin{array}{l} \forall s \in \mathsf{supp}(\pi).\, s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s \xrightarrow{\hat{\alpha}}_\ell \pi''_s \\ \pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s, \\ \pi'' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi''_s. \end{array}\right.$$

Proof. First we prove that

$$\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi' \iff \left[\begin{array}{l} s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s \text{ for all } s \in \mathsf{supp}(\pi) \\ \pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s. \end{array}\right. \qquad (1)$$

In order to prove (1), we show the two implications separately.

<u>CASE "$\Longrightarrow$"</u>. We proceed by induction over the number n of $\hat{\tau}$-transitions giving rise to $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'$.

<u>Base case: $n = 1$</u>. We have $\pi' = \pi$, therefore $\pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\delta(s)$. For all $s \in \mathsf{supp}(\pi)$, by definition of $\xrightarrow{\hat{\tau}}_\ell$, we have $s \xrightarrow{\hat{\tau}}_\ell \delta(s)$, which implies $s \xrightarrow{\hat{\varepsilon}}_\ell \delta(s)$, thus implying that the property $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s$ for all $s \in \mathsf{supp}(\pi)$ and $\pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s$ follows for $\pi'_s = \delta(s)$.

<u>Inductive step: $n > 1$</u>. The sequence $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'$ is obtained by sequences $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'''$ and $\pi''' \xrightarrow{\hat{\tau}}_\ell \pi'$, where $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'''$ consists in n $\hat{\tau}$-transitions. By the inductive hypothesis we have $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'''_s$ for all $s \in \mathsf{supp}(\pi)$ and $\pi''' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'''_s$. By Lemma 1, from $\pi''' \xrightarrow{\hat{\tau}}_\ell \pi'$ we infer $\pi'''_s \xrightarrow{\hat{\tau}}_\ell \pi'_s$ for all $s \in \mathsf{supp}(\pi)$, with $\pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s$. It remains to prove that $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s$ for all $s \in \mathsf{supp}(\pi)$. This follows by $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'''_s$ and $\pi'''_s \xrightarrow{\hat{\tau}}_\ell \pi'_s$.

<u>CASE "$\Longleftarrow$"</u>. Let n_s be the number of $\hat{\tau}$-transitions giving rise to $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s$. We proceed by induction over $n = \max_{s\in\mathsf{supp}(\pi)} n_s$.

<u>Base case: $n = 1$</u>. For all $s \in \mathsf{supp}(\pi)$ we have $s \xrightarrow{\hat{\tau}}_\ell \pi'_s$. By Definition 4 we infer $\pi \xrightarrow{\hat{\tau}}_\ell \pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi'_s$, thus giving $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'$.

<u>Inductive step: $n > 1$</u>. For all $s \in \mathsf{supp}(\pi)$ we have that the sequence $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s$ is obtained by sequences $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'''_s$ and $\pi'''_s \xrightarrow{\hat{\tau}}_\ell \pi'_s$, where each $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'''_s$ consists in at most n $\hat{\tau}$-transitions. By the inductive hypothesis, we derive that $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'''$ for $\pi''' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'''_s$. By Lemma 1, we have that $\pi'''_s \xrightarrow{\hat{\tau}}_\ell \pi'_s$ implies $\pi''' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'''_s \xrightarrow{\hat{\tau}}_\ell \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s = \pi'$, which, together with $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'''$ gives $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'$.

Now we exploit (1) to prove the thesis. We prove the two implications separately.

<u>CASE "$\Longrightarrow$"</u> Assume that $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi' \xrightarrow{\hat{\alpha}}_\ell \pi''$. By (1), $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'$ implies $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s$ for all $s \in \mathsf{supp}(\pi)$ and $\pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s$. By Lemma 1, $\pi' \xrightarrow{\hat{\alpha}}_\ell \pi''$ implies $\pi'_s \xrightarrow{\hat{\alpha}}_\ell \pi''_s$ with $\pi'' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi''_s$. Summarising, $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s$ and $\pi'_s \xrightarrow{\hat{\alpha}}_\ell \pi''_s$, $\pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s$ and $\pi'' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi''_s$, as required.

<u>CASE '$\Longleftarrow$"</u> Assume that $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s \xrightarrow{\hat{\alpha}}_\ell \pi''_s$ for all $s \in \mathsf{supp}(\pi)$, $\pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s$, $\pi'' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi''_s$. The first two properties imply that $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'$, by (1). Then, by Lemma 1, $\pi'_s \xrightarrow{\hat{\alpha}}_\ell \pi''_s$ implies that we have $\pi' = \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi'_s \xrightarrow{\hat{\alpha}}_\ell \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi''_s$. Summarizing, $\pi \xrightarrow{\hat{\varepsilon}}_\ell \pi'$ and $\pi' \xrightarrow{\hat{\alpha}}_\ell \sum_{s\in\mathsf{supp}(\pi)} \pi(s)\pi''_s = \pi''$, as required. □

Lemma 2 allows us to show that semi-branching bisimulations are preserved by the relation lifting. Before we adapt the notion of semi-branching bisimulation to distributions.

Definition 7. A binary relation $\mathcal{S} \subseteq \Delta(\mathbf{S}) \times \Delta(\mathbf{S})$ is a semi-branching bisimulation if whenever $\pi_1 \;\mathcal{S}\; \pi_2$ then

- whenever $\pi_1 \xrightarrow{\alpha}_\ell \pi'_1$ then $\pi_2 \xrightarrow{\hat{\varepsilon}}_\ell \pi''_2 \xrightarrow{\hat{\alpha}}_\ell \pi'_2$ with $\pi_1 \;\mathcal{S}\; \pi''_2$ and $\pi'_1 \;\mathcal{S}\; \pi'_2$
- whenever $\pi_2 \xrightarrow{\alpha}_\ell \pi'_2$ then $\pi_1 \xrightarrow{\hat{\varepsilon}}_\ell \pi''_1 \xrightarrow{\hat{\alpha}}_\ell \pi'_1$ with $\pi''_1 \;\mathcal{S}\; \pi_2$ and $\pi'_1 \;\mathcal{S}\; \pi'_2$.

Clearly, also semi-branching bisimulations over distributions are closed under union and the union of all semi-branching bisimulations is the greatest one.

Lemma 3. *If a binary relation* $\mathcal{S} \subseteq \mathbf{S} \times \mathbf{S}$ *is a semi-branching bisimulation, then so is* $\mathcal{S}^\ell \subseteq \Delta(\mathbf{S}) \times \Delta(\mathbf{S})$.

Proof. We need to show that whenever $\pi_1 \, \mathcal{S}^\ell \, \pi_2$ then

$$\pi_1 \xrightarrow{\alpha}_\ell \pi_1' \text{ implies } \left[\begin{array}{l} \pi_2 \xrightarrow{\hat{\varepsilon}}_\ell \pi_2'' \xrightarrow{\hat{\alpha}}_\ell \pi_2' \text{ with} \\ \pi_1 \, \mathcal{S}^\ell \, \pi_2'' \wedge \pi_1' \, \mathcal{S}^\ell \, \pi_2' \end{array}\right. \quad (2)$$

$$\pi_2 \xrightarrow{\alpha}_\ell \pi_2' \text{ implies } \left[\begin{array}{l} \pi_1 \xrightarrow{\hat{\varepsilon}}_\ell \pi_1'' \xrightarrow{\hat{\alpha}}_\ell \pi_1' \text{ with} \\ \pi_1'' \, \mathcal{S}^\ell \, \pi_2 \wedge \pi_1' \, \mathcal{S}^\ell \, \pi_2'. \end{array}\right. \quad (3)$$

We expand only the proof of Equation (2), since the other one can be obtained by applying a symmetric argument. First of all, we recall that by Definition 3, $\pi_1 \, \mathcal{S}^\ell \, \pi_2$ implies the existence of a matching $\mathfrak{w} \in \mathfrak{W}(\pi_1, \pi_2)$ such that whenever $\mathfrak{w}(s,t) > 0$ then $s \, \mathcal{S} \, t$. By Proposition 2 this is equivalent to state that there are a finite set of indexes I and weights $p_i \in (0,1]$ such that $\pi_1 = \sum_{i \in I} p_i \delta(s_i)$, $\pi_2 = \sum_{i \in I} p_i \delta(t_i)$ and $s_i \, \mathcal{S} \, t_i$ for all $i \in I$. Assume that $\pi_1 \xrightarrow{\alpha}_\ell \pi_1'$. By Definition 4 (last case) this implies that for all $i \in I$, $s_i \xrightarrow{\alpha}_\ell \pi_i$ and $\pi_1' = \sum_{i \in I} p_i \pi_i$. From $s_i \, \mathcal{S} \, t_i$, we infer $t_i \xrightarrow{\hat{\varepsilon}}_\ell \pi_i'' \xrightarrow{\hat{\alpha}}_\ell \pi_i'$ with $\delta(s_i) \, \mathcal{S}^\ell \, \pi_i''$ and $\pi_i \, \mathcal{S}^\ell \, \pi_i'$. As this holds for all $i \in I$, since $\pi_2 = \sum_{i \in I} p_i \delta(t_i)$, by Lemma 2 we get $\pi_2 \xrightarrow{\hat{\varepsilon}}_\ell \pi_2'' \xrightarrow{\hat{\alpha}}_\ell \pi_2'$ with $\pi_2'' = \sum_{i \in I} p_i \pi_i''$ and $\pi_2 = \sum_{i \in I} p_i \pi_i'$. To conclude, it remains to show that $\pi_1 \, \mathcal{S}^\ell \, \pi_2''$ and $\pi_1' \, \mathcal{S}^\ell \, \pi_2'$. Both relations follow by Proposition 1. In fact, for the former it is enough to notice that $\pi_1 = \sum_{i \in I} p_i \delta(s_i)$, $\pi_2'' = \sum_{i \in I} p_i \pi_i''$ and $\delta(s_i) \, \mathcal{S}^\ell \, \pi_i''$ for all $i \in I$. For the latter, we have $\pi_1' = \sum_{i \in I} p_i \pi_i$, $\pi_2' = \sum_{i \in I} p_i \pi_i'$ and $\pi_i \, \mathcal{S}^\ell \, \pi_i'$ for all $i \in I$. □

3.2. Simulation of sequences of silent transitions

The next lemma exploits Lemma 3 and generalizes the transfer condition of semi-branching bisimilarity to sequences of silent transitions. In its proof, we use $s \to^n \pi$ as a shorthand for $s \xrightarrow{\hat{\tau}}_\ell \pi_1 \xrightarrow{\hat{\tau}}_\ell \ldots \xrightarrow{\hat{\tau}}_\ell \pi_n = \pi$.

Lemma 4. *Assume that* $s \, \mathcal{S} \, t$ *for some probabilistic semi-branching bisimulation* $\mathcal{S}$. *Then:*

1. *Whenever* $s \xrightarrow{\hat{\varepsilon}}_\ell \pi_s$, *then* $t \xrightarrow{\hat{\varepsilon}}_\ell \pi_t$ *with* $\pi_s \, \mathcal{S}^\ell \, \pi_t$.
2. *Whenever* $t \xrightarrow{\hat{\varepsilon}}_\ell \pi_t$, *then* $s \xrightarrow{\hat{\varepsilon}}_\ell \pi_s$ *with* $\pi_s \, \mathcal{S}^\ell \, \pi_t$.

Proof. We expand only the proof of the first item. The proof for the second item can be obtained by a symmetric argument. We proceed by induction over the length $n \in \mathbb{N}$ of the sequence of transitions $s \xrightarrow{\hat{\varepsilon}}_\ell \pi_s$.

Base case: $n = 1$. In this case $s \xrightarrow{\hat{\varepsilon}}_\ell \pi_s$ derives from $s \xrightarrow{\hat{\tau}}_\ell \pi_s$. Being $\mathcal{S}$ a semi-branching bisimulation, we have that $t \xrightarrow{\hat{\varepsilon}}_\ell \pi \xrightarrow{\hat{\tau}}_\ell \pi_t$, thus giving $t \xrightarrow{\hat{\varepsilon}}_\ell \pi_t$, with $\delta(s) \, \mathcal{S}^\ell \, \pi$ and $\pi_s \, \mathcal{S}^\ell \, \pi_t$.

Inductive step: $n > 1$. Assume that $s \to^n \pi_s$. Clearly, this is equivalent to $s \to^{n-1} \pi_s' \xrightarrow{\hat{\tau}}_\ell \pi_s$. By the inductive hypothesis, there is π_t' such that $t \xrightarrow{\hat{\varepsilon}}_\ell \pi_t'$ and $\pi_s' \, \mathcal{S}^\ell \, \pi_t'$ (notice that the length of the sequence of transitions taken by t does not depend on n). Since, by Lemma 3, $\mathcal{S}^\ell$ is a semi-branching bisimulation, from $\pi_s' \xrightarrow{\hat{\tau}}_\ell \pi_s$ we infer that $\pi_t' \xrightarrow{\hat{\varepsilon}}_\ell \pi_t'' \xrightarrow{\hat{\tau}}_\ell \pi_t$ for some $\pi_s' \, \mathcal{S}^\ell \, \pi_t''$ and $\pi_s \, \mathcal{S}^\ell \, \pi_t$. Thus, we can concatenate the two sequences of lifted transitions and obtain $t \xrightarrow{\hat{\varepsilon}}_\ell \pi_t$ with $\pi_s \, \mathcal{S}^\ell \, \pi_t$. □

3.3. Relation lifting distributes over relation composition

Next we show that relation lifting distributes over relation composition.

Lemma 5. *Let* $\mathcal{R}_1, \mathcal{R}_2$ *be two relations over processes. Then* $\mathcal{R}_1^\ell \circ \mathcal{R}_2^\ell = (\mathcal{R}_1 \circ \mathcal{R}_2)^\ell$.

Proof. We prove first the inclusion $\mathcal{R}_1^\ell \circ \mathcal{R}_2^\ell \subseteq (\mathcal{R}_1 \circ \mathcal{R}_2)^\ell$.

Assume $\pi_1 (\mathcal{R}_1^\ell \circ \mathcal{R}_2^\ell) \pi_3$. Thus, there is a distribution π_2 such that $\pi_1 \, \mathcal{R}_1^\ell \, \pi_2$ and $\pi_2 \, \mathcal{R}_2^\ell \, \pi_3$. Let $\mathfrak{w}_{12} \in \mathfrak{W}(\pi_1, \pi_2)$ and $\mathfrak{w}_{23} \in \mathfrak{W}(\pi_2, \pi_3)$ be the matchings obtained from such relations according to Definition 3. We define

$$\mathfrak{w}_{13}(s,u) = \sum_{t \in \mathsf{supp}(\pi_2)} \frac{\mathfrak{w}_{12}(s,t) \cdot \mathfrak{w}_{23}(t,u)}{\pi_2(t)}.$$

It is immediate to verify that $\mathfrak{w}_{13}$ is a well defined matching for π_1 and π_3. Moreover, $\mathfrak{w}_{13}(s,u) > 0$ if and only if $\mathfrak{w}_{12}(s,t) > 0$ and $\mathfrak{w}_{23}(t,u) > 0$ for some process t. By the choices of $\mathfrak{w}_{12}, \mathfrak{w}_{23}$, this implies $s \, \mathcal{R}_1 t$ and $t \, \mathcal{R}_2 u$, thus giving $s \, (\mathcal{R}_1 \circ \mathcal{R}_2) \, u$ whenever $\mathfrak{w}_{13}(s,u) > 0$. Therefore, we can conclude that $\pi_1 \, (\mathcal{R}_1 \circ \mathcal{R}_2)^\ell \, \pi_3$.

Next, we prove the inclusion $\mathcal{R}_1^\ell \circ \mathcal{R}_2^\ell \supseteq (\mathcal{R}_1 \circ \mathcal{R}_2)^\ell$.

Assume $\pi_1 (\mathcal{R}_1 \circ \mathcal{R}_2)^\ell \pi_3$. Then, by Definition 3, there is a matching $\mathfrak{w}_{13} \in \mathfrak{W}(\pi_1, \pi_3)$ such that whenever $\mathfrak{w}_{13}(s,u) > 0$ then $s \, (\mathcal{R}_1 \circ \mathcal{R}_2) \, u$. This implies that for such pairs s, u there is (at least) one process t such that $s \, \mathcal{R}_1 \, t$ and $t \, \mathcal{R}_2 \, u$. For each pair s, u we choose one of such processes t, which we will denote by $t_{s,u}$. We need to exhibit a distribution π_2 such that $\pi_1 \, \mathcal{R}_1^\ell \, \pi_2$ and $\pi_2 \, \mathcal{R}_2^\ell \, \pi_3$. Clearly, to this end, the support of π_2 should be constituted by the processes $t_{s,u}$. Hence, we define

$$\pi_2(t) = \begin{cases} \mathfrak{w}_{13}(s,u) & \text{if } t = t_{s,u} \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathfrak{w}_{13} \in \mathfrak{W}(\pi_1, \pi_3)$, we get

$$\pi_1 = \sum_{s \in \mathsf{supp}(\pi_1), u \in \mathsf{supp}(\pi_3)} \mathfrak{w}_{13}(s,u) \cdot \delta(s)$$

$$\pi_2 = \sum_{s \in \mathsf{supp}(\pi_1), \, u \in \mathsf{supp}(\pi_3)} \mathfrak{w}_{13}(s,u) \cdot \delta(t_{s,u})$$

$$\pi_3 = \sum_{s \in \mathsf{supp}(\pi_1), \, u \in \mathsf{supp}(\pi_3)} \mathfrak{w}_{13}(s,u) \cdot \delta(u).$$

Let I be the set $I = \{(s,u) \mid \mathfrak{w}_{13}(s,u) > 0\}$. For each $i \in I$, we let $s_i = s$ whenever $i = (s,u)$ for some u. Similarly we let $u_i = u$ whenever $i = (s,u)$ for some s. Moreover, we let $t_i = t_{s,u}$ whenever $i = (s,u)$. Finally, we define $p_i = \mathfrak{w}_{13}(s,u)$ whenever $i = (s,u)$. Then we get

$$\pi_1 = \sum_{i\in I} p_i\delta(s_i) \qquad \pi_2 = \sum_{i\in I} p_i\delta(t_i) \qquad \pi_3 = \sum_{i\in I} p_i\delta(u_i)$$

with $s_i \,\mathcal{R}_1\, t_i$ and $t_i \,\mathcal{R}_2\, u_i$ for all $i \in I$. By Proposition 2 we can conclude that $\pi_1 \,\mathcal{R}_1^\ell\, \pi_2$ and $\pi_2 \,\mathcal{R}_2^\ell\, \pi_3$. ☐

3.4. *Probabilistic semi-branching bisimilarity is an equivalence relation*

By exploiting Lemma 3, Lemma 4 and Lemma 5, we can now prove that the relation composition of two probabilistic semi-branching bisimulations is again a probabilistic semi-branching bisimulation.

Proposition 3. *Assume two probabilistic semi-branching bisimulations $\mathcal{S}_1$, $\mathcal{S}_2 \subseteq \mathbf{S}\times\mathbf{S}$. Then, their relation composition $\mathcal{S}_1 \circ \mathcal{S}_2$ is a probabilistic semi-branching bisimulation.*

Proof. Assume $s\,(\mathcal{S}_1 \circ \mathcal{S}_2)\,u$. Thus, there is some process t for which we have $s\ \mathcal{S}_1 t$ and $t\ \mathcal{S}_2 u$. We prove that whenever $s \xrightarrow{\alpha}_\ell \pi_s$, there are distributions π_u' and π_u such that $u \xrightarrow{\hat\varepsilon}_\ell \pi_u' \xrightarrow{\hat\alpha}_\ell \pi_u$, $\delta(s)\,(\,\mathcal{S}_1 \circ\ \mathcal{S}_2)^\ell\,\pi_u'$ and $\pi_s\,(\,\mathcal{S}_1 \circ \mathcal{S}_2)^\ell\,\pi_u$. The symmetric case can be proved by analogous arguments. Since $s\ \mathcal{S}_1\, t$, then there are distributions π_t', π_t such that $t \xrightarrow{\hat\varepsilon}_\ell \pi_t' \xrightarrow{\hat\alpha}_\ell \pi_t$, $\delta(s)\ \mathcal{S}_1^\ell\,\pi_t'$ and $\pi_s\ \mathcal{S}_1^\ell\,\pi_t$. Then, as $t\,\mathcal{S}_2 u$, by Lemma 4 $t \xrightarrow{\hat\varepsilon}_\ell \pi_t'$ implies $u \xrightarrow{\hat\varepsilon}_\ell \pi_u''$ with $\pi_t'\ \mathcal{S}_2^\ell\,\pi_u''$. By Lemma 3, since $\mathcal{S}_2$ is a semi-branching bisimulation then also $\mathcal{S}_2^\ell$ is a semi-branching bisimulation. Therefore, from $\pi_t'\ \mathcal{S}_2^\ell\,\pi_u''$ and $\pi_t' \xrightarrow{\hat\alpha}_\ell \pi_t$ we get $\pi_u'' \xrightarrow{\hat\varepsilon}_\ell \pi_u' \xrightarrow{\hat\alpha}_\ell \pi_u$ with $\pi_t'\,\mathcal{S}_2^\ell\,\pi_u'$ and $\pi_t\ \mathcal{S}_2^\ell\,\pi_u$. We have therefore obtained that $u \xrightarrow{\hat\varepsilon}_\ell \pi_u' \xrightarrow{\hat\alpha}_\ell \pi_u$ with $\delta(s)\,(\mathcal{S}_1^\ell \circ\ \mathcal{S}_2^\ell)\,\pi_u'$ and $\pi_s\,(\,\mathcal{S}_1^\ell \circ\ \mathcal{S}_2^\ell)\,\pi_u$. Finally, by applying Lemma 5 we can conclude that $\delta(s)\,(\mathcal{S}_1 \circ \mathcal{S}_2)^\ell\,\pi_u'$ and $\pi_s\,(\mathcal{S}_1 \circ \mathcal{S}_2)^\ell\,\pi_u$, thus obtaining the thesis. ☐

We finally have all the ingredients necessary to state that semi-branching bisimilarity is an equivalence relation.

Theorem 1. *Probabilistic semi-branching bisimilarity is an equivalence relation over PTSs.*

Proof. Reflexivity and symmetry follow from Definition 6. Transitivity follows from Proposition 3. ☐

4. Probabilistic branching bisimilarity is an equivalence on divergence-free PTSs

We dedicate the rest of the paper to lift Theorem 1 to probabilistic branching bisimilarity. The result is proved for divergence-free PTSs. Once again, we let [5] guide us.

Firstly, we need to relate probabilistic semi-branching bisimilarity and probabilistic branching bisimilarity. In the fully nondeterministic case it was argued in [5, 19, 20] that each semi-branching bisimulation satisfying the *stuttering property* is also a branching bisimulation. Therefore, we need to introduce the equivalent to the *stuttering lemma* of [19, 20] for the probabilistic case.

Definition 8 (Probabilistic stuttering property)**.** A binary relation $\mathcal{R} \subseteq \mathbf{S}\times\mathbf{S}$ has the *probabilistic stuttering property* if and only if for all distributions $\pi_1,\ldots,\pi_n \in \Delta(\mathbf{S})$ it holds that:

- whenever $s\,\mathcal{R}\,t$, $t \xrightarrow{\hat\tau}_\ell \pi_1 \xrightarrow{\hat\tau}_\ell \ldots \xrightarrow{\hat\tau}_\ell \pi_n$ and $\delta(s)\,\mathcal{R}^\ell\,\pi_n$, then $\delta(s)\,\mathcal{R}^\ell\,\pi_i$ for all $i = 1,\ldots,n-1$;
- whenever $s\,\mathcal{R}\,t$, $s \xrightarrow{\hat\tau}_\ell \pi_1 \xrightarrow{\hat\tau}_\ell \ldots \xrightarrow{\hat\tau}_\ell \pi_n$ and $\pi_n\,\mathcal{R}^\ell\,\delta(t)$, then $\pi_i\,\mathcal{R}^\ell\,\delta(t)$ for all $i = 1,\ldots,n-1$.

Clearly, any probabilistic semi-branching bisimulation $\mathcal{R}$ with the probabilistic stuttering property, is a probabilistic branching bisimulation. Informally, given processes $s\,\mathcal{R}\,t$, the definition of probabilistic semi-branching bisimulation admits that a transition $s \xrightarrow{\tau} \pi_s$ is mimicked by a transition $t \xrightarrow{\hat\varepsilon}_\ell \pi_t$ with $\delta(s)\,\mathcal{R}^\ell\,\pi_t$ and $\pi_s\,\mathcal{R}^\ell\,\pi_t$, which is not enough, in general, to conclude that $\mathcal{R}$ is a probabilistic branching bisimulation. Assume that $t \xrightarrow{\hat\varepsilon}_\ell \pi_t$ is derived from $t \xrightarrow{\hat\varepsilon}_\ell \pi_t'$ and $\pi_t' \xrightarrow{\hat\tau}_\ell \pi_t$. The stuttering property of $\mathcal{R}$ together with $s\,\mathcal{R}\,t$ and $\delta(s)\,\mathcal{R}^\ell\,\pi_t$ imply $\delta(s)\,\mathcal{R}^\ell\,\pi_t'$. Summarising, we have $t \xrightarrow{\hat\varepsilon}_\ell \pi_t' \xrightarrow{\hat\tau}_\ell \pi_t$ with $\delta(s)\,\mathcal{R}^\ell\,\pi_t'$ and $\pi_s\,\mathcal{R}^\ell\,\pi_t$, which is a correct way to mimic $s \xrightarrow{\tau} \pi_s$ according to the definition of probabilistic branching bisimulation.

Hence, the next step consists in showing that probabilistic semi-branching bisimilarity satisfies the probabilistic stuttering property, thus allowing us to infer that it coincides with probabilistic branching bisimilarity. This result requires that PTSs are divergence-free.

To this purpose, we consider the equivalent formulation to the stuttering lemma proposed in [10] (as Lemma 2.3.2).

Definition 9 (Probabilistic stuttering property)**.** An equivalence relation $\mathcal{R} \subseteq \mathbf{S}\times\mathbf{S}$ has the *probabilistic stuttering property* if and only if $s \xrightarrow{\hat\tau}_\ell \pi_1 \xrightarrow{\hat\tau}_\ell \ldots \xrightarrow{\hat\tau}_\ell \pi_n$ for some $n \in \mathbb{N}$ and $\pi_1,\ldots,\pi_n \in \Delta(\mathbf{S})$, then $\delta(s)\,\mathcal{R}^\ell\,\pi_n$ implies $\delta(s)\,\mathcal{R}^\ell\,\pi_i$ for all $i = 1,\ldots,n-1$.

We prove that for equivalence relations, Definition 8 and Definition 9 are equivalent.

Proposition 4. *For an equivalence relation $\mathcal{R}$, Definition 8 and Definition 9 are equivalent.*

Proof. Assume first that $\mathcal{R}$ satisfies Definition 8. We have to prove that $\mathcal{R}$ satisfies Definition 9, namely $s \xrightarrow{\hat\tau}_\ell \pi_1 \xrightarrow{\hat\tau}_\ell \ldots \xrightarrow{\hat\tau}_\ell \pi_n$ and $\delta(s)\,\mathcal{R}^\ell\,\pi_n$ imply $\delta(s)\,\mathcal{R}^\ell\,\pi_i$ for all $i = 1,\ldots,n-1$. Since $\mathcal{R}$ is an equivalence, we have $s\,\mathcal{R}\,s$. Therefore, by Definition 8, from $s\,\mathcal{R}\,s$, $s \xrightarrow{\hat\tau}_\ell \pi_1 \xrightarrow{\hat\tau}_\ell \ldots \xrightarrow{\hat\tau}_\ell \pi_n$ and $\delta(s)\,\mathcal{R}^\ell\,\pi_n$ we infer $\delta(s)\,\mathcal{R}^\ell\,\pi_i$ for all $i = 1,\ldots,n-1$, as required.

Assume now that $\mathcal{R}$ satisfies Definition 9. We have to prove that $\mathcal{R}$ satisfies Definition 8, namely $s\,\mathcal{R}\,t$, $t \xrightarrow{\hat{\tau}}_\ell \pi_1 \xrightarrow{\hat{\tau}}_\ell \ldots \xrightarrow{\hat{\tau}}_\ell \pi_n$ and $\delta(s)\,\mathcal{R}^\ell\,\pi_n$ imply $\delta(s)\,\mathcal{R}^\ell\,\pi_i$ for all $i = 1,\ldots,n-1$. Since $\mathcal{R}$ is an equivalence relation, if $s\,\mathcal{R}\,t$ and $\delta(s)\,\mathcal{R}^\ell\,\pi_n$ we get $\delta(t)\,\mathcal{R}^\ell\,\pi_n$. Since $\mathcal{R}$ satisfies Definition 9, from $t \xrightarrow{\hat{\tau}}_\ell \pi_1 \xrightarrow{\hat{\tau}}_\ell \ldots \xrightarrow{\hat{\tau}}_\ell \pi_n$ and $\delta(t)\,\mathcal{R}^\ell\,\pi_n$ we get $\delta(t)\,\mathcal{R}^\ell\,\pi_i$ for all $i = 1,\ldots,n-1$. Since $\mathcal{R}$ is an equivalence and $s\,\mathcal{R}\,t$ we get $\delta(s)\,\mathcal{R}^\ell\,\pi_i$ for all $i = 1,\ldots,n-1$, as required. □

Since semi-branching bisimilarity is an equivalence relation (Theorem 1), to show that it satisfies the stuttering property it is enough to show that it satisfies the constraints in Definition 9.

Proposition 5. *On divergence-free PTSs, semi-branching bisimilarity satisfies the stuttering property.*

Proof. Assume that $s \xrightarrow{\hat{\tau}}_\ell \pi_1 \xrightarrow{\hat{\tau}}_\ell \ldots \xrightarrow{\hat{\tau}}_\ell \pi_n$, for some $n \in \mathbb{N}$ and $\pi_1,\ldots,\pi_n \in \Delta(\mathbf{S})$ with $\delta(s) \approx^\ell_{\mathrm{sb}} \pi_n$. Each π_i is of the form $\pi_i = \sum_{j\in J_i} p_j \delta(s_j)$ for some finite set of indexes J_i and weights $p_j \in (0,1]$ with $\sum_{j\in J_i} p_j = 1$. We aim at proving that $\delta(s) \approx^\ell_{\mathrm{sb}} \pi_i$ for all $i \in \{1,\ldots,n-1\}$. Due to Definition 3, this is equivalent to show that $s \approx_{\mathrm{sb}} s_j$ for all $j \in J_i$, $i \in \{1,\ldots,n-1\}$. According to Definition 4, for each $i \in \{1,\ldots,n-1\}$ and $j \in J_i$, we can distinguish two cases:

1. $s_j \in \mathsf{supp}(\pi_n)$. In this case, $s \approx_{\mathrm{sb}} s_j$ directly follows from $\delta(s) \approx^\ell_{\mathrm{sb}} \pi_n$, which, by Definition 3, implies $s \approx_{\mathrm{sb}} s'$ for all $s' \in \mathsf{supp}(\pi_n)$.

2. $s_j \notin \mathsf{supp}(\pi_n)$, which means that s_j performs at least one τ-move in the sequence $\pi_i \xrightarrow{\hat{\varepsilon}}_\ell \pi_n$. Let us prove that $s \approx_{\mathrm{sb}} s_j$ by verifying the two constraints of Definition 6.

   Assume first that $s \xrightarrow{\alpha} \pi_s$. Since $\delta(s) \approx^\ell_{\mathrm{sb}} \pi_n$ and by Lemma 3 $\approx^\ell_{\mathrm{sb}}$ is a semi-branching bisimulation, we infer that $\pi_n \xrightarrow{\hat{\varepsilon}}_\ell \pi'' \xrightarrow{\hat{\alpha}}_\ell \pi'$ with $\delta(s) \approx^\ell_{\mathrm{sb}} \pi''$ and $\pi_s \approx^\ell_{\mathrm{sb}} \pi'$. Since $\pi_i \xrightarrow{\hat{\varepsilon}}_\ell \pi_n$, we infer $\pi_i \xrightarrow{\hat{\varepsilon}}_\ell \pi'' \xrightarrow{\hat{\alpha}}_\ell \pi'$. By Definition 4, $\pi_i \xrightarrow{\hat{\varepsilon}}_\ell \pi''$ implies that, for each $j \in J_i$, $s_j \xrightarrow{\hat{\varepsilon}}_\ell \pi''_j$ and $\pi'' = \sum_{j\in J_i} p_j \pi''_j$. Then $\delta(s) \approx^\ell_{\mathrm{sb}} \pi''$ gives $\sum_{j\in J_i} p_j \delta(s) \approx^\ell_{\mathrm{sb}} \sum_{j\in J_i} p_j \pi''_j$, namely $\delta(s) \approx^\ell_{\mathrm{sb}} \pi''_j$ for all $j \in J_i$. Therefore, $\delta(s) \xrightarrow{\alpha}_\ell \pi_s$ implies $\pi''_j \xrightarrow{\hat{\varepsilon}}_\ell \pi'_j \xrightarrow{\hat{\alpha}}_\ell \pi_j$ with $\delta(s) \approx^\ell_{\mathrm{sb}} \pi'_j$ and $\pi_s \approx^\ell_{\mathrm{sb}} \pi_j$. We have therefore obtained that whenever $s \xrightarrow{\alpha} \pi_s$ then $s_j \xrightarrow{\hat{\varepsilon}}_\ell \pi'_j \xrightarrow{\hat{\alpha}}_\ell \pi_j$ with $\delta(s) \approx^\ell_{\mathrm{sb}} \pi'_j$ and $\pi_s \approx^\ell_{\mathrm{sb}} \pi_j$.

   Assume now that $s_j \xrightarrow{\alpha} \pi_j$. Suppose by contradiction that there is no pair of distributions π'_s, π_s with $s \xrightarrow{\hat{\varepsilon}}_\ell \pi'_s \xrightarrow{\hat{\alpha}}_\ell \pi_s$, $\pi'_s \approx^\ell_{\mathrm{sb}} \delta(s_j)$ and $\pi_s \approx^\ell_{\mathrm{sb}} \pi_j$, namely that $s \not\approx_{\mathrm{sb}} s_j$. We show that this gives a contradiction with s being divergence-free. In fact, $s \not\approx_{\mathrm{sb}} s_j$ gives $\delta(s) \not\approx^\ell_{\mathrm{sb}} \pi_i$, as by Definition 3 we would have $\delta(s) \approx^\ell_{\mathrm{sb}} \pi_i$ if and only if $s \approx_{\mathrm{sb}} s_j$ for all $j \in J_i$. This implies that at least one τ-move is performed in the sequence $s \xrightarrow{\hat{\varepsilon}}_\ell \pi_i$. In particular, $s \xrightarrow{\tau} \pi'$ for some π' with $\delta(s) \not\approx^\ell_{\mathrm{sb}} \pi'$. Similarly, $\pi_i \xrightarrow{\hat{\varepsilon}}_\ell \pi_n$ and $\delta(s) \approx^\ell_{\mathrm{sb}} \pi_n$ imply that at least one τ-move is performed in this sequence, so that $\pi_i \xrightarrow{\hat{\varepsilon}}_\ell \pi'' \xrightarrow{\hat{\varepsilon}}_\ell \pi_n$ for some π'' with $\pi_i \not\approx^\ell_{\mathrm{sb}} \pi''$. Therefore, the sequence $s \xrightarrow{\hat{\varepsilon}}_\ell \pi_n$ contains at least two τ-moves that, as $\delta(s) \approx^\ell_{\mathrm{sb}} \pi_n$, must be mimicked by π_n and by its $\delta(s)$-equivalent derivatives, namely s can perform an infinite sequence of τ-moves, and it is therefore not divergence-free, thus giving the desired contradiction.

□

We can give an example showing that the divergence-free assumption is needed in order to have Proposition 5.

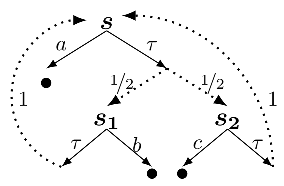

Figure 1

Let nil be the process that cannot perform any move and consider process s in Figure 1, where $\bullet$ represents the distribution $\delta(\text{nil})$. Clearly s has divergence, since by performing a τ-step it reaches the distribution $\pi = 1/2\,\delta(s_1) + 1/2\,\delta(s_2)$ that can in turn perform a τ-lifted transition and reach $\delta(s)$ (as $s_i \xrightarrow{\tau} \delta(s)$ for $i = 1, 2$). One can now notice that semi branching bisimilarity does not satisfy the stuttering property in Definition 9: for $n = 2$, we have $s \xrightarrow{\hat{\tau}}_\ell \pi \xrightarrow{\hat{\tau}}_\ell \delta(s)$ and clearly $\delta(s)\,\mathcal{S}^\ell\,\delta(s)$. However, $\delta(s)$ is not related to π by $\mathcal{S}^\ell$ as s is semi branching bisimilar to neither s_1 nor s_2. In fact, $s_1 \xrightarrow{b} \delta(\text{nil})$, but there is no distribution reachable from s via a sequence of τ-lifted transitions such that all the processes in its support can perform a b-move, and thus there are no π', π'' such that $s \xrightarrow{\hat{\varepsilon}}_\ell \pi' \xrightarrow{b}_\ell \pi''$. Similarly for the c-move by s_2.

By Proposition 5 we can directly infer that on divergence-free PTSs two processes are probabilistic branching bisimilar if and only if they are probabilistic semi-branching bisimilar.

Corollary 1. *On divergence-free PTSs, probabilistic branching bisimilarity coincides with probabilistic semi-branching bisimilarity.*

Proof. Firstly, we recall that $\approx_{\mathrm{b}}$ and $\approx_{\mathrm{sb}}$ are defined, respectively, as the largest branching bisimulation and the largest semi-branching bisimulation. By Proposition 5 we get that $\approx_{\mathrm{sb}}$ satisfies the stuttering property and it is therefore a branching bisimulation, giving $\approx_{\mathrm{sb}} \subseteq \approx_{\mathrm{b}}$. Conversely, by Definition 6 it immediately follows that $\approx_{\mathrm{b}}$ is a semi-branching bisimulation. Hence $\approx_{\mathrm{b}} \subseteq \approx_{\mathrm{sb}}$ holds. From the two inclusions, we derive $\approx_{\mathrm{b}} = \approx_{\mathrm{sb}}$ as required. □

As an immediate consequence, we obtain that probabilistic branching bisimilarity is an equivalence relation over nondeterministic probabilistic processes.

Theorem 2. *Probabilistic branching bisimilarity is an equivalence relation over divergence-free PTSs.*

Proof. The result is a direct consequence of Theorem 1 and Corollary 1. □

5. Remarks on divergence

We have proved that probabilistic branching bisimilarity is an equivalence over PTSs that are divergence-free.

Technically, since a probabilistic semi-branching bisimiulation satisfying the stuttering property is also a probabilistic branching bisimulation, in order to prove our result we have shown that probabilistic semi-branching bisimilarity is an equivalence (also when divergence is admitted, Theorem 1) and that on divergence-free PTSs semi-branching bisimilarity is stuttering (Proposition 5).

Then, we have provided an example of a PTS with divergence where semi-branching bisimilarity is not stuttering. The arguments we used are valid also for branching bisimilarity. Hence, the example shows that when divergence is admitted then neither probabilistic semi branching bisimilarity nor probabilistic branching bisimilarity are guaranteed to be stuttering. This is an interesting insight since in the non probabilistic case, both these equivalences are stuttering even if divergence is admitted. The example clearly shows the reason of this discrepancy between the probabilistic and the non-probabilistic case: in the former case we may have a loop like $\delta(s) \xrightarrow{\tau} \pi \xrightarrow{\tau} \delta(s)$ where the processes in the support of π, like s_1 and s_2, have different observable behaviour, thus implying that $\delta(s)$ and π cannot be related by any behavioral relation. This kind of situation cannot be replicated in the latter case.

The discrepancy between the probabilistic and the non-probabilistic case arises only on PTSs with divergence. Indeed, in the proof of Proposition 5 (second part of item 2) we have formally proved that a process on which probabilistic (semi)branching bisimilarity is not stuttering always has divergence. However, the converse implication does not hold in general, since there are cases of processes with divergence on which the relations are stuttering. Intuitively, if the divergence is only due to the presence of self-loops (the process has probability 1 to go back to itself with a single τ-move) then clearly the stuttering property is satisfied. Moreover, notice that since the probability of the loop is 1 we are also guaranteed that the infinite sequence of silent-moves would not modify the probabilities of observable events to be eventually observed (we refer the interested reader to [36] for further details). We conjecture that, modulo branching bisimilarity, this is the only case in which the stuttering property and divergence can coexist. However, a proof of this fact is beyond the scope of this paper.

Nonetheless, we would like to point out that in [35] no distinction between processes with or without divergence was made, but the definition of branching bisimilarity came with an additional constraint (which can be also found in, e.g, [36, 38]). More precisely, using our notation, in the definition of branching bisimulation the sequence of lifted transitions $t \xrightarrow{\hat{\varepsilon}}_{\ell} \pi \xrightarrow{\alpha}_{\ell} \pi_t$ with which process t mimics the transition $s \xrightarrow{\alpha} \pi_s$, has to satisfy the *branching condition*: either $\alpha = \tau$ and $\pi_t = \delta(t)$, or for all distributions π' reached in the lifted sequence $t \xrightarrow{\hat{\varepsilon}}_{\ell} \pi$ it holds that $\delta(t)\ \mathcal{B}^{\ell}\ \pi'$. In other words, branching bisimilarity is established only on sequences of silent steps inducing the stuttering property of the relation.

Finally, we remark that to prove that semi-branching bisimilarity is an equivalence, we have never used the assumption that processes are divergence-free. In many occasions, the relation defined as branching bisimilarity is actually the semi-branching bisimilarity. But to be formally correct, we have distinguished the two relations and used the stuttering lemma to relate them. As argued above, the stuttering property is sensible to divergence.

Acknowledgements: V. Castiglioni is supported by the project 'Open Problems in the Equational Logic of Processes' (OPEL) of the Icelandic Research Fund (grant nr. 196050-051).

References

[1] Andova, S., Georgievska, S., & Trcka, N. (2012). Branching bisimulation congruence for probabilistic systems. *TCS*, *413*, 58–72. doi:10.1016/j.tcs.2011.07.020.

[2] Andova, S., & Willemse, T. A. (2006). Branching bisimulation for probabilistic systems: characteristics and decidability. *Theor. Comput. Sci.*, *356*, 325–355. doi:10.1016/j.tcs.2006.02.010.

[3] Baier, C. (1998). *On Algorithmic Verification Methods for Probabilistic Systems*. Ph.D. thesis Fakultät für Mathematik and Informatik Uviversität Mannheim.

[4] Baier, C., & Hermanns, H. (1997). Weak bisimulation for fully probabilistic processes. In *Proceedings of CAV '97* (pp. 119–130). volume 1254 of *Lecture Notes in Computer Science*. doi:10.1007/3-540-63166-6_14.

[5] Basten, T. (1996). Branching bisimilarity is an equivalence indeed! *Inf. Process. Lett.*, *58*, 141–147. doi:10.1016/0020-0190(96)00034-8.

[6] Belder, T., ter Beek, M. H., & de Vink, E. P. (2015). Coherent branching feature bisimulation. In *Proc. FMSPLE@ETAPS 2015* (pp. 14–30). doi:10.4204/EPTCS.182.2.

[7] Bergstra, J. A., Ponse, A., & van der Zwaag, M. (2003). Branching time and orthogonal bisimulation equivalence. *Theor. Comput. Sci.*, *309*, 313–355. doi:10.1016/S0304-3975(03)00277-9.

[8] Castiglioni, V., Lanotte, R., & Tini, S. (2014). A specification format for rooted branching bisimulation. *Fundam. Inform.*, *135*, 355–369. doi:10.3233/FI-2014-1128.

[9] Castiglioni, V., & Tini, S. (2020). Probabilistic divide & congruence: Branching bisimilarity. *Theor. Comput. Sci.*, *802*, 147–196. doi:10.1016/j.tcs.2019.09.037.

[10] De Nicola, R., Montanari, U., & Vaandrager, F. W. (1990). Back and forth bisimulations. In *Proc. CONCUR '90* (pp. 152–165). volume 458 of *Lecture Notes in Computer Science*. doi:10.1007/BFb0039058.

[11] De Nicola, R., & Vaandrager, F. W. (1995). Three logics for branching bisimulation. *J. ACM*, *42*, 458–487. doi:10.1145/201019.201032.

[12] Deng, Y., & Du, W. (2011). Logical, metric, and algorithmic characterisations of probabilistic bisimulation. *CoRR*, *abs/1103.4577*.

[13] Deng, Y., & van Glabbeek, R. J. (2010). Characterising probabilistic processes logically - (extended abstract). In *Proc. LPAR-17* (pp. 278–293). volume 6397 of *LNCS*. doi:10.1007/978-3-642-16242-8_20.

[14] Deng, Y., van Glabbeek, R. J., Hennessy, M., & Morgan, C. (2008). Characterising testing preorders for finite probabilistic processes. *Logical Methods in Computer Science*, *4*. doi:10.2168/LMCS-4(4:4)2008.

[15] Fokkink, W. J. (2000). Rooted branching bisimulation as a congruence. *J. Comput. Syst. Sci.*, *60*, 13–37. doi:10.1006/jcss.1999.1663.

[16] Fokkink, W. J., van Glabbeek, R. J., & de Wind, P. (2012). Divide and congruence: From decomposition of modal formulas to preservation of branching and η-bisimilarity. *Inf. Comput.*, *214*, 59–85. doi:10.1016/j.ic.2011.10.011.

[17] van Glabbeek, R. J. (1993). The linear time - branching time spectrum II. In *Proc. CONCUR '93* (pp. 66–81). doi:10.1007/3-540-57208-2_6.

[18] van Glabbeek, R. J., Smolka, S. A., & Steffen, B. (1995). Reactive, generative and stratified models of probabilistic processes. *Inf. Comput.*, *121*, 59–80. doi:10.1006/inco.1995.1123.

[19] van Glabbeek, R. J., & Weijland, W. P. (1989). Branching time and abstraction in bisimulation semantics (extended abstract). In *IFIP Congress* (pp. 613–618).

[20] van Glabbeek, R. J., & Weijland, W. P. (1996). Branching time and abstraction in bisimulation semantics. *J. ACM*, *43*, 555–600. doi:10.1145/233551.233556.

[21] Groote, J. F., Jansen, D. N., Keiren, J. J. A., & Wijs, A. (2017). An $O(m \log n)$ algorithm for computing stuttering equivalence and branching bisimulation. *ACM Trans. Comput. Log.*, *18*, 13:1–13:34. doi:10.1145/3060140.

[22] Groote, J. F., & Vaandrager, F. W. (1990). An efficient algorithm for branching bisimulation and stuttering equivalence. In *Proc. ICALP '90* (pp. 626–638). volume 443 of *Lecture Notes in Computer Science*. doi:10.1007/BFb0032063.

[23] Hansson, H., & Jonsson, B. (1994). A logic for reasoning about time and reliability. *Formal Aspects of Computing*, *6*, 512–535. doi:10.1007/BF01211866.

[24] Hansson, H. A. (1992). Time and probabilities in specification and verification of real-time systems. In *Proc. RTS 1992* (pp. 92–97). doi:10.1109/EMWRT.1992.637477.

[25] Hennessy, M. (2012). Exploring probabilistic bisimulations, part I. *Formal Asp. Comput.*, *24*, 749–768. doi:10.1007/s00165-012-0242-7.

[26] Hennessy, M., & Regan, T. (1995). A Process Algebra for Timed Systems. *Information and Computation*, *117*, 221–239.

[27] Jou, C., & Smolka, S. A. (1990). Equivalences, congruences, and complete axiomatizations for probabilistic processes. In *Proceedings of CONCUR '90* (pp. 367–383). volume 458 of *Lecture Notes in Computer Science*. doi:10.1007/BFb0039071.

[28] Keller, R. M. (1976). Formal verification of parallel programs. *Commun. ACM*, *19*, 371–384. doi:10.1145/360248.360251.

[29] Lee, M. D., & de Vink, E. P. (2015). Rooted branching bisimulation as a congruence for probabilistic transition systems. In *Proc. QAPL 2015* (pp. 79–94). volume 194 of *EPTCS*. doi:10.4204/EPTCS.194.6.

[30] Lynch, N. A., Segala, R., & Vaandrager, F. W. (2007). Observing branching structure through probabilistic contexts. *SIAM J. Comput.*, *37*, 977–1013. doi:10.1147/S0097539704446487.

[31] Milner, R. (1980). *A Calculus of Communicating Systems* volume 92 of *Lecture Notes in Computer Science*. doi:10.1007/3-540-10235-3.

[32] Park, D. M. (1981). Concurrency and automata of infinite sequences. In 5^{th} *GI Conference* (pp. 167–183). volume 104 of *LNCS*. doi:10.1007/BFb0017309.

[33] Philippou, A., Lee, I., & Sokolsky, O. (2000). Weak bisimulation for probabilistic systems. In *Proc. CONCUR 2000* (pp. 334–349). volume 1877 of *Lecture Notes in Computer Science*. doi:10.1007/3-540-44618-4_25.

[34] Segala, R. (1995). *Modeling and Verification of Randomized Distributed Real-Time Systems*. Ph.D. thesis MIT.

[35] Segala, R., & Lynch, N. A. (1994). Probabilistic simulations for probabilistic processes. In *Proceedings of CONCUR '94* (pp. 481–496). volume 836 of *Lecture Notes in Computer Science*. doi:10.1007/978-3-540-48654-1_35.

[36] Timmer, M., Katoen, J., van de Pol, J., & Stoelinga, M. (2016). Confluence reduction for markov automata. *Theor. Comput. Sci.*, *655*, 193–219. doi:10.1016/j.tcs.2016.01.017.

[37] Trcka, N. (2009). Strong, weak and branching bisimulation for transition systems and Markov reward chains: A unifying matrix approach. In *Proc. QFM 2009* (pp. 55–65). volume 13 of *EPTCS*. doi:10.4204/EPTCS.13.5.

[38] Turrini, A., & Hermanns, H. (2015). Polynomial time decision algorithms for probabilistic automata. *Inf. Comput.*, *244*, 134–171. doi:10.1016/j.ic.2015.07.004.

[39] Yang, X., Katoen, J., Lin, H., Liu, G., & Wu, H. (2018). Branching bisimulation and concurrent object verification. In *Proc. DSN 2018* (pp. 267–278). doi:10.1109/DSN.2018.00037.