SOS rule formats for convex and abstract probabilistic bisimulations*

Pedro R. D'Argenio Matias David Lee
FaMAF, Universidad Nacional de Córdoba – CONICET.
Ciudad Universitaria, X5000HUA – Córdoba, Argentina.
dargenio@famaf.unc.edu.ar lee@famaf.unc.edu.ar

Daniel Gebler
Department of Computer Science, VU University Amsterdam,
De Boelelaan 1081a, NL-1081 HV Amsterdam, The Netherlands
e.d.gebler@vu.nl

Probabilistic transition system specifications (PTSSs) in the *ntμfθ/ntμxθ* format provide structural operational semantics for Segala-type systems that exhibit both probabilistic and nondeterministic behavior and guarantee that bisimilarity is a congruence for all operator defined in such format. Starting from the *ntμfθ/ntμxθ*, we obtain restricted formats that guarantee that three coarser bisimulation equivalences are congruences. We focus on (i) Segala's variant of bisimulation that considers combined transitions, which we call here *convex bisimulation*; (ii) the bisimulation equivalence resulting from considering Park & Milner's bisimulation on the usual stripped probabilistic transition system (translated into a labelled transition system), which we call here *probability obliterated bisimulation*; and (iii) a *probability abstracted bisimulation*, which, like bisimulation, preserves the structure of the distributions but instead, it ignores the probability values. In addition, we compare these bisimulation equivalences and provide a logic characterization for each of them.

1 Introduction

Structural operational semantics (SOS for short) [24] is a powerful tool to provide semantics to programming languages. In SOS, process behavior is described using transition systems and the behavior of a composite process is given in terms of the behavior of its components. SOS has been formalized using an algebraic framework as *Transition Systems Specifications (TSS)* [6, 7, 14, 15, 23, etc.]. Basically, a TSS contains a signature, a set of actions or labels, and a set of rules. The signature defines the terms in the language. The set of actions represents all possible activities that a process (i.e., a term over the signature) can perform. The rules define how a process should behave (i.e., perform certain activities) in terms of the behavior of its subprocesses, that is, the rules define compositionally the transition system associated to each term of the language. A particular focus of these formalizations was to provide a meta-theory that ensures a diversity of semantic properties by simple inspection on the form of the rules. (See [1, 2, 23] for overviews.) One of such kind of properties is to ensure that a given equivalence relation is a congruence for all operators whose semantics is defined in a TSS whose rules complies to a particular format. These so called *congruence theorems* have been proved for a variety of equivalences in the non-probabilistic case [6, 14, 15, etc.].

*Supported by ANPCyT project PICT-2012-1823, SeCyT-UNC program 05/BP12 and their related projects, and EU 7FP grant agreement 295261 (MEALS).

S. Crafa and D. Gebler (Eds.): Combined Workshop on Expressiveness in Concurrency and Structural Operational Semantics (EXPRESS/SOS 2015)
EPTCS 190, 2015, pp. 31–45, doi:10.4204/EPTCS.190.3

The introduction of probabilistic process algebras motivated the need for a theory of structural operational semantics to define *probabilistic* transition systems. Few earlier results appeared in this direction [4, 5, 17, 18] presenting congruence theorems for Larsen & Skou bisimulation equivalence [19]. Most of these formats have complicated restrictions that extend to sets of rules due to the fact that they considered transitions labeled both with an action and a probability value. By using a more modern view of probabilistic transition systems (where the target of the transition is a probability distribution on states) we manage to obtained the most general format for bisimulation equivalence, which we called *ntμfθ/ntμxθ*, following the nomenclature of [14, 15].

Starting from the *ntμfθ/ntμxθ* format, in this paper we define formats to guarantee that three coarser versions of bisimulation equivalence are congruences for all operator definable in the respective format. The first relation we focus on is Segala's variant of bisimulation that considers combined transitions, here called *convex bisimulation* [25]. The second relation we explore originates here and we call it *probability abstracted bisimulation*. Like bisimulation and unlike convex bisimulation, it preserves the structure of the distributions of each transition, but instead, it ignores the probability values. This relation preserves the fairness introduced by the probability distributions. Finally, we study the bisimulation equivalence resulting from considering Park & Milner's bisimulation [22] on the usual stripped probabilistic transition system (translated into a labeled transition system). Here we call it *probability obliterated bisimulation*. This is the usual way to abstract probabilities, but it has the drawback that it breaks the basic fairness provided by probabilistic choices.

Apart from presenting congruence theorems for all previously mentioned bisimulation equivalences, we briefly study alternative definitions of these bisimulations, compare them with each other, and provide logical characterizations, which are particularly new here for probability abstracted and probability obliterated bisimulation equivalences.

The paper is organized as follows. Sec. 2 recalls the type of algebraic structure and Sec. 3 provides the basic notions and results of probabilistic transition system specifications (PTSS). Sec. 4 presents the different bisimulation equivalences and a brief study of them, including their logical characterizations. The study of all the PTSS formats and the respective congruence theorems is given in Sec. 5. The paper concludes in Sec 6.

2 Preliminaries

Let $S = \{s,d\}$ be a set denoting two sorts. Elements of sort $s \in S$ are intended to represent states in the transition system, while elements of sort $d \in S$ will represent distributions over states. We let σ range over S. An S*-sorted signature* is a structure (F,ar), where (i) F is a set of *function names*, and (ii) $\mathsf{ar}\colon F \to (S^* \times S)$ is the *arity function*. The rank of $f \in F$ is the number of arguments of f, defined by $\mathsf{rk}(f) = n$ if $\mathsf{ar}(f) = \sigma_1 \dots \sigma_n \to \sigma$. (We write "$\sigma_1 \dots \sigma_n \to \sigma$" instead of "$(\sigma_1 \dots \sigma_n, \sigma)$" to highlight that function f maps to sort σ.) Function f is a *constant* if $\mathsf{rk}(f) = 0$. To simplify the presentation we will write an S-sorted signature (F,ar) as a pair of disjoint signatures (Σ_s, Σ_d) where Σ_s is the set of operations that map to s and Σ_d is the set of operations that map to d. Let $\mathcal{V}$ and $\mathcal{V}_d$ be two infinite sets of S-sorted variables where $\mathcal{V}, \mathcal{V}_d, F$ are all mutually disjoint. We use x,y,z (with possible sub- or super-scripts) to range over $\mathcal{V}$, μ,ν to range over $\mathcal{V}_d$ and ζ to range over $\mathcal{V} \cup \mathcal{V}_d$.

Definition 1. *Let Σ_s and Σ_d be two signatures as before and let $V \subseteq \mathcal{V}$ and $D \subseteq \mathcal{V}_d$. We simultaneously define the sets of* state terms $T(\Sigma_s, V, D)$ *and* distribution terms $T(\Sigma_d, V, D)$ *as the smallest sets satisfying: (i) $V \subseteq T(\Sigma_s, V, D)$; (ii) $D \subseteq T(\Sigma_d, V, D)$; (iii) $f(\xi_1, \cdots, \xi_{\mathsf{rk}(f)}) \in T(\Sigma_\sigma, V, D)$, if $\mathsf{ar}(f) = \sigma_1 \dots \sigma_n \to \sigma$ and $\xi_i \in T(\Sigma_{\sigma_i}, V, D)$.*

We let $\mathbb{T}(\Sigma) = T(\Sigma_s,\mathcal{V},\mathcal{V}_d) \cup T(\Sigma_d,\mathcal{V},\mathcal{V}_d)$ denote the set of all *open terms* and distinguish the sets $\mathbb{T}(\Sigma_s) = T(\Sigma_s,\mathcal{V},\mathcal{V}_d)$ of *open state terms* and $\mathbb{T}(\Sigma_d) = T(\Sigma_d,\mathcal{V},\mathcal{V}_d)$ of *open distribution terms*. Similarly, we let $\mathsf{T}(\Sigma) = T(\Sigma_s,\emptyset,\emptyset) \cup T(\Sigma_d,\emptyset,\emptyset)$ denote the set of all *closed terms* and distinguish the sets $\mathsf{T}(\Sigma_s) = T(\Sigma_s,\emptyset,\emptyset)$ of *closed state terms* and $\mathsf{T}(\Sigma_d) = T(\Sigma_d,\emptyset,\emptyset)$ of *closed distribution terms*. We let t, t', t_1,... range over state terms, θ, θ', θ_1,... range over distribution terms, and ξ, ξ', ξ_1,... range over any kind of terms. With $\mathcal{V}(\xi) \subseteq \mathcal{V} \cup \mathcal{V}_d$ we denote the set of variables occurring in term ξ.

Let $\Delta(\mathsf{T}(\Sigma_s))$ denote the set of all (discrete) probability distributions on $\mathsf{T}(\Sigma_s)$. We let π range over $\Delta(\mathsf{T}(\Sigma_s))$. For each $t \in \mathsf{T}(\Sigma_s)$, let $\delta_t \in \Delta(\mathsf{T}(\Sigma_s))$ denote the *Dirac distribution*, i.e., $\delta_t(t) = 1$ and $\delta_t(t') = 0$ if t and t' are not syntactically equal. For $X \subseteq \mathsf{T}(\Sigma_s)$ we define $\pi(X) = \sum_{t\in X} \pi(t)$. The convex combination $\sum_{i\in I} p_i\pi_i$ of a family $\{\pi_i\}_{i\in I}$ of probability distributions with $p_i \in (0,1]$ and $\sum_{i\in I} p_i = 1$ is defined by $(\sum_{i\in I} p_i\pi_i)(t) = \sum_{i\in I}(p_i\pi_i(t))$.

The type of signatures we consider has a particular construction. We start from a signature Σ_s of functions mapping into sort s and construct the signature Σ_d of functions mapping into d as follows. For each $f \in F_s$ we include a function symbol $\boldsymbol{f} \in F_d$ with $\mathsf{ar}(\boldsymbol{f}) = d\ldots d \to d$ and $\mathsf{rk}(\boldsymbol{f}) = \mathsf{rk}(f)$. We call $\boldsymbol{f}$ the *probabilistic lifting* of f. (We use boldface fonts to indicate that a function in Σ_d is the probabilistic lifting of another in Σ_s.) Moreover Σ_d may include any of the following additional operators: (i) δ with arity $\mathsf{ar}(\delta) = s \to d$, and (ii) $\bigoplus_{i\in I}[p_i]\_$ with I being a finite or countable infinite index set, $\sum_{i\in I} p_i = 1$, $p_i \in (0,1]$ for all $i \in I$, and $\mathsf{ar}\big(\bigoplus_{i\in I}[p_i]\_\big) = d^{|I|} \to d$. Notice that if I is countably infinite, $\bigoplus_{i\in I}[p_i]\_$ is an infinitary operator.

Operators δ and $\bigoplus_{i\in I}[p_i]\_$ are used to construct discrete probability functions of countable support: $\delta(t)$ is interpreted as a distribution that assigns probability 1 to the state term t and probability 0 to any other term t' (syntactically) different from t, and $\bigoplus_{i\in I}[p_i]\theta_i$ represents a distribution that weights with p_i the distribution represented by the term θ_i. Moreover, a probabilistically lifted operator $\boldsymbol{f}$ is interpreted by properly lifting the probabilities of the operands to terms composed with the operator f.

Formally, the algebra associated with a probabilistically lifted signature $\Sigma = (\Sigma_s,\Sigma_d)$ is defined as follows. For sort s, it is the freely generated algebraic structure $\mathsf{T}(\Sigma_s)$. For sort d, it is defined by the carrier $\Delta(\mathsf{T}(\Sigma_s))$ and the following interpretation: $[\![\delta(t)]\!] = \delta_t$ for all $t \in \mathsf{T}(\Sigma_s)$, $[\![\bigoplus_{i\in I}[p_i]\theta_i]\!] = \sum_{i\in I} p_i[\![\theta_i]\!]$ for $\{\theta_i \mid i \in I\} \subseteq \mathsf{T}(\Sigma_d)$, $[\![\boldsymbol{f}(\theta_1,\ldots,\theta_{\mathsf{rk}(f)})]\!](f(\xi_1,\ldots,\xi_{\mathsf{rk}(f)})) = \prod_{\sigma_i=s}[\![\theta_i]\!](\xi_i)$ if for all j s.t. $\sigma_j = d$, $\theta_j = \xi_j$, and $[\![\boldsymbol{f}(\theta_1,\ldots,\theta_{\mathsf{rk}(f)})]\!](f(\xi_1,\ldots,\xi_{\mathsf{rk}(f)})) = 0$ otherwise. Here it is assumed that $\prod\emptyset = 1$. Notice that in the semantics of a lifted function $\boldsymbol{f}$, the big product only considers the distributions related to the s-sorted positions in f, while the distribution terms corresponding to the d-sorted positions in f should match exactly to the parameters of $\boldsymbol{f}$.

A *substitution* ρ is a map $\mathcal{V} \cup \mathcal{V}_d \to \mathbb{T}(\Sigma)$ such that $\rho(x) \in \mathbb{T}(\Sigma_s)$, for all $x \in \mathcal{V}$, and $\rho(\mu) \in \mathbb{T}(\Sigma_d)$, for all $\mu \in \mathcal{V}_d$. A substitution is closed if it maps each variable to a closed term. A substitution extends to a mapping from terms to terms as usual.

Finally, we remark a general property of distribution terms: let $f \in \Sigma_s$ with $\mathsf{ar}(f) = \sigma_1\ldots\sigma_n \to s$, and let $\sigma_j = s$; then $\boldsymbol{f} \in \Sigma_d$ is distributive w.r.t. $\oplus$ in the position j, i.e. $[\![\rho(\boldsymbol{f}(\ldots,\xi_{j-1},\bigoplus_{i\in I}[p_i]\theta_i,\xi_{j+1},\ldots))]\!] = [\![\rho(\bigoplus_{i\in I}[p_i]\boldsymbol{f}(\ldots,\xi_{j-1},\theta_i,\xi_{j+1},\ldots))]\!]$ for any closed substitution ρ. The proof follows from the definition of $[\![\_]\!]$. However, notice that $\boldsymbol{f}$ *does not* distribute w.r.t. $\oplus$ in a position k such that $\sigma_k = d$.

3 Probabilistic Transition System Specifications

A (probabilistic) transition relation prescribes which possible activity can be performed by a term in a signature. Such activity is described by the label of the action and a probability distribution on terms that indicates the probability to reach a particular new term. We will follow the probabilistic automata style of probabilistic transitions [25] which are a generalization of the so-called reactive model [19].

Definition 2 (PTS)**.** *A* probabilistic labeled transition system *(PTS) is a triple* $(T(\Sigma_s), A, \rightarrow)$*, where* $\Sigma = (\Sigma_s, \Sigma_d)$ *is a probabilistically lifted signature, A is a countable set of actions, and* $\rightarrow \subseteq T(\Sigma_s) \times A \times \Delta(T(\Sigma_s))$*, is a transition relation. We write* $t \xrightarrow{a} \pi$ *for* $(t, a, \pi) \in \rightarrow$*.*

Transition relations are usually defined by means of structured operational semantics in Plotkin's style [24]. For PTS, algebraic characterizations of this style were provided in [8, 9, 21] where the term *probabilistic transition system specification* was used and which we adopt in our paper.

Definition 3 (PTSS)**.** *A* probabilistic transition system specification *(PTSS) is a triple* $P = (\Sigma, A, R)$ *where* Σ *is a probabilistically lifted signature, A is a set of labels, and R is a set of rules of the form:*

$$\frac{\{t_k \xrightarrow{a_k} \theta_k \mid k \in K\} \cup \{t_l \stackrel{b_l}{\not\rightarrow} \mid l \in L\} \cup \{\theta_j(T_j) \bowtie_j q_j \mid j \in J\}}{t \xrightarrow{a} \theta}$$

where K, L, J *are index sets,* $t, t_k, t_l \in \mathbb{T}(\Sigma_s)$*,* $a, a_k, b_l \in A$*,* $T_j \subseteq T(\Sigma_s)$*,* $\bowtie_j \in \{>, \geq, <, \leq\}$*,* $q_j \in [0,1]$ *and* $\theta_j, \theta_k, \theta \in \mathbb{T}(\Sigma_d)$*.*

Expressions of the form $t \xrightarrow{a} \theta$, $t \stackrel{a}{\not\rightarrow}$, and $\theta(T) \bowtie p$ are called *positive literal*, *negative literal*, and *quantitative literal*, respectively. For any rule $r \in R$, literals above the line are called *premises*, notation prem(r); the literal below the line is called *conclusion*, notation conc(r). We denote with pprem(r), nprem(r), and qprem(r) the sets of positive, negative, and quantitative premises of the rule r, respectively. In general, we allow the sets of positive, negative, and quantitative premises to be infinite.

Substitutions provide instances to the rules of a PTSS that, together with some appropriate machinery, allow us to define probabilistic transition relations. Given a substitution ρ, it extends to literals as follows: $\rho(t \stackrel{a}{\not\rightarrow}) = \rho(t) \stackrel{a}{\not\rightarrow}$, $\rho(\theta(T) \bowtie p) = \rho(\theta)(\rho(T)) \bowtie p$ (where $\rho(T) = \{\rho(t) \mid t \in T\}$), and $\rho(t \xrightarrow{a} \theta) = \rho(t) \xrightarrow{a} \rho(\theta)$.

We say that r' is a (closed) instance of a rule r if there is a (closed) substitution ρ so that $r' = \rho(r)$. We say that ρ is a *proper substitution of r* if for all quantitative premises $\theta(T) \bowtie p$ of r and all $t \in T$, $[\![\rho(\theta)]\!](\rho(t)) > 0$ holds. We use only this kind of substitution in the paper.

In the rest of the paper, we will deal with models as *symbolic* transition relations in the set $\mathsf{T}(\Sigma_s) \times A \times \mathsf{T}(\Sigma_d)$ rather than the *concrete* transition relations in $\mathsf{T}(\Sigma_s) \times A \times \Delta(\mathsf{T}(\Sigma_s))$ required by a PTS. Hence we will mostly refer with the term "transition relation" to the symbolic transition relation. In any case, a symbolic transition relation induces always a unique concrete transition relation by interpreting every target distribution term as the distribution it defines; that is, the symbolic transition $t \xrightarrow{a} \theta$ is interpreted as the concrete transition $t \xrightarrow{a} [\![\theta]\!]$. If the symbolic transition relation turns out to be a model of a PTSS P, we say that the induced concrete transition relation defines a PTS associated to P.

To define an appropriate notion of model we consider *3-valued models*. A 3-valued model partitions the set $\mathsf{T}(\Sigma_s) \times A \times \mathsf{T}(\Sigma_d)$ in three sets containing, respectively, the transition that are known to hold, that are known not to hold, and those whose validity is unknown. Thus, a 3-valued model can be presented as a pair $\langle \mathsf{CT}, \mathsf{PT} \rangle$ of transition relations $\mathsf{CT}, \mathsf{PT} \subseteq \mathsf{T}(\Sigma_s) \times A \times \mathsf{T}(\Sigma_d)$, with $\mathsf{CT} \subseteq \mathsf{PT}$, where CT is the set of transitions that *certainly* hold and PT is the set of transitions that *possibly* hold. So, transitions in $\mathsf{PT} \setminus \mathsf{CT}$ are those whose validity is unknown and transitions in $(\mathsf{T}(\Sigma_s) \times A \times \mathsf{T}(\Sigma_d)) \setminus \mathsf{PT}$ are those that certainly do not hold. A 3-valued model $\langle \mathsf{CT}, \mathsf{PT} \rangle$ that is justifiably compatible with the proof system defined by a PTSS P is said to be *stable* for P. (See Def. 5.)

Before formally defining the notions of proof and 3-valued stable model we introduce some notation. Given a transition relation $\mathsf{Tr} \subseteq \mathsf{T}(\Sigma_s) \times A \times \mathsf{T}(\Sigma_d)$, $t \xrightarrow{a} \theta$ *holds in* Tr, notation $\mathsf{Tr} \models t \xrightarrow{a} \theta$, if $t \xrightarrow{a} \theta \in \mathsf{Tr}$; $t \stackrel{a}{\not\rightarrow}$ *holds in* Tr, notation $\mathsf{Tr} \models t \stackrel{a}{\not\rightarrow}$, if for all $\theta \in \mathsf{T}(\Sigma_d)$, $t \xrightarrow{a} \theta \notin \mathsf{Tr}$. A closed quantitative constraint

$\theta(T) \bowtie p$ *holds in* Tr, notation $\mathsf{Tr} \models \theta(T) \bowtie p$, if $[\![\theta]\!](T) \bowtie p$. Notice that the satisfaction of a quantitative constraint does not depend on the transition relation. We nonetheless use this last notation as it turns out to be convenient. Given a set of literals H, we write $\mathsf{Tr} \models H$ if for all $\phi \in H$, $\mathsf{Tr} \models \phi$.

Definition 4 (Proof)**.** *Let $P = (\Sigma, A, R)$ be a PTSS. Let ψ be a positive literal and let H be a set of literals. A* proof *of a transition rule $\frac{H}{\psi}$ from P is a well-founded, upwardly branching tree where each node is a literal such that: (i) the root is ψ; and (ii) if χ is a node and K is the set of nodes directly above χ, then one of the following conditions holds: (a) $K = \emptyset$ and $\chi \in H$, or (b) $\chi = (\theta(T) \bowtie p)$ is a closed quantitative literal such that $[\![\theta]\!](T) \bowtie p$ holds, or (c) $\frac{K}{\chi}$ is a valid substitution instance of a rule from R.*

$\frac{H}{\psi}$ is provable from *P, notation $P \vdash \frac{H}{\psi}$, if there exists a proof of $\frac{H}{\psi}$ from P.*

Before, we said that a 3-valued stable model $\langle \mathsf{CT}, \mathsf{PT} \rangle$ for a PTSS P has to be *justifiably compatible* with the proof system defined by P. By "compatible" we mean that $\langle \mathsf{CT}, \mathsf{PT} \rangle$ has to be consistent with every provable rule. With "justifiable" we require that for each transition in CT and PT there is actually a proof that justifies it. More precisely, we require that (a) for every certain transition in CT there is a proof in P such that all negative hypotheses of the proof are known to hold (i.e. there is no possible transition in PT denying a negative hypothesis), and (b) for every possible transition in PT there is a proof in P such that all negative hypotheses possibly hold (i.e. there is no certain transition in CT denying a negative hypothesis). This is formally stated in the next definition.

Definition 5 (3-valued stable model)**.** *Let $P = (\Sigma, A, R)$ be a PTSS. A tuple $\langle \mathsf{CT}, \mathsf{PT} \rangle$ with $\mathsf{CT} \subseteq \mathsf{PT} \subseteq T(\Sigma_s) \times A \times T(\Sigma_d)$ is a* 3-valued stable model *for P if for every closed positive literal ψ,*

(a) $\psi \in \mathsf{CT}$ iff there is a set N of closed negative literals such that $P \vdash \frac{N}{\psi}$ and $\mathsf{PT} \models N$

(b) $\psi \in \mathsf{PT}$ iff there is a set N of closed negative literals such that $P \vdash \frac{N}{\psi}$ and $\mathsf{CT} \models N$.

The least 3-valued stable model of a PTSS can be constructed using induction [8, 11, 12].

Lemma 1. *Let P be a PTSS. For each ordinal α, define the pair $\langle \mathsf{CT}_\alpha, \mathsf{PT}_\alpha \rangle$ as follows:*

- $\mathsf{CT}_0 = \emptyset$ *and* $\mathsf{PT}_0 = T(\Sigma_s) \times A \times T(\Sigma_d)$.
- *For every non-limit ordinal $\alpha > 0$, define:*

$$\mathsf{CT}_\alpha = \left\{ t \xrightarrow{a} \theta \,|\, \textit{for some set } N \textit{ of negative literals, } P \vdash \frac{N}{t \xrightarrow{a} \theta} \textit{ and } \mathsf{PT}_{\alpha-1} \models N \right\}$$
$$\mathsf{PT}_\alpha = \left\{ t \xrightarrow{a} \theta \,|\, \textit{for some set } N \textit{ of negative literals, } P \vdash \frac{N}{t \xrightarrow{a} \theta} \textit{ and } \mathsf{CT}_{\alpha-1} \models N \right\}$$

- *For every limit ordinal α, define $\mathsf{CT}_\alpha = \bigcup_{\beta<\alpha} \mathsf{CT}_\beta$ and $\mathsf{PT}_\alpha = \bigcap_{\beta<\alpha} \mathsf{PT}_\beta$.*

Then: 1. if $\beta \leq \alpha$, $\mathsf{CT}_\beta \subseteq \mathsf{CT}_\alpha$ and $\mathsf{PT}_\beta \supseteq \mathsf{PT}_\alpha$, and 2. there is an ordinal λ such that $\mathsf{CT}_\lambda = \mathsf{CT}_{\lambda+1}$ and $\mathsf{PT}_\lambda = \mathsf{PT}_{\lambda+1}$. Moreover, $\langle \mathsf{CT}_\lambda, \mathsf{PT}_\lambda \rangle$ is the least 3-valued stable model for P.

PTSSs with least 3-valued stable model that are also a 2-valued model are particularly interesting, since this model is actually the only 3-valued stable model [7, 13]. A PTSS P is said to be *complete* if its least 3-valued stable model $\langle \mathsf{CT}, \mathsf{PT} \rangle$ satisfies that $\mathsf{CT} = \mathsf{PT}$ (i.e., the model is also 2-valued). We associate a probabilistic transition system to each complete PTSS.

Definition 6. *Let P be a complete PTSS and let $\langle Tr, Tr \rangle$ be its unique 3-valued stable model. We say that Tr is the* transition relation associated to *P. We also define the* PTS associated to *P as the unique PTS $(T(\Sigma_s), A, \rightarrow)$ such that $t \xrightarrow{a} \pi$ if and only if $t \xrightarrow{a} \theta \in Tr$ and $[\![\theta]\!] = \pi$ for some $\theta \in T(\Sigma_d)$.*

The different examples that we give in the rest of the papers are in terms of a basic probabilistic process algebra. We introduce it here, but address the reader to [8] for an example of a PTSS with richer operators. Signature Σ_s contains the constant 0, representing the stop process, for each action $a \in A$, a unary probabilistic prefix operators $a.\_$ with arity $\mathsf{ar}(a) = d \to s$, and a binary operator +, the alternative composition or sum, with arity $\mathsf{ar}(+) = ss \to s$, while Σ_d contains the respective lifted signature, δ, and all binary operators $\bigoplus_{i\in\{1,2\}}[p_i]\theta_i$ which we denote by $\oplus_p$. The semantics is defined with the usual rules:

$$\frac{}{a.\mu \xrightarrow{a} \mu} \qquad \frac{x \xrightarrow{a} \mu}{x+y \xrightarrow{a} \mu} \qquad \frac{y \xrightarrow{a} \mu}{x+y \xrightarrow{a} \mu}$$

4 Bisimulation relations

This work revolves around four different types of bisimulation relations: (i) the usual *(strong) bisimulation* [19] relation on probabilistic system, in which each probabilistic transition should be matched with a single probabilistic transition so that the distributions of both transitions agree on the probabilities of jumping into equivalent states; (ii) the *convex bisimulation* [25] relation, in which the matching is performed instead with a convex combination of transition relations; (iii) the *probability abstracted bisimulation*, in which the matching is performed by a single transition so that the distributions of both transitions agree on jumping to the same equivalent classes of states but not necessarily with the same probability value; and (iv) the *probabilistic obliterated bisimulation*, which represents the usual bisimulation [22] once the probabilistic transition system is abstracted into a traditional labeled transition system in the usual way.

To our knowledge, the probability abstracted bisimulation originates here. Its intention is to strictly preserve the probabilistic structure of a system without caring about the probability values. Thus, probability abstracted bisimulation is consistent with any bisimulation preserving quantitative properties that only tests for positive quantifications, rather than a particular value. Instead, this kind of properties are not preserved by the probabilistic obliterated bisimulation as it is shown below in this section.

In the following we introduce all these relations and discuss their relationship as well as alternative definitions. For the rest of the section we assume given a PTS $P = (\mathsf{T}(\Sigma_s), A, \to)$.

Given a relation $\mathsf{R} \subseteq \mathsf{T}(\Sigma_s)\times\mathsf{T}(\Sigma_s)$, a set $Q \subseteq \mathsf{T}(\Sigma_s)$ is *R-closed* if for all $t \in Q$ and $t' \in \mathsf{T}(\Sigma_s)$, $t \mathrel{\mathsf{R}} t'$ implies $t' \in Q$ (i.e. $\mathsf{R}(Q) \subseteq Q$). It is easy to verify that if two relations $\mathsf{R}, \mathsf{R}' \subseteq \mathsf{T}(\Sigma_s)\times\mathsf{T}(\Sigma_s)$ are such that $\mathsf{R}' \subseteq \mathsf{R}$, then if $Q \subseteq \mathsf{T}(\Sigma_s)$ is R-closed, it is also R′-closed.

Definition 7. *A relation $R \subseteq T(\Sigma_s)\times T(\Sigma_s)$ is a* bisimulation *if it is symmetric and for all $t,t' \in T(\Sigma_s)$, $a \in A$, and $\pi \in \Delta(T(\Sigma_s))$, $t \mathrel{R} t'$ and $t \xrightarrow{a} \pi$ imply that there exists $\pi' \in \Delta(T(\Sigma_s))$ s.t. $t' \xrightarrow{a} \pi'$ and $\pi \mathrel{R} \pi'$, where $\pi \mathrel{R} \pi'$ if and only if for all R-closed $Q \subseteq T(\Sigma_s)$, $\pi(Q) = \pi'(Q)$. The relation $\sim$, called* bisimilarity *or* bisimulation equivalence*, is defined as the smallest relation that includes all bisimulations.*

A *combined transition $t \xrightarrow{a}_c \pi$* is defined whenever there is a family $\{\pi_i\}_{i\in I} \subseteq \Delta(\mathsf{T}(\Sigma_s))$ and a family $\{p_i\}_{i\in I} \subseteq [0,1]$ such that $t \xrightarrow{a} \pi_i$ for all $i \in I$, $\sum_{i\in I} p_i = 1$ and $\pi = \sum_{i\in I} p_i\pi_i$.

Definition 8. *A relation $R \subseteq T(\Sigma_s)\times T(\Sigma_s)$ is a* convex bisimulation *if it is symmetric and for all $t,t' \in T(\Sigma_s)$, $a \in A$, and $\pi \in \Delta(T(\Sigma_s))$, $t \mathrel{R} t'$ and $t \xrightarrow{a} \pi$ imply that there exists $\pi' \in \Delta(T(\Sigma_s))$ s.t. $t' \xrightarrow{a}_c \pi'$ and $\pi \mathrel{R} \pi'$. The relation $\sim_c$, called* convex bisimilarity *or* convex bisimulation equivalence*, is defined as the smallest relation that includes all convex bisimulations.*

Definition 9. *A relation $R \subseteq T(\Sigma_s)\times T(\Sigma_s)$ is a* probability abstracted bisimulation *if it is symmetric and for all $t,t' \in T(\Sigma_s)$, $a \in A$, and $\pi \in \Delta(T(\Sigma_s))$, $t \mathrel{R} t'$ and $t \xrightarrow{a} \pi$ imply that there exists $\pi' \in \Delta(T(\Sigma_s))$ s.t. $t' \xrightarrow{a} \pi'$*

and for all R*-closed* $Q \subseteq T(\Sigma_s)$*,* $\pi(Q) > 0$ *iff* $\pi'(Q) > 0$*. The relation* $\sim_a$*, called* probability abstracted bisimilarity *or* probability abstracted bisimulation equivalence*, is defined as the smallest relation that includes all probability abstracted bisimulations.*

Notice that the transfer property in this last case follows the same structure as the bisimulation, only that it only requires that $\pi(Q) > 0$ iff $\pi'(Q) > 0$ for all R-closed, instead of $\pi(Q) = \pi'(Q)$.

Definition 10. *A relation* $R \subseteq T(\Sigma_s) \times T(\Sigma_s)$ *is a* probability obliterated bisimulation *if it is symmetric and for all* $t, t' \in T(\Sigma_s)$*,* $a \in A$*, and* $\pi \in \Delta(T(\Sigma_s))$*,* $t\, R\, t'$ *and* $t \xrightarrow{a} \pi$ *imply that for all* R*-closed* $Q \subseteq T(\Sigma_s)$ *with* $\pi(Q) > 0$*, there exists* $\pi' \in \Delta(T(\Sigma_s))$ *s.t.* $t' \xrightarrow{a} \pi'$ *and* $\pi'(Q) > 0$*. The relation* $\sim_o$*, called* probability obliterated bisimilarity *or* probability obliterated bisimulation equivalence*, is defined as the smallest relation that includes all probability obliterated bisimulations.*

Compare this last definition with Def. 9. While for probability abstracted bisimulation we require that there is a single matching transition $t' \xrightarrow{a} \pi'$ so that π' gives some positive probability to all R-closed sets exactly whenever π does, the definition of probability obliterated bisimulation permits to choose different matching transitions for each R-closed set that measures positively on π.

It is well known that $\sim$ and $\sim_c$ are equivalences relations and that they also are, respectively, a bisimulation relation and a convex bisimulation relation. The fact that $\sim_o$ is also an equivalence relation and itself a probability obliterated bisimulation follows from Lemma 4 which state that it agrees with Park & Milner's bisimulation. The same properties can be proven for probability abstracted bisimulation:

Lemma 2. $\sim_a$ *is an equivalence relation and is itself a probability abstracted bisimulation.*

Similarly to the bisimulation [3, Prop 3.4.4], the probability abstracted bisimulation has a characterization in terms of an *abstract* weight function. This alternative characterization is the one used in the proof of Theorem 4 and that is why we present it in this paper.

Given a relation $\mathsf{R} \subseteq \mathsf{T}(\Sigma_s) \times \mathsf{T}(\Sigma_s)$, we define $\equiv^w_{\mathsf{R}} \in \Delta(\mathsf{T}(\Sigma_s)) \times \Delta(\mathsf{T}(\Sigma_s))$ as follows. For all $\pi, \pi' \in \Delta(\mathsf{T}(\Sigma_s))$, $\pi \equiv^w_{\mathsf{R}} \pi'$ if there is an *abstract weight function* $w : (\mathsf{T}(\Sigma_s) \times \mathsf{T}(\Sigma_s)) \to [0,1]$ s.t. for all $t, t' \in \mathsf{T}(\Sigma_s)$, (i) $w(t, \mathsf{T}(\Sigma_s)) > 0$ iff $\pi(t) > 0$, (ii) $w(\mathsf{T}(\Sigma_s), t') > 0$ iff $\pi'(t') > 0$, and (iii) $w(t,t') > 0$ implies $t\, \mathsf{R}\, t'$.

Lemma 3. *For all* $t, t' \in T(\Sigma_s)$*,* $t \sim_a t'$ *if and only if there is a symmetric relation* $R \subseteq T(\Sigma_s) \times T(\Sigma_s)$ *with* $t\, R\, t'$ *such that for all* $t_1, t_2 \in T(\Sigma_s)$*,* $a \in A$*, and* $\pi_1 \in \Delta(T(\Sigma_s))$*,* $t_1\, R\, t_2$ *and* $t_1 \xrightarrow{a} \pi_1$ *imply that there exists* $\pi_2 \in \Delta(T(\Sigma_s))$ *s.t.* $t_2 \xrightarrow{a} \pi_2$ *and* $\pi_1 \equiv^w_R \pi_2$*.*

The next lemma shows that the probability obliterated bisimulation agrees with Park & Milner's bisimulation. Denote $t \overset{a}{\rightsquigarrow} t'$ iff there is π such that $t \xrightarrow{a} \pi$ and $\pi(t') > 0$. Notice that this notation precisely defines the usual abstraction of probabilistic transition systems into labeled transition systems in which all information regarding the probability distribution is lost except from the fact that one state can reach another state with positive probability after a transition.

Lemma 4. *For all* $t, t' \in T(\Sigma_s)$*,* $t \sim_o t'$ *iff there is a symmetric relation* $R \subseteq T(\Sigma_s) \times T(\Sigma_s)$ *with* $t\, R\, t'$ *s.t. for all* $t_1, t_2, t'_1 \in T(\Sigma_s)$ *and* $a \in A$*,* $t_1\, R\, t_2$ *and* $t_1 \overset{a}{\rightsquigarrow} t'_1$ *imply that there exists* $t'_2 \in T(\Sigma_s)$ *s.t.* $t_2 \overset{a}{\rightsquigarrow} t'_2$ *and* $t'_1\, R\, t'_2$*.*

Finally we state the relation among the different bisimulations

Lemma 5. *The following inclusions hold and are proper:* $\sim \subsetneq \sim_c \subsetneq \sim_o$ *and* $\sim \subsetneq \sim_a \subsetneq \sim_o$*. Besides* $\sim_c$ *and* $\sim_a$ *are incomparable.*

In fact the results can be proved stronger as we explain in the following. Any bisimulation relation is also a convex bisimulation, which follow from the fact that $t \xrightarrow{a} \pi$ implies $t \xrightarrow{a}_c \pi$. Any convex bisimulation is also a probability obliterated bisimulation since $t \xrightarrow{a}_c \pi$ with $\pi(Q) > 0$ implies that there is a

π' such that $t \xrightarrow{a} \pi'$ and $\pi'(Q) > 0$. Any bisimulation is also a probability abstracted bisimulation since $\pi \mathsf{R} \pi'$ implies $\pi(Q) > 0$ iff $\pi'(Q) > 0$ for all R-closed Q. Finally, any probability abstracted bisimulation is also a probability obliterated bisimulation since, for a given π and R, the existence of a π' s.t. $t' \xrightarrow{a} \pi'$ and $\pi(Q) > 0$ iff $\pi'(Q) > 0$ for all R-closed Q, guarantees that, for all R-closed Q with $\pi(Q) > 0$ there is a π' s.t. $t' \xrightarrow{a} \pi'$ and $\pi'(Q) > 0$.

Notice that $\mathbf{t}_1 = a.(\boldsymbol{b}.\mathbf{0}) + a.(\boldsymbol{c}.\mathbf{0})$ and $\mathbf{t}_2 = \mathbf{t}_1 + a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ are convex bisimilar but not probability abstracted bisimilar. Besides, notice that $\mathbf{t}_3 = a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ and $\mathbf{t}_4 = a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.1} \boldsymbol{c}.\mathbf{0})$ are probability abstracted bisimilar but not convex bisimilar. These examples not only show that $\sim_c$ and $\sim_a$ are incomparable, but also that all stated inclusions are proper.

In the rest of the section we present logical characterizations for the different bisimulation equivalences. This work has already been done for bisimulation [10, 16] and convex bisimulation [16]. We adopt here the two-level logic style of [10].

We define the logic $\mathcal{L}_b$ as the set of all formulas with the following syntax:

$$\phi := \top \mid \langle a\rangle\psi \mid \langle a\rangle_c\psi \mid \textstyle\bigwedge_{i\in I}\phi_i \mid \neg\phi \qquad\qquad \psi := [\phi]_p \mid \textstyle\bigsqcap_{i\in I}\psi_i$$

where $a \in A$, $p \in [0,1]\cap\mathbb{Q}$, and I is any index set. The logic $\mathcal{L}_c$ contains all formulas of $\mathcal{L}_b$ without the modality $\langle a\rangle\_$. The logic $\mathcal{L}_a$ contains all formulas of $\mathcal{L}_b$ without modalities $\langle a\rangle_c\_$ and $[\_]_p$ for all $p > 0$ (i.e. it only accepts $[\_]_0$ among this type of modalities.) Finally, the logic $\mathcal{L}_o$ contains all formulas of $\mathcal{L}_a$ without $\bigsqcap_{i\in I}\_$.

The semantics of $\mathcal{L}_b$ is defined with the satisfaction relation $\models$ on a PTS $P = (\mathsf{T}(\Sigma_s), A, \rightarrow)$ as follows.

(i) $t \models \top$ for all $t \in \mathsf{T}(\Sigma_s)$

(ii) $t \models \langle a\rangle\psi$ if there is $t \xrightarrow{a} \pi$ s.t. $\pi \models \psi$

(iii) $t \models \langle a\rangle_c\psi$ if there is $t \xrightarrow{a}_c \pi$ s.t. $\pi \models \psi$

(iv) $t \models \bigwedge_{i\in I}\phi_i$ if $t \models \phi_i$ for all $i \in I$

(v) $t \models \neg\phi$ if $t \not\models \phi$

(vi) $\pi \models [\phi]_p$ if $\pi(\{t \in \mathsf{T}(\Sigma_s) \mid t \models \phi\}) > p$

(vii) $\pi \models \bigsqcap_{i\in I}\psi_i$ if $\pi \models \psi_i$ for all $i \in I$

The semantics of the other logics is defined in the same way but restricted to the respective operators.

For $\chi \in \{b,c,a,o\}$, let $\mathcal{L}_\chi(t) = \{\phi \in \mathcal{L}_\chi \mid t \models \phi\}$, for all $t \in \mathsf{T}(\Sigma_s)$, and $\mathcal{L}_\chi(\pi) = \{\psi \in \mathcal{L}_\chi \mid \pi \models \psi\}$, for all $\pi \in \Delta(\mathsf{T}(\Sigma_s))$. We write $t_1 \sim_{\mathcal{L}_\chi} t_2$ iff $\mathcal{L}_\chi(t_1) = \mathcal{L}_\chi(t_2)$ and $\pi_1 \sim_{\mathcal{L}_\chi} \pi_2$ iff $\mathcal{L}_\chi(\pi_1) = \mathcal{L}_\chi(\pi_2)$. Then, we have the following characterization theorem.

Theorem 1. *For all $\chi \in \{b,c,a,o\}$ and for all $t_1,t_2 \in \mathsf{T}(\Sigma_s)$, $t_1 \sim_\chi t_2$ iff $t_1 \sim_{\mathcal{L}_\chi} t_2$ (where $\sim_b = \sim$).*

Let $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$, and $\mathbf{t}_4$ be as before. Recall $\mathbf{t}_1 \sim_c \mathbf{t}_2$ and $\mathbf{t}_3 \sim_a \mathbf{t}_4$. Notice that $\langle a\rangle([\langle b\rangle\top]_{0.5} \sqcap [\langle c\rangle\top]_{0.5})$ distinguish $\mathbf{t}_1$ from $\mathbf{t}_2$, while $\langle a\rangle_c([\langle b\rangle\top]_{0.5} \sqcap [\langle c\rangle\top]_{0.5})$ is satisfied by both $\mathbf{t}_1$ and $\mathbf{t}_2$. That is why $\langle a\rangle\_$ is not an operator of $\mathcal{L}_c$. Notice $[\langle b\rangle\top]_{0.5}$ distinguishes the distribution $[\![b.\mathbf{0} \oplus_{0.5} c.\mathbf{0}]\!]$ from $[\![b.\mathbf{0} \oplus_{0.1} c.\mathbf{0}]\!]$, while $[\langle b\rangle\top]_0$ does not (but it does distinguish them from e.g. $[\![c.\mathbf{0}]\!]$). Thus $\langle a\rangle[\langle b\rangle\top]_{0.5}$ distinguishes $\mathbf{t}_3$ from $\mathbf{t}_4$. That is why $[\_]_p$ is not an operator of $\mathcal{L}_a$ if $p > 0$. Finally, notice that $\langle a\rangle([\langle b\rangle\top]_0 \sqcap [\langle c\rangle\top]_0)$ distinguishes $\mathbf{t}_5 = a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ from $\mathbf{t}_6 = a.\boldsymbol{b}.\mathbf{0} + a.\boldsymbol{c}.\mathbf{0}$, and observe that $\mathbf{t}_5 \sim_o \mathbf{t}_6$. However, neither $\langle a\rangle[\langle b\rangle\top]_0$ nor $\langle a\rangle[\langle c\rangle\top]_0$ can distinguish them. That is why $\bigsqcap_{i\in I}\_$ is not an operator of $\mathcal{L}_o$.

5 Formats

In this section we introduce rule and specification formats that guarantee that each bisimulation equivalences discussed in the previous section is a congruence for every operator whose semantics is defined within the respective rule of the specification format. In particular, the format *ntμfθ/ntμxθ*, which ensures that bisimulation equivalence is a congruence for all operator in such format, has been already introduced in [9] and finally revised in [8]. We present here its more general form.

The following definition is important to ensure a symmetric treatment of variables and terms within the format. Let $\{Y_l\}_{l\in L}$ be a family of sets of state term variables with the same cardinality. The l-th element of a tuple $\vec{y}$ is denoted by $\vec{y}(l)$. For a set of tuples $T = \{\vec{y_i} \mid i \in I\}$ we denote the l-th projection by $\Pi_l(T) = \{\vec{y_i}(l) \mid i \in I\}$. Fix a set $\mathsf{Diag}\{Y_l\}_{l\in L} \subseteq \prod_{l\in L} Y_l$ such that: (i) for all $l \in L$, $\Pi_l(\mathsf{Diag}\{Y_l\}_{l\in L}) = Y_l$; and (ii) for all $\vec{y},\vec{y'} \in \mathsf{Diag}\{Y_l\}_{l\in L}$, $(\exists l \in L : \vec{y}(l) = \vec{y'}(l)) \Rightarrow \vec{y} = \vec{y'}$. Property (ii) ensures that different tuples $\vec{y},\vec{y'} \in \mathsf{Diag}\{Y_l\}_{l\in L}$ differ in all positions, and by property (i) every variable of every Y_l is used in (exactly) one $\vec{y} \in \mathsf{Diag}\{Y_l\}_{l\in L}$. Diag stands for "diagonal", following the intuition that each $\vec{y}$ represents a coordinate in the space $\prod_{l\in L} Y_l$, so that $\mathsf{Diag}\{Y_l\}_{l\in L}$ can be seen as the line that traverses the main diagonal of the space. Therefore, notice that, for $Y_l = \{y_l^0, y_l^1, y_l^2, \ldots\}$, a possible definition for the set $\mathsf{Diag}\{Y_l\}_{l\in L}$ is $\{(y_1^0, y_2^0, \ldots, y_L^0), (y_1^1, y_2^1, \ldots, y_L^1), (y_1^2, y_2^2, \ldots, y_L^2), \ldots\}$. In addition, we use the following notation: $t(\zeta_1, \ldots, \zeta_n)$ denotes a term that only has variables in the set $\{\zeta_1, \ldots, \zeta_n\}$, that is $\mathcal{V}(t(\zeta_1, \ldots, \zeta_n)) \subseteq \{\zeta_1, \ldots, \zeta_n\}$, and moreover, $t(\zeta'_1, \ldots, \zeta'_n)$ denotes the same term as $t(\zeta_1, \ldots, \zeta_n)$ in which each variable ζ_i has been replaced by ζ'_i.

Definition 11. *Let $P = (\Sigma, A, R)$ be a PTSS. A rule $r \in R$ is in ntμfθ* format *if it has the following form*

$$\frac{\bigcup_{m\in M}\{t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}} \mid \vec{z} \in \mathcal{Z}\} \;\cup\; \bigcup_{n\in N}\{t_n(\vec{z}) \stackrel{b_n}{\not\longrightarrow} \mid \vec{z} \in \mathcal{Z}\} \cup \{\theta_l(Y_l) \trianglerighteq_{l,k} p_{l,k} \mid l \in L, k \in K_l\}}{f(\zeta_1, \ldots, \zeta_{\mathsf{rk}(f)}) \xrightarrow{a} \theta}$$

with $\trianglerighteq_{l,k} \in \{>, \geq\}$ for all $l \in L$ and $k \in K_l$, and $\mathcal{Z} = \mathsf{Diag}\{Y_l\}_{l\in L} \times \prod_{\zeta\in W}\{\zeta\}$, with $W \subseteq \mathcal{V} \cup \mathcal{V}_d \setminus \bigcup_{l\in L} Y_l$, In addition, it has to satisfy the following conditions:

1. *Each set Y_l should be at least countably infinite, for all $l \in L$, and the cardinality of L should be strictly smaller than that of the Y_l's.*
2. *All variables $\zeta_1, \ldots, \zeta_{\mathsf{rk}(f)}$ are different.*
3. *All variables $\mu_m^{\vec{z}}$, with $m \in M$ and $\vec{z} \in \mathcal{Z}$, are different and $\{\zeta_1, \ldots, \zeta_{\mathsf{rk}(f)}\} \cap \{\mu_m^{\vec{z}} \mid \vec{z} \in \mathcal{Z}, m \in M\} = \emptyset$.*
4. *For all $l \in L$, $Y_l \cap \{\zeta_1, \ldots, \zeta_{\mathsf{rk}(f)}\} = \emptyset$, and $Y_l \cap Y_{l'} = \emptyset$ for all $l' \in L$, $l \neq l'$.*
5. *For all $m \in M$, the set $\{\mu_m^{\vec{z}} \mid \vec{z} \in \mathcal{Z}\} \cap (\mathcal{V}(\theta) \cup (\bigcup_{l\in L} \mathcal{V}(\theta_l)) \cup W)$ is finite.*
6. *For all $l \in L$, the set $Y_l \cap (\mathcal{V}(\theta) \cup \bigcup_{l'\in L} \mathcal{V}(\theta_{l'}))$ is finite.*

A rule $r \in R$ is in ntμxθ format *if its form is like above but has a conclusion of the form $x \xrightarrow{a} \theta$ and, in addition, it satisfies the same conditions as above, except that whenever we write $\{\zeta_1, \ldots, \zeta_{\mathsf{rk}(f)}\}$, we should write $\{x\}$. P is in ntμfθ* format *if all its rules are in ntμfθ format. P is in ntμfθ/ntμxθ* format *if all its rules are in either ntμfθ format or ntμxθ format.*

The rationale behind each of the restrictions are discussed in [8] in depth (see also [9]). In the following we briefly summarize it. Variables $\zeta_1, \ldots, \zeta_{\mathsf{rk}(f)}$ in the source of the conclusion, all variables $\mu_m^{\vec{z}}$ in the target of the positive premises, and all variables in the sets Y_l, $l \in L$, as part of the measurable sets in the quantitative premises, are binding. That is why all of them are requested to be different, which is stated in conditions 2, 3, and 4. If Y_l is finite, quantitative premises will allow to count the minimum number of terms that gather certain probabilities. This goes against the spirit of bisimulation that measures equivalence classes of terms regardless of the size of them. Therefore Y_l needs to be infinite (condition 1). Condition 5 ensures that, for each $m \in M$ there are sufficiently many distribution variables in the set $\{\mu_m^{\vec{z}} \mid \vec{z} \in \mathcal{Z}\}$ to be freely instantiated. The use of a distribution variable in a quantitative premise may disclose part of the structural nature of the distribution term that substitutes such variable. Thus, for instance, if all variables $\mu_m^{\vec{z}}$ are used in different quantitative premises together with some lookahead, we may restrict the syntactic form of the eventually substituted distribution terms, hence revealing unwanted differences. A similar situation arises with the use of variables in Y_l for all $l \in L$, hence condition 6. The precise understanding of conditions 5 and 6 requires a rather lengthy explanation that is beyond the scope of this paper. The reader is referred to [8,9] for details.

All congruence theorems in this article apply only to PTSSs whose rules are *well-founded*. A rule r is *well-founded* if there is no infinite backward chain in the *dependency graph* $G_r = (V, E)$ of r defined by $V = \mathcal{V} \cup \mathcal{V}_d$ and $E = \{\langle \zeta, \mu \rangle \mid (t \xrightarrow{a} \mu) \in \text{pprem}(r), \zeta \in \mathcal{V}(t)\} \cup \{\langle \zeta, y \rangle \mid (\theta(Y) \trianglerighteq p) \in \text{qprem}(r), \zeta \in \mathcal{V}(\theta), y \in Y\}$. A PTSS is called *well-founded* if all its rules are well-founded.

The full proof of the following theorem can be found in [8].

Theorem 2. *Let $P = (\Sigma, A, R)$ be a complete well-founded PTSS in ntμfθ/ntμxθ format. Then, the bisimulation equivalence is a congruence for all operators defined in P.*

The *ntμfθ/ntμxθ* format is still too general to preserve the other (weaker) bisimulation equivalences presented in Sec. 4. In the reminder of the section, we will discuss through appropriate examples how the *ntμfθ/ntμxθ* format should be further restricted or modified so that the other bisimulation equivalences are congruences for the resulting restricted formats.

We focus first on convex bisimulation. For this consider the terms $\mathsf{t}_1 = a.(\boldsymbol{b}.\mathbf{0}) + a.(\boldsymbol{c}.\mathbf{0})$ and $\mathsf{t}_2 = \mathsf{t}_1 + a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$. Notice that $\mathsf{t}_1 \sim_c \mathsf{t}_2$. Consider a possible extension of our running example with a unary operator f with the following *ntμfθ* rule:

$$\frac{x \xrightarrow{a} \mu \quad \mu(Y) \geq 0.5 \quad \{y \xrightarrow{b} \mu_y \mid y \in Y\} \quad \mu(Y') \geq 0.5 \quad \{y' \xrightarrow{c} \mu_{y'} \mid y' \in Y'\}}{f(x) \xrightarrow{a} \mathbf{0}} \tag{1}$$

Since $\mathsf{t}_2 \xrightarrow{a} (\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$, $f(\mathsf{t}_2) \xrightarrow{a} \mathbf{0}$. However it is easy to see that $f(\mathsf{t}_1)$ cannot perform any transition. Therefore $f(\mathsf{t}_1) \not\sim_c f(\mathsf{t}_2)$.

The problem arises precisely because, in order to show that $\mathsf{t}_1 \sim_c \mathsf{t}_2$, transition $\mathsf{t}_2 \xrightarrow{a} (\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ is matched with the appropriate convex combination of the transitions $\mathsf{t}_1 \xrightarrow{a} \boldsymbol{b}.\mathbf{0}$ and $\mathsf{t}_1 \xrightarrow{a} \boldsymbol{c}.\mathbf{0}$. Thus, we need that a quantitative premise guarantees that the test is produced on a convex combination of target distributions rather than on a single target distribution. An appropriate modification of such rule would be to replace it by a family of rules of the form

$$\frac{\{x \xrightarrow{a} \mu_n \mid n \in \mathbb{N}\} \quad \left(\bigoplus_{n\in\mathbb{N}}[p_n]\mu_n\right)(Y) \geq 0.5 \quad \{y \xrightarrow{b} \mu_y \mid y \in Y\} \quad \left(\bigoplus_{n\in\mathbb{N}}[p_n]\mu_n\right)(Y') \geq 0.5 \quad \{y' \xrightarrow{c} \mu_{y'} \mid y' \in Y'\}}{f(x) \xrightarrow{a} \mathbf{0}},$$

one for each $\{p_n\}_{n\in\mathbb{N}}$ such that $\sum_{n\in\mathbb{N}} p_n = 1$ and each $p_i \in [0,1] \cap \mathbb{Q}$.

Consider now that the semantic of f is defined by the rule

$$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \boldsymbol{a}.\mu} \tag{2}$$

and notice that $f(\mathsf{t}_2) \xrightarrow{a} \boldsymbol{a}.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$. However, the only two possible transitions for $f(\mathsf{t}_1)$ are $f(\mathsf{t}_1) \xrightarrow{a} \boldsymbol{a}.\boldsymbol{b}.\mathbf{0}$ and $f(\mathsf{t}_1) \xrightarrow{a} \boldsymbol{a}.\boldsymbol{c}.\mathbf{0}$, and there is no $p \in [0,1]$ such that $\boldsymbol{a}.\boldsymbol{b}.\mathbf{0} \oplus_p \boldsymbol{a}.\boldsymbol{c}.\mathbf{0} \sim_c \boldsymbol{a}.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$. For this reason, we will require that a target of a positive premise does not appear in a d-sorted position of a subterm in the target of the conclusion.

For the next example, we consider an additional unary d-sorted operator g and the following rules

$$\frac{x \xrightarrow{a} \mu \qquad g(\mu) \xrightarrow{b} \mu'}{f(x) \xrightarrow{a} \mathbf{0}} \qquad\qquad \frac{\mu(Y) > 0 \quad \{y \xrightarrow{b} \mu \mid y \in Y\} \quad \mu(Y') > 0 \quad \{y' \xrightarrow{c} \mu' \mid y' \in Y'\}}{g(\mu) \xrightarrow{b} \mathbf{0}} \tag{3}$$

Notice that $g(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0}) \xrightarrow{b} \mathbf{0}$. Therefore $f(\mathsf{t}_2) \xrightarrow{a} \mathbf{0}$. However, neither $g(\boldsymbol{b}.\mathbf{0})$ nor $g(\boldsymbol{c}.\mathbf{0})$ can perform any transition, and as a consequence $f(\mathsf{t}_1)$ cannot perform any transition either. Hence, $f(\mathsf{t}_1) \not\sim_c f(\mathsf{t}_2)$. For this reason we will require that a target of a positive premise does not appear in the source of a positive or negative premise.

Suppose now that g is a binary s-sorted operator and consider the following rules

$$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \boldsymbol{g}(\mu,\mu)} \qquad \frac{x_1 \xrightarrow{b} \mu_1 \quad x_2 \xrightarrow{c} \mu_2}{g(x_1,x_2) \xrightarrow{a} \mathbf{0}} \tag{4}$$

Notice that the only possible transitions for $f(\mathsf{t}_1)$ are $f(\mathsf{t}_1) \xrightarrow{a} \boldsymbol{g}(\boldsymbol{b}.\mathbf{0},\boldsymbol{b}.\mathbf{0})$ and $f(\mathsf{t}_1) \xrightarrow{a} \boldsymbol{g}(\boldsymbol{c}.\mathbf{0},\boldsymbol{c}.\mathbf{0})$. Moreover, notice that $\boldsymbol{g}(\boldsymbol{b}.\mathbf{0},\boldsymbol{b}.\mathbf{0}) \sim_c \boldsymbol{g}(\boldsymbol{c}.\mathbf{0},\boldsymbol{c}.\mathbf{0}) \sim_c \mathbf{0}$. However, $f(\mathsf{t}_2) \xrightarrow{a} \boldsymbol{g}(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0}, \boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$, and it is not difficult to see that $\boldsymbol{g}(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0}, \boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0}) \sim_c (\boldsymbol{a}.\mathbf{0} \oplus_{0.25} \mathbf{0})$. Therefore, $f(\mathsf{t}_1) \not\sim_c f(\mathsf{t}_2)$. In this case, the problem seems to arise because the same distribution variable occurs in the target of the conclusion of the first rule in two different s-sorts positions of the target distribution term. However, the problem is not so general. Notice that if the target in the conclusion is replaced by the term $\boldsymbol{g}(\mu,\boldsymbol{c}.\mathbf{0}) \oplus_p \boldsymbol{g}(\boldsymbol{b}.\mathbf{0},\mu)$ we would have $f(\mathsf{t}_1) \sim_c f(\mathsf{t}_2)$. The difference arises from the fact that in the interpretation of $\boldsymbol{g}(\theta,\theta)$ the probability distribution $[\![\theta]\!]$ multiplies with itself. This is not the case in the interpretation of $\boldsymbol{g}(\theta,\boldsymbol{c}.\mathbf{0}) \oplus_p \boldsymbol{g}(\boldsymbol{b}.\mathbf{0},\theta)$ where the two instances of $[\![\theta]\!]$ are summed up. Thus, we will actually request that the target of the conclusion is *linear* with respect to each distribution variable on a target of a positive premise.

Definition 12. *A distribution term $\theta \in \mathbb{T}(\Sigma_d)$ is* linear for a set $V \subseteq \mathcal{V}_d$ *if (i)* $\theta \in T(\Sigma_d) \cup \mathcal{V}_d \cup \{\delta(x) \mid x \in \mathcal{V}\}$. *(ii)* $\theta = \bigoplus_{i\in I}[p_i]\theta_i$ *and* θ_i *is linear for* V, *for all* $i \in I$, *(iii)* $\theta = f(\theta_1,\dots,\theta_n)$, *for all* $i \in I$, θ_i *is linear for* V, *and* $\mathcal{V}(\theta_i) \cap \mathcal{V}(\theta_j) \cap V = \emptyset$, *for all* $i, j \in \{1,\dots,n\}$ *and* $i \neq j$,

Definition 13. *Let* $P = (\Sigma, A, R)$ *be a PTSS. A rule* $r \in R$ *is in* convex $nt\mu f\theta$ format *if has the form*

$$\frac{\begin{array}{cc}\bigcup_{m\in M}\{t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}} \mid \vec{z} \in \mathcal{Z}\} & \bigcup_{n\in N}\{t_n(\vec{z}) \stackrel{b_n}{\not\longrightarrow} \mid \vec{z} \in \mathcal{Z}\} \\ \bigcup_{\tilde{m}\in\tilde{M}}\{t_{\tilde{m}}(\vec{z}_{\tilde{m}}) \xrightarrow{a_{\tilde{m}}} \mu_i^{\tilde{m}} \mid i \in \mathbb{N}\} & \bigcup_{\tilde{m}\in\tilde{M}}\{(\bigoplus_{i\in\mathbb{N}}[p_i^{\tilde{m}}]\mu_i^{\tilde{m}})(Y_l) \trianglerighteq_{l,k} p_{l,k} \mid l \in L_{\tilde{m}}, k \in K_l\}\end{array}}{f(\zeta_1,\dots,\zeta_{\mathsf{rk}(f)}) \xrightarrow{a} \theta}$$

with $L = \cup_{\tilde{m}\in\tilde{M}} L_{\tilde{m}}$, $L_{\tilde{m}} \cap L_{\tilde{m}'} = \emptyset$ *whenever* $\tilde{m} \neq \tilde{m}'$, $\trianglerighteq_{l,k} \in \{>,\geq\}$ *for all* $l \in L$ *and* $k \in K_l$, $\mathcal{Z} = \mathsf{Diag}\{Y_l\}_{l\in L} \times \prod_{\zeta\in W}\{\zeta\}$, *with* $W \subseteq \mathcal{V} \cup \mathcal{V}_d \setminus \bigcup_{l\in L} Y_l$. *In addition, it should also satisfy conditions 1 to 6 in Def. 11 and the following extra conditions:*

7. *For every* $\tilde{m} \in \tilde{M}$, *the family* $\{p_i^{\tilde{m}}\}_{i\in\mathbb{N}} \subseteq [0,1] \cap \mathbb{Q}$ *and* $\sum_{i\in\mathbb{N}} p_i^{\tilde{m}} = 1$
8. *For every* $\tilde{m} \in \tilde{M}$, *there is exactly one* $j \in \mathbb{N}$ *such that* $\mu_j^{\tilde{m}} = \mu_m^{\vec{z}}$ *for some* $m \in M$ *and* $\vec{z} \in \mathcal{Z}$, *in which case also* $t_{\tilde{m}}(\vec{z}_{\tilde{m}}) \xrightarrow{a_{\tilde{m}}} \mu_j^{\tilde{m}} = t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}}$. *Moreover* $\{\mu_i^{\tilde{m}} \mid i \in \mathbb{N}\} \cap \{\mu_i^{\tilde{m}'} \mid i \in \mathbb{N}\} = \emptyset$ *for all* $\tilde{m} \neq \tilde{m}'$, *and* $\{\mu_i^{\tilde{m}} \mid i \in \mathbb{N}\} \cap \{\zeta_1,\dots,\zeta_{\mathsf{rk}(f)}\} = \emptyset$.
9. *No variable* $\mu_m^{\vec{z}}$, *with* $m \in M$ *and* $\vec{z} \in \mathcal{Z}$, *appears in the source of a premise (i.e. in the set* W*) or in a d-sorted position of a subterm in the target of the conclusion* θ.
10. θ *is linear for* $\{\mu_m^{\vec{z}} \mid m \in M, \vec{z} \in \mathcal{Z}\}$.

A rule $r \in R$ *is in* convex $nt\mu x\theta$ format *if its form is like above but has a conclusion of the form* $x \xrightarrow{a} \theta$ *and it satisfies the same conditions, except that whenever we write* $\{\zeta_1,\dots,\zeta_{\mathsf{rk}(f)}\}$, *we should write* $\{x\}$. *A set of convex* $nt\mu f\theta/nt\mu x\theta$ *rules* R *is* convex closed *if for all* $r \in R$, *for any term* $\bigoplus_{i\in\mathbb{N}}[p_i^{\tilde{m}}]\mu_i^{\tilde{m}}$ *appearing in a quantitative premise of* r *and any family* $\{q_i\}_{i\in\mathbb{N}} \subseteq [0,1] \cap \mathbb{Q}$ *such that* $\sum_{i\in\mathbb{N}} q_i = 1$, *then the rule* r' *obtained by replacing each occurrence of* $\bigoplus_{i\in\mathbb{N}}[p_i^{\tilde{m}}]\mu_i^{\tilde{m}}$ *in* r *by* $\bigoplus_{i\in\mathbb{N}}[q_i]\mu_i^{\tilde{m}}$ *is also in* R. *A PTSS* $P = (\Sigma, A, R)$ *is in* convex $nt\mu f\theta/nt\mu x\theta$ format *if all rules in* R *are in convex* $nt\mu f\theta/nt\mu x\theta$ *format and* R *is convex closed.*

The problem indicated in rule (1) is attacked with the requirement of having sets $\{t_{\tilde{m}}(\vec{z}_{\tilde{m}}) \xrightarrow{a_{\tilde{m}}} \mu_i^{\tilde{m}} \mid i \in \mathbb{N}\}$ as positive premises with which the convex closures $\bigoplus_{i\in\mathbb{N}}[p_i^{\tilde{m}}]\mu_i^{\tilde{m}}$ can be constructed, plus the request that the set of rules is convex closed. Notice that condition 8 states that these sets of positive premises are only used to construct such distribution terms and are only linked to the "actual" positive premises in $\bigcup_{m\in M}\{t_m(\vec{z}) \xrightarrow{a_m} \mu_m^{\vec{z}} \mid \vec{z} \in \mathcal{Z}\}$ through a single transition $t_{\tilde{m}}(\vec{z}_{\tilde{m}}) \xrightarrow{a_{\tilde{m}}} \mu_j^{\tilde{m}}$.

Rules like (2) and the left rule on (3) are excluded on condition 9 since no variable of a positive premise can be used in the source of a premise (excluding (3)) or in a d-sort position in the target of the conclusion (excluding (2)). Finally, rules like on the left of (4) are excluded by requesting that the target of the conclusion is linear (condition 10).

Now, we can state the congruence theorem for convex bisimulation equivalence.

Theorem 3. *Let P be a complete well-founded PTSS in convex ntμfθ/ntμxθ format. Then, convex bisimulation equivalence is a congruence for all operators defined by P.*

We focus now on the probability abstracted bisimulation. Notice that the terms $\mathsf{t}_3 = a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ and $\mathsf{t}_4 = a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.1} \boldsymbol{c}.\mathbf{0})$ are probability abstracted bisimilar, i.e., $\mathsf{t}_3 \sim_a \mathsf{t}_4$. Consider now the unary operator f whose semantics is defined with rule (1). It should not be difficult so see that $f(\mathsf{t}_3) \xrightarrow{a} \mathbf{0}$ while $f(\mathsf{t}_4)$ cannot perform any transition. Therefore $f(\mathsf{t}_3) \not\sim_a f(\mathsf{t}_4)$. The problem is a consequence of the fact that the quantitative premises are tested against non-zero values which may distinguish distributions with the same support set but mapping into different probability values. Thus, in order to preserve probability abstracted bisimulation equivalence, the only extra restriction that we ask to a rule in *ntμfθ/ntμxθ* format is that none of its quantitative premises test against a value different from 0.

Definition 14. *A PTSS $P = \langle \Sigma, A, R\rangle$ is in* probability abstracted *ntμfθ/ntμxθ* format *if it is in ntμfθ/ntμxθ format and for every rule $r \in R$ and quantitative premise $\theta(Y) \trianglerighteq p \in qprem(r)$, $p = 0$.*

The proof of the congruence theorem for probability abstracted bisimulation equivalence (Theorem 4 below) follows closely the lines of the proof of Theorem 2 as given in [8].

Theorem 4. *Let P be a complete well-founded PTSS in probability abstracted ntμfθ/ntμxθ format. Then, the probability abstracted bisimulation equivalence is a congruence for all operators defined in P.*

Given the alternative definition of the probability obliterated bisimulation provided by Lemma 4, we will now consider simpler definitions for the quantitative premises for the rule format associated to this relation. Thus, we consider quantitative premises of the form $\theta(\{y\}) \trianglerighteq p$ rather than $\theta(Y) \trianglerighteq p$.

Taking t_3 and t_4 as before, we have that $\mathsf{t}_3 \sim_o \mathsf{t}_4$. The same example of the unary operator f, whose semantics is defined with a conveniently modify rule (1), shows that $f(\mathsf{t}_3) \not\sim_a f(\mathsf{t}_4)$ and hence the need that the quantitative premises can only be tested against 0.

Let $\mathsf{t}_5 = a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ and $\mathsf{t}_6 = a.\boldsymbol{b}.\mathbf{0} + a.\boldsymbol{c}.\mathbf{0}$, and observe that $\mathsf{t}_5 \sim_o \mathsf{t}_6$. Take rule (2) as the semantic definition for f. Notice that $f(\mathsf{t}_5) \xrightarrow{a} \boldsymbol{a}.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ is the only transition for $f(\mathsf{t}_5)$, while the only possible transitions for $f(\mathsf{t}_6)$ are $f(\mathsf{t}_6) \xrightarrow{a} \boldsymbol{a}.\boldsymbol{b}.\mathbf{0}$ and $f(\mathsf{t}_6) \xrightarrow{a} \boldsymbol{a}.\boldsymbol{c}.\mathbf{0}$. Since $a.\boldsymbol{b}.\mathbf{0} \not\sim_o a.(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0}) \not\sim_o a.\boldsymbol{c}.\mathbf{0}$, $f(\mathsf{t}_5) \not\sim_o f(\mathsf{t}_6)$. Like for the convex bisimulation case, this shows that the target of a positive premise cannot appear in a d-sorted position of a subterm in the target of the conclusion.

Suppose now that the semantics of f is defined with the rule

$$\frac{x \xrightarrow{a} \mu \qquad \mu(\{y_1\}) > 0 \qquad \mu(\{y_2\}) > 0 \qquad y_1 \xrightarrow{b} \mu_1 \qquad y_2 \xrightarrow{c} \mu_2}{f(x) \xrightarrow{a} \mathbf{0}} \tag{5}$$

Notice that $f(\mathsf{t}_5) \not\sim_o f(\mathsf{t}_6)$ since $f(\mathsf{t}_5) \xrightarrow{a} \mathbf{0}$ while $f(\mathsf{t}_6)$ cannot perform any transition. This is due to the fact that, by allowing the same distribution variable μ to occur in different quantitative premises, we gain some knowledge of the structure of (the instance of) μ, in particular of its support set.

Consider now that f is defined with the left rule in (3) and g with an appropriate modification of the right rule in (3). Notice that $f(\mathtt{t}_5) \xrightarrow{a} \mathbf{0}$ but $f(\mathtt{t}_6)$ cannot perform any transition. Thus $f(\mathtt{t}_5) \not\sim_o f(\mathtt{t}_6)$. In this case, we are also gaining knowledge of the support set of μ, but this time through the rule associated to the operator g. Therefore we require that a target of a positive premise does not appear in the source of a positive or negative premise.

Consider now the rules

$$\frac{x \xrightarrow{a} \mu \quad \mu(\{y\}) > 0 \quad y \xrightarrow{b} \mu'}{f(x) \xrightarrow{a} \boldsymbol{g}(\mu)} \qquad \frac{x \xrightarrow{c} \mu}{g(x) \xrightarrow{c} \mathbf{0}} \tag{6}$$

Notice that the only transition for $f(\mathtt{t}_5)$ is $f(\mathtt{t}_5) \xrightarrow{a} \boldsymbol{g}(\boldsymbol{b}.\mathbf{0} \oplus_{0.5} \boldsymbol{c}.\mathbf{0})$ and the only transition for $f(\mathtt{t}_6)$ is $f(\mathtt{t}_6) \xrightarrow{a} \boldsymbol{g}(\boldsymbol{b}.\mathbf{0})$. Then $f(\mathtt{t}_5) \overset{a}{\rightsquigarrow} g(c.\mathbf{0}) \overset{c}{\rightsquigarrow} 0$ while $f(\mathtt{t}_6) \overset{a}{\rightsquigarrow} g(b.\mathbf{0})$ is the only possible "obliterated" transition for $f(\mathtt{t}_6)$. Then $f(\mathtt{t}_5) \not\sim_o f(\mathtt{t}_6)$. This is an alternative way of gaining information on the support set of a possible instance of μ in (6): on the one hand, by the quantitative premise on the first rule, we deduce that such instance has an element in the support set that performs a b-transition and, on the other hand, by having μ as an argument in the target of the conclusion, we may gather extra information from the same instance of μ through the rules for the semantics of the target of the conclusion (in this case, that μ has another element in the support set that performs a c-transition.) Therefore, we forbid that the target of a positive premise is both tested in a quantitative premise and used in the target of the conclusion.

Notice that the example in rules (4) also apply for probability obliterated bisimulation since $\mathtt{t}_1 \sim_o \mathtt{t}_2$ but $f(\mathtt{t}_1) \not\sim_o f(\mathtt{t}_2)$ with exactly the same explanation. Thus, we also request that the target of the conclusion is linear for all distribution variables on targets of positive premises.

Finally, consider a modification (4) where the left rule is instead

$$\frac{x \xrightarrow{a} \mu \quad \boldsymbol{g}(\mu,\mu)(\{y\}) > 0 \quad y \xrightarrow{a} \mu'}{f(x) \xrightarrow{a} \mathbf{0}} \tag{7}$$

It should not be difficult to observe that $f(\mathtt{t}_5) \xrightarrow{a} \mathbf{0}$ but $f(\mathtt{t}_6)$ cannot perform any transition. Thus $f(\mathtt{t}_5) \not\sim_o f(\mathtt{t}_6)$. For this reason we also require that the quantitative premises only allow linear distribution terms.

Definition 15. *Let $P = (\Sigma, A, R)$ be a well-founded PTSS. A rule $r \in R$ is in* probability obliterated *ntμfθ* format *if it has the form*

$$\frac{\bigcup_{m\in M}\{t_m \xrightarrow{a_m} \mu_m\} \quad \cup \quad \bigcup_{n\in N}\{t_n \overset{b_n}{\not\longrightarrow}\} \quad \cup \quad \bigcup_{l\in L}\{\theta_l(\{y_l\}) > 0\}}{f(\zeta_1,\ldots,\zeta_{\mathsf{rk}(f)}) \xrightarrow{a} \theta}$$

where all variables $\zeta_1,\ldots,\zeta_{\mathsf{rk}(f)}$, μ_m, with $m \in M$, and y_l, with $l \in L$, are different and the following restrictions are satisfied:

1. *For all $m \in M$, $\mathcal{V}(t_m) \cap \{\mu_{m'} \mid m' \in M\} = \emptyset$. Similarly, for all $n \in N$, $\mathcal{V}(t_n) \cap \{\mu_{m'} \mid m' \in M\} = \emptyset$.*
2. *For all $l \in L$, θ_l is linear for $\{\mu_{m'} \mid m' \in M\}$ and, moreover, for all $l,l' \in L$ with $l \neq l'$, $\mathcal{V}(\theta_l) \cap \mathcal{V}(\theta_{l'}) \cap \{\mu_m \mid m \in M\} = \emptyset$.*
3. *θ is linear for $\{\mu_{m'} \mid m' \in M\}$, $\mathcal{V}(\theta) \cap (\bigcup_{l\in L}\mathcal{V}(\theta_l)) \cap \{\mu_m \mid m \in M\} = \emptyset$, and no variable μ_m appear in a d-sorted position of a subterm of the target of the conclusion θ.*

A rule is in probability obliterated *ntμxθ* format *if its form is like above but has a conclusion of the form $x \xrightarrow{a} \theta$. P is in* probability obliterated *ntμfθ/ntμxθ* format *if all its rules are in probability obliterated ntμfθ/ntμxθ format.*

Condition 1 limits the form to exclude rules like the one on the left of (3). Condition 2 requires that the distribution terms on the quantitative premises are linear (excluding (6)), and that they do not share distributions variables on the target of positive premises (excluding (5)). Finally, condition 3 request that the target of the conclusion is linear (excluding (4)) and does not have targets of positive premises on d-sorted positions (excluding (2)) nor if they are used in quantitative premises (excluding (6)).

Finally, we state the congruence theorem for probability obliterated bisimulation equivalence.

Theorem 5. *Let P be a complete well-founded PTSS in probability obliterated ntμfθ/ntμxθ format. Then, probability obliterated bisimulation equivalence is a congruence for all operators defined by P.*

6 Conclusion and Future Work

In this article, we presented three new rule formats that preserve three different bisimulation equivalences coarser than Larsen & Skou's bisimulation. These formats are more restricted variants of the *ntμfθ/ntμxθ* format and notably, all of them can be seen as generalizations of the non-probabilistic *ntyft/ntyxt* format [7, 14]. For completeness we mention two other similar results on PTSSs that fall out of Larsen & Skou's bisimulation. They are [20], that presents a format for rooted branching bisimulation, and [26], that presents a format for non-expansiveness of ϵ-bisimulations.

Prior to the congruence theorems, we presented the different bisimulation equivalences, compare them, and, in particular, we gave a logic characterization for each of them. The intention of presenting these logic characterizations is to use them as the basis for the proof of full abstraction theorems (see, e.g., [8, 9, 14, 15].) Full abstraction theorems are somewhat dual to the congruence theorems. An equivalence relation is fully abstract with respect to a particular format and an equivalence relation $\equiv$ if it is the largest relation included in $\equiv$ that is a congruence for all operators definable by any PTSS in that format. In particular we are interested when $\equiv$ is the coarsest reasonable behavioral equivalence, namely, (possibilistic) trace equivalence. We are busy now on trying to prove this results for the formats presented here using the logic characterization as a means to construct the so called *testers*. As the current point of our investigation, we do not foresee major problems for all relations except for convex bisimulation equivalences, for which we may need to relax some of the conditions of the convex *ntμfθ/ntμxθ* format.

References

[1] L. Aceto, W. Fokkink & C. Verhoef (2001): *Conservative Extension in Structural Operational Semantics*. In: Current Trends in Theoretical Computer Science, pp. 504–524, doi:10.1142/9789812810403_0004.

[2] L. Aceto, W. Fokkink & C. Verhoef (2001): *Structural operational semantics*. In: Handbook of Process Algebra, Elsevier, pp. 197–292, doi:10.1016/B978-044482830-9/50021-7.

[3] C. Baier (1998): *On Algorithmics Verification Methods for Probabilistic Systems*. Habilitation thesis, University of Mannheim.

[4] F. Bartels (2002): *GSOS for Probabilistic Transition Systems*. Electr. Notes Theor. Comput. Sci. 65(1), pp. 29–53, doi:10.1016/S1571-0661(04)80358-X.

[5] F. Bartels (2004): *On Generalised Coinduction and Probabilistic Specification Formats*. Ph.D. thesis, VU University Amsterdam.

[6] B. Bloom, S. Istrail & A.R. Meyer (1995): *Bisimulation can't be traced*. J. ACM 42, pp. 232–268, doi:10.1145/200836.200876.

[7] R. Bol & J.F. Groote (1996): *The meaning of negative premises in transition system specifications*. J. ACM 43, pp. 863–914, doi:10.1145/234752.234756.

[8] P.R. D'Argenio, D. Gebler & M.D. Lee (2015): *A general SOS theory for the specification of probabilistic transition systems*. Available at `http://www.cs.famaf.unc.edu.ar/~lee/publications/DArgenioGeblerLee.pdf`. Submitted.

[9] P.R. D'Argenio & M.D. Lee (2012): *Probabilistic Transition System Specification: Congruence and Full Abstraction of Bisimulation*. In: *Proc. FoSSaCS'12*, *LNCS* 7213, Springer, pp. 452–466, doi:10.1007/978-3-642-28729-9_30.

[10] P.R. D'Argenio, P. Sánchez Terraf & N. Wolovick (2012): *Bisimulations for non-deterministic labelled Markov processes*. *Mathematical Structures in Computer Science* 22(1), pp. 43–68, doi:10.1017/S0960129511000454.

[11] W. Fokkink (2000): *Rooted Branching Bisimulation as a Congruence*. *J. Comput. Syst. Sci.* 60(1), pp. 13–37, doi:10.1006/jcss.1999.1663.

[12] W. Fokkink & C. Verhoef (1998): *A Conservative Look at Operational Semantics with Variable Binding*. *Inf. Comput.* 146(1), pp. 24–54, doi:10.1006/inco.1998.2729.

[13] R.J. van Glabbeek (2004): *The meaning of negative premises in transition system specifications II*. *J. Log. Algebr. Program.* 60-61, pp. 229–258, doi:10.1016/j.jlap.2004.03.007.

[14] J.F. Groote (1993): *Transition system specifications with negative premises*. *Theoretical Computer Science* 118(2), pp. 263–299, doi:10.1016/0304-3975(93)90111-6.

[15] J.F. Groote & F.W. Vaandrager (1992): *Structured Operational Semantics and Bisimulation as a Congruence*. *Inf. Comput.* 100(2), pp. 202–260, doi:10.1016/0890-5401(92)90013-6.

[16] H. Hermanns, A. Parma, R. Segala, B. Wachter & L. Zhang (2011): *Probabilistic Logical Characterization*. *Inf. Comput.* 209(2), pp. 154–172, doi:10.1016/j.ic.2010.11.024.

[17] R. Lanotte & S. Tini (2005): *Probabilistic Congruence for Semistochastic Generative Processes*. In: *Foundations of Software Science and Computational Structures, 8th International Conference, FOSSACS 2005.*, pp. 63–78, doi:10.1007/978-3-540-31982-5_4.

[18] R. Lanotte & S. Tini (2009): *Probabilistic bisimulation as a congruence*. *ACM Trans. Comput. Log.* 10(2), doi:10.1145/1462179.1462181.

[19] K.G. Larsen & A/ Skou (1991): *Bisimulation through Probabilistic Testing*. *Inf. Comput.* 94(1), pp. 1–28, doi:10.1016/0890-5401(91)90030-6.

[20] M.D. Lee & E. de Vink (2015): *Rooted branching bisimulation as a congruence for probabilistic transition systems*. In: *Proc. QAPL'15*, EPTCS. To appear.

[21] M.D. Lee, D. Gebler & P.R. D'Argenio (2012): *Tree Rules in Probabilistic Transition System Specifications with Negative and Quantitative Premises*. In: *Proc. EXPRESS/SOS'12*, *EPTCS* 89, pp. 115–130, doi:10.4204/EPTCS.89.9.

[22] R. Milner (1989): *Communication and Concurrency*. Prentice-Hall.

[23] M.R. Mousavi, M.A. Reniers & J.F. Groote (2007): *SOS formats and meta-theory: 20 years after*. *Theor. Comput. Sci.* 373(3), pp. 238–272, doi:10.1016/j.tcs.2006.12.019.

[24] G. Plotkin (1981): *A structural approach to operational semantics*. Report DAIMI FN-19, Aarhus University, doi:doi:10.1016/j.jlap.2004.05.001. Reprinted in *J. Log. Algebr. Program.*, 60-61:17-139, 2004.

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

[26] S. Tini (2010): *Non-expansive ϵ-bisimulations for probabilistic processes*. *Theor. Comput. Sci.* 411(22-24), pp. 2202–2222, doi:10.1016/j.tcs.2010.01.027.