Probabilistic Divide & Congruence: Branching Bisimilarity

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

[a] *Reykjavik University, Iceland*
[b] *INRIA Saclay - Ile de France, Paris, France*
[c] *University of Insubria, Como, Italy*

Abstract

Since the seminal paper by Bloom, Fokkink and van Glabbeek, the *Divide and Congruence* technique allows for the derivation of compositional properties of nondeterministic processes from the SOS-based decomposition of their modal properties. In an earlier paper, we extended their technique to deal also with quantitative aspects of process behavior: we proved the (pre)congruence property for strong (bi)simulations on processes with nondeterminism and probability. In this paper we further extend our decomposition method to favor compositional reasoning with respect to probabilistic weak semantics. In detail, we consider *probabilistic branching* and *rooted probabilistic branching bisimilarity*, and we propose logical characterizations for them. These are strongly based on the modal operator $\langle \varepsilon \rangle$ which combines quantitative information and weak semantics by introducing a sort of *probabilistic lookahead* on process behavior. Our enhanced method will exploit *distribution specifications*, an SOS-like framework defining the probabilistic behavior of processes, to decompose this particular form of lookahead. We will show how we can apply the proposed decomposition method to derive *congruence formats* for the considered equivalences from their logical characterizations.

Keywords: Modal decomposition, Nondeterministic probabilistic transition systems, SOS, Congruence formats, Probabilistic branching bisimulation

1. Introduction

Structural Operational Semantics (SOS) [68] is the standard framework to define the operational semantics of processes. Briefly, processes are represented as *terms* over a proper *algebra*, giving the *abstract syntax* of the considered language, and their operational behavior is expressed by *transition steps* that are derived from a *transition system specification* (TSS) [68], namely a set of inference rules of the form $\frac{premises}{conclusion}$ whose intuitive meaning is that whenever the premises are satisfied, then the transition step constituting the conclusion can be deduced. The set of transitions steps that can be deduced, or *proved*, from the TSS constitutes the *labeled transition system (LTS) generated by the TSS* [46]. Usually, *behavioral relations*, such as *preorders* or *equivalences*, are defined on the LTS in order to compare the behavior of processes, possibly abstracting away from details that are irrelevant in a given context.

Equipping processes with a semantics, however, is not the only application of the SOS framework. One of the main targets in the development of a meta-theory of process description languages is to support *compositional reasoning*, which requires language operators to be compatible with the behavioral relation $\mathcal{R}$ chosen for the application context. In algebraic terms, this compatibility is known as the *congruence property* of $\mathcal{R}$ with respect to all language operators, which consists in verifying whether

$$f(t_1, \dots, t_\mathfrak{n}) \,\mathcal{R}\, f(t'_1, \dots, t'_\mathfrak{n}) \text{ for any } \mathfrak{n}\text{-ary operator } f \text{ whenever } t_i \,\mathcal{R}\, t'_i \text{ for all } i = 1, \dots, \mathfrak{n}.$$

[1]This author is partially supported by the project 'Open Problems in the Equational Logic of Processes' (OPEL) of the Icelandic Research Fund (grant nr. 196050-051).

The congruence property guarantees that the *substitution* of a component of a system with an $\mathcal{R}$-equivalent one does not affect the behavior of that system. The SOS framework plays a crucial role in supporting compositional reasoning and verification: a *rule* (or *specification*) *format* (see the survey papers [1, 65]) is a set of *syntactical constraints* over inference rules ensuring the desired semantic properties of the generated LTS. Thus, one can prove useful results, as the congruence property, for a whole class of languages at the same time. For instance, the *De Simone* format [70] ensures that trace equivalence is a congruence, the *GSOS* format [11] works for bisimilarity and in [10] the *ntyft/ntxt* format [49] is reduced to the *ready trace* format and its variants to guarantee that decorated trace preorders are precongruences.

1.1. Probabilistic process algebras

Quantitative phenomena occur whenever the behavior of a system is not deterministic and it is characterized by a variety of uncertainties, which can be either empirical, and thus due to incomplete knowledge on system design, or physical, due to random (physical) events and the interaction of the system with the external environment. *Probability* is one of the most important measures of uncertainty and is employed, e.g., in networks [38, 50], security [2, 51], embedded systems [30].

Probabilistic process algebras [21, 54], such as the probabilistic versions of CCS [7, 23, 54], CSP [7, 19, 27, 54] and ACP [3, 4], extend classical process algebras with suitable operators allowing us to express probability distributions over sets of possible events or behaviors. The most general semantic model extending LTSs to capture probabilistic behavior is that of *nondeterministic probabilistic labeled transition systems* (PTSs) [69], where, essentially, transition steps take processes to *distributions* over processes.

The SOS framework has been applied also in the probabilistic setting. For instance, considering only congruence formats proposed on TSSs generating PTSs, the *PGSOS* format [7] and the *ntμfθ/ntμxθ* format [19] ensure the congruence property of probabilistic bisimilarity [59]. Then, in [61] a probabilistic version of the RBB safe format from [31] for (rooted) branching bisimilarity is proposed.

1.2. *The* Divide & Congruence *technique*

Irrespective of probability being considered or not, one of the main questions that needs to be answered is "*How can we derive compositional results for a behavioral relation from a rule format?*"

One possible answer, the one that will be pursued in this paper, is to exploit the *logical characterization* of the considered equivalence relation, which consists in exhibiting a class of *modal formulae* $\mathcal{L}$ such that two processes are equivalent if and only if they satisfy the same formulae in $\mathcal{L}$, thus expressing that two processes cannot be distinguished by the observations that we can make on their behavior [53]. Hence the congruence property becomes equivalent to state that the set of formulae satisfied by any process $f(t_1, \dots, t_n)$ can be determined from the sets of formulae satisfied by the subprocesses $t_1, \dots, t_n$.

Concretely, in [60] a *compositional proof system* for the Hennessy-Milner logic (HML) [53] is provided. In order to obtain system implementations that are correct with respect to their specifications, the authors propose an *implementation by contexts* reasoning: instead of extracting an implementation for the complete system from the specification, it is easier, and hence preferable, to implement first the behavior of subcomponents. Thus, to obtain the correctness of the whole system we need to establish what properties each subcomponent should satisfy in order to guarantee that the system in which they are combined (or more generally their *context*) will satisfy some given property (by the specification). Since the specification of a system can be expressed in terms of modal formulae, the above statement can be reduced to establish whether given a formula ϕ and a context C there are formulae $\phi_1 \dots \phi_n$ such that

$$\text{whenever } t_i \models \phi_i \text{ for each } i = 1, \dots, n \text{ then } C[t_1, \dots, t_n] \models \phi. \qquad (1)$$

The analogy with the congruence property is immediate. To obtain it, [60] exploits an SOS machinery to specify contexts [58]: by means of *action transducers* they reformulate a given TSS into a TSS in De Simone format from which the formulae ϕ_i required in (1) are derived by means of *property transformers*.

Inspired by this work, [10] introduces the *modal decomposition method*. The underlying idea is the same: reducing the satisfaction problem of a formula for a process to verifying whether its subprocesses satisfy certain formulae obtained from its decomposition. This is obtained by the notions of *ruloids* [11] (an enhanced version of the action transducers of [60]), namely derived inference rules deducing the behavior of process terms directly from the behavior of the variables occurring in them, and of *decomposition mappings* (the property transformers of [60]) associating to each pair term-formula (t, ϕ) the set of formulae that the variables in t have to satisfy to guarantee that t satisfies ϕ. The contribution of [10] and the subsequent works [32–37] is not only related to the definition of the decomposition methods but also to their application. In fact, they show that by combining the logical characterization of a relation, the decomposition of such a logic and a rule format for the relation it is possible to systematically derive a congruence format for that relation directly from its modal characterization. Briefly, it is enough to guarantee that the construction of ruloids from the considered TSS preserves the syntactical constraints of the format. At the same time the modal decomposition has to preserve the logical characterization, that is formulae in the characterizing class $\mathcal{L}$ have to be decomposed into formulae in $\mathcal{L}$. Then, from the compositional result (1) related to the modal decomposition, the congruence property follows.

For what concerns probabilistic processes, in [39] the model of *reactive* probabilistic LTSs [48] is considered and a method for decomposing formulae from a probabilistic version of HML [66] characterizing probabilistic bisimilarity with respect to a probabilistic TSS in the format of [57] is proposed.

1.3. Our goal: the Probabilistic Divide & Congruence

In our earlier papers [13, 14] we extended the *Divide and Congruence* technique to *nondeterministic and probabilistic processes* in the PTS model. We defined a decomposition method for formulae in the probabilistic extension of HML from [24] and we derived congruence formats for probabilistic *strong* (bi)simulations. To obtain the decomposition, we developed an SOS-like machinery, called *distribution specification*, by which we modeled *syntactically* the behavior of probability distributions over process terms. This allowed us to *decompose* the formulae capturing the *quantitative behavior* of processes.

Our aim is now to generate congruence formats for *probabilistic weak semantics*, which allows us to abstract from *unobservable* (also called *internal*, or *silent*) computation steps by processes, which, as usual, are represented in the transition system by transitions labeled with the special symbol τ. The target is therefore to pave the way to the development of a *Probabilistic Divide and Congruence* technique allowing us to deal with the interplay of *nondeterminism, probability and weak semantics*.

Considering the vast amount of technical definitions and results that we will need, we decided to present only the congruence formats for *probabilistic branching bisimulation* and its *rooted* version, so that we can support our results with examples and thorough explanations. Accordingly, we will consider only process terms defined in the PGSOS specification format [17], instead of the more general ntμfθ/ntμxθ format [19]. This will allow us to give an inductive definition of ruloids. Moreover, the PGSOS format allows for specifying nearly all probabilistic process algebras operators, so that our results will still have a wide range of applications, as shown in Section 8. We delay till Section 9 an in depth discussion about the potential extension of our results to the more general ntμfθ/ntμxθ format.

1.4. Implementation of the goal

The derivation of the congruence property for probabilistic (rooted) branching bisimilarity is obtained in three main steps: 1. First, we have to provide a logical characterization for the considered relations. 2. Then, we have to develop a decomposition method for the characterizing class of formulae. 3. Finally, we have to define the formats for the considered relations and verify that the constraints of both the formats and the logical characterization are preserved in the decomposition procedure. In detail:

1. *Logical characterization* (Section 3).

   We introduce a class of modal formulae $\mathbb{L}$, obtained by extending HML with special formulae $\langle\tau\rangle\psi$, $\langle\varepsilon\rangle\psi$ in which the modalities $\langle\tau\rangle$ and $\langle\varepsilon\rangle$ express the execution of (sequences of) silent steps, while a formula ψ is defined via the probabilistic choice operator $\bigoplus$ from [24] and expresses the probabilistic behavior of a process. In particular, the interplay of probability and weak semantics is specially captured by formulae of the form $\langle\varepsilon\rangle\psi$ which express a form of *probabilistic lookahead*. Informally, a process satisfies $\langle\varepsilon\rangle\psi$ if via the execution of an arbitrary number of silent steps it reaches a probability distribution

π that satisfies ψ. Such distribution π is obtained by combining all the probabilistic choices that are performed during the execution of the silent steps. Hence, to define π we need to keep memory of the probabilistic choices that have been performed. For this reason, we shall say that a formula $\langle\varepsilon\rangle\psi$ gives us constraints on future behavior without loosing all the information on probabilistic step-by-step behavior. Thus, similarly to what happens in the classic lookahead, to verify the desired probabilistic behavior we need to check the entire sequence of weak steps and probabilistic choices leading to that particular final distribution π.

As a first result, we prove that two fragments of $\mathbb{L}$, denoted by $\mathbb{L}_{\mathrm{b}}$ and $\mathbb{L}_{\mathrm{rb}}$, allow us to characterize, respectively, branching bisimilarity and its rooted version (Theorem 1).

2. *The decomposition method* (Sections 5–6).

   The decomposition of a formula $\varphi \in \mathbb{L}$ with respect to any term $f(t_1, \ldots, t_n)$ is defined as the set of formulae that the subterms $t_1, \ldots, t_n$ must satisfy to guarantee that $f(t_1, \ldots, t_n)$ satisfies φ. Hence, we first need to find a method to infer the behavior of open terms directly from that of the variables occurring in them. This is obtained via the notion of *ruloids*. In particular, we define *ruloids* from PGSOS rules to derive the behavior of process terms and we exploit the *distribution specification* introduced in [13] to define *distribution ruloids* dealing with the behavior of open *distribution terms*, namely the syntactic expressions we use to denote probability distributions. Then, by means of PGSOS ruloids we obtain the decomposition of formulae expressing the reactive and nondeterministic behavior of processes and by distribution ruloids we derive the decomposition of formulae expressing the probabilistic behavior of processes.

   The main novelty of our method with respect to those in [13, 37] is in the decomposition of formulae of the form $\langle\varepsilon\rangle\psi$, as the probabilistic lookahead introduced by them is strictly related to the interplay of probability and weak semantics. It has never been addressed in previous work on modal decomposition. We will combine PGSOS ruloids and the distribution ruloids to decompose this form of lookahead. Informally, we introduce formulae of the form $\psi^{\langle\varepsilon\rangle}$, in which the superscript $\cdot^{\langle\varepsilon\rangle}$ is a marker that allows us to record, in the decomposition procedure, that ψ occurs in the scope of $\langle\varepsilon\rangle$. Then, we alternate the decomposition of a τ-step in the ε-sequence, via a PGSOS ruloid, and the decomposition of a (proper) formula $\tilde{\psi}^{\langle\varepsilon\rangle}$, via a distribution ruloid, until we consume the entire sequence of τ-steps and reach a distribution term on which ψ can be directly decomposed.

   Ours is the first proposal of a decomposition method capturing nondeterminism, probability and weak behavior (Theorem 4). Our technique applies to other notions of behavioral equivalence.

3. *The congruence formats* (Sections 7–8).

   We introduce the congruence formats on PGSOS specifications for probabilistic branching bisimilarity and its rooted version, the *probabilistic branching bisimulation format* (PBB) and the *probabilistic rooted branching bisimulation format* (PRBB), respectively. These, can be considered as the extensions to the probabilistic setting of the corresponding formats in [37].

   Then we prove the congruence result. Firstly, we prove that the syntactic constraints of the formats are preserved in the construction of PGSOS ruloids (Theorem 5). Then, we prove that our logical characterizations are preserved by the decomposition, namely formulae in $\mathbb{L}_{\mathrm{b}}$ (resp. $\mathbb{L}_{\mathrm{rb}}$) are decomposed into formulae that are still in $\mathbb{L}_{\mathrm{b}}$ (resp. $\mathbb{L}_{\mathrm{rb}}$) (Theorems 6, 7). Finally, by combining these results with the decomposition theorem (Theorem 4), we get that (rooted) branching bisimilarity is a congruence for a PGSOS specification in our PBB (resp. PRBB) format (Theorems 8, 9).

   Finally, we apply the PRBB and the PBB formats to the operators of P_{PA}, a probabilistic process algebra from [40] that includes most of probabilistic operators. We observe that the majority the operators in P_{PA} are included in our formats, thus showing that these formats are not too demanding. Moreover, operators out of the formats do not enjoy the congruence property.

1.5. Summary of results

Our contributions can then be summarized as follows:

1. We provide a logical characterization of probabilistic branching bisimilarity and its rooted version via proper subclasses of the modal logic $\mathbb{L}$, a probabilistic extension of HML tailored to express weak semantics.
2. We define a modal decomposition method for formulae in $\mathbb{L}$. To decompose formulae expressing probabilistic behavior, our method will make use of distribution specifications, an SOS-like machinery that we introduced in [13] to specify the behavior of open distribution terms. In particular, the distribution specification allows us to decompose the probabilistic lookahead introduced by formulae of the form $\langle\varepsilon\rangle\psi$. As this special form of lookahead depends only on the interplay of probability and weak semantics, the decomposition of formulae of the form $\langle\varepsilon\rangle\psi$ is a unique feature of our method. To the best of our knowledge, ours is the first proposal of a modal decomposition for formulae expressing nondeterministic, probabilistic and weak behavior of processes.
3. We define the PRBB and the PBB formats guaranteeing, respectively, that probabilistic rooted branching bisimilarity and probabilistic branching bisimilarity are congruences for nondeterministic probabilistic processes.
4. We prove that ruloids derived from PGSOS rules preserve the syntactic constraints of the formats.
5. We prove that the decomposition method preserves the logical characterization, namely formulae belonging to a particular class are decomposed into formulae belonging to the same class.
6. We obtain the congruence result as a direct consequence of previous results.

1.6. Outline of the paper

We start by recalling the necessary notions on the PTS model and the PGSOS framework in Section 2. In Section 3 we introduce the modal logic $\mathbb{L}$ as well as the characterization results. In Section 4 we review the distribution specification introduced in [13] and we define the ruloids in Section 5. We dedicate Section 6 to the definition of the decomposition method. Our main results, the probabilistic (rooted) branching bisimulation format and the congruence results, are presented in Section 7, whereas Section 8 comes with some examples of their application. In Section 9 we discuss the extension of our decomposition method to specifications in the ntμfθ/ntμxθ format. Finally, we conclude the paper in Section 10 by discussing related and future work.

2. Background

2.1. The PTS model

In the process algebra setting, processes are inductively defined as terms over a suitable *signature*, namely a countable set Σ of *operators*. We let f range over Σ and $\mathfrak{n}$ denote the *rank* of $f \in \Sigma$. Operators with rank 0 are called *constants*. We assume a set of (state) *variables* $\mathbf{V}_s$ disjoint from Σ and ranged over by $x, y, \ldots$. The set $\mathsf{T}(\Sigma, V)$ of *terms* over Σ and a subset of variables $V \subseteq \mathbf{V}_s$ is the least set such that: (i) $V \subseteq \mathsf{T}(\Sigma, V)$, and (ii) $f(t_1, \ldots, t_\mathfrak{n}) \in \mathsf{T}(\Sigma, V)$ whenever $f \in \Sigma$ is of rank $\mathfrak{n}$ and $t_1, \ldots, t_\mathfrak{n} \in \mathsf{T}(\Sigma, V)$. By $\mathbf{T}(\Sigma)$ we denote the set of the *closed terms* $\mathsf{T}(\Sigma, \emptyset)$, also called *processes*. By $\mathbb{T}(\Sigma)$ we denote the set of all *open terms* $\mathsf{T}(\Sigma, \mathbf{V}_s)$. By $\mathrm{var}(t)$ we denote the set of the variables occurring in term t. We say that a term t is *univariate* if each variable in $\mathrm{var}(t)$ occurs exactly once in it. Otherwise, t is called *multivariate*.

A very general semantic model for probabilistic processes is that of nondeterministic probabilistic labeled transition systems (PTSs) [69], which combine classic labeled transition systems [55] and discrete-time Markov chains (MCs) [52], allowing us to model, at the same time, reactive behavior, nondeterminism and probability. The state space is given by the set of the processes $\mathbf{T}(\Sigma)$. Each transition has a label describing the underlying computation step and takes a process to a probability distribution over $\mathbf{T}(\Sigma)$, i.e. a mapping

$\pi\colon \mathbf{T}(\Sigma) \to [0,1]$ with $\sum_{t \in \mathbf{T}(\Sigma)} \pi(t) = 1$. By $\Delta(\mathbf{T}(\Sigma))$ we denote the set of all probability distributions over $\mathbf{T}(\Sigma)$. We assume a set of *action labels* $\mathcal{A}$ and a *non observable*, or *silent*, action $\tau \notin \mathcal{A}$. The set $\mathcal{A} \cup \{\tau\}$ is denoted with $\mathcal{A}_\tau$. We let $a, b, \dots$ range over $\mathcal{A}$ and $\alpha, \beta, \dots$ over $\mathcal{A}_\tau$.

Definition 1 (PTS, [69])**.** A PTS is a triple $(\mathbf{T}(\Sigma), \mathcal{A}_\tau, \to)$, where: (i) Σ is a signature, (ii) $\mathcal{A}$ is a countable set of action labels with $\tau \notin \mathcal{A}$, and (iii) $\to \subseteq \mathbf{T}(\Sigma) \times \mathcal{A}_\tau \times \Delta(\mathbf{T}(\Sigma))$ is a *transition relation*.

As usual, a *transition* $(t, \alpha, \pi) \in \to$ is denoted $t \xrightarrow{\alpha} \pi$. Then, $t \stackrel{\alpha}{\not\longrightarrow}$ denotes that $t \xrightarrow{\alpha} \pi$ holds for no π.

We need some notation for distributions. For $\pi \in \Delta(\mathbf{T}(\Sigma))$, $\mathsf{supp}(\pi) = \{t \in \mathbf{T}(\Sigma) \mid \pi(t) > 0\}$ is the *support* of distribution π. For a process $t \in \mathbf{T}(\Sigma)$, δ_t is the *Dirac distribution* with $\delta_t(t) = 1$ and $\delta_t(s) = 0$ for all processes $s \neq t$. For $f \in \Sigma$ and $\pi_i \in \Delta(\mathbf{T}(\Sigma))$, $f(\pi_1, \dots, \pi_\mathfrak{n})$ is the distribution with $f(\pi_1, \dots, \pi_\mathfrak{n})(f(t_1, \dots, t_\mathfrak{n})) = \prod_{i=1}^{\mathfrak{n}} \pi_i(t_i)$ and $f(\pi_1, \dots, \pi_\mathfrak{n})(t) = 0$ for all t not in the form $f(t_1, \dots, t_\mathfrak{n})$. The convex combination $\sum_{i \in I} p_i \pi_i$ of a family of distributions $\{\pi_i\}_{i \in I} \subseteq \Delta(\mathbf{T}(\Sigma))$ with $p_i \in (0,1]$ and $\sum_{i \in I} p_i = 1$ is defined by $(\sum_{i \in I} p_i \pi_i)(t) = \sum_{i \in I} (p_i \pi_i(t))$ for all $t \in \mathbf{T}(\Sigma)$. We may write $p\pi_1 + (1-p)\pi_2$ for $\sum_{i \in \{1,2\}} p_i \pi_i$ with $p = p_1$ and $(1-p) = p_2$. Note that $\delta_{f(t_1, \dots, t_\mathfrak{n})} = f(\delta_{t_1}, \dots, \delta_{t_\mathfrak{n}})$.

In a PTS, given an infinite set I, a *divergence* is a sequence of processes $\{t_i\}_{i \in I}$ and distributions $\{\pi_i\}_{i \in I}$ with $t_i \xrightarrow{\tau} \pi_i$ and $t_{i+1} \in \mathsf{supp}(\pi_i)$, for all $i \in I$. To avoid the use of additional technical expedients in our paper, we consider only *divergence-free* PTSs, namely those having no divergence.

2.2. Probabilistic branching bisimulation

A *probabilistic branching bisimulation* is an equivalence relation over $\mathbf{T}(\Sigma)$ that equates two terms if they can mimic each others *observable* transitions and evolve to distributions related, in turn, by the same equivalence. To formalize this intuition, we need to *lift* relations over terms 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 arbitrary 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:

(i) $\sum_{y \in Y} \mathfrak{w}(x,y) = \pi_X(x)$, for all $x \in X$ (ii) $\sum_{x \in X} \mathfrak{w}(x,y) = \pi_Y(y)$, 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.

In the literature, we can find several equivalent definitions of lifting for relations over terms. We recall here that in [69], other proposals will be recalled in the Appendix. In the upcoming proofs, we will use the most suitable definition of relation lifting among the proposed ones.

Definition 3 (Relation lifting, [69])**.** The *lifting* of a relation $\mathcal{R} \subseteq \mathbf{T}(\Sigma) \times \mathbf{T}(\Sigma)$ is the relation $\mathcal{R}^\dagger \subseteq \Delta(\mathbf{T}(\Sigma)) \times \Delta(\mathbf{T}(\Sigma))$ with $\pi \, \mathcal{R}^\dagger \, \pi'$ iff there is a matching $\mathfrak{w} \in \mathfrak{W}(\pi, \pi')$ s.t. $s' \, \mathcal{R} \, t'$ whenever $\mathfrak{w}(s', t') > 0$.

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 [61, 63] and stems from [26].

Definition 4 (Lifted transition)**.** Assume a PTS $(\mathbf{T}(\Sigma), \mathcal{A}_\tau, \to)$. We define the relation $\xrightarrow{\hat{\cdot}} : (\mathbf{T}(\Sigma) \cup \Delta(\mathbf{T}(\Sigma))) \times \mathcal{A}_\tau \times \Delta(\mathbf{T}(\Sigma))$ from $\to$ as follows:

- for $a \in \mathcal{A}$, $t \xrightarrow{\hat{a}} \pi$ iff $t \xrightarrow{a} \pi$;
- $t \xrightarrow{\hat{\tau}} \pi$ iff either $t \xrightarrow{\tau} \pi$ or $\pi = \delta_t$;
- for $\alpha \in \mathcal{A}_\tau$, $\pi \xrightarrow{\hat{\alpha}} \pi'$ iff (i) $t \xrightarrow{\hat{\alpha}} \pi_t$ for all $t \in \mathsf{supp}(\pi)$, and (ii) $\pi' = \sum_{t \in \mathsf{supp}(\pi)} \pi(t) \pi_t$.

As usual we write $t \xrightarrow{\hat{\varepsilon}} \pi$ for the reflexive-transitive closure of the relation $\xrightarrow{\hat{\tau}}$. We remark that, since we are considering divergence-free PTSs, a lifted transition $t \xrightarrow{\hat{\tau}} \delta_t$ can never be inferred from any transition $t \xrightarrow{\tau} \delta_t$. Therefore, $t \xrightarrow{\hat{\tau}} \delta_t$ is syntactic sugar to denote that no silent-move took place. As in [26], this notation allows us to write $\pi \xrightarrow{\hat{\tau}} \pi'$ when only some of the processes in the support of π have the τ-transition.

Our definition of branching bisimulation is equivalent to the *scheduler-free* version defined in [5, 61].

Definition 5 (Probabilistic branching bisimulation)**.** For a PTS $(\mathbf{T}(\Sigma), \mathcal{A}_\tau, \rightarrow)$, a symmetric relation $\mathcal{B} \subseteq \mathbf{T}(\Sigma) \times \mathbf{T}(\Sigma)$ is a *probabilistic branching bisimulation* if whenever $s\, \mathcal{B}\, t$ and $s \xrightarrow{\alpha} \pi_s$ then

- either $\alpha = \tau$ and $\pi_s\ \mathcal{B}^\dagger\ \delta_t$;
- or $t \xrightarrow{\hat{\varepsilon}} \pi \xrightarrow{\hat{\alpha}} \pi_t$ with $\delta_s\ \mathcal{B}^\dagger\ \pi$ and $\pi_s\ \mathcal{B}^\dagger\ \pi_t$.

Of course, in order to guarantee compositionality with respect to the nondeterminisitc choice operator, which is offered by most process algebras, the rootedness condition is necessary.

Definition 6 (Probabilistic rooted branching bisimulation)**.** Assume a PTS $(\mathbf{T}(\Sigma), \mathcal{A}_\tau, \rightarrow)$ and a branching bisimulation $\mathcal{B}$ on $\mathbf{T}(\Sigma)$. A symmetric relation $\mathcal{R} \subseteq \mathbf{T}(\Sigma) \times \mathbf{T}(\Sigma)$ is a *probabilistic rooted branching bisimulation* if, whenever $s\,\mathcal{R}\,t$, then if $s \xrightarrow{\alpha} \pi_s$ then there is a transition $t \xrightarrow{\alpha} \pi_t$ such that $\pi_s\ \mathcal{B}^\dagger\ \pi_t$.

The union of all (rooted) branching bisimulations is the greatest (rooted) branching bisimulation, denoted $\approx_{\mathrm{b}}$ (resp. $\approx_{\mathrm{rb}}$), called (*rooted*) *branching bisimilarity*. In [8] it is proved that in the classical non-probabilistic case, branching bisimilarity is an equivalence relation. A similar result is given in [5] for the alternating model of probabilistic processes [67]. In order to infer that the result holds also in the case of the PTS model we consider in the present paper, it is enough to rephrase the proof in [8].

2.3. PGSOS specification

PTSs are defined by means of SOS rules, which are syntax-driven inference rules allowing us to inductively infer the behavior of terms with respect to their structure. Here we consider rules in the *probabilistic GSOS* (PGSOS) format [17, 18], which allows for specifying most probabilistic process algebras [40, 44]. These rules are based on expressions of the form $t \xrightarrow{\alpha} \Theta$, with t a term and Θ a *distribution term*, which will instantiate to transitions through *substitutions*.

Distribution terms are defined assuming a countable set of *distribution variables* $\mathbf{V}_d$ disjoint from Σ and $\mathbf{V}_s$. We use $\mu, \nu, \ldots$ to range over $\mathbf{V}_d$ and ζ, ζ' to range over $\mathbf{V}_s \cup \mathbf{V}_d$. The set of *distribution terms* over Σ and the subsets of variables $V_s \subseteq \mathbf{V}_s$ and $V_d \subseteq \mathbf{V}_d$, notation $\mathsf{DT}(\Sigma, V_s, V_d)$, is the least set satisfying:

- $\{\delta_t \mid t \in \mathsf{T}(\Sigma, V_s)\} \subseteq \mathsf{DT}(\Sigma, V_s, V_d)$,
- $V_d \subseteq \mathsf{DT}(\Sigma, V_s, V_d)$,
- $f(\Theta_1, \ldots, \Theta_n) \in \mathsf{DT}(\Sigma, V_s, V_d)$ whenever $f \in \Sigma$ and $\Theta_i \in \mathsf{DT}(\Sigma, V_s, V_d)$,
- $\sum_{i \in I} p_i \Theta_i \in \mathsf{DT}(\Sigma, V_s, V_d)$ whenever $\Theta_i \in \mathsf{DT}(\Sigma, V_s, V_d)$ and $p_i \in (0, 1]$ with $\sum_{i \in I} p_i = 1$ and I finite.

We write $\mathbb{DT}(\Sigma)$ for $\mathsf{DT}(\Sigma, \mathbf{V}_s, \mathbf{V}_d)$, i.e. the set of all *open distribution terms*, and $\mathbf{DT}(\Sigma)$ for $\mathsf{DT}(\Sigma, \emptyset, \emptyset)$, i.e. the set of the *closed distribution terms*. We may write $p\Theta_1 + (1-p)\Theta_2$ for $\sum_{i \in \{1,2\}} p_i \Theta_i$ with $p = p_1$ and $(1-p) = p_2$. Notice that closed distribution terms denote probability distributions over $\mathsf{T}(\Sigma)$. We denote by $\mathrm{var}(\Theta)$ the set of variables occurring in distribution term Θ.

A *positive* (resp. *negative*) *literal* is an expression of the form $t \xrightarrow{\alpha} \Theta$ (resp. $t \stackrel{\alpha}{\not\longrightarrow}$) with $t \in \mathbb{T}(\Sigma)$, $\alpha \in \mathcal{A}_\tau$ and $\Theta \in \mathbb{DT}(\Sigma)$. The literals $t \xrightarrow{\alpha} \Theta$ and $t \stackrel{\alpha}{\not\longrightarrow}$, with the same term t in the left-hand side and the same action label α, are said to *deny* each other. In the following, we let $\mathbf{H}$ denote an arbitrary set of literals.

$$r_1 = \frac{x \xrightarrow{\alpha} \mu \quad y \stackrel{\alpha}{\not\longrightarrow}}{x +_p y \xrightarrow{\alpha} \mu} \qquad r_2 = \frac{x \stackrel{\alpha}{\not\longrightarrow} \quad y \xrightarrow{\alpha} \nu}{x +_p y \xrightarrow{\alpha} \nu} \qquad r_3 = \frac{x \xrightarrow{\alpha} \mu \quad y \xrightarrow{\alpha} \nu}{x +_p y \xrightarrow{\alpha} p\mu + (1-p)\nu}$$

$$r_4 = \frac{x \xrightarrow{\alpha} \mu \quad \alpha \notin B}{x \parallel_B y \xrightarrow{\alpha} \mu \parallel_B \delta_y} \qquad r_5 = \frac{y \xrightarrow{\alpha} \nu \quad \alpha \notin B}{x \parallel_B y \xrightarrow{\alpha} \delta_x \parallel_B \nu} \qquad r_6 = \frac{x \xrightarrow{\alpha} \mu \quad y \xrightarrow{\alpha} \nu \quad \alpha \in B}{x \parallel_B y \xrightarrow{\alpha} \mu \parallel_B \nu}$$

Table 1: The probabilistic alternative choice $+_p$ and the CSP-like parallel composition $\parallel_B$ operators.

Definition 7 (PGSOS rules, [17, 18])**.** A *PGSOS rule* r over a signature Σ has the form:

$$\frac{\{x_i \xrightarrow{\alpha_{i,m}} \mu_{i,m} \mid i \in I, m \in M_i\} \;\; \{x_i \stackrel{\alpha_{i,n}}{\not\longrightarrow} \mid i \in I, n \in N_i\}}{f(x_1, \ldots, x_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta}$$

where $f \in \Sigma$ is an operator of rank $\mathfrak{n}$, $I \subseteq \{1, \ldots, \mathfrak{n}\}$, M_i and N_i are finite sets of indexes, $\alpha_{i,m}, \alpha_{i,n}, \alpha \in \mathcal{A}_\tau$ are action labels, $x_i \in \mathbf{V}_s$ are state variables, $\mu_{i,m} \in \mathbf{V}_d$ are distribution variables and $\Theta \in \mathbb{DT}(\Sigma)$ is a distribution term. Furthermore: (i) all variables $x_1, \ldots, x_{\mathfrak{n}}$ are distinct, (ii) all distribution variables $\mu_{i,m}$ with $i \in I$ and $m \in M_i$ are distinct, (iii) $\mathrm{var}(\Theta) \subseteq \{\mu_{i,m} \mid i \in I, m \in M_i\} \cup \{x_1, \ldots, x_{\mathfrak{n}}\}$.

Constraints (i)–(iii) are inherited by the classic GSOS rules in [11] and are necessary to ensure that strong probabilistic bisimulation [60] is a congruence [13, 17, 18].

Definition 8 (PGSOS-TSS)**.** A *PGSOS-transition system specification (PGSOS-TSS)* is a tuple $P = (\Sigma, \mathcal{A}_\tau, R)$, with Σ a signature, $\mathcal{A}_\tau$ a countable set of actions and R a set of PGSOS rules.

For a PGSOS rule r, the positive (resp. negative) literals above the line are called the *positive* (resp. *negative*) *premises*, notation pprem(r) (resp. nprem(r)). The literal $f(x_1, \ldots, x_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta$ is called the *conclusion*, notation conc(r), the term $f(x_1, \ldots, x_{\mathfrak{n}})$ is called the *source* of the rule, notation src(r), and the distribution term Θ is called the *target* of the rule, notation trg(r).

***Example* 1.** The operators of probabilistic alternative composition $+_p$, with $p \in (0, 1]$, and of the CSP-like parallel composition with multi-party synchronization on actions in $B \subseteq \mathcal{A}$, are specified in Table 1. The probabilistic alternative composition $t +_p t'$ evolves to the probabilistic choice between a distribution reached by t (with probability p) and a distribution reached by t' (with probability $1 - p$) for actions which can be performed by both processes. For actions that can be performed by either only t or only t', the probabilistic alternative composition $t +_p t'$ behaves just like the nondeterministic alternative composition $t + t'$. The term $t \parallel_B t'$ describes multi-party synchronization where t and t' synchronize on actions in B and evolve independently for all other actions. Notice that since $B \subseteq \mathcal{A}$ it follows that $\tau \notin B$. ◀

A PTS is derived from a TSS through the notions of substitution and proof.

Let $\mathbf{V}$ denote the set of all variables $\mathbf{V} = \mathbf{V}_s \cup \mathbf{V}_d$. A *substitution* is a mapping $\sigma \colon \mathbf{V} \to \mathbb{T}(\Sigma) \cup \mathbb{DT}(\Sigma)$ with $\sigma(x) \in \mathbb{T}(\Sigma)$, if $x \in \mathbf{V}_s$, and $\sigma(\mu) \in \mathbb{DT}(\Sigma)$, if $\mu \in \mathbf{V}_d$. It extends to terms, distribution terms, literals and rules by element-wise application. A substitution is *closed* if it maps variables to closed terms. Henceforth, a *closed substitution* maps an open term in $\mathbb{T}(\Sigma)$ to a process and an open distribution term in $\mathbb{DT}(\Sigma)$ to a distribution over processes. A closed substitution instance of a literal is called a *closed literal.*

Definition 9 (Proof)**.** A *proof* from a PGSOS-TSS $P = (\Sigma, \mathcal{A}_\tau, R)$ of a closed literal ℓ is a well-founded, upwardly branching tree, with nodes labeled by closed literals, such that the root is labeled ℓ and, if ℓ' is the label of a node and $\mathbf{K}$ is the set of labels of the nodes directly above it, then:

- either ℓ' is positive and $\frac{\mathbf{K}}{\ell'}$ is a closed substitution instance of a rule in R,
- or ℓ' is negative and for each closed substitution instance of a rule in R whose conclusion denies ℓ', a literal in $\mathbf{K}$ denies one of its premises.

A closed literal ℓ is *provable* from P, notation $P \vdash \ell$, if there is a proof from P of ℓ.

We have that each PGSOS-TSS P is *strictly stratifiable* [46] which implies that P *induces a unique model* corresponding to the PTS $(\mathbf{T}(\Sigma), \mathcal{A}_\tau, \to)$ whose transition relation $\to$ contains exactly the closed positive literals provable from P. Moreover, the existence of a stratification implies that P is also *complete* [46], thus giving that for any term $t \in \mathbf{T}(\Sigma)$ and action label $\alpha \in \mathcal{A}_\tau$ either $P \vdash t \xrightarrow{\alpha} \pi$ for some $\pi \in \Delta(\mathbf{T}(\Sigma))$ or $P \vdash t \xrightarrow{\alpha}\!\!\!\!\not\,\,\,$. In particular, the notion of provability in Definition 9 (which is called *supported* in [46]) subsumes the *negation as failure* principle of [16] for the derivation of negative literals: for each closed term t we have that $P \vdash t \xrightarrow{\alpha}\!\!\!\!\not\,\,\,$ iff $P \not\vdash t \xrightarrow{\alpha} \pi$ for any distribution $\pi \in \Delta(\mathbf{T}(\Sigma))$, namely the PTS induced by P contains literals that do not deny each other [11].

Next, we introduce the notion of *disjoint extension* of a TSS which allows us to introduce new operators without affecting the behavior of those already specified.

Definition 10 (Disjoint extension)**.** A PGSOS-TSS $P' = (\Sigma', \mathcal{A}, R')$ is a *disjoint extension* of a PGSOS-TSS $P = (\Sigma, \mathcal{A}, R)$ if $\Sigma \subseteq \Sigma'$, $R \subseteq R'$ and R' introduces no new rule for any operator in Σ.

We recall now the notion of *liquid* and *frozen* arguments of an operator with respect to a given predicate. We remark that in these preliminary definitions we will refer to a generic predicate Γ on operators. In upcoming Sections 7 and 8 we will concretely instantiate Γ with special predicates $\aleph$ and Λ from [37], where Λ marks as liquid the running processes, namely the processes that have already started their execution, and $\aleph$ marks as liquid the ones that can start their execution immediately. For instance, in [37] both arguments of the classic nondeterministic choice operator $\_ + \_$ are marked as Λ-frozen, since in a given process $s + t$ both s and t cannot have already performed any transition, and $\aleph$-liquid, since both s and t are able to start running immediately. If we take the sequential operator $\_;\_$, then the first argument is $\aleph$-liquid and Λ-liquid, whereas the second one is $\aleph$-frozen, since in a given process $s;t$ it is possible that s is running, in the sense that it is the result of a transition, and that it can continue its execution, and we are sure that t cannot move, since it has to wait for the termination of s.

Definition 11 (Liquid and frozen arguments of operators [10])**.** Let Γ be an unary predicate on $\{(f, i) \mid f \in \Sigma, 1 \le i \le \mathfrak{n}\}$. If $\Gamma(f, i)$ then argument i of f is labeled as Γ-*liquid*, otherwise it is Γ-*frozen*. An occurrence of a variable x in a term t is Γ-*liquid* if either $t = x$, or $t = f(t_1, \dots, t_\mathfrak{n})$ and the occurrence is Γ-liquid in t_i for a Γ-liquid argument i of f; otherwise the occurrence of x is Γ-frozen.

A *context* is a term having one occurrence of the context symbol $[\,]$ as a subterm. A Γ-liquid *context* is a term where $[\,]$ appears at a Γ-liquid position. A predicate Γ is *universal* if it holds for all arguments of all operators in the considered signature.

In particular, in Sections 7 and 8, we will use the liquid / frozen differentiation of arguments of operators with respect to the predicate Λ to distinguish those process arguments that must satisfy the rootedness condition (namely the Λ-*frozen* ones) from those that can simply satisfy the weak branching condition (namely the Λ-*liquid* ones). Moreover, we will mark as $\aleph$-*frozen* the arguments of an operator f that cannot be tested in the premises of the rules defining f.

***Example* 2.** We mark the arguments of the probabilistic alternative choice operator $+_p$ both Λ-frozen and $\aleph$-liquid. In fact, considering the rules defining $+_p$ in Table1, they can both start their execution ($\aleph$-liquid) but none of them can be a process that already started its execution (Λ-frozen). Conversely, we mark both arguments of the multi-party synchronous parallel operator $\|_B$ as $(\aleph \cap \Lambda)$-liquid. ◀

Finally, we introduce *patience rules*, expressing that a term can mimic the τ-moves of its arguments.

Definition 12 (Patience rule)**.** A PGSOS rule is a *patience rule* for argument i of f if it is of the form

$$\frac{x_i \xrightarrow{\tau} \mu}{f(x_1, \dots, x_\mathfrak{n}) \xrightarrow{\tau} f(\delta_{x_1}, \dots, \delta_{x_{i-1}}, \mu, \delta_{x_{i+1}}, \dots, \delta_{x_\mathfrak{n}})}.$$

Given a predicate Γ, a patience rule for the i-th argument of f is called a Γ-*patient rule* if $\Gamma(f, i)$. A PTSS is said to be Γ-*patient* if it contains all the Γ-patience rules.

***Example* 3.** Consider the PGSOS rules r_4, r_5 defining $\|_B$ in Table 1 and recall that we mark both arguments of this operator as $(\aleph \cap \Lambda)$-liquid (cf. Example 2). As by definition $\tau \notin B$, by instantiating $\alpha = \tau$ such rules are $(\aleph \cap \Lambda)$-patient rules for the arguments of $\|_B$. ◀

3. Logical characterization

In this section we introduce the class of modal formulae that will allow us to characterize probabilistic (rooted) branching bisimilarity. We combine the logic used in [37] for the characterization of (rooted) branching bisimilarity in the non-probabilistic setting with the probabilistic choice operator of the probabilistic extension of HML from [24].

Definition 13 (Modal logic $\mathbb{L}$)**.** The modal logic $\mathbb{L} = \mathbb{L}^{\mathrm{s}} \cup \mathbb{L}^{\mathrm{d}}$ is given by the classes of *state formulae* $\mathbb{L}^{\mathrm{s}}$ and *distribution formulae* $\mathbb{L}^{\mathrm{d}}$ over $\mathcal{A}_\tau$ defined by

$$\mathbb{L}^{\mathrm{s}}: \quad \varphi ::= \quad \top \mid \neg\varphi \mid \textstyle\bigwedge_{j \in J} \varphi_j \mid \langle\alpha\rangle\psi \mid \langle\hat{\tau}\rangle\psi \mid \langle\varepsilon\rangle\psi$$
$$\mathbb{L}^{\mathrm{d}}: \quad \psi ::= \quad \textstyle\bigoplus_{i \in I} r_i\varphi_i \mid \textstyle\bigwedge_{j \in J} \psi_j$$

where: (i) φ ranges over $\mathbb{L}^{\mathrm{s}}$, (ii) ψ ranges over $\mathbb{L}^{\mathrm{d}}$, (iii) $\alpha \in \mathcal{A}_\tau$, (iv) I and J are at most countable sets of indexes, and (v) $r_i \in (0, 1]$ for each $i \in I$ and $\sum_{i \in I} r_i = 1$.

We let ϕ range over $\mathbb{L}$ and we write $\phi_1 \wedge \phi_2$ in place of $\bigwedge_{j \in \{1,2\}} \phi_j$, $r_1\varphi_1 \oplus r_2\varphi_2$ in place of $\bigoplus_{i \in I} r_i\varphi_i$ with $I = \{1, 2\}$, and $\langle\cdot\rangle\varphi$ for $\langle\cdot\rangle \bigoplus_{i \in I} r_i\varphi_i$ with $I = \{i\}$, $r_i = 1$ and $\varphi_i = \varphi$. Notice that instead of using $\top$ we could use $\bigwedge_\emptyset$. We decided to use $\top$ to improve readability.

Formulae are interpreted over a PTS. The meaning of Hennessy-Milner operators is as usual [24, 53].

The special modalities $\langle\hat{\tau}\rangle\psi$ and $\langle\varepsilon\rangle\psi$ allow us to capture the weak semantics. In detail, a process t satisfies the formula $\langle\hat{\tau}\rangle\psi$ iff either the distribution δ_t satisfies ψ or there is a transition $t \xrightarrow{\tau} \pi$ such that the distribution π satisfies ψ. This can be summarized by saying that there is a lifted transition $t \xrightarrow{\hat{\tau}} \pi$ with π satisfying ψ. We stress that the difference between state formulae $\langle\tau\rangle\psi$ and $\langle\hat{\tau}\rangle\psi$ is in that the former is satisfied only if the process *actually performs* a τ-move, whereas the latter does not require such a step. In particular, $\langle\tau\rangle\psi$ will be used to guarantee the rootedness condition on the silent step τ.

The formulae of the form $\langle\varepsilon\rangle\psi$ will be used to capture the *weak* property of branching bisimulation: processes may execute an arbitrary, and potentially different, number of τ steps before reaching equivalent distributions. Moreover, the latter ones can be obtained as a convex combination of the distributions reached via the sequence of τ steps and thus we consider lifted transitions. Hence, t satisfies $\langle\varepsilon\rangle\psi$ if and only if there is a (possibly empty) sequence of n lifted transitions $t \xrightarrow{\hat{\tau}} \pi_1 \xrightarrow{\hat{\tau}} \pi_2 \ldots \xrightarrow{\hat{\tau}} \pi_n$, for some $n \in \mathbb{N}$, such that π_n satisfies the formula ψ. Notice that this can be restated as $t \xrightarrow{\hat{\varepsilon}} \pi$ for a distribution π with $\pi \models \psi$, or, equivalently, $t \xrightarrow{\hat{\tau}} \pi'$ for a distribution π' such that $\pi' \xrightarrow{\hat{\varepsilon}} \pi$ for a distribution π with $\pi \models \psi$.

Finally, a probability distribution π satisfies the distribution formula $\psi = \bigoplus_{i \in I} r_i\varphi_i$ if π can be rewritten as a convex combination of distributions π_i, using the r_i as weights of the combination, such that all the processes in the support of π_i satisfy the formula φ_i. This is formalized by requiring that there is a matching $\mathfrak{w}$ for π and ψ such that $t \models \varphi_i$ whenever $\mathfrak{w}(t, \varphi_i) > 0$.

Definition 14 (Semantics of $\mathbb{L}$)**.** The *satisfaction relation* $\models \subseteq (\mathbf{T}(\Sigma) \times \mathbb{L}^{\mathrm{s}}) \cup (\Delta(\mathbf{T}(\Sigma)) \times \mathbb{L}^{\mathrm{d}})$ is defined by structural induction on formulae by:

- $t \models \top$ always;
- $t \models \neg\varphi$ iff $t \models \varphi$ does not hold;
- $t \models \bigwedge_{j \in J} \varphi_j$ iff $t \models \varphi_j$ for all $j \in J$;
- $t \models \langle\alpha\rangle\psi$ iff there is a transition $t \xrightarrow{\alpha} \pi$ for a distribution $\pi \in \Delta(\mathbf{T}(\Sigma))$ with $\pi \models \psi$;

- $t \models \langle\hat{\tau}\rangle\psi$ iff there is a lifted transition $t \xrightarrow{\hat{\tau}} \pi$ for a distribution $\pi \in \Delta(\mathbf{T}(\Sigma))$ with $\pi \models \psi$;
- $t \models \langle\varepsilon\rangle\psi$ iff there is a lifted transition $t \xrightarrow{\hat{\varepsilon}} \pi$ for a distribution $\pi \in \Delta(\mathbf{T}(\Sigma))$ with $\pi \models \psi$;
- $\pi \models \bigoplus_{i\in I} r_i\varphi_i$ iff there is a matching $\mathfrak{w} \in \mathfrak{W}(\pi, \bigoplus_{i\in I} r_i\varphi_i)$ with $t \models \varphi_i$ whenever $\mathfrak{w}(t,\varphi_i) > 0$;
- $\pi \models \bigwedge_{j\in J} \psi_j$ iff $\pi \models \psi_j$ for all $j \in J$.

As usual we denote by $\mathbb{L}^{\mathrm{s}}(t)$ the set of state formulae in $\mathbb{L}^{\mathrm{s}}$ which are satisfied by t, namely $\mathbb{L}^{\mathrm{s}}(t) = \{\varphi \in \mathbb{L}^{\mathrm{s}} \mid t \models \varphi\}$. Moreover, we denote by $\mathbb{L}^{\mathrm{d}}(\pi)$ the set of distribution formulae in $\mathbb{L}^{\mathrm{d}}$ which are satisfied by π, namely $\mathbb{L}^{\mathrm{d}}(\pi) = \{\psi \in \mathbb{L}^{\mathrm{d}} \mid \pi \models \psi\}$. Then, we write $\varphi \equiv \varphi'$ if φ and φ' are *equivalent*, namely if it holds that $t \models \varphi$ iff $t \models \varphi'$ for all $t \in \mathbf{T}(\Sigma)$. Similarly, $\psi \equiv \psi'$ holds if $\pi \models \psi$ iff $\pi \models \psi'$ for all $\pi \in \Delta(\mathbf{T}(\Sigma))$. We let $\mathbb{L}^{\equiv}$ denote the quotient of the class of formulae $\mathbb{L}$ with respect to $\equiv$.

***Example* 4.** Consider process s and formulae $\varphi_1, \varphi_2, \varphi_3$ in Figure 1. We have $s \models \varphi_1$, $s \models \varphi_2$, and $s \not\models \varphi_3$.

Let us start with φ_1. It is enough to notice that $s \xrightarrow{\hat{\tau}} 1/2\delta_{s_1} + 1/2\delta_{s_2} \xrightarrow{\hat{\tau}} 1/2\delta_{s_3} + 1/2\delta_{s_2} = \pi_1$, where δ_{s_2} is not forced to perform any $\hat{\tau}$-move together with δ_{s_1}. It is immediate to verify that $s_3 \models \langle a\rangle\top$ and $s_2 \models \langle b\rangle\top$, so that we can infer $\pi_1 \models 1/2\langle a\rangle\top \oplus 1/2\langle b\rangle\top$. Hence, $s \xrightarrow{\hat{\varepsilon}} \pi_1$ allows us to infer that $s \models \varphi_1$.

Consider the formula φ_2. It requires s to perform a silent step (notice that the first τ does not have a hat) and reach a distribution satisfying the formula $\psi_2 = 1/2\langle\hat{\tau}\rangle\langle a\rangle\top \oplus 1/2\langle\hat{\tau}\rangle\langle b\rangle\top$. Clearly, $s \xrightarrow{\tau} 1/2\delta_{s_1} + 1/2\delta_{s_2} = \pi_2$. Now, $s_1 \xrightarrow{\tau} \delta_{s_3} \xrightarrow{\hat{a}} \delta_{\mathrm{nil}}$ and therefore $s_1 \models \langle\hat{\tau}\rangle\langle a\rangle\top$. Similarly, $s_2 \xrightarrow{b} \delta_{\mathrm{nil}}$ implies $s_2 \xrightarrow{\hat{\tau}} \delta_{s_2} \xrightarrow{\hat{b}} \delta_{\mathrm{nil}}$ (recall that the lifted transition $s_2 \xrightarrow{\hat{\tau}} \delta_{s_2}$ is only syntactic sugar) so that $s_2 \models \langle\hat{\tau}\rangle\langle b\rangle\top$. As π_2 assigns probability $1/2$ to both s_1 and s_2, we can conclude that $\pi_2 \models \psi_2$ and thus $s \models \varphi_2$.

Finally, we deal with φ_3. Notice that such formula does impose to s to perform a τ-move (the first τ in fact has a hat) but requires the processes in the eventually reached distribution to perform either a silent step followed by an a, or a silent step followed by a b, both with probability $1/2$. We can distinguish two cases:

- s does not perform the silent step, and thus $s \xrightarrow{\hat{\tau}} \delta_s$. Hence δ_s should satisfy the distribution formula $\psi_3 = 1/2\langle\tau\rangle\langle a\rangle\top \oplus 1/2\langle\tau\rangle\langle b\rangle\top$ which, implicitly, means that s has to satisfy both $\varphi_3^a = \langle\tau\rangle\langle a\rangle\top$ and $\varphi_3^b = \langle\tau\rangle\langle b\rangle\top$. However, s satisfies neither of them as it only has probability $1/2$ to perform b after the τ, and thus $s \not\models \varphi_3^b$, and it needs two τ-steps to reach a process that performs a and again with only probability $1/2$, and thus $s \not\models \varphi_3^a$. Hence $\delta_s \not\models \psi_3$.
- s performs the silent step, and thus $s \xrightarrow{\tau} \pi_2$, with π_2 the distribution $\pi_2 = 1/2\delta_{s_1} + 1/2\delta_{s_2}$. Therefore, we need to verify whether $\pi_2 \models \psi_3$. Clearly, $s_1 \models \langle\tau\rangle\langle a\rangle\top$. However, $s_2 \not\models \langle\tau\rangle\langle b\rangle\top$, since it cannot perform the required silent step. Hence $\pi_2 \not\models \psi_3$.

From these two cases we infer $s_3 \not\models \varphi_3$. ◀

Now we introduce two subclasses of $\mathbb{L}$: the class $\mathbb{L}_{\mathrm{b}}$ of *branching formulae* and the class $\mathbb{L}_{\mathrm{rb}}$ of *rooted branching formulae*, that characterize branching and rooted branching bisimilarity, respectively.

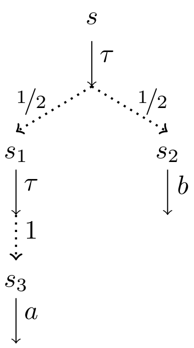
Definition 15 (Subclasses $\mathbb{L}_{\mathrm{b}}$ and $\mathbb{L}_{\mathrm{rb}}$)**.** The classes of *branching formulae* $\mathbb{L}_{\mathrm{b}} = \mathbb{L}_{\mathrm{b}}^{\mathrm{s}} \cup \mathbb{L}_{\mathrm{b}}^{\mathrm{d}}$ and of *rooted branching formulae* $\mathbb{L}_{\mathrm{rb}} = \mathbb{L}_{\mathrm{rb}}^{\mathrm{s}} \cup \mathbb{L}_{\mathrm{b}}^{\mathrm{d}}$ over $\mathcal{A}_\tau$ are defined by:

$$\mathbb{L}_{\mathrm{b}}^{\mathrm{s}}: \quad \varphi ::= \quad \top \mid \neg\varphi \mid \textstyle\bigwedge_{j\in J}\varphi_j \mid \langle\varepsilon\rangle 1\big(\varphi \wedge \langle a\rangle\psi\big) \mid \langle\varepsilon\rangle 1\big(\varphi \wedge \langle\hat{\tau}\rangle\psi\big) \quad (\text{with } \psi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{d}})$$

$$\mathbb{L}_{\mathrm{b}}^{\mathrm{d}}: \quad \psi ::= \quad \textstyle\bigoplus_{i\in I} r_i\varphi_i \mid \bigwedge_{j\in J}\psi_j \quad (\text{with } \varphi_i \in \mathbb{L}_{\mathrm{b}}^{\mathrm{s}})$$

and

$$\mathbb{L}_{\mathrm{rb}}^{\mathrm{s}}: \quad \varphi ::= \quad \top \mid \neg\varphi \mid \textstyle\bigwedge_{j\in J}\varphi_j \mid \langle\alpha\rangle\psi \mid \tilde{\varphi} \quad (\text{with } \tilde{\varphi} \in \mathbb{L}_{\mathrm{b}}^{\mathrm{s}} \text{ and } \psi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{d}})$$



$$\varphi_1 = \langle \varepsilon \rangle \, (1/2\langle a \rangle \top \oplus 1/2\langle b \rangle \top)$$

$$\varphi_2 = \langle \tau \rangle \, (1/2\langle \hat{\tau} \rangle\langle a \rangle \top \oplus 1/2\langle \hat{\tau} \rangle\langle b \rangle \top)$$

$$\varphi_3 = \langle \hat{\tau} \rangle \, (1/2\langle \tau \rangle\langle a \rangle \top \oplus 1/2\langle \tau \rangle\langle b \rangle \top)$$

Figure 1: Process s satisfies φ_1 and φ_2 but not φ_3. An arrow $u \xrightarrow{a}$ with no target models an a-step by u to the Dirac distribution δ_{nil}, with nil the process that can execute no action.

Informally, in a formula of the form $\langle \varepsilon \rangle 1(\varphi \wedge \langle a \rangle \psi)$ the formula φ plays the role of a guard for branching bisimulation. Given processes $s \approx_{\mathrm{b}} t$, if $s \xrightarrow{a} \pi$ then all the processes (recall that the probability weight is 1) in the support of the distribution π' reached by t after the possibly empty sequence of $\hat{\tau}$-moves have to be branching bisimilar to δ_s. This is due to the notion of lifting of a relation, in Definition 3. Thus, they all have to: (i) satisfy the same formulae satisfied by s, of which φ is a representative, and (ii) mimic the execution of a by s reaching a distribution satisfying the same formulae satisfied by π, of which ψ is a representative. The idea behind formulae $\langle \varepsilon \rangle 1(\varphi \wedge \langle \hat{\tau} \rangle \psi)$ is similar, the only difference being that equivalent processes have to mimic the execution of the *lifted* transition $\xrightarrow{\hat{\tau}}$. We notice that we can combine formulae of this kind to obtain the same expressive power of the linear formula $\varphi \, \mathbb{U} \, \langle a \rangle \psi$, using the 'until' operator $\mathbb{U}$, whose non-probabilistic version was introduced in [20] to characterize branching bisimilarity.

Theorem 1 (Logical characterization)**.** *Let $P = (\mathbf{T}(\Sigma), \mathcal{A}, \rightarrow)$ be a PTS and $s, t \in \mathbf{T}(\Sigma)$. Then:*

1. *$s \approx_{\mathrm{b}} t$ if and only if $\mathbb{L}_{\mathrm{b}}(s) = \mathbb{L}_{\mathrm{b}}(t)$.*
2. *$s \approx_{\mathrm{rb}} t$ if and only if $\mathbb{L}_{\mathrm{rb}}(s) = \mathbb{L}_{\mathrm{rb}}(t)$.*

We delay the comparison with the logical characterization of (rooted) branching bisimilarity proposed in [62] until Section 10.

4. Distribution specifications

The decomposition of state formulae, expressing properties over terms, will be given in Section 6 and relies on a collection of rules extracted from the TSS, called *ruloids* [11]. Essentially, a ruloid is an SOS rule with an arbitrary term not necessarily of the form $f(x_1, \ldots, x_{\mathfrak{n}})$ as source and with variables occurring in that term as source variables. Intuitively, ruloids relate the behavior of any term with the behavior of its variables. Based on ruloids, any formula for a term can be decomposed into formulae for its variables.

In order to have a similar method for distribution formulae, expressing properties over distribution terms, we need a notion of ruloid relating the *behavior* of any distribution term with the behavior of its variables, so that any formula for a distribution term can be decomposed into formulae for its variables. By behavior of a distribution term we intend the probability weights it assigns to terms. Notice that the behavior of *closed* distribution terms may be defined by a simple function, defined inductively starting from Dirac distributions and having as inductive steps those for convex combinations and for lifting of operators in Σ to distributions. However, this would not allow us to express the relations between the behavior of *open* distribution terms and the behavior of their variables. To this purpose, we need an SOS-like machinery. We developed such a machinery in [14], and we dedicate this section to recall it.

The first target is to have a set of SOS inference rules allowing us to infer the expression $\Theta \xrightarrow{q} t$ whenever a closed distribution term Θ assigns probability weight q to a closed term t.

A *distribution literal* is an expression of the form $\Theta \xrightarrow{q} t$, with $\Theta \in \mathbb{DT}(\Sigma)$, $q \in (0,1]$ and $t \in \mathbb{T}(\Sigma)$. A set of distribution literals $\{\Theta \xrightarrow{q_i} t_i \mid i \in I\}$ is called a *distribution over terms* if $\sum_{i \in I} q_i = 1$ and the terms t_i are all syntactically distinct. This expresses that Θ is the distribution over $\mathbb{T}(\Sigma)$ giving weight q_i to t_i.

Then, Σ-*distribution rules* are inference rules allowing us to infer distributions over terms $\{\Theta \xrightarrow{q_i} t_i \mid i \in I\}$ inductively with respect to the structure of Θ. Let $\delta_{\mathbf{V}_s} := \{\delta_x \mid x \in \mathbf{V}_s\}$ denote the set of all Dirac distribution terms with a variable as term, and $\vartheta, \vartheta_i, \ldots$ denote distribution terms in $\mathbb{DT}(\Sigma)$ ranging over $\mathbf{V}_d \cup \delta_{\mathbf{V}_s}$. Then, for arbitrary sets $S_1, \ldots, S_n$, we denote by $\bigtimes_{i=1}^{n} S_i$ the set of tuples $k = [s_1, \ldots, s_n]$ with $s_i \in S_i$. The i-th element of k is denoted $k(i)$.

Definition 16 (Σ-distribution rule and Σ-distribution specification [13])**.** Assume a signature Σ. The set R_Σ of the Σ-*distribution rules* consists of the least set containing the following inference rules:

1. $\dfrac{}{\{\delta_x \xrightarrow{1} x\}}$ for any state variable $x \in \mathbf{V}_s$;

2. $$\frac{\bigcup_{i=1,\ldots,\mathfrak{n}} \Big\{\vartheta_i \xrightarrow{q_{i,j}} x_{i,j} \mid j \in J_i, \sum_{j \in J_i} q_{i,j} = 1\Big\}}{\Big\{f(\vartheta_1, \ldots, \vartheta_{\mathfrak{n}}) \xrightarrow{q_k} f(x_{1,k(1)}, \ldots, x_{\mathfrak{n},k(\mathfrak{n})}) \mid q_k = \prod_{i=1,\ldots,\mathfrak{n}} q_{i,k(i)},\ k \in \bigtimes_{i=1,\ldots,\mathfrak{n}} J_i\Big\}}$$
where $f \in \Sigma$, the distribution terms $\vartheta_i \in \mathbf{V}_d \cup \delta_{\mathbf{V}_s}$ are all distinct, and the state variables $x_{i,j}$ with $j \in J_i$ and $i = 1, \ldots, \mathfrak{n}$ are pairwise distinct;

3. $$\frac{\bigcup_{i \in I} \Big\{\vartheta_i \xrightarrow{q_{i,j}} x_{i,j} \mid j \in J_i, \sum_{j \in J_i} q_{i,j} = 1\Big\}}{\Big\{\sum_{i \in I} p_i \vartheta_i \xrightarrow{q_x} x \mid q_x = \sum_{i \in I, j \in J_i \text{ s.t. } x_{i,j} = x} p_i \cdot q_{i,j} \text{ and } x \in \{x_{i,j} \mid j \in J_i, i \in I\}\Big\}}$$
where I is a countable set of indexes, the distribution terms $\vartheta_i \in \mathbf{V}_d \cup \delta_{\mathbf{V}_s}$ are all distinct, and the state variables $x_{i,j}$ with $j \in J_i$ and $i \in I$ are pairwise distinct.

Then, the Σ-*distribution specification* (Σ-*DS*) is the pair $D_\Sigma = (\Sigma, R_\Sigma)$.

For each Σ-distribution rule r_{D}, all sets above the line are called *premises*, notation $\mathrm{prem}(r_{\mathsf{D}})$, and the set below the line is called *conclusion*, notation $\mathrm{conc}(r_{\mathsf{D}})$. Clearly, all premises and the conclusion are distributions over terms. This was formally showed in [14].

***Example* 5.** An example of a Σ-distribution rule having as source the distribution term $\mu \parallel_B \nu$, with $\parallel_B$ the parallel composition operator explained in Example 1, is the following:

$$\frac{\{\mu \xrightarrow{1/4} x_1, \quad \mu \xrightarrow{3/4} x_2\} \quad \{\nu \xrightarrow{1/3} y_1, \quad \nu \xrightarrow{2/3} y_2\}}{\{\mu \parallel_B \nu \xrightarrow{1/12} x_1 \parallel_B y_1, \quad \mu \parallel_B \nu \xrightarrow{1/6} x_1 \parallel_B y_2, \quad \mu \parallel_B \nu \xrightarrow{1/4} x_2 \parallel_B y_1, \quad \mu \parallel_B \nu \xrightarrow{1/2} x_2 \parallel_B y_2\}}$$

◀

The following notion of reduction with respect to a substitution allows us to extend the notion of substitution to distributions over terms and to Σ-distribution rules.

Definition 17 (Reduction with respect to a substitution)**.** Assume a substitution σ and a distribution over terms $L = \{\Theta \xrightarrow{q_i} t_i \mid i \in I\}$. We say that σ *reduces* L to the set of distribution literals $L' = \{\sigma(\Theta) \xrightarrow{q_j} t_j \mid j \in J\}$, or that L' is the *reduction with respect to* σ of L, notation $\sigma(L) = L'$, if:

1. for each $j \in J$ there is at least one $i \in I$ with $\sigma(t_i) = t_j$;

2. the terms $\{t_j \mid j \in J\}$ are pairwise distinct;

3. for each $j \in J$, it holds $q_j = \sum_{\{i \in I \mid \sigma(t_i) = t_j\}} q_i$.

For a Σ-distribution rule r_{D}, we call its reduction with respect to σ the *reduced instance* of r_{D} wrt. σ.

Definition 18 (Proof from the Σ-DS)**.** A *proof* from the Σ-DS D_Σ of a closed distribution over terms L is a well-founded, upwardly branching tree, whose nodes are labeled by closed distributions over terms, such that the root is labeled L, and, if β is the label of a node and $\mathbf{K}$ is the set of labels of the nodes directly above it, then $\frac{\mathbf{K}}{\beta}$ is a closed reduced instance of a Σ-distribution rule in R_Σ.
A closed distribution over terms L is *provable* from D_Σ, written $D_\Sigma \vdash L$, if there is a proof from D_Σ for L.

Since Σ-distribution rules have only positive premises, the set of the distributions over terms provable from the Σ-DS is trivially unique. In [14] we proved that the set of the distributions over terms provable from the Σ-DS give the correct behavior of all closed distribution terms.

Proposition 1 ([14])**.** *Assume a signature Σ. Let $\pi \in \mathbf{DT}(\Sigma)$ be a closed distribution term and $\{t_m\}_{m\in M} \subseteq \mathbf{T}(\Sigma)$ a set of pairwise distinct closed terms. Then*

$$D_\Sigma \vdash \{\pi \xrightarrow{q_m} t_m \mid m \in M\} \Leftrightarrow \textit{for all } m \in M \textit{ it holds that } \pi(t_m) = q_m \textit{ and } \sum_{m\in M} q_m = 1.$$

We conclude this section by introducing the novel notion of *liquid* and *frozen* arguments of a distribution term with respect to a given predicate. This is obtained by lifting the labeling of arguments of operators to arguments of distribution terms.

Definition 19 (Liquid and frozen arguments of distribution terms)**.** Let Γ be a unary predicate on $\{(f,i) \mid f \in \Sigma, 1 \le i \le \mathfrak{n}\}$, Θ a distribution term, $x \in \mathbf{V}_s$ a state variable and $\mu \in \mathbf{V}_d$ a distribution variable.
An occurrence of x in Θ is Γ-*liquid* if: (i) either $\Theta = \delta_t$ and x occurs Γ-liquid in t, or (ii) $\Theta = f(\Theta_1,\dots,\Theta_{\mathfrak{n}})$ and the occurrence of x is Γ-liquid in Θ_i for a Γ-liquid argument i of f, or (iii) $\Theta = \sum_{i\in I} p_i\Theta_i$ and the occurrence of x is Γ-liquid in some Θ_i. Otherwise, the occurrence of x in Θ is Γ-*frozen*.
Then, an occurrence of μ in Θ is Γ-*liquid* if: (i) either $\Theta = \mu$, or (ii) $\Theta = f(\Theta_1,\dots,\Theta_{\mathfrak{n}})$ and the occurrence of μ is Γ-liquid in Θ_i for a Γ-liquid argument i of f, or (iii) $\Theta = \sum_{i\in I} p_i\Theta_i$ and the occurrence of μ is Γ-liquid in some Θ_i. Otherwise, the occurrence of μ in Θ is Γ-*frozen*.

***Example* 6.** By using the marking for the arguments of operator $\|_B$ given in Example 2, we mark the occurrences of μ and ν in the term $\mu \|_B \nu$ used in Example 5 as $(\aleph \cap \Lambda)$-liquid. ◀

A *distribution context* is a distribution term having one occurrence of the context symbol [ ] as a subterm. A Γ-liquid *distribution context* is a distribution term where [ ] appears at a Γ-liquid position. The notion of a Γ-liquid distribution context generalizes the notion of *w-nested context* from [61].

Remark 1. Consider any Σ-distribution rule $\frac{\{\mu \xrightarrow{q_i} x_i \mid i\in I\} \cup \mathbf{H}}{\{\Theta \xrightarrow{q_m} t_m \mid m\in M\}}$ and assume that μ occurs Γ-liquid in Θ. Then, for each $i \in I$, there is at least one target t_m such that x_i occurs Γ-liquid in t_m.

5. Ruloids and distribution ruloids

We dedicate this section to the presentation of *ruloids* [10, 11, 13] and *distribution ruloids* [13], which will play a central role in the modal decomposition in Section 6. Informally, as anticipated in Section 4, (distribution) ruloids are derived (distribution) rules, with an arbitrary (distribution) term as source, allowing us to infer the behavior of that (distribution) term directly from the behavior of the variables occurring in it. Both classes of ruloids have been proved to *sound and specifically witnessing* according to [11]: a closed literal ℓ (resp. a distribution over terms L_{D}) is provable from a PGSOS-TSS (resp. the Σ-DS) iff ℓ (resp. L_{D}) is an instance of the conclusion of a ruloid (resp. distribution ruloid) (Theorems 2–3 below).

We remark that the definition of a ruloid is in general quite technical. However, our choice of considering PGSOS specifications, in place of the more general ntμfθ/ntμxθ format of [18], allows us to give an explicit inductive construction technique for them, thus simplifying their presentation. Besides, we also notice that, as in the non-probabilistic setting, to derive ruloids from an ntμfθ/ntμxθ specification we would first have to transform such a specification into an equivalent PGSOS-like one (cf. [10]).

5.1. Ruloids

Ruloids are defined inductively from PGSOS rules. All PGSOS rules are ruloids. Then, from a rule r and a substitution σ, a ruloid ρ with conclusion $\sigma(\mathrm{conc}(r))$ is built as follows: 1) for each positive premise ℓ in $\sigma(r)$, we take any ruloid having ℓ as conclusion and we put its premises among the premises of ρ; 2) for each negative premise ℓ in $\sigma(r)$ and for each ruloid ρ' having a literal denying ℓ as conclusion, we select any premise ℓ_1 of ρ', we take any literal ℓ_2 denying ℓ_1, and we put ℓ_2 among the premises of ρ (we recall that two literals deny each other if they have the same term in the left-hand side, have the same label, one of them is positive and other is negative).

Definition 20 (Ruloids, [13])**.** Let $P = (\Sigma, \mathcal{A}_\tau, R)$ be a PGSOS-TSS. The set of P-*ruloids* $\Re^P$ is the smallest set such that:

- $\dfrac{x \xrightarrow{\alpha} \mu}{x \xrightarrow{\alpha} \mu}$ is a P-ruloid for all $x \in \mathbf{V}_s$, $\alpha \in \mathcal{A}_\tau$ and $\mu \in \mathbf{V}_d$;
- $\dfrac{\bigcup_{i \in I} \big(\bigcup_{m \in M_i} \mathbf{H}_{i,m} \cup \bigcup_{n \in N_i} \mathbf{H}_{i,n} \big)}{f(t_1, \dots, t_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta}$ is a P-ruloid if there is a PGSOS rule $r \in R$

$$\frac{\{x_i \xrightarrow{\alpha_{i,m}} \mu_{i,m} \mid i \in I, m \in M_i\} \;\; \{x_i \xrightarrow{\alpha_{i,n}}\!\!\!\!\!/\;\; \mid i \in I, n \in N_i\}}{f(x_1, \dots, x_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta'}$$

  together with a substitution σ, with $\sigma(x_i) = t_i$ for $i = 1, \dots, \mathfrak{n}$ and $\sigma(\Theta') = \Theta$, such that:

  - For every positive premise $x_i \xrightarrow{\alpha_{i,m}} \mu_{i,m}$ of r:
    - either $\sigma(x_i)$ is a variable and $\mathbf{H}_{i,m} = \{\sigma(x_i) \xrightarrow{\alpha_{i,m}} \sigma(\mu_{i,m})\}$,
    - or there is a P-ruloid $\rho_{i,m} = \dfrac{\mathbf{H}_{i,m}}{\sigma(x_i) \xrightarrow{\alpha_{i,m}} \sigma(\mu_{i,m})}$.
  - Right-hand side variables $\mathrm{rhs}(\rho_{i,m})$ are all pairwise disjoint.
  - For every negative premise $x_i \xrightarrow{\alpha_{i,n}}\!\!\!\!\!/\;\;$ of r:
    - either $\sigma(x_i)$ is a variable and $\mathbf{H}_{i,n} = \{\sigma(x_i) \xrightarrow{\alpha_{i,n}}\!\!\!\!\!/\;\;\}$,
    - or $\mathbf{H}_{i,n} = \mathrm{opp}(\mathrm{pick}(\Re^P_{(\alpha_{i,n})}))$, where
      i. $\Re^P_{(\alpha_{i,n})} \in \mathcal{P}(\mathcal{P}(\mathrm{Lit}(P)))$ is given by $\Re^P_{(\alpha_{i,n})} = \{\mathrm{prem}(\rho) \mid \rho \in \Re^P, \mathrm{conc}(\rho) = \sigma(x_i) \xrightarrow{\alpha_{i,n}} \theta, \theta \in \mathbb{DT}(\Sigma)\}$ of the sets of the premises of all P-ruloids with conclusion $\sigma(x_i) \xrightarrow{\alpha_{i,n}} \theta$, for some $\theta \in \mathbb{DT}(\Sigma)$;
      ii. pick: $\mathcal{P}(\mathcal{P}(\mathrm{Lit}(P))) \to \mathcal{P}(\mathrm{Lit}(P))$ is any mapping such that for any sets of literals L_k with $k \in K$, $\mathrm{pick}(\{L_k \mid k \in K\}) = \{l_k \mid k \in K \wedge l_k \in L_k\}$;
      iii. opp: $\mathcal{P}(\mathrm{Lit}(P)) \to \mathcal{P}(\mathrm{Lit}(P))$ is any mapping satisfying $\mathrm{opp}(t' \xrightarrow{\alpha} \theta) = t' \xrightarrow{\alpha}\!\!\!\!\!/\;\;$, and $\mathrm{opp}(t' \xrightarrow{\alpha}\!\!\!\!\!/\;\;) = t' \xrightarrow{\alpha} \theta$ for some fresh distribution term θ.

Example 7. Let P be any PGSOS-TSS containing the rules in Table 1. Assume the operator $\|_B$ with $B = \{a\}$. All the ruloids having as source the term $x +_p (y \,\|_{\{a\}}\, z)$ are in Table 2. We describe in detail the construction of the P-ruloids ρ_1 and ρ_6. Graphically, assuming $\beta \neq a$, their construction can be detailed as follows:

$$(\rho_1)\quad \frac{x \xrightarrow{\beta} \mu \qquad \dfrac{y \xrightarrow{\beta}\!\!\!\!\!/\;\; \quad z \xrightarrow{\beta}\!\!\!\!\!/\;\;}{y \,\|_{\{a\}}\, z \xrightarrow{\beta}\!\!\!\!\!/\;\;}}{x +_p (y \,\|_{\{a\}}\, z) \xrightarrow{\beta} \mu} \qquad\qquad (\rho_6)\quad \frac{x \xrightarrow{a} \mu \qquad \dfrac{y \xrightarrow{a}\!\!\!\!\!/\;\;}{y \,\|_{\{a\}}\, z \xrightarrow{a}\!\!\!\!\!/\;\;}}{x +_p (y \,\|_{\{a\}}\, z) \xrightarrow{a} \mu}$$

$$\rho_1 = \frac{x \xrightarrow{\beta} \mu \quad y \stackrel{\beta}{\not\longrightarrow} \quad z \stackrel{\beta}{\not\longrightarrow} \quad \beta \neq a}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{\beta} \mu} \qquad \rho_2 = \frac{x \stackrel{\beta}{\not\longrightarrow} \quad y \xrightarrow{\beta} \nu \quad \beta \neq a}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{\beta} \nu \parallel_{\{a\}} \delta_z} \qquad \rho_3 = \frac{x \stackrel{\beta}{\not\longrightarrow} \quad z \xrightarrow{\beta} \upsilon \quad \beta \neq a}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{\beta} \delta_y \parallel_{\{a\}} \upsilon}$$

$$\rho_4 = \frac{x \xrightarrow{\beta} \mu \quad y \xrightarrow{\beta} \nu \quad \beta \neq a}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{\beta} p\mu + (1-p)(\nu \parallel_{\{a\}} \delta_z)} \qquad \rho_5 = \frac{x \xrightarrow{\beta} \mu \quad z \xrightarrow{\beta} \upsilon \quad \beta \neq a}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{\beta} p\mu + (1-p)(\delta_y \parallel_{\{a\}} \upsilon)}$$

$$\rho_6 = \frac{x \xrightarrow{a} \mu \quad y \stackrel{a}{\not\longrightarrow}}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{a} \mu} \qquad \rho_7 = \frac{x \xrightarrow{a} \mu \quad z \stackrel{a}{\not\longrightarrow}}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{a} \mu} \qquad \rho_8 = \frac{x \stackrel{a}{\not\longrightarrow} \quad y \xrightarrow{a} \nu \quad z \xrightarrow{a} \upsilon}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{a} \nu \parallel_{\{a\}} \upsilon}$$

$$\rho_9 = \frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu \quad z \xrightarrow{a} \upsilon}{x +_p (y \parallel_{\{a\}} z) \xrightarrow{a} p\mu + (1-p)(\nu \parallel_{\{a\}} \upsilon)}$$

Table 2: The P-ruloids for the term $x +_p (y \parallel_{\{a\}} z)$.

Both ruloids are inferred starting from the rule r_1 in Table 1, with the source variable x in r_1 instantiated as x and the source variable y instantiated as $y \parallel_{\{a\}} z$. Then, the action α in r_1 is instantiated with any $\beta \neq a$ in ρ_1, whereas $\alpha = a$ in ρ_6. Both the premise $x \xrightarrow{\beta} \mu$ in the first case, and the premise $x \xrightarrow{a} \mu$ in the second case, have a variable as left-hand side, thus implying that the premise $x \xrightarrow{\beta} \mu$ appears in ρ_1 and the premise $x \xrightarrow{a} \mu$ appears in ρ_6. Conversely, the left-hand sides of the negative premises $y \parallel_{\{a\}} z \stackrel{\beta}{\not\longrightarrow}$ and $y \parallel_{\{a\}} z \stackrel{a}{\not\longrightarrow}$ are more structured terms and cannot directly appear in the ruloids. By Definition 20 we need to consider all the PGSOS rules having as conclusion instance $y \parallel_{\{a\}} z \xrightarrow{\beta} \Theta$, in the first case, and $y \parallel_{\{a\}} z \xrightarrow{a} \Theta$, in the second case, for some Θ in $\mathbb{DT}(\Sigma)$, namely any proper instance of rules $r_4 - r_6$ in Table 1. Hence, we need to distinguish two cases:

- $\alpha = \beta \neq a$. In this case, both rules r_4 and r_5 could be used in the derivation of literal $y \parallel_{\{a\}} z \xrightarrow{\beta} \Theta$. Hence we need to choose, and deny, one premise instance for each of those rules. Since r_4 and r_5 have a single premise each, we deny both of them, thus getting $y \stackrel{\beta}{\not\longrightarrow}$ and $z \stackrel{\beta}{\not\longrightarrow}$. Moreover, since the left-hand sides of both $y \stackrel{\beta}{\not\longrightarrow}$ and $z \stackrel{\beta}{\not\longrightarrow}$ are variables, these negative literals are premises in ρ_1.
- $\alpha = a$. In this case, the literal $y \parallel_{\{a\}} z \xrightarrow{a} \Theta$ can be derived only through the synchronization of terms y and z over action a, namely we need to apply rule r_6. Thus, we choose one of the premises for such rule, for instance the one having y as left-hand side, and we deny it. In our example, from this construction we obtain the single negative premise $y \stackrel{a}{\not\longrightarrow}$ whose left-hand side is a variable and can be a premise of the ruloid ρ_6. Notice that by taking $z \stackrel{a}{\not\longrightarrow}$ instead of $y \stackrel{a}{\not\longrightarrow}$ we get ρ_7 in Table 2.

◀

In [14] we showed that ruloids are sound and specifically witnessing.

Theorem 2 ([14])**.** *Assume a PGSOS-TSS P and a closed substitution σ. Then $P \vdash \sigma(t) \xrightarrow{a} \Theta'$ for any term $t \in \mathbb{T}(\Sigma)$ and closed distribution term $\Theta' \in \mathbf{DT}(\Sigma)$ iff there are a P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{a} \Theta}$ and a closed substitution σ' with $P \vdash \sigma'(\mathbf{H})$, $\sigma'(t) = \sigma(t)$ and $\sigma'(\Theta) = \Theta'$.*

Remark 2. By looking at Definition 20, we note that treatment of negative premises allows us to get also ruloids of the form $\frac{\mathbf{H}}{t \stackrel{a}{\not\longrightarrow}}$ that are usually called *non-standard* rules [10, 37]. Then, it is easy to see that, according to the negation as failure principle, by means of non-standard rules we obtain the analogous of Theorem 2 for closed instances of negative literals: $P \vdash \sigma(t) \stackrel{a}{\not\longrightarrow}$ for $t \in \mathbb{T}(\Sigma)$ and σ closed substitution iff there are a non-standard rule $\frac{\mathbf{H}}{t \stackrel{a}{\not\longrightarrow}}$ and a closed substitution σ' with $P \vdash \sigma'(\mathbf{H})$ and $\sigma'(t) = \sigma(t)$.

The notion of patience for rules can be extended to ruloids. Given a predicate Γ and a P-ruloid ρ with conclusion labeled τ, we say that ρ is Γ-*patient* iff it is of the form $\rho = \frac{x \xrightarrow{\tau} \mu}{C[x] \xrightarrow{\tau} C[\mu]}$, for a Γ-liquid context $C[\,]$. Otherwise, ρ is Γ-*impatient*. Notice that all Γ-patient PGSOS rules are Γ-patient ruloids.

5.2. Σ-distribution ruloids

Σ-distribution ruloids are a generalization of Σ-distribution rules and capture the behavior of arbitrary open distribution terms. More precisely, they allow us to infer the behavior of a (possibly open) distribution term as a probability distribution over terms from the distribution over terms that characterize the behavior of the variables occurring in it. For instance, distribution ruloids allow us to infer the behavior of a distribution term of the form $\frac{2}{3}\mu + \frac{1}{3}(\nu \parallel_B \upsilon)$ from the behavior of the variables μ, ν and υ. Notice that distribution rules are not enough to meet this purpose, since in the source of distribution rules only one operator over distributions is admitted, and therefore there is no Σ-distribution rule with source $\frac{2}{3}\mu + \frac{1}{3}(\nu \parallel_B \upsilon)$. Similarly to P-ruloids, a Σ-distribution ruloid is defined by an inductive composition of Σ-distribution rules and the left-hand sides of its premises are the variables occurring in the source, which is an arbitrary open distribution term. As the Σ-DS is positive, the definition of Σ-distribution ruloids is technically simpler than that of P-ruloids.

Definition 21 (Distribution ruloids, [13])**.** Let $D_\Sigma = (\Sigma, R_\Sigma)$ be the Σ-DS. The set of Σ-*distribution ruloids* $\Re^\Sigma$ is the smallest set such that:

- $\dfrac{\{\delta_x \xrightarrow{1} x\}}{\{\delta_x \xrightarrow{1} x\}}$ is a Σ-distribution ruloid in $\Re^\Sigma$ for any state variable $x \in \mathbf{V}_s$;
- $\dfrac{\{\mu \xrightarrow{q_i} x_i \mid \sum_{i \in I} q_i = 1\}}{\{\mu \xrightarrow{q_i} x_i \mid i \in I\}}$ is a Σ-distribution ruloid in $\Re^\Sigma$ for any distribution variable $\mu \in \mathbf{V}_d$;
- $\dfrac{\bigcup_{i=1,\dots,\mathfrak{n}} \mathbf{H}_i}{\Big\{f(\Theta_1,\dots,\Theta_\mathfrak{n}) \xrightarrow{Q_m} f(t_{1,m},\dots,t_{\mathfrak{n},m}) \mid m \in M\Big\}}$ is a Σ-distribution ruloid in $\Re^\Sigma$ if there are a substitution σ with $\sigma(\vartheta_i) = \Theta_i$ and a Σ-distribution rule $r_\mathsf{D} \in R_\Sigma$ as in Definition 16.2 such that:
  - $\sigma(r_\mathsf{D}) = \dfrac{\bigcup_{i=1,\dots,\mathfrak{n}} \{\Theta_i \xrightarrow{q_{i,h}} t_{i,h} \mid h \in H_i, \sum_{h \in H_i} q_{i,h} = 1\}}{\Big\{f(\Theta_1,\dots,\Theta_\mathfrak{n}) \xrightarrow{q_\kappa} f(t_{1,\kappa(1)},\dots,t_{\mathfrak{n},\kappa(\mathfrak{n})}) \mid q_\kappa = \prod_{i=1,\dots,\mathfrak{n}} q_{i,\kappa(i)},\ \kappa \in \bigtimes_{i=1,\dots,\mathfrak{n}} H_i\Big\}}$
  - there is a bijection $\mathfrak{f}\colon \bigtimes_{i=1}^{\mathfrak{n}} H_i \to M$ with $t_{i,\kappa(i)} = t_{i,\mathfrak{f}(\kappa)}$ and $q_\kappa = Q_{\mathfrak{f}(\kappa)}$;
  - for every Θ_i, for $i = 1,\dots,\mathfrak{n}$
    * either $\Theta_i \in \mathbf{V}_d \cup \delta_{\mathbf{V}_s}$ and $\mathbf{H}_i = \{\Theta_i \xrightarrow{q_{i,h}} t_{i,h} \mid h \in H_i, \sum_{h \in H_i} q_{i,h} = 1\}$,
    * or there is a Σ-distribution ruloid $\rho_i^\mathsf{D} = \dfrac{\mathbf{H}_i}{\{\Theta_i \xrightarrow{q_{i,h}} t_{i,h} \mid h \in H_i\}}$;
  - right-hand sides of premises are pairwise distinct.
- $\dfrac{\bigcup_{i \in I} \mathbf{H}_i}{\Big\{\sum_{i \in I} p_i \Theta_i \xrightarrow{Q_m} t_m \mid m \in M\Big\}}$ is a Σ-distribution ruloid in $\Re^\Sigma$ if there are a substitution σ with $\sigma(\vartheta_i) = \Theta_i$ for $i \in I$ and a Σ-distribution rule $r_\mathsf{D} \in R_\Sigma$ as in Definition 16.3 such that:
  - $\sigma(r_\mathsf{D}) = \dfrac{\bigcup_{i \in I} \{\Theta_i \xrightarrow{q_{i,h}} t_{i,h} \mid h \in H_i, \sum_{h \in H_i} q_{i,h} = 1\}}{\Big\{\sum_{i \in I} p_i \Theta_i \xrightarrow{q_u} u \mid q_u = \sum_{i \in I, h \in H_i \text{ s.t. } t_{i,h} = u} p_i \cdot q_{i,h},\ u \in \{t_{i,h} \mid h \in H_i, i \in I\}\Big\}}$
  - there is a bijection $\mathfrak{f}\colon \{t_{i,h} \mid h \in H_i, i \in I\} \to M$ with $u = t_{\mathfrak{f}(u)}$, and $q_u = Q_{\mathfrak{f}(u)}$;

– for every Θ_i, for $i \in I$

  * either $\Theta_i \in \mathbf{V}_d \cup \delta_{\mathbf{V}_s}$ and $\mathbf{H}_i = \{\Theta_i \xrightarrow{q_{i,h}} t_{i,h} \mid h \in H_i, \sum_{h \in H_i} q_{i,h} = 1\}$,
  * or there is a Σ-distribution ruloid $\rho_i^{\mathsf{D}} = \dfrac{\mathbf{H}_i}{\{\Theta_i \xrightarrow{q_{i,h}} t_{i,h} \mid h \in H_i\}}$;

– right-hand sides of premises are pairwise distinct.

***Example* 8.** Consider the distribution term $\frac{2}{3}\mu + \frac{1}{3}(\nu \parallel_B \delta_z)$ (which is an instance of the target of the P-ruloid ρ_4 in Table 2, for $p = 2/3$). Then, we can build the following Σ-distribution ruloid ρ^{D}:

$$\frac{\{\mu \xrightarrow{1/4} x_1 \quad \mu \xrightarrow{3/4} x_2\} \qquad \dfrac{\{\nu \xrightarrow{1/2} y_1, \quad \nu \xrightarrow{1/2} y_2\} \quad \{\delta_z \xrightarrow{1} z\}}{\{\nu \parallel_B \delta_z \xrightarrow{1/2} y_1 \parallel_B z \quad \nu \parallel_B \delta_z \xrightarrow{1/2} y_2 \parallel_B z\}}}{\left\{\frac{2}{3}\mu + \frac{1}{3}(\nu \parallel_B \delta_z) \xrightarrow{\frac{1}{6}} x_1,\ \frac{2}{3}\mu + \frac{1}{3}(\nu \parallel_B \delta_z) \xrightarrow{\frac{1}{2}} x_2,\ \frac{2}{3}\mu + \frac{1}{3}(\nu \parallel_B \delta_z) \xrightarrow{\frac{1}{6}} y_1 \parallel_B z,\ \frac{2}{3}\mu + \frac{1}{3}(\nu \parallel_B \delta_z) \xrightarrow{\frac{1}{6}} y_2 \parallel_B z\right\}}.$$

◀

In [14] we showed that also distribution ruloids are sound and specifically witnessing.

Theorem 3 ([14])**.** *Consider the Σ-DS D_Σ and a closed substitution σ. Then $D_\Sigma \vdash \{\sigma(\Theta) \xrightarrow{q_m} t_m \mid m \in M\}$ for a distribution term $\Theta \in \mathbb{DT}(\Sigma)$ and closed terms $t_m \in \mathbf{T}(\Sigma)$ pairwise distinct iff there are a Σ-distribution ruloid $\frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}}$ and a closed substitution σ' with $D_\Sigma \vdash \sigma'(\mathbf{H})$, $\sigma'(\Theta) = \sigma(\Theta)$ and $\sigma'(u_m) = t_m$ for all $m \in M$.*

The following two technical lemmas will support some of our proofs. They state that the conclusion of a Σ-distribution ruloid is a distribution over terms and that all variables appearing in a Σ-distribution ruloid already appear in its premises, respectively.

Lemma 1. *The conclusion of a Σ-distribution ruloid is a distribution over terms.*

Lemma 2. *Any Σ-distribution ruloid $\frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}}$ is such that*

1. *for all distribution variabels $\mu \in \mathbf{V}_d$, $\mu \in \mathrm{var}(\Theta)$ iff μ is the left-hand side of a premise in $\mathbf{H}$;*
2. *for all state variables $x \in \mathbf{V}_s$, $x \in \mathrm{var}(\Theta)$ iff δ_x is the left-hand side of a premise in $\mathbf{H}$;*
3. *$\bigcup_{m \in M} \mathrm{var}(t_m) = \mathrm{rhs}(\mathbf{H})$.*

6. Decomposition of modal formulae

In this section we present our method for decomposing formulae in $\mathbb{L}$, which exploits the two classes of ruloids introduced in Section 5. The idea behind the decomposition of a state (resp. distribution) formula ϕ with respect to a term t (resp. distribution term Θ) is to establish the constraints that the closed instances of the variables occurring in t (resp. Θ) must satisfy in order to ensure that the closed instance of t (resp. Θ) satisfies ϕ. Therefore, the decomposition method is firmly based on ruloids (resp. Σ-distribution ruloids), since they relate the behavior of t (resp. Θ) with that of its variables.

The decomposition of state formulae consists in assigning to each term $t \in \mathbb{T}(\Sigma)$ and formula $\varphi \in \mathbb{L}^{\mathrm{s}}$ a set of functions $\xi \colon \mathbf{V}_s \to \mathbb{L}^{\mathrm{s}}$, called *decomposition mappings*, assigning to each state variable x in t a proper formula $\xi(x)$ in $\mathbb{L}^{\mathrm{s}}$ such that for any closed substitution σ it holds that $\sigma(t) \models \varphi$ if and only if for some for some of the decomposition mappings ξ for ϕ and t it holds that $\sigma(x) \models \xi(x)$ for each $x \in \mathrm{var}(t)$ (Theorem 4). Each mapping ξ will be defined on a P-ruloid having t as source, P being the considered PGSOS-TSS, and on a predicate Γ on arguments of operators. The motivation for using Γ is that, besides having a correct decomposition, we also aim to infer a congruence result, which requires that decomposition preserves the modal class of formulae, namely formulae in $\mathbb{L}_{\mathrm{b}}$ (resp. $\mathbb{L}_{\mathrm{rb}}$) are decomposed into formulae in $\mathbb{L}_{\mathrm{b}}$ (resp. $\mathbb{L}_{\mathrm{rb}}$).

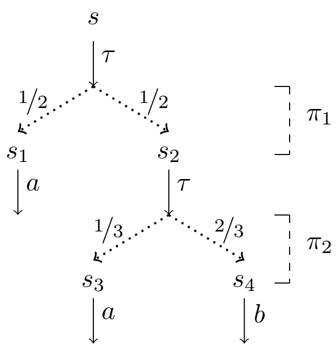


Figure 2: Process s satisfies the formula $\langle\varepsilon\rangle(2/3\langle a\rangle\top \oplus 1/3\langle b\rangle\top)$.

To this purpose, we will instantiate Γ as $\Gamma = \aleph \cap \Lambda$, for the predicates $\aleph$ and Λ discussed in Section 2.3. This particular instance will allow us to distinguish the arguments for which we must test the rootedness condition, namely the $(\aleph \cap \Lambda)$-frozen ones, from those for which the branching property by itself is sufficient, namely the $(\aleph \cap \Lambda)$-liquid ones.

Similarly, the decomposition of distribution formulae consists in assigning to each distribution term $\Theta \in \mathbb{DT}(\Sigma)$ and distribution formula $\psi \in \mathbb{L}^{\mathrm{d}}$ a set of decomposition mappings $\eta\colon \mathbf{V} \to \mathbb{L}^{\mathrm{d}} \cup \mathbb{L}^{\mathrm{s}}$ such that for any closed substitution σ we get that $\sigma(\Theta) \models \psi$ if and only if for some of the decomposition mappings η for Θ and ψ it holds that $\sigma(\zeta) \models \eta(\zeta)$ for each $\zeta \in \mathrm{var}(\Theta)$ (Theorem 4). Each mapping η will be defined on a Σ-distribution ruloid having Θ as source.

6.1. Decomposing $\langle\varepsilon\rangle\psi$

Our decomposition method is inspired by those in [37], dealing with *weak* semantics in the classic *nondeterministic* case, and [14], dealing with *strong* semantics in the *probabilistic* case. The main technical challenge caused by the interplay of weak semantics with probability is in dealing with formulae $\langle\varepsilon\rangle\psi$.

In the non-probabilistic setting of [37], the $\langle\varepsilon\rangle$-modality occurred in formulae of the form $\langle\varepsilon\rangle\phi$, with ϕ a *state formula*, and expressed a sequence of τ steps leading to a *process* satisfying ϕ. Hence, the decomposition of $\langle\varepsilon\rangle\phi$ was obtained from both Γ-patient and Γ-impatient P-ruloids allowing one to infer that sequence of τ-steps originating from the process satisfying $\langle\varepsilon\rangle\phi$. Our case is technically different since ϕ is a *distribution* formula that should be satisfied by a *distribution* whose support contains processes that may be reached by sequences of τ-steps of possibly different length originating from the state satisfying $\langle\varepsilon\rangle\phi$. More concretely, these sequences of silent steps give rise to a lifted transition $s \xrightarrow{\hat{\varepsilon}} \pi_n$ of the form $s \xrightarrow{\hat{\tau}} \pi_1 \xrightarrow{\hat{\tau}} \pi_2 \xrightarrow{\hat{\tau}} \dots \xrightarrow{\hat{\tau}} \pi_n$, where each π_i "memorizes" the probabilistic weights of the distributions reached so far through the execution of part of the sequence of $\hat{\tau}$-steps.

Consider, for instance, process $s \in \mathbf{T}(\Sigma)$ in Figure 2 for which we have $P \vdash s \xrightarrow{\tau} \pi_1$ with $\pi_1 = 1/2\delta_{s_1} + 1/2\delta_{s_2}$, $P \vdash s_1 \xrightarrow{a} \delta_{\mathrm{nil}}$, $P \vdash s_2 \xrightarrow{\tau} \pi_2$ with $\pi_2 = 1/3\delta_{s_3} + 2/3\delta_{s_4}$, $P \vdash s_3 \xrightarrow{a} \delta_{\mathrm{nil}}$ and $P \vdash s_4 \xrightarrow{b} \delta_{\mathrm{nil}}$. We have that $s \models \langle\varepsilon\rangle\psi$, for $\psi = 2/3\langle a\rangle\top \oplus 1/3\langle b\rangle\top$. This is given by the lifted transition $s \xrightarrow{\hat{\varepsilon}} \pi'$ with $\pi' = 1/2\delta_{s_1} + 1/6\delta_{s_3} + 1/3\delta_{s_4}$ for which $\pi' \models \psi$ clearly holds. We can observe that process s_1 must "wait" for s_2 to finish the execution of the sequence of τ-steps, thus reaching s_3 and s_4, and, then, s_1, s_3 and s_4 contribute to the satisfaction of the distribution formula ψ accordingly to the probabilistic weights "inherited" from π_1 and π_2, that is $1/2$ for s_1, $1/2 \cdot 1/3$ for s_3 and $1/2 \cdot 2/3$ for s_4.

In a broader sense, a formula $\langle\varepsilon\rangle\psi$ expresses a form of *probabilistic lookahead* on the behavior of processes, which is formally represented by sequences of lifted transitions of the form $\pi \xrightarrow{\hat{\tau}} \pi'$: we allow processes in the support of a distribution to perform some $\hat{\tau}$ steps in such a way that the global distribution that will be obtained, as described in Definition 4, will satisfy ψ.

The decomposition of formulae $\langle\varepsilon\rangle\psi$ must follow the same principle. To this aim, we recall that, as we discussed in Section 3, the semantics of $\langle\varepsilon\rangle\psi$ could be expressed by splitting the sequence of $\hat{\tau}$ steps into a first $\hat{\tau}$ transition $t \xrightarrow{\hat{\tau}} \pi$ performed by the process t, and a sequence of lifted transitions $\pi \xrightarrow{\hat{\varepsilon}} \pi'$ with $\pi' \models \psi$. Hence, we abuse of notation and introduce the distribution formulae $\psi^{\langle\varepsilon\rangle}$, with $\cdot^{\langle\varepsilon\rangle}$ a marker allowing us to record, during the decomposition, that the distribution formula ψ occurred in the scope of $\langle\varepsilon\rangle$. Hence, $\psi^{\langle\varepsilon\rangle}$ could be either of the form $(\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$ or $(\bigwedge_{j\in J} \bigoplus_{i_j\in I_j} r_{i_j}\varphi_{i_j})^{\langle\varepsilon\rangle}$, and we will use this kind of formulae to express that a closed distribution term performs a sequence of $\hat{\tau}$ lifted transitions by which it reaches a distribution satisfying the formula ψ. Actually, we could rewrite the semantics of $\langle\varepsilon\rangle\psi$ in Definition 14 as:

- $t \models \langle\varepsilon\rangle\psi$ iff $t \xrightarrow{\hat{\tau}} \pi$ and $\pi \models \psi^{\langle\varepsilon\rangle}$, for some $\pi \in \Delta(\mathbf{T}(\Sigma))$;
- $\pi \models \psi^{\langle\varepsilon\rangle}$ iff $\pi \xrightarrow{\hat{\varepsilon}} \pi'$ and $\pi' \models \psi$, for some $\pi' \in \Delta(\mathbf{T}(\Sigma))$.

The decomposition of $\langle\varepsilon\rangle\psi$ with respect to term t will be obtained via a P-ruloid allowing us to derive a silent step $t \xrightarrow{\hat{\tau}} \Theta$ and the decomposition of the formula $\psi^{\langle\varepsilon\rangle}$ with respect to Θ. The latter will be obtained via a Σ-distribution ruloid allowing us to derive the probabilistic behavior $\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}$ of Θ and, for each $m \in M$, the decomposition of a formula $\langle\varepsilon\rangle\psi_m$ with respect to term u_m. Here, $\langle\varepsilon\rangle$ allows for the derivation of the (possible) next τ-step in the sequence and ψ_m captures the probabilistic behavior that the distribution reached by u_m via the entire sequence of silent steps has to show in order to guarantee that the probabilistic behavior of the distribution reached by the initial term t via the entire sequence of silent steps is as in ψ. The construction of the formulae ψ_m will exploit the matching from Definition 2 to *keep memory* of the probabilistic behavior of Θ. The alternation in the decomposition of formulae of the form $\langle\varepsilon\rangle\tilde{\psi}$ and $\tilde{\psi}^{\langle\varepsilon\rangle}$ so obtained will allow us to properly decompose the probabilistic lookahead introduced by $\langle\varepsilon\rangle\psi$.

We remark that formulae $\psi^{\langle\varepsilon\rangle}$ occur only in the decomposition of formulae of the form $\langle\varepsilon\rangle\psi$ and play no role in the decomposition of other state or distribution formulae. We also stress that we do not allow a formula $\psi^{\langle\varepsilon\rangle}$ to be constructed over a nesting of conjunctions of distribution formulae. Such restriction has the only purpose to simplify the presentation of the decomposition method, and moreover it does not limit the expressive power of the logic, in that it is always possible to unfold the nested conjunctions in order to obtain a single conjunction on distribution formulae defined via the probabilistic choice operator.

6.2. The decomposition method

Given any term $t \in \mathbb{T}(\Sigma)$ and variables $x \in \mathrm{var}(t)$ and $\mu \in \mathbf{V}_d$, we denote by $t[\mu/x]$ the distribution term obtained by substituting x with μ and y with δ_y for all state variables $y \in \mathrm{var}(t) \setminus \{x\}$ in the term t.

Then, we consider the notion *matching for a distribution over terms and a distribution formula.*

Definition 22. Assume a distribution term $\Theta \in \mathbb{DT}(\Sigma)$, a distribution over terms $\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}$ and a formula $\psi = \bigoplus_{i\in I} r_i\varphi_i \in \mathbb{L}^{\mathrm{d}}$. Then a *matching* for Θ and ψ is a distribution over the product space $\mathfrak{w} \in \Delta(\mathbb{T}(\Sigma) \times \mathbb{L}^{\mathrm{s}})$ having Θ and ψ as left and right marginals, that is

- $\sum_{i\in I} \mathfrak{w}(t_m, \varphi_i) = q_m$, for all $m \in M$, and
- $\sum_{m\in M} \mathfrak{w}(t_m, \varphi_i) = r_i$, for all $i \in I$.

We denote by $\mathfrak{W}(\Theta, \psi)$ the set of all matchings for Θ and ψ.

To favor readability, we will provide a formal discussion of the decomposition method after its technical definition (Definition 23 below).

Definition 23 (Decomposition of formulae in $\mathbb{L}$)**.** Let $P = (\Sigma, \mathcal{A}, R)$ be a Γ-patient PGSOS-TSS and let D_Σ be the Σ-DS. We define the mappings

- $\cdot^{-1}\colon \mathbb{T}(\Sigma) \to (\mathbb{L}^{\mathrm{s}} \to \mathcal{P}(\mathbf{V}_s \to \mathbb{L}^{\mathrm{s}}))$, and
- $\cdot^{-1}\colon \mathbb{DT}(\Sigma) \to (\mathbb{L}^{\mathrm{d}} \to \mathcal{P}(\mathbf{V} \to \mathbb{L}))$

as follows. For each term $t \in \mathbb{T}(\Sigma)$ and state formula $\varphi \in \mathbb{L}^{\mathrm{s}}$, $t^{-1}(\varphi) \in \mathcal{P}(\mathbf{V}_s \to \mathbb{L}^{\mathrm{s}})$ is the set of *decomposition mappings* $\xi\colon \mathbf{V}_s \to \mathbb{L}^{\mathrm{s}}$ such that for any univariate term t we have:

1. $\xi \in t^{-1}(\top)$ iff $\xi(x) = \top$ for all state variables $x \in \mathbf{V}_s$.
2. $\xi \in t^{-1}(\neg\varphi)$ iff there is a function $\mathfrak{f}\colon t^{-1}(\varphi) \to \mathrm{var}(t)$ such that

$$\xi(x) = \begin{cases} \bigwedge_{\xi' \in \mathfrak{f}^{-1}(x)} \neg\xi'(x) & \text{if } x \in \mathrm{var}(t) \\ \top & \text{otherwise.} \end{cases}$$

3. $\xi \in t^{-1}(\bigwedge_{j\in J} \varphi_j)$ iff there is a decomposition mappings $\xi_j \in t^{-1}(\varphi_j)$ for each $j \in J$ such that

$$\xi(x) = \bigwedge_{j\in J} \xi_j(x), \text{ for all } x \in \mathbf{V}_s.$$

4. $\xi \in t^{-1}(\langle\alpha\rangle\psi)$ iff there are a P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\alpha} \Theta}$ and a decomposition mapping $\eta \in \Theta^{-1}(\psi)$ such that

$$\xi(x) = \begin{cases} \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \wedge \bigwedge_{x \not\xrightarrow{\gamma} \in \mathbf{H}} \neg\langle\gamma\rangle\top \wedge \eta(x) & \text{if } x \in \mathrm{var}(t) \\ \top & \text{otherwise.} \end{cases}$$

5. $\xi \in t^{-1}(\langle\hat{\tau}\rangle\psi)$ iff one of the following three cases holds:
   (a) there is a decomposition mapping $\eta \in (\delta_t)^{-1}(\psi)$ such that

$$\xi(x) = \begin{cases} \top & \text{if } x \notin \mathrm{var}(t) \\ \langle\hat{\tau}\rangle\eta(x) & \text{if } x \text{ occurs } \Gamma\text{-liquid in } t \\ \eta(x) & \text{otherwise.} \end{cases}$$

   (b) there are a variable y occurring Γ-liquid in t, a Γ-patient ruloid $\frac{y \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/y]}$ and a decomposition mapping $\eta \in (t[\mu/y])^{-1}(\psi)$ such that

$$\xi(x) = \begin{cases} \top & \text{if } x \notin \mathrm{var}(t) \\ \langle\hat{\tau}\rangle\eta(\mu) & \text{if } x = y \\ \eta(x) & \text{otherwise;} \end{cases}$$

   (c) there are a Γ-impatient P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and a decomposition mapping $\eta \in \Theta^{-1}(\psi)$ with

$$\xi(x) = \begin{cases} \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \wedge \bigwedge_{x \not\xrightarrow{\gamma} \in \mathbf{H}} \neg\langle\gamma\rangle\top \wedge \eta(x) & \text{if } x \in \mathrm{var}(t) \\ \top & \text{otherwise.} \end{cases}$$

6. $\xi \in t^{-1}(\langle\varepsilon\rangle\psi)$ iff one of the following three cases holds:
   (a) there is a decomposition mapping $\eta \in (\delta_t)^{-1}(\psi)$ such that

$$\xi(x) = \begin{cases} \top & \text{if } x \notin \mathrm{var}(t) \\ \langle\varepsilon\rangle\eta(x) & \text{if } x \text{ occurs } \Gamma\text{-liquid in } t \\ \eta(x) & \text{otherwise.} \end{cases}$$

(b) there are a variable y occurring Γ-liquid in t, a Γ-patient ruloid $\frac{y \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/y]}$ and a decomposition mapping $\eta \in (t[\mu/y])^{-1}(\psi^{\langle\varepsilon\rangle})$ such that

$$\xi(x) = \begin{cases} \top & \text{if } x \notin \mathrm{var}(t) \\ \langle\varepsilon\rangle\eta(\mu) & \text{if } x = y \\ \eta(x) & \text{otherwise.} \end{cases}$$

(c) there are a Γ-impatient P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and a decomposition mapping $\eta \in \Theta^{-1}(\psi^{\langle\varepsilon\rangle})$ with

$$\xi(x) = \begin{cases} \top & \text{if } x \notin \mathrm{var}(t) \\ \langle\varepsilon\rangle 1 \left(\bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \wedge \bigwedge_{x \xrightarrow{\gamma}\!\!\!\!\!/\;\, \in \mathbf{H}} \neg\langle\gamma\rangle\top \wedge \eta(x) \right) & \text{if } x \text{ occurs } \Gamma\text{-liquid in } t \\ \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \wedge \bigwedge_{x \xrightarrow{\gamma}\!\!\!\!\!/\;\, \in \mathbf{H}} \neg\langle\gamma\rangle\top \wedge \eta(x) & \text{otherwise.} \end{cases}$$

7. $\xi \in (\sigma(t))^{-1}(\varphi)$ for a non injective substitution $\sigma\colon \mathrm{var}(t) \to \mathbf{V}_s$ iff there is a $\xi' \in t^{-1}(\varphi)$ such that

$$\xi(x) = \begin{cases} \bigwedge_{y \in \sigma^{-1}(x)} \xi'(y) & \text{if } x \in \mathrm{var}(t) \\ \top & \text{otherwise.} \end{cases}$$

Then, for each distribution term $\Theta \in \mathbb{DT}(\Sigma)$ and distribution formula $\psi \in \mathbb{L}^{\mathrm{d}}$, $\Theta^{-1}(\psi) \in \mathcal{P}(\mathbf{V} \to \mathbb{L})$ is the set of *decomposition mappings* $\eta\colon \mathbf{V} \to \mathbb{L}$ such that for any univariate distribution term Θ we have:

8. $\eta \in \Theta^{-1}(\bigoplus_{i \in I} r_i\varphi_i)$ iff there are a Σ-distribution ruloid $\frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}}$ and a matching $\mathfrak{w} \in \mathfrak{W}(\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}, \bigoplus_{i \in I} r_i\varphi_i)$ such that for all $m \in M$ and $i \in I$ there is a decomposition mapping $\xi_{m,i}$ defined by $\begin{cases} \xi_{m,i} \in u_m^{-1}(\varphi_i) & \text{if } \mathfrak{w}(u_m, \varphi_i) > 0 \\ \xi_{m,i} \in u_m^{-1}(\top) & \text{otherwise} \end{cases}$ and such that:

   (a) for any distribution variable $\mu \in \mathbf{V}_d$:

$$\eta(\mu) = \begin{cases} \bigoplus_{\{\mu \xrightarrow{q_j} x_j\} \in \mathbf{H}} q_j \bigwedge_{i \in I, m \in M} \xi_{m,i}(x_j) & \text{if } \mu \in \mathrm{var}(\Theta) \\ 1\top & \text{otherwise.} \end{cases}$$

   (b) for any state variable $x \in \mathbf{V}_s$:

$$\eta(x) = \begin{cases} \bigwedge_{i \in I, m \in M} \xi_{m,i}(x) & \text{if } x \in \mathrm{var}(\Theta) \\ \top & \text{otherwise.} \end{cases}$$

9. $\eta \in \Theta^{-1}(\bigwedge_{j \in J} \psi_j)$ iff there exist decomposition mappings $\eta_j \in \Theta^{-1}(\psi_j)$ for all $j \in J$ such that

$$\eta(\zeta) = \bigwedge_{j \in J} \eta_j(\zeta), \text{ for all variables } \zeta \in \mathbf{V}.$$

10. $\eta \in \Theta^{-1}((\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle})$ iff one of the following two cases holds:

   (a) there is a decomposition mapping $\eta' \in \Theta^{-1}(\bigoplus_{i\in I} r_i\varphi_i)$ with $\eta(\zeta) = \eta'(\zeta)$ for all $\zeta \in \mathrm{var}(\Theta)$;

   (b) there are a Σ-distribution ruloid $\frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m\in M\}}$ and a matching $\mathfrak{w} \in \mathfrak{W}(\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}, \bigoplus_{i\in I} r_i\varphi_i)$ such that for all $m \in M$ there is a decomposition mapping $\xi_m \in u_m^{-1}(\langle\varepsilon\rangle\psi_m)$, where ψ_m is the formula

$$\psi_m = \bigoplus_{i\in I_m} \frac{\mathfrak{w}(u_m,\varphi_i)}{q_m}\varphi_i, \text{ with } I_m = \{i \in I \mid \mathfrak{w}(u_m,\varphi_i) > 0\},$$

   for which we have

   i. for any distribution variable $\mu \in \mathbf{V}_d$, $\eta(\mu) = \begin{cases} \bigoplus_{\{\mu \xrightarrow{q_j} x_j\}\in\mathbf{H}} q_j \bigwedge_{m\in M} \xi_m(x_j) & \text{if } \mu \in \mathrm{var}(\Theta) \\ 1\top & \text{otherwise.} \end{cases}$

   ii. for any state variable $x \in \mathbf{V}_s$, $\eta(x) = \begin{cases} \bigwedge_{m\in M} \xi_m(x) & \text{if } x \in \mathrm{var}(\Theta) \\ \top & \text{otherwise.} \end{cases}$

11. $\eta \in \Theta^{-1}((\bigwedge_{j\in J}\psi_j)^{\langle\varepsilon\rangle})$, for $\psi_j = \bigoplus_{i_j\in I_j} r_{i_j}\varphi_{i_j}$ for all $j \in J$, iff one of the following two cases holds:

   (a) there is a decomposition mapping $\eta' \in \Theta^{-1}(\bigwedge_{j\in J}\psi_j)$ with $\eta(\zeta) = \eta'(\zeta)$ for all $\zeta \in \mathrm{var}(\Theta)$;

   (b) there are a Σ-distribution ruloid $\frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m\in M\}}$ and, for each $m \in M$, a decomposition mapping $\xi_m \in u_m^{-1}(\langle\varepsilon\rangle \bigwedge_{j\in J} \psi_{m,j})$, with $\psi_{m,j}$ defined from $\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}$ and ψ_j as in previous item 10b, i.e., we consider a matching $\mathfrak{w}_j \in \mathfrak{W}(\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}, \psi_j)$ and define

$$\psi_{m,j} = \bigoplus_{i_j\in I_{m,j}} \frac{\mathfrak{w}_j(u_m,\varphi_{i_j})}{q_m}\varphi_i, \text{ with } I_{m,j} = \{i_j \in I_j \mid \mathfrak{w}(u_m,\varphi_{i_j}) > 0\},$$

   for which we have

   i. for any distribution variable $\mu \in \mathbf{V}_d$, $\eta(\mu) = \begin{cases} \bigoplus_{\{\mu \xrightarrow{q_h} x_h\}\in\mathbf{H}} q_h \bigwedge_{m\in M} \xi_m(x_h) & \text{if } \mu \in \mathrm{var}(\Theta) \\ 1\top & \text{otherwise.} \end{cases}$

   ii. for any state variable $x \in \mathbf{V}_s$, $\eta(x) = \begin{cases} \bigwedge_{m\in M} \xi_m(x) & \text{if } x \in \mathrm{var}(\Theta) \\ \top & \text{otherwise.} \end{cases}$

12. $\eta \in (\sigma(\Theta))^{-1}(\psi)$ for a non injective substitution $\sigma\colon \mathrm{var}(\Theta) \to \mathbf{V}$ iff there is a decomposition mapping $\eta' \in \Theta^{-1}(\psi)$ such that

$$\eta(\zeta) = \begin{cases} \bigwedge_{\zeta'\in\sigma^{-1}(\zeta)} \eta'(\zeta') & \text{if } \zeta \in \mathrm{var}(\Theta) \\ \top & \text{otherwise.} \end{cases}$$

We refer the reader to [37] for an explanation of the decomposition of the standard Hennessy-Milner operators and to [13] for an explanation of the decomposition of distribution formulae of the form $\bigoplus_{i\in I} r_i\varphi_i$. Here we detail our decomposition method for formulae of the form $\langle\hat{\tau}\rangle\psi$ (item 5 of Definition 23), $\langle\varepsilon\rangle\psi$ (item 6 of Definition 23) and $\psi^{\langle\varepsilon\rangle}$ (items 10 and 11 of Definition 23).

We start with $\langle\hat{\tau}\rangle\psi$. Given any term $t \in \mathbb{T}(\Sigma)$ and closed substitution σ, we aim to identify in $\xi \in t^{-1}(\langle\hat{\tau}\rangle\psi)$ which properties $\sigma(x)$ has to satisfy in order to guarantee $\sigma(t) \models \langle\hat{\tau}\rangle\psi$, for all $x \in \mathrm{var}(t)$. According to Definition 14 we can distinguish two cases for $\sigma(t) \models \langle\hat{\tau}\rangle\psi$: either $\delta_{\sigma(t)} \models \psi$, or $P \vdash \sigma(t) \xrightarrow{\tau} \pi$ for a distribution $\pi \in \Delta(\mathbf{T}(\Sigma))$ with $\pi \models \psi$. The former case, in which $\sigma(t)$ does not execute any τ step, motivates item 5a in Definition 23. In the latter case, by Theorem 2 the τ-transition by $\sigma(t)$ is inferred by a ruloid $\rho = \frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$. Item 5b applies if ρ is Γ-patient whereas item 5c applies if ρ is Γ-impatient. In detail:

<u>Definition 23.5a</u>: To have $\delta_{\sigma(t)} \models \psi$ we need that $\sigma(\zeta) \models \eta(\zeta)$ for all variables $\zeta \in \mathrm{var}(\delta_t)$, for a suitable decomposition mapping $\eta \in (\delta_t)^{-1}(\psi)$. Since $\mathrm{var}(\delta_t) = \mathrm{var}(t)$, this is equivalent to have $\sigma(x) \models \eta(x)$ for all state variables $x \in \mathrm{var}(t)$. However, for x occurring Γ-liquid in t, we define the decomposed formula $\xi(x)$ as the less demanding $\xi(x) = \langle\hat{\tau}\rangle\eta(x)$. This is to guarantee that the decomposed formula does not discriminate processes by their ability of performing a τ step.

<u>Definition 23.5b</u>: Since ρ is a Γ-patient ruloid, there is a variable y that occurs Γ-liquid in t, $\mathbf{H}$ is of the form $\mathbf{H} = y \xrightarrow{\tau} \mu$ and Θ is of the form $\Theta = t[\mu/y]$, namely $\rho = \frac{y \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/y]}$. By tuning Theorem 2 on this particular case, the closed substitution σ' satisfying $\sigma'(t) = \sigma(t)$ must be such that: (i) $P \vdash \sigma'(y) \xrightarrow{\tau} \sigma'(\mu)$ and (ii) $\sigma'(t[\mu/y]) \models \psi$. Hence, given $\eta \in (t[\mu/y])^{-1}(\psi)$, the first condition is satisfied if $\sigma'(y) \models \langle\tau\rangle\eta(\mu)$ and the second if $\sigma'(x) \models \eta(x)$ for all $x \in \mathrm{var}(t)$ with $x \neq y$. The reason why we have $\xi(y) = \langle\hat{\tau}\rangle\eta(\mu)$ instead of the more demanding property $\xi(y) = \langle\tau\rangle\eta(\mu)$ is that this allows us to obtain formulae in $\mathbb{L}_{\mathrm{b}}$ whenever we decompose formulae in $\mathbb{L}_{\mathrm{b}}$ (Theorem 6 below). In particular, since $\langle\hat{\tau}\rangle\eta(\mu)$ does not impose the execution of the silent step, we are guaranteed that the decomposed formula does not discriminate processes by their ability of performing a τ step.

<u>Definition 23.5c</u> By tuning Theorem 2 to this particular case we get a substitution σ' with $\sigma'(t) = \sigma(t)$, that must satisfy: (i) $P \vdash \sigma'(\mathbf{H})$, and (ii) $\sigma'(\Theta) \models \psi$. Hence, given $\eta \in \Theta^{-1}(\psi)$, the validity of condition (i) follows if, for each x occurring in t, $\sigma'(x)$ is such that $\sigma'(x) \models \langle b\rangle\eta(\mu)$ for each $x \xrightarrow{b} \mu \in \mathbf{H}$, and $\sigma'(x) \models \neg\langle c\rangle\top$ for each $x \stackrel{c}{\not\to} \in \mathbf{H}$. The validity of condition (ii) then follows if $\sigma'(x) \models \eta(x)$ for the variables x occurring in Θ. Notice that, since we are considering a Γ-impatient ruloid, the decomposed formula is obtained from the premises of the ruloid without relaxing any constraint on the execution of the silent steps. This is to guarantee the preservation of the rootedness condition, so that to obtain formulae in $\mathbb{L}_{\mathrm{rb}}$ whenever we decompose formulae in $\mathbb{L}_{\mathrm{rb}}$ (Theorem 7 below).

Now consider $\langle\varepsilon\rangle\psi$. Accordingly to Definition 14 we can distinguish two cases for $\sigma(t) \models \langle\varepsilon\rangle\psi$: either $\delta_{\sigma(t)} \models \psi$, or $P \vdash \sigma(t) \xrightarrow{\hat{\varepsilon}} \pi$ for a distribution $\pi \in \Delta(\mathbf{T}(\Sigma))$ with $\pi \models \psi$. The former case, in which $\sigma(t)$ does not execute any τ step, motivates item 6a in Definition 23. In the latter case, $\sigma(t)$ executes an arbitrary number of lifted transitions with label $\hat{\tau}$ before reaching a distribution that satisfies ψ. However, since we are considering *lifted* transitions, such a distribution may be obtained as a convex combination of the distributions that have been reached through the execution of each $\hat{\tau}$-step. Hence this case motivates items 6b and 6c, and the introduction of formulae of the form $\psi^{\langle\varepsilon\rangle}$. In particular, item 6b applies when the *first* τ-transition by $\sigma(t)$ is inferred by a Γ-patient P-ruloid and item 6c applies when the *first* τ-transition by $\sigma(t)$ is inferred by a Γ-impatient P-ruloid. Formulae of the form $\psi^{\langle\varepsilon\rangle}$, and their decomposition that we will detail below, allow us to proceed in the decomposition of the sequence of $\hat{\tau}$ steps by also keeping track of the probabilistic behavior of the process. In detail:

<u>Definition 23.6a</u>: this case is analogous to that of Definition 23.5a.

<u>Definition 23.6b</u>: this case is analogous to that of Definition 23.5b. The main difference is in that we must ensure that after having derived the first τ-step, we can continue the decomposition of the remaining silent steps in the $\langle\varepsilon\rangle$ sequence. Therefore, we exploit the formula $\psi^{\langle\varepsilon\rangle}$ and we look for the mapping $\eta \in (t[\mu/y])^{-1}(\psi^{\langle\varepsilon\rangle})$, such that $\sigma'(y) \models \langle\varepsilon\rangle\eta(\mu)$ and $\sigma'(x) \models \eta(x)$ for all $x \neq y$ with $x \in \mathrm{var}(t)$.

Definition 23.6c: this case is analogous to that of Definition 23.5c to which the same observations made in the case of Definition 23.6b apply. Hence, we consider a mapping $\eta \in \Theta^{-1}(\psi^{\langle\varepsilon\rangle})$ to build the decomposition. Moreover, we notice that if x occurs Γ-liquid in t then we have to admit that certain silent moves of $\sigma'(x)$ have to be executed in order to enable the steps required in $\mathbf{H}$. Hence, we define $\xi(x)$ as the formula $\langle\varepsilon\rangle\bar{\psi}$ where $\bar{\psi}$ is the distribution formula that assigns probability 1 to the state formula obtained for Γ-frozen variables x. Notice that the weight 1 guarantees that all the processes in the support of the distribution reached via the sequence of silent steps must satisfy the state formula obtained from the premises of $\mathbf{H}$. This is fundamental to guarantee that such premises are enabled and that no modification in the probabilistic behavior of the subprocess occurs in the initial sequence of silent steps.

We proceed now to discuss the decomposition of distribution formulae of the form $(\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$. We remark, once again, that such formulae occur only in the decomposition of state formulae of the form $\langle\varepsilon\rangle \bigoplus_{i\in I} r_i\varphi_i$ and their decomposition is tailored to capture the *probabilistic lookahead* introduced by such state formulae. This is the reason why here we cannot apply the same decomposition method proposed for distribution formulae of the form $\bigoplus_{i\in I} r_i\varphi_i$.

Given any distribution term $\Theta \in \mathbb{DT}(\Sigma)$ and a closed substitution σ, our purpose is to identify in $\eta \in \Theta^{-1}((\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle})$ which properties each $\sigma(\zeta)$, with $\zeta \in \mathrm{var}(\Theta)$, has to satisfy in order to guarantee that $\sigma(\Theta) \models (\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$. Informally, we should find the requirements guaranteeing that either $\sigma(\Theta)$ satisfies $\bigoplus_{i\in I} r_i\varphi_i$, or it will satisfy it after the execution of an arbitrary number of $\hat{\tau}$-steps performed by *all* the processes in its support. The former case motivates item 10a of Definition 23, whereas the latter case leads to item 10b of Definition 23. The interesting case is the second one. Here, a decomposition mapping $\eta \in \Theta^{-1}((\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle})$ is built on the derivation of a lifted transition $\sigma(\Theta) \xrightarrow{\hat{\tau}} \pi$, where π satisfies, in turn, $(\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$. Decomposing $(\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$ in terms of the decomposition of $(\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$ itself is correct since we assume that the PTS is divergence-free. To obtain that $P \vdash \sigma(\Theta) \xrightarrow{\hat{\tau}} \pi$, for the desired π, we must have suitable processes t_m, distributions π_m and probability weights q_m such that:

1. $D_\Sigma \vdash \{\sigma(\Theta) \xrightarrow{q_m} t_m \mid \sum_{m\in M} q_m = 1 \text{ and } t_m \in \mathbf{T}(\Sigma)\}$;
2. $P \vdash t_m \xrightarrow{\hat{\tau}} \pi_m$ for all $m \in M$;
3. $\pi = \sum_{m\in M} q_m\pi_m \models (\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$.

Firstly, we notice that by Theorem 3 (distribution ruloids are sound and specifically witnessing) and the construction of lifted transitions, items 1 and 2 are equivalent to say that there are a Σ-distribution ruloid $\rho^{\mathsf{D}} = \frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m\in M\}}$ and a closed substitution σ' with

4. $\sigma'(\Theta) = \sigma(\Theta)$;
5. $D_\Sigma \vdash \sigma'(\mathbf{H})$, and
6. $P \vdash \sigma'(u_m) \xrightarrow{\hat{\tau}} \pi_m$, with $\sigma'(u_m) = t_m$ for all $m \in M$.

Then we have to understand which formula should be satisfied by each π_m in order to obtain the validity of item 3. We rely on some matching $\mathfrak{w} \in \mathfrak{W}(\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}, \bigoplus_{i\in I} r_i\varphi_i)$ to establish which processes of those in the support of the distribution π' reached by $\sigma'(\Theta)$ via the sequence of $\hat{\tau}$-steps, knowing that for each $m \in M$ the first step takes $\sigma'(u_m)$ to π_m, will have to satisfy the formula φ_i. Informally, for each $i \in I$, $\frac{\mathfrak{w}(u_m,\varphi)}{q_m}$ expresses the probability that $\sigma'(u_m)$ has to reach, via the sequence of silent steps, (a subset of) such processes. The formula $\langle\epsilon\rangle\psi_m$, with ψ_m assigning weight $\frac{\mathfrak{w}(u_m,\varphi_i)}{q_m}$ to φ_i, is therefore the formula that must be satisfied by $\sigma'(u_m)$. Notice that the division by q_m, namely the probability weight assigned to $\sigma'(u_m)$ by $\sigma'(\Theta)$, derives from the conditioning on the first $\hat{\tau}$-step performed by $\sigma'(\Theta)$ and it guarantees both that ψ_m is a well-defined distribution formula, and that π' satisfies $\sum_{i\in I} r_i\varphi_i$. Then, for each $m \in M$, we consider a decomposition mapping $\xi_m \in u_m^{-1}(\langle\varepsilon\rangle\psi_m)$ that will allow us to assign to each variable x occurring in u_m a proper formula $\xi_m(x)$ such that whenever $\sigma'(x) \models \xi_m(x)$ for all $x \in \mathrm{var}(u_m)$, then $\sigma'(u_m) \models \langle\varepsilon\rangle\psi_m$, namely $P \vdash \sigma'(u_m) \xrightarrow{\hat{\tau}} \pi_m$ and $\pi_m \models \psi_m^{\langle\varepsilon\rangle}$.

Once the distribution formulae ψ_m are built, consider any $\mu \in \mathrm{var}(\Theta)$. Then there is a distribution over terms $\{\mu \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}$ such that $D_\Sigma \vdash \{\sigma'(\mu) \xrightarrow{q_j} \sigma'(x_j) \mid j \in J\}$. Since the weights q_m of $\sigma'(\Theta)$ are univocally determined by the distributions over terms in $\mathbf{H}$, and thus also by those of $\sigma'(\mu)$, we define the decomposed formula $\eta(\mu)$ as the distribution formula that has as weights exactly the q_j of the distribution over terms for μ in $\mathbf{H}$. Then, to ensure the validity of item 6, we assign to each variable x_j the conjunction of the formulae obtained from the decomposition mappings ξ_m when applied to x_j. Similarly, for each variable $x \in \mathrm{var}(\Theta)$, we define $\eta(x)$ as the state formula obtained from the conjunction over $m \in M$ of the decomposed formulae $\xi_m(x)$. Intuitively, the conjunction over $m \in M$ is needed since the same variable y may occur in more than one term u_m, and moreover, from the construction of the mappings ξ_m described above, we can be sure that $\sigma'(y) \models \bigwedge_{m\in M} \xi_m(y)$, for all variables y occurring as right-hand sides of distributions over terms in $\mathbf{H}$ or in $\mathrm{var}(\Theta)$, gives that $P \vdash \sigma'(u_m) \xrightarrow{\hat{\tau}} \pi_m$ and $\pi_m \models \psi_m^{\langle\varepsilon\rangle}$ for all $m \in M$. By the construction of the formulae ψ_m, this implies the validity of item 3.

The decomposition of formulae $(\bigwedge_{j\in J} \psi_j)^{\langle\varepsilon\rangle}$, which allow us to obtain a proper decomposition for formulae of the form $\langle\varepsilon\rangle \bigwedge_{j\in J} \psi_j$, follows the same ideas of that of $(\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$. Informally, we decompose all the formulae ψ_j at the same time with respect to a single distribution term. This is to avoid to consider distinct probability distributions possibly reached via the execution of distinct silent steps by the same process. Hence, from each formula ψ_j in the conjunction and the distribution term Θ, we construct the formulae $\psi_{m,j}$ as described above, and we consider the decomposition mappings $\xi_m \in u_m^{-1}(\langle\varepsilon\rangle \bigwedge_{j\in J} \psi_{m,j})$ will allow us to assign to each variable x occurring in u_m a proper formula $\xi_m(x)$ such that whenever $\sigma'(x) \models \xi_m(x)$ for all $x \in \mathrm{var}(u_m)$, then $\sigma'(u_m) \models \langle\varepsilon\rangle \bigwedge_{j\in J} \psi_{m,j}$, namely $P \vdash \sigma'(u_m) \xrightarrow{\hat{\tau}} \pi_m$ and $\pi_m \models (\bigwedge_{j\in J} \psi_{m,j})^{\langle\varepsilon\rangle}$. As in the previous case, the construction of the formulae $\psi_{m,j}$ guarantees that $\sigma(\Theta) \models (\bigwedge_{j\in J} \psi_j)^{\langle\varepsilon\rangle}$, thus giving us a well-defined decomposition.

***Example* 9.** We show now how we can decompose the state formula $\varphi = \langle\varepsilon\rangle\psi$, with ψ the distribution formula $\psi = 1/4\langle a\rangle\top \oplus 1/4\langle\tau\rangle\top \oplus 1/2\neg\langle b\rangle\top$, with respect to the term $t = x +_{2/3} (y \parallel_{\{a\}} z)$. For sake of simplicity, we propose the construction of a single decomposition mapping $\xi \in t^{-1}(\varphi)$. In particular, we consider $\Gamma = \aleph \cap \Lambda$ and we will choose arbitrarily the ruloids that have to be used in the construction, in order to enlighten the peculiarities of our decomposition method. Moreover, in the formulae we will omit all occurrences of conjuncts of the form $\wedge\top$.

We recall that in Example 2 we discussed that both arguments of probabilistic alternative composition should be marked as $\aleph$-liquid and Λ-frozen. Hence, any derivation of a silent move for the term t will be obtained via a $(\aleph\cap\Lambda)$-impatient ruloid. For instance, we consider the P-ruloid ρ_4 in Table 2, that for $\beta = \tau$ and $p = 2/3$ constitutes a $\aleph \cap \Lambda$-impatient ruloid for t, and we derive a decomposition mapping by following Definition 23.6c. Let $\Theta = 2/3\mu + 1/3(\nu \parallel_{\{a\}} \delta_z)$ denote the target of such an instance of ρ_4. Then, for $\eta \in \Theta^{-1}(\psi^{\langle\varepsilon\rangle})$, from the premises of ρ_4 we get

$$\xi(x) = \langle\tau\rangle\eta(\mu) \wedge \eta(x) \qquad \xi(y) = \langle\tau\rangle\eta(\nu) \wedge \eta(y) \qquad \xi(z) = \eta(z).$$

We build the decomposition mapping $\eta \in \Theta^{-1}(\psi^{\langle\varepsilon\rangle})$ by applying Definition 23.10b with the Σ-distribution ruloid ρ^{D} in Example 8 and the matching $\mathfrak{w} \in \mathfrak{W}(\mathrm{conc}(\rho^{\mathsf{D}}), \psi)$ for the distribution over terms $\mathrm{conc}(\rho^{\mathsf{D}})$ and the distribution formula ψ defined by

$$\mathfrak{w}(x_1, \langle a\rangle\top) = 1/12 \qquad \mathfrak{w}(x_1, \neg\langle b\rangle\top) = 1/12 \qquad \mathfrak{w}(x_2, \langle\tau\rangle\top) = 1/8 \qquad \mathfrak{w}(x_2, \neg\langle b\rangle\top) = 3/8$$

$$\mathfrak{w}(y_1 \parallel_{\{a\}} z, \langle a\rangle\top) = 1/6 \qquad \mathfrak{w}(y_2 \parallel_{\{a\}} z, \langle\tau\rangle\top) = 1/8 \qquad \mathfrak{w}(y_2 \parallel_{\{a\}} z, \neg\langle b\rangle\top) = 1/24$$

and giving value 0 in all other cases. Following Definition 23.10b, we can now construct, for each term t' in the support of Θ, the proper formula $\psi_{t'}$ allowing us to proceed in the decomposition of $\psi^{\langle\varepsilon\rangle}$:

$$\psi_{x_1} = \frac{1/12}{1/6}\langle a\rangle\top \oplus \frac{1/12}{1/6}\neg\langle b\rangle\top = \frac{1}{2}\langle a\rangle\top \oplus \frac{1}{2}\neg\langle b\rangle\top$$
$$\psi_{x_2} = \frac{1/8}{1/2}\langle \tau\rangle\top \oplus \frac{3/8}{1/2}\neg\langle b\rangle\top = \frac{1}{4}\langle \tau\rangle\top \oplus \frac{3}{4}\neg\langle b\rangle\top$$
$$\psi_{y_1\|_{\{a\}}z} = \frac{1/6}{1/6}\langle a\rangle\top = 1\langle a\rangle\top$$
$$\psi_{y_2\|_{\{a\}}z} = \frac{1/8}{1/6}\langle \tau\rangle\top \oplus \frac{1/24}{1/6}\neg\langle b\rangle\top = \frac{3}{4}\langle \tau\rangle\top \oplus \frac{1}{4}\neg\langle b\rangle\top.$$

Accordingly to Definition 23.10b, to obtain the decomposition mapping η we need to consider, for each term t', a decomposition mapping $\xi_{t'} \in (t')^{-1}(\langle\varepsilon\rangle\psi_{t'})$ so that

$$\eta(\mu) = \frac{1}{4}\xi_{x_1}(x_1) \oplus \frac{3}{4}\xi_{x_2}(x_2) \qquad \eta(\nu) = \frac{1}{2}\xi_{y_1\|_{\{a\}}z}(y_1) \oplus \frac{1}{2}\xi_{y_2\|_{\{a\}}z}(y_2) \qquad \eta(z) = \xi_{y_1\|_{\{a\}}z}(z) \wedge \xi_{y_2\|_{\{a\}}z}(z).$$

To simplify reasoning, we assume that both ξ_{x_1} and ξ_{x_2} are derived by applying Definition 23.6a, namely that, for $i = 1, 2$, there is a mapping $\eta_{x_i} \in (\delta_{x_i})^{-1}(\psi_{x_i})$ such that $\xi_{x_i}(x_i) = \eta_{x_i}(x_i)$. Hence, by Definition 23.8b and thus by combining the Σ-distribution ruloids having a Dirac delta as source and the P-ruloids having a variable as source, we can directly infer that

$$\xi_{x_1}(x_1) = \langle a\rangle\top \wedge \neg\langle b\rangle\top \qquad\qquad \xi_{x_2}(x_2) = \langle \tau\rangle\top \wedge \neg\langle b\rangle\top$$

where, in particular, we get the conjuncts $\neg\langle b\rangle\top$ by applying Definition 23.2 to the P-ruloids $\frac{x_i \xrightarrow{b} \mu_i}{x_i \xrightarrow{b} \mu_i}$.

Let us focus now on the construction of the mapping $\xi_{y_1\|_{\{a\}}z} \in (y_1 \|_{\{a\}} z)^{-1}(\langle\varepsilon\rangle 1\langle a\rangle\top)$. We remark that since the distribution formula assigns weight 1 to the state formula $\langle a\rangle\top$ then, no matter how may $\hat{\tau}$-steps can be derived for $y_1 \|_{\{a\}} z$ and how clunky can the probability distributions reached via those steps be, *all* the processes in the support of the distribution π s.t. $y_1 \|_{\{a\}} z \xrightarrow{\hat{\varepsilon}} \pi$ will have to satisfy the formula $\langle a\rangle\top$. Hence, for sake of simplicity (ant to keep this example with an acceptable length) we construct $\xi_{y_1\|_{\{a\}}z}$ by applying Definition 23.6a. This, considering that Definition 23.8b, is equivalent to say that we build the decomposition mapping $\xi'_{y_1\|_{\{a\}}z} \in (y_1 \|_{\{a\}} z)^{-1}(\langle a\rangle\top)$ and $\xi_{y_1\|_{\{a\}}z}(w) = \xi'_{y_1\|_{\{a\}}z}(w)$ for all $w \in \mathbf{V}_s$. Since $\|_{\{a\}}$ forces the synchronization on action a, to decompose $\langle a\rangle\top$ we use the P-ruloid corresponding to the PGSOS-rule r_6 in Table 1 from which we infer

$$\xi_{y_1\|_{\{a\}}z}(y_1) = \xi_{y_1\|_{\{a\}}z}(z) = \langle a\rangle\top.$$

Finally, we consider the mapping $\xi_{y_2\|_{\{a\}}z} \in (y_2 \|_{\{a\}} z)^{-1}(\langle\varepsilon\rangle\psi_{y_2\|_{\{a\}}z})$. As shown in Example 2, both arguments of $\|_{\{a\}}$ are marked as $\aleph \cap \Lambda$-liquid. Thus, for $\alpha = \tau$, rules r_4 and r_5 in Table 1 are $\aleph \cap \Lambda$-patient ruloids for the considered term. So, we construct $\xi_{y_2\|_{\{a\}}z}$ by applying Definition 23.6b with the proper instance of rule r_4 as $\aleph \cap \Lambda$-patient ruloid. Let $\frac{y_2 \xrightarrow{\tau} \nu'}{y_2\|_{\{a\}}z \xrightarrow{\tau} \nu'\|_{\{a\}}\delta_z}$ be such an instance. Hence, we need now to look for a decomposition mapping $\eta' \in (\nu' \|_{\{a\}} \delta_z)^{-1}(\psi^{\langle\varepsilon\rangle}_{y_2\|_{\{a\}}z})$ s.t.

$$\xi_{y_2\|_{\{a\}}z}(y_2) = \langle\varepsilon\rangle\eta'(\nu') \qquad\qquad \xi_{y_2\|_{\{a\}}z}(z) = \eta'(z).$$

Assume that η' is built by applying Definition 23.10a and thus we need to decompose the distribution formula $\psi_{y_2\|_{\{a\}}z}$ wrt. the term $\nu' \|_{\{a\}} \delta_z$. Following Definition 23.8, we consider the Σ-distribution ruloid $\rho^{\mathsf{D}}_1 = \frac{\{\nu' \xrightarrow{1/3} y_3,\quad \nu' \xrightarrow{2/3} y_4\} \qquad \{\delta_z \xrightarrow{1} z\}}{\{\nu'\|_{\{a\}}\delta_z \xrightarrow{1/3} y_3\|_{\{a\}}z,\quad \nu'\|_{\{a\}}\delta_z \xrightarrow{2/3} y_4\|_{\{a\}}z\}}$ and the matching $\mathfrak{w}_1 \in \mathfrak{W}(\mathrm{conc}(\rho^{\mathsf{D}}_1), \psi_{y_2\|_{\{a\}}z})$ defined by

$$\mathfrak{w}_1(y_3 \|_{\{a\}} z, \langle\tau\rangle\top) = 1/12 \qquad \mathfrak{w}(y_3 \|_{\{a\}} z, \neg\langle b\rangle\top) = 1/4 \qquad \mathfrak{w}(y_4 \|_{\{a\}} z, \langle\tau\rangle\top) = 2/3$$

and assumes value 0 in all other cases. To conclude we need to construct the decomposition mappings $\xi_1 \in (y_3 \parallel_{\{a\}} z)^{-1}(\langle\tau\rangle\top)$, $\xi_2 \in (y_3 \parallel_{\{a\}} z)^{-1}(\neg\langle b\rangle\top)$ and $\xi_3 \in (y_4 \parallel_{\{a\}} z)^{-1}(\langle\tau\rangle\top)$ s.t.

$$\eta'(\nu') = \frac{1}{3}(\xi_i(y_3) \wedge \xi_2(y_3)) \oplus \xi_3(y_4) \qquad \eta'(z) = \xi_1(z) \wedge \xi_2(z) \wedge \xi_3(z).$$

We can assume that ξ_1 and ξ_3 are obtained from Definition 23.4 applied to proper instances of, respectively, rule r_4 and rule r_5 in Table 1, so that

$$\xi_1(y_3) = \langle\tau\rangle\top \quad \xi_1(z) = \top \qquad \xi_3(y_4) = \top \quad \xi_3(z) = \langle\tau\rangle\top.$$

Conversely, to obtain the mapping ξ_2, we need to exploit Definition 23.2. Thus, first of all we need to construct the set of decomposition mappings $(y_3 \parallel_{\{a\}} z)^{-1}(\langle b\rangle\top)$. Since there is no synchronization on action b, we get $(y_3 \parallel_{\{a\}} z)^{-1}(\langle b\rangle\top) = \{\chi_1, \chi_2\}$ with $\chi_1(y_3) = \langle b\rangle\top, \chi_1(z) = \top$ and $\chi_2(y_3) = \top, \chi_2(z) = \langle b\rangle\top$. Then we consider the function $\mathfrak{f}\colon (y_3 \parallel_{\{a\}} z)^{-1}(\langle b\rangle\top) \to \{y_3, z\}$ defined by $\mathfrak{f}(\chi_1) = y_3$ and $\mathfrak{f}(\chi_2) = z$. By applying Definition 23.2 with such function $\mathfrak{f}$ we infer

$$\xi_2(y_3) = \neg\chi_1(y_3) = \neg\langle b\rangle\top \qquad \xi_2(z) = \neg\chi_2(z) = \neg\langle b\rangle\top.$$

Finally, if we trace back the construction of all the decomposition mappings, we obtain

$$\begin{aligned}
\xi(x) &= \langle\tau\rangle \left[\frac{1}{4}(\langle a\rangle\top \wedge \neg\langle b\rangle\top) \oplus \frac{3}{4}(\langle\tau\rangle\top \wedge \neg\langle b\rangle\top)\right] \\
\xi(y) &= \langle\tau\rangle \left[\frac{1}{2}\langle a\rangle\top \oplus \frac{1}{2}\langle\varepsilon\rangle\Big(\frac{1}{3}(\langle\tau\rangle\top \wedge \neg\langle b\rangle\top) \oplus \frac{2}{3}\top\Big)\right] \\
\xi(z) &= \langle a\rangle\top \wedge \langle\tau\rangle\top \wedge \neg\langle b\rangle\top.
\end{aligned}$$

◄

6.3. The decomposition theorem

We can show now that our decomposition method is correct. Anyhow, the proof of the *decomposition theorem* (Theorem 4) will make use of some auxiliary results that we present here.

Firstly, we have that the decomposition mappings correctly assign state formulae to state variables and distribution formulae to distribution variables.

Lemma 3. *Assume the terms $t \in \mathbb{T}(\Sigma)$ and $\Theta \in \mathbb{DT}(\Sigma)$ and the formulae $\varphi \in \mathbb{L}^{\mathrm{s}}$ and $\psi \in \mathbb{L}^{\mathrm{d}}$.*

1. *For all state variables $x \in \mathbf{V}_s$ we have that $\xi(x) \in \mathbb{L}^{\mathrm{s}}$ for each $\xi \in t^{-1}(\varphi)$.*
2. *For all distribution variables $\mu \in \mathbf{V}_d$ we have that $\eta(\mu) \in \mathbb{L}^{\mathrm{d}}$ for each $\eta \in \Theta^{-1}(\psi)$.*
3. *For all state variables $x \in \mathbf{V}_s$ we have that $\eta(x) \in \mathbb{L}^{\mathrm{s}}$ for each $\eta \in \Theta^{-1}(\psi)$.*

Then, we notice that the decomposition mappings are not sensible to α-conversion.

Lemma 4. *Consider terms $t \in \mathbb{T}(\Sigma)$ and $\Theta \in \mathbb{DT}(\Sigma)$ and formulae $\varphi \in \mathbb{L}^{\mathrm{s}}$ and $\psi \in \mathbb{L}^{\mathrm{d}}$. Moreover, let $\sigma\colon \mathbf{V} \to \mathbf{V}$ be any bijective renaming of variables.*

1. *Let $\xi \in t^{-1}(\varphi)$. Then $\xi(x) \equiv \top$ for all $x \notin \mathrm{var}(t)$.*
2. *Let $\xi \in \sigma(t)^{-1}(\varphi)$. Then there is a $\xi' \in t^{-1}(\varphi)$ with $\xi'(x) \equiv \xi(\sigma(x))$ for all $x \in \mathbf{V}_s$.*
3. *Let $\eta \in \sigma(\Theta)^{-1}(\psi)$. Then there is an $\eta' \in \Theta^{-1}(\psi)$ with $\eta'(\zeta) \equiv \eta(\sigma(\zeta))$ for all $\zeta \in \mathbf{V}$.*

Next, we show that the formulae ψ_m and $\psi_{m,j}$ used, respectively, in Definition 23.10 and Definition 23.11 are well-defined with respect to the satisfaction of distribution formulae.

Proposition 2. *Let $\Theta \in \mathbb{DT}(\Sigma)$ and σ be a closed substitution with $\sigma(\Theta) = \sum_{m \in M} q_m \delta_{t_m}$ for $t_m \in \mathbf{T}(\Sigma)$ pairwise distinct. Then*

1. *$\sigma(\Theta) \models (\bigoplus_{i \in I} r_i \varphi_i)^{\langle \varepsilon \rangle}$ iff there is a matching $\tilde{\mathfrak{w}} \in \mathfrak{W}(\sigma(\Theta), \bigoplus_{i \in I} r_i \varphi_i)$ such that for each $m \in M$ it holds that $t_m \models \langle \varepsilon \rangle \psi_m$ with $\psi_m = \bigoplus_{i \in I_m} \frac{\tilde{\mathfrak{w}}(t_m, \varphi_i)}{q_m} \varphi_i$ and $I_m = \{i \in I \mid \tilde{\mathfrak{w}}(t_m, \varphi_i) > 0\}$.*

2. *$\sigma(\Theta) \models (\bigwedge_{j \in J} \psi_j)^{\langle \varepsilon \rangle}$ iff for each $m \in M$ it holds that $t_m \models \langle \varepsilon \rangle \bigwedge_{j \in J} \psi_{m,j}$ with $\psi_{m,j}$ defined from Θ and ψ_j as in Definition 23.10.*

Proof. We expand only the prof of the first item. The proof for the second item follows by similar arguments. We prove the two implications separately.

($\Rightarrow$) Assume that $\sigma(\Theta) \models (\bigoplus_{i \in I} r_i \varphi_i)^{\langle \varepsilon \rangle}$. Then, we can distinguish two cases.

1. $\sigma(\Theta) \models \bigoplus_{i \in I} r_i \varphi_i$. By Definition 14, this implies that there is a matching $\mathfrak{w} \in \mathfrak{W}(\sigma(\Theta), \bigoplus_{i \in I} r_i \varphi_i)$ such that whenever $\mathfrak{w}(t_m, \varphi_i) > 0$ then $t_m \models \varphi_i$. It is then immediate to verify that the thesis follows for ψ_m constructed using the matching $\mathfrak{w}$ as $\tilde{\mathfrak{w}}$.

2. $P \vdash \sigma(\Theta) \xrightarrow{\hat{\varepsilon}} \pi$ and $\pi \models \bigoplus_{i \in I} r_i \varphi_i$, which implies the existence of a matching $\mathfrak{w} \in \mathfrak{W}(\pi, \bigoplus_{i \in I} r_i \varphi_i)$ such that whenever $\mathfrak{w}(t, \varphi_i) > 0$ then $t \models \varphi_i$. By Definition 4, $\sigma(\Theta) \xrightarrow{\hat{\varepsilon}} \pi$ implies that, for each $m \in M$, $P \vdash t_m \xrightarrow{\hat{\varepsilon}} \pi_m$ and $\pi = \sum_{m \in M} q_m \pi_m$. To simplify the reasoning, we assume that the supports of the distributions π_m are all distinct. If this is not the case, then we can reason on the support of π as the multi-set obtained by the union of the supports of the π_m. Define

$$\tilde{\mathfrak{w}}(t_m, \varphi_i) = \sum_{t \in \mathsf{supp}(\pi_m)} \mathfrak{w}(t, \varphi_i).$$

The proof that $\tilde{\mathfrak{w}}$ is a well-defined matching for $\sigma(\Theta)$ and $\bigoplus_{i \in I} r_i \varphi_i$ is straightforward. Hence, to conclude the proof, we need to show that for each $m \in M$ we have $\pi_m \models \psi_m = \bigoplus_{i \in I_m} \frac{\tilde{\mathfrak{w}}(t_m, \varphi_i)}{q_m} \varphi_i$. Define, for each $m \in M$

$$\mathfrak{w}_m(t, \varphi_i) = \begin{cases} \dfrac{\mathfrak{w}(t, \varphi_i)}{q_m} & \text{if } t \in \mathsf{supp}(\pi_m) \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\sum_{t \in \mathbf{T}(\Sigma)} \mathfrak{w}_m(t, \varphi_i) = \sum_{t \in \mathsf{supp}(\pi_m)} \mathfrak{w}_m(t, \varphi_i) = \frac{\sum_{t \in \mathsf{supp}(\pi_m)} \mathfrak{w}(t, \varphi_i)}{q_m} = \frac{\tilde{\mathfrak{w}}(t_m, \varphi_i)}{q_m}$$

$$\sum_{i \in I_m} \mathfrak{w}_m(t, \varphi_i) = \sum_{i \in I} \mathfrak{w}_m(t, \varphi_i) = \frac{\sum_{i \in I} \mathfrak{w}(t, \varphi_i)}{q_m} = \frac{q_m \cdot \pi_m(t)}{q_m} = \pi_m(t)$$

and thus $\mathfrak{w}_m$ is a well-defined matching for π_m and ψ_m. Moreover, notice that $\mathfrak{w}_m(t, \varphi_i) > 0$ iff $\mathfrak{w}(t, \varphi_i) > 0$ and therefore we gather $t \models \varphi_i$ whenever $\mathfrak{w}_m(t, \varphi_i) > 0$. We can therefore conclude that the matching $\tilde{\mathfrak{w}}$ is such that $t_m \models \langle \varepsilon \rangle \psi_m$ for each $m \in M$.

($\Leftarrow$) Assume now that there is a matching $\tilde{\mathfrak{w}} \in \mathfrak{W}(\sigma(\Theta), \bigoplus_{i \in I} r_i \varphi_i)$ such that $t_m \models \langle \varepsilon \rangle \psi_m$ for each $m \in M$. By Definition 14, for each $m \in M$ we have that $P \vdash t_m \xrightarrow{\hat{\varepsilon}} \pi_m$ with $\pi_m \models \psi_m$. Hence, by Definition 4 we can infer that $P \vdash \sigma(\Theta) \xrightarrow{\hat{\varepsilon}} \pi$ for $\pi = \sum_{m \in M} q_m \pi_m$. To conclude the proof, we need to show that $\pi \models \bigoplus_{i \in I} r_i \varphi_i$. For each $m \in M$, from $\pi_m \models \psi_m$ we get that there is a matching $\mathfrak{w}_m \in \mathfrak{W}(\pi_m, \psi_m)$ such that $t \models \varphi_i$ whenever $\mathfrak{w}_m(t, \varphi_i) > 0$. Define

$$\bar{\mathfrak{w}}(t, \varphi_i) = \sum_{m \in M} q_m \cdot \mathfrak{w}_m(t, \varphi_i).$$

Then we have

$$\sum_{t\in\mathbf{T}(\Sigma)} \bar{\mathfrak{w}}(t,\varphi_i) = \sum_{t\in\mathbf{T}(\Sigma)} \left(\sum_{m\in M} q_m \cdot \mathfrak{w}_m(t,\varphi_i)\right) = \sum_{m\in M} q_m \cdot \left(\sum_{t\in\mathsf{supp}(\pi_m)} \mathfrak{w}_m(t,\varphi_i)\right) = \sum_{m\in M} q_m \cdot r_i = r_i$$

$$\sum_{i\in I} \bar{\mathfrak{w}}(t,\varphi_i) = \sum_{i\in I} (q_m \cdot \mathfrak{w}_m(t,\varphi_i)) = \sum_{m\in M} q_m \cdot \left(\sum_{i\in I} \mathfrak{w}_m(t,\varphi_i)\right) = \sum_{m\in M} q_m \cdot \pi_m(t) = \pi(t).$$

Hence, $\bar{\mathfrak{w}}$ is a well-defined matching for π and $\bigoplus_{i\in I} r_i\varphi_i$. Then, note that $\bar{\mathfrak{w}}(t,\varphi_i) > 0$ iff $\mathfrak{w}_m(t,\varphi_i) > 0$ for at least one $m \in M$. Hence we infer that $t \models \varphi_i$ whenever $\bar{\mathfrak{w}}(t,\varphi_i) > 0$ and conclude $\pi \models \bigoplus_{i\in I} r_i\varphi_i$. □

We can now present the decomposition theorem.

Theorem 4 (Decomposition theorem)**.** *Let $P = (\Sigma, \mathcal{A}, R)$ be a Γ-patient PGSOS-TSS and let D_Σ be the Σ-DS. For any term $t \in \mathbb{T}(\Sigma)$, closed substitution σ and state formula $\varphi \in \mathbb{L}^{\mathrm{s}}$ we have*

$$\sigma(t) \models \varphi \Leftrightarrow \exists\, \xi \in t^{-1}(\varphi) \text{ such that for all state variables } x \in \mathrm{var}(t) \text{ it holds } \sigma(x) \models \xi(x) \tag{2}$$

and for any distribution term $\Theta \in \mathbb{DT}(\Sigma)$, closed substitution σ and distribution formula $\psi \in \mathbb{L}^{\mathrm{d}}$ we have

$$\sigma(\Theta) \models \psi \Leftrightarrow \exists\, \eta \in \Theta^{-1}(\psi) \text{ such that for all variables } \zeta \in \mathrm{var}(\Theta) \text{ it holds } \sigma(\zeta) \models \eta(\zeta). \tag{3}$$

Proof. We start with univariate terms, by proceeding by structural induction over formula $\phi \in \mathbb{L}$ to prove both statements at the same time. Showing the two statements at the same time is necessary since state formula are constructed on distribution formulae, and, analogously, distribution formulae are constructed on state formulae. For each inductive case we prove both implications. Then, we will conclude with showing how the result for univariate terms can be extended to multivariate terms.

Proof of the base case $\phi = \top$

Then by Definition 23.1 we have that $\xi \in t^{-1}(\top)$ iff $\xi(x) = \top$ for all $x \in \mathbf{V}_s$. Then Equation (2) directly follows from Definition 14.

Proof of the inductive step $\phi = \neg\varphi$

We have

$$\begin{aligned}
&\sigma(t) \models \neg\varphi \\
\Leftrightarrow\ &\sigma(t) \not\models \varphi \\
\Leftrightarrow\ &\forall\, \xi \in t^{-1}(\varphi)\ \exists\, x \in \mathrm{var}(t)\ \mathrm{s}\sigma(x) \not\models \xi(x) \\
\Leftrightarrow\ &\exists\, \mathfrak{f} \colon t^{-1}(\varphi) \to \mathrm{var}(t) \text{ s.t. } \forall\, \xi' \in t^{-1}(\varphi) \text{ it holds } \sigma(\mathfrak{f}(\xi')) \not\models \xi'(\mathfrak{f}(\xi')) \\
\Leftrightarrow\ &\exists\, \mathfrak{f} \colon t^{-1}(\varphi) \to \mathrm{var}(t) \text{ s.t. } \forall\, x \in \mathrm{var}(t) \text{ it holds } \sigma(x) \models \bigwedge_{\xi' \in \mathfrak{f}^{-1}(x)} \neg\xi'(x) \\
\Leftrightarrow\ &\exists\, \xi \in t^{-1}(\neg\varphi) \text{ s.t. } \forall\, x \in \mathrm{var}(t) \text{ it holds } \sigma(x) \models \xi(x)
\end{aligned}$$

where the second relation follows by the inductive hypothesis and the last relation follows by construction of $t^{-1}(\neg\varphi)$ (Definition 23.2). Hence, Equation (2) holds also in this case.

Proof of the inductive step $\phi = \bigwedge_{j\in J} \varphi_j$

We have

$$\sigma(t) \models \bigwedge_{j\in J} \varphi_j$$

$$\Leftrightarrow \sigma(t) \models \varphi_j, \text{ for all } j \in J$$
$$\Leftrightarrow \exists \xi_j \in t^{-1}(\varphi_j) \text{ s.t. } \forall x \in \mathrm{var}(t) \text{ it holds } \sigma(x) \models \xi_j(x), \text{ for all } j \in J$$
$$\Leftrightarrow \exists \xi_j \in t^{-1}(\varphi_j) \text{ for all } j \in J \text{ s.t. } \forall x \in \mathrm{var}(t) \text{ it holds } \sigma(x) \models \bigwedge_{j \in J} \xi_j(x)$$
$$\Leftrightarrow \exists \xi \in t^{-1}(\bigwedge_{j \in J} \varphi_j) \text{ s.t. } \forall x \in \mathrm{var}(t) \text{ it holds } \sigma(x) \models \xi(x)$$

where the second relation follows by the inductive hypothesis and the last relation follows by construction of $t^{-1}(\bigwedge_{j \in J} \varphi_j)$ (Definition 23.3). Hence, Equation (2) holds also in this case.

Proof of the inductive step $\phi = \bigoplus_{i \in I} r_i \varphi_i$

Since $\phi \in \mathbb{L}^{\mathrm{d}}$, we need to show Equation (3). To this aim, we prove the two implications separately.

($\Rightarrow$) Assume first that $\sigma(\Theta) \models \bigoplus_{i \in I} r_i \varphi_i$ and that $\sigma(\Theta) = \sum_{m \in M} q_m \delta_{t_m}$ with $t_m \in \mathbf{T}(\Sigma)$ pairwise distinct. By Definition 14, there is a matching $\tilde{\mathfrak{w}} \in \mathfrak{W}(\sigma(\Theta), \bigoplus_{i \in I} r_i \varphi_i)$ with $t_m \models \varphi_i$ whenever $\tilde{\mathfrak{w}}(t_m, \varphi_i) > 0$. Moreover, by Proposition 1, we get $D \vdash \{\sigma(\Theta) \xrightarrow{q_m} t_m \mid m \in M\}$ which, by Theorem 3, implies that there are a Σ-distribution ruloid $\rho^{\mathsf{D}} = \frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}}$ and a closed substitution σ' with $D_\Sigma \vdash \sigma'(\mathbf{H})$, $\sigma'(\Theta) = \sigma(\Theta)$ and $\sigma'(u_m) = t_m$ for each $m \in M$. Define $\mathfrak{w} \in \mathfrak{W}(\mathrm{conc}(\rho^{\mathsf{D}}), \bigoplus_{i \in I} r_i \varphi_i)$ as $\mathfrak{w}(u_m, \varphi_i) = \tilde{\mathfrak{w}}(\sigma'(u_m), \varphi_i)$ for all $m \in M, i \in I$. Then we can infer:

1. from $\sigma'(\Theta) = \sigma(\Theta)$ we obtain that $\sigma'(\zeta) = \sigma(\zeta)$ for all variables $\zeta \in \mathrm{var}(\Theta)$;
2. whenever $\mathfrak{w}(u_m, \varphi_i) > 0$ it holds that $\sigma'(u_m) \models \varphi_i$. By the inductive hypothesis we derive that there is a decomposition mapping $\xi_{m,i} \in u_m^{-1}(\varphi_i)$ s.t. $\sigma'(x) \models \xi_{m,i}(x)$ for all $x \in \mathrm{var}(u_m)$;
3. from $D_\Sigma \vdash \sigma'(\mathbf{H})$ we obtain that for all premises $\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}$ we have $D_\Sigma \vdash \{\sigma'(\zeta) \xrightarrow{q_h} t'_h \mid h \in H\}$, where $\{\sigma'(\zeta) \xrightarrow{q_h} t'_h \mid h \in H\}$ is $\sigma'(\{\zeta \xrightarrow{q_j} x_j \mid j \in J\})$, for a suitable set of indexes H and proper terms t'_h. By Proposition 1, $D_\Sigma \vdash \{\sigma'(\zeta) \xrightarrow{q_h} t'_h \mid h \in H\}$ iff $\sigma'(\zeta)(t'_h) = q_h$ and $\sum_{h \in H} q_h = 1$. Hence, as the t'_h are pairwise distinct, we have that
$$\sigma'(\zeta) = \sum_{h \in H} q_h \delta_{t'_h} = \sum_{h \in H} (\sum_{j \in J, \sigma'(x_j) = t'_h} q_j) \delta_{t'_h} = \sum_{h \in H} (\sum_{j \in J, \sigma'(x_j) = t'_h} q_j \delta_{\sigma'(x_j)}) = \sum_{j \in J} q_j \delta_{\sigma'(x_j)}$$

Let $\eta \in \Theta^{-1}(\bigoplus_{i \in I} r_i \varphi_i)$ be the decomposition mapping defined as in Definition 23.8 by means of the Σ-distribution ruloid $\frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}}$ and the decomposition mappings $\xi_{m,i}$ as in item (2) above for each $m \in M$ and $i \in I$ s.t. $\mathfrak{w}(u_m, \varphi_i) > 0$, and $\xi_{m,i}$ defined by $\xi_{m,i}(x) = \top$ for all $x \in \mathbf{V}_s$ for those m, i s.t. $\mathfrak{w}(u_m, \varphi_i) = 0$. We aim to show that for this η it holds that $\sigma'(\zeta) \models \eta(\zeta)$ for each $\zeta \in \mathrm{var}(\Theta)$. By construction,

$$\eta(\zeta) = \begin{cases} \displaystyle\bigoplus_{\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}} q_j \bigwedge_{\substack{m \in M \\ i \in I}} \xi_{m,i}(x_j) & \text{if } \zeta \in \mathbf{V}_d \\ \displaystyle\bigwedge_{\substack{m \in M \\ i \in I}} \xi_{m,i}(x) & \text{if } \zeta = x \in \mathbf{V}_s. \end{cases}$$

For each $y \in \{x_j \mid j \in J\} \cup \{x\}$ and for each $m \in M$ and $i \in I$, we distinguish three cases:

4. $y \in \mathrm{var}(u_m)$ and $\mathfrak{w}(u_m, \varphi_i) > 0$. Then, by item (2) above, we have $\sigma'(y) \models \xi_{m,i}(y)$.
5. $y \in \mathrm{var}(u_m)$ and $\mathfrak{w}(u_m, \varphi_i) = 0$. Then by construction $\xi_{m,i}(y) = \top$, thus giving that $\sigma'(y) \models \xi_{m,i}(y)$ holds trivially also in this case.

6. $y \notin \text{var}(u_m)$. Then, whichever is the value of $\mathfrak{w}(u_m, \varphi_i)$, we have $\xi_{m,i}(y) = \top$ (see Definition 23) and consequently $\sigma'(y) \models \xi_{m,i}(y)$ holds trivially also in this case.

Since these considerations apply to each $m \in M$ and $i \in I$ we can conclude that if $\zeta \in \mathbf{V}_d$ then for all $\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}$ it holds that for each x_j with $j \in J$ we have $\sigma'(x_j) \models \bigwedge_{m \in M, i \in I} \xi_{m,i}(x_j)$. Furthermore, by item (3) above, if $\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}$ then $D_\Sigma \vdash \sigma'(\mathbf{H})$ gives $\sigma'(\zeta) = \sum_{j \in J} q_j \delta_{\sigma'(x_j)}$, from which we can conclude that

$$\sigma'(\zeta) \models \bigoplus_{j \in J} q_j \bigwedge_{i \in I, m \in M} \xi_{m,i}(x_j), \text{ namely } \sigma'(\zeta) \models \eta(\zeta).$$

Similarly, if $\zeta = x \in \mathbf{V}_s$ then

$$\sigma'(x) \models \bigwedge_{m \in M, i \in I} \xi_{m,i}(x), \text{ namely } \sigma'(x) \models \eta(x).$$

Thus, we can conclude that for each $\zeta \in \text{var}(\Theta)$ it holds that $\sigma'(\zeta) \models \eta(\zeta)$. Since moreover $\sigma(\zeta) = \sigma'(\zeta)$ (item (1) above), we can conclude that $\sigma(\zeta) \models \eta(\zeta)$ as required.

($\Leftarrow$) Assume now that there is a decomposition mapping $\eta \in \Theta^{-1}(\bigoplus_{i \in I} r_i \varphi_i)$ s.t. $\sigma(\zeta) \models \eta(\zeta)$ for all $\zeta \in \text{var}(\Theta)$. Following Definition 23.8, the existence of such a decomposition mapping η entails the existence of a Σ-distribution ruloid $\rho^{\mathsf{D}} = \frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}}$ with $\sum_{m \in M} q_m = 1$ (Lemma 1) and of a matching $\mathfrak{w}$ for $\text{conc}(\rho^{\mathsf{D}})$ and $\bigoplus_{i \in I} r_i \varphi_i$ from which we can build the following decomposition mappings:

$$\begin{cases} \xi_{m,i} \in t_m^{-1}(\varphi_i) & \text{if } \mathfrak{w}(t_m, \varphi_i) > 0 \\ \xi_{m,i} \in t_m^{-1}(\top) & \text{otherwise.} \end{cases}$$

In particular, we have that for each $\mu \in \text{var}(\Theta)$

$$\eta(\mu) = \bigoplus_{\{\mu \xrightarrow{q_j} x_j \mid \sum_{j \in J} q_j = 1\} \in \mathbf{H}} q_j \bigwedge_{i \in I, m \in M} \xi_{m,i}(x_j)$$

and for each $x \in \text{var}(\Theta)$

$$\eta(x) = \bigwedge_{i \in I, m \in M} \xi_{m,i}(x).$$

We define a closed substitution σ' s.t. $\sigma'(\zeta) = \sigma(\zeta)$ for each $\zeta \in \text{var}(\Theta)$ and $\sigma'(x) = \sigma(x)$ for each $x \in \text{rhs}(\mathbf{H})$. Then, the following properties hold:

(a) From $\sigma'(\zeta) = \sigma(\zeta)$ and $\sigma(\zeta) \models \eta(\zeta)$ we derive $\sigma'(\zeta) \models \eta(\zeta)$. In particular we obtain that $\sigma'(x) \models \bigwedge_{i \in I, m \in M} \xi_{m,i}(x)$ for each $x \in \text{var}(\Theta)$.

(b) As $\sigma'(\mu) \models \eta(\mu)$ for each $\mu \in \text{var}(\Theta)$, and, by Definition 23.8a, the weights of the distribution formula $\eta(\mu)$ coincide with the weights of the distribution literals in $\{\mu \xrightarrow{q_j} x_j \mid \sum_{j \in J} q_j = 1\} \in \mathbf{H}$, we gather $\sigma'(x_j) \models \bigwedge_{i \in I, m \in M} \xi_{m,i}(x_j)$, for each $j \in J$.

(c) From $\sigma(\zeta) = \sigma'(\zeta)$ for each $\zeta \in \text{var}(\Theta)$ we infer that $\sigma'(\Theta) = \sigma(\Theta)$. Moreover, by Lemma 2.3 we have that $\text{rhs}(\mathbf{H}) = \bigcup_{m \in M} \text{var}(t_m)$, so that $\sigma'(x) = \sigma(x)$ for each $x \in \text{rhs}(\mathbf{H})$ implies $\sigma'(t_m) = \sigma(t_m)$ for each $m \in M$.

From items (a), (b) above and by structural induction we gather $\sigma'(t_m) \models \varphi_i$ for each $m \in M, i \in I$ with $\mathfrak{w}(t_m, \varphi_i) > 0$. Moreover, from $\sigma'(\zeta) \models \eta(\zeta)$ for each $\zeta \in \text{var}(\Theta)$, item (a) above, we obtain that $D_\Sigma \vdash \sigma'(\mathbf{H})$, namely D_Σ proves the reduced instance w.r.t. σ' of each set of distribution premises $\{\zeta \xrightarrow{q_j}$

$x_j \mid \sum_{j\in J} q_j = 1\} \in \mathbf{H}$. This fact taken together with item (c) above and Theorem 3 gives that D_Σ proves the reduced instance of $\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}$ wrt. σ, that is $D_\Sigma \vdash \{\sigma(\Theta) \xrightarrow{q_h} t'_h \mid h \in H\}$ for a suitable set of indexes H and a proper set of closed terms t'_h s.t. for each $h \in H$ there is at least one $m \in M$ s.t. $t'_h = \sigma'(t_m)$ and moreover $q_h = \sum_{\{m\in M \mid \sigma'(t_m)=t'_h\}} q_m$ (Definition 17). In addition, by Proposition 1 it follows that $q_h = \sigma(\Theta)(t'_h)$ for each $h \in H$ and $\sum_{h\in H} q_h = 1$. Since moreover $q_h \in (0,1]$ for each $h \in H$, this is equivalent to say that $\sigma(\Theta) = \sum_{h\in H} q_h \delta_{t'_h}$.

To conclude, we exhibit a matching in $\mathfrak{W}(\sigma(\Theta), \bigoplus_{i\in I} r_i\varphi_i)$ s.t. $t'_h \models \varphi_i$ whenever $\tilde{\mathfrak{w}}(t'_h, \varphi_i) > 0$. Define $\tilde{\mathfrak{w}}(t_h, \varphi_i) = \sum_{\{m\in M \mid \sigma'(t_m)=t_h\}} \mathfrak{w}(t_m, \varphi_i)$. As $\mathfrak{w} \in \mathfrak{W}(\mathrm{conc}(\rho^{\mathsf{D}}), \bigoplus_{i\in I} r_i\varphi_i)$, we have

$$\sum_{h\in H} \tilde{\mathfrak{w}}(t'_h, \varphi_i) = \sum_{m\in M} \mathfrak{w}(t_m, \varphi_i) = r_i$$

$$\sum_{i\in I} \tilde{\mathfrak{w}}(t'_h, \varphi_i) = \sum_{\{m\in M \mid \sigma'(t_m)=t'_h\}} \left(\sum_{i\in I} \mathfrak{w}(t_m, \varphi_i) \right) = \sum_{\{m\in M \mid \sigma'(t_m)=t'_h\}} q_m = q_h.$$

Hence $\tilde{\mathfrak{w}}$ is a well-defined matching for $\sigma(\Theta)$ and $\bigoplus_{i\in I} r_i\varphi_i$. Moreover, notice that $\tilde{\mathfrak{w}}(t_h, \varphi_i) > 0$ iff $\mathfrak{w}(t_m, \varphi_i) > 0$ for at least one index $m \in M$ with $\sigma'(t_m) = t_h$. Since $\mathfrak{w}(t_m, \varphi_i) > 0$ implies $\sigma'(t_m) \models \varphi_i$, we can infer that $t_h \models \varphi_i$ whenever $\tilde{w}(t_h, \varphi_i) > 0$. Therefore, we can conclude that $\sigma(\Theta) \models \bigoplus_{i\in I} r_i\varphi_i$ as requested.

Hence, Equation (3) follows from the two implications.

Proof of the inductive step $\phi = \bigwedge_{j\in J} \psi_j$

Since $\phi \in \mathbb{L}^{\mathrm{d}}$, we need to show Equation (3). We have

$$\begin{aligned}
&\sigma(\Theta) \models \bigwedge_{j\in J} \psi_j \\
\Leftrightarrow\ & \sigma(\Theta) \models \psi_j, \text{ for all } j \in J \\
\Leftrightarrow\ & \exists\, \eta_j \in \Theta^{-1}(\psi_j) \text{ s.t. } \forall\, \zeta \in \mathrm{var}(\Theta) \text{ it holds } \sigma(\zeta) \models \eta_j(\zeta), \text{ for all } j \in J \\
\Leftrightarrow\ & \exists\, \eta_j \in \Theta^{-1}(\psi_j) \text{ for all } j \in J \text{ s.t. } \forall\, \zeta \in \mathrm{var}(\Theta) \text{ it holds } \sigma(\zeta) \models \bigwedge_{j\in J} \eta_j(\zeta) \\
\Leftrightarrow\ & \exists\, \eta \in \Theta^{-1}(\bigwedge_{j\in J} \psi_j) \text{ s.t. } \forall\, \zeta \in \mathrm{var}(\Theta) \text{ it holds } \sigma(\zeta) \models \eta(\zeta)
\end{aligned}$$

where the second relation follows by the inductive hypothesis and the last relation follows by construction of $\Theta^{-1}(\bigwedge_{j\in J} \psi_j)$ (Definition 23.9). Hence, Equation (3) holds also in this case.

Proof of the inductive step $\phi = \left(\bigoplus_{i\in I} r_i\varphi_i\right)^{\langle\varepsilon\rangle}$

We prove the two implications separately.

($\Rightarrow$) Assume first that $\sigma(\Theta) \models \bigoplus_{i\in I} r_i\varphi_i$ and that $\sigma(\Theta) = \sum_{m\in M} q_m \delta_{t_m}$ with $t_m \in \mathbf{T}(\Sigma)$ pairwise distinct. By Proposition 1, we get $D \vdash \{\sigma(\Theta) \xrightarrow{q_m} t_m \mid m \in M\}$ which, by Theorem 3, implies that there are a Σ-distribution ruloid $\rho^{\mathsf{D}} = \dfrac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m\in M\}}$ and a closed substitution σ' with $D_\Sigma \vdash \sigma'(\mathbf{H})$, $\sigma'(\Theta) = \sigma(\Theta)$ and $\sigma'(u_m) = t_m$ for each $m \in M$. Then we can infer that:

1. from $\sigma'(\Theta) = \sigma(\Theta)$ we obtain that $\sigma'(\zeta) = \sigma(\zeta)$ for all variables $\zeta \in \mathrm{var}(\Theta)$;
2. from $\sigma'(u_m) = t_m$ and Proposition 2, we get that there is a matching $\tilde{\mathfrak{w}} \in \mathfrak{W}(\sigma(\Theta), \bigoplus_{i\in I} r_i\varphi_i)$ s.t. for each $m \in M$ we have $\sigma'(u_m) \models \langle\varepsilon\rangle\psi_m$ for ψ_m built on $\tilde{\mathfrak{w}}$ as in Definition 23.10. By the inductive hypothesis we derive that, for each $m \in M$, there is a decomposition mapping $\xi_m \in u_m^{-1}(\langle\varepsilon\rangle\psi_m)$ s.t. $\sigma'(x) \models \xi_m(x)$ for all $x \in \mathrm{var}(u_m)$;

3. from $D_\Sigma \vdash \sigma'(\mathbf{H})$ we obtain that for all premises $\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}$ we have $D_\Sigma \vdash \{\sigma'(\zeta) \xrightarrow{q_h} t'_h \mid h \in H\}$, where $\{\sigma'(\zeta) \xrightarrow{q_h} t'_h \mid h \in H\}$ is $\sigma'(\{\zeta \xrightarrow{q_j} x_j \mid j \in J\})$, for a suitable set of indexes H and proper terms t'_h. By Proposition 1, $D_\Sigma \vdash \{\sigma'(\zeta) \xrightarrow{q_h} t'_h \mid h \in H\}$ iff $\sigma'(\zeta)(t'_h) = q_h$ and $\sum_{h \in H} q_h = 1$. Hence, as the t'_h are pairwise distinct, we have

$$\sigma'(\zeta) = \sum_{h \in H} q_h \delta_{t'_h} = \sum_{h \in H} \Big(\sum_{j \in J, \sigma'(x_j) = t'_h} q_j \Big) \delta_{t'_h} = \sum_{h \in H} \Big(\sum_{j \in J, \sigma'(x_j) = t'_h} q_j \delta_{\sigma'(x_j)} \Big) = \sum_{j \in J} q_j \delta_{\sigma'(x_j)}.$$

Let $\eta \in \Theta^{-1}\left((\bigoplus_{i \in I} r_i \varphi_i)^{\langle \varepsilon \rangle}\right)$ be the decomposition mapping defined as in Definition 23.10 by means of the Σ-distribution ruloid $\frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} u_m \mid m \in M\}}$ and the decomposition mappings ξ_m as in item (2) above for each $m \in M$ and $i \in I$ s.t. $\mathfrak{w}(u_m, \varphi_i) > 0$, and ξ_m defined by $\xi_m(x) = \top$ for all $x \in \mathbf{V}_s$ for those m, i s.t. $\mathfrak{w}(u_m, \varphi_i) = 0$. We aim to show that for this η it holds that $\sigma'(\zeta) \models \eta(\zeta)$ for each $\zeta \in \mathrm{var}(\Theta)$. By construction,

$$\eta(\zeta) = \begin{cases} \displaystyle\bigoplus_{\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}} q_j \bigwedge_{m \in M} \xi_m(x_j) & \text{if } \zeta \in \mathbf{V}_d \\ \displaystyle\bigwedge_{m \in M} \xi_m(x) & \text{if } \zeta = x \in \mathbf{V}_s. \end{cases}$$

For each variable $y \in \{x_j \mid j \in J\} \cup \{x\}$ and for each $m \in M$ and $i \in I$, we distinguish three cases:

4. $y \in \mathrm{var}(u_m)$ and $\mathfrak{w}(u_m, \varphi_i) > 0$. Then, by item (2) above, we have $\sigma'(y) \models \xi_m(y)$.
5. $y \in \mathrm{var}(u_m)$ and $\mathfrak{w}(u_m, \varphi_i) = 0$. Then by construction $\xi_m(y) = \top$, thus giving that $\sigma'(y) \models \xi_m(y)$ holds trivially also in this case.
6. $y \notin \mathrm{var}(u_m)$. Then, whichever is the value of $\mathfrak{w}(u_m, \varphi_i)$, we have $\xi_m(y) = \top$ (see Definition 23) and consequently $\sigma'(y) \models \xi_m(y)$ holds trivially also in this case.

Since these considerations apply to each $m \in M$ and $i \in I$ we can conclude that if $\zeta \in \mathbf{V}_d$ then for all $\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}$ it holds that for each x_j with $j \in J$ we have $\sigma'(x_j) \models \bigwedge_{m \in M} \xi_m(x_j)$. Furthermore, by item (3) above, if $\{\zeta \xrightarrow{q_j} x_j \mid j \in J\} \in \mathbf{H}$ then $D_\Sigma \vdash \sigma'(\mathbf{H})$ gives $\sigma'(\zeta) = \sum_{j \in J} q_j \delta_{\sigma'(x_j)}$, from which we can conclude that

$$\sigma'(\zeta) \models \bigoplus_{j \in J} q_j \bigwedge_{m \in M} \xi_m(x_j), \text{ namely } \sigma'(\zeta) \models \eta(\zeta).$$

Similarly, if $\zeta = x \in \mathbf{V}_s$ then

$$\sigma'(x) \models \bigwedge_{m \in M} \xi_m(x), \text{ namely } \sigma'(x) \models \eta(x).$$

Thus, we can conclude that for each $\zeta \in \mathrm{var}(\Theta)$ it holds that $\sigma'(\zeta) \models \eta(\zeta)$. Since moreover $\sigma(\zeta) = \sigma'(\zeta)$ (item (1) above), we can conclude that $\sigma(\zeta) \models \eta(\zeta)$ as required.

($\Leftarrow$) Assume now that there is a decomposition mapping $\eta \in \Theta^{-1}\left((\bigoplus_{i \in I} r_i \varphi_i)^{\langle \varepsilon \rangle}\right)$ s.t. $\sigma(\zeta) \models \eta(\zeta)$ for all $\zeta \in \mathrm{var}(\Theta)$. Following Definition 23.10, the existence of such a decomposition mapping η entails the existence of a Σ-distribution ruloid $\rho^{\mathsf{D}} = \frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}}$ with $\sum_{m \in M} q_m = 1$ (Lemma 1) and of a matching $\mathfrak{w}$ for $\mathrm{conc}(\rho^{\mathsf{D}})$ and $\bigoplus_{i \in I} r_i \varphi_i$ from which we can build the following decomposition mappings:

$$\xi_m \in t_m^{-1}(\langle \varepsilon \rangle \psi_m) \text{ with } \psi_m = \bigoplus_{i \in I_m} \frac{\mathfrak{w}(t_m, \varphi_i)}{q_m} \varphi_i$$

In particular, we have that for each $\mu \in \mathrm{var}(\Theta)$

$$\eta(\mu) = \bigoplus_{\{\mu \xrightarrow{q_j} x_j \mid \sum_{j\in J} q_j = 1\} \in \mathbf{H}} q_j \bigwedge_{m\in M} \xi_m(x_j)$$

and for each $x \in \mathrm{var}(\Theta)$

$$\eta(x) = \bigwedge_{m\in M} \xi_m(x).$$

We define a closed substitution σ' s.t. $\sigma'(\zeta) = \sigma(\zeta)$ for each $\zeta \in \mathrm{var}(\Theta)$ and $\sigma'(x) = \sigma(x)$ for each $x \in \mathrm{rhs}(\mathbf{H})$. Then, the following properties hold:

(a) From $\sigma'(\zeta) = \sigma(\zeta)$ and $\sigma(\zeta) \models \eta(\zeta)$ we derive $\sigma'(\zeta) \models \eta(\zeta)$. In particular we obtain that $\sigma'(x) \models \bigwedge_{m\in M} \xi_m(x)$ for each $x \in \mathrm{var}(\Theta)$.

(b) As $\sigma'(\mu) \models \eta(\mu)$ for each $\mu \in \mathrm{var}(\Theta)$, and, by Definition 23.10b, the weights of the distribution formula $\eta(\mu)$ coincide with the weights of the distribution literals in $\{\mu \xrightarrow{q_j} x_j \mid \sum_{j\in J} q_j = 1\} \in \mathbf{H}$, we gather $\sigma'(x_j) \models \bigwedge_{m\in M} \xi_m(x_j)$, for each $j \in J$.

(c) From $\sigma(\zeta) = \sigma'(\zeta)$ for each $\zeta \in \mathrm{var}(\Theta)$ we infer that $\sigma'(\Theta) = \sigma(\Theta)$. Moreover, by Lemma 2.3 we have that $\mathrm{rhs}(\mathbf{H}) = \bigcup_{m\in M} \mathrm{var}(t_m)$, so that $\sigma'(x) = \sigma(x)$ for each $x \in \mathrm{rhs}(\mathbf{H})$ implies $\sigma'(t_m) = \sigma(t_m)$ for each $m \in M$.

From items (a), (b) above and by structural induction we gather $\sigma'(t_m) \models \langle\varepsilon\rangle\psi_m$ for each $m \in M$. Moreover, from $\sigma'(\zeta) \models \eta(\zeta)$ for each $\zeta \in \mathrm{var}(\Theta)$, item (a) above, we obtain that $D_\Sigma \vdash \sigma'(\mathbf{H})$, namely D_Σ proves the reduced instance w.r.t, σ' of each set of distribution premises $\{\zeta \xrightarrow{q_j} x_j \mid \sum_{j\in J} q_j = 1\} \in \mathbf{H}$. This fact taken together with item (c) above and Theorem 3 gives that D_Σ proves the reduced instance of $\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}$ wrt. σ. Finally, notice that by Proposition 2, from $\sigma'(t_m) \models \langle\varepsilon\rangle\psi_m$ for each $m \in M$, we can conclude that $\sigma(\Theta) \models \big(\bigoplus_{i\in I} r_i\varphi_i\big)^{\langle\varepsilon\rangle}$ as requested.

Proof of the inductive step $\phi = (\bigwedge_{j\in J} \psi_j)^{\langle\varepsilon\rangle}$

The proof for this case is analogous to the one for the inductive step $\phi = (\bigoplus_{i\in I} r_i\varphi_i)^{\langle\varepsilon\rangle}$.

Proof of the inductive step $\phi = \langle\alpha\rangle\psi$

Since $\phi \in \mathbb{L}^{\mathrm{s}}$, we need to show Equation (2). To this aim, we prove the two implications separately.

($\Rightarrow$) Assume first that $\sigma(t) \models \langle\alpha\rangle\psi$. Then, by Definition 14, there exists a probability distribution $\pi \in \Delta(\mathbf{T}(\Sigma))$ with $P \vdash \sigma(t) \xrightarrow{\alpha} \pi$ and $\pi \models \psi$. By Theorem 2, $P \vdash \sigma(t) \xrightarrow{\alpha} \pi$ implies that there are a P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\alpha} \Theta}$ and a closed substitution σ' with $P \vdash \sigma'(\mathbf{H})$, $\sigma'(t) = \sigma(t)$ and $\sigma'(\Theta) = \pi$. We infer that: the following facts:

1. from $\sigma'(t) = \sigma(t)$ we obtain that $\sigma'(x) = \sigma(x)$ for all $x \in \mathrm{var}(t)$;

2. from $\sigma'(\Theta) = \pi$ and $\pi \models \psi$, we gather $\sigma'(\Theta) \models \psi$ and by the inductive hypothesis we obtain that there exists a $\eta \in \Theta^{-1}(\psi)$ s.t. $\sigma'(\zeta) \models \eta(\zeta)$ for all $\zeta \in \mathrm{var}(\Theta)$;

3. from $P \vdash \sigma'(\mathbf{H})$ we obtain that whenever $x \xrightarrow{\beta} \mu \in \mathbf{H}$ we have $P \vdash \sigma'(x) \xrightarrow{\beta} \sigma'(\mu)$. Then, if $\mu \in \mathrm{var}(\Theta)$, by previous item (2), we get $\sigma'(\mu) \models \eta(\mu)$. Otherwise, if $\mu \notin \mathrm{var}(\Theta)$, we have $\eta(\mu) = \top$ thus giving $\sigma'(\mu) \models \eta(\mu)$ also in this case. Hence, $\sigma'(\mu) \models \eta(\mu)$ and $\sigma'(x) \models \langle\beta\rangle\eta(\mu)$ in all cases.

4. from $P \vdash \sigma'(\mathbf{H})$ we obtain that whenever $x \stackrel{\gamma}{\not\rightarrow} \in \mathbf{H}$ we have $P \vdash \sigma'(x) \stackrel{\gamma}{\not\rightarrow}$, namely $P \not\vdash \sigma'(x) \xrightarrow{\gamma} \upsilon$ for any $\upsilon \in \mathbf{DT}(\Sigma)$, giving $\sigma'(x) \models \neg\langle\gamma\rangle\top$.

Let $\xi \in t^{-1}(\langle\alpha\rangle\psi)$ be defined as in Definition 23.4 by means of the P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\alpha} \Theta}$ and the decomposition mapping η introduced in item (2) above. We aim to show that for this ξ it holds that $\sigma'(x) \models \xi(x)$ for each $x \in \mathrm{var}(t)$. By construction,

$$\xi(x) = \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \wedge \bigwedge_{x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}} \neg\langle\gamma\rangle\top \wedge \eta(x).$$

By item (3) above we have $\sigma'(x) \models \langle\beta\rangle\eta(\mu)$ for each $x \xrightarrow{\beta} \mu \in \mathbf{H}$. By item (4) above we have $\sigma'(x) \models \neg\langle\gamma\rangle\top$ for each $x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}$. Finally, if $x \in \mathrm{var}(\Theta)$ by item (2) above we get $\sigma'(x) \models \eta(x)$. If $x \notin \mathrm{var}(\Theta)$ then we have $\eta(x) = \top$ (Definition 23.8b) thus giving $\sigma'(x) \models \eta(x)$ also in this case. Hence, $\sigma'(x) \models \eta(x)$ in all cases. Thus, we can conclude that $\sigma'(x) \models \xi(x)$. Since, by item (1) above, $\sigma(x) = \sigma'(x)$ we can conclude that $\sigma(x) \models \xi(x)$ as required.

($\Leftarrow$) Assume now that there is a $\xi \in t^{-1}(\langle\alpha\rangle\psi)$ s.t. $\sigma(x) \models \xi(x)$ for all $x \in \mathrm{var}(t)$. Following Definition 23.4, we construct ξ in terms of some P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\alpha} \Theta}$ and decomposition mapping $\eta \in \Theta^{-1}(\psi)$. In particular, we have that for each $x \in \mathrm{var}(t)$

$$\xi(x) = \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \wedge \bigwedge_{x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}} \neg\langle\gamma\rangle\top \wedge \eta(x).$$

We define a closed substitution σ' s.t. the following properties hold:

(a) $\sigma'(x) = \sigma(x)$ for all $x \in \mathrm{var}(t)$. As a consequence, from $\sigma(x) \models \xi(x)$ we derive $\sigma'(x) \models \xi(x)$.

(b) As $\sigma'(x) \models \xi(x)$, by previous item (a), we derive that $\sigma'(x) \models \langle\beta\rangle\eta(\mu)$ for each $x \xrightarrow{\beta} \mu \in \mathbf{H}$. This implies that for each positive premise in $\mathbf{H}$ there exists a probability distribution $\pi_{\beta,\mu}$ s.t. $P \vdash \sigma'(x) \xrightarrow{\beta} \pi_{\beta,\mu}$ and $\pi_{\beta,\mu} \models \eta(\mu)$. We define $\sigma'(\mu) = \pi_{\beta,\mu}$ thus obtaining that for each $x \xrightarrow{\beta} \mu \in \mathbf{H}$ we have $P \vdash \sigma'(x) \xrightarrow{\beta} \sigma'(\mu)$ and $\sigma'(\mu) \models \eta(\mu)$.

(c) As $\sigma'(x) \models \xi(x)$, by previous item (a), we derive that $\sigma'(x) \models \neg\langle c\rangle\top$ for each $x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}$. Therefore, we obtain that $P \vdash \sigma'(x) \stackrel{\gamma}{\not\longrightarrow}$ for each $x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}$.

(d) Since $\mathrm{var}(\Theta) \subseteq \mathrm{var}(t) \cup \mathrm{rhs}(\mathbf{H})$, previous items (b) and (c) we obtain that $\sigma'(\mu) \models \eta(\mu)$ for each $\mu \in \mathbf{V}_d$.

(e) $\sigma'(x) \models \eta(x)$ for each $x \in \mathrm{var}(\Theta)$.

From items (d), (e) and structural induction, we gather $\sigma'(\Theta) \models \psi$. Moreover, items (b) and (c) give $P \vdash \sigma'(\mathbf{H})$. Hence, by Theorem 2 we obtain $P \vdash \sigma'(t) \xrightarrow{\alpha} \sigma'(\Theta)$. From item (a) we have that $\sigma'(t) = \sigma(t)$ and, therefore, we can conclude that $\sigma(t) \models \langle\alpha\rangle\psi$.

Hence, Equation (2) follows from the two implications.

Proof of the inductive step $\phi = \langle\hat{\tau}\rangle\psi$

We prove the two implications separately.

($\Rightarrow$) Assume first that $\sigma(t) \models \langle\hat{\tau}\rangle\psi$. Then, by Definition 14, there is a distribution π s.t. $\sigma(t) \xrightarrow{\hat{\tau}} \pi$ and $\pi \models \psi$. We can distinguish two cases:

1. Either $\pi = \delta_{\sigma(t)}$. Then $\delta_{\sigma(t)} \models \psi$ and structural induction imply that there is a decomposition mapping $\eta \in \delta_t^{-1}(\psi)$ s.t. $\sigma(x) \models \eta(x)$ for each $x \in \mathrm{var}(t)$. Then, define $\xi \in t^{-1}(\langle\hat{\tau}\rangle\psi)$ by means of η as in Definition 23.5a, namely $\xi(x) = \langle\hat{\tau}\rangle\eta(x)$ for all x occurring Γ-liquid in t and $\xi(y) = \eta(y)$ for all other variables. Clearly, for Γ-liquid variables $\sigma(x) \models \eta(x)$ implies $\sigma(x) \models \langle\hat{\tau}\rangle\eta(x)$ and thus $\sigma(x) \models \xi(x)$ follows for all $x \in \mathrm{var}(t)$.

2. Or $P \vdash \sigma(t) \xrightarrow{\tau} \pi$. By Theorem 2, $P \vdash \sigma(t) \xrightarrow{\tau} \pi$ implies that there are a P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and a closed substitution σ' with $P \vdash \sigma'(\mathbf{H})$, $\sigma'(t) = \sigma(t)$ and $\sigma'(\Theta) = \pi$. We infer that:
   (a) from $\sigma'(t) = \sigma(t)$ we obtain that $\sigma'(x) = \sigma(x)$ for all $x \in \mathrm{var}(t)$;
   (b) from $\sigma'(\Theta) = \pi$ and $\pi \models \psi$, we gather $\sigma'(\Theta) \models \psi$ and by the inductive hypothesis we obtain that there exists a $\eta \in \Theta^{-1}(\psi)$ s.t. $\sigma'(\zeta) \models \eta(\zeta)$ for all $\zeta \in \mathrm{var}(\Theta)$.

   We can distinguish two cases:
   - $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ is Γ-patient, namely it is of the form $\frac{x_0 \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/x_0]}$ and the unique occurrence of μ in $t[\mu/x_0]$ is Γ-liquid. Consider the decomposition mapping η obtained in item 2b above and define from it $\xi \in t^{-1}(\langle\hat{\tau}\rangle\psi)$ as in Definition 23.5b. Clearly, $\sigma(\mu) \models \eta(\mu)$ giving that $\sigma(x_0) \models \langle\hat{\tau}\rangle\eta(\mu)$ and $\sigma(x) \models \eta(x)$ for all other variables. Hence we have obtained that $\sigma(x) \models \xi(x)$ for all $x \in \mathrm{var}(t)$.
   - $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ is Γ-impatient. In this case the thesis follows by applying the same arguments used in the proof of the inductive step $\phi = \langle\alpha\rangle\psi$ to a decomposition mapping built as in Definition 23.5c.

($\Leftarrow$) Assume now that there is a $\xi \in t^{-1}(\langle\hat{\tau}\rangle\psi)$ s.t. $\sigma(x) \models \xi(x)$ for all $x \in \mathrm{var}(t)$. Following Definition 23.5, we can distinguish three cases.

1. ξ is constructed in terms of a decomposition mapping $\eta \in \delta_t^{-1}(\psi)$ as in Definition 23.5a. This implies that $\sigma(x) \models \eta(x)$ for all $x \in \mathrm{var}(t)$ and thus, by induction we gather that $\delta_{\sigma(t)} \models \psi$. By Definition 14, we can therefore infer that $\sigma(t) \models \langle\hat{\tau}\rangle\psi$.

2. ξ is constructed in terms of a Γ-liquid argument x_0 of t, a Γ-patient ruloid $\frac{x_0 \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/x_0]}$ and of a decomposition mapping $\eta \in t[\mu/x_0]^{-1}(\psi)$ as in Definition 23.5b. In this case the thesis follows by applying the same arguments used in the proof of the inductive step $\phi = \langle\alpha\rangle\psi$.

3. ξ is constructed in terms of a Γ-impatient P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and of a decomposition mapping $\eta \in \Theta^{-1}(\psi)$ as in Definition 23.5c. Also in this case the thesis follows by applying the same arguments used in the proof of the inductive step $\phi = \langle\alpha\rangle\psi$.

Hence, Equation (2) follows from the two implications.

Proof of the inductive step $\phi = \langle\varepsilon\rangle\psi$

We prove the two implications separately.

($\Rightarrow$) Assume first that $\sigma(t) \models \langle\varepsilon\rangle\psi$. Then, by Definition 14, there are distributions $\pi_0, \ldots, \pi_n$, for some $n \in \mathbb{N}$ s.t. $\sigma(t) \xrightarrow{\hat{\tau}} \pi_0 \xrightarrow{\hat{\tau}} \pi_2 \xrightarrow{\hat{\tau}} \ldots \xrightarrow{\hat{\tau}} \pi_n$ and $\pi_n \models \psi$. We can distinguish two cases.

1. $\pi_0 = \delta_{\sigma(t)} \models \psi$ and structural induction imply that there is a decomposition mapping $\eta \in \delta_t^{-1}(\psi)$ s.t. $\sigma(x) \models \eta(x)$ for each $x \in \mathrm{var}(t)$. Then, define $\xi \in t^{-1}(\langle\hat{\varepsilon}\rangle\psi)$ by means of η as in Definition 23.6a. Clearly, $\sigma(x) \models \xi(x)$ follows for all $x \in \mathrm{var}(t)$.

2. We have $P \vdash \sigma(t) \xrightarrow{\tau} \pi$ and $\pi \models \psi^{\langle\varepsilon\rangle}$. By Theorem 2, $P \vdash \sigma(t) \xrightarrow{\tau} \pi$ implies that there are a P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and a closed substitution σ' with $P \vdash \sigma'(\mathbf{H})$, $\sigma'(t) = \sigma(t)$ and $\sigma'(\Theta) = \pi$. We infer the following facts:
   (a) from $\sigma'(t) = \sigma(t)$ we obtain that $\sigma'(x) = \sigma(x)$ for all $x \in \mathrm{var}(t)$;
   (b) Assume wlog. that $\mathsf{supp}(\sigma'(\Theta)) = \{\sigma'(t_m) \mid m \in M\} \subseteq \mathbf{T}(\Sigma)$. From $\sigma'(\Theta) = \pi$, $\pi \models \psi^{\langle\varepsilon\rangle}$ and structural induction we get that there is a decomposition mapping $\eta \in \Theta^{-1}(\psi^{\langle\varepsilon\rangle})$ s.t. $\sigma'(\zeta) \models \eta(\zeta)$ for all $\zeta \in \mathrm{var}(\Theta)$.

We can distinguish two cases:

- $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ is Γ-patient, namely it is of the form $\frac{x_0 \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/x_0]}$ and the unique occurrence of μ in $t[\mu/x_0]$ is Γ-liquid. Consider the decomposition mapping η obtained in item 2b above and define from it $\xi \in t^{-1}(\langle\varepsilon\rangle\psi)$ as in Definition 23.6b. Clearly, $\sigma(\mu) \models \eta(\mu)$ giving that $\sigma(x_0) \models \langle\varepsilon\rangle\eta(\mu)$ (as $P \vdash \sigma(x) \xrightarrow{\tau} \sigma'(\mu)$), and $\sigma(x) \models \eta(x)$ for all other variables. Hence we have obtained that $\sigma(x) \models \xi(x)$ for all $x \in \mathrm{var}(t)$.
- $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ is Γ-impatient. In this case the thesis follows by applying the same arguments used in the proof of the inductive step $\phi = \langle\alpha\rangle\psi$ to a decomposition mapping ξ built as in Definition 23.6c. We simply need to notice that for a Γ-liquid argument x in t, whenever $\sigma(x) \models \xi(x)$ then $\sigma(x) \models \langle\varepsilon\rangle 1\xi(x)$.

($\Leftarrow$) Assume now that there is a $\xi \in t^{-1}(\langle\varepsilon\rangle\psi)$ s.t. $\sigma(x) \models \xi(x)$ for all $x \in \mathrm{var}(t)$. Following Definition 23.6, we can distinguish three cases.

1. ξ is constructed in terms of a decomposition mapping $\eta \in \delta_t^{-1}(\psi)$ as in Definition 23.6a. This implies that $\sigma(x) \models \eta(x)$ for all $x \in \mathrm{var}(t)$ and thus, by induction we gather that $\delta_{\sigma(t)} \models \psi$. By Definition 14, we can therefore infer that $\sigma(t) \models \langle\varepsilon\rangle\psi$.

2. ξ is constructed in terms of a Γ-liquid argument x_0 of t, a Γ-patient ruloid $\frac{x_0 \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/x_0]}$ and of a decomposition mapping $\eta \in t[\mu/x_0]^{-1}(\psi^{\langle\varepsilon\rangle})$ as in Definition 23.6b. We define a substitution σ' s.t. the following properties hold:
   (a) $\sigma'(x) = \sigma(x)$ for all $x \in \mathrm{var}(t)$. As a consequence, from $\sigma(x) \models \xi(x)$ we derive $\sigma'(x) \models \xi(x)$.
   (b) As $\sigma'(x) \models \xi(x)$, by previous item (2a), we derive that if x occurs Γ-frozen in t then $\sigma'(x) \models \eta(x)$.
   (c) As $\sigma'(x_0) \models \xi(x_0)$, by previous item (2a), we derive that if $x = x_0$, and thus x occurs Γ-liquid in t, then $\sigma'(x) \models \langle\varepsilon\rangle\eta(\mu)$ This implies that there is a probability distribution π_x s.t. $P \vdash \sigma'(x) \xrightarrow{\hat{\varepsilon}} \pi_x$ and $\pi_x \models \eta(\mu)$. We define $\sigma'(\mu) = \pi_x$. Since P is Γ-patient, we obtain that $P \vdash \sigma'(x) \xrightarrow{\hat{\varepsilon}} \sigma'(\mu)$ and $\sigma'(\mu) \models \eta(\mu)$.
   (d) Since $\mathrm{var}(t[\mu/x_0]) = \mathrm{var}(t) \setminus \{x_0\} \cup \{\mu\}$, by previous items (2b) and (2c) we obtain $\sigma'(\zeta) \models \eta(\zeta)$ for each $\zeta \in \mathrm{var}(t[\mu/x_0])$. Hence, by induction, $\sigma'(t[\mu/x_0]) \models \psi$.

   From items (2a),(2c) we obtain that $P \vdash \sigma'(x_0) \xrightarrow{\tau} \sigma'(\mu)$. Hence by Theorem 2 we obtain that $P \vdash \sigma'(t) \xrightarrow{\tau} \sigma'(t[\mu/x_0])$. As P is Γ-patient we can infer that $P \vdash \sigma'(t) \xrightarrow{\hat{\varepsilon}} \sigma'(t[\mu/x_0])$ and since form item (2a) we have $\sigma'(t) = \sigma(t)$ and from item (2d) $\sigma'(t[\mu/x_0]) \models \psi^{\langle\varepsilon\rangle}$ we can conclude that $\sigma(t) \models \langle\varepsilon\rangle\psi$.

3. ξ is constructed in terms of a Γ-impatient P-ruloid $\frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and of a decomposition mapping $\eta \in \Theta^{-1}(\psi)$ as in Definition 23.5c. In this case the thesis follows by combining the arguments used in the proof of the inductive step $\phi = \langle\alpha\rangle\psi$ with the ones proposed in the previous item.

Hence, Equation (2) follows from the two implications.

<u>The case of multivariate terms</u>

Assume first that t is not univariate, namely $t = \varsigma(s)$ for some univariate s and non-injective substitution $\varsigma\colon \mathrm{var}(s) \to \mathbf{V}_s$. Then, $\sigma(\varsigma(s)) \models \varphi$ iff there exists a decomposition mapping $\xi' \in s^{-1}(\varphi)$ s.t. $\sigma(\varsigma(y)) \models \xi'(y)$, which by Definition 23.7 is equivalent to require that there exists a decomposition mapping $\xi' \in s^{-1}(\varphi)$ s.t. for each $x \in \mathrm{var}(t)$ we have $\sigma(x) \models \bigwedge_{y \in \varsigma^{-1}(x)} \xi'(y)$. By defining the decomposition mapping $\xi \in t^{-1}(\varphi)$ as $\xi(x) = \bigwedge_{y \in \varsigma^{-1}(x)} \xi'(y)$, we obtain the thesis.

The case for Θ not univariate, namely $\Theta = \varsigma(\Theta_1)$ for some univariate Θ_1 and non-injective substitution $\varsigma\colon \mathrm{var}(\Theta_1) \to \mathbf{V}_d \cup \delta_{\mathbf{V}_s}$, is analogous. □

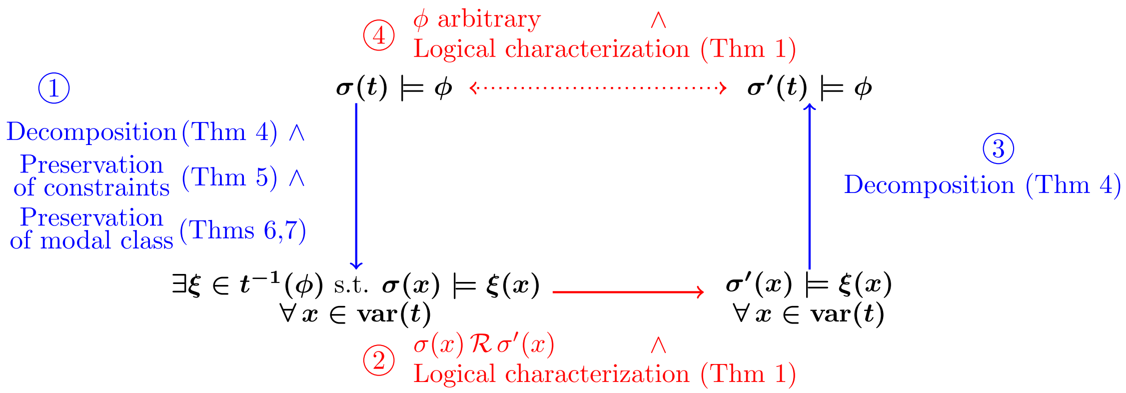


Figure 3: General schema to prove that $\sigma(x)\,\mathcal{R}\,\sigma'(x)$ for all $x \in \mathrm{var}(t)$ implies $\sigma(t)\,\mathcal{R}\,\sigma'(t)$, by combining a logical characterization of $\mathcal{R}$ with the related modal decomposition.

7. A congruence format for (rooted) branching bisimilarity

Since [70], a successful approach to study systematically a behavioural property of interest that should be satisfied by processes, is the structural analysis of SOS specification (see [1, 65] for surveys). In this approach one realizes that the property of interest depends on the patterns of the specifying SOS rules and proposes *syntactical constraints* on the form of these SOS rules, called a *format*, that ensure by construction the considered *behavioral property*. In particular, we are interested in the *congruence property* of behavioral equivalences, which is fundamental for compositional reasoning on systems.

Definition 24 (Congruence)**.** A behavioral relation $\mathcal{R}$ is a *congruence* for the operators in a signature Σ if for any term $t \in \mathbb{T}(\Sigma)$ and closed substitutions σ, σ' we have that

$$\text{whenever } \sigma(x)\,\mathcal{R}\,\sigma'(x) \text{ for all } x \in \mathrm{var}(t), \text{ then } \sigma(t)\,\mathcal{R}\,\sigma'(t).$$

Our aim is to derive a congruence format for (rooted) branching bisimilarity by exploiting the logical characterization given in Section 3 and the decomposition method given in Section 6. The underlying idea is that a congruence format for (rooted) branching bisimilarity must ensure that the formulae in the class characterizing it are always decomposed into formulae in the same class, so that by combining the logical characterizations of the equivalence (Theorem 1) and the decomposition theorem (Theorem 4) we can derive the congruence result, as sketched in Figure 3. Informally, given any term t and closed substitutions σ and σ', assume that $\sigma(x)\,\mathcal{R}\,\sigma'(x)$ for all x in $\mathrm{var}(t)$. By the characterization theorem, to conclude that $\sigma(t)\,\mathcal{R}\,\sigma'(t)$ it is enough to prove that for all formulae ϕ in the class characterizing $\mathcal{R}$ we have that $\sigma(t) \models \phi$ implies $\sigma'(t) \models \phi$ (step (4) in Figure 3), and viceversa. Assume then that $\sigma(t) \models \phi$ for an arbitrary formula ϕ in the class characterizing $\mathcal{R}$. The decomposition theorem ensures that there is a mapping $\xi \in t^{-1}(\phi)$ such that $\sigma(x) \models \xi(x)$ for all variables $x \in \mathrm{var}(t)$ (step (1) in Figure 3). If, moreover, the decomposition preserves the class characterizing $\mathcal{R}$, then all the decomposed formulae $\xi(x)$ are in that class. Therefore, by the hypothesis $\sigma(x)\,\mathcal{R}\,\sigma'(x)$ and the characterization theorem, we infer $\sigma'(x) \models \xi(x)$ for all x in t (step (2) in Figure 3). By the decomposition theorem we conclude $\sigma'(t) \models \phi$ (step (3) in Figure 3). Summarizing, we can conclude that $\sigma(t) \models \phi$ implies $\sigma'(t) \models \phi$.

7.1. The PRBB format

Firstly we need to identify the syntactic constraints characterizing our format. Interestingly, they coincide with those proposed in [31, 37] to obtain a format for (rooted) branching bisimilarity in the non-probabilistic setting. Clearly, the syntactical constraints for the probabilistic version of any equivalence cannot be less demanding that those for the classical non-probabilistic version, since any labelled transition system can be viewed as a particular PTS with only Dirac distributions. We can informally explain why the constraints in [31, 37] apply also to the probabilistic case.

First of all we notice that both the format in [37] and the PGSOS format do not allow *lookahead*, namely the ability to testing for two consecutive moves of a process. A consequence of this inability, is that in our case probability is never involved in the derivation of nondeterministic transitions. Equivalently, we are guaranteed that Σ-distribution rules (and ruloids) are never used to determine the provability of a closed literal. Moreover, the constraints on the probability weights in the definition of behavioral relations do not depend on the syntactical definition of processes and thus they are independent from the constraints of rule formats. We refer the interested reader to a comparison of the RBB safe format of [31] and the format in [61] as a further evidence of this fact. In addition, the syntactical representation of distributions through Σ-distribution rules together with the notion of $\_$-liquid contexts on them, further simplify the reasoning over PGSOS-rules. Hence, we can follow [37] and build our formats on the predicates $\aleph$ and Λ introduced in Section 2. Informally, these predicates allow us to properly control the weak behavior of processes, since Λ-liquid arguments are those that can satisfy the weak branching condition, the Λ-frozen ones are those for which we need the rootedness condition and $\aleph$-frozen arguments are those that cannot be tested.

We remark that although the source terms of a PGSOS rule are always univariate, we define the format for general multivariate terms. This generalization will simplify the reasoning when lifting the format to P-ruloids (Theorem 5 below) and, moreover, can be easily applied to the generalization of our decomposition method to the ntμfθ format discussed in upcoming Section 9.

Definition 25 (PRBB rule)**.** A PGSOS rule r is *probabilistic rooted branching bisimulation safe (PRBB)* with respect to predicates $\aleph$ and Λ if it satisfies the following conditions.

1. Right-hand sides of positive premises occur only Λ-liquid in target $\mathrm{trg}(r)$.
2. If x occurs only Λ-liquid in source $\mathrm{src}(r)$, then x occurs only Λ-liquid in the rule r.
3. If x occurs only $\aleph$-frozen in source $\mathrm{src}(r)$, then x does not occur in $\mathrm{prem}(r)$.
4. If x has exactly one $\aleph$-liquid occurrence in source $\mathrm{src}(r)$, which is also Λ-liquid, then x has at most one occurrence in $\mathrm{prem}(r)$, which must be in a positive premise. If moreover this premise is labeled τ, then r must be $\aleph \cap \Lambda$-patient.

Since the two definitions coincide, we refer the interested reader to [37, 61] for the counterexamples showing that the constraints in Definition 25 are necessary for the congruence result. Here we briefly give an intuition for them. In item 1, since a positive premise $x \xrightarrow{a} \mu$ clearly implies that all processes in the support of any closed instance of μ are running, it is required that μ appears in the target only at Λ-liquid positions, which are those dedicated to running processes. Item 2 guarantees that running processes, namely those marked by predicate Λ, maintain their mark and thus the possibility to execute. Item 3 prevents the testing of processes that cannot execute namely those that are at $\aleph$-frozen positions in the source. Finally, item 4 regulates the testing of processes at $(\aleph \cap \Lambda)$-liquid positions, namely those for which the rootedness property of $\approx_{\mathrm{rb}}$ does not need to hold. For an immediate intuition, consider processes $s = a.\delta_{\mathrm{nil}} \parallel_{\{c\}} b.\delta_{\mathrm{nil}}$ and $t = a.\delta_{\mathrm{nil}} \parallel_{\{c\}} \tau.\delta_{t_1}$, for $t_1 = b.\delta_{\mathrm{nil}}$. Clearly, $s \approx_{\mathrm{b}} t$. Since s can perform both a and b, whereas t cannot, two premises testing the ability to perform both a and b are able to discriminate them. The same happens with a premise testing for the inability to perform b or for the ability to perform τ. For these reasons, item 4 forbids double testing and negative premises for process arguments that may be branching but not rooted branching bisimilar and allows for these arguments the testing of τ-moves only in patience rules.

Notice that the first two constraints on the conservation of predicate Λ are the only ones that involve distribution terms. However, there is no need to translate these into constraints over Σ-distribution rules, as they are implicit in the definition of Γ-liquid contexts on distribution terms (Definition 19).

Definition 26 (PRBB format)**.** A PGSOS-TSS is in *probabilistic rooted branching bisimulation (PRBB) format* if, for some predicates $\aleph$ and Λ, it is $(\aleph \cap \Lambda)$-patient and it only contains PRBB rules. It is in *probabilistic branching bisimulation (PBB) format* if moreover the predicate Λ is universal.

7.2. Preservation of syntactic restrictions

Since the decomposition method in Definition 23 is not defined in terms of PGSOS rules but of P-ruloids, we need to guarantee that the syntactic constraints imposed by the PRBB format are preserved in the construction of ruloids from PGSOS rules fitting the format. Further, as differently from PGSOS rules a P-ruloid can have a negative literal as conclusion, we need to extend the notion of PRBB rule to non-standard rules. Briefly, no constraint on right-hand sides of premises is necessary and we need to drop also item 4 of Definition 25. This is because non-standard rules are built by denying premises of other rules and thus no further control on the occurrences of $(\aleph \cap \Lambda)$-liquid arguments can be guaranteed.

Definition 27 (PRBB non-standard rule)**.** A non-standard rule $\frac{\mathbf{H}}{t \xrightarrow{\alpha}\!\!\!\!\!/\;}$ is *probabilistic rooted branching bisimulation safe (PRBB)* with respect to predicates $\aleph$ and Λ if it satisfies the following conditions.

1. If x occurs only Λ-liquid in t, then x occurs only Λ-liquid in $\mathbf{H}$.
2. If x occurs only $\aleph$-frozen in t, then x does not occur in $\mathbf{H}$.

Theorem 5. *Let P be a PGSOS-PTSS in PRBB format with respect to predicates $\aleph$ and Λ. Then each P-ruloid is rooted branching bisimulation safe with respect to $\aleph$ and Λ.*

Proof. Assume that term t is the source of ruloid. We proceed by induction over the structure of t.

Base case $t = x$. Then the P-ruloids having t as source are of the form

$$\frac{t \xrightarrow{\alpha} \mu}{t \xrightarrow{\alpha} \mu} \quad \text{or} \quad \frac{t \xrightarrow{\alpha}\!\!\!\!\!/\;}{t \xrightarrow{\alpha}\!\!\!\!\!/\;}.$$

Both are PRBB safe wrt. $\aleph$ and Λ.

Inductive step $t = f(t_1, \ldots, t_{\mathfrak{n}})$. We distinguish two cases.

- Consider any P-ruloid

$$\rho = \frac{\mathbf{H}}{f(t_1, \ldots, t_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta}$$

Then there are a substitution σ and a PGSOS-rule

$$r = \frac{\{x_i \xrightarrow{\alpha_{i,m}} \mu_{i,m} \mid m \in M_i, i \in I\} \quad \{x_i \xrightarrow{\alpha_{i,n}}\!\!\!\!\!/\; \mid n \in N_i, i \in I\}}{f(x_1, \ldots, x_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta'}$$

with $\sigma(x_i) = t_i$ and $\sigma(\Theta') = \Theta$ s.t. $\mathbf{H} = \bigcup_{m \in M_i, i \in I} \mathbf{H}_{i,m} \cup \bigcup_{n \in N_i, i \in I} \mathbf{H}_{i,n}$ and, by induction over each $\sigma(x_i)$,

  – for each positive premise, the P-ruloid $\rho_{i,m} = \frac{\mathbf{H}_{i,m}}{\sigma(x_i) \xrightarrow{\alpha_{i,m}} \sigma(\mu_{i,m})}$ is a PRBB safe rule wrt. $\aleph$ and Λ, and

  – for each negative premise, the P-ruloid $\rho_{i,n} = \frac{\mathbf{H}_{i,n}}{\sigma(x_i) \xrightarrow{\alpha_{i,n}}\!\!\!\!\!/\;}$ is a PRBB safe non-standard rule wrt. $\aleph$ and Λ.

We prove that ρ satisfies the four conditions of Definition 25.

1. Let $\mu \in \text{rhs}(\mathbf{H})$. As $\text{rhs}(\mathbf{H})$ are all distinct by Definition 20, there is a particular pair of indexes $\tilde{i}, \tilde{m}$ s.t. $\mu \in \text{rhs}(\mathbf{H}_{\tilde{i},\tilde{m}})$ Since $\rho_{\tilde{i},\tilde{m}}$ is a PRBB safe rule wrt. $\aleph$ and Λ, by item 1 of Definition 25 we have that μ occurs only Λ-liquid in $\sigma(\mu_{\tilde{i},\tilde{m}})$. Since moreover, r is PRBB safe, the same condition gives that $\mu_{\tilde{i},\tilde{m}}$ occurs only Λ-liquid in Θ' and thus $\sigma(\mu_{\tilde{i},\tilde{m}})$ occurs only Λ-liquid in $\sigma(\Theta')$. As a consequence, μ occurs only Λ-liquid in Θ.

2. Assume that $x \in \mathrm{var}(t)$ occurs only Λ-liquid in t. Let $I(x) = \{i \in I \mid x \in \mathrm{var}(\sigma(x_i))\}$. As $t = \sigma(f(x_1, \dots, x_{\mathfrak{n}}))$, for each $i \in I(x)$ we have $\Lambda(f,i)$ and x occurs only Λ-liquid in $\sigma(x_i)$. Since r is PRBB safe wrt. $\aleph$ and Λ, by item 2 of Definition 25 for all $i \in I(x)$ it holds that x_i occurs only Λ-liquid in r. Hence x occurs only Λ-liquid in $\sigma(\mathrm{prem}(r))$. Since $\rho_{i,m}$ and $\rho_{i,n}$ are PRBB safe wrt. $\aleph$ and Λ for all $i \in I, m \in M_i, n \in N_i$, we can infer that x occurs only Λ-liquid in these ruloids, thus implying that x occurs only Λ-liquid in $\mathbf{H}$. Moreover, x occurs only Λ-liquid in $\sigma(\mu_{i,m})$ for all $i \in I, m \in M_i$ and by item 1 of Definition 25 $\mu_{i,m}$ occurs only Λ-liquid in Θ'. Furthermore, for each $i \in I(x)$, $\Lambda(f,i)$ implies that x_i occurs only Λ-liquid in Θ'. Therefore, we can infer that x occurs only Λ-liquid in Θ, and we can conclude that x occurs only Λ-liquid in ρ.
3. Assume that x occurs only $\aleph$-frozen in t. For each $h = 1, \dots, \mathfrak{n}$, we have that either $\neg\aleph(f,h)$ or x occurs only $\aleph$-frozen in $\sigma(x_h)$. Notice that in the first case, since r is PRBB safe wrt. $\aleph$ and Λ, by item 3 of Definition 25 x_h does not occur in $\mathrm{prem}(r)$, which clearly implies that also x is never tested. Hence, in both cases, since $\rho_{i,m}$ and $\rho_{i,n}$ are PRBB safe wrt. $\aleph$ and Λ for all $i \in I, m \in M_i, n \in N_i$, we can infer that x does not occur in the premises of these ruloids, thus implying that x does not occur in $\mathbf{H}$.
4. Assume that x has exactly one $\aleph$-liquid occurrence in t, which is also Λ-liquid. Then there is an $\tilde{i} \in \{1, \dots, \mathfrak{n}\}$ with $\aleph(f,\tilde{i})$ and $\Lambda(f,\tilde{i})$ s.t. x has exactly one $\aleph$-liquid occurrence in $\sigma(x_{\tilde{i}})$, which is also Λ-liquid. Moreover, for each $h \in \{1, \dots, \mathfrak{n}\} \setminus \{\tilde{i}\}$, either $\neg\aleph(f,h)$ or x occurs only $\aleph$-frozen in $\sigma(x_h)$. Since r is PRBB safe wrt. $\aleph$ and Λ, by item 3 of Definition 25 if $\neg\aleph(f,h)$ then x_h does not occur in $\mathrm{prem}(r)$. Further, by item 4 of Definition 25 $x_{\tilde{i}}$ has at most one occurrence in $\mathrm{prem}(r)$, which must be in a positive premise. By item 2 of Definition 25 this occurrence is also Λ-liquid. Therefore, we have that there is one $\tilde{m} \in M_{\tilde{i}}$ s.t. x has exactly one $\aleph$-liquid occurrence in $\mathbf{H}_{\tilde{i},\tilde{m}}$, which is also Λ-liquid. Since $\rho_{i,m}$ and $\rho_{i,n}$ are PRBB safe wrt. $\aleph$ and Λ for all $i \in I, m \in M_i, n \in N_i$, by item 4 of Definition 25 we can infer that such an occurrence of x in $\mathbf{H}_{\tilde{i},\tilde{m}}$ must be in the left-hand side of a positive premise. Thus, x has at most one occurrence in a premise of $\mathbf{H}$ which must be positive.
Assume now that this positive premise is labeled τ. Since $\rho_{\tilde{i},\tilde{m}}$ is PRBB safe wrt. $\aleph$ and Λ, by item 4 of Definition 25 it must be $\aleph\cap\Lambda$-patient. Hence, $\aleph\cap\Lambda(f,\tilde{i})$ gives that $x_{\tilde{i}}$ has an $\aleph\cap\Lambda$-liquid occurrence in $\mathrm{prem}(r)$ and thus, as r is PRBB safe wrt. $\aleph$ and Λ, by item 4 of Definition 25 r must be an $\aleph \cap \Lambda$-patience rule. Therefore, by construction of P-ruloids for P $\aleph \cap \Lambda$-patient, we can conclude that ρ is $\aleph \cap \Lambda$-patient.

- Consider any P-ruloid
$$\rho = \frac{\mathbf{H}}{f(t_1, \dots, t_{\mathfrak{n}}) \stackrel{\alpha}{\not\longrightarrow}}.$$
where $\mathbf{H} = \mathrm{opp}(\mathrm{pick}(\Re^P_\alpha))$ is built from the P-ruloids having $f(t_1, \dots, t_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta$ as conclusion, for some $\Theta \in \mathbb{DT}(\Sigma)$, as described in Definition20. In particular we note that the transformation operated by the mappings pick and opp does not affect the labeling of arguments of operators. Moreover, as we have shown is previous item, all P-ruloids having $f(t_1, \dots, t_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta$ as conclusion are PRBB safe wrt. $\aleph$ and Λ. We prove that ρ satisfies the two conditions of Definition 27.

1. Assume that $x \in \mathrm{var}(t)$ occurs only Λ-liquid in t. Let $I(x) = \{i \in I \mid x \in \mathrm{var}(\sigma(x_i))\}$. As $t = \sigma(f(x_1, \dots, x_{\mathfrak{n}}))$, for each $i \in I(x)$ we have $\Lambda(f,i)$ and x occurs only Λ-liquid in $\sigma(x_i)$. Since all P-ruloids having $f(t_1, \dots, t_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta$, for some $\Theta \in \mathbb{DT}(\Sigma)$, as conclusion are PRBB safe wrt. $\aleph$ and Λ, by item 2 of Definition 25 we can infer that x occurs only Λ-liquid in these ruloids, thus implying that x occurs only Λ-liquid in their premises and therefore in $\mathbf{H}$
2. Assume that x occurs only $\aleph$-frozen in t. For each $i = 1, \dots, \mathfrak{n}$, we have that either $\neg\aleph(f,i)$ or x occurs only $\aleph$-frozen in $\sigma(x_i)$. In both cases, since all P-ruloids having $f(t_1, \dots, t_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta$ as conclusion are PRBB safe wrt. $\aleph$ and Λ, by item 3 of Definition 25 we have that x does not occur in their premises. This implies that x does not occur in $\mathbf{H}$.

□

7.3. Preservation of logical characterizations

Our next task is to ensure that the decomposition of formulae in a chosen class preserves the syntactic restrictions of that class. In particular we are interested in showing that a formula in $\mathbb{L}_{\mathrm{rb}}$ (resp. $\mathbb{L}_{\mathrm{b}}$) is decomposed into formulae in $\mathbb{L}_{\mathrm{rb}}^{\equiv}$ (resp. $\mathbb{L}_{\mathrm{b}}^{\equiv}$). More precisely, in Theorem 7 we will show that given a term $t \in \mathbb{T}(\Sigma)$ (resp. distribution term $\Theta \in \mathbb{DT}(\Sigma)$), if $\varphi \in \mathbb{L}_{\mathrm{rb}}$ (resp. $\psi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{d}}$) then $\xi(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$ (resp. $\eta(\zeta) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$) for all $x \in \mathrm{var}(t)$ and $\xi \in t^{-1}(\varphi)$ (resp. $\zeta \in \mathrm{var}(\Theta)$ and $\eta \in \Theta^{-1}(\psi)$). Furthermore, in Theorem 6 we prove that for (distribution) variables occurring Λ-liquid in t (resp. Θ) formulae decomposed from formulae in $\mathbb{L}_{\mathrm{b}}$ are in $\mathbb{L}_{\mathrm{b}}^{\equiv}$.

Theorem 6. *Assume a PGSOS-TSS $P = (\mathbf{T}(\Sigma), \mathcal{A}, \rightarrow)$ in PRBB format with respect to predicates $\aleph$ and Λ. For any term t and variable x that occurs Λ-liquid in t*

$$\textit{whenever } \varphi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{s}} \textit{ then for all decomposition mapping } \xi \in t^{-1}(\varphi) \textit{ we have } \xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv} \tag{4}$$

and for any distribution term Θ and variable ζ that occurs Λ-liquid in Θ

$$\textit{whenever } \psi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{d}} \textit{ then for all decomposition mapping } \eta \in \Theta^{-1}(\psi) \textit{ we have } \eta(\zeta) \in \mathbb{L}_{\mathrm{b}}^{\equiv}\,. \tag{5}$$

Proof. We start with univariate terms $t \in \mathbb{T}(\Sigma)$ and $\Theta \in \mathbb{DT}(\Sigma)$. We proceed by induction on the structure of $\phi \in \mathbb{L}_{\mathrm{b}}$ to prove both statements at the same time.

- Base case $\phi = \top$. By Definition 23.1 $\xi(x) = \top \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. Moreover, we notice that by Lemma 4.1 we have that whenever $x \notin \mathrm{var}(t)$, then $\xi(x) \equiv \top \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. Hence, Equation (4) follows in this case.
- Inductive step $\phi = \neg\varphi$. By Definition 23.2 there is a function $\mathfrak{f}\colon t^{-1}(\varphi) \rightarrow \mathrm{var}(t)$ s.t.

$$\xi(x) = \bigwedge_{\xi' \in \mathfrak{f}^{-1}(x)} \neg\xi'(x).$$

  By structural induction $\xi'(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $\xi' \in \mathfrak{f}^{-1}(x)$. Therefore, $\neg\xi'(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $\xi' \in \mathfrak{f}^{-1}(x)$ and thus $\xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. Hence, Equation (4) follows in this case.
- Inductive step $\phi = \bigwedge_{j \in J} \varphi_j$. By Definition 23.3 $\xi(x) = \bigwedge_{j \in J} \xi_j(x)$ for some $\xi_j \in t^{-1}(\varphi_j)$ for all $j \in J$. By structural induction, $\xi_j(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $j \in J$, thus giving that $\xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. Hence, Equation (4) follows in this case.
- Inductive step $\phi = \bigoplus_{i \in I} r_i\varphi_i$. By Definition 23.8 there are a Σ-distribution ruloid $\rho^{\mathsf{D}} = \frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}}$ and a matching $\mathfrak{w} \in \mathfrak{W}(\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}, \bigoplus_{i \in I} r_i\varphi_i)$ s.t. for all $m \in M$ and $i \in I$ there is a decomposition mapping $\xi_{m,i}$ with $\begin{cases} \xi_{m,i} \in t_m^{-1}(\varphi_i) & \text{if } \mathfrak{w}(t_m, \varphi_i) > 0 \\ \xi_{m,i} \in t_m^{-1}(\top) & \text{otherwise} \end{cases}$ and we can distinguish two cases:

  1. $\zeta = \mu \in \mathbf{V}_d$. Then by Definition 23.8a we have

$$\eta(\mu) = \begin{cases} \displaystyle\bigoplus_{\mu \xrightarrow{q_j} x_j \in \mathbf{H}} q_j \bigwedge_{\substack{i \in I \\ m \in M}} \xi_{m,i}(x_j) & \text{if } \mu \in \mathrm{var}(\Theta) \\ 1\top & \text{otherwise} \end{cases}$$

     Clearly, $1\top \in \mathbb{L}_{\mathrm{b}}^{\equiv}$, hence assume that $\mu \in \mathrm{var}(\Theta)$. By Definition 21, the right-hand sides of ρ^{D} are pairwise distinct and thus we can infer that since μ occurs only Λ-liquid in Θ then for each $j \in J$, x_j occurs only Λ-liquid in $\mathrm{trg}(\rho^{\mathsf{D}})$. Thus, structural induction gives that $\xi_{m,i}(x_j) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $m \in M, i \in I, j \in J$. Hence $\bigwedge_{\substack{i \in I \\ m \in M}} \xi_{m,i}(x_j) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $j \in J$ from which we can conclude that $\eta(\mu) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

2. $\zeta = x \in \mathbf{V}_s$. Then by Definition 23.8b we have

$$\eta(x) = \begin{cases} \bigwedge_{\substack{i \in I \\ m \in M}} \xi_{m,i}(x) & \text{if } x \in \mathrm{var}(\Theta) \\ \top & \text{otherwise.} \end{cases}$$

Clearly, $\top \in \mathbb{L}_{\mathrm{b}}^{\equiv}$, hence assume that $x \in \mathrm{var}(\Theta)$. By Definition 21, the right-hand sides of ρ^{D} are pairwise distinct and thus we can infer that since x occurs only Λ-liquid in Θ then x occurs only Λ-liquid in $\mathrm{trg}(\rho^{\mathsf{D}})$. Thus, structural induction gives that $\xi_{m,i}(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $m \in M, i \in I$. Hence $\bigwedge_{\substack{i \in I \\ m \in M}} \xi_{m,i}(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$, namely $\eta(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

Hence, Equation (5) follows from both cases.

- Inductive step $\phi = \bigwedge_{j \in J} \psi_j$. By Definition 23.9, $\eta(\zeta) = \bigwedge_{j \in J} \eta_j(\zeta)$ for some $\eta_j \in \Theta^{-1}(\psi_j)$ for all $j \in J$. By structural induction, $\eta_j(\zeta) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $j \in J$, thus giving that $\eta(\zeta) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. Hence, Equation (5) follows in this case.

- Inductive step $\phi = \left(\bigoplus_{i \in I} r_i \varphi_i\right)^{\langle \varepsilon \rangle}$. By Definition 23.10, we can distinguish two cases:

1. Either there is a decomposition mapping $\eta' \in \Theta^{-1}(\bigoplus_{i \in I} r_i \varphi_i)$ such that $\eta(\zeta) = \eta'(\zeta)$ for all $\zeta \in \mathrm{var}(\Theta)$. Then the thesis follows as in the case of $\phi = \bigoplus_{i \in I} r_i \varphi_i$.
2. Or there are a Σ-distribution ruloid $\rho^{\mathsf{D}} = \frac{\mathbf{H}}{\{\Theta \xrightarrow{q_m} t_m \mid m \in M\}}$ and a matching $\mathfrak{w}$ for $\mathrm{conc}(\rho^{\mathsf{D}})$ and $\bigoplus_{i \in I} r_i \varphi_i$ s.t. for all $m \in M$ there is a decomposition mapping $\xi_m \in u_m^{-1}(\langle \varepsilon \rangle \psi_m)$, where $\psi_m = \bigoplus_{i \in I_m} \frac{\mathfrak{w}(u_m, \varphi_i)}{q_m} \varphi_i$, with $I_m = \{i \in I \mid \mathfrak{w}(u_m, \varphi_i) > 0\}$. We can distinguish two cases:

   (a) $\zeta = \mu \in \mathbf{V}_d$. Then by Definition 23.10(b)i we have

$$\eta(\mu) = \begin{cases} \bigoplus_{\mu \xrightarrow{q_j} x_j \in \mathbf{H}} q_j \bigwedge_{m \in M} \xi_m(x_j) & \text{if } \mu \in \mathrm{var}(\Theta) \\ 1\top & \text{otherwise} \end{cases}$$

   Clearly, $1\top \in \mathbb{L}_{\mathrm{b}}^{\equiv}$, hence assume that $\mu \in \mathrm{var}(\Theta)$. By Definition 21, the right-hand sides of ρ^{D} are pairwise distinct and thus we can infer that since μ occurs only Λ-liquid in Θ then for each $j \in J$, x_j occurs only Λ-liquid in $\mathrm{trg}(\rho^{\mathsf{D}})$. Thus, structural induction gives that $\xi_m(x_j) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $m \in M, j \in J$. Hence $\bigwedge_{m \in M} \xi_m(x_j) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $j \in J$ from which we can conclude that $\eta(\mu) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

   (b) $\zeta = x \in \mathbf{V}_s$. Then by Definition 23.10(b)ii we have

$$\eta(x) = \begin{cases} \bigwedge_{m \in M} \xi_m(x) & \text{if } x \in \mathrm{var}(\Theta) \\ \top & \text{otherwise.} \end{cases}$$

   Clearly, $\top \in \mathbb{L}_{\mathrm{b}}^{\equiv}$, hence assume that $x \in \mathrm{var}(\Theta)$. By Definition 21, the right-hand sides of ρ^{D} are pairwise distinct and thus we can infer that since x occurs only Λ-liquid in Θ then x occurs only Λ-liquid in $\mathrm{trg}(\rho^{\mathsf{D}})$. Thus, structural induction gives that $\xi_m(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $m \in M$. Hence $\bigwedge_{m \in M} \xi_{m,i}(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$, namely $\eta(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

   Hence, Equation (5) follows from both cases.

- Inductive step $\phi = \langle \varepsilon \rangle 1(\varphi \wedge \langle a \rangle \psi)$. We can distinguish three cases:

1. $\xi(x)$ is defined on the basis of Definition 23.6a. Then there is a $\eta' \in \delta_t^{-1}(1(\varphi \wedge \langle a \rangle \psi))$ s.t. $\xi(x) = \langle \varepsilon \rangle \eta'(x)$ if x occurs $\aleph \cap \Lambda$-liquid in t and $\xi(x) = \eta'(x)$ otherwise. As we are considering δ_t as distribution term, we have that there is a $\xi' \in t^{-1}(\varphi \wedge \langle a \rangle \psi)$ s.t. $\eta'(x) = \xi'(x)$. In particular, we have that by Definition 23.3 $\xi'(x) = \xi_1(x) \wedge \xi_2(x)$ with $\xi_1 \in t^{-1}(\varphi)$ and $\xi_2 \in t^{-1}(\langle a \rangle \psi)$. By structural induction we can immediately conclude that $\xi_1(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. Then, by Definition 23.4 there are a P ruloid $\rho = \frac{\mathbf{H}}{t \xrightarrow{a} \Theta}$ and a $\eta \in \Theta^{-1}(\psi)$ s.t.

$$\xi_2(x) = \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle \beta \rangle \eta(\mu) \quad \wedge \quad \bigwedge_{x \xrightarrow{\gamma}\!\!\!\!\not\;\; \in \mathbf{H}} \neg \langle \gamma \rangle \top \quad \wedge \quad \eta(x).$$

By Theorem 5, ρ is PRBB safe wrt. $\aleph$ and Λ and therefore
   - rhs($\mathbf{H}$) occur only Λ-liquid in Θ (Definition 25.1), and
   - x occurs only Λ-liquid in Θ (Definition 25.2).

   Thus, by structural induction over ψ we get that $\eta(\zeta) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for any $\zeta \in \{x\} \cup \mathrm{rhs}(\mathbf{H})$. We distinguish two cases:

   (a) The occurrence of x in t is $\aleph$-liquid. Then $\xi(x) \equiv \langle \varepsilon \rangle 1(\xi_1(x) \wedge \xi_2(x))$. The P-ruloid ρ is PRBB safe wrt. $\aleph$ and Λ and thus (by Definition 25.4) x is the left-hand side of at most one premise in $\mathbf{H}$, which must be positive. Hence, either $\xi_2(x) = \eta(x)$ or $\xi_2(x) = \eta(x) \wedge \langle \beta \rangle \eta(\mu)$ for $x \xrightarrow{\beta} \mu \in \mathbf{H}$. As $a \neq \tau$ we get that $\beta \neq \tau$ (Definition 25.4). Therefore,

$$\text{either } \xi(x) \equiv \langle \varepsilon \rangle 1\big(\xi_1(x) \wedge \eta(x)\big)$$
$$\text{or } \xi(x) \equiv \langle \varepsilon \rangle 1\big((\xi_1(x) \wedge \eta(x)) \wedge \langle \beta \rangle \eta(\mu)\big)$$

   we can conclude that $\xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

   (b) The occurrence of x in t is $\aleph$-frozen. Then $\xi(x) = \xi_1(x) \wedge \xi_2(x)$. The P-ruloid ρ is PRBB safe wrt. $\aleph$ and Λ and thus (by Definition 25.3) x does not occur in $\mathbf{H}$, thus implying that $\xi_2(x) = \eta(x)$. We can therefore conclude that $\xi(x) = \xi_1(x) \wedge \eta(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

2. $\xi(x)$ is defined on the basis of Definition 23.6b. Then there are a Γ-patient ruloid $\rho = \frac{x \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/x]}$ and an $\eta \in t[\mu/x]^{-1}\left((1(\varphi \wedge \langle a \rangle \psi))^{\langle \varepsilon \rangle}\right)$ s.t.

$$\xi(x) = \begin{cases} \langle \varepsilon \rangle \eta(\mu) & \text{if } x \text{ occurs } \aleph \cap \Lambda\text{-liquid in } t \\ \eta(x) & \text{otherwise.} \end{cases}$$

   The case for x occurring $\aleph$-frozen in t follows by the same arguments used in the proof of previous item. Hence, let us consider the case of x occurring $\aleph$-liquid in t. By Theorem 5, ρ is PRBB safe wrt. $\aleph$ and Λ and therefore μ occur only Λ-liquid in $t[\mu/x]$ (Definition 25.1), and thus it is $\aleph \cap \Gamma$-liquid in $t[\mu/x]$. By Definition 23.10(b)i, considering that the distribution formula assigns weight 1 to the state formula $\varphi \wedge \langle a \rangle \psi$, we have that

$$\eta(\mu) = \bigoplus_{\mu \xrightarrow{q_j} x_j \in \mathbf{H}} q_j \xi'(x_j)$$

   for a distribution ruloid

$$\rho^{\mathsf{D}} = \frac{\{\mu \xrightarrow{q_j} x_j\} \cup \bigcup_{x' \in \mathrm{var}(t) \setminus \{x\}} \{\delta'_x \xrightarrow{1} x'\}}{\{t[\mu/x] \xrightarrow{q_j} t[x_j/x] \mid j \in J\}}$$

and decomposition mappings $\xi' \in t^{-1}(\langle\varepsilon\rangle 1(\varphi \wedge \langle a\rangle\psi))$ for each . In particular we have that $\eta(\mu) \equiv 1\xi'(x)$. Therefore, we can proceed as in the previous item for x occurring $\aleph \cap \Gamma$-liquid in t to prove

$$\text{either } \xi(x) \equiv \langle\varepsilon\rangle 1\big(\xi_1(x) \wedge \eta(x)\big)$$
$$\text{or } \xi(x) \equiv \langle\varepsilon\rangle 1\big((\xi_1(x) \wedge \eta(x)) \wedge \langle\beta\rangle\eta(\mu)\big)$$

and thus concluding that $\xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

3. $\xi(x)$ is defined on the basis of Definition 23.6c, namely in terms of a Γ-impatient ruloid $\rho = \frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and an $\eta \in \Theta^{-1}((1(\varphi \wedge \langle a\rangle\psi))^{\langle\varepsilon\rangle})$. By Theorem 5, ρ is PRBB safe wrt. $\aleph$ and Λ and therefore
   - rhs($\mathbf{H}$) occur only Λ-liquid in Θ (Definition 25.1), and
   - x occurs only Λ-liquid in Θ (Definition 25.2).

   Since moreover $\phi \in \mathbb{L}_{\mathrm{b}}$, by structural induction we can immediately infer that $\eta(\zeta) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $\zeta \in \{x\} \cup \mathrm{rhs}(\mathbf{H})$. We can distinguish two cases:

   (a) The occurrence of x in t is $\aleph$-liquid. Then

   $$\xi(x) = \langle\varepsilon\rangle 1\Big(\bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \quad\wedge \bigwedge_{x \stackrel{\gamma}{\not\rightarrow} \in \mathbf{H}} \neg\langle\gamma\rangle\top \quad\wedge\quad \eta(x)\Big).$$

   By Definition 25.4, $\mathbf{H}$ has at most one premise of the form $x \xrightarrow{\beta} \mu$, for which moreover $\beta \neq \tau$, and it has no negative premises having x has left-hand side. Thus

   $$\text{either } \xi(x) = \langle\varepsilon\rangle 1\eta(x)$$
   $$\text{or } \xi(x) = \langle\varepsilon\rangle 1\big(\eta(x) \wedge \langle\beta\rangle\eta(\mu)\big)$$

   In the first case there is nothing more to prove. In the second case we have that both $\eta(x), \eta(\mu) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ and thus we can conclude that $\xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

   (b) The occurrence of x in t is $\aleph$-frozen. Then

   $$\xi(x) = \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \quad\wedge \bigwedge_{x \stackrel{\gamma}{\not\rightarrow} \in \mathbf{H}} \neg\langle\gamma\rangle\top \quad\wedge\quad \eta(x).$$

   By Definition 25.3, x does not occur in $\mathbf{H}$ and thus $\xi(x) = \eta(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

Hence, Equation (4) follows in this case.

- Inductive step $\phi = \langle\varepsilon\rangle 1(\varphi \wedge \langle\hat{\tau}\rangle\psi)$. This case is analogous to the previous one for $\phi = \langle\varepsilon\rangle 1(\varphi \wedge \langle a\rangle\psi)$. Hence, Equation (4) follows also in this case.

Finally, let us deal with multivariate terms.

Assume first that t is not univariate, namely $t = \sigma(s)$ for some univariate term s and non-injective substitution $\sigma\colon \mathrm{var}(s) \to \mathbf{V}_s$. Then, by Definition 23.7 each $\xi \in t^{-1}(\varphi)$, for some $\varphi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{s}}$, is built in terms of a $\xi' \in s^{-1}(\varphi)$ s.t. $\xi(x) = \bigwedge_{y \in \sigma^{-1}(x)} \xi'(x)$. As s is univariate and for each $y \in \sigma^{-1}(x)$ the occurrence of y in s is Λ-liquid we can infer that $\xi'(y) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ for all $y \in \sigma^{-1}(x)$. Therefore, we can conclude that $\xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. Hence, Equation (4) follows also in this case.

The case for Θ not univariate, namely $\Theta = \sigma(\theta)$ for some univariate distribution term θ and a non-injective substitution $\sigma\colon \mathrm{var}(\theta) \to \mathbf{V}$, is analogous. □

Theorem 7. *Assume a PGSOS-TSS $P = (\mathbf{T}(\Sigma), \mathcal{A}, \rightarrow)$ in PRBB format with respect to predicates $\aleph$ and Λ. For any term t and variable x*

$$\textit{whenever } \varphi \in \mathbb{L}_{\mathrm{rb}} \textit{ then for all decomposition mapping } \xi \in t^{-1}(\varphi) \textit{ we have } \xi(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv} \qquad (6)$$

and for any distribution term Θ *and variable* ζ

$$\text{whenever } \psi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{d}} \text{ then for all decomposition mapping } \eta \in \Theta^{-1}(\psi) \text{ we have } \eta(\zeta) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}. \tag{7}$$

Proof. We start with univariate terms $t \in \mathbb{T}(\Sigma)$ and $\Theta \in \mathbb{DT}(\Sigma)$. We proceed by induction on the structure of $\phi \in \mathbb{L}_{\mathrm{rb}}$ to prove both statements at the same time.

- The proof for the base case $\phi = \top$ and for the inductive steps $\phi = \neg\varphi$, $\phi = \bigwedge_{j\in J}\varphi_j$, $\phi = \bigoplus_{i\in I} r_i\varphi_i$ and $\phi = \bigwedge_{j\in J}\psi_j$ follow as in the analogous cases of the proof of Theorem 6.
- Inductive step $\phi = \langle\alpha\rangle\psi$. Notice that since $\mathbb{L}_{\mathrm{b}} \subseteq \mathbb{L}_{\mathrm{rb}}$, we can apply induction over ψ. By Definition 23.4 there are a P-ruloid $\rho = \frac{\mathbf{H}}{t \xrightarrow{\alpha} \Theta}$ and a decomposition mapping $\eta \in \Theta^{-1}(\psi)$ s.t.

$$\xi(x) = \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \wedge \bigwedge_{x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}} \neg\langle\gamma\rangle\top \wedge \eta(x)$$

  By structural induction we have that $\eta(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$. Moreover, by Theorem 5 ρ is PRBB safe and thus, by Definition 25.1, we have that all variables in rhs($\mathbf{H}$) occur only Λ-liquid in Θ. Hence, by Theorem 6 we obtain that $\eta(\mu) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$. This implies that $\langle\beta\rangle\eta(\mu) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$ for all $x \xrightarrow{\beta} \mu \in \mathbf{H}$. Further, we have that $\neg\langle\gamma\rangle\top \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$ for all $x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}$ directly by definition. Therefore, we can conclude that $\xi(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$. Hence, Equation (6) follows also in this case.
- Inductive step $\phi \in \mathbb{L}_{\mathrm{b}}$. The cases $\phi = \top$, $\phi = \neg\varphi$, $\phi = \bigwedge_{j\in J}\varphi_j$, $\phi = \bigoplus_{i\in I} r_i\varphi_i$ and $\phi = \bigwedge_{j\in J}\psi_j$ follow as in the analogous cases of the proof of Theorem 6. Moreover, if the occurrence of x in t is Λ-liquid, then $\xi(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$ follows from Theorem 6. Hence, assume that the occurrence of x in t is Λ-frozen.
  - $\phi = \langle\varepsilon\rangle 1\big(\varphi \wedge \langle a\rangle\psi\big)$. We can distinguish three cases.
    1. $\xi(x)$ is defined on the basis of Definition 23.6a. Since x occurs Λ-frozen in t then there is a $\eta' \in \delta_t^{-1}(1(\varphi \wedge \langle a\rangle\psi))$ s.t. $\xi(x) = \eta'(x)$. As we are considering δ_t as distribution term, we have that there is a $\xi' \in t^{-1}(\varphi \wedge \langle a\rangle\psi)$ s.t. $\eta'(x) = \xi'(x)$. In particular, we have that by Definition 23.3 $\xi'(x) = \xi_1(x) \wedge \xi_2(x)$ with $\xi_1 \in t^{-1}(\varphi)$ and $\xi_2 \in t^{-1}(\langle a\rangle\psi)$. Since, $\mathbb{L}_{\mathrm{b}}^{\mathrm{s}} \subseteq \mathbb{L}_{\mathrm{rb}}$ and $\langle \alpha\rangle\psi \in \mathbb{L}_{\mathrm{rb}}$, by structural induction we can immediately conclude that $\xi_1(x), \xi_2(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$. Therefore, we can conclude that $\xi(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$.
    2. $\xi(x)$ is defined on the basis of Definition 23.6b. Then there are a Γ-patient ruloid $\rho = \frac{x \xrightarrow{\tau} \mu}{t \xrightarrow{\tau} t[\mu/x]}$ and an $\eta \in t[\mu/x]^{-1}\left((1(\varphi \wedge \langle a\rangle\psi))^{\langle\varepsilon\rangle}\right)$ s.t.

$$\xi(x) = \begin{cases} \langle\varepsilon\rangle\eta(\mu) & \text{if } x \text{ occurs } \aleph \cap \Lambda\text{-liquid in } t \\ \eta(x) & \text{otherwise.} \end{cases}$$

       Since x occurs Λ-frozen in t the proof follows by the same arguments used in the proof of previous item.
    3. $\xi(x)$ is defined on the basis of Definition 23.6c, namely in terms of a Γ-impatient ruloid $\rho = \frac{\mathbf{H}}{t \xrightarrow{\tau} \Theta}$ and an $\eta \in \Theta^{-1}\left((1(\varphi \wedge \langle a\rangle\psi))^{\langle\varepsilon\rangle}\right)$. By Theorem 5, ρ is PRBB safe wrt. $\aleph$ and Λ and therefore
       - rhs($\mathbf{H}$) occur only Λ-liquid in Θ (Definition 25.1), and
       - Since x occurs Λ-frozen in t then x does not occur in $\mathbf{H}$ (Definition 25.3).

       Since $\phi \in \mathbb{L}_{\mathrm{b}}$, by structural induction we can immediately infer that $\eta(\zeta) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$ for all $\zeta \in \{x\} \cup \mathrm{rhs}(\mathbf{H})$. Moreover, as x occurs Λ-frozen in t we have

$$\xi(x) = \bigwedge_{x \xrightarrow{\beta} \mu \in \mathbf{H}} \langle\beta\rangle\eta(\mu) \quad\wedge \bigwedge_{x \stackrel{\gamma}{\not\longrightarrow} \in \mathbf{H}} \neg\langle\gamma\rangle\top \quad\wedge\quad \eta(x)$$

       and since x does not occur in $\mathbf{H}$, we can conclude that $\xi(x) = \eta(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$.

– $\phi = \langle\varepsilon\rangle 1(\varphi \wedge \langle\hat{\tau}\rangle\psi)$. This case is analogous to the previous one for $\phi = \langle\varepsilon\rangle 1(\varphi \wedge \langle a\rangle\psi)$.

Hence, Equation (6) follows also in this case.

The proof for multivariate terms follows by the same arguments used in the proof of Theorem 6. □

7.4. The congruence theorems

We can now exploit the schema in Figure 3 to derive the congruence results. Firstly, from Theorem 1.1, Theorem 4 and Theorem 6 we get that $\approx_{\mathrm{b}}$ is a congruence for all operators defined in a PGSOS-PTSS in PBB format.

Theorem 8. *Let P be a complete PGSOS-PTSS in branching bismulation format. Then $\approx_{\mathrm{b}}$ is a congruence for all operators defined by P.*

Proof. Let $t \in \mathbb{T}(\Sigma)$ and let σ, σ' be two closed substitutions. We aim to show that

$$\text{whenever } \sigma(x) \approx_{\mathrm{b}} \sigma'(x) \text{ for each } x \in \mathrm{var}(t) \text{ then } \sigma(t) \approx_{\mathrm{b}} \sigma'(t). \tag{8}$$

Considering the characterization result of $\mathbb{L}_{\mathrm{b}}$ for probabilistic branching bisimilarity (Theorem 1.1), to prove Equation (8) we simply have to show that $\sigma(t)$ and $\sigma'(t)$ satisfy the same formulae in $\mathbb{L}_{\mathrm{b}}$. By Definition 26 each rule in P is PRBB safe wrt. some predicate $\aleph$ and a universal predicate Λ and P is $\aleph \cap \Lambda$-patient. Assume that $\sigma(t) \models \varphi$, for some state formula $\varphi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{s}}$. By Theorem 4, (taking $\Gamma = \aleph \cap \Lambda$) there is a decomposition mapping $\xi \in t^{-1}(\varphi)$ s.t. $\sigma(x) \models \xi(x)$ for each $x \in \mathrm{var}(t)$. Since Λ is universal, we have that each $x \in \mathrm{var}(t)$ occurs Λ-liquid in t and thus by Theorem 6 we gather that $\xi(x) \in \mathbb{L}_{\mathrm{b}}^{\equiv}$ and moreover by Thm 1.1 from $\sigma(x) \approx_{\mathrm{b}} \sigma'(x)$ we obtain that $\sigma'(x) \models \xi(x)$ for each $x \in \mathrm{var}(t)$. By applying Theorem 4 once again, we obtain that $\sigma'(t) \models \varphi$, thus proving Equation (8). □

Similarly, from Theorem 1.2, Theorem 4 and Theorem 7 we obtain that $\approx_{\mathrm{rb}}$ is a congruence for all operators defined in a PGSOS-PTSS in PRBB format.

Theorem 9. *Let P be a complete PGSOS-PTSS in rooted branching bismulation format. Then $\approx_{\mathrm{rb}}$ is a congruence for all operators defined by P.*

Proof. Let $t \in \mathbb{T}(\Sigma)$ and let σ, σ' be two closed substitutions. We aim to show that

$$\text{whenever } \sigma(x) \approx_{\mathrm{rb}} \sigma'(x) \text{ for each } x \in \mathrm{var}(t) \text{ then } \sigma(t) \approx_{\mathrm{rb}} \sigma'(t). \tag{9}$$

Considering the characterization result of $\mathbb{L}_{\mathrm{rb}}$ for probabilistic rooted branching bisimilarity (Theorem 1.2), to prove Equation (9) we simply have to show that $\sigma(t)$ and $\sigma'(t)$ satisfy the same formulae in $\mathbb{L}_{\mathrm{rb}}$. By Definition 26 each rule in P is PRBB safe wrt. some predicates $\aleph$ and Λ and P is $\aleph \cap \Lambda$-patient. Assume that $\sigma(t) \models \varphi$, for some state formula $\varphi \in \mathbb{L}_{\mathrm{rb}}$. By Theorem 4, (taking $\Gamma = \aleph \cap \Lambda$) there is a decomposition mapping $\xi \in t^{-1}(\varphi)$ s.t. $\sigma(x) \models \xi(x)$ for each $x \in \mathrm{var}(t)$. By Theorem 7 we gather that $\xi(x) \in \mathbb{L}_{\mathrm{rb}}^{\equiv}$ and moreover by Thm 1.2 from $\sigma(x) \approx_{\mathrm{rb}} \sigma'(x)$ we obtain that $\sigma'(x) \models \xi(x)$ for each $x \in \mathrm{var}(t)$. By applying Theorem 4 once again, we obtain that $\sigma'(t) \models \varphi$, thus proving Equation (9). □

8. Application

In general, an SOS specification format guaranteeing a given property is relevant only if it is not too demanding, namely if it captures an interesting bulk of operators. In this section, we consider the operators of the probabilistic process algebra P_{PA} from [40] and we apply to them our congruence formats. P_{PA} extends the *basic process algebra* $BPA_{\varepsilon,\tau}$ [6] with probabilistic operators from (probabilistic) CCS [54] and (probabilistic) CSP [19, 26, 54] as well as with recursion [7, 64]. More precisely, P_{PA} is obtained from the disjoint extension of the four PTSSs represented in Table 1 and Tables 3–5 below, where we let α range over $\mathcal{A}_\tau \cup \{\surd\}$, for $\surd$ the special action from BPA denoting successful termination. We will show that by assigning a proper liquid/frozen labeling, with respect to predicates Λ and $\aleph$, of arguments of operators

$$\frac{}{\varepsilon \xrightarrow{\surd} \delta_{\mathrm{nil}}} \qquad \frac{}{\alpha.\bigoplus_{i=1}^{n}[p_i]x_i \xrightarrow{\alpha} \sum_{i=1}^{n} p_i \delta_{x_i}} \qquad \frac{x \xrightarrow{\alpha} \mu \quad a \neq \surd}{x;y \xrightarrow{\alpha} \mu;\delta_y} \qquad \frac{x \xrightarrow{\surd} \mu \quad y \xrightarrow{a} \nu}{x;y \xrightarrow{\alpha} \nu} \qquad \frac{x \xrightarrow{\alpha} \mu}{x+y \xrightarrow{\alpha} \mu} \qquad \frac{y \xrightarrow{\alpha} \nu}{x+y \xrightarrow{\alpha} \nu}$$

$$\frac{x \xrightarrow{\alpha} \mu \quad \alpha \neq \surd}{x \parallel y \xrightarrow{\alpha} \mu \parallel \delta_y} \qquad \frac{y \xrightarrow{\alpha} \nu \quad \alpha \neq \surd}{x \parallel y \xrightarrow{\alpha} \delta_x \parallel \nu} \qquad \frac{x \xrightarrow{\surd} \mu \quad y \xrightarrow{\surd} \nu}{x \parallel y \xrightarrow{\surd} \delta_{\mathrm{nil}}}$$

$$\frac{x \xrightarrow{\tau} \mu}{x \mid y \xrightarrow{\tau} \mu \mid \delta_y} \qquad \frac{y \xrightarrow{\tau} \nu}{x \mid y \xrightarrow{\tau} \delta_x \mid \nu} \qquad \frac{x \xrightarrow{\alpha} \mu \quad y \xrightarrow{\alpha} \nu \quad \alpha \neq \tau, \surd}{x \mid y \xrightarrow{\alpha} \mu \mid \nu} \qquad \frac{x \xrightarrow{\surd} \mu \quad y \xrightarrow{\surd} \nu}{x \mid y \xrightarrow{\surd} \delta_{\mathrm{nil}}}$$

$$\frac{x \xrightarrow{\alpha} \mu \quad \alpha \notin B \cup \{\surd\}}{x \parallel_B y \xrightarrow{\alpha} \mu \parallel_B \delta_y} \qquad \frac{y \xrightarrow{\alpha} \nu \quad \alpha \notin B \cup \{\surd\}}{x \parallel_B y \xrightarrow{\alpha} \delta_x \parallel_B \nu} \qquad \frac{x \xrightarrow{\alpha} \mu \quad y \xrightarrow{\alpha} \nu \quad \alpha \in B \setminus \{\surd\}}{x \parallel_B y \xrightarrow{\alpha} \mu \parallel_B \nu} \qquad \frac{x \xrightarrow{\surd} \mu \quad y \xrightarrow{\surd} \nu}{x \parallel_B y \xrightarrow{\surd} \delta_{\mathrm{nil}}}$$

Table 3: Standard non-recursive operators.

in P_{PA} specified in Table 1 and Tables 3–5 our formats provide congruence results for probabilistic rooted branching bisimilarity (Corollary 1–2) and also for probabilistic branching bisimilarity (Corollary 1). We also show that the operator in Table 5 is outside of our format.

First of all, notice that the PTSSs in Tables 3–5 are in PGSOS-format.

In Table 3 we have the PGSOS rules specifying standard non-recursive operators. The probabilistic prefix operator expresses that the process $\alpha.\bigoplus_{i=1}^{n}[p_i]t_i$ can perform action α and evolves to process t_i with probability p_i. We mark all arguments of this operator as $(\aleph \cap \Lambda)$-frozen) since each process t_i has not yet started to run and cannot start to run. The constraints in Definition 25 are trivially satisfied. The sequential composition $t;t'$ and the alternative composition $t+t'$ are as usual and so is the labeling of their arguments: the first argument of sequential composition is Λ-liquid, whereas its second argument and both arguments of alternative composition are Λ-frozen, exactly as those for its probabilistic variant already discussed in Example 2. In fact, the first argument of sequential composition can be a process that has already started its execution, whereas the second argument has to wait for the first one to end its execution before starting its own. Notice that the two rules for sequential composition satisfy constraints 1, 2 of Definition 25 since x, μ, ν occur only Λ-liquid in them. Moreover, we set both arguments of both operators to be $\aleph$-liquid as they can all start their execution. Thus, the rules for both operators trivially satisfy constraint 3 of Definition 25. Finally, we need to check whether the two rules for sequential composition satisfy constraint 4 of the same Definition with respect to the first argument which is $(\aleph \cap \Lambda)$-liquid. We have that in both rules x occurs as left-hand side of a single positive premise. Moreover, in the first rule, if $\alpha = \tau$ then the rule becomes a patient-rule for the operator.

In Example 2 we already marked both arguments of CSP-like parallel composition operator $\parallel_B$ as $(\aleph \cap \Lambda)$-liquid. The same marking holds for interleaving operator $\parallel$ and the synchronous parallel composition operator $\mid$, which are special cases of $\parallel_B$, with $B = \emptyset$ and $B = \mathcal{A}$, respectively. Constraints 1 and 3 of Definition 25 are then easy to check for all these three operators. As regards constraint 2, we notice that, by Definition 19, x (resp. y) occur always Λ-liquid in δ_x (resp. δ_y) which in turns occurs at a Λ-liquid position in the targets of the considered rules. Finally, the same arguments used to verify constraint 4 of Definition 25 for the first argument of sequential composition hold also in this case.

Notice that if we exclude the rules for sequential and alternative composition from Table 3 and those for probabilistic alternative composition in Table 1 then predicate Λ is universal on the remaining rules in the two Tables. Also notice that the PTSSs in Tables 1 and 3 are $\aleph \cap \Lambda$-patient.

Corollary 1. *Probabilistic rooted branching bisimilarity is a congruence for all operators specified in Table 1 and Table 3. Moreover, probabilistic branching bisimilarity is a congruence for prefixing and all parallel composition operators specified in Table 1 and Table 3.*

Table 4 presents the PGSOS rules specifying (probabilistic) recursive operators. The infinite iteration t^ω of process t expresses that t is performed infinitely many times in a row. The binary Kleene-star expresses

$$\frac{x \xrightarrow{\alpha} \mu \quad a \neq \surd}{x^{\omega} \xrightarrow{\alpha} \mu; \delta_{x^{\omega}}} \qquad \frac{x \xrightarrow{\alpha} \mu \quad a \neq \surd}{x^{*}y \xrightarrow{\alpha} \mu; \delta_{x^{*}y}} \qquad \frac{y \xrightarrow{\alpha} \nu}{x^{*}y \xrightarrow{\alpha} \nu} \qquad \frac{x \xrightarrow{\alpha} \mu \quad \alpha \neq \surd}{!x \xrightarrow{\alpha} \mu \parallel \delta_{!x}} \qquad \frac{x \xrightarrow{\alpha} \mu \quad \alpha \neq \surd}{!_p x \xrightarrow{\alpha} \mu \oplus_p (\mu \parallel \delta_{!_p x})}$$

$$\frac{x \xrightarrow{\alpha} \mu \quad y \xrightarrow{\alpha} \nu \quad \alpha \neq \surd}{x^{*_p}y \xrightarrow{\alpha} \nu \oplus_p \mu; \delta_{x^{*_p}y}} \qquad \frac{x \xrightarrow{\alpha} \mu \quad y \stackrel{\alpha}{\not\longrightarrow} \quad \alpha \neq \surd}{x^{*_p}y \xrightarrow{\alpha} \mu; \delta_{x^{*_p}y}} \qquad \frac{x \stackrel{\alpha}{\not\longrightarrow} \quad y \xrightarrow{\alpha} \nu \quad \alpha \neq \surd}{x^{*_p}y \xrightarrow{\alpha} \nu} \qquad \frac{y \xrightarrow{\surd} \nu}{x^{*_p}y \xrightarrow{\surd} \nu}$$

Table 4: Recursive operators.

$$\frac{x \xrightarrow{\tau} \mu}{x \parallel_p y \xrightarrow{\tau} \mu \parallel_p \delta_y} \qquad \frac{y \xrightarrow{\tau} \nu}{x \parallel_p y \xrightarrow{\tau} \delta_x \parallel_p \nu} \qquad \frac{x \xrightarrow{\alpha} \mu \quad y \stackrel{\alpha}{\not\longrightarrow} \quad \alpha \neq \tau, \surd}{x \parallel_p y \xrightarrow{\alpha} \mu \parallel_p \delta_y} \qquad \frac{x \stackrel{\alpha}{\not\longrightarrow} \quad y \xrightarrow{\alpha} \nu \quad \alpha \neq \tau, \surd}{x \parallel_p y \xrightarrow{\alpha} \delta_x \parallel_p \nu}$$

$$\frac{x \xrightarrow{\alpha} \mu \quad y \xrightarrow{\alpha} \nu \quad \alpha \neq \tau, \surd}{x \parallel_p y \xrightarrow{\alpha} \mu \parallel_p \delta_y \oplus_p \delta_x \parallel_p \nu} \qquad \frac{x \xrightarrow{\surd} \mu \quad y \xrightarrow{\surd} \nu}{x \parallel_p y \xrightarrow{\surd} \delta_{\mathrm{nil}}}$$

Table 5: Probabilistic parallel composition is outside our format.

for $t_1{}^{*}t_2$ that either t_1 is performed infinitely often in sequel, or t_1 is performed a finite number of times in sequel, followed by t_2. The bang operator $!t$ expresses that infinitely many copies of t evolve asynchronously. The probabilistic Kleene iteration expresses that $t_1{}^{*_p}t_2$ evolves to a probabilistic choice (with, resp., the probability p and $1-p$) between the two nondeterministic choices of the Kleene star operation $t_1{}^{*}t_2$ for actions which can be performed by both t_1 and t_2. For actions that can be performed by either only t_1 or only t_2, $t_1{}^{*_p}t_2$ behaves just like $t_1{}^{*}t_2$. The probabilistic bang replication expresses that $!_p t$ replicates the argument process t with probability $1-p$ and behave like t with probability p.

The arguments of all these operators are labeled Λ-frozen, as they do not contain running processes, and $\aleph$-liquid, as they can all start their execution immediately. Thus, constraints 2–4 of Definition 25 are trivially satisfied by the rules in Table 4. To check constraint 1, notice that for each rule whenever the right-hand side of a positive premise occurs in the target of the rule it does it in a Λ-liquid position, accordingly to the liquid/frozen labeling of arguments with respect to predicates $\aleph$ and Λ discussed for the operators specified by the rules in Table 3. Moreover, the PTSS in Table 4 is $(\aleph \cap \Lambda)$-patient.

Corollary 2. *Probabilistic rooted branching bisimilarity is a congruence for all operators specified in Table 4.*

Table 5 contains the PGSOS rules defining the probabilistic parallel composition $\parallel_p$, which falls outside of our format. For actions that can be performed by both t and t', $t \parallel_p t'$ evolves to a probabilistic choice (with probability weights p and $1-p$) between the two targets of the asynchronous parallel composition $t \parallel t'$. For actions that can be performed by either only t or only t', the probabilistic parallel composition $t \parallel_p t'$ behaves just like the asynchronous parallel composition $t \parallel t'$. However, notice that differently from the rules for $\parallel$ in Table 3, the asynchronous moves for t, t' in $t \parallel_p t'$ are derived by testing also negative premises: if t does an α-move, then t' must not be able to perform α in order to derive the asynchronous α-move for $t \parallel_p t'$. The presence of negative premises causes the probabilistic parallel composition operator to fail our format. In fact, both arguments of $\parallel_p$ have to be labeled $(\aleph \cap \Lambda)$-liquid since they can both contain processes that are running (as in asynchronous parallel composition) as well as processes that can immediately start their execution. The testing of negative premises for one of the two arguments then violates constraint 4 of Definition 25.

Notice that operator $\parallel_p$ breaks rooted branching bisimilarity. For instance, given processes $s = a.\tau.a.\varepsilon$ and $t = a.a.\varepsilon$ we have $s \approx_{\mathrm{rb}} t$ but $s \parallel_p t \not\approx_{\mathrm{rb}} t \parallel_p t$.

9. Extending the format to ntμfθ/ntμxθ specifications

As previously outlined, the choice of considering transition system specifications in the PGSOS format was motivated by the fact that it allowed us to present a constructive definition of ruloids as an inductive composition of PGSOS rules. Although technical, Definition 20 is way more simpler and intuitive than the one proposed in, e.g., [10] for a transition system specification (TSS) in the ntyft/ntyxt format, which involves the transformation of the TSS into an equivalent TSS (in the sense of provability) in nxytt format. However, the reader may wonder whether the use of PGSOS rules led to an oversimplified framework and thus how general and relevant our results can be.

In this Section we address this issue at an informal level and we show how our results can be extended to *complete* PTSSs in *decent* ntμfθ/ntμxθ format. The choice of discussing this section only in an informal fashion relies on the fact that the main results of this paper already come with a such heavy amount of technical definitions and proofs that we prefer not to add more technicalities to the paper. Moreover, most of the results that we will discuss in this section can be adapted from their non-probabilistic counterparts in [10] in a straightforward manner.

The key idea is that all the results in Sections 5–7 of [10] on TSSs in ntyft/ntyxt format *without lookahead* can be extended to PTSSs in the *simple* ntμfθ/ntμxθ format from [41]. We remark that the latter format does not allow lookahead by definition (see Definition 28 below). As the absence of lookahead in the rules is required by all the aforementioned results in [10], this feature is neither a restriction nor a simplification of the format.

We recall that a variable occurring in an inference rule is said to be *free* if it occurs neither in the source nor in the right-hand sides of the positive premises of the rule.

Definition 28 (ntμtθ rules)**.** A *simple ntμfθ rule* has the form

$$\frac{\{t_i \xrightarrow{\alpha_i} \mu_i \mid i \in I\} \quad \{t_j \xrightarrow{\alpha_j}\!\!\!\!\!\not\;\; \mid j \in J\}}{f(x_1,\dots,x_{\mathfrak{n}}) \xrightarrow{\alpha} \Theta}$$

with $t_i, t_j \in \mathbb{T}(\Sigma) \cup \mathbb{DT}(\Sigma)$, $\alpha_i, \alpha_j, \alpha \in \mathcal{A}_\tau$, $\mu_i \in \mathbf{V}_d$, $f \in \Sigma$, $x_1,\dots,x_{\mathfrak{n}} \in \mathbf{V}_s$, $\Theta \in \mathbb{DT}(\Sigma)$, and constraints:

- all μ_i, for $i \in I$, are pairwise different;
- all $x_1,\dots,x_{\mathfrak{n}}$ are pairwise different.

Then, starting from a simple ntμfθ rule we define:

- *simple ntμxθ rules* having as source of their conclusion a variable $x \in \mathbf{V}_s$ instead of $f(x_1,\dots,x_{\mathfrak{n}})$;
- *simple ntμtθ rules* having as source of their conclusion an arbitrary term $t \in \mathbb{T}(\Sigma)$ instead of $f(x_1,\dots,x_{\mathfrak{n}})$;
- *simple nxμfθ (resp. nxμtθ) rules* are ntμfθ (resp. ntμtθ) rules in which left-hand sides of premises are variables;
- *simple xμnfθ (resp. xμntθ) rules* are ntμfθ (resp. ntμtθ) rules in which left-hand sides of positive premises are variables.

A *simple ntμfθ/ntμxθ rule* is either a simple ntμfθ rule or a simple ntμxθ rule. For x $\in$ {ntμfθ, ntμxθ, ntμtθ, nxμfθ, nxμtθ, xμnfθ, xμntθ}, a simple x rule is *decent* if no free variable occurs in it. A PTSS is in *(decent) simple* x *format* if all its rules are (decent) simple x rules.

Notice that PGSOS rules are decent simple nxμfθ rules.

For a PTSS P in simple ntμfθ/ntμxθ format, the notion of P-ruloid is strictly related to that of *irredundant proof.*

Definition 29 (Irredundant proof, [10])**.** Assume a PTSS P. An *irredundant proof* of an inference rule $\frac{\mathbf{H}}{\ell}$ from P is a well-founded, upwardly branching tree of which the nodes are labeled by literals, and some of the leaves are marked 'hypothesis', such that:

- the root is labeled by ℓ,
- $\mathbf{H}$ is the set of labels of the hypothesis, and

- if ℓ' is the label of a node which is not an hypothesis and $\mathbf{K}$ is the set of labels of the nodes directly above it, then $\frac{\mathbf{K}}{\ell'}$ is a substitution instance of a transition rule in P.

A *proof* of $\frac{\mathbf{K}}{\ell}$ from P is an irredundant proof of $\frac{\mathbf{H}}{\ell}$ from P with $\mathbf{H} \subseteq \mathbf{K}$. We say that $\frac{\mathbf{K}}{\ell}$ is *provable* (resp. *irredundantly provable*) from P, notation $P \Vdash \frac{\mathbf{K}}{\ell}$ (resp. $P \Vdash_{\mathrm{irr}} \frac{\mathbf{K}}{\ell}$), if an proof (resp. irredundant proof) of $\frac{\mathbf{K}}{\ell}$ from P exists.

We recall that the notion of provability $\vdash$ in Definition 9 is called *well-supported provability* in [10, 46]. Then, to obtain the ruloids for a PTSS in ntμfθ/ntμxθ format we would need the following results:

1. Firstly we show that for any standard PTSS P in simple ntμfθ/ntμxθ format, there exists a standard PTSS P' in decent xμntθ format such that $P \vdash \ell$ iff $P' \vdash \ell$ for all closed literals ℓ.
2. We show that the notion of well-supported provability from P' coincides with the less demanding notion of *supported provability* [46] from P'.
3. From P' we build a PTSS P^+ in decent simple ntμfθ format such that supported provability from P' coincides with the provability $\Vdash$ from P^+. We remark that P^+ is not guaranteed to contain standard rules only and, moreover, that the transformation form P' to P^+ performed in the non-probabilistic case [10] subsumes the same construction proposed in Definition 20 for the definition of ruloids having a negative literal as conclusion.
4. We prove that in P^+, for each operator f there is a collection of decent simple nxμfθ rules that are irredundantly provable from P^+ and that allow us to derive all the closed literals having $f(t_1, \dots, t_{\mathtt{n}})$ as left-hand side that are provable from P^+.
5. Finally, P-ruloids are defined as such decent simple ntμfθ/ntμxθ rules that are irredundantly provable form P^+.

Interestingly, all these results can be proved by following the same reasoning and arguments proposed in [10]. In fact, we notice that the right-hand sides of positive premises basically play no role in the proofs in [10], that can thus be easily adapted to deal with distribution variables and terms. Clearly, the decomposition method would be defined as in Definition 23 by exploiting the ruloids obtained from the simple ntμfθ/ntμxθ rules. Then, we can extend the PRBB and PBB formats to PTSSs in decent simple ntμfθ/ntμxθ format with respect to predicates $\aleph$ and Λ. To this purpose we need to slightly modify Definition 25 in order to deal with the presence of arbitrary terms as left-hand sides of premises.

Definition 30 (Extended PRBB rule)**.** An ntμtθ rule r is *probabilistic rooted branching bisimulation safe (PRBB)* with respect to predicates $\aleph$ and Λ if it as in Definition 25, with items 3–4 rewritten as:

3. If x occurs only $\aleph$-frozen in source src(r), then x occurs only $\aleph$-frozen in prem(r).
4. If x has exactly one $\aleph$-liquid occurrence in source src(r), which is also Λ-liquid, then x has at most one $\aleph$-liquid occurrence in prem(r), which must be in a positive premise. If moreover this premise is labeled τ, then r must be $\aleph \cap \Lambda$-patient.

Definition 31 (Extended format)**.** A PTSS P is in *probabilistic rooted branching bisimulation (PRBB) format* if, for some predicates $\aleph$ and Λ, it is $(\aleph \cap \Lambda)$-patient, it is in decent simple ntμfθ/ntμxθ format and it only contains PRBB rules. It is in *probabilistic branching bisimulation (PBB) format* if moreover the predicate Λ is universal.

The proof of the congruence results then follows as discussed in our Section 7:

A. We must show that the syntactic constraints imposed by the format on the rules of P are preserved by the ruloids. This is done in three steps: i) We prove that the syntactic constraints on P are preserved by the PTSS P' introduced in item 1 above. ii) We prove that the syntactic constraints on P' are preserved by the PTSS P^+ introduced in item 3 above. iii) We prove that the syntactic constraints on P^+ are preserved by the decent simple nxμfθ rules that are irredundantly provable from P^+.

B. We need to show that the logical characterization is preserved by the decomposition method. This is proved by applying the same reasoning of the proof of Theorems 6 and 7.

C. We remark that the Σ-DS is independent from the choice of the format for inference rules on terms and thus all the results about Σ-distribution rules and ruloids, and the decomposition of distribution formulae would still hold in this general setting.

D. The congruence property is then derived by following the schema in Figure 3.

10. Conclusions

We have defined the PRBB and the PBB formats guaranteeing, respectively, that probabilistic rooted branching bisimilarity and probabilistic branching bisimilarity are congruences for nondeterministic probabilistic processes specified via PGSOS rules.

Up to our knowledge, the only other proposal of a format for probabilistic weak semantics is the RBB safe specification format of [61]. Although the constraints on inference rules are almost the same, the construction of our PRBB format results technically simpler. In [61], terms are built on a two sorted signature, mixing state and distribution terms. Thus, to obtain a proper labeling of arguments of operators, defined in terms of *wild/tame* labels as in [9, 31], they need to define *w-nesting graphs* which allow them to keep track of wild arguments of operators in the derivation of transitions. The use of unary predicates and the liquid/frozen labeling of arguments of operators, together with the use of Σ-distribution rules naturally substitute this technical construction.

To derive the congruence results we have developed an SOS-based decomposition method for modal formulae expressing both probabilistic and weak behavior of processes.

Formally, we have introduced the modal logic $\mathbb{L}$ which extends HML by providing the probabilistic choice operator $\bigoplus$ from [24] to capture probabilistic behavior, and the modalities $\langle\hat{\tau}\rangle$ and $\langle\varepsilon\rangle$ to capture the weak behavior. In particular, formulae of the form $\langle\varepsilon\rangle\psi$, with ψ a formula capturing the probabilistic behavior, express a form of *probabilistic lookahead*: it specifies the probabilistic behavior that the processes have to show after the execution of an arbitrary sequence of silent steps and such behavior can be verified only without loosing all the information of probabilistic step-by-step behavior.

As a first result we have proved that two fragments of $\mathbb{L}$, denoted by $\mathbb{L}_{\mathrm{b}}$ and $\mathbb{L}_{\mathrm{rb}}$, allow us to characterize, respectively, probabilistic branching bisimilarity and its rooted version.

Recently, [62] proposed an alternative characterization of branching bisimilarity obtained on a modal logic equipped with a quantitative modality $[\phi]_r$ from [66] that simulates the quantitative diamond modality from [59]: a probability distribution π satisfies $[\phi]_r$ iff $\pi(\{t \in \mathbf{T}(\Sigma) \mid t \models \phi\}) \geq r$. Moreover, since the hyper-transitions in [62] always subsume the silent moves, they do not need to test the ε sequences, namely our formulae $\langle\varepsilon\rangle 1(\varphi \wedge \langle a\rangle\psi)$ and $\langle\varepsilon\rangle 1(\varphi \wedge \langle\hat{\tau}\rangle\psi)$ are replaced, resp., by $\varphi \wedge \langle a\rangle\varphi'$ and $\varphi \wedge \langle\hat{\tau}\rangle\varphi'$ for φ' generic formula in their logic. However, the use of this kind of hyper-transitions would have ended up in a much more technical dissertation of our results.

To define a modal decomposition for formulae in $\mathbb{L}$ we have exploited the ruloids and SOS-like machinery for distribution terms that we proposed in [13] that, in particular, allowed us to decompose formulae of the form $\langle\varepsilon\rangle\psi$ and thus the probabilistic lookahead introduced by them. Then, we have introduced the set of syntactical constraints of the RBB format from [37] on PGSOS rules and we have proved that they are preserved in the construction of ruloids. Further we have shown that the logical characterizations are preserved in the decomposition: by decomposing formulae in $\mathbb{L}_{\mathrm{b}}$ (resp. $\mathbb{L}_{\mathrm{rb}}$) we obtain formulae in $\mathbb{L}_{\mathrm{b}}$ (resp. $\mathbb{L}_{\mathrm{rb}}$). We have then obtained the congruence theorems as direct consequences of these results.

As a final remark, we recall that differently from [10, 33, 36, 37] our decomposition method and format are defined on a PGSOS specification instead of the more general ntμfθ/ntμxθ specification [19]. Undoubtedly, this choice resulted in a simplified framework. However, we remark that all the technical results proved in [10] and exploited in [33, 36, 37] to the derive congruence formats from the decomposition, were aimed at transforming a ntyft/ntyxt specification into a GSOS specification without double testing, on which the ruloids are defined (cf. [47]). Thus, to simplify reasoning and improve the readability of our paper, we decided to start directly from a PGSOS PTSS. Notice that this choice also allowed us to present the concrete construction of P-ruloids. Nevertheless, in Section 9 we have sketched how it would possible to extend our formats and results to ntμfθ/ntμxθ specifications.

As future work, we aim to extend the formats in [10, 33, 37] to the probabilistic setting and to apply our decomposition method to systematically derive congruence formats for other relations in the weak spectrum [45]. This can be done in a straightforward fashion for the probabilistic relations of η, *delay* and *weak bisimilarity*. Intuitively, the definition of ruloids and the decomposition method would be the same that we have proposed in this paper. The logical characterizations would be obtained on proper sub-classes of $\mathbb{L}$ and the syntactic constraints of the formats could be derived from those of the corresponding formats in [33, 37].

Conversely, the definition of formats for weak relations respecting *stability* or with *divergence* [35] would be more challenging as the introduction of syntactic constraints also on probabilistic behavior, and thus on distribution specifications, seems inevitable.

Moreover, in [12, 15] we provided a logical characterization of the *bisimulation metric* [22, 28, 71]. Inspired by this result, we aim to extend the divide & congruence approach to derive the compositional properties of a behavioral pseudometric from the modal decomposition of formulae characterizing it. As the metric semantics provide notions of *distance* over processes, the formats for them guarantee that a small variance in the behavior of the subprocesses leads to a bounded small variance in the behavior of the composed processes (*uniform continuity* [40]). We aim to use the decomposition method to re-obtain the formats for the bisimilarity metric proposed in [42–44] and to automatically derive original formats for weak metric semantics [29, 56], and metric variants of branching bisimulation equivalence.

References

[1] Aceto, L., Fokkink, W. J., & Verhoef, C. (2001). Structural operational semantics. In *Handbook of Process Algebra* (pp. 197–292). Elsevier.

[2] Aldini, A., & Pierro, A. D. (2004). A quantitative approach to noninterference for probabilistic systems. *ENTCS*, *99*, 155–182. doi:10.1016/j.entcs.2004.02.007.

[3] Andova, S. (1999). Process algebra with probabilistic choice. In *Proceedings of ARTS'99* (pp. 111–129). volume 1601 of *LNCS*. doi:10.1007/3-540-48778-6_7.

[4] Andova, S. (2002). *Probabilistic Process Algebra*. Ph.D. thesis Eindhoven University of Technology.

[5] 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.

[6] Baeten, J. C. M., & Weijland, W. P. (1990). *Process Algebra*. Cambridge University Press.

[7] Bartels, F. (2004). *On Generalised Coinduction and Probabilistic Specification Formats: Distributive laws in coalgebraic modelling*. Ph.D. thesis VU University Amsterdam.

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

[9] Bloom, B. (1995). Structural operational semantics for weak bisimulations. *Theor. Comput. Sci.*, *146*, 25–68. doi:10.1016/0304-3975(94)00152-9.

[10] Bloom, B., Fokkink, W. J., & van Glabbeek, R. J. (2004). Precongruence formats for decorated trace semantics. *ACM Trans. Comput. Log.*, *5*, 26–78. doi:10.1145/963927.963929.

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

[12] Castiglioni, V., Gebler, D., & Tini, S. (2016). Logical characterization of bisimulation metrics. In *Proceedings of QAPL 2016* (pp. 44–62). volume 227 of *EPTCS*. doi:10.4204/EPTCS.227.4.

[13] Castiglioni, V., Gebler, D., & Tini, S. (2016). Modal decomposition on nondeterministic probabilistic processes. In *Proceedings of CONCUR 2016* (pp. 36:1–36:15). volume 59 of *LIPIcs*. doi:10.4230/LIPIcs.CONCUR.2016.36.

[14] Castiglioni, V., Gebler, D., & Tini, S. (2018). SOS-based modal decomposition on nondeterministic probabilistic processes. *Logical Methods in Computer Science*, *14*. doi:10.23638/LMCS-14(2:18)2018.

[15] Castiglioni, V., & Tini, S. (2019). Logical characterization of branching metrics for nondeterministic probabilistic transition systems. *Inf. Comput.*, . doi:10.1016/j.ic.2019.06.001.

[16] Clark, K. L. (1977). Negation as failure. In *Logic and Data Bases* (pp. 293–322).

[17] D'Argenio, P. R., Gebler, D., & Lee, M. D. (2014). Axiomatizing bisimulation equivalences and metrics from probabilistic SOS rules. In *Proceedings of FoSSaCS 2014* (pp. 289–303). volume 8412 of *LNCS*. doi:10.1007/978-3-642-54830-7_19.

[18] D'Argenio, P. R., Gebler, D., & Lee, M. D. (2016). A general SOS theory for the specification of probabilistic transition systems. *Inf. Comput.*, *249*. doi:10.1016/j.ic.2016.03.009.

[19] D'Argenio, P. R., & Lee, M. D. (2012). Probabilistic transition system specification: Congruence and full abstraction of bisimulation. In *Proceedings of FoSSaCS 2012* (pp. 452–466). volume 7213 of *LNCS*. doi:10.1007/978-3-642-28729-9_30.

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

[21] Deng, Y. (2015). *Semantics of Probabilistic Processes: An Operational Approach*. Springer.

[22] Deng, Y., Chothia, T., Palamidessi, C., & Pang, J. (2006). Metrics for action-labelled quantitative transition systems. *ENTCS*, *153*, 79–96. doi:10.1016/j.entcs.2005.10.033.

[23] Deng, Y., & Du, W. (2007). Probabilistic barbed congruence. *ENTCS*, *190*, 185–203. doi:10.1016/j.entcs.2007.07.011.

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

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

[26] 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.

[27] Deng, Y., van Glabbeek, R. J., Hennessy, M., Morgan, C., & Zhang, C. (2007). Remarks on testing probabilistic processes. *ENTCS*, *172*, 359–397. doi:10.1016/j.entcs.2007.02.013.

[28] Desharnais, J., Gupta, V., Jagadeesan, R., & Panangaden, P. (2004). Metrics for labelled Markov processes. *Theoret. Comput. Sci.*, *318*, 323–354. doi:10.1016/j.tcs.2003.09.013.

[29] Desharnais, J., Jagadeesan, R., Gupta, V., & Panangaden, P. (2002). The metric analogue of weak bisimulation for probabilistic processes. In *Proceedings of LICS 2002* (pp. 413–422). doi:10.1109/LICS.2002.1029849.

[30] Florescu, O., de Hoon, M., Voeten, J., & Corporaal, H. (2006). Probabilistic modelling and evaluation of soft real-time embedded systems. In *Proceedings of SAMOS 2006* (pp. 206–215). volume 4017 of *LNCS*. doi:10.1007/11796435_22.

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

[32] Fokkink, W. J., & van Glabbeek, R. J. (2016). Divide and congruence II: delay and weak bisimilarity. In *Proceedings of LICS 2016* (pp. 778–787). doi:10.1145/2933575.2933590.

[33] Fokkink, W. J., & van Glabbeek, R. J. (2017). Divide and congruence II: from decomposition of modal formulas to preservation of delay and weak bisimilarity. *Inf. Comput.*, *257*, 79–113. doi:10.1016/j.ic.2017.10.003.

[34] Fokkink, W. J., & van Glabbeek, R. J. (2017). Precongruence formats with lookahead through modal decomposition. In *Proceedings of CSL 2017* (pp. 25:1–25:20). volume 82 of *LIPIcs*. doi:10.4230/LIPIcs.CSL.2017.25.

[35] Fokkink, W. J., van Glabbeek, R. J., & Luttik, B. (2017). Divide and congruence III: stability & divergence. In *Proceedings of CONCUR 2017* (pp. 15:1–15:16). volume 85 of *LIPIcs*. doi:10.4230/LIPIcs.CONCUR.2017.15.

[36] Fokkink, W. J., van Glabbeek, R. J., & de Wind, P. (2006). Compositionality of Hennessy-Milner logic by structural operational semantics. *Theoret. Comput. Sci.*, *354*, 421–440. doi:10.1016/j.tcs.2005.11.035.

[37] 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.

[38] Foster, N., Kozen, D., Mamouras, K., Reitblatt, M., & Silva, A. (2016). Probabilistic netkat. In *Proceedings of ESOP 2016* (pp. 282–309). volume 9632 of *LNCS*. doi:10.1007/978-3-662-49498-1_12.

[39] Gebler, D., & Fokkink, W. J. (2012). Compositionality of probabilistic Hennessy-Milner logic through structural operational semantics. In *Proceedings of CONCUR 2012* (pp. 395–409). volume 7454 of *LNCS*. doi:10.1007/978-3-642-32940-1_28.

[40] Gebler, D., Larsen, K. G., & Tini, S. (2016). Compositional metric reasoning with Probabilistic Process Calculi. *Logical Methods in Computer Science*, *12*. doi:10.2168/LMCS-12(4:12)2016.

[41] Gebler, D., & Tini, S. (2013). Compositionality of approximate bisimulation for probabilistic systems. In *Proceedings of EXPRESS/SOS 2013* (pp. 32–46). volume 120 of *EPTCS*. doi:10.4204/EPTCS.120.4.

[42] Gebler, D., & Tini, S. (2014). Fixed-point characterization of compositionality properties of probabilistic processes combinators. In *Proceedings of EXPRESS/SOS 2014* (pp. 63–78). doi:10.4204/EPTCS.160.7.

[43] Gebler, D., & Tini, S. (2015). SOS specifications of probabilistic systems by uniformly continuous operators. In *Proceedings of CONCUR 2015* (pp. 155–168). volume 42 of *LIPIcs*. doi:10.4230/LIPIcs.CONCUR.2015.155.

[44] Gebler, D., & Tini, S. (2018). SOS specifications for uniformly continuous operators. *J. Comput. Syst. Sci.*, *92*, 113–151. doi:10.1016/j.jcss.2017.09.011.

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

[46] van Glabbeek, R. J. (1996). The meaning of negative premises in transition system specifications II. In *Proceedings of ICALP'96* (pp. 502–513). doi:10.1007/3-540-61440-0_154.

[47] van Glabbeek, R. J. (2011). On cool congruence formats for weak bisimulations. *Theor. Comput. Sci.*, *412*, 3283–3302. doi:10.1016/j.tcs.2011.02.036.

[48] 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.

[49] Groote, J. F. (1993). Transition system specifications with negative premises. *Theoret. Comput. Sci.*, *118*, 263–299. doi:10.1016/0304-3975(93)90111-6.

[50] Gutiérrez-Reina, D., Marín, S. L. T., Johnson, P., & Barrero, F. (2015). A survey on probabilistic broadcast schemes for wireless ad hoc networks. *Ad Hoc Networks*, *25*, 263–292. doi:10.1016/j.adhoc.2014.10.001.

[51] Hamadou, S., Palamidessi, C., & Sassone, V. (2017). Quantifying leakage in the presence of unreliable sources of information. *J. Comput. Syst. Sci.*, *88*, 27–52. doi:10.1016/j.jcss.2017.03.013.

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

[53] Hennessy, M., & Milner, R. (1985). Algebraic laws for nondeterminism and concurrency. *J. ACM*, *32*, 137–161. doi:10.1145/2455.2460.

[54] Jonsson, B., Yi, W., & Larsen, K. G. (2001). Probabilistic extensions of process algebras. In *Handbook of Process Algebra* (pp. 685–710). North-Holland / Elsevier. doi:10.1016/b978-044482830-9/50029-1.

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

[56] Lanotte, R., Merro, M., & Tini, S. (2017). Compositional weak metrics for group key update. In *Proceedings of MFCS 2017* (pp. 72:1–72:16). volume 83 of *LIPIcs*. doi:10.4230/LIPIcs.MFCS.2017.72.

[57] Lanotte, R., & Tini, S. (2009). Probabilistic bisimulation as a congruence. *ACM TOCL*, *10*, 1–48. doi:10.1145/1462179.1462181.

[58] Larsen, K. G. (1986). *Context-dependent bisimulation between processes*. Ph.D. thesis University of Edinburgh.

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

[60] Larsen, K. G., & Xinxin, L. (1991). Compositionality through an operational semantics of contexts. *J. Log. Comput.*, *1*, 761–795. doi:10.1093/logcom/1.6.761.

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

[62] Lee, M. D., & de Vink, E. P. (2016). Logical characterization of bisimulation for transition relations over probability distributions with internal actions. In *Proceedings of MFCS 2016* (pp. 29:1–29:14). volume 58 of *LIPIcs*. doi:10.4230/LIPIcs.MFCS.2016.29.

[63] 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.

[64] Mio, M., & Simpson, A. (2013). A proof system for compositional verification of probabilistic concurrent processes. In *Proceedings of FoSSsCS 2013* (pp. 161–176). volume 7794 of *LNCS*. doi:10.1007/978-3-642-37075-5_11.

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

[66] Parma, A., & Segala, R. (2007). Logical characterizations of bisimulations for discrete probabilistic systems. In *Proceedings of FoSSaCS 2007* (pp. 287–301). volume 4423 of *LNCS*. doi:10.1007/978-3-540-71389-0_21.

[67] Philippou, A., Lee, I., & Sokolsky, O. (2000). Weak bisimulation for probabilistic systems. In *Proceedings of CONCUR 2000* (pp. 334–349). volume 1877 of *LNCS*. doi:10.1007/3-540-44618-4_25.

[68] Plotkin, G. D. (1980). An operational semantics for CSO. In *Proceedings of Logics of Programs and Their Applications* (pp. 250–252). doi:10.1007/3-540-11981-7_17.

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

[70] de Simone, R. (1985). Higher-level synchronising devices in MEIJE-SCCS. *Theor. Comput. Sci.*, *37*, 245–267. doi:10.1016/0304-3975(85)90093-3.

[71] van Breugel, F., & Worrell, J. (2001). An algorithm for quantitative verification of probabilistic transition systems. In *Proceedings of CONCUR 2001* (pp. 336–350). volume 2154 of *LNCS*. doi:10.1007/3-540-44685-0_23.

Appendix A. Relation lifting

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

Proposition 3 ([24]). *Consider a relation $\mathcal{R} \subseteq X \times Y$. Then $\mathcal{R}^\dagger \subseteq \Delta(X) \times \Delta(Y)$ is the smallest relation satisfying*

1. *$x \,\mathcal{R}\, y$ implies $\delta_x \,\mathcal{R}^\dagger\, \delta_y$;*
2. *$\pi_i \,\mathcal{R}^\dagger\, \pi_i'$ implies $(\sum_{i\in I} p_i\pi_i) \,\mathcal{R}^\dagger\, (\sum_{i\in I} p_i\pi_i')$, for any set of indexes I with $\sum_{i\in I} p_i = 1$.*

Proposition 4 ([25, Proposition 1]). *Consider two sets X, Y, Let $\pi \in \Delta(X), \pi' \in \Delta(Y)$ and $\mathcal{R} \subseteq X \times Y$. Then $\pi \,\mathcal{R}^\dagger\, \pi'$ iff there are a set of indexes I and of weights $p_i \in (0,1]$ with $\sum_{i\in I} p_i = 1$, such that*

(i) $\pi = \sum_{i\in I} p_i\delta_{x_i}$,

(ii) $\pi' = \sum_{i\in I} p_i\delta_{y_i}$ *and*

(iii) $x_i \;\mathcal{R}\; y_i$ *for all* $i \in I$.

Appendix B. Proofs of results in Section 3

In order to prove Theorem 1, we need to introduce the following auxiliary Lemma.

Lemma 5. *Assume a family of distributions $\{\pi_j \mid j \in J\}$ such that $\pi_j \models \psi$ for each $j \in J$. Then, whenever $\pi = \sum_{j\in J} p_j\pi_j$, for weights $p_j \in (0,1]$ with $\sum_{j\in J} p_j = 1$, then $\pi \models \psi$.*

The proof of Theorem 1 is based on the analogous result for nonprobabilistic processes in [37].

Proof of Theorem 1. We prove that $s \approx_{\mathrm{b}} t$ if and only if $\mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(s) = \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(t)$. The proof that $s \approx_{\mathrm{rb}} t$ if and only if $\mathbb{L}_{\mathrm{rb}}(s) = \mathbb{L}_{\mathrm{rb}}(t)$ is analogous. We proceed as follows:

1. First we prove that $s \approx_{\mathrm{b}} t$ implies $\mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(s) = \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(t)$ by showing that for any formula $\phi \in \mathbb{L}$ we have:
   (i) if ϕ is a state formula and $s \approx_{\mathrm{b}} t$ then we have $\phi \in \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(s)$ if and only if $\phi \in \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(t)$, and
   (ii) if ϕ is a distribution formula and $\pi_1 \approx^\dagger_{\mathrm{b}} \pi_2$ then we have $\phi \in \mathbb{L}^{\mathrm{d}}_{\mathrm{b}}(\pi_1)$ if and only if $\phi \in \mathbb{L}^{\mathrm{d}}_{\mathrm{b}}(\pi_2)$.
2. Then we prove that $\mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(s) = \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(t)$ implies $s \approx_{\mathrm{b}} t$ by showing that the relation
$$\mathcal{R} = \{(s,t) \in \mathbf{T}(\Sigma) \times \mathbf{T}(\Sigma) \mid \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(s) = \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(t)\}$$
   is a probabilistic branching bisimulation.

1. We prove 1i and 1ii in parallel, by induction over ϕ.
   - Base case $\varphi = \top$. In this case there is nothing to prove as $s \models \top$ and $t \models \top$ always hold.
   - Inductive step $\varphi = \neg\varphi'$. By Definition 14, $s \models \neg\varphi'$ implies $s \not\models \varphi'$, which by the inductive hypothesis and the symmetry of $\approx_{\mathrm{b}}$ gives $t \not\models \varphi'$ and thus, by Definition 14, $t \models \neg\varphi'$.
   - Inductive step $\varphi = \bigwedge_{j\in J}\varphi_j$. By Definition 14, $s \models \bigwedge_{j\in J}\varphi_j$ implies that $s \models \varphi_j$ for all $j \in J$, which by the inductive hypothesis gives $t \models \varphi_j$ for all $j \in J$ and thus, by Definition 14, $t \models \bigwedge_{j\in J}\varphi_j$.
   - Inductive step $\psi = \bigoplus_{i\in I} r_i\varphi_i$. On one hand, by Definition 3, $\pi_1 \approx^\dagger_{\mathrm{b}} \pi_2$ implies the existence of a matching $\mathfrak{w} \in \mathfrak{W}(\pi_1, \pi_2)$ s.t.
     (a) for each $s' \in \mathsf{supp}(\pi_1)$, $\pi_1(s') = \sum_{t'\in\mathsf{supp}(\pi_2)} \mathfrak{w}(s',t')$;
     (b) for each $t' \in \mathsf{supp}(\pi_2)$, $\pi_2(t') = \sum_{s'\in\mathsf{supp}(\pi_1)} \mathfrak{w}(s',t')$;

(c) $s' \approx_{\mathrm{b}} t'$ whenever $\mathfrak{w}(s', t') > 0$.

Item (1c) and the inductive hypothesis give $\mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(s') = \mathbb{L}^{\mathrm{s}}_{\mathrm{b}}(t')$ whenever $\mathfrak{w}(s', t') > 0$. On the other hand, by Definition 14, $\pi_1 \models \psi$ implies the existence of a matching $\mathfrak{w}' \in \mathfrak{W}(\pi_1, \bigoplus_{i\in I} r_i\varphi_i)$ s.t.

(d) for each $s' \in \mathsf{supp}(\pi_1)$, $\pi_1(s') = \sum_{i\in I} \mathfrak{w}'(s', \varphi_i)$;

(e) for each $i \in I$, $r_i = \sum_{s'\in\mathsf{supp}(\pi_1)} \mathfrak{w}'(s', \varphi_i)$;

(f) $s' \models \varphi_i$ whenever $\mathfrak{w}'(s', \varphi_i) > 0$.

We need to exhibit a matching $\tilde{\mathfrak{w}} \in \mathfrak{W}(\pi_2, \bigoplus_{i\in I} r_i\varphi_i)$ s.t. $t' \models \varphi_i$ whenever $\tilde{\mathfrak{w}}(t', \varphi_i) > 0$. Define, for all $t' \in \mathsf{supp}(\pi_2)$ and $i \in I$,

$$\tilde{\mathfrak{w}}(t', \varphi_i) = \sum_{s'\in\mathsf{supp}(\pi_1)} \frac{\mathfrak{w}(s', t') \cdot \mathfrak{w}'(s', \varphi_i)}{\pi_1(s')}.$$

First we show that $\tilde{\mathfrak{w}}$ is a well defined matching in $\mathfrak{W}(\pi_2, \bigoplus_{i\in I} r_i\varphi_i)$. We have

$$\begin{aligned}
\sum_{i\in I} \tilde{\mathfrak{w}}(t', \varphi_i) &= \sum_{i\in I} \left(\sum_{s'\in\mathsf{supp}(\pi_1)} \frac{\mathfrak{w}(s', t') \cdot \mathfrak{w}'(s', \varphi_i)}{\pi_1(s')} \right) \\
&= \sum_{s'\in\mathsf{supp}(\pi_1)} \frac{\mathfrak{w}(s', t')}{\pi_1(s')} \cdot \left(\sum_{i\in I} \mathfrak{w}'(s', \varphi_i) \right) \\
&= \sum_{s'\in\mathsf{supp}(\pi_1)} \frac{\mathfrak{w}(s', t')}{\pi_1(s')} \cdot \pi_1(s') && \text{(by item (1d))} \\
&= \sum_{s'\in\mathsf{supp}(\pi_1)} \mathfrak{w}(s', t') \\
&= \pi_2(t') && \text{(by item (1b))}
\end{aligned}$$

and, similarly,

$$\begin{aligned}
\sum_{t'\in\mathsf{supp}(\pi_2)} \tilde{\mathfrak{w}}(t', \varphi_i) &= \sum_{t'\in\mathsf{supp}(\pi_2)} \left(\sum_{s'\in\mathsf{supp}(\pi_1)} \frac{\mathfrak{w}(s', t') \cdot \mathfrak{w}'(s', \varphi_i)}{\pi_1(s')} \right) \\
&= \sum_{s'\in\mathsf{supp}(\pi_1)} \frac{\mathfrak{w}'(s', \varphi_i)}{\pi_1(s')} \cdot \left(\sum_{t'\in\mathsf{supp}(\pi_2)} \mathfrak{w}(s', t') \right) \\
&= \sum_{s'\in\mathsf{supp}(\pi_1)} \frac{\mathfrak{w}'(s', \varphi_i)}{\pi_1(s')} \cdot \pi_1(s') && \text{(by item (1a))} \\
&= \sum_{s'\in\mathsf{supp}(\pi_1)} \mathfrak{w}'(s', \varphi_i) \\
&= r_i && \text{(by item (1e)).}
\end{aligned}$$

Finally, notice that $\tilde{\mathfrak{w}}(t', \varphi_i) > 0$ if and only if $\mathfrak{w}(s', t') > 0$ and $\mathfrak{w}'(s', \varphi_i) > 0$, which, together with items (1c) and (1f), gives that $t' \models \varphi_i$ whenever $\tilde{\mathfrak{w}}(t', \varphi_i) > 0$. We can therefore conclude that $\pi_2 \models \bigoplus_{i\in I} r_i\varphi_i$.

- Inductive step $\psi = \bigwedge_{j\in J} \psi_j$. By Definition 14, $\pi_1 \models \bigwedge_{j\in J} \psi_j$ implies that $\pi_1 \models \psi_j$ for each $j \in J$. By the inductive hypothesis, this entails that $\pi_2 \models \psi_j$ for each $j \in J$, thus implying $\pi_2 \models \bigwedge_{j\in J} \psi_j$.
- Inductive step $\varphi = \langle\varepsilon\rangle 1\big(\varphi' \wedge \langle a\rangle\psi\big)$. By Definition 14, $s \models \varphi$ implies that for some $n \in \mathbb{N}$ there are distributions $\pi_0, \ldots, \pi_n \in \Delta(\mathbf{T}(\Sigma))$ with

(a) $s \xrightarrow{\hat{\tau}} \pi_0$,

(b) $\pi_k \xrightarrow{\hat{\tau}} \pi_{k+1}$ for $k = 0, \dots, n-1$ and

(c) $\pi_n \models 1\big(\varphi' \wedge \langle a \rangle \psi\big)$.

We proceed by induction over $n \in \mathbb{N}$.

– Base case $n = 0$. Then $\pi_0 = \delta_s$ and $\delta_s \models 1\big(\varphi' \wedge \langle a \rangle \psi\big)$ which means that $s \models \varphi'$ and moreover there is a distribution π_s s.t. $s \xrightarrow{a} \pi_s$ and $\pi_s \models \psi$. As $s \approx_{\mathrm{b}} t$, by Definition 5 we have $t \xrightarrow{\hat{\varepsilon}} \pi \xrightarrow{\hat{a}} \pi_t$ with $\pi \approx_{\mathrm{b}}^{\dagger} \delta_s$ and $\pi_t \approx_{\mathrm{b}}^{\dagger} \pi_s$. Assume wlog. that $\pi = \sum_{j \in J} p_j \delta_{t_j}$. By Definition 4, $\pi \xrightarrow{\hat{a}} \pi_t$ implies that there are distributions π_j s.t. $t_j \xrightarrow{a} \pi_j$, for all $j \in J$, and $\pi_t = \sum_{j \in J} p_j \pi_j$. Then, by Definition 3, $\pi \approx_{\mathrm{b}}^{\dagger} \delta_s$ implies $t_j \approx_{\mathrm{b}} s$ for all $j \in J$. Thus, structural induction and $s \models \varphi'$ give

$$t_j \models \varphi' \text{ for all } j \in J. \tag{B.1}$$

Moreover, $\pi \xrightarrow{\hat{a}} \pi_t$, $t_j \approx_{\mathrm{b}} s$ and $s \xrightarrow{\hat{a}} \pi_s$ imply that for each $j \in J$ there is a distribution π_j s.t. $t_j \xrightarrow{\hat{a}} \pi_j$ and $\pi_j \approx_{\mathrm{b}}^{\dagger} \pi_s$. By structural induction, $\pi_s \models \psi$ and $\pi_s \approx_{\mathrm{b}}^{\dagger} \pi_j$ implies that $\pi_j \models \psi$ for all $j \in J$. Hence, by Lemma 5, from $\pi_t = \sum_{j \in J} p_j \pi_j$ we can conclude that $\pi_t \models \psi$. This gives that

$$t_j \models \langle a \rangle \psi \text{ for all } j \in J. \tag{B.2}$$

Summarizing, we have obtained:

* $t \xrightarrow{\hat{\varepsilon}} \pi \xrightarrow{\hat{a}} \pi_t$,
* Equations (B.1),(B.2) give $\pi \models 1(\varphi' \wedge \langle a \rangle \psi)$.

Therefore we can conclude that $t \models \langle \varepsilon \rangle 1(\varphi' \wedge \langle a \rangle \psi)$.

– Inductive step $n > 0$. As $s \xrightarrow{\hat{\tau}} \pi_0$ and $s \approx_{\mathrm{b}} t$ according to Definition 5 there are two possibilities.

(a) Either $\pi_0 \approx_{\mathrm{b}}^{\dagger} \delta_t$. Since $\pi_0 \models 1(\varphi' \wedge \langle a \rangle \psi)$, then $s' \models \langle \varepsilon \rangle 1(\varphi' \wedge \langle a \rangle \psi)$ for all $s' \in \mathsf{supp}(\pi_0)$. Moreover, $\pi_0 \approx_{\mathrm{b}}^{\dagger} \delta_t$ implies that $s' \approx_{\mathrm{b}} t$ for all $s' \in \mathsf{supp}(\pi_0)$. Hence, by induction over n we gather that $t \models \langle \varepsilon \rangle 1(\varphi' \wedge \langle a \rangle \psi)$.

(b) Or $t \xrightarrow{\hat{\varepsilon}} \pi \xrightarrow{\hat{a}} \pi_t$ with $\pi \approx_{\mathrm{b}}^{\dagger} \pi_0$. Since $s' \models \langle \varepsilon \rangle 1(\varphi' \wedge \langle a \rangle \psi)$ for all $s' \in \mathsf{supp}(\pi_0)$ and $\pi_0 \approx_{\mathrm{b}}^{\dagger} \pi$, by induction over n we gather that $t' \models \langle \varepsilon \rangle 1(\varphi' \wedge \langle a \rangle \psi)$ for all $t' \in \mathsf{supp}(\pi)$. Hence, we can conclude that $t \models \langle \varepsilon \rangle 1(\varphi' \wedge \langle a \rangle \psi)$ as well.

• Inductive step $\varphi = \langle \varepsilon \rangle 1(\varphi' \wedge \langle \hat{\tau} \rangle \psi)$. This case can be treated similarly to the previous case of $\varphi = \langle \varepsilon \rangle 1(\varphi' \wedge \langle a \rangle \psi)$.

2. Assume that $s \xrightarrow{\hat{\alpha}} \pi_s$. Firstly we notice that whenever $\alpha = \tau$ and $\pi_s \mathcal{R}^{\dagger} \delta_t$, then the first item of Definition 5 is satisfied. Hence, we can assume now that either $\alpha \neq \tau$ or $\pi_s \not\mathcal{R}^{\dagger} \delta_t$. We introduce two sets:

$$D_1 = \Big\{ \pi' \in \Delta(\mathbf{T}(\Sigma)) \mid t \xrightarrow{\hat{\varepsilon}} \pi' \wedge \delta_s \not\mathcal{R}^{\dagger} \pi' \Big\}$$

$$D_2 = \Big\{ \pi'' \in \Delta(\mathbf{T}(\Sigma)) \mid \exists \, \pi' \text{ s.t. } t \xrightarrow{\hat{\varepsilon}} \pi' \xrightarrow{\hat{\alpha}} \pi'' \wedge \pi_s \not\mathcal{R}^{\dagger} \pi'' \Big\}.$$

For each $\pi' \in D_1$, let $\varphi_{\pi'} \in \mathbb{L}_{\mathrm{b}}^{\mathrm{s}}$ be s.t. $\delta_s \models 1\varphi_{\pi'}$ and $\pi' \not\models 1\varphi_{\pi'}$. We define

$$\varphi = \bigwedge_{\pi' \in D_1} \varphi_{\pi'}.$$

For each $\pi'' \in D_2$, let $\psi_{\pi''} \in \mathbb{L}_{\mathrm{b}}^{\mathrm{d}}$ be s.t. $\pi_s \models \psi_{\pi''}$ and $\pi'' \not\models \psi_{\pi''}$. We define

$$\psi = \bigwedge_{\pi'' \in D_2} \psi_{\pi''}.$$

Clearly, $\varphi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{s}}, \psi \in \mathbb{L}_{\mathrm{b}}^{\mathrm{d}}, s \models \varphi$ and $\pi_s \models \psi$. We distinguish two cases.

- $\alpha \neq \tau$. Since $s \models \langle \varepsilon \rangle 1\big(\varphi \wedge \langle \alpha \rangle \psi\big)$ and $s \,\mathcal{R}\, t$, then also $t \models \langle \varepsilon \rangle 1\big(\varphi \wedge \langle \alpha \rangle \psi\big)$. Hence $t \xrightarrow{\hat{\varepsilon}} \pi_1 \xrightarrow{\hat{\alpha}} \pi_2$ with, in particular $\pi_1 \models 1\varphi$ and $\pi_2 \models \psi$. By construction of φ and ψ we can infer, resp., that $\delta_s \,\mathcal{R}^\dagger\, \pi_1$ and $\pi_s \,\mathcal{R}^\dagger\, \pi_2$.
- $\alpha = \tau$ and $\pi_s \not\mathcal{R}^\dagger\, \delta_t$. Let ψ' be any formula in $\mathbb{L}_{\mathrm{b}}^{\mathrm{d}}$ s.t. $\pi_s \models \psi'$ and $\delta_s, \delta_t \not\models \psi'$. Since $s \models \langle \varepsilon \rangle 1\big(\varphi \wedge \langle \hat{\tau} \rangle (\psi' \wedge \psi)\big)$ and $s \,\mathcal{R}\, t$, then also $t \models \langle \varepsilon \rangle 1\big(\varphi \wedge \langle \hat{\tau} \rangle (\psi' \wedge \psi)\big)$. Hence $t \xrightarrow{\hat{\varepsilon}} \pi_1$ with $\pi_1 \models 1\big(\varphi \wedge \langle \hat{\tau} \rangle (\psi' \wedge \psi)\big)$. In particular this gives that $\pi_1 \models 1\varphi$ which, by construction of φ allows us to infer that $\delta_s \,\mathcal{R}^\dagger\, \pi_1$. Therefore, $\delta_s \not\models \psi'$ implies that also $\pi_1 \not\models \psi'$ and thus $\pi_1 \models 1\langle \hat{\tau} \rangle (\psi' \wedge \psi)$ only if $\pi_1 \xrightarrow{\hat{\tau}} \pi_2$ for a distribution π_2 s.t. $\pi_2 \models \psi' \wedge \psi$. By construction of ψ we can thus infer that $\pi_s \,\mathcal{R}^\dagger\, \pi_2$.

In both cases we have that the second item of Definition 5 is satisfied, and thus we can conclude that $\mathcal{R}$ is a branching bisimulation as required.

$\square$