



Fully Syntactic Uniform Continuity Formats for Bisimulation Metrics

Valentina Castiglioni^{1,2(✉)}, Ruggero Lanotte³, and Simone Tini³

¹ INRIA Saclay - Ile de France, Palaiseau, France
vale.castiglioni@gmail.com

² Reykjavik University, Reykjavik, Iceland

³ University of Insubria, Como, Italy

Abstract. Behavioral metrics play a fundamental role in the analysis of probabilistic systems. They allow for a robust comparison of the behavior of processes and provide a formal tool to study their performance, privacy and security properties. Gebler, Larsen and Tini showed that the *bisimilarity metric* is also suitable for *compositional reasoning*, expressed in terms of *continuity properties* of the metric. Moreover, Gebler and Tini provided *semantic formats* guaranteeing, respectively, the *non-extensiveness*, *non-expansiveness* and *Lipschitz continuity* of this metric. In this paper, starting from their work, we define three *specification formats* for the bisimilarity metric, one for each continuity property, namely sets of *syntactic constraints* over the SOS rules defining process operators that guarantee the desired continuity property of the metric.

1 Introduction

With the ever-increasing interest in probabilistic processes, robust notions and formal tools are needed to analyze their behavior. It has been argued many times by now that *behavioral metrics* [2, 6, 7, 13, 14, 20, 23, 34] are preferable to behavioral relations in this setting in that they overcome the high sensitivity of equivalences and preorders with respect to tiny variations in the values of probabilities. Instead of stating whether the behavior of two processes is exactly the same or not, behavioral metrics *measure* the disparities in their behavior.

A fundamental aspect for process specification and verification is *compositional reasoning*, namely to prove the compatibility of the language operators with the chosen behavioral semantics. In [16] the authors proposed to express the compositional properties of the *bisimilarity metric* [6, 13, 14], i.e., the quantitative analogue to bisimulation equivalence, in terms of its *continuity properties*. In detail, a notion of *uniform continuity* was proposed subsuming the properties of *non-extensiveness* [4], *non-expansiveness* [14] and *Lipschitz continuity* of the metric. Informally, an operator is non-extensive wrt. a metric if the distance of two systems defined via that operator is not greater than the *maximum* of the pairwise distances of their components. The notion of non-expansiveness relaxes that of non-extensiveness by allowing the distance of the two systems to be

bounded by the *sum* of the pairwise distances of their components. Lipschitz continuity extends non-expansiveness by considering as bound a *multiplicative factor* of the sum of pairwise distances of the components of the two systems.

Since [33], a *specification format* is a set of *syntactic constraints* on the form of the SOS inference rules [29, 30] used to define the operational semantics of processes, ensuring by construction a given compositionality property (see [1, 28] for surveys). A first step in the definition of specification formats for uniform continuity properties of bisimilarity metric can be found in [19]. However, the formats in [19] are not fully syntactic since they do not simply require an inspection of the pattern of the rules, but they require to compute how many copies of a process argument can be spawned by each operator. Such computation is necessarily recursive since the SOS rules define operators in terms of operators.

Our Contribution. In this paper we provide a classic syntactic format for uniform continuity properties for the bisimilarity metric. In detail, we propose three formats: the *non-extensiveness*, the *non-expansiveness*, and the *Lipschitz continuity* format. These formats build on the PGSOS format [11, 12] and on *syntactic constraints* regulating the form of the *targets of rules*. We recall that the PGSOS format guarantees that probabilistic bisimilarity, i.e., the kernel of the bisimilarity metric, is a congruence [12]. Hence, to obtain the desired continuity properties it is fundamental to start from the PGSOS format, since it ensures that by composing processes which are at distance 0, i.e., bisimilar, we obtain processes which are at distance 0. We will see that less demanding continuity properties require less demanding rule constraints.

2 Background

In the process algebra setting, *processes* are constructed inductively as *closed terms* over a suitable *signature*, namely a countable set Σ of *operators*. We let f range over Σ and \mathbf{n} range over the *rank* of $f \in \Sigma$. Operators with rank 0 are called *constants*. We let Σ^0 denote the subset of constants in the signature Σ .

Assume a set of (*state*, or *process*) *variables* \mathbf{V}_s disjoint from Σ and ranged over by x, y, \dots . The set $\mathbf{T}(\Sigma, V)$ of *terms* over Σ and a set of variables $V \subseteq \mathbf{V}_s$ is the least set satisfying: (i) $V \subseteq \mathbf{T}(\Sigma, V)$, and (ii) $f(t_1, \dots, t_n) \in \mathbf{T}(\Sigma, V)$ whenever $f \in \Sigma$ and $t_1, \dots, t_n \in \mathbf{T}(\Sigma, V)$. By $\mathbf{T}(\Sigma)$ we denote the set of the *closed terms* $\mathbf{T}(\Sigma, \emptyset)$, or *processes*. By $\mathbf{T}(\Sigma)$ we denote the set of all *open terms* $\mathbf{T}(\Sigma, \mathbf{V}_s)$. By $\text{var}(t)$ we denote the set of the variables occurring in the term t .

To equip processes with a semantics we rely on *nondeterministic probabilistic labeled transition systems* (PTSs) [31], which allow for modeling reactive behavior, nondeterminism and probability. The state space in a PTS is the set of processes $\mathbf{T}(\Sigma)$. Transitions take processes to discrete probability distributions over processes, i.e., mappings $\pi: \mathbf{T}(\Sigma) \rightarrow [0, 1]$ with $\sum_{t \in \mathbf{T}(\Sigma)} \pi(t) = 1$. The *support* of a distribution π is $\text{supp}(\pi) = \{t \in \mathbf{T}(\Sigma) \mid \pi(t) > 0\}$. By $\Delta(\mathbf{T}(\Sigma))$ we denote the set of all *finitely supported* distributions over $\mathbf{T}(\Sigma)$.

Definition 1 (PTS, [31]). A PTS is a triple $(\mathbf{T}(\Sigma), \mathcal{A}, \rightarrow)$, where Σ is a signature, \mathcal{A} a countable set of action labels, and $\rightarrow \subseteq \mathbf{T}(\Sigma) \times \mathcal{A} \times \Delta(\mathbf{T}(\Sigma))$ a transition relation.

As usual, a transition $(t, \alpha, \pi) \in \rightarrow$ is denoted $t \xrightarrow{\alpha} \pi$. Then, $t \not\xrightarrow{\alpha}$ denotes that there is no distribution π with $t \xrightarrow{\alpha} \pi$ and is called a *negative transition*.

Distributions and Distribution Terms. 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 $s \neq t$. For $f \in \Sigma$ and $\pi_i \in \Delta(\mathbf{T}(\Sigma))$, $f(\pi_1, \dots, \pi_n)$ is the distribution defined by $f(\pi_1, \dots, \pi_n)(f(t_1, \dots, t_n)) = \prod_{i=1}^n \pi_i(t_i)$ and $f(\pi_1, \dots, \pi_n)(t) = 0$ for all t not in the form $f(t_1, \dots, t_n)$. 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 $t \in \mathbf{T}(\Sigma)$.

Then, we need the notion of *distributions term*. Assume a countable set of *distribution variables* \mathbf{V}_d disjoint from Σ and \mathbf{V}_s and ranged over by μ, ν, \dots . Intuitively, any $\mu \in \mathbf{V}_d$ may instantiate to a distribution in $\Delta(\mathbf{T}(\Sigma))$. The set of *distribution terms* over Σ and the sets of variables $V_s \subseteq \mathbf{V}_s$ and $V_d \subseteq \mathbf{V}_d$, notation $\mathbf{DT}(\Sigma, V_s, V_d)$, is the least set satisfying: (i) $\{\delta_t \mid t \in \mathbf{T}(\Sigma, V_s)\} \subseteq \mathbf{DT}(\Sigma, V_s, V_d)$, (ii) $V_d \subseteq \mathbf{DT}(\Sigma, V_s, V_d)$, (iii) $f(\theta_1, \dots, \theta_n) \in \mathbf{DT}(\Sigma, V_s, V_d)$ whenever $f \in \Sigma$ and $\theta_i \in \mathbf{DT}(\Sigma, V_s, V_d)$, (iv) $\sum_{j \in J} p_j \theta_j \in \mathbf{DT}(\Sigma, V_s, V_d)$ whenever $\theta_j \in \mathbf{DT}(\Sigma, V_s, V_d)$ and $p_j \in (0, 1]$ with $\sum_{j \in J} p_j = 1$. We write $\mathbb{DT}(\Sigma)$ for $\mathbf{DT}(\Sigma, \mathbf{V}_s, \mathbf{V}_d)$, i.e. the set of all *open distribution terms*, and $\mathbf{DT}(\Sigma)$ for $\mathbf{DT}(\Sigma, \emptyset, \emptyset)$, i.e. the set of the *closed distribution terms*. Notice that closed distribution terms denote distributions. We denote by $\text{var}(\theta)$ the set of the variables occurring in distribution term θ . We let $\delta_{\mathbf{V}_s}$ denote the set $\{\delta_x \mid x \in \mathbf{V}_s\}$.

A distribution term is in *normal form* if the Dirac operator is applied only to single variables δ_x and not to non-variable terms δ_t with $t \notin \mathbf{V}_s$ and, then, either there is no convex combination or there is only one convex combination as outermost operation.

Definition 2 (Normal form, [19]). A distribution term $\theta \in \mathbb{DT}(\Sigma)$ is in normal form iff: either (i) $\theta = \mu \in \mathbf{V}_d$, or (ii) $\theta = \delta_x$ for $x \in \mathbf{V}_s$, or (iii) $\theta = f(\theta_1, \dots, \theta_n)$ where all the θ_i are in normal form and no convex combination occurs in any of the θ_i , or (iv) $\theta = \sum_{j \in J} p_j \theta_j$ where all the θ_j are in normal form and no convex combination occurs in any of the θ_j .

A *substitution* is a mapping $\sigma: \mathbf{V}_s \cup \mathbf{V}_d \rightarrow \mathbf{T}(\Sigma) \cup \mathbb{DT}(\Sigma)$ with $\sigma(x) \in \mathbf{T}(\Sigma)$ for all $x \in \mathbf{V}_s$ and $\sigma(\mu) \in \mathbb{DT}(\Sigma)$ for all $\mu \in \mathbf{V}_d$. A substitution σ is *closed* if it maps terms to closed terms in $\mathbf{T}(\Sigma) \cup \mathbf{DT}(\Sigma)$.

We say that two distribution terms θ_1, θ_2 are *equivalent*, written $\theta_1 \equiv \theta_2$ if under all closed substitutions σ , the closed distribution terms $\sigma(\theta_1)$ and $\sigma(\theta_2)$ express the same probability distribution. Each distribution term is equivalent to a normal form.

Proposition 1 ([19]). For each distribution term $\theta \in \mathbb{DT}(\Sigma)$ there is a distribution term $\theta' \in \mathbb{DT}(\Sigma)$ in normal form such that $\theta \equiv \theta'$.

PGSOS Specifications. PTSs are defined by means of SOS rules [30], which are syntax-driven inference rules allowing us to infer the behavior of processes inductively with respect to their structure. They are based on expressions of the form $t \xrightarrow{\alpha} \Theta$ and $t \not\xrightarrow{\alpha}$, called, resp., *positive* and *negative literals* with t a term and Θ a distribution term, which will instantiate to transitions through substitutions. The literals $t \xrightarrow{\alpha} \Theta$ and $t \not\xrightarrow{\alpha}$ with the same term t in the left hand side and the same label α , are said to *deny* each other.

We assume SOS rules in the *probabilistic GSOS* (PGSOS) format [12], which allows for specifying most of probabilistic process algebras operators [16]. As examples, all SOS rules in Tables 1, 2 and 3 are PGSOS rules according to the following definition.

Definition 3 (PGSOS rules, [12]). A PGSOS rule r has the form:

$$\frac{\{x_h \xrightarrow{\alpha_{h,m}} \mu_{h,m} \mid h \in H, m \in M_h\} \quad \{x_h \not\xrightarrow{\alpha_{h,n}} \mid h \in H, n \in N_h\}}{f(x_1, \dots, x_n) \xrightarrow{\alpha} \Theta} \quad (1)$$

where $f \in \Sigma$ is an operator, $H \subseteq \{1, \dots, n\}$, M_h and N_h are finite sets of indexes, $\alpha_{h,m}, \alpha_{h,n}, \alpha \in \mathcal{A}$ are action labels, $x_h \in \mathbf{V}_s$ are state variables, $\mu_{h,m} \in \mathbf{V}_d$ are distribution variables and $\Theta \in \mathbb{DT}(\Sigma)$ is a distribution term. Furthermore: (i) all distribution variables $\mu_{h,m}$ with $h \in H$ and $m \in M_h$ are distinct, (ii) all variables x_1, \dots, x_n are distinct, (iii) $\text{var}(\Theta) \subseteq \{\mu_{h,m} \mid h \in H, m \in M_h\} \cup \{x_1, \dots, x_n\}$.

Constraints (i)–(iii) are inherited by the classical GSOS rules in [5] and are necessary to ensure that probabilistic bisimulation is a congruence [8, 9, 12]. For a PGSOS rule r , the positive (resp. negative) literals above the line are called the *positive* (resp. *negative*) *premises*, notation $\text{pprem}(r)$ (resp. $\text{nprem}(r)$). We denote by $\text{der}(x, r) = \{\mu \mid x \xrightarrow{\alpha} \mu \in \text{pprem}(r)\}$ the set of *derivatives* of x in r . The literal $f(x_1, \dots, x_n) \xrightarrow{\alpha} \Theta$ is called the *conclusion*, notation $\text{conc}(r)$, the term $f(x_1, \dots, x_n)$ is called the *source*, notation $\text{src}(r)$, and the distribution term Θ is called the *target*, notation $\text{trg}(r)$.

Definition 4 (PTSS, [12]). A PGSOS-transition system specification (PTSS) is a tuple $P = (\Sigma, \mathcal{A}, R)$, with Σ a signature, \mathcal{A} a set of actions and R a set of PGSOS rules.

We conclude with the notion of *disjoint extension* for a PTSS which allows us to introduce new operators without affecting the behavior of those already specified.

Definition 5 (Disjoint extension). A PTSS $P' = (\Sigma', \mathcal{A}, R')$ is a disjoint extension of a PTSS $P = (\Sigma, \mathcal{A}, R)$, written $P \sqsubseteq P'$, if $\Sigma \subseteq \Sigma'$, $R \subseteq R'$ and R' introduces no new rule for any operator in R .

3 Behavioral Metrics and Their Compositional Properties

Behavioral equivalences, such as *probabilistic bisimulation* [27,32], answer the question of whether two processes behave precisely the same way or not with respect to the *observations* we can make on them. *Behavioral metrics* [2,6,13,14,34] answer the more general question of measuring the differences in processes behavior. They are usually defined as 1-bounded *pseudometrics* expressing the behavioral distance on processes, namely they quantify the disparities in the observations that we can make on them.

A 1-bounded *pseudometric* on a set X is a function $d: X \times X \rightarrow [0, 1]$ such that: (i) $d(x, x) = 0$, (ii) $d(x, y) = d(y, x)$, and (iii) $d(x, y) \leq d(x, z) + d(z, y)$, for all $x, y, z \in X$. The *kernel* of a pseudometric d on X consists in the set of the pairs of elements in X that are at distance 0, namely $\ker(d) = \{(x, y) \in X \times X \mid d(x, y) = 0\}$. For simplicity, we denote by $\mathcal{D}(X)$ the set of 1-bounded pseudometrics on the set X .

As elsewhere in the literature, we will sometimes use the term *metric* in place of pseudometric. Below we recall the notion of *bisimilarity metric*, which is one of the most studied behavioral metrics in the literature, and we discuss its compositionality.

The notion of *bisimulation metric* [6,13,14] bases on the quantitative analogous to the *bisimulation game*: two processes can be at some given distance $\varepsilon < 1$ only if they can mimic each other's transitions and evolve to distributions that are, in turn, at a distance $\leq \varepsilon$. To formalize this intuition, we need to lift pseudometrics on processes to pseudometrics on distributions. To this purpose, we rely on the notions of *matching* (also known as *coupling* or *weight function*) and *Kantorovich lifting*.

Definition 6 (Matching). *Assume two sets X and Y . A matching for distributions $\pi \in \Delta(X)$ and $\pi' \in \Delta(Y)$ is a distribution over the product space $\mathfrak{w} \in \Delta(X \times Y)$ with π and π' as left and right marginal, namely: (i) $\sum_{y \in Y} \mathfrak{w}(x, y) = \pi(x)$, for all $x \in X$, and (ii) $\sum_{x \in X} \mathfrak{w}(x, y) = \pi'(y)$, for all $y \in Y$. We let $\mathfrak{M}(\pi, \pi')$ denote the set of all matchings for π and π' .*

Definition 7 (Kantorovich metric, [21]). *Given $d \in \mathcal{D}(X)$, the Kantorovich lifting of d is the pseudometric $\mathbf{K}(d): \Delta(X) \times \Delta(X) \rightarrow [0, 1]$ defined for all $\pi, \pi' \in \Delta(X)$ by*

$$\mathbf{K}(d)(\pi, \pi') = \min_{\mathfrak{w} \in \mathfrak{M}(\pi, \pi')} \sum_{x, y \in X} \mathfrak{w}(x, y) \cdot d(x, y).$$

Bisimulation metrics are normally parametric with respect to a *discount factor* allowing us to specify how much the distance of future transitions is mitigated [3,14]. Informally, any difference that can be observed only after a long sequence of computation steps does not have the same impact of the differences that can be witnessed at the beginning of the computation. In our context, the discount factor is a value $\lambda \in (0, 1]$ such that the distance arising at step n is mitigated by λ^n ($\lambda = 1$ means no discount).

Definition 8 (Bisimulation metric, [13]). Let $\lambda \in (0, 1]$. We say that $d \in \mathcal{D}(\mathbf{T}(\Sigma))$ is a bisimulation metric if for all $s, t \in \mathbf{T}(\Sigma)$ with $d(s, t) < 1$, we have that:

$$\forall s \xrightarrow{a} \pi_s \exists t \xrightarrow{a} \pi_t \text{ such that } \lambda \cdot \mathbf{K}(d)(\pi_s, \pi_t) \leq d(s, t).$$

Notice that any bisimulation metric is a pseudometric by definition. In [13] it was proved that the *smallest* bisimulation metric exists and it is called *bisimilarity metric*, denoted by \mathbf{b}^λ . Moreover, its kernel induces an equivalence relation that coincides with probabilistic bisimilarity, namely $\mathbf{b}^\lambda(s, t) = 0$ if and only if s and t are bisimilar [6, 13].

3.1 Compositionality

In order to specify and verify systems in a compositional manner, it is necessary that the behavioral semantics be compatible with all operators of the language that describe these systems. In the behavioral metric approach, we must guarantee that the distance between two composed systems $f(s_1, \dots, s_n)$ and $f(t_1, \dots, t_n)$ depends on the distance between all pairs of arguments s_i, t_i , so that the closer all pairs s_i and t_i the closer the composed systems $f(s_1, \dots, s_n)$ and $f(t_1, \dots, t_n)$.

Following [18, 19], we define compositionality properties for bisimulation metric relying on the notion of modulus of continuity for an operator with respect to a distance.

Definition 9 (Modulus of continuity). Let $P = (\Sigma, \mathcal{A}, R)$ be a PTSS, $f \in \Sigma$ some n -ary operator and $d \in \mathcal{D}(\mathbf{T}(\Sigma))$. Then, a mapping $\mathbf{m}: [0, 1]^n \rightarrow [0, 1]$ is a modulus of continuity for f with respect to d iff:

– for all processes $s_i, t_i \in T(\Sigma)$ we have:

$$d(f(s_1, \dots, s_n), f(t_1, \dots, t_n)) \leq \mathbf{m}(d(s_1, t_1), \dots, d(s_n, t_n)),$$

- \mathbf{m} is continuous at $(0, \dots, 0)$, i.e. $\lim_{(\varepsilon_1 \dots \varepsilon_n) \rightarrow (0 \dots 0)} \mathbf{m}(\varepsilon_1, \dots, \varepsilon_n) = \mathbf{m}(0, \dots, 0)$,
- $\mathbf{m}(0, \dots, 0) = 0$.

In [16, 18, 19] an operator is called *uniformly continuous* if this operator admits any modulus of continuity with respect to \mathbf{b}^λ . Intuitively, a uniformly continuous n -ary operator f ensures that for any non-zero bisimulation distance ε (understood as the admissible tolerance from the operational behavior of the composed processes $f(s_1, \dots, s_n)$ and $f(t_1, \dots, t_n)$) there are non-zero bisimulation distances ε_i such that the distance between the composed processes $f(s_1, \dots, s_n)$ and $f(t_1, \dots, t_n)$ is at most ε whenever the component t_i is in distance of at most ε_i from s_i .

Definition 10 (Uniformly continuous operator). Let $P = (\Sigma, \mathcal{A}, R)$ be a PTSS. We say that an n -ary operator $f \in \Sigma$ is

- non extensive if $\mathbf{m}(\varepsilon_1, \dots, \varepsilon_n) = \max_{i=1}^n \varepsilon_i$ is a modulus of continuity for f ;
- non expansive if $\mathbf{m}(\varepsilon_1, \dots, \varepsilon_n) = \sum_{i=1}^n \varepsilon_i$ is a modulus of continuity for f ;
- Lipschitz continuous if $\mathbf{m}(\varepsilon_1, \dots, \varepsilon_n) = L \cdot \sum_{i=1}^n \varepsilon_i$ is a modulus of continuity for f for some $L > 1$.

Non-extensiveness and non-expansiveness are the most studied notion of compositionality (see, e.g., [4, 10, 13–15, 17, 22, 34, 35]). However, they do not allow for reasoning on recursive processes [16, 19, 24]. Essentially, recursive operators are Lipschitz continuous but not non-expansive [16].

We conclude this section by showing that the PGSOS format is necessary to obtain uniform continuity properties of operators. In the following Example we show an operator f defined by a rule which does not respect the PGSOS format, and for which no modulus of continuity among those in Definition 10 can be established wrt. \mathbf{b}^λ .

Example 1 (Counterexample for PGSOS format). For simplicity of notation, in all the Examples we will let $a.t$ stand for $a.\delta_t$. Assume the prefix operator $a.\bigoplus_{i \in I} [p_i]$ in Table 1 and a PTSS containing the following rules

$$\frac{x \parallel_{\mathcal{A}} x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} f(\mu)} \quad \frac{}{\eta \xrightarrow{\surd} \delta_{\text{nil}}} \quad \frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu \quad a \neq \surd}{x \parallel_{\mathcal{A}} y \xrightarrow{a} \mu \parallel_{\mathcal{A}} \nu} \quad \frac{x \xrightarrow{a} \surd \quad y \xrightarrow{a} \surd}{x \parallel_{\mathcal{A}} y \xrightarrow{\surd} \delta_{\text{nil}}}$$

in which $\parallel_{\mathcal{A}}$ denotes the fully synchronous parallel composition operator, η denotes the *skip* process, *nil* is the processes that cannot perform any action and \surd is a special action in \mathcal{A} denoting successful termination. Clearly, the rule for f falls out of the PGSOS format since an arbitrary term different from a single process variable occurs as left hand side of the premise. Given any process $u \in \mathbf{T}(\Sigma)$ s.t. $\mathbf{b}^\lambda(u, \eta) = 1$, consider processes $s_1 = a.u$ and $t_1 = a.([1-\varepsilon]u \oplus [\varepsilon]\eta)$ and for all $k > 1$ define $s_k = a.s_{k-1}$ and $t_k = a.t_{k-1}$. Clearly, for all $k \geq 1$ we have $\mathbf{b}^\lambda(s_k, t_k) = \varepsilon \cdot \lambda^k$. We now aim at evaluating $\mathbf{b}^\lambda(f(s_k), f(t_k))$, for all $k \geq 1$. Since the target of the rule for f is of the form $f(\mu)$, where μ is derived from the synchronous parallel composition of the source variable with itself, we can infer that $f(s_k)$ via a sequence of k a -labeled transitions reaches a distribution that assigns probability 1 to the parallel composition of 2^k copies of process u . Similarly, after k a -labeled transitions $f(t_k)$ reaches a distribution assigning probability $(1-\varepsilon)^{2^k}$ to the process corresponding to the parallel composition of 2^k copies of u ; probability ε^{2^k} to the parallel composition of 2^k copies of η , and the remaining probability $1 - (1-\varepsilon)^{2^k} - \varepsilon^{2^k}$ to processes that cannot execute any action (being the parallel composition of copies of u and η). Therefore we get $\mathbf{b}^\lambda(f(s_k), f(t_k)) = \lambda^k(1 - (1-\varepsilon)^{2^k})$. This implies that for any k with $2^k > L$, $\mathbf{b}^\lambda(f(s), f(t))/\mathbf{b}^\lambda(s, t) = (1 - (1-\varepsilon)^{2^k})/\varepsilon > L$ holds for $s = s_k$, $t = t_k$ and $\varepsilon \in (0, (2^k - L)/(2^{k-1}(2^k - 1)))$. We infer that no Lipschitz constant can be found for f and thus none of the continuity properties in Definition 10 can be established for f .

4 The Non-extensiveness Format

In this section we provide a specification format ensuring that all operators in PTSSs respecting the syntactic constraints imposed by the format are non-extensive. Briefly, the *non-extensiveness format* admits only convex combinations of variables and constants as targets of the rules. In the following, we rely on Proposition 1 and without loss of generality we always assume that the targets of all PGSOS-rules are in normal form.

Definition 11 (Specification format for non-extensiveness). *A PTSS $P = (\Sigma, \mathcal{A}, R)$ is in non-extensiveness format if all rules $r \in R$ are PGSOS rules as in Definition 3 and $\Theta = \text{trg}(r)$ is of the form $\Theta = \sum_{j \in J} p_j \Theta_j$, with $\Theta_j \in \mathbf{V}_d \cup \delta \mathbf{V}_s \cup \Sigma^0$ for all $j \in J$.*

Then, all operators in Σ are called P -non-extensive.

Essentially, in each element Θ_j we have at most one occurrence of one process argument x_h or of one of its derivatives $\mu_{h,m}$, thus implying that Θ_j contributes to the distance $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n))$ by $\mathbf{b}^\lambda(s_h, t_h)$. As a consequence, we have that $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n)) \leq \sum_{j \in J} p_j \max_{i=1}^n \mathbf{b}^\lambda(s_i, t_i) = \max_{i=1}^n \mathbf{b}^\lambda(s_i, t_i)$.

We notice that all rules in Table 1 respect Definition 11. Conversely, for each operator defined in Tables 2 and 3 there is at least one rule violating the constraints in Definition 11. For instance, the target of the first rule for operator of sequentialization $x; y$ contains both the derivative μ of the process argument x and the process argument y . If we consider the first rule for the operator of parallel composition $x \parallel_B y$, we note that the target contains one derivative for each of the two process arguments.

To obtain the non-extensiveness property of bisimilarity metric from our format, we need to apply a quantitative analogous to the standard technique used to prove the congruence property of behavioural equivalences. Informally, we introduce the notion of *non-extensive closure* of a pseudometric, seen as the quantitative analogous to the congruence closure of behavioral relations. Briefly, the non-extensive closure allows us to lift a pseudometric to a pseudometric that satisfies the non-extensiveness property.

Definition 12 (Non-extensive closure). *Assume a PTSS (Σ, \mathcal{A}, R) . The non-extensive closure $\text{Ext} : \mathcal{D}(\mathbf{T}(\Sigma)) \rightarrow \mathcal{D}(\mathbf{T}(\Sigma))$ is defined for $d \in \mathcal{D}(\mathbf{T}(\Sigma))$ and $s, t \in \mathbf{T}(\Sigma)$ by*

$$\text{Ext}(d)(s, t) = \begin{cases} \min\{d(s, t), \max_{i=1, \dots, n} \text{Ext}(d)(s_i, t_i)\} & \text{if } s = f(s_1, \dots, s_n), \\ & t = f(t_1, \dots, t_n) \text{ and } f \in \Sigma \\ d(s, t) & \text{otherwise.} \end{cases}$$

In particular we denote by \mathbf{Ext}^λ the non-extensive closure of \mathbf{b}^λ .

Clearly, focusing on the non-extensive closure of bisimilarity metric \mathbf{Ext}^λ , we observe that $\mathbf{Ext}^\lambda(s, t) \leq \mathbf{b}^\lambda(s, t)$ for all $s, t \in \mathbf{T}(\Sigma)$. Moreover, notice that each operator f admits its modulus of continuity for non-extensiveness wrt. \mathbf{Ext}^λ . Therefore, the proof of the *non-extensiveness theorem* (Theorem 1 below), that states the correctness of the non-extensiveness format (Definition 11), consists in proving that under the constraints of our format the metric \mathbf{Ext}^λ is itself a bisimulation metric. As a consequence, from $\mathbf{Ext}^\lambda(s, t) \leq \mathbf{b}^\lambda(s, t)$ and \mathbf{b}^λ being the least bisimulation metric, we can infer that $\mathbf{Ext}^\lambda = \mathbf{b}^\lambda$ and thus that \mathbf{b}^λ is non-extensive wrt. all the operators satisfying our non-extensiveness format.

Theorem 1 (Non-extensiveness). *Let $\lambda \in (0, 1]$ and $P = (\Sigma, \mathcal{A}, R)$ be a PTSS in non-extensiveness format. Then the operators in Σ are non-extensive wrt. \mathbf{b}^λ .*

Proof sketch. We present only a sketch of the proof. The proofs of Theorems 2 and 3 follow the same reasoning schema.

As $\mathbf{Ext}^\lambda \leq \mathbf{b}^\lambda$ by definition and \mathbf{b}^λ is the least bisimulation metric, the thesis is equivalent to prove that \mathbf{Ext}^λ is a bisimulation metric. We need to show that

$$\begin{aligned} & \text{whenever } \mathbf{Ext}^\lambda(\sigma(t), \sigma'(t)) < 1 \text{ and } P \vdash \sigma(t) \xrightarrow{a} \pi \\ & \text{then } P \vdash \sigma'(t) \xrightarrow{a} \pi' \text{ and } \lambda \cdot \mathbf{K}(\mathbf{Ext}^\lambda)(\pi, \pi') \leq \mathbf{Ext}^\lambda(\sigma(t), \sigma'(t)). \end{aligned} \quad (2)$$

This is proved by induction over the structure of t .

For the non-trivial inductive step with $t = f(t_1, \dots, t_n)$ and $\mathbf{Ext}^\lambda(\sigma(t), \sigma'(t)) = \max_{i=1, \dots, n} \mathbf{Ext}^\lambda(\sigma(t_i), \sigma'(t_i))$ the proof can be sketched as follows. Assume that $P \vdash \sigma(t) \xrightarrow{a} \pi$, so that there are a non-extensive rule

$$r = \frac{\{x_h \xrightarrow{\alpha_{h,m}} \mu_{h,m} \mid h \in H, m \in M_h\} \{x_h \xrightarrow{\alpha_{h,n}} \nu_{h,n} \mid h \in H, n \in N_h\}}{f(x_1, \dots, x_n) \xrightarrow{a} \sum_{j \in J} p_j \Theta_j}$$

and a closed substitution σ_1 , such that $\sigma_1(x_i) = \sigma(x_i)$ for all $i = 1, \dots, n$ and $\sigma_1(\sum_{j \in J} p_j \Theta_j) = \pi$, from which we can derive such a transition.

For each $h \in H, m \in M_h$, the inductive hypothesis gives that $\sigma'(t_h) \xrightarrow{\alpha_{h,m}} \pi_{h,m}$ and $\lambda \cdot \mathbf{K}(\mathbf{Ext}^\lambda)(\sigma_1(\mu_{h,m}), \pi_{h,m}) \leq \mathbf{Ext}^\lambda(\sigma(t_h), \sigma'(t_h))$. For the closed substitution σ_2 with $\sigma_2(x_i) = \sigma'(t_i)$ for all $i = 1, \dots, n$ and $\sigma_2(\mu_{h,m}) = \pi_{h,m}$ for all $h \in H, m \in M_h$, we get that $\{\sigma_2(x_h) \xrightarrow{\alpha_{h,m}} \sigma_2(\mu_{h,m}) \mid h \in H, m \in M_h\}$ is provable from P . For each $h \in H, n \in N_h$, by the inductive hypothesis we can infer that $\sigma'(t_h) \xrightarrow{\alpha_{h,n}} \nu_{h,n}$ for all $h \in H, n \in N_h$. For σ_2 defined as above, we get that the set of negative premises $\{\sigma_2(x_h) \xrightarrow{\alpha_{h,n}} \nu_{h,n} \mid h \in H, n \in N_h\}$ is provable from P . Hence, all the premises of the closed instance of r with respect to σ_2 are provable from P , thus implying that $\sigma'(t) \xrightarrow{a} \sigma'(\sum_{j \in J} p_j \Theta_j)$ is provable from P .

Then, $\lambda \cdot \mathbf{K}(\mathbf{Ext}^\lambda)(\sum_{j \in J} p_j \sigma_1(\Theta_j), \sum_{j \in J} p_j \sigma_2(\Theta_j)) \leq \mathbf{Ext}^\lambda(\sigma(t), \sigma'(t))$ follows by: 1. the convexity properties of the Kantorovich metric, 2. the form of the Θ_j in the target of a non-extensive rule, and 3. the inductive hypothesis. \square

Table 1. Non-extensive process algebra operators. $\mu \oplus_p \nu$ stays for $p \cdot \mu + (1 - p) \cdot \nu$.

$$\begin{array}{c}
 \frac{}{a. \bigoplus_{i=1}^n [p_i] x_i \xrightarrow{a} \sum_{i=1}^n p_i \delta_{x_i}} \qquad \frac{x \xrightarrow{a} \mu}{x + y \xrightarrow{a} \mu} \qquad \frac{y \xrightarrow{a} \nu}{x + y \xrightarrow{a} \nu} \\
 \\
 \frac{x \xrightarrow{a} \mu \quad y \not\xrightarrow{a}}{x +_p y \xrightarrow{a} \mu} \qquad \frac{x \not\xrightarrow{a} \quad y \xrightarrow{a} \nu}{x +_p y \xrightarrow{a} \nu} \qquad \frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu}{x +_p y \xrightarrow{a} \mu \oplus_p \nu}
 \end{array}$$

By applying Theorem 1, we get the non-extensiveness of all operators in Table 1.

Corollary 1. *The operators defined by the PGSOS rules in Table 1 are non-extensive.*

Since we already know from [16] that none of the operators in Tables 2 and 3 is non-extensive, we can conclude that our format is not too restrictive, meaning that no non-extensive operator among those used in standard probabilistic processes algebras is out of the format. Moreover, we can also infer that the constraints of the format cannot be relaxed in any immediate way.

5 The Non-expansiveness Format

In this section we focus on the non-expansiveness property. We show that by relaxing the constraints of the non-extensiveness format, we obtain a specification format for non-expansiveness. In detail, we allow the target of each rule to be a convex combination of distribution terms in which also non-constant operators can appear and can be applied to source variables and to their derivatives, provided that for each source variables at most one occurrence of that variable or of its derivatives occurs.

Definition 13 (Specification format for non-expansiveness). *Assume a PTSS $P_1 = (\Sigma_1, \mathcal{A}, R_1)$ in non-extensiveness format. A PTSS $P_2 = (\Sigma_2, \mathcal{A}, R_2)$ with $P_1 \sqsubseteq P_2$ is in non-expansiveness format if all rules $r \in R_2 \setminus R_1$ are of the form*

$$\frac{\{x_h \xrightarrow{\alpha_{h,m}} \mu_{h,m} \mid h \in H, m \in M_h\} \quad \{x_h \not\xrightarrow{\alpha_{h,n}} \mid h \in H, n \in N_h\}}{f(x_1, \dots, x_n) \xrightarrow{\alpha} \Theta}$$

where:

1. the target Θ is of the form $\Theta = \sum_{j \in J} p_j \Theta_j$,
2. for all $j \in J$ and $h \in H$, at most one among the variables in $(\{x_h\} \cup \{\mu_{h,m} \mid m \in M_h\})$ occurs in Θ_j and there must be at most one such occurrence, and
3. for all $j \in J$ and $i \in \{1, \dots, n\} \setminus H$, the variable x_i can occur at most once in Θ_j .

Moreover, all operators defined by rules in $R_2 \setminus R_1$ are called P_2 -non-expansive.

Since in each Θ_j we can have at most one occurrence of a source variable x_h or of one of its derivatives $\{\mu_{h,m} \mid m \in M_h\}$, we can infer that Θ_j contributes to $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n))$ by, at most, $\sum_{i=1}^n \mathbf{b}^\lambda(s_i, t_i)$. Consequently, we obtain $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n)) \leq \sum_{j \in J} p_j \sum_{i=1}^n \mathbf{b}^\lambda(s_i, t_i) = \sum_{i=1}^n \mathbf{b}^\lambda(s_i, t_i)$.

Notice that all non-extensive operators are also non-expansive. Further, all operators defined by rules in Table 2, as well as those in Table 1, respect the non-expansiveness format. Conversely, the operators in Table 3, which are known to be expansive [16], have at least one rule falling out of the format. For instance, the rule for the bang operator $!x$, which stands for the parallel composition of infinitely many copies of the process argument, violates Definition 13 since both the source variable x and its derivative μ occur in the same distribution term in the target.

As in the case of non-extensiveness, the formal proof of the correctness of the non-expansiveness format bases on a proper notion of closure of a pseudometric. We define the *non-expansiveness closure* of a pseudometric as its lifting to a pseudometric which is non-extensive on operators satisfying Definition 11, and non-expansive on the others.

Definition 14 (Non-expansive closure). Assume the PTSSs $P_1 = (\Sigma_1, \mathcal{A}, R_1)$ in non-extensiveness format and $P_2 = (\Sigma_2, \mathcal{A}, R_2)$ with $P_1 \sqsubseteq P_2$. The non-expansive closure $\text{Exp}: \mathcal{D}(\mathbf{T}(\Sigma_2)) \rightarrow \mathcal{D}(\mathbf{T}(\Sigma_2))$ is defined for $d \in \mathcal{D}(\mathbf{T}(\Sigma_2))$ and $s, t \in \mathbf{T}(\Sigma_2)$ by

$$\text{Exp}(d)(s, t) = \begin{cases} \min\{d(s, t), \max_{i=1, \dots, n} \text{Exp}(d)(s_i, t_i)\} & \text{if } s = f(s_1, \dots, s_n), \\ & t = f(t_1, \dots, t_n), f \in \Sigma_1 \\ \min\{d(s, t), \sum_{i=1}^n \text{Exp}(d)(s_i, t_i)\} & \text{if } s = f(s_1, \dots, s_n), \\ & t = f(t_1, \dots, t_n), f \in \Sigma_2 \setminus \Sigma_1 \\ d(s, t) & \text{otherwise.} \end{cases}$$

In particular we denote by \mathbf{Exp}^λ the non-expansive closure of \mathbf{b}^λ .

We can then show that under the constraints of Definition 13, the non-expansive closure \mathbf{Exp}^λ of \mathbf{b}^λ is indeed a bisimulation metric. Since by definition $\mathbf{Exp}^\lambda(s, t) \leq \mathbf{b}^\lambda(s, t)$ for all $s, t \in \mathbf{T}(\Sigma)$ and \mathbf{b}^λ is the least bisimulation metric, we can conclude that $\mathbf{Exp}^\lambda = \mathbf{b}^\lambda$ and the operators in a PTSS in non-expansiveness format are non-expansive with respect to \mathbf{b}^λ , as stated in Theorem 2 below.

Theorem 2 (Non-expansiveness). Let $\lambda \in (0, 1]$ and (Σ, \mathcal{A}, R) be a PTSS in non-expansiveness format. Then the operators in Σ are non-expansive with respect to \mathbf{b}^λ .

By applying Theorem 2 we get the non-expansiveness of all operators in Table 2.

Corollary 2. *The operators defined by the PGSOS rules in Tables 1 and 2 are non-expansive with respect to \mathbf{b}^λ .*

Since we already know from [16] that none of the operators in Table 3 is non-expansive, we can conclude that our format is not too restrictive, meaning that no non-expansive operator among those used in standard probabilistic processes algebras is out of the format. Moreover, the following examples show that the constraints in Definition 13 cannot be relaxed in any trivial way.

Example 2 (Counterexample for Definition 13.2, I). Consider (i) $s_1 = a.([1/2]u_1 \oplus [1/2]u_2)$; (ii) $t_1 = a.([1/3]u_1 \oplus [1/3]u_2 \oplus [1/3]u_3)$; (iii) $u_1 = b.\eta$; (iv) $u_2 = c.\eta$; (v) $u_3 = u_1 + u_2$. Clearly we have $\mathbf{b}^\lambda(s_1, t_1) = 1/3 \cdot \lambda$, due to u_3 . Let \parallel_\emptyset denote \parallel_\emptyset . Consider the operator f defined by the following PGSOS rule

$$\frac{x \xrightarrow{a} \mu}{f(x) \xrightarrow{a} \delta_x \parallel \mu}$$

which violates the non-expansiveness format of Definition 13 since both x and μ occur in the target. Let us evaluate $\mathbf{b}^\lambda(f(s_1), f(t_1))$. Let $\pi_{s_1} = 1/2\delta(s_1 \parallel u_1) + 1/2\delta(s_1 \parallel u_2)$ and $\pi_{t_1} = 1/3\delta(t_1 \parallel u_1) + 1/3\delta(t_1 \parallel u_2) + 1/3\delta(t_1 \parallel u_3)$. We have $f(s_1) \xrightarrow{a} \pi_{s_1}$, $f(t_1) \xrightarrow{a} \pi_{t_1}$ and $\mathbf{b}^\lambda(f(s_1), f(t_1)) \geq \lambda \cdot \mathbf{K}(\mathbf{b}^\lambda)(\pi_{s_1}, \pi_{t_1})$. Then $\mathbf{b}^\lambda(s_1 \parallel u_1, t_1 \parallel u_3) = 1 = \mathbf{b}^\lambda(s_1 \parallel u_2, t_1 \parallel u_3)$, whereas $\mathbf{b}^\lambda(s_1 \parallel u_1, t_1 \parallel u_1) = \mathbf{b}^\lambda(s_1 \parallel u_2, t_1 \parallel u_2) = \mathbf{b}^\lambda(s_1, t_1)$. Hence, $\mathbf{b}^\lambda(f(s_1), f(t_1)) = 1/3 \cdot \lambda + 2/9 \cdot \lambda^2 > \mathbf{b}^\lambda(s_1, t_1)$.

Example 3 (Counterexample for Definition 13.2, II). Consider process u_1 from Example 2 and processes $s_2 = a.([1/2]u_1 \oplus [1/2]\eta)$ and $t_2 = a.u_1$. Clearly, $\mathbf{b}^\lambda(s_2, t_2) = 1/2 \cdot \lambda$. Consider the following PGSOS rule defining the behavior of operator g

$$\frac{x \xrightarrow{a} \mu}{g(x) \xrightarrow{a} \mu \parallel_{\mathcal{A}} \mu}$$

which violates the non-expansiveness format as variable μ occurs in the right hand side of premise and twice in the target. Let us evaluate $\mathbf{b}^\lambda(g(s_2), g(t_2))$. We have that $g(s_2) \xrightarrow{a} 1/4\delta(u_1 \parallel_{\mathcal{A}} u_1) + 1/4\delta(\eta \parallel_{\mathcal{A}} \eta) + 1/4\delta(u_1 \parallel_{\mathcal{A}} \eta) + 1/4\delta(\eta \parallel_{\mathcal{A}} u_1)$. Conversely, $g(t_2) \xrightarrow{a} \delta(u_1 \parallel_{\mathcal{A}} u_1)$. Therefore, since $\parallel_{\mathcal{A}}$ corresponds to the fully synchronous parallel composition, we can directly infer that $\mathbf{b}^\lambda(g(s_2), g(t_2)) \geq 3/4 \cdot \lambda > \mathbf{b}^\lambda(s_2, t_2)$.

Example 4 (Counterexample for Definition 13.3). Consider process u_1 from Example 2 and processes s_2, t_2 from Example 3. Consider the following PGSOS rule defining the behavior of operator h

$$\frac{x \xrightarrow{b} \mu}{h(x, y) \xrightarrow{b} \delta_y \parallel_{\mathcal{A}} \delta_y}$$

Table 2. Non-expansive process algebra operators. The symmetric version of the first rule for \parallel_B and the one for \parallel_p have been omitted.

$$\begin{array}{c}
\frac{}{\eta \xrightarrow{\checkmark} \delta_{\text{nil}}} \quad \frac{x \xrightarrow{a} \mu \quad a \neq \checkmark}{x; y \xrightarrow{a} \mu; \delta_y} \quad \frac{x \xrightarrow{\checkmark} \mu \quad y \xrightarrow{a} \nu}{x; y \xrightarrow{a} \nu} \\
\\
\frac{x \xrightarrow{a} \mu \quad a \notin B \cup \{\checkmark\}}{x \parallel_B y \xrightarrow{a} \mu \parallel_B \delta_y} \quad \frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu \quad a \in B \setminus \{\checkmark\}}{x \parallel_B y \xrightarrow{a} \mu \parallel_B \nu} \quad \frac{x \xrightarrow{\checkmark} \mu \quad y \xrightarrow{\checkmark} \nu}{x \parallel_B y \xrightarrow{\checkmark} \delta_{\text{nil}}} \\
\\
\frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \checkmark \quad a \neq \checkmark}{x \parallel_p y \xrightarrow{a} \mu \parallel_p \delta_y} \quad \frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu \quad a \neq \checkmark}{x \parallel_p y \xrightarrow{a} \mu \parallel_p \delta_y \oplus_p \delta_x \parallel_p \nu} \quad \frac{x \xrightarrow{\checkmark} \mu \quad y \xrightarrow{\checkmark} \nu}{x \parallel_p y \xrightarrow{\checkmark} \delta_{\text{nil}}}
\end{array}$$

which violates the non-expansiveness format in Definition 13 as variable y occurs twice in the target. From similar calculations to those applied in Example 3, we can infer that $\mathbf{b}^\lambda(h(u_1, s_2), h(u_1, t_2)) \geq 3/4 \cdot \lambda^2$, which is greater than $\mathbf{b}^\lambda(s_2, t_2)$ for $\lambda > 2/3$.

6 The Lipschitz Continuity Format

We dedicate this section to the definition of a specification *format for Lipschitz continuity*. This is obtained by relaxing the constraints of the non-expansiveness format in Definition 13 as follows: firstly, for each operator f we fix $n_f, m_f \in \mathbb{N}$. Then, we allow the targets of the rules for f to be convex combinations of distribution terms in which (i) the number of occurrences of source variables and their derivatives is bounded by n_f ; (ii) the number of nested operators in which a source variable can occur is bounded by m_f ; (iii) derivatives can only occur in the scope of non-expansive operators.

Definition 15 (Specification format for Lipschitz continuity). *Assume a PTSS $P_1 = (\Sigma_1, \mathcal{A}, R_1)$ in non-expansiveness format. A PTSS $P_2 = (\Sigma_2, \mathcal{A}, R_2)$ with $P_1 \sqsubseteq P_2$ is in Lipschitz continuity format if for all operators $f \in \Sigma_2 \setminus \Sigma_1$ there are naturals n_f, m_f and all rules $r \in R_2 \setminus R_1$ are of the form*

$$\frac{\{x_h \xrightarrow{\alpha_{h,m}} \mu_{h,m} \mid h \in H, m \in M_h\} \quad \{x_h \xrightarrow{\alpha_{h,n}} \checkmark \mid h \in H, n \in N_h\}}{f(x_1, \dots, x_n) \xrightarrow{\alpha} \Theta}$$

where:

1. the target Θ is of the form $\Theta = \sum_{j \in J} p_j \Theta_j$,
2. for all $j \in J$ and $i \in \{1, \dots, n\}$, x_i occurs at most n_f times in Θ_j and such occurrences appear in the scope of at most m_f operators, and
3. for all $j \in J$ and $h \in H$, there are at most n_f occurrences of variables in $\{\mu_{h,m} \mid m \in M_h\}$ in Θ_j , which appear only in the scope of non-expansive operators in Σ_1 .

Moreover, all operators defined by rules in $R_2 \setminus R_1$ are called P_2 -Lipschitz.

Informally, for each $h \in H$, bounding to n_f the number of occurrences of derivatives $\{\mu_{h,m} \mid m \in M_h\}$ that may occur in Θ_j and requiring them in the scope of non-expansive operators, guarantees that process argument x_h and Θ_j contribute to $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n))$ by, at most, $n_f \cdot \mathbf{b}^\lambda(s_h, t_h)$. Considering instead source variables, constraint 2 of Definition 15 guarantees that for each $i = 1, \dots, n$ there is a constant $L_i \geq 1$ such that process argument x_i and Θ_j contribute to distance $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n))$ by, at most, $\lambda \cdot n_f \cdot L_i^{m_f} \cdot \mathbf{b}^\lambda(s_i, t_i)$, where L_i is the maximal Lipschitz constant among those of operators applied on top of x_i , and factor λ arises since x_i is delayed by one computation step. Therefore, we can infer that for each $j \in J$ there is a constant $L_j \geq n_f$ such that Θ_j contributes to $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n))$ by, at most, $L_j \sum_{i=1}^n \mathbf{b}^\lambda(s_i, t_i)$. Finally, we obtain $\mathbf{b}^\lambda(f(s_1, \dots, s_n), f(t_1, \dots, t_n)) \leq \sum_{j \in J} p_j L_j \cdot \sum_{i=1}^n \mathbf{b}^\lambda(s_i, t_i) = L \cdot \sum_{i=1}^n \mathbf{b}^\lambda(s_i, t_i)$ for some constant $L \geq 1$.

Notice that the rules in Table 3, as those in Tables 1 and 2, respect the constraints in Definition 15. We observe that Table 3 contains the specification of standard recursive operators x^ω , $!x$ and x^*y which were ruled out by Definitions 11 and 13.

Next, we formalize the definition of a *Lipschitz factor* for an operator f , namely the constant L_f wrt. which f is Lipschitz continuous, and its construction over the rules for f and their targets. Since this construction is recursive for recursive operators, in order to have well-defined Lipschitz factors for these operators we need a discount factor strictly less than one.

Definition 16 (Lipschitz factor). Assume $\lambda \in (0, 1)$, a PTSS $P_1 = (\Sigma_1, \mathcal{A}, R_1)$ in non-extensiveness format and a PTSS $P_2 = (\Sigma_2, \mathcal{A}, R_2)$ with $P_1 \sqsubseteq P_2$. For each operator $f \in \Sigma_2 \setminus \Sigma_1$ let $R_f \subseteq R_2$ denote the set of rules defining the behavior. Then, the set of Lipschitz factors of operators in $\Sigma_2 \setminus \Sigma_1$ is defined as the set of minimum reals $L_f \geq 1$ satisfying the system of inequalities of the form

$$L_f \geq \max_{r \in R_f} \max_{x \in \text{src}(r)} \left\{ \ell_{\text{trg}(r)}(x) + \sum_{\mu \in \text{der}(x,r)} \ell_{\text{trg}(r)}(\mu) \right\}$$

for each $f \in \Sigma_2 \setminus \Sigma_1$ where, for any distribution term $\Theta \in \text{DT}(\Sigma)$ in normal form, the factor ℓ_Θ is defined inductively as follows

$$\ell_\Theta(x) = \begin{cases} 0 & \text{if } x \notin \text{var}(\Theta) \\ \lambda & \text{if } \Theta = \delta_x \\ \max_{i=1, \dots, n} \ell_{\Theta_i}(x) & \text{if } \Theta = f(\Theta_1, \dots, \Theta_n) \text{ and } f \in \Sigma_1 \\ L_f \sum_{i=1}^n \ell_{\Theta_i}(x) & \text{if } \Theta = f(\Theta_1, \dots, \Theta_n) \text{ and } f \in \Sigma_2 \setminus \Sigma_1 \\ \sum_{j \in J} p_j \ell_{\Theta_j}(x) & \text{if } \Theta = \sum_{j \in J} p_j \Theta_j \end{cases}$$

$$\ell_{\Theta}(\mu) = \begin{cases} 0 & \text{if } \mu \notin \text{var}(\Theta) \\ 1 & \text{if } \Theta = \mu \\ \max_{i=1, \dots, n} \ell_{\Theta_i}(\mu) & \text{if } \Theta = f(\Theta_1, \dots, \Theta_n) \text{ and } f \in \Sigma_1 \\ L_f \sum_{i=1}^n \ell_{\Theta_i}(\mu) & \text{if } \Theta = f(\Theta_1, \dots, \Theta_n) \text{ and } f \in \Sigma_2 \setminus \Sigma_1 \\ \sum_{j \in J} p_j \ell_{\Theta_j}(\mu) & \text{if } \Theta = \sum_{j \in J} p_j \Theta_j \end{cases}$$

The following proposition states that Lipschitz factors are well-defined if the PTSS is in Lipschitz format and $\lambda \in (0, 1)$. Moreover, $L_f = 1$ if f is non-expansive.

Proposition 2. *Let $\lambda \in (0, 1)$. Assume a PTSS $P_1 = (\Sigma_1, \mathcal{A}, R_1)$ in non-extensiveness format and an extension $P_1 \sqsubseteq P_2 = (\Sigma_2, \mathcal{A}, R_2)$ in Lipschitz format. Then:*

- for each operator $f \in \Sigma_2$, the Lipschitz factor $L_f < +\infty$ is well-defined.
- moreover, if $f \in \Sigma_2$ is non-expansive then $L_f = 1$.

The following two examples show that if P_2 falls out of the Lipschitz continuity format, then Proposition 2 does not hold.

Example 5. Consider operators f and g specified by the following rules for all $k \in \mathbb{N}$:

$$\frac{x \xrightarrow{a_k} \mu}{f(x) \xrightarrow{a_k} \underbrace{\mu \parallel_{\mathcal{A}} \dots \parallel_{\mathcal{A}} \mu}_{k\text{-times}}} \quad \frac{x \xrightarrow{a_k} \mu}{g(x) \xrightarrow{a_k} \underbrace{\delta(h(\dots h(x)))}_{k\text{-times}}} \quad \frac{x \xrightarrow{a_k} \mu}{h(x) \xrightarrow{a_k} \mu \parallel_{\mathcal{A}} \mu}$$

Notice that rules for both f and g are out of the Lipschitz format since we cannot define any n_f and m_g accordingly to Definition 15. Then, the inductive definition of Lipschitz factors in Definition 16 gives that $L_{\parallel_{\mathcal{A}}} = 1$, $L_h = 2$ and, then, $L_f \geq k$ and $L_g \geq \lambda \cdot 2^k$ for all $k \in \mathbb{N}$, thus implying that $L_f = L_g = +\infty$.

Example 6. Consider the *copy* operator [5] defined by the PGSOS rules below

$$\frac{x \xrightarrow{a} \mu \quad a \notin \{1, r\}}{\text{cp}(x) \xrightarrow{a} \mu} \quad \frac{x \xrightarrow{1} \mu \quad x \xrightarrow{r} \nu}{\text{cp}(x) \xrightarrow{s} \text{cp}(\mu) \parallel_{\mathcal{A}} \text{cp}(\nu)}.$$

Firstly, notice that the second rule violates the non-expansiveness format since we have two occurrences of derivatives of x in the same distribution term (Definition 13.2). Consequently, it also violates the Lipschitz continuity format in that μ and ν occur in the scope of an expansive operator (Definition 15.3). By Definition 16 we would have $L_{\text{cp}} \geq L_{\parallel_{\mathcal{A}}} \cdot L_{\text{cp}} + L_{\parallel_{\mathcal{A}}} \cdot L_{\text{cp}} = 2 \cdot L_{\text{cp}}$ which is satisfied only by $L_{\text{cp}} = +\infty$.

Then, the following example shows that, due to the presence of recursive operators, we must require λ to be strictly lesser than one to guarantee the validity of Proposition 2.

Example 7. Consider the rule for the bang operator $!_n$ in Table 3. By Definition 16 we obtain that $L_! \geq \lambda \cdot L_! + 1$, thus giving $L_! \geq \frac{1}{1-\lambda}$, which does not hold for $\lambda = 1$.

We remark that by restricting to non-recursive operators, such as x^n and $!^n$ in Table 3, Proposition 2 would hold also for $\lambda = 1$.

As one can expect, the proof of the correctness of our Lipschitz continuity format bases on the notion of *Lipschitz closure* of a pseudometric that lifts it to a pseudometric which is non-extensive on non-extensive operators, non-expansive on non-expansive operators and Lipschitz continuous on Lipschitz operators.

Definition 17 (Lipschitz closure). Let $\lambda \in (0, 1)$. Assume a PTSS $P_1 = (\Sigma_1, \mathcal{A}, R_1)$ in non-extensiveness format and a PTSS $P_2 = (\Sigma_2, \mathcal{A}, R_2)$ with $P_1 \sqsubseteq P_2$. Then, the Lipschitz closure $\text{Lip}: \mathcal{D}(\mathbf{T}(\Sigma_2)) \rightarrow \mathcal{D}(\mathbf{T}(\Sigma_2))$ is defined for all $d \in \mathcal{D}(\mathbf{T}(\Sigma_2))$ and $s, t \in \mathbf{T}(\Sigma_2)$ by

$$\text{Lip}(d)(s, t) = \begin{cases} \min\{d(s, t), \max_{i=1, \dots, n} \text{Lip}(d)(s_i, t_i)\} & \text{if } s = f(s_1, \dots, s_n), \\ & t = f(t_1, \dots, t_n), f \in \Sigma_1 \\ \min\{d(s, t), L_f \cdot \sum_{i=1}^n \text{Lip}(d)(s_i, t_i)\} & \text{if } s = f(s_1, \dots, s_n), \\ & t = f(t_1, \dots, t_n), f \in \Sigma_2 \setminus \Sigma_1 \\ d(s, t) & \text{otherwise.} \end{cases}$$

In particular we denote by Lip^λ the Lipschitz closure of \mathbf{b}^λ .

We can show that on a PTSS respecting Definition 15 the Lipschitz closure Lip^λ of \mathbf{b}^λ is a bisimulation metric and thus the operators in the PTSS are Lipschitz continuous wrt. \mathbf{b}^λ . In particular, the constraints of the Lipschitz continuity format combined with the definition of the Lipschitz factor L_f of operator f (Definition 16) guarantee that f is L_f -Lipschitz wrt. \mathbf{b}^λ .

Theorem 3 (Lipschitz continuity). Let (Σ, \mathcal{A}, R) be a PTSS in Lipschitz continuity format and let $\lambda \in (0, 1)$. Then the operators in Σ are Lipschitz continuous wrt. \mathbf{b}^λ . In particular, for a Lipschitz operator f , we have that f is L_f -Lipschitz wrt. \mathbf{b}^λ .

As an application of Theorem 3, we obtain that \mathbf{b}^λ is compositional wrt. all the operators defined in Table 3 in the sense of Lipschitz continuity.

Corollary 3. Let $\lambda \in (0, 1)$. The operators defined by the PGSOS rules in Tables 1, 2 and 3 are Lipschitz continuous with respect to \mathbf{b}^λ .

Again, we remark that in Corollary 3 the condition λ strictly lesser than one is required by the operators x^ω , $!x$ and x^*y .

Since all operators that have been proved in [16] to be Lipschitz continuous are in Table 3, we can infer that our format is not too restrictive as it does not exclude any known Lipschitz continuous operator. Moreover, the following two examples show that the constraint of Definition 15 cannot be trivially relaxed.

Table 3. Lipschitz continuous process algebra operators. We recall that $\|$ denotes $\|\cdot\|_\emptyset$.

$$\begin{array}{c}
\frac{x \xrightarrow{a} \mu \quad a \neq \surd}{x^{n+1} \xrightarrow{a} \mu; \delta_{x^n}} \quad \frac{x \xrightarrow{\surd} \mu}{x^{n+1} \xrightarrow{\surd} \mu} \quad \frac{}{x^0 \xrightarrow{\surd} \delta_{\text{nil}}} \quad \frac{x \xrightarrow{\surd} \mu \quad x \xrightarrow{a} \nu \quad a \neq \surd \quad n > m}{x^n \xrightarrow{a} \nu; \delta_{x^m}} \\
\frac{x \xrightarrow{a} \mu \quad a \neq \surd}{x^\omega \xrightarrow{a} \mu; \delta_{x^\omega}} \quad \frac{x \xrightarrow{a} \mu \quad a \neq \surd}{x^*y \xrightarrow{a} \mu; \delta_{x^*y}} \quad \frac{y \xrightarrow{a} \nu}{x^*y \xrightarrow{a} \nu} \\
\frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu \quad a \neq \surd}{x^{*p}y \xrightarrow{a} \nu \oplus_p (\mu; \delta_{x^{*p}y})} \quad \frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu \quad a \neq \surd}{x^{*p}y \xrightarrow{a} \mu; \delta_{x^{*p}y}} \quad \frac{x \xrightarrow{a} \mu \quad y \xrightarrow{a} \nu \quad a \neq \surd}{x^{*p}y \xrightarrow{a} \nu} \quad \frac{y \xrightarrow{\surd} \nu}{x^{*p}y \xrightarrow{\surd} \nu} \\
\frac{x \xrightarrow{a} \mu \quad a \neq \surd}{!^{n+1}x \xrightarrow{a} \mu \parallel \delta_{!^n x}} \quad \frac{x \xrightarrow{\surd} \mu}{!^{n+1}x \xrightarrow{\surd} \mu} \quad \frac{}{!^0x \xrightarrow{\surd} \delta_{\text{nil}}} \\
\frac{x \xrightarrow{a} \mu \quad a \neq \surd}{!x \xrightarrow{a} \mu \parallel \delta_{!x}} \quad \frac{x \xrightarrow{a} \mu \quad a \neq \surd}{!_p x \xrightarrow{a} \mu \oplus_p (\mu \parallel \delta_{!_p x})}
\end{array}$$

Example 8 (Counterexample for bounded number of occurrences). Assume $\mathcal{A} = \{a_k \mid k \in \mathbb{N}\}$. Consider the operators f and g as in Example 5, where we already noticed that the rules defining them violate Definition 15. Let $\varepsilon \in (0, 1)$. Following [19], consider processes $u_k = a_k.a_k.\text{nil}$ and $v_k = a_k.([1-\varepsilon]a_k.\text{nil} \oplus [\varepsilon]\text{nil})$ with $k \in \mathbb{N}$. Clearly, $\mathbf{b}^\lambda(u_k, v_k) = \lambda\varepsilon$ for all $k \in \mathbb{N}$.

Consider first operator f . We have $f(u_k) \xrightarrow{a_k} \pi_{u,k}$ and $f(v_k) \xrightarrow{a_k} \pi_{v,k}$, with $\pi_{u,k}(a_k.\text{nil} \parallel_{\mathcal{A}} \dots \parallel_{\mathcal{A}} a_k.\text{nil}) = 1$, $\pi_{v,k}(a_k.\text{nil} \parallel_{\mathcal{A}} \dots \parallel_{\mathcal{A}} a_k.\text{nil}) = (1-\varepsilon)^k$ and $\pi_{v,k}$ giving the remaining probability $1 - (1-\varepsilon)^k$ to processes that cannot move. We get $\mathbf{b}^\lambda(f(u_k), f(v_k)) = \lambda(1 - (1-\varepsilon)^k)$. Hence, $\sup_{k \in \mathbb{N}} \mathbf{b}^\lambda(f(u_k), f(v_k)) = \sup_{k \in \mathbb{N}} \lambda(1 - (1-\varepsilon)^k) = \lambda$, thus giving that the distance between f -composed processes is bounded by $\mathbf{m}(\varepsilon) = \lambda$ if $\varepsilon > 0$ and $\mathbf{m}(0) = 0$, which is not a modulus of continuity since it is not continuous at 0. Hence, f is not uniformly continuous accordingly to Definition 10.

Consider now operator g . We have $g(u_k) \xrightarrow{a_k} \delta(h(\dots h(u_k)\dots))$ (with k nested occurrences of h) and $g(v_k) \xrightarrow{a_k} \delta(h(\dots h(v_k)\dots))$. Then, $h(\dots h(u_k)\dots) \xrightarrow{a_k} \pi'_{u,k}$ and $h(\dots h(v_k)\dots) \xrightarrow{a_k} \pi'_{v,k}$, with $\pi'_{u,k}(a_k.\text{nil} \parallel_{\mathcal{A}} \dots \parallel_{\mathcal{A}} a_k.\text{nil}) = 1$ (parallel composition of 2^k copies of $a_k.\text{nil}$), $\pi'_{v,k}(a_k.\text{nil} \parallel_{\mathcal{A}} \dots \parallel_{\mathcal{A}} a_k.\text{nil}) = (1-\varepsilon)^{2^k}$ and $\pi'_{v,k}$ giving probability $1 - (1-\varepsilon)^{2^k}$ to processes that cannot move. We get $\mathbf{b}^\lambda(g(u_k), g(v_k)) = \lambda^2(1 - (1-\varepsilon)^{2^k})$ and, then, $\sup_{k \in \mathbb{N}} \mathbf{b}^\lambda(g(u_k), g(v_k)) = \sup_{k \in \mathbb{N}} \lambda^2(1 - (1-\varepsilon)^{2^k}) = \lambda^2$. As for operator f , we conclude that g is not uniformly continuous.

Example 9 Counterexample for occurrences in non-expansive operators. In Example 6 we noticed that the rules for the copy operator cp violate Definition 15. Let $\varepsilon \in (0, 1)$. Following [19], consider processes $s_1 = l.([1-\varepsilon]a.\text{nil} \oplus [\varepsilon]\text{nil}) + r.([1-\varepsilon]a.\text{nil} \oplus [\varepsilon]\text{nil})$ and $t_1 = l.a.\text{nil} + r.a.\text{nil}$ for some $a \neq l, r$. Then, for $k > 1$, let $s_k = l.s_{k-1} + r.s_{k-1}$ and $t_k = l.t_{k-1} + r.t_{k-1}$.

Clearly, $\mathbf{b}^\lambda(s_k, t_k) = \lambda^k \cdot \varepsilon$. Similar calculations to those in Example 1 give $\mathbf{b}^\lambda(\text{cp}(s_k), \text{cp}(t_k)) = \lambda^k(1 - (1 - \varepsilon)^{2^k})$. Thus, for any k with $2^k > L$, $\mathbf{b}^\lambda(\text{cp}(s), \text{cp}(t))/\mathbf{b}^\lambda(s, t) = (1 - (1 - \varepsilon)^{2^k})/\varepsilon > L$ holds for $s = s_k$, $t = t_k$ and $\varepsilon \in (0, (2^k - L)/(2^{k-1}(2^k - 1)))$. We can conclude that cp is not Lipschitz continuous.

7 Conclusions

We have provided three specification formats for the bisimilarity metric. These formats guarantee, respectively, the *non-extensiveness*, *non-expansiveness* and *Lipschitz continuity* of the operators satisfying the related format wrt. the bisimilarity metric.

To the best of our knowledge, the only other formats for the bisimilarity metric proposed in the literature are those in [19]. As outlined in the Introduction, the difference between the two proposals is in that our formats are purely syntactic, whereas the ones in [19] are *semantic*, in the sense that they require a recursive computation of how many copies of a process argument can be spawned by each operator. However, we remark that we have applied the same technique of [19] to the evaluation of Lipschitz factors of operators. This is due to the fact that from the syntactic constraints on PGSOS rules for an operator f we can only guarantee that f is Lipschitz continuous *for some* L_f and that $L_f \geq n_f$, but no further information can be inferred. This is reasonable since the validity of the Lipschitz continuity property does not depend on the value of the Lipschitz constant, but only on its existence. The actual value of L_f can be established, in general, only from the semantics of f .

As future work, we plan to provide specification formats for *weak* bisimulation metrics [15], whose applicability has been demonstrated in [24–26]. Intuitively, we need to consider TSSs in a format guaranteeing that the *kernel* of the chosen metric is a *congruence* and enrich our three formats with constraints allowing us to deal with silent moves.

Acknowledgements. We thank the anonymous reviewers for their detailed comments and feedback. V. Castiglioni has been partially supported by the project ‘Open Problems in the Equational Logic of Processes’ (OPEL) of the Icelandic Research Fund (grant nr. 196050-051). We thank Carlos, Kostas, Mario and Frank for inviting us to submit this paper. Last but not least, we extend our heartiest wishes to Catuscia. Valentina is indebted to Catuscia for her support, mentoring and friendship during the period spent at INRIA, in her team COMETE. We wish Catuscia a very happy birthday!

References

1. Aceto, L., Fokink, W.J., Verhoef, C.: Structural operational semantics. In: Handbook of Process Algebra, pp. 197–292. Elsevier (2001)
2. de Alfaro, L., Faella, M., Stoelinga, M.: Linear and branching system metrics. IEEE Trans. Softw. Eng. **35**(2), 258–273 (2009)

3. de Alfaro, L., Henzinger, T.A., Majumdar, R.: Discounting the future in systems theory. In: Baeten, J.C.M., Lenstra, J.K., Parrow, J., Woeginger, G.J. (eds.) ICALP 2003. LNCS, vol. 2719, pp. 1022–1037. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-45061-0_79
4. Bacci, G., Bacci, G., Larsen, K.G., Mardare, R.: Computing behavioral distances, compositionally. In: Chatterjee, K., Sgall, J. (eds.) MFCS 2013. LNCS, vol. 8087, pp. 74–85. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40313-2_9
5. Bloom, B., Istrail, S., Meyer, A.R.: Bisimulation can't be traced. *J. ACM* **42**(1), 232–268 (1995)
6. van Breugel, F., Worrell, J.: A behavioural pseudometric for probabilistic transition systems. *Theor. Comput. Sci.* **331**(1), 115–142 (2005)
7. Castiglioni, V.: Trace and testing metrics on nondeterministic probabilistic processes. In: Proceedings of the EXPRESS/SOS 2018. EPTCS, vol. 276, pp. 19–36 (2018)
8. Castiglioni, V., Gebler, D., Tini, S.: Modal decomposition on nondeterministic probabilistic processes. In: Proceedings of the CONCUR 2016, pp. 36:1–36:15 (2016)
9. Castiglioni, V., Gebler, D., Tini, S.: SOS-based modal decomposition on nondeterministic probabilistic processes. *Logical Methods Comput. Sci.* **14**(2) (2018)
10. Chatzikokolakis, K., Gebler, D., Palamidessi, C., Xu, L.: Generalized bisimulation metrics. In: Baldan, P., Gorla, D. (eds.) CONCUR 2014. LNCS, vol. 8704, pp. 32–46. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44584-6_4
11. D'Argenio, P.R., Gebler, D., Lee, M.D.: A general SOS theory for the specification of probabilistic transition systems. *Inf. Comput.* **249**, 76–109 (2016)
12. D'Argenio, P.R., Gebler, D., Lee, M.D.: Axiomatizing bisimulation equivalences and metrics from probabilistic SOS rules. In: Muscholl, A. (ed.) FoSSaCS 2014. LNCS, vol. 8412, pp. 289–303. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-54830-7_19
13. Deng, Y., Chothia, T., Palamidessi, C., Pang, J.: Metrics for action-labelled quantitative transition systems. *Electr. Notes Theor. Comput. Sci.* **153**(2), 79–96 (2006)
14. Desharnais, J., Gupta, V., Jagadeesan, R., Panangaden, P.: Metrics for labelled markov processes. *Theor. Comput. Sci.* **318**(3), 323–354 (2004)
15. Desharnais, J., Jagadeesan, R., Gupta, V., Panangaden, P.: The metric analogue of weak bisimulation for probabilistic processes. In: Proceedings of the LICS 2002, pp. 413–422 (2002)
16. Gebler, D., Larsen, K.G., Tini, S.: Compositional bisimulation metric reasoning with probabilistic process calculi. *Logical Methods Comput. Sci.* **12**(4) (2016)
17. Gebler, D., Tini, S.: Compositionality of approximate bisimulation for probabilistic systems. In: Proceedings of the EXPRESS/SOS 2013, pp. 32–46 (2013)
18. Gebler, D., Tini, S.: Fixed-point characterization of compositionality properties of probabilistic processes combinators. In: Proceedings of the EXPRESS/SOS 2014, pp. 63–78 (2014)
19. Gebler, D., Tini, S.: SOS specifications for uniformly continuous operators. *J. Comput. Syst. Sci.* **92**, 113–151 (2018)
20. Giacalone, A., Jou, C.C., Smolka, S.A.: Algebraic reasoning for probabilistic concurrent systems. In: Proceedings of the IFIP Work, Conference on Programming, Concepts and Methods, pp. 443–458 (1990)
21. Kantorovich, L.V.: On the transfer of masses. *Dokl. Akad. Nauk SSSR* **37**(2), 227–229 (1942)

22. Katoen, J.P., Baier, C., Latella, D.: Metric semantics for true concurrent real time. *TCS* **254**(1–2), 501–542 (2001)
23. Kwiatkowska, M., Norman, G.: Probabilistic metric semantics for a simple language with recursion. In: Penczek, W., Szałas, A. (eds.) *MFCS 1996*. LNCS, vol. 1113, pp. 419–430. Springer, Heidelberg (1996). https://doi.org/10.1007/3-540-61550-4_167
24. Lanotte, R., Merro, M., Tini, S.: Compositional weak metrics for group key update. In: *MFCS 2017*. LIPIcs, vol. 42, pp. 72:1–72:16 (2017)
25. Lanotte, R., Merro, M., Tini, S.: Equational reasonings in wireless networkgossip protocols. *Logical Methods Comput. Sci.* **14**(3) (2018)
26. Lanotte, R., Merro, M., Tini, S.: Towards a formal notion of impact metric for cyber-physical attacks. In: Furia, C.A., Winter, K. (eds.) *IFM 2018*. LNCS, vol. 11023, pp. 296–315. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98938-9_17
27. Larsen, K.G., Skou, A.: Bisimulation through probabilistic testing. *Inf. Comput.* **94**(1), 1–28 (1991)
28. Mousavi, M.R., Reniers, M.A., Groote, J.F.: Sos formats and meta-theory: 20 years after. *Theoret. Comput. Sci.* **373**(3), 238–272 (2007)
29. Plotkin, G.: A structural approach to operational semantics. Report DAIMI FN-19, Aarhus University (1981). Reprinted in *JLAP*, 60–61:17–139 (2004)
30. Plotkin, G.D.: A structural approach to operational semantics. Report DAIMI FN-19, Aarhus University (1981)
31. Segala, R.: Modeling and verification of randomized distributed real-time systems. Ph.D. thesis, MIT (1995)
32. Segala, R., Lynch, N.A.: Probabilistic simulations for probabilistic processes. *Nord. J. Comput.* **2**(2), 250–273 (1995)
33. de Simone, R.: Higher-level synchronising devices in MEIJE-SCCS. *Theor. Comput. Sci.* **37**, 245–267 (1985)
34. Song, L., Deng, Y., Cai, X.: Towards automatic measurement of probabilistic processes. In: *Proceedings of the QSIK 2007*, pp. 50–59 (2007)
35. Tracol, M., Desharnais, J., Zhioua, A.: Computing distances between probabilistic automata. In: *Proceedings of the QAPL 2011*, pp. 148–162 (2011)