Complete Axiomatization for the Bisimilarity Distance on Markov Chains*

Giorgio Bacci[1], Giovanni Bacci[2], Kim G. Larsen[3], and Radu Mardare[4]

1 **Dept. of Computer Science, Aalborg University, Denmark.**
`grbacci@cs.aau.dk`
2 **Dept. of Computer Science, Aalborg University, Denmark.**
`giovbacci@cs.aau.dk`
3 **Dept. of Computer Science, Aalborg University, Denmark.**
`klg@cs.aau.dk`
4 **Dept. of Computer Science, Aalborg University, Denmark.**
`mardare@cs.aau.dk`

Abstract

In this paper we propose a complete axiomatization of the bisimilarity distance of Desharnais et al. for the class of finite labelled Markov chains. Our axiomatization is given in the style of a quantitative extension of equational logic recently proposed by Mardare, Panangaden, and Plotkin (LICS'16) that uses equality relations $t \equiv_\varepsilon s$ indexed by rationals, expressing that "t is approximately equal to s up to an error ε". Notably, our quantitative deductive system extends in a natural way the equational system for probabilistic bisimilarity given by Stark and Smolka by introducing an axiom for dealing with the Kantorovich distance between probability distributions.

1998 ACM Subject Classification F.3.2 Algebraic Approaches to Semantics.

Keywords and phrases Markov chains, Behavioral distances, Axiomatization.

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2016.21

1 Introduction

A very attractive approach toward the study of the behavior of systems consists in expressing behavioral properties in an equational algebraic fashion. The attractiveness of the equational reasoning comes from the fact that one can deal with different notions of behaviors (such as non-deterministic, probabilistic, etc.) in a compositional way, by introducing new algebraic operators and their corresponding axioms as a sequence of successive refinements.

There is a well-established literature considering complete axiomatizations of several semantic theories [15, 3, 19, 4, 1, 16, 6, 18]. Amongst the aforementioned references, the studies [19, 1, 16, 18] consider operators for the definitions of recursive behaviors and offer implicational equational proof systems for probabilistic bisimulation equivalence. It is well-known that for reasoning about the behavior of probabilistic system, a notion of distance measuring the dissimilarities of two systems is preferable to that of equivalence, since the latter is not robust w.r.t. small variations of numerical values (see e.g. [7] for more details).

* Work supported by the EU 7th Framework Programme (FP7/2007-13) under Grants Agreement nr.318490 (SENSATION), nr.601148 (CASSTING), the Sino-Danish Basic Research Center IDEA4CPS funded by Danish National Research Foundation and National Science Foundation China, the ASAP Project (4181-00360) funded by the Danish Council for Independent Research, the ERC Advanced Grant LASSO, and the Innovation Fund Denmark center DiCyPS.

The first proposals for a complete axiomatization of behavioral distances for quantitative systems are due to Larsen et al. [12] and D'Argenio et al. [5], respectively axiomatizing the weighted and probabilistic bisimulation metrics. The approaches pursued in these works, however, are rather specific and based on *ad hoc* assumptions.

Recently, Mardare, Panangaden, and Plotkin [14] —with the purpose of developing a general research program for a quantitative algebraic theory of effects [17]— proposed the concepts of *quantitative equational theory* and *quantitative algebra* as models for these theories. The key idea behind their approach is to use equations of the form $t \equiv_\varepsilon s$ annotated with a rational number ε to be interpreted as "t is approximately equal to s up to an error ε". Their main result is that completeness for a quantitative theory always holds on the freely generated algebra of terms equipped with a metric that is freely induced by the axioms. Due to this result, they were able to prove completeness for many interesting axiomatizations, such as the Hausdorff, the total variation, the p-Wasserstein, and the Kantorovich metrics.

In this paper, we contribute to the quest of complete axiomatizations of behavioral metrics, by proposing a quantitative equational theory in the sense of [14] that is proved to be complete w.r.t. the bisimilarity distance of Desharnais et al. [7] for finitely presentable labelled Markov chains. The signature of operators that we consider is the one of [19], consisting of a prefix operator, a binary probabilistic choice operator, and a recursion operator. The set of axioms we use is that of barycentric algebras relative to the probabilistic choice operator and Milner's axioms for recursion [15]. To deal with the Kantorovich distance —that is the basic ingredient for the definition of the bisimilarity distance— we use the axiom (IB) from [14]. The resulting axiomatization is simpler than the one presented in [5] for probabilistic transition systems and it extends [5] by allowing recursive behaviors.

For the proof of completeness we could not apply the general proof technique of [14], since the recursion operator is not sound w.r.t. the axiom of non-expansiveness, that is required to fit within the quantitative algebraic framework of [14]. To prove completeness we needed to appeal to specific properties of the functional operator used to define the distance, namely, that it preserves infima of countable decreasing chains, a.k.a. ω-cocontinuity. Interestingly, the proof technique we use seems to be generic on the functional operator that defines the distance, provided that it is ω-cocontinuous.

Moreover, we show that the class of expressible behaviors, namely those that can be described as syntactic terms of this signature, corresponds up to bisimilarity to the class of finite and finitely supported labelled Markov chains. This establishes a strong correspondence between syntactic terms and a clearly defined semantic class of probabilistic systems.

2 Preliminaries and Notation

For $R \subseteq X \times X$ an equivalence relation, we denote by $X/_R$ its quotient set. For two sets X and Y, we denote by $X \uplus Y$ their disjoint union.

A *discrete sub-probability* on X is a function $\mu \colon X \to [0,1]$, such that $\mu(X) \leq 1$, where, for $E \subseteq X$, $\mu(E) = \sum_{x \in E} \mu(x)$; it is a *probability distribution* if $\mu(X) = 1$. The support of μ is the set $supp(\mu) = \{x \in X \mid \mu(x) > 0\}$. We denote by $\Delta(X)$ and $\mathcal{D}(X)$ the set of probability and finitely-supported sub-probability distributions on X, respectively.

A 1-bounded *pseudometric* on X is a function $d \colon X \times X \to [0,1]$ such that, for any $x, y, z \in X$, $d(x,x) = 0$, $d(x,y) = d(y,x)$ and $d(x,y) + d(y,z) \geq d(x,z)$; d is a *metric* if, in addition, $d(x,y) = 0$ implies $x = y$. The pair (X, d) is called *(pseudo)metric space*. For $n \in \mathbb{N}$, the *n-th product (pseudo)metric space* of (X,d) is defined as (X^n, d') where $d'((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max_{i=1}^{n} d(x_i, y_i)$. The kernel of a (pseudo)metric d is the set $\ker(d) = \{(x,y) \mid d(x,y) = 0\}$.

3 Quantitative Algebras and their Equational Theories

We recall the notions of quantitative equational theory and quantitative algebras from [14].

Let Σ be an algebraic signature of function symbols $f\colon n \in \Sigma$ of arity $n \in \mathbb{N}$. Fix a countable set of *metavariables* X, ranged over by $x, y, z, \ldots \in X$. We denote by $\mathbb{T}(\Sigma, X)$ the set of Σ-terms freely generated over X; terms will be ranged over by $t, s, u, \ldots$ A *substitution of type* Σ is a function $\sigma\colon X \to \mathbb{T}(\Sigma, X)$ that is homomorphically extended to terms as $\sigma(f(t_1, \ldots, t_n)) = f(\sigma(t_1), \ldots, \sigma(t_n))$; by $\mathcal{S}(\Sigma)$ we denote the set of substitutions of type Σ.

A *quantitative equation of type* Σ is an expression of the form $t \equiv_\varepsilon s$, where $t, s \in \mathbb{T}(\Sigma, X)$ and $\varepsilon \in \mathbb{Q}_+$. Let $\mathcal{E}(\Sigma)$ denote the set of quantitative equations of type Σ and let range over its subsets by $\Gamma, \Theta, \Pi, \ldots \subseteq \mathcal{E}(\Sigma)$.

Let $\vdash\, \subseteq 2^{\mathcal{E}(\Sigma)} \times \mathcal{E}(\Sigma)$ be a binary relation from the powerset of $\mathcal{E}(\Sigma)$ to $\mathcal{E}(\Sigma)$. We write $\Gamma \vdash t \equiv_\varepsilon s$ if $(\Gamma, t \equiv_\varepsilon s) \in\, \vdash$, and $\Gamma \nvdash t \equiv_\varepsilon s$ otherwise; by $\vdash t \equiv_\varepsilon s$ we denote $\emptyset \vdash t \equiv_\varepsilon s$, and by $\Gamma \vdash \Theta$ we mean that $\Gamma \vdash t \equiv_\varepsilon s$, for all $t \equiv_\varepsilon s \in \Theta$. The relation $\vdash$ is called *quantitative deduction system of type* Σ if it satisfies the following axioms and rules

$$
\begin{array}{rl}
(\mathsf{Refl}) & \vdash t \equiv_0 t\,, \\
(\mathsf{Symm}) & \{t \equiv_\varepsilon s\} \vdash s \equiv_\varepsilon t\,, \\
(\mathsf{Triang}) & \{t \equiv_\varepsilon u, u \equiv_{\varepsilon'} s\} \vdash t \equiv_{\varepsilon+\varepsilon'} s\,, \\
(\mathsf{Max}) & \{t \equiv_\varepsilon s\} \vdash t \equiv_{\varepsilon+\varepsilon'} s\,, \text{ for all } \varepsilon' > 0\,, \\
(\mathsf{Arch}) & \{t \equiv_{\varepsilon'} s \mid \varepsilon' > \varepsilon\} \vdash t \equiv_\varepsilon s\,, \\
(\mathsf{NExp}) & \{t_1 =_\varepsilon s_1, \ldots, t_n =_\varepsilon s_n\} \vdash f(t_1, \ldots, t_n) \equiv_\varepsilon f(s_1, \ldots, s_n)\,, \text{ for all } f\colon n \in \Sigma\,, \\
(\mathsf{Subst}) & \text{If } \Gamma \vdash t \equiv_\varepsilon s, \text{ then } \sigma(\Gamma) \vdash \sigma(t) \equiv_\varepsilon \sigma(t), \text{ for all } \sigma \in \mathcal{S}(\Sigma)\,, \\
(\mathsf{Cut}) & \text{If } \Gamma \vdash \Theta \text{ and } \Theta \vdash t \equiv_\varepsilon s, \text{ then } \Gamma \vdash t \equiv_\varepsilon s\,, \\
(\mathsf{Assum}) & \text{If } t \equiv_\varepsilon s \in \Gamma, \text{ then } \Gamma \vdash t \equiv_\varepsilon s\,.
\end{array}
$$

where $\sigma(\Gamma) = \{\sigma(t) \equiv_\varepsilon \sigma(s) \mid t \equiv_\varepsilon s \in \Gamma\}$.

An expression of the form $\{t_1 \equiv_{\varepsilon_1} s_1, \ldots, t_n \equiv_{\varepsilon_n} s_n\} \vdash t \equiv_\varepsilon s$ —i.e., with finite set of hypotheses— is called *basic quantitative inference*. A *quantitative equational theory* is a set $\mathcal{U}$ of basic quantitative inferences closed under $\vdash$-deducibility. A set $\mathcal{A}$ of basic inferences is said to axiomatize a quantitative equational theory $\mathcal{U}$, if $\mathcal{U}$ is the smallest quantitative equational theory containing $\mathcal{A}$. A theory $\mathcal{U}$ is called *inconsistent* if $\vdash x \equiv_0 y \in \mathcal{U}$, for distinct metavariables $x, y \in X$, it is called *consistent* otherwise[1]. The models of quantitative equational theories are given by the following structures.

▶ **Definition 1** (Quantitative Algebra). A *quantitative* Σ*-algebra* is a tuple $\mathcal{A} = (A, \Sigma^{\mathcal{A}}, d^{\mathcal{A}})$, consisting of a pseudometric space $(A, d^{\mathcal{A}})$, with $d^{\mathcal{A}}\colon A \times A \to [0, \infty]$, and a set of non-expansive *interpretations* $\Sigma^{\mathcal{A}} = \big\{f^{\mathcal{A}}\colon A^n \to A \mid f\colon n \in \Sigma\big\}$ for the function symbols in Σ.

Quantitative Σ-algebras extend standard Σ-algebras with a notion of distance. Morphisms of quantitative algebras are non-expansive homomorphisms.

A quantitative algebra $\mathcal{A} = (A, \Sigma^{\mathcal{A}}, d^{\mathcal{A}})$ *satisfies* the quantitative inference $\Gamma \vdash t \equiv_\varepsilon s$, written $\Gamma \models_{\mathcal{A}} t \equiv_\varepsilon s$, if for any assignment of the meta-variables $\iota\colon X \to A$,

$$\big(\text{for all } t' \equiv_{\varepsilon'} s' \in \Gamma,\ d^{\mathcal{A}}(\iota(t'), \iota(s')) \le \varepsilon'\big) \qquad \text{implies} \qquad d^{\mathcal{A}}(\iota(t), \iota(s)) \le \varepsilon\,,$$

[1] Note that for an inconsistent theory $\mathcal{U}$, by Subst, we have $\vdash t \equiv_0 s \in \mathcal{U}$, for all $t, s \in \mathbb{T}(\Sigma, X)$.

where, for a term $t \in \mathbb{T}(\Sigma, X)$, $\iota(t)$ denotes the homomorphic interpretation of t in $\mathcal{A}$. A quantitative algebra $\mathcal{A}$ is said to be a *model* for a quantitative theory $\mathcal{U}$, if $\Gamma \models_{\mathcal{A}} t \equiv_\varepsilon s$, for all $\Gamma \vdash t \equiv_\varepsilon s \in \mathcal{U}$.

In [14] it is shown that any quantitative theory $\mathcal{U}$ has a universal model $\mathcal{T}_{\mathcal{U}}$ (the freely generated $\vdash$-model) satisfying exactly those quantitative equations belonging to $\mathcal{U}$. Moreover, in [14] it is proven a strong completeness theorem for quantitative equational theories $\mathcal{U}$, stating that a basic inference is satisfied by all the algebras satisfying $\mathcal{U}$ iff it belongs to $\mathcal{U}$.

Furthermore, in [14] several interesting examples of quantitative equational theories have been proposed. The one we will focus on later in this paper is the so called interpolative barycentric equational theory (see §10 in [14]).

4 The Quantitative Algebra of Probabilistic Behaviors

In this section we present the quantitative algebra of open Markov chains. Open Markov chains extend the familiar notion of discrete-time labelled Markov chain with "open" states taken from a fixed countable set $\mathcal{X}$ of names ranged over by $X, Y, Z, \ldots \in \mathcal{X}$. Names indicate states at which the behavior of the Markov chain can be extended by substitution of another Markov chain, in a way which will be made precise later.

4.1 Open Markov Chains and Bisimilarity Distance

In what follows we fix a set $\mathcal{L}$ of labels, ranged over by $a, b, c, \ldots \in \mathcal{L}$. Recall that $\mathcal{D}(M)$ denotes the set of *finitely supported* discrete sub-probability distributions over a set M.

▶ **Definition 2** (Open Markov Chain). An *open Markov chain* $\mathcal{M} = (M, \tau)$ consists of a set M of *states* and a *transition probability function* $\tau \colon M \to \mathcal{D}((\mathcal{L} \times M) \uplus \mathcal{X})$.

Intuitively, if $\mathcal{M}$ is in a state $m \in M$ it moves with action $a \in \mathcal{L}$ to a state $n \in M$, with probability $\tau(m)(a, n)$ and to a name $X \in \mathcal{X}$ with probability $\tau(m)(X)$. A name $X \in \mathcal{X}$ is said to be *unguarded* in a state $m \in M$, if $\tau(m)(X) > 0$. Clearly, $\mathcal{L}$-labelled sub-probabilistic Markov chains are encoded as open Markov chains by letting $\tau(m)(\mathcal{X}) = 0$, for all $m \in M$.

A *pointed open Markov chain*, denoted by $(\mathcal{M}, m)$, is a Markov chain $\mathcal{M} = (M, \tau)$ with a distinguished *initial* state $m \in M$. We use $\mathcal{M} = (M, \tau)$ and $\mathcal{N} = (N, \theta)$ to range over open Markov chains and $(\mathcal{M}, m)$, $(\mathcal{N}, n)$ to range over the set **OMC** of pointed open Markov chains. In the following we will often refer to the constituents of $\mathcal{M}$ and $\mathcal{N}$ implicitly.

Next we recall the probabilistic bisimulation of Larsen and Skou [13].

▶ **Definition 3** (Bisimulation). An equivalence relation $R \subseteq M \times M$ is a *bisimulation* on $\mathcal{M}$ if whenever $m \mathrel{R} m'$, then, for all $a \in \mathcal{L}$, $X \in \mathcal{X}$ and $C \in M/_R$,
- $\tau(m)(X) = \tau(m')(X)$,
- $\tau(m)(\{a\} \times C) = \tau(m')(\{a\} \times C)$.

Two states $m, m' \in M$ are *bisimilar* w.r.t. $\mathcal{M}$, written $m \sim_{\mathcal{M}} m'$, if there exists a bisimulation relation on $\mathcal{M}$ relating them.

We say that two pointed open Markov chains $(\mathcal{M}, m), (\mathcal{N}, n) \in \mathbf{OMC}$ are bisimilar, written $(\mathcal{M}, n) \sim (\mathcal{N}, n)$, if m and n are bisimilar w.r.t. the disjoint union of $\mathcal{M}$ and $\mathcal{N}$, defined as expected. The bisimilarity relation $\sim \subseteq \mathbf{OMC} \times \mathbf{OMC}$ is an equivalence (see e.g. [2]).

The notion of bisimulation can be lifted to pseudometrics by means of a straightforward extension of the bisimilarity distance of Desharnais et al. [7] over open Markov chains —we refer the interested reader to [7, 22] for more details about its properties— that is based on the Kantorovich distance $\mathcal{K}(d)(\mu, \nu) = \min \left\{ \int d \, \mathrm{d}\omega \mid \omega \in \Omega(\mu, \nu) \right\}$ between probability measures $\mu, \nu \in \Delta(A)$ w.r.t. the underlying distance d on A. In the definition, $\Omega(\mu, \nu)$ is the set of *couplings* for (μ, ν), i.e., a probability distributions $\omega \in \Delta(A \times A)$ such that, for all $E \subseteq A$, $\omega(E \times A) = \mu(E)$ and $\omega(A \times E) = \nu(E)$.

▶ Remark. The definition of $\mathcal{K}(d)$ above is tailored on probability distributions, whereas in the present setting we are dealing with sub-probability distributions $\mu \in \mathcal{D}(A)$. To use $\mathcal{K}(d)$ on $\mathcal{D}(A)$ it is standard to add a bottom element $\perp$ in A, that is assumed to be at maximum distance from all elements $a \in A$, written $A_\perp$; and define $\mu^* \in \Delta(A_\perp)$ as the unique probability distribution such that, for all $E \subseteq A$, $\mu^*(E) = \mu(E)$ and, $\mu^*(\perp) = 1 - \mu(A)$. ◀

The set of 1-bounded pseudometrics over a set M ordered point-wise by $d \sqsubseteq d'$ iff for all $m, n \in M$, $d(m, n) \leq d'(m, n)$ is a complete partial order, with bottom given by the 0-constant pseudometric $\mathbf{0}$ and join being the point-wise supremum. We define the bisimilarity pseudometric $\mathbf{d}_\mathcal{M} \colon M \times M \to [0, 1]$ over $\mathcal{M} = (M, \tau)$ as the least fixed-point of the following functional operator on 1-bounded pseudometrics

$$\Psi_\mathcal{M}(d)(m, m') = \mathcal{K}(\Lambda(d))(\tau^*(m), \tau^*(m')) \qquad \text{(Kantorovich Operator)}$$

where $\Lambda(d)$ is the greatest 1-bounded pseudometric on $\big((\mathcal{L} \times M) \uplus \mathcal{X}\big)_\perp$ such that, for all $a \in \mathcal{L}$ and $t, s \in \mathbb{T}$, $\Lambda(d)((a, t), (a, s)) = d(t, s)$. Hereafter, whenever $\mathcal{M}$ is clear from the context we will simply write $\mathbf{d}$ and Ψ in place of $\mathbf{d}_\mathcal{M}$ and $\Psi_\mathcal{M}$, respectively.

The well definition of $\mathbf{d}$ is guaranteed by the first half of the next lemma and Knaster-Tarski fixed-point theorem. We also prove that Ψ is ω-continuous, i.e., it preserves suprema of countable increasing chains. Note that by this and Kleene fixed-point theorem, the bisimilarity distance can be alternatively characterized as $\mathbf{d} = \bigsqcup_{n \in \omega} \Psi^n(\mathbf{0})$.

▶ **Lemma 4.** *The operator Ψ is monotonic and ω-continuous.*

Proof. Monotonicity of Ψ follows from the monotonicity of $\mathcal{K}$ and Λ. ω-continuity follows from [21, Theorem 1] by showing that Ψ is non expansive, i.e., for all $d, d' \colon M \times M \to [0, 1]$, $\|\Psi(d') - \Psi(d)\| \leq \|d' - d\|$, where $\|f\| = \sup_x |f(x)|$ is the supremum norm. It suffices to prove that for all $d \sqsubseteq d'$ and $m, m' \in M$, $\Psi(d')(m, m') - \Psi(d)(m, m') \leq \|d' - d\|$:

$$\begin{aligned} &\Psi(d')(m, m') - \Psi(d)(m, m') \\ &= \mathcal{K}(\Lambda(d'))(\tau^*(m), \tau^*(m')) - \mathcal{K}(\Lambda(d))(\tau^*(m), \tau^*(m')) && \text{(by def. } \Psi) \end{aligned}$$

by choosing $\omega \in \Omega(\tau^*(m), \tau^*(m'))$ such that $\mathcal{K}(\Lambda(d))(\tau^*(m), \tau^*(m')) = \int \Lambda(d) \, d\omega$,

$$\begin{aligned} &= \mathcal{K}(\Lambda(d'))(\tau^*(m), \tau^*(m')) - \textstyle\int \Lambda(d) \, d\omega \\ &\leq \textstyle\int \Lambda(d') \, d\omega - \int \Lambda(d) \, d\omega && \text{(by def. of } \mathcal{K}(\Lambda(d'))) \\ &= \textstyle\int (\Lambda(d') - \Lambda(d)) \, d\omega && \text{(by linearity)} \end{aligned}$$

and since, for all $(\alpha, \beta) \notin E = \{((a, n), (a, n')) \mid a \in \mathcal{L}, n, n' \in M\}$, $\Lambda(d')(\alpha, \beta) = \Lambda(d)(\alpha, \beta)$,

$$\begin{aligned} &= \textstyle\int_E (\Lambda(d') - \Lambda(d)) \, d\omega \\ &\leq \textstyle\int_E \|d' - d\| \, d\omega && \text{(by def. } \Lambda) \\ &\leq \|d' - d\| \, . && \text{(by linearity and } \textstyle\int_E 1 \, d\omega \leq 1) \end{aligned}$$

◀

The next Lemma states that $\mathbf{d}$ lifts the bisimilarity relation to a pseudometric.

▶ **Lemma 5.** $\mathbf{d}(m, m') = 0$ *iff* $m \sim m'$.

Proof. ($\Leftarrow$) We prove that $R = \{(m, m') \mid \mathbf{d}(m, m') = 0\}$ (i.e., $\ker(\mathbf{d})$) is a bisimulation. Clearly, R is an equivalence, and also $\ker(\Lambda(d))$ is so. Assume $(m, m') \in R$. By definition of Ψ, we have that $\mathcal{K}(\Lambda(\mathbf{d}))(\tau^*(m), \tau^*(m')) = 0$. By [8, Lemma 3.1], for all $\ker(\Lambda(d))$-equivalence classes $D \subseteq ((\mathcal{L} \times M) \uplus \mathcal{X})_\perp$, $\tau^*(m)(D) = \tau^*(m')(D)$. By definition of Λ, this implies that, for all $a \in \mathcal{L}$, $X \in \mathcal{X}$ and $C \in M/_R$, $\tau(m)(X) = \tau(m')(X)$ and, moreover, $\tau(m)(\{a\} \times C) = \tau(m')(\{a\} \times C)$.

($\Rightarrow$) Let $R \subseteq M \times M$ be a bisimulation on $\mathcal{M}$, and define $d_R \colon M \times M \to [0, 1]$ by $d_R(m, m') = 0$ if $(m, m') \in R$ and $d_R(m, m') = 1$ otherwise. We show that $\Psi(d_R) \sqsubseteq d_R$. If $(m, m') \notin R$, then $d_R(m, m') = 1 \geq \Psi(d_R)(m, m')$. If $(m, m') \in R$, then for all $a \in \mathcal{L}$, $X \in \mathcal{X}$ and $C \in M/_R$, $\tau(m)(X) = \tau(m')(X)$, $\tau(m)(\{a\} \times C) = \tau(m')(\{a\} \times C)$. This implies that for all $\ker(\Lambda(d_R))$-equivalence class $D \subseteq ((\mathcal{L} \times M) \uplus \mathcal{X})_\perp$, $\tau^*(m)(D) = \tau^*(m')(D)$. By [8, Lemma 3.1], we have $\mathcal{K}(\Lambda(d_R))(\tau^*(m), \tau^*(m')) = 0$. This implies that $\Psi(d_R) \sqsubseteq d_R$. Since $\sim$ is a bisimulation, $\Psi(d_\sim) \sqsubseteq d_\sim$, so that, by Tarski's fixed point theorem, $\mathbf{d} \sqsubseteq d_\sim$. By definition of $d_\sim$ and $\mathbf{d} \sqsubseteq d_\sim$, $m \sim m'$ implies $\mathbf{d}(m, m') = 0$. ◀

The definition above can be extended to the collection **OMC** of open Markov chains as $\mathbf{d_{OMC}} \colon \mathbf{OMC} \times \mathbf{OMC} \to [0, 1]$ by using the bisimilarity distance on the disjoint union of their open Markov chains structures and by taking the distance between their initial states.

4.2 The Algebra of Open Markov Chains

Next we turn to simple algebra of pointed Markov chains. The signature of this algebra is defined as follows,

$$\begin{aligned} \Sigma = {} & \{X \colon 0 \mid X \in \mathcal{X}\} \cup && \text{(NAMES)} \\ & \{a.(\cdot) \colon 1 \mid a \in \mathcal{L}\} \cup && \text{(PREFIX)} \\ & \{+_e \colon 2 \mid e \in [0, 1]\} \cup && \text{(PROBABILISTIC CHOICE)} \\ & \{\mathsf{rec}\, X \colon 1 \mid X \in \mathcal{X}\} \,, && \text{(RECURSION)} \end{aligned}$$

consisting of a constant X for each name in $\mathcal{X}$; prefix $a.\cdot$ and a recursion $\mathsf{rec}\, X$ unary operators, for each $a \in \mathcal{L}$ and $X \in \mathcal{X}$; and a probabilistic choice $+_e$ binary operator for each $e \in [0, 1]$. For $t \in \mathbb{T}(\Sigma, M)$, $fn(t)$ denotes the set of free names in t, where the notions of *free* and *bound name* are defined in the standard way, with $\mathsf{rec}\, X$ acting as a binding construct. A term is *closed* if it does not contain any free variable. Throughout the paper we consider two terms as syntactically identical if they are identical up to renaming of their bound names. For $t, s_1, \ldots, s_n \in \mathbb{T}(\Sigma, M)$ and an n-vector $\overline{X} = (X_1, \ldots, X_n)$ of distinct names, $t[\overline{s}/\overline{X}]$ denotes the simultaneous *capture avoiding substitution* of X_i in t with s_i, for $i = 1, \ldots, n$. A name X is *guarded*[2] in a term t if every free occurrence of X in t occurs within a context the following forms: $a.[\cdot]$, $s +_1 [\cdot]$, or $[\cdot] +_0 s$.

Since from now on we will only refer to terms constructed over the signature Σ, we will simply write $\mathbb{T}(M)$ and $\mathbb{T}$, in place of $\mathbb{T}(\Sigma, M)$ and $\mathbb{T}(\Sigma, \emptyset)$, respectively.

Before giving the interpretation for these operations in **OMC**, we define an operator on open Markov chains, taking $\mathcal{M} = (M, \tau)$ to the open Markov chain $\mathbb{U}(\mathcal{M}) = (\mathbb{T}(M), \mu_{\mathcal{M}})$, where $\mu_{\mathcal{M}} \colon \mathbb{T}(M) \to \mathcal{D}((\mathcal{L} \times \mathbb{T}(M)) \uplus \mathcal{X})$ is defined as the least solution (over the complete

[2] This notion, coincides with the one in [19], though our definition may seem more involved due to the fact that we allow the probabilistic choice operators $+_e$ with e ranging in the closed interval $[0, 1]$.

partial order of the set of transition probability functions over $\mathbb{T}(M)$, ordered point-wise as $\tau \sqsubseteq \tau'$ iff $\tau(t)(E) \leq \tau'(t)(E)$, for all $t \in \mathbb{T}(M)$, and $E \subseteq (\mathcal{L} \times \mathbb{T}(M)) \uplus \mathcal{X})$ of the equation

$$\mu_{\mathcal{M}} = \mathcal{P}_{\mathcal{M}}(\mu_{\mathcal{M}}) \,,$$

where $\mathcal{P}_{\mathcal{M}}$ is defined by induction on $\mathbb{T}(M)$, for arbitrary transition probability functions θ over $\mathbb{T}(M)$, as follows:

$$\mathcal{P}_{\mathcal{M}}(\theta)(m) = \tau(m) \qquad \mathcal{P}_{\mathcal{M}}(\theta)(a.t) = \delta_{(a,t)}$$
$$\mathcal{P}_{\mathcal{M}}(\theta)(X) = \delta_X \qquad \mathcal{P}_{\mathcal{M}}(\theta)(t +_e s) = e\theta(t) + (1-e)\theta(s)$$
$$\mathcal{P}_{\mathcal{M}}(\theta)(\mathsf{rec}\, X.t) = \theta(t[\mathsf{rec}\, X.t/X]) \,,$$

where δ_X and $\delta_{(a,t)}$ denote the Dirac distributions pointed at $X \in \mathcal{X}$ and $(a,t) \in \mathcal{L} \times \mathbb{T}(M)$, respectively. The definition of $\mu_{\mathcal{M}}$ corresponds essentially to the transition probability of the operational semantics of probabilistic processes given by Stark and Smolka in [19]. The only difference with their definition is that $\mu_{\mathcal{M}}$ is defined over $\mathbb{T}(M)$ rather than $\mathbb{T}(\emptyset)$; and that our formulation simplifies theirs by skipping the definition of a labelled transition system. We refer the interested reader to [19] for more information on the definition of $\mu_{\mathcal{M}}$. Here we limit ourself to recalling that $\mu_{\mathcal{M}}(\mathsf{rec}\, X.X)((a,t)) = 0$, for all $a \in \mathcal{L}$ and $t \in \mathbb{T}(M)$ and $\mu_{\mathcal{M}}(\mathsf{rec}\, X.X)(Y) = 0$, for all $Y \in \mathcal{X}$, that is, $\mathsf{rec}\, X.X$ is a terminating state in $\mathbb{U}(\mathcal{M})$.

▶ **Definition 6** (Universal open Markov chain)**.** Let $\mathcal{M}_\emptyset = (\emptyset, \tau_\emptyset)$ be the open Markov chain with $\tau_\emptyset$ the empty transition function. The *universal open Markov chain* is $\mathbb{U}(\mathcal{M}_\emptyset)$.

The reason why it is called universal will be clarified later. As for now just note that $\mathbb{U}(\mathcal{M}_\emptyset)$ has $\mathbb{T}$ as the set of states, and that its transition probability function corresponds to the one defined in [19]. To ease the notation we will denote $\mathbb{U}(\mathcal{M}_\emptyset)$ as $\mathbb{U} = (\mathbb{T}, \mu_{\mathbb{T}})$.

Next we give an algebraic interpretation over **OMC** to the operations in Σ. For arbitrary $(\mathcal{M}, m), (\mathcal{N}, n) \in$ **OMC** and $f \colon n \in \Sigma$ define $f^{\mathsf{omc}} \colon \mathbf{OMC}^n \to \mathbf{OMC}$ as follows:

$$X^{\mathsf{omc}} = (\mathbb{U}, X) \,, \qquad (\mathcal{M}, m) +_e^{\mathsf{omc}} (\mathcal{N}, n) = (\mathbb{U}(\mathcal{M} \oplus \mathcal{N}), m +_e n) \,,$$
$$(a.(\mathcal{M}, m))^{\mathsf{omc}} = (\mathbb{U}(\mathcal{M}), a.m) \,, \qquad (\mathsf{rec}\, X.(M, m))^{\mathsf{omc}} = (\mathbb{U}(\mathcal{M}^*_{X,m}), \mathsf{rec}\, X.m) \,,$$

where $\mathcal{M} \oplus \mathcal{N}$ denotes the disjoint union of $\mathcal{M}$ and $\mathcal{N}$, and for $\mathcal{M} = (M, \tau)$, $\mathcal{M}^*_{X,m}$ is the open Markov chain $(M, \tau^*_{X,m})$ with transition function defined, for all $m' \in M$ and $E \subseteq (\mathcal{L} \times M) \uplus \mathcal{X}$, as $\tau^*_{X,m}(m')(E) = \tau(m')(X)\tau(m)(E \setminus \{X\}) + (1 - \tau(m')(X))\tau(m')(E \setminus \{X\})$. Intuitively, $\tau^*_{X,m}$ modifies τ by removing the name $X \in \mathcal{X}$ from the support of all $\tau(m')$ and replacing the removed probability mass with the probabilistic behavior of m.

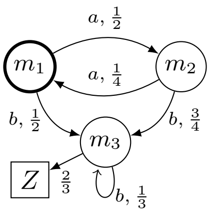
▶ **Definition 7.** The *quantitative algebra of open Markov chains* is $(\mathbf{OMC}, \Sigma^{\mathsf{omc}}, \mathbf{d_{OMC}})$.

The (initial) semantics for terms $t \in \mathbb{T}$ to pointed open Markov chains is given via the function $[\![\cdot]\!] \colon \mathbb{T} \to$ **OMC**, defined by induction on terms as follows

$$[\![X]\!] = X^{\mathsf{omc}} \qquad [\![t +_e s]\!] = [\![t]\!] +_e^{\mathsf{omc}} [\![s]\!] \,,$$
$$[\![a.t]\!] = (a.[\![t]\!])^{\mathsf{omc}} \,, \qquad [\![\mathsf{rec}\, X.t]\!] = (\mathsf{rec}\, X.[\![t]\!])^{\mathsf{omc}} \,. \qquad \text{(SEMANTICS)}$$

For an example of how a term is interpreted to a pointed open Markov chain see Figure 1.

Note that the freely-generated algebra of Σ-terms, namely $(\mathbb{T}, \Sigma)$ can be turned into a quantitative algebra as $\mathbb{U} = (\mathbb{T}, \Sigma, \mathbf{d}_{\mathbb{U}})$, where $\mathbf{d}_{\mathbb{U}}$ is the bisimilarity distance defined over the universal open Markov chain $\mathbb{U}$. The next result states the universality of $\mathbb{U}$.



$$t_1 = \mathsf{rec}\, X.(a.t_2 +_{\frac{1}{2}} b.t_3)$$
$$t_2 = a.X +_{\frac{1}{4}} b.t_3$$
$$t_3 = \mathsf{rec}\, Y.(b.Y +_{\frac{1}{3}} Z)$$

Figure 1 The term t_1 is interpreted to the pointed open Markov chain $[\![t_1]\!] \sim (\mathcal{M}, m_1)$ depicted on the left (restricted only to the states reachable from t_1).

▶ **Theorem 8** (Universality). *Let $t, s \in \mathbb{T}$. Then $[\![t]\!] \sim (\mathbb{U}, t)$ and $\mathbf{d}_{\mathbf{OMC}}([\![t]\!], [\![s]\!]) = \mathbf{d}_{\mathbb{U}}(t, s)$.*

Proof (sketch). The proof of $[\![t]\!] \sim (\mathbb{U}, t)$ is by induction on t. The base case is trivial. The cases for the prefix and probabilistic choice operations are completely routine from the definition of the interpretations and the operator $\mathbb{U}\colon \mathbf{OMC} \to \mathbf{OMC}$ (in each case a bisimulation can be constructed from those given by the inductive hypothesis). The only nontrivial case is when $t = \mathsf{rec}\, X.t'$. The proof carries over in two steps. First one shows that $(\mathbb{U}, \mathsf{rec}\, X.t') \sim (\mathsf{rec}\, X.(\mathbb{U}, t'))^{\mathsf{omc}}$; then, by using the inductive hypothesis $[\![t']\!] \sim (\mathbb{U}, t')$, that $(\mathsf{rec}\, X.(\mathbb{U}, t'))^{\mathsf{omc}} \sim (\mathsf{rec}\, X.[\![t']\!])^{\mathsf{omc}}$. Since $[\![\mathsf{rec}\, X.t']\!] = (\mathsf{rec}\, X.[\![t']\!])^{\mathsf{omc}}$, by transitivity of the bisimilarity relation $[\![\mathsf{rec}\, X.t']\!] \sim (\mathbb{U}, \mathsf{rec}\, X.t')$.

The proof of $\mathbf{d}_{\mathbf{OMC}}([\![t]\!], [\![s]\!]) = \mathbf{d}_{\mathbb{U}}(t, s)$ follows by Lemma 5 and the above result. Indeed

$$\begin{aligned}
\mathbf{d}_{\mathbf{OMC}}([\![t]\!], [\![s]\!]) &= \mathbf{d}([\![t]\!], [\![s]\!]) && \text{(def. } \mathbf{d}) \\
&\leq \mathbf{d}([\![t]\!], (\mathbb{U}, t)) + \mathbf{d}((\mathbb{U}, t), (\mathbb{U}, s)) + \mathbf{d}((\mathbb{U}, s), [\![s]\!]) && \text{(triangular ineq.)} \\
&= \mathbf{d}((\mathbb{U}, t), (\mathbb{U}, s)) && ([\![t]\!] \sim (\mathbb{U}, t),\ [\![s]\!] \sim (\mathbb{U}, s) \ \&\ \text{Lemma 5}) \\
&= \mathbf{d}_{\mathbb{U}}(t, s)\,. && \text{(def. } \mathbf{d})
\end{aligned}$$

By a similar argument we also have $\mathbf{d}_{\mathbf{OMC}}([\![t]\!], [\![s]\!]) \geq \mathbf{d}_{\mathbb{U}}(t, s)$, hence the thesis. ◀

The above result states that it is totally equivalent to reason about the behavior of $[\![t]\!]$ by just considering the state corresponding to the term t in the universal model $\mathbb{U}$. Hence, due to Theorem 8, in the rest of the paper whenever we refer to the distance between two terms we will use $\mathbf{d}_{\mathbb{U}}$, often simply denoted as $\mathbf{d}$. Similarly, $\Gamma \models_{\mathbf{OMC}} t \equiv_{\varepsilon} s$ is equivalent to $\Gamma \models_{\mathbb{U}} t \equiv_{\varepsilon} s$, and it will be denoted just by $\Gamma \models t \equiv_{\varepsilon} s$.

▶ Remark. We already noted that the universal open Markov chain $\mathbb{U}$ corresponds to the operational semantics of probabilistic expressions given by Stark and Smolka [19]. In the light of Theorem 8, the soundness and completeness results for axiomatic equational system w.r.t. probabilistic bisimilarity over probabilistic expressions given in [19], can be moved without further efforts to the class of open Markov chains of the form $[\![t]\!]$.

5 Axiomatization of the Bisimilarity Distance

This section presents a quantitative deductive system, namely the one satisfying the axioms in Figure 2, and prove it to be sound and complete w.r.t. the bisimilarity distance $\mathbf{d}$.

The axioms (B1), (B2), (SC), (SA) are that of barycentric algebras [20], used to axiomatize probability distributions. The axioms (Unfold), (Unguard), (Fix), (Cong) are used to axiomatize recursive behaviors and correspond to those proposed by Milner [15]. All together, these axioms have been used by Stark and Smolka [19] to provide a complete axiomatization of probabilistic bisimilarity. To this set of axioms we add the axiom (IB)

$$
\begin{aligned}
&(\mathsf{B1}) && \vdash t +_1 s \equiv_0 t\,,\\
&(\mathsf{B2}) && \vdash t +_e t \equiv_0 t\,,\\
&(\mathsf{SC}) && \vdash t +_e s \equiv_0 s +_{1-e} t\,,\\
&(\mathsf{SA}) && \vdash (t +_e s) +_{e'} u \equiv_0 t +_{ee'} (s +_{\frac{e'-ee'}{1-ee'}} u)\,, \text{ for } e, e' \in [0,1)\,,\\
&(\mathsf{Unfold}) && \vdash \mathsf{rec}\, X.t \equiv_0 t[\mathsf{rec}\, X.t/X]\,,\\
&(\mathsf{Unguard}) && \vdash \mathsf{rec}\, X.t +_e X \equiv_0 \mathsf{rec}\, X.t\,,\\
&(\mathsf{Fix}) && \{s \equiv_0 t[s/X]\} \vdash s \equiv_0 \mathsf{rec}\, X.t\,, \text{ for } X \text{ guarded in } t,\\
&(\mathsf{Cong}) && \{t \equiv_0 s\} \vdash \mathsf{rec}\, X.t \equiv_0 \mathsf{rec}\, X.s\,,\\
&(\mathsf{Top}) && \vdash t \equiv_1 s\,,\\
&(\mathsf{IB}) && \{t \equiv_\varepsilon s, t' \equiv_{\varepsilon'} s'\} \vdash t +_e t' \equiv_{\varepsilon''} s +_e s'\,, \text{ for } \varepsilon'' \geq e\varepsilon + (1-e)\varepsilon'.
\end{aligned}
$$

Figure 2 Quantitative axioms for the bisimilarity pseudometric.

of [14], that, in combination with the barycentric axioms, axiomatizes the Kantorovich distance between finitely-supported probability distributions (see §10 in [14] for more details). Finally the axiom (Top) is used to bound the distance between terms.

A significant difference w.r.t. the original framework of quantitative deductive systems of Mardare, Panangaden, and Plotkin, recalled in Section 3, is that we do not impose non-expansiveness for the operator $\mathsf{rec}\, X$ (i.e., the axiom (NExp) associated to $\mathsf{rec}\, X$ is dropped). This is replaced by the weaker axiom (Cong). The intuitive reason why (NExp) is not sound for $\mathsf{rec}\, X$ is that the recursion magnifies the differences of the behaviors of its arguments. We refer the interested reader to [9] for an exhaustive explanation of this phenomenon.

By [20, Theorem 2], any barycentric algebra has a one-to-one embedding into a convex subset of a suitable vector space. By this result, we can conveniently its elements as n-ary convex combinations of terms $t_1, \dots, t_n \in \mathbb{T}$, as $\sum_{i=1}^n e_i \cdot t_i \in \mathbb{T}$, provided that $e_i \in [0,1]$ and $\sum_{i=1}^n e_i = 1$. We refer the reader to [19, 10, 11] for an analytic discussion of this notation.

5.1 Soundness

In this section we show the soundness of our quantitative deductive system w.r.t. the bisimilarity distance. As noticed in Remark 4.2, the soundness of the axioms already present in the deductive system of Stark and Smolka follow without further changes from [19], due to Theorem 8 and Lemma 5.

▶ **Theorem 9** (Soundness). *If* $\vdash t \equiv_\varepsilon s$ *then* $\models t \equiv_\varepsilon s$.

Proof. As usual, we must show that each axiom and rule of inference is valid. The axioms (Refl), (Symm), (Triang), (Max), and (Arch) are sound since $\mathbf{d}$ is a pseudometric (Lemma 5). The soundness of the classical logical deduction rules (Subst), (Cut), and (Assum) is immediate. By Lemma 5, the kernel of $\mathbf{d}$ is $\sim$. Hence the axioms of barycentric algebras (B1), (B2), (SC), and (SA) all along with the axioms (Unfold), (Unguard), (Cong), and (Fix) follow directly by the soundness theorem proven in [19]. The axiom (Top) is immediate consequence of the fact that $\mathbf{d}$ is 1-bounded. Note that (IB) subsumes the axiom (NExp$+_e$) —the two coincide when $\varepsilon = \varepsilon'$. It only remains to show the soundness of (NExp-pref) for the prefix operator and (IB). To prove (NExp-pref) it suffices to show that $\mathbf{d}(t,s) \geq \mathbf{d}(a.t, a.s)$:

$$
\begin{aligned}
\mathbf{d}(a.t, a.s) &= \mathcal{K}(\Lambda(\mathbf{d}))(\mu^*_{\mathbb{T}}(a.t), \mu^*_{\mathbb{T}}(a.s)) && (\mathbf{d} \text{ fixed-point \& def. } \Psi) \\
&= \mathcal{K}(\Lambda(\mathbf{d}))(\delta_{(a,t)}, \delta_{(a,s)}) && (\text{def. } \mu_{\mathbb{T}} \text{ \& } \mathcal{P}_{\mathbb{U}}) \\
&= \Lambda(\mathbf{d})((a,t),(a,s)) && (\text{def. } \mathcal{K}) \\
&= \mathbf{d}(t,s)\,. && (\text{def. } \Lambda)
\end{aligned}
$$

Finally, the soundness of (IB) follows by $e\mathbf{d}(t,s) + (1-e)\mathbf{d}(t',s') \geq \mathbf{d}(t +_e t', s +_e s')$

$$
\begin{aligned}
&e\mathbf{d}(t,s) + (1-e)\mathbf{d}(t',s') \\
&\quad = e\Psi(\mathbf{d})(t,s) + (1-e)\Psi(\mathbf{d})(t',s') && (\mathbf{d} \text{ fixed point}) \\
&\quad = e\mathcal{K}(\Lambda(\mathbf{d}))(\mu^*_{\mathbb{T}}(t), \mu^*_{\mathbb{T}}(s)) + (1-e)\mathcal{K}(\Lambda(\mathbf{d}))(\mu^*_{\mathbb{T}}(t'), \mu^*_{\mathbb{T}}(s')) && (\text{def. } \Psi)
\end{aligned}
$$

then, for $\omega \in \Omega(\mu^*_{\mathbb{T}}(t), \mu^*_{\mathbb{T}}(s))$ and $\omega' \in \Omega(\mu^*_{\mathbb{T}}(t'), \mu^*_{\mathbb{T}}(s'))$ optimal couplings for $\mathcal{K}(\Lambda(\mathbf{d}))$, and by noticing that $e\omega + (1-e)\omega' \in \Omega(e\mu^*_{\mathbb{T}}(t) + (1-e)\mu^*_{\mathbb{T}}(t'), e\mu^*_{\mathbb{T}}(s) + (1-e)\mu^*_{\mathbb{T}}(s'))$ we have

$$
\begin{aligned}
&= e \textstyle\int \Lambda(\mathbf{d})\,\mathrm{d}\omega + (1-e)\int \Lambda(\mathbf{d})\,\mathrm{d}\omega' \\
&= \textstyle\int \Lambda(\mathbf{d})\,\mathrm{d}(e\omega + (1-e)\omega') && (\text{linearity}) \\
&\geq \mathcal{K}(\Lambda(\mathbf{d}))(e\mu^*_{\mathbb{T}}(t) + (1-e)\mu_{\mathbb{T}}(t'), e\mu^*_{\mathbb{T}}(s) + (1-e)\mu_{\mathbb{T}}(s')) && (\text{def. } \mathcal{K} \text{ and above}) \\
&= \mathcal{K}(\Lambda(\mathbf{d}))(\mathcal{P}_{\mathbb{U}}(\mu_{\mathbb{T}})^*(t +_e t'), \mathcal{P}_{\mathbb{U}}(\mu_{\mathbb{T}})^*(s +_e s')) && (\text{def. } \mathcal{P}_{\mathbb{U}}) \\
&= \mathcal{K}(\Lambda(\mathbf{d}))(\mu^*_{\mathbb{T}}(t +_e t'), \mu^*_{\mathbb{T}}(s +_e s')) && (\mu_{\mathbb{T}} \text{ fixed-point}) \\
&= \mathbf{d}(t +_e t', s +_e s') && (\text{def. } \Psi \text{ \& } \delta \text{ fixed-point})
\end{aligned}
$$

The above concludes the proof. ◀

5.2 Completeness

This section is devoted to prove the completeness of our axiomatization w.r.t. the bisimilarity distance. The proof relies on the completeness theorem of the axiomatization of probabilistic bisimilarity in [19], and the one for interpolative barycentric algebras in [14].

The next theorem, due to Milner [15], and restated in the probabilistic setting by Stark and Smolka [19, Theorem 1] is essential for proving the completeness of our axiomatization.

▶ **Theorem 10** (Unique Solution of Equations). *Let $\overline{X} = (X_1, \dots, X_k)$ and $\overline{Y} = (Y_1, \dots, Y_h)$ be distinct names, and $\overline{t} = (t_1, \dots, t_k)$ terms with free names in $(\overline{X}, \overline{Y})$ in which each X_i is guarded. Then there exist terms $\overline{s} = (s_1, \dots, s_k)$ with free names in $\overline{Y}$ such that*

$$\vdash s_i \equiv_0 t[\overline{s}/\overline{X}]\,, \qquad \text{for all } i \leq k.$$

Moreover, if for some terms $\overline{u} = (u_1, \dots, u_k)$ with free variables in $\overline{Y}$, $\vdash u_i \equiv_0 t[\overline{u}/\overline{X}]$, for all $i \leq k$, then $\vdash s_i \equiv_0 u_i$, for all $i \leq k$.

The next theorem is the equational characterization theorem of Stark and Smolka. In our formulation the statement is simpler than [19, Theorem 2] since in our axiomatization we have the unit laws for $+_1$ and $+_0$, derivable from the axioms (B1) and (SC).

▶ **Theorem 11** (Equational Characterization). *For any term t, with free names in $\overline{Y}$, there exist terms $t_1, \dots, t_k$ with free names in $\overline{Y}$, such that $\vdash t \equiv_0 t_1$ and*

$$\vdash t_i \equiv_0 \textstyle\sum_{j=1}^{h(i)} p_{ij} \cdot s_{ij} + \sum_{j=1}^{l(i)} q_{ij} \cdot Y_{g(i,j)}\,, \qquad \text{for all } i \leq k,$$

where the terms s_{ij} and names $Y_{g(i,j)}$ are enumerated without repetitions, and s_{ij} is either $\mathsf{rec}\,X.X$ or has the form $a_{ij}.t_{f(i,j)}$.

Recall that (NExp) is not sound for the recursion operator. Nevertheless, the completeness of the axiomatization can be carried out regardless, thanks to the fact that the bisimilarity distance can alternatively be obtained as $\delta = \bigsqcap_{k\in\omega} \tilde{\Psi}^k(\mathbf{1})$, i.e., as the ω-limit of the decreasing sequence $\mathbf{1} \sqsupseteq \tilde{\Psi}(\mathbf{1}) \sqsupseteq \tilde{\Psi}^2(\mathbf{1}) \sqsupseteq \dots$ of the operator

$$\tilde{\Psi}(d)(m,m') = \begin{cases} 0 & \text{if } m \sim m', \\ \Psi(d)(m,m') & \text{otherwise.} \end{cases}$$

▶ **Lemma 12.** *The operator $\tilde{\Psi}$ is monotone and ω-cocontinuous. Moreover, $\delta = \bigsqcap_{k\in\omega} \tilde{\Psi}^k(\mathbf{1})$.*

Proof. Monotonicity and ω-cocontinuity follow similarly to Lemma 4 and [21, Theorem 1]. By ω-cocontinuity $\bigsqcap_{k\in\omega} \tilde{\Psi}^k(\mathbf{1})$ is a fixed point. By Lemma 5 and $\delta = \Psi(\delta)$, also δ is a fixed point of $\tilde{\Psi}$. We show that they coincide by proving that $\tilde{\Psi}$ has a unique fixed point.

Assume that $\tilde{\Psi}$ has two fixed points d and d' such that $d \sqsubset d'$. Define $R \subseteq M \times M$ as $m\,R\,m'$ iff $d'(m,m') - d(m,m') = \|d' - d\|$. By the assumption made on d and d' we have that $\|d' - d\| > 0$ and $R \cap \sim\, = \emptyset$. Consider arbitrary $m, m' \in M$ such that $m\,R\,m'$, then

$$\begin{aligned} \|d' - d\| &= d'(m,m') - d(m,m') \\ &= \tilde{\Psi}(d')(m,m') - \tilde{\Psi}(d)(m,m') && \text{(by } d = \tilde{\Psi}(d) \text{ and } d' = \tilde{\Psi}(d')) \\ &= \Psi(d')(m,m') - \Psi(d)(m,m') && \text{(by } m \not\sim m' \text{ and def. } \tilde{\Psi}) \\ &\le \textstyle\int_E (\Lambda(d') - \Lambda(d))\, \mathrm{d}\omega\,, && \text{(as proved in Lemma 4)} \end{aligned}$$

where we recall that $E = \{((a,n),(a,n')) \mid a \in \mathcal{L}, n, n' \in M\}$.

Observe that $(\Lambda(d') - \Lambda(d))((a,n),(a,n')) = d'(n) - d(n') \le \|d' - d\|$, for all $n, n' \in M$ and $a \in \mathcal{L}$. Since $\|d' - d\| > 0$ the inequality $\|d' - d\| \le \int_E (\Lambda(d') - \Lambda(d))\, \mathrm{d}\omega \le \|d' - d\|$ holds only if the support of ω is included in $E_R = \{((a,n),(a,n')) \mid a \in \mathcal{L} \text{ and } n\,R\,n'\}$. Since the argument holds for arbitrary $m, m' \in M$ such that $m\,R\,m'$, we have that R is a bisimulation, which is in contradiction with the initial assumptions. ◀

Now we are ready to prove the main result of this section.

▶ **Theorem 13** (Completeness). *If $\models t \equiv_\varepsilon s$, then $\vdash t \equiv_\varepsilon s$.*

Proof. Let $t, s \in \mathbb{T}$ and $\varepsilon \in \mathbb{Q}_+$. We have to show that if $\mathbf{d}(t,s) \le \varepsilon$ then $\vdash t \equiv_\varepsilon s$. The case $\varepsilon \ge 1$ trivially follows by (Top) and (Max). Let $\varepsilon < 1$. By Theorem 11, there exist terms $t_1, \dots, t_k$ and $s_1, \dots, s_r$ with free names in $\overline{X}$ and $\overline{Y}$, respectively, such that $\vdash t \equiv_0 t_1$, $\vdash s \equiv_0 s_1$, and

$$\vdash t_i \equiv_0 \textstyle\sum_{j=1}^{h(i)} p_{ij} \cdot t'_{ij} + \sum_{j=1}^{l(i)} q_{ij} \cdot X_{g(i,j)}\,, \qquad \text{for all } i \le k, \tag{1}$$

$$\vdash s_u \equiv_0 \textstyle\sum_{v=1}^{n(u)} e_{uv} \cdot s'_{uv} + \sum_{v=1}^{m(u)} d_{uv} \cdot Y_{w(u,v)}\,, \qquad \text{for all } u \le r, \tag{2}$$

where the terms t'_{ij} (resp. s'_{uv}) and names $X_{g(i,j)}$ (resp. $Y_{w(u,v)}$) are enumerated without repetitions, and t'_{ij} (resp. s'_{uv}) have either the form $a_{ij}.t'_{f(i,j)}$ (resp. $b_{uv}.s'_{z(u,v)}$) or rec $Z.Z$. By induction on $\alpha \in \mathbb{N}$, we prove that

$$\vdash t_i \equiv_\varepsilon s_u\,, \qquad \text{for all } i \le k,\ u \le r, \text{ and } \varepsilon \ge \tilde{\Psi}^\alpha(\mathbf{1})(t_i, s_u)\,. \tag{3}$$

(Base case: $\alpha = 0$) $\tilde{\Psi}^0(\mathbf{1})(t_i, s_u) = \mathbf{1}(t_i, s_u)$. Since $\mathbf{1}(t_i, s_u) = 0$ whenever $t_i = s_u$ and $\mathbf{1}(t_i, s_u) = 1$ if $t_i \neq s_u$, then (3) follows by the axioms (Refl), (Top) and (Max).
(Inductive step: $\alpha \ge 0$). Assume that (3) holds for α. We want to show $\vdash t_i \equiv_\varepsilon s_u$, for all $\varepsilon \ge \tilde{\Psi}^{\alpha+1}(\mathbf{1})(t_i, s_u)$. Since our deduction system includes the one of Stark and Smolka,

whenever $t_i \sim_{\mathbb{U}} s_u$, by completeness w.r.t. $\sim_{\mathbb{U}}$, namely [19, Theorem 3], we obtain $\vdash t_i \equiv_0 s_u$. By (Max) $\vdash t_i \equiv_\varepsilon s_u$, for all $\varepsilon \geq \tilde{\Psi}^{\alpha+1}(\mathbf{1})(t_i, s_u) = 0$. Let consider the case $t_i \not\sim_{\mathbb{U}} s_u$. By inductive hypothesis, we have that (3) holds. For each name $X_{g(i,j)}$ and $Y_{w(u,v)}$ occurring in (1) and (2), respectively, by (Top) we have $\vdash X_{g(i,j)} \equiv_1 Y_{w(u,v)}$ whenever $X_{g(i,j)} \neq Y_{w(u,v)}$, and by (Refl) we have $\vdash X_{g(i,j)} \equiv_0 Y_{w(u,v)}$ whenever $X_{g(i,j)} = Y_{w(u,v)}$. For each term $a_{ij}.t'_{f(i,j)}$ and $b_{uv}.s'_{z(u,v)}$ occurring in (1) and (2), respectively, by inductive hypothesis and (NExp) we can deduce $\vdash a_{ij}.t'_{f(i,j)} \equiv_\varepsilon b_{uv}.s'_{z(u,v)}$, for all $\varepsilon \geq \tilde{\Psi}^{\alpha}(\mathbf{1})(t'_{f(i,j)}, s'_{z(u,v)})$, whenever $a_{ij} = b_{uv}$. If $a_{ij} \neq b_{uv}$, by (Top) we get $\vdash a_{ij}.t'_{f(i,j)} \equiv_1 b_{uv}.s'_{z(u,v)}$. As for $\mathsf{rec}\, Z.Z$, by (Top) we have $\vdash \mathsf{rec}\, Z.Z \equiv_1 \beta$ for all terms $\beta \neq \mathsf{rec}\, Z.Z$ occurring in the right hand side of (1) and (2); and by (Refl) we have $\vdash \mathsf{rec}\, Z.Z \equiv_0 \mathsf{rec}\, Z.Z$.

Note that in this manner —possibly using (Max)— we have deduced $\vdash \beta \equiv_{\varepsilon'} \gamma$, for all $\varepsilon' \geq \Lambda(\tilde{\Psi}^{\alpha}(\mathbf{1}))(\beta, \gamma)$, where β and γ are arbitrary terms occurring in the right hand side of (1) and (2), respectively. Since our quantitative deductive system includes all the axioms of interpolative barycentric algebras in the sense of [14], by completeness w.r.t. the Kantorovich distance (see §10 in [14]), for all $t_i \not\sim_{\mathbb{U}} s_u$,

$$\vdash t_i \equiv_\varepsilon s_u\,, \qquad \text{for all } \varepsilon \geq \mathcal{K}(\Lambda(\tilde{\Psi}^{\alpha}(\mathbf{1})))(\mu^*_{\mathbb{T}}(t_i), \mu^*_{\mathbb{T}}(s_u)) = \tilde{\Psi}^{\alpha+1}(\mathbf{1})(t_i, s_u)\,. \tag{4}$$

By Lemma 12 and (3), applying (Arch) we have $\vdash t_i \equiv_\varepsilon s_u$, for all $\varepsilon \geq \mathbf{d}(t_i, s_u)$. By $\vdash t \equiv_0 t_1$, $\vdash s \equiv_0 s_1$, and (Triang), we deduce $\vdash t \equiv_\varepsilon s$, for all $\varepsilon \geq \mathbf{d}(t, s)$. ◀

6 The Class of Expressible Open Markov Chains

In this last section we show that the class of expressible open Markov chains corresponds up to bisimilarity to the class of finite (and finitely supported) open Markov chains. Specifically, this means that any finite open Markov chain (hence, also "closed" Markov chains) can be represented, up to bisimilarity, as Σ-terms; so that by Theorems 9 and 13 we can reason about their quantitative operational semantics in a purely algebraic way via the axiomatic system presented in Section 5[3].

A pointed Markov chain $(\mathcal{M}, m)$ is said *expressible* if there exists a term $t \in \mathbb{T}$ such that $[\![t]\!] \sim (\mathcal{M}, m)$. The next result is a corollary of Theorems 8, 10, and 9.

▶ **Corollary 14.** *If $(\mathcal{M}, m)$ is finite then it is expressible.*

Proof. We have to show that there exists $t \in \mathbb{T}$ such that $[\![t]\!] \sim (\mathcal{M}, m)$. Since the set of states $M = \{m_1, \dots, m_k\}$ is finite and, for each $m_i \in M$, $\tau(m_i)$ is finitely supported, then the sets of unguarded names $\{Y^i_1, \dots, Y^i_{h(i)}\} = supp(\tau(m_i)) \cap \mathcal{X}$ and labelled transitions $\{\alpha^i_1, \dots, \alpha^i_{l(i)}\} = supp(m_i) \cap (\mathcal{L} \times M)$ of m_i are finite. Let us associate with each α^i_j a name X^i_j, for all $i \leq k$ and $j \leq l(i)$. For each $i \leq k$, we define the terms

$$t_i = \textstyle\sum_{j=1}^{l(i)} \tau(m_i)(\alpha^i_j) \cdot a^i_j.X^i_j + \sum_{j=1}^{h(i)} \tau(m_i)(Y^i_j) \cdot Y^i_j\,,$$

where $\alpha^i_j = (a^i_j, m^i_j)$, for all $i \leq k$ and $i \leq l(i)$. By Theorem 10, for $i \leq k$, there exists terms $\overline{s^i} = (s^i_1, \dots, s^i_{l(i)})$ such that $\vdash s_i \equiv_0 t_i[\overline{s^i}/\overline{X^i}]$, so that by soundness (Theorem 9), $[\![s_i]\!] \sim [\![t_i[\overline{s^i}/\overline{X^i}]]\!]$. Hence, by Theorem 8, we have $(\mathbb{U}, s_i) \sim (\mathbb{U}, t_i[\overline{s^i}/\overline{X^i}])$.

Let $\overline{m^i} = (m^i_1, \dots, m^i_{l(i)})$ and $\overline{X^i} = (X^i_1, \dots, X^i_{l(i)})$, for $i \leq k$. It is a routine check to prove that the smallest equivalence relation R_i containing $\{(m_i, t_i[\overline{m^i}/\overline{X^i}]) \mid i \leq k\}$ is

[3] The results in this section can be alternatively obtained as in [18] by observing that open Markov chains are coalgebras of a quantitative functor.

a bisimulation for $(\mathcal{M}, m_i)$ and $(\mathbb{U}(\mathcal{M}), t_i[\overline{m^i}/\overline{X^i}])$, hence $(\mathcal{M}, m_i) \sim (\mathbb{U}(\mathcal{M}), t_i[\overline{m^i}/\overline{X^i}])$. Similarly, one can prove $(\mathbb{U}(\mathcal{M}), t_i[\overline{m^i}/\overline{X^i}]) \sim (\mathbb{U}, t_i[\overline{s^i}/\overline{X^i}])$ by taking the smallest equivalence relation containing $\{(t_i[\overline{m^i}/\overline{X^i}], t_i[\overline{s^i}/\overline{X^i}]) \mid i \leq k\}$ and $\{(m^i_j, s^i_j) \mid i \leq k, j \leq l(i)\}$. By transitivity of $\sim$, $(\mathcal{M}, m_i) \sim [\![s_i]\!]$, for all $i \leq k$, hence $(\mathcal{M}, m)$ is expressible. ◀

The converse (up to bisimilarity) of the above result can also be proved, and it follows as a corollary of Theorems 8, 9, and 11.

▶ **Corollary 15.** *If $(\mathcal{M}, n)$ is expressible then it is finite up-to-bisimilarity.*

Proof. Let $t \in \mathbb{T}$. We have to show that there exists $(\mathcal{M}, m) \in \mathbf{OMC}$ with a finite set of states such that $[\![t]\!] \sim (\mathcal{M}, m)$. From Theorem 11, there exist $t_1, \dots, t_k$ with free names in $\overline{Y}$, such that $\vdash t \equiv_0 t_1$ and

$$\vdash t_i \equiv_0 \textstyle\sum_{j=1}^{h(i)} p_{ij} \cdot s_{ij} + \sum_{j=1}^{l(i)} q_{ij} \cdot Y_{g(i,j)}\,, \qquad \text{for all } i \leq k,$$

where the terms s_{ij} and names $Y_{g(i,j)}$ are enumerated without repetitions, and s_{ij} is either $\mathsf{rec}\,X.X$ or has the form $a_{ij}.t_{f(i,j)}$. Let $Z_1, \dots, Z_k$ be fresh names distinct from $\overline{Y}$, and define t'_i as the term obtained by replacing in the right end side of the equation above each occurrence of t_i with Z_i. Then, clearly $\vdash t_i \equiv_0 t'_i[\overline{t}/\overline{Z}]$. By soundness (Theorem 9), we have that $[\![t_i]\!] \sim [\![t'_i[\overline{t}/\overline{Z}]]\!]$, so that, by Theorem 8, $(\mathbb{U}, t_i) \sim (\mathbb{U}, t'_i[\overline{t}/\overline{Z}])$.

Define $\mathcal{M} = (M, \tau)$ by setting $M = \{t_1, \dots, t_k\}$, $m = t_1$, and, for all $i \leq k$, taking as $\tau(t_i)$ the smallest sub-probability distribution on $(\mathcal{L} \times M) \uplus \mathcal{X}$ such that $\tau(t_i)((a_{ij}, t_{f(i,j)})) = p_{ij}$ and $\tau(t_i)(Y_{g(i,e)}) = q_{ie}$, for all $i \leq k$, $j \leq h(i)$, and $e \leq l(i)$. Notice that since the equation above is without repetitions, τ is well defined. Moreover, $1 - \tau(m_i)((\mathcal{L} \times M) \uplus \mathcal{X}) = p_{iw}$ whenever there exists $w \leq h(i)$ such that $s_{iw} = \mathsf{rec}\,X.X$. It is not difficult to prove that $(\mathcal{M}, t_i) \sim (\mathbb{U}, t'_i[\overline{t}/\overline{Z}])$ (take the smallest equivalence relation containing the pairs $(t_i, t'_i[\overline{t}/\overline{Z}])$, for $i \leq k$), so that by transitivity of $\sim$, $(\mathcal{M}, t_i) \sim [\![t_i]\!]$, for all $i \leq k$. By $\vdash t \equiv_0 t_1$ and Theorem 9, we also have $[\![t]\!] \sim [\![t_1]\!]$, so that $[\![t]\!] \sim (\mathcal{M}, m)$. ◀

7 Conclusions and Future Work

In this paper we proposed a complete axiomatization for the bisimilarity distance of Desharnais et al. The axiomatic system comes as a natural generalization of the one proposed by Stark and Smolka [19] for probabilistic bisimilarity, where we only added the axiom (IB) from [14] for dealing with the Kantorovich distance. Although the use of the recursion operator does not fit the general framework of Mardare et al. [14], we believe that the proof technique employed in the present paper may be general enough to accommodate the axiomatization of other behavioral distances for probabilistic systems, such as the total variation distance. Moreover, in the light of the results in Section 6, it would be interesting to see to what extent one could approach infinitary behaviors by means of finitary ones, and how such an axiomatization would look like. These questions are left open for future work.

Acknowledgments. We thank the anonymous reviewers for their comments and suggestions. The first two authors are indebted to Lotte Legarth for the support provided and the delicious meals.

References

1 Luca Aceto, Zoltán Ésik, and Anna Ingólfsdóttir. Equational axioms for probabilistic bisimilarity. In *AMAST 2002*, LNCS, pages 239–253, 2002.

2 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. Bisimulation on Markov Processes over Arbitrary Measurable Spaces. In *Horizons of the Mind. A Tribute to Prakash Panangaden*, volume 8464 of *LNCS*, pages 76–95. Springer, 2014.
3 Jos C. M. Baeten, Jan A. Bergstra, and Scott A. Smolka. Axiomatizing probabilistic processes: ACP with generative probabilities. *Inf. Comput.*, 121(2):234–255, 1995.
4 Emanuele Bandini and Roberto Segala. Axiomatizations for probabilistic bisimulation. In *ICALP*, LNCS, pages 370–381, 2001.
5 Pedro R. D'Argenio, Daniel Gebler, and Matias D. Lee. Axiomatizing bisimulation equivalences and metrics from probabilistic SOS rules. In *FoSSaCS*, LNCS, pages 289–303, 2014.
6 Yuxin Deng and Catuscia Palamidessi. Axiomatizations for probabilistic finite-state behaviors. *Theor. Comput. Sci.*, 373(1-2):92–114, 2007.
7 Josee Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.
8 Norm Ferns, Prakash Panangaden, and Doina Precup. Metrics for finite Markov Decision Processes. In *Proceedings of the 20th conference on Uncertainty in Artificial Intelligence*, UAI, pages 162–169. AUAI Press, 2004.
9 Daniel Gebler, Kim Guldstrand Larsen, and Simone Tini. Compositional metric reasoning with probabilistic process calculi. In *FoSSaCS*, LNCS, pages 230–245, 2015.
10 C. Jones and Gordon D. Plotkin. A probabilistic powerdomain of evaluations. In *LICS*, pages 186–195. IEEE Computer Society, 1989.
11 Klaus Keimel and Gordon Plotkin. Mixed powerdomains for probability and nondeterminism. *Logical Methods in Computer Science*, page ??, 2015. (to appear).
12 Kim G. Larsen, Uli Fahrenberg, and Claus R. Thrane. Metrics for weighted transition systems: Axiomatization and complexity. *Theor. Comput. Sci.*, 412(28):3358–3369, 2011.
13 Kim Guldstrand Larsen and Arne Skou. Bisimulation Through Probabilistic Testing. In *POPL*, pages 344–352, 1989.
14 Radu Mardare, Prakash Panangaden, and Gordon Plotkin. Quantitative algebraic reasoning. In *LICS*, page ?? IEEE Computer Society, 2016. (to appear).
15 Robin Milner. A complete inference system for a class of regular behaviours. *J. Comput. Syst. Sci.*, 28(3):439–466, 1984.
16 Michael Mislove, Joël Ouaknine, and James Worrell. Axioms for probability and nondeterminism. *Electronic Notes in Theoretical Computer Science*, 96:7–28, 2004.
17 Gordon D. Plotkin and John Power. Semantics for algebraic operations. *Electr. Notes Theor. Comput. Sci.*, 45:332–345, 2001.
18 Alexandra Silva, Filippo Bonchi, Marcello M. Bonsangue, and Jan J. M. M. Rutten. Quantitative kleene coalgebras. *Inf. Comput.*, 209(5):822–849, 2011.
19 Eugene W. Stark and Scott A. Smolka. A complete axiom system for finite-state probabilistic processes. In *Proof, Language, and Interaction, Essays in Honour of Robin Milner*, MIT Press, pages 571–596, 2000.
20 Marshall H. Stone. Postulates for the barycentric calculus. *Annali di Matematica Pura ed Applicata*, 29(1):25–30, 1949.
21 Franck van Breugel. On behavioural pseudometrics and closure ordinals. *Inf. Process. Lett.*, 112(19):715–718, 2012.
22 Franck van Breugel and James Worrell. Towards Quantitative Verification of Probabilistic Transition Systems. In *ICALP*, volume 2076 of *LNCS*, pages 421–432, 2001.