Behavioral Metrics via Functor Lifting

Paolo Baldan[1], Filippo Bonchi[2], Henning Kerstan[3], and Barbara König[3]

1 **Dipartimento di Matematica, Università di Padova, Italy**
baldan@math.unipd.it
2 **CNRS, ENS Lyon, Université de Lyon, France**
filippo.bonchi@ens-lyon.fr
3 **Universität Duisburg-Essen, Germany**
henning.kerstan@uni-due.de, barbara_koenig@uni-due.de

Abstract

We study behavioral metrics in an abstract coalgebraic setting. Given a coalgebra $\alpha\colon X \to FX$ in **Set**, where the functor F specifies the branching type, we define a framework for deriving pseudometrics on X which measure the behavioral distance of states.

A first crucial step is the lifting of the functor F on **Set** to a functor $\overline{F}$ in the category **PMet** of pseudometric spaces. We present two different approaches which can be viewed as generalizations of the Kantorovich and Wasserstein pseudometrics for probability measures. We show that the pseudometrics provided by the two approaches coincide on several natural examples, but in general they differ.

Then a final coalgebra for F in **Set** can be endowed with a behavioral distance resulting as the smallest solution of a fixed-point equation, yielding the final $\overline{F}$-coalgebra in **PMet**. The same technique, applied to an arbitrary coalgebra $\alpha\colon X \to FX$ in **Set**, provides the behavioral distance on X. Under some constraints we can prove that two states are at distance 0 if and only if they are behaviorally equivalent.

1998 ACM Subject Classification F.3.1 Specifying and Verifying and Reasoning about Programs, D.2.4 Software/Program Verification

Keywords and phrases behavioral metric, functor lifting, pseudometric, coalgebra

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2014.403

1 Introduction

Increasingly, modelling formalisms are equipped with quantitative information, such as probability, time or weight. Such quantitative information should be taken into account when reasoning about behavioral equivalence of system states, such as bisimilarity. In this setting the appropriate notion is not necessarily equivalence, but a behavioral metric that measures the distance of the behavior of two states. In a quantitative setting, it is often unreasonable to assume that two states have exactly the same behavior, but it makes sense to express that their behavior differs by some (small) value ε.

The above considerations led to the study of behavioral metrics which aims at quantifying the the distance between the behavior of states. Since different states can have exactly the same behavior it is quite natural to consider *pseudometrics*, which allow different elements to be at zero distance.

Earlier contributions defined behavioral metrics in the setting of probabilistic systems [9, 23] and of metric transition systems [6]. Our aim is to generalize these ideas and to study behavioral metrics in a general coalgebraic setting. The theory of coalgebra [17] is nowadays

34th Int'l Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2014).
Editors: Venkatesh Raman and S. P. Suresh; pp. 403–415
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

a well-established tool for defining and reasoning about various state based transition systems such as deterministic, nondeterministic, weighted or probabilistic automata. Hence, it is the appropriate setting to ask and answer general questions about behavioral metrics.

- *How can we define behavioral metrics for transition systems with different branching types?* We provide a coalgebraic framework in the category of pseudometric spaces **PMet** that allows to define and reason about such metrics.
- *Are the behavioral metrics canonical in some way?* We provide a natural way to define metrics by lifting functors from **Set** to the category of pseudometric spaces. In fact, we study two liftings: the Kantorovich and the Wasserstein lifting and observe that they coincide in many cases. This provides us with a notion of canonicity and justification for the choice of metrics.
- *Does the measurement of distances affect behavioral equivalence?* If we start by considering coalgebras in **PMet** (as, e.g., in [23]), it is not entirely clear a priori whether the richer categorical structure influences the notion of behavioral equivalence. In our setting we start with coalgebras in **Set** and put distance measurements "on top", showing that, under some mild constraints, the original notion of behavioral equivalence is not compromised, in the sense that two states are behaviorally equivalent iff their distance is 0.
- *Are there generic algorithms to compute metrics?* Coalgebra is a valuable tool to define generic methods that can be instantiated to concrete cases in order to obtain prototype algorithms. In our case we give a (high-level) procedure for computing behavioral distances on a given coalgebra, based on determining the smallest solution of a fixed-point equation.

A central contribution of this paper is the lifting of a functor F from **Set** to **PMet**. Given a pseudometric space (X, d), the goal is to define a suitable pseudometric on FX. Such liftings of metrics have been extensively studied in transportation theory [24], e.g. for the case of the (discrete) probability distribution functor, which comes with a nice analogy: assume several cities (with fixed distances between them) and two probability distributions s, t on cities, representing supply and demand (in units of mass). The distance between s, t can be measured in two ways: the first is to set up an optimal transportation plan with minimal costs (in the following also called coupling) to transport goods from cities with excess supply to cities with excess demand. The cost of transport is determined by the product of mass and distance. In this way we obtain the Wasserstein distance. A different view is to imagine a logistics firm that is commissioned to handle the transport. It sets prices for each city and buys and sells for this price at every location. However, it has to ensure that the price function is nonexpansive, i.e., the difference of prices between two cities is smaller than the distance of the cities, otherwise it will not be worthwhile to outsource this task. This firm will attempt to maximize its profit, which can be considered as the Kantorovich distance of s, t. The Kantorovich-Rubinstein duality informs us that these two views lead to the exactly same result, a very good argument for the canonicity of this notion of distance.

It is our observation that these two notions of distance lifting can analogously be defined for arbitrary functors F, leading to a rich general theory. The lifting has an evaluation function as parameter. As concrete examples, besides the probability distribution functor, we study the (finite) powerset functor (resulting in the Hausdorff metric) and the coproduct and product bifunctors. In the case of the product bifunctor we consider different evaluation functions, each leading to a well-known product metric. The Kantorovich-Rubinstein duality holds for these functors, but it does not hold in general (we provide a counterexample).

After discussing functor liftings, we define coalgebraic behavioral pseudometrics and answer the questions above. Specifically we show how to compute distances on the final coalgebra as well as on arbitrary coalgebras via fixed-point iteration and we prove that the pseudometric obtained on the final coalgebra is indeed a metric. In [3] we discuss a fibrational perspective on our work and we compare with [13]. All proofs for our results are in [3].

2 Preliminaries, Notation and Evaluation Functions

We assume that the reader is familiar with the basic notions of category theory, especially with the definitions of functor, product, coproduct and weak pullbacks.

For a function $f\colon X \to Y$ and sets $A \subseteq X$, $B \subseteq Y$ we write $f[A] := \{f(a) \mid a \in A\}$ for the *image* of A and $f^{-1}[B] = \{a \in A \mid f(x) \in B\}$ for the *preimage* of B. If $Y \subseteq [0,\infty]$ and $f, g\colon X \to Y$ are functions we write $f \le g$ when $\forall x \in X : f(x) \le g(x)$.

Given a natural number $n \in \mathbb{N}$ and a family $(X_i)_{i=1}^n$ of sets X_i we denote the projections of the (cartesian) product of the X_i by $\pi_i^n\colon \prod_{i=1}^n X_i \to X_i$, or just by π_i if n is clear from the context. For a source $(f_i\colon X \to X_i)_{i=1}^n$ we denote the unique mediating arrow to the product by $\langle f_1, \dots, f_n\rangle\colon X \to \prod_{i=1}^n X_i$. Similarly, given a family of arrows $(f_i\colon X_i \to Y_i)_{i=1}^n$, we write $f_1 \times \cdots \times f_n = \langle f_1 \circ \pi_1, \dots, f_n \circ \pi_n\rangle\colon \prod_{i=1}^n X_i \to \prod_{i=1}^n Y_i$.

We quickly recap the basic ideas of coalgebras. Let F be an endofunctor on the category **Set** of sets and functions. An F-coalgebra is just a function $\alpha\colon X \to FX$. Given another F-coalgebra $\beta\colon Y \to FY$ a coalgebra homomorphism from α to β is a function $f\colon A \to B$ such that $\beta \circ f = Ff \circ \alpha$. We call an F-coalgebra $\kappa\colon \Omega \to F\Omega$ *final* if for any other coalgebra $\alpha\colon X \to FX$ there is a unique coalgebra homomorphism $[\![\_]\!]\colon X \to \Omega$. The final coalgebra need not exist but if it does it is unique up to isomorphism. It can be considered as the universe of all possible behaviors. If we have an endofunctor F such that a final coalgebra $\kappa\colon \Omega \to F\Omega$ exists then for any coalgebra $\alpha\colon X \to FX$ two states $x_1, x_2 \in X$ are said to be *behaviorally equivalent* if and only if $[\![x_1]\!] = [\![x_2]\!]$.

We now introduce some preliminaries about (pseudo)metric spaces. Our (pseudo)metrics assume values in a closed interval $[0,\top]$, where $\top \in (0,\infty]$ is a fixed maximal element (for our examples we will use $\top = 1$ or $\top = \infty$). In this way the set of (pseudo)metrics over a fixed set with pointwise order is a complete lattice (since $[0,\top]$ is) and the resulting category of pseudometric spaces is complete and cocomplete.

▶ **Definition 2.1** (Pseudometric, Pseudometric Space)**.** Given a set X, a *pseudometric on* X is a function $d\colon X \times X \to [0,\top]$ such that for all $x, y, z \in X$, the following axioms hold: $d(x,x) = 0$ (*reflexivity*), $d(x,y) = d(y,x)$ (*symmetry*), $d(x,z) \le d(x,y) + d(y,z)$ (*triangle inequality*). If additionally $d(x,y) = 0$ implies $x = y$, d is called a *metric*. A *(pseudo)metric space* is a pair (X,d) where X is a set and d is a (pseudo)metric on X.

By $d_e\colon [0,\top]^2 \to [0,\top]$ we denote the ordinary Euclidean distance on $[0,\top]$, i.e., $d_e(x,y) = |x-y|$ for $x, y \in [0,\top] \setminus \{\infty\}$, and – where appropriate – $d_e(x,\infty) = \infty$ if $x \neq \infty$ and $d_e(\infty,\infty) = 0$. Addition is defined in the usual way, in particular $x + \infty = \infty$ for $x \in [0,\infty]$.

Hereafter, we only consider those functions between pseudometric spaces that do not increase distances.

▶ **Definition 2.2** (Nonexpansive Function, Isometry)**.** Let (X, d_X), (Y, d_Y) be pseudometric spaces. A function $f\colon X \to Y$ is called *nonexpansive* if $d_Y \circ (f \times f) \le d_X$. In this case we write $f\colon (X, d_X) \xrightarrow{1} (Y, d_Y)$. If equality holds, f is called an *isometry*.

For our purposes it will turn out to be useful to consider the following alternative characterization of the triangle inequality using the concept of nonexpansive functions.

▶ **Lemma 2.3.** *A reflexive and symmetric function $d\colon X^2 \to [0,\top]$ satisfies the triangle inequality iff for all $x \in X$ the function $d(x,\_)\colon X \to [0,\top]$ is nonexpansive.*

As stated before, our definition of a pseudometric gives rise to a suitably rich category.

▶ **Definition 2.4** (Category of Pseudometric Spaces)**.** For a fixed $\top \in (0,\infty]$ we denote by **PMet** the category of all pseudometric spaces and nonexpansive functions.

This category is complete and cocomplete (see [3]) and, in particular, it has products and coproducts as we will see in Examples 5.1 and 5.2. We now introduce two motivating examples borrowed from [23] and [6].

▶ **Example 2.5** (Probabilistic Transition Systems and Behavioral Distance)**.** We regard probabilistic transition systems as coalgebras of the form $\alpha\colon X \to \mathcal{D}(X+\mathbf{1})$, where $\mathcal{D}$ is the probability distribution functor (with finite support) which maps a set X to the set $\mathcal{D}X = \{P\colon X \to [0,1] \mid \sum_{x\in X} P(x) = 1, P \text{ has finite support}\}$ and a function $f\colon X \to Y$ to the function $\mathcal{D}f\colon \mathcal{D}X \to \mathcal{D}Y, P \mapsto \lambda y. \sum_{x\in f^{-1}[\{y\}]} P(x)$. Here $\alpha(x)(y)$, for $x,y \in X$, denotes the probability of a transition from a state x to y and $\alpha(x)(\checkmark)$ stands for the probability of terminating from x (we use $\checkmark$ for the single element of the set $\mathbf{1}$).

In [23] a metric for the continuous version of these systems is introduced, by considering a discount factor $c \in (0,1)$. In the discrete case we obtain the behavioral distance $d\colon X^2 \to [0,1]$, defined as the least solution of the equation $d(x,y) = \overline{d}(\alpha(x),\alpha(y))$, where $x,y \in X$ and $\overline{d}\colon (\mathcal{D}(X+\mathbf{1}))^2 \to [0,1]$ is defined in two steps: First, $\hat{d}\colon (X+\mathbf{1})^2 \to [0,1]$ is defined as $\hat{d}(x,y) = c\cdot d(x,y)$ if $x,y \in X$, $\hat{d}(\checkmark,\checkmark) = 0$ and 1 otherwise. Then, for all $P_1,P_2 \in \mathcal{D}(X+\mathbf{1})$, $\overline{d}(P_1,P_2)$ is defined as the supremum of all values $\sum_{x\in X+\mathbf{1}} f(x)\cdot\big(P_1(x) - P_2(x)\big)$, with $f\colon (X+\mathbf{1},\hat{d}) \xrightarrow{1} ([0,1],d_e)$ being an arbitrary nonexpansive function. As we will further discuss in Example 3.3, $\overline{d}$ is the Kantorovich pseudometric given by the space $(X+\mathbf{1},\hat{d})$.

We consider a concrete example from [23], illustrated on the left of Figure 1. The behavioral distance of u and z is $d(u,z) = 1$ and hence $d(x,y) = c\cdot\varepsilon$.

▶ **Example 2.6** (Metric Transition Systems and Propositional Distances)**.** We give another example based on the notions of [6]. A finite set $\Sigma = \{r_1,\dots,r_n\}$ of propositions is given and each proposition $r \in \Sigma$ is associated with a pseudometric space (M_r,d_r). A valuation u is a function with domain Σ that assigns to each $r \in \Sigma$ an element of M_r. We denote the set of all valuations by $\mathcal{U}[\Sigma]$. A metric transition system is a tuple $M = (S,\tau,\Sigma,[\cdot])$ with a set S of states, a transition relation $\tau \subseteq S\times S$, a finite set Σ of propositions and a valuation $[s]$ for each state $s \in S$. We write $\tau(s)$ for $\{s' \in S \mid (s,s') \in \tau\}$ and require that $\tau(s)$ is finite.

In [6] the propositional distance between two valuations is given by $\overline{pd}(u,v) = \max_{r\in\Sigma} d_r(u(r),v(r))$ for $u,v \in \mathcal{U}[\Sigma]$. The (undirected) branching distance $d\colon S\times S \to \mathbb{R}_0^+$ is defined as the smallest fixed-point of the following equation, where $s,t \in S$:

$$d(s,t) = \max\left\{\overline{pd}([s],[t]), \max_{s'\in\tau(s)}\min_{t'\in\tau(t)} d(s',t'), \max_{t'\in\tau(t)}\min_{s'\in\tau(s)} d(s',t')\right\} \tag{1}$$

Note that, apart from the first argument, this coincides with the Hausdorff distance.

We consider an example which appears similarly in [6] (see Figure 1, right) with a single proposition $r \in \Sigma$, where $M_r = [0,1]$ is equipped with the Euclidean distance d_e. According to (1), $d(x_1,y_1)$ equals the Hausdorff distance of the reals associated with the sets of successors, which is 0.3 (since this is the maximal distance of any successor to the closest successor in the other set of successors, here: the distance from y_3 to x_3).

In order to model such transition systems as coalgebras we define the following n-ary auxiliary functor: $G(X_1,\dots,X_n) = \{u\colon \Sigma \to X_1+\dots+X_n \mid u(r_i) \in X_i\}$. Then coalgebras

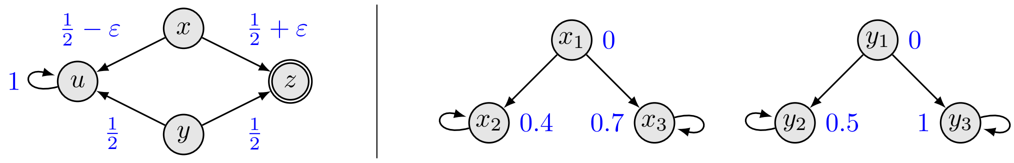

Figure 1 A probabilistic transition system (left) and a metric transition system (right).

are of the form $\alpha\colon S \to G(M_{r_1},\ldots,M_{r_n}) \times \mathcal{P}_{\mathit{fin}}(S)$, where $\mathcal{P}_{\mathit{fin}}$ is the finite powerset functor and $\alpha(s) = ([s], \tau(s))$. As we will see later in Example 6.7, the right-hand side of (1) can be seen as lifting a metric d on X to a metric on $G(M_{r_1},\ldots,M_{r_n}) \times \mathcal{P}_{\mathit{fin}}(X)$.

Generalizing from the examples, we now establish a general framework for deriving such behavioral distances. In both cases, the crucial step is to find, for a functor F, a way to lift a pseudometric on X to a pseudometric on FX. Based on this, one can set up a fixed-point equation and define behavioral distance as its smallest solution. Hence, in the next sections we describe how to lift an endofunctor F on **Set** to an endofunctor on **PMet**.

▶ **Definition 2.7** (Lifting). Let $U\colon \mathbf{PMet} \to \mathbf{Set}$ be the forgetful functor which maps every pseudometric space to its underlying set. A functor $\overline{F}\colon \mathbf{PMet} \to \mathbf{PMet}$ is called a *lifting* of a functor $F\colon \mathbf{Set} \to \mathbf{Set}$ if it satisfies $U \circ \overline{F} = F \circ U$.

It is not difficult to prove that such a lifting is always monotone on pseudometrics over a common set, i.e. for any two pseudometrics $d_1 \le d_2$ on the same set X, we also have $d_1^F \le d_2^F$ where d_i^F are the pseudometrics on FX obtained by applying $\overline{F}$ to (X, d_i) (see [3]). Similarly to predicate lifting of coalgebraic modal logic [18], liftings on **PMet** can be conveniently defined via an evaluation function.

▶ **Definition 2.8** (Evaluation Function and Evaluation Functor). Let F be an endofunctor on **Set**. An *evaluation function* for F is a function $ev_F\colon F[0,\top] \to [0,\top]$. Given such a function, we define the *evaluation functor* to be the endofunctor $\widetilde{F}$ on $\mathbf{Set}/[0,\top]$, the slice category[1] over $[0,\top]$, via $\widetilde{F}(g) = ev_F \circ Fg$ for all $g \in \mathbf{Set}/[0,\top]$. On arrows $\widetilde{F}$ is defined as F.

3 Lifting Functors to Pseudometric Spaces à la Kantorovich

Let us now consider an endofunctor F on **Set** with an evaluation function ev_F. Given a pseudometric space (X,d), our first approach will be to take the smallest possible pseudometric d^F on FX such that, for all nonexpansive functions $f\colon (X,d) \xrightarrow{1} ([0,\top], d_e)$, also $\widetilde{F}f\colon (FX, d^F) \xrightarrow{1} ([0,\top], d_e)$ is nonexpansive again, i.e. we want to ensure that for all $t_1, t_2 \in FX$ we have $d_e(\widetilde{F}f(t_1), \widetilde{F}f(t_2)) \le d^F(t_1,t_2)$. This idea immediately leads us to the next definition.

▶ **Definition 3.1** (Kantorovich Pseudometric and Kantorovich Lifting). Let $F\colon \mathbf{Set} \to \mathbf{Set}$ be a functor with an evaluation function ev_F. For every pseudometric space (X,d) the *Kantorovich pseudometric* on FX is the function $d^{\uparrow F}\colon FX \times FX \to [0,\top]$, where for all $t_1, t_2 \in FX$:

$$d^{\uparrow F}(t_1,t_2) := \sup\{d_e(\widetilde{F}f(t_1), \widetilde{F}f(t_2)) \mid f\colon (X,d) \xrightarrow{1} ([0,\top], d_e)\}\,.$$

[1] The slice category $\mathbf{Set}/[0,\top]$ has as objects all functions $g\colon X \to [0,\top]$ where X is an arbitrary set. Given g as before and $h\colon Y \to [0,\top]$, an arrow from g to h is a function $f\colon X \to Y$ satisfying $h \circ f = g$.

The *Kantorovich lifting* of the functor F is the functor $\overline{F}\colon \mathbf{PMet} \to \mathbf{PMet}$ defined as $\overline{F}(X,d) = (FX, d^{\uparrow F})$ and $\overline{F}f = Ff$.

It is easy to show that $d^{\uparrow F}$ is indeed a pseudometric. Since $\overline{F}$ inherits the preservation of identities and composition of morphisms from F we can prove that nonexpansive functions are mapped to nonexpansive functions and isometries to isometries.

▶ **Proposition 3.2.** *The Kantorovich lifting $\overline{F}$ of a functor F preserves isometries.*

We chose the name *Kantorovich* because our definition is reminiscent of the Kantorovich pseudometric in probability theory. If we take the proper combination of functor and evaluation function, we can recover that pseudometric (in the discrete case) as the first instance for our framework.

▶ **Example 3.3** (Probability Distribution Functor)**.** We take $\top = 1$, the probability distribution functor $\mathcal{D}$ from Example 2.5 and define $ev_{\mathcal{D}}\colon \mathcal{D}[0,1] \to [0,1]$, $ev_{\mathcal{D}}(P) = \mathbb{E}_P[\mathrm{id}_{[0,1]}] = \sum_{x\in[0,1]} x\cdot P(x)$ yielding $\widetilde{\mathcal{D}}g(P) = \mathbb{E}_P[g] = \sum_{x\in[0,1]} g(x)\cdot P(x)$ for all $g\colon X \to [0,1]$. For every pseudometric space (X,d) we obtain the Kantorovich pseudometric $d^{\uparrow\mathcal{D}}\colon (\mathcal{D}X)^2 \to [0,1]$, $d^{\uparrow\mathcal{D}}(P_1,P_2) = \sup\{\sum_{x\in X} f(x)\cdot \big(P_1(x) - P_2(x)\big) \mid f\colon (X,d) \xrightarrow{1} ([0,1], d_e)\}$.

In general Kantorovich liftings do not preserve metrics, as shown by the following example.

▶ **Example 3.4.** Let $F\colon \mathbf{Set} \to \mathbf{Set}$ be given as $FX = X \times X$ on sets and $Ff = f \times f$ on functions and take $\top = \infty$, $ev_F\colon F[0,\infty] \to [0,\infty]$, $ev_F(r_1,r_2) = r_1 + r_2$. For a metric space (X,d) with $|X| \geq 2$ let $t_1 = (x_1,x_2) \in FX$ with $x_1 \neq x_2$ and define $t_2 := (x_2,x_1)$. Clearly $t_1 \neq t_2$ but for every nonexpansive function $f\colon (X,d) \xrightarrow{1} ([0,\top], d_e)$ we have $\widetilde{F}f(t_1) = f(x_1) + f(x_2) = f(x_2) + f(x_1) = \widetilde{F}f(t_2)$ and thus $d^{\uparrow F}(t_1,t_2) = 0$.

4 Wasserstein Pseudometric and Kantorovich-Rubinstein Duality

We have seen that our first lifting approach bears close resemblance to the original Kantorovich pseudometric on probability measures. In that context there exists another pseudometric, the Wasserstein pseudometric, which under certain conditions coincides with the Kantorovich pseudometric. We will define a generalized version of the Wasserstein pseudometric and compare it with our generalized Kantorovich pseudometric. To do that we first need to define how we can couple elements of FX.

▶ **Definition 4.1** (Coupling)**.** Let $F\colon \mathbf{Set} \to \mathbf{Set}$ be a functor and $n \in \mathbb{N}$. Given a set X and $t_i \in FX$ for $1 \leq i \leq n$ we call an element $t \in F(X^n)$ such that $F\pi_i(t) = t_i$ a *coupling* of the t_i (with respect to F). We write $\Gamma_F(t_1, t_2, \ldots, t_n)$ for the set of all these couplings.

If F preserves weak pullbacks, we can define new couplings based on given ones.

▶ **Lemma 4.2** (Gluing Lemma)**.** *Let $F\colon \mathbf{Set} \to \mathbf{Set}$ be a weak pullback preserving functor, X a set, $t_1, t_2, t_3 \in FX$, $t_{12} \in \Gamma_F(t_1,t_2)$, and $t_{23} \in \Gamma_F(t_2,t_3)$ be couplings. Then there is a coupling $t_{123} \in \Gamma_F(t_1,t_2,t_3)$ such that $F(\langle\pi_1^3, \pi_2^3\rangle)(t_{123}) = t_{12}$ and $F(\langle\pi_2^3,\pi_3^3\rangle)(t_{123}) = t_{23}$.*

This lemma already hints at the fact that our new lifting will only work for weak pullback preserving functors, which is a standard requirement in coalgebra. In addition to that we have to impose three extra conditions on the evaluation functions.

▶ **Definition 4.3** (Well-Behaved Evaluation Function)**.** Let ev_F be an evaluation function for a functor $F\colon \mathbf{Set} \to \mathbf{Set}$. We call ev_F *well-behaved* if it satisfies the following conditions:

1. $\widetilde{F}$ is monotone, i.e., for $f, g\colon X \to [0, \top]$ with $f \le g$, we have $\widetilde{F}f \le \widetilde{F}g$.
2. For each $t \in F([0,\top]^2)$ it holds that $d_e(ev_F(t_1), ev_F(t_2)) \le \widetilde{F}d_e(t)$ for $t_i := F\pi_i(t)$.
3. $ev_F^{-1}[\{0\}] = Fi[F\{0\}]$ where $i\colon \{0\} \hookrightarrow [0,\top]$ is the inclusion map.

While the first condition of this definition is quite natural, the other two need to be explained. Condition 2 is needed to ensure that $\widetilde{F}\mathrm{id}_{[0,\top]} = ev_F\colon F[0,\top] \to [0,\top]$ is nonexpansive once d_e is lifted to $F[0,\top]$ (cf. the intuition behind the Kantorovich lifting, where we ensure that $\widetilde{F}f$ is nonexpansive whenever f is nonexpansive). Furthermore Condition 3 intuitively says that exactly the elements of $F\{0\}$ are mapped to 0 via ev_F. Before we define the Wasserstein pseudometric and the corresponding lifting, we take a look at an example of a functor together with a well-behaved evaluation function.

▶ **Example 4.4** (Finite Powerset Functor)**.** Let $\top = \infty$. We take the finite powerset functor $\mathcal{P}_{\mathit{fin}}$ with evaluation function $\max\colon \mathcal{P}_{\mathit{fin}}([0,\infty]) \to [0,\infty]$ with $\max \emptyset = 0$. This evaluation function is well-behaved whereas $\min\colon \mathcal{P}_{\mathit{fin}}([0,\infty]) \to [0,\infty]$ is not well-behaved.

▶ **Definition 4.5** (Wasserstein Pseudometric and Wasserstein Lifting)**.** Let $F\colon \mathbf{Set} \to \mathbf{Set}$ be a weak pullback preserving functor with well-behaved evaluation function ev_F. For every pseudometric space (X, d) the *Wasserstein pseudometric* on FX is the function $d^{\downarrow F}\colon FX \times FX \to [0,\top]$ given by, for all $t_1, t_2 \in FX$,

$$d^{\downarrow F}(t_1, t_2) := \inf\{\widetilde{F}d(t) \mid t \in \Gamma_F(t_1, t_2)\}\,.$$

We define the *Wasserstein lifting* of F to be the functor $\overline{F}\colon \mathbf{PMet} \to \mathbf{PMet}$, $\overline{F}(X, d) = (FX, d^{\downarrow F})$, $\overline{F}f = Ff$.

This time it is not straightforward to prove that $d^{\downarrow F}$ is a pseudometric, so we explicitly provide the following result. Its proof relies on all properties of well-behavedness of ev_F and uses Lemma 4.2 which explains why we need a weak pullback preserving functor.

▶ **Proposition 4.6.** *The Wasserstein pseudometric is a well-defined pseudometric on FX.*

It is not hard to show functoriality of $\overline{F}$ and, as before, the lifted functor preserves isometries.

▶ **Proposition 4.7.** *The Wasserstein lifting $\overline{F}$ of a functor F preserves isometries.*

In contrast to our previous approach, metrics are preserved in certain situations.

▶ **Proposition 4.8** (Preservation of Metrics)**.** *Let (X, d) be a metric space and F be a functor. If the infimum in Definition 4.5 is a minimum for all $t_1, t_2 \in FX$ where $d^{\downarrow F}(t_1, t_2) = 0$ then $d^{\downarrow F}$ is a metric, thus also $\overline{F}(X, d) = (FX, d^{\downarrow F})$ is a metric space.*

Please note that a similar restriction for the Kantorovich lifting (i.e. requiring that the supremum in Definition 3.1 is a maximum) does not yield preservation of metrics: In Example 3.4 the supremum is always a maximum but we do not get a metric.

Let us now compare both lifting approaches. Whenever it is defined, the Wasserstein pseudometric is an upper bound for the Kantorovich pseudometric.

▶ **Proposition 4.9.** *Let F be a weak pullback preserving functor with well-behaved evaluation function. Then for all pseudometric spaces (X, d) it holds that $d^{\uparrow F} \le d^{\downarrow F}$.*

In general this inequality may be strict in general, as the following example shows.

▶ **Example 4.10.** The functor of Example 3.4 preserves weak pullbacks and the evaluation function is well-behaved. We continue the example and take $t_1 = (x_1, x_2)$, $t_2 = (x_2, x_1)$. The unique coupling $t \in \Gamma_F(t_1, t_2)$ is $t = (x_1, x_2, x_2, x_1)$. Using that d is a metric we conclude that $d^{\downarrow F}(t_1, t_2) = \widetilde{F}d(t) = d(x_1, x_2) + d(x_2, x_1) = 2d(x_1, x_2) > 0 = d^{\uparrow F}(t_1, t_2)$.

When the inequality can be replaced by an equality we will in the following say that the Kantorovich-Rubinstein duality holds. In this case we obtain a canonical notion of distance on FX, given a pseudometric space (X, d). To calculate the distance of $t_1, t_2 \in FX$ it is then enough to find a nonexpansive function $f\colon (X, d) \xrightarrow{1} ([0, \top], d_e)$ and a coupling $t \in \Gamma_F(t_1, t_2)$ such that $d_e(\widetilde{F}f(t_1), \widetilde{F}f(t_2)) = \widetilde{F}d_e(t)$. Then, due to Proposition 4.9, this value equals $d^{\uparrow F}(t_1, t_2) = d^{\downarrow F}(t_1, t_2)$. We will now take a look at some examples where the duality holds.

▶ **Example 4.11** (Identity Functor)**.** Take $F = \mathrm{Id}$ with the identity evaluation map $ev_{\mathrm{Id}} = \mathrm{id}_{[0,\top]}$. For any $t_1, t_2 \in X$, $t := (t_1, t_2)$ is the unique coupling of t_1, t_2. Hence, $d^{\downarrow F}(t_1, t_2) = d(t_1, t_2)$. With the function $d(t_1, \_)\colon (X, d) \xrightarrow{1} ([0, \top], d_e)$ we obtain duality because we have $d(t_1, t_2) = d_e(d(t_1, t_1), d(t_1, t_2)) \le d^{\uparrow F}(t_1, t_2) \le d^{\downarrow F}(t_1, t_2) = d(t_1, t_2)$ and thus equality. Similarly, if we define $ev_{\mathrm{Id}}(r) = c \cdot r$ for $r \in [0, \top]$, $0 < c \le 1$, the Kantorovich and Wasserstein liftings coincide and we obtain the discounted distance $d^{\uparrow F}(t_1, t_2) = d^{\downarrow F}(t_1, t_2) = c \cdot d(t_1, t_2)$.

▶ **Example 4.12** (Probability Distribution Functor)**.** The functor $\mathcal{D}$ of Example 3.3 preserves weak pullbacks [19] and the evaluation function $ev_{\mathcal{D}}$ is well-behaved. We recover the usual Wasserstein pseudometric $d^{\downarrow \mathcal{D}}(P_1, P_2) = \inf\{\sum_{x_1, x_2 \in X} d(x_1, x_2) \cdot P(x_1, x_2) \mid P \in \Gamma_{\mathcal{D}}(P_1, P_2)\}$ and the Kantorovich-Rubinstein duality [24] from transportation theory for the discrete case.

▶ **Example 4.13** (Finite Powerset Functor and Hausdorff Pseudometric)**.** Let $\top = \infty$, $F = \mathcal{P}_{\mathit{fin}}$ with evaluation map $ev_{\mathcal{P}_{\mathit{fin}}}\colon \mathcal{P}_{\mathit{fin}}([0, \infty]) \to [0, \infty]$, $ev_{\mathcal{P}_{\mathit{fin}}}(R) = \max R$ with $\max \emptyset = 0$ (as in Example 4.4). In this setting we obtain duality and both pseudometrics are equal to the Hausdorff pseudometric d_H on $\mathcal{P}_{\mathit{fin}}(X)$ which is defined as, for all $X_1, X_2 \in \mathcal{P}_{\mathit{fin}}(X)$,

$$d_H(X_1, X_2) = \max\left\{\max_{x_1 \in X_1} \min_{x_2 \in X_2} d(x_1, x_2), \max_{x_2 \in X_2} \min_{x_1 \in X_1} d(x_1, x_2)\right\}.$$

Note that the distance is ∞, if either X_1 or X_2 is empty.

It is also illustrative to consider the countable powerset functor. Using the supremum as evaluation function, one obtains again the Hausdorff pseudometric (with supremum/infimum replacing maximum/minimum). However, in this case the Hausdorff distance of different countable sets might be 0, even if we lift a metric. This shows that in general the Wasserstein lifting does not preserve metrics but we need an extra condition, e.g. the one in Proposition 4.8.

5 Lifting Multifunctors

Our two approaches can easily be generalized[2] to lift a multifunctor $F\colon \mathbf{Set}^n \to \mathbf{Set}$ (for $n \in \mathbb{N}$) in a similar sense as given by Definition 2.7 to a multifunctor $\overline{F}\colon \mathbf{PMet}^n \to \mathbf{Set}$. The only difference is that we start with n pseudometric spaces instead of one. Now we need an *evaluation function* $ev_F\colon F([0, \top], \dots, [0, \top]) \to [0, \top]$ which we call *well-behaved* if it satisfies conditions similar to Definition 4.3 and which gives rise to an evaluation multifunctor $\widetilde{F}\colon (\mathbf{Set}/[0, \top])^n \to \mathbf{Set}/[0, \top]$. Given $t_1, t_2 \in F(X_1, \dots, X_n)$ we

[2] The details are spelled out in [3], here we provide just the basic ideas.

write again $\Gamma_F(t_1,t_2) \subseteq F(X_1^2,\ldots,X_n^2)$ for the set of couplings which is defined analogously to Definition 4.1. For pseudometrics $d_i\colon X_i^2 \to [0,\top]$, we can then define the *Kantorovich/Wasserstein pseudometric* $d^{\uparrow F}_{1,\ldots,n}, d^{\downarrow F}_{1,\ldots,n}\colon F(X_1,\ldots,X_n) \times F(X_1,\ldots,X_n) \to [0,\top]$, as $d^{\uparrow F}_{1,\ldots,n}(t_1,t_2) := \sup\{d_e(\widetilde{F}(f_1,\ldots,f_n)(t_1), \widetilde{F}(f_1,\ldots,f_n)(t_2)) \mid f_i\colon (X_i,d_i) \xrightarrow{1} ([0,\top],d_e)\}$ and $d^{\downarrow F}_{1,\ldots,n}(t_1,t_2) := \inf\{\widetilde{F}(d_1,\ldots,d_n)(t) \mid t \in \Gamma_F(t_1,t_2)\}$. This setting grants us access to new examples such as the product and the coproduct bifunctors.

▶ **Example 5.1** (Product Bifunctor). For the product bifunctor $F\colon \mathbf{Set}^2 \to \mathbf{Set}$ where $F(X_1,X_2) = X_1 \times X_2$ and $F(f_1,f_2) = f_1 \times f_2$ we consider the evaluation function $\max\colon [0,\top]^2 \to [0,\top]$ and for fixed parameters $c_1,c_2 \in (0,1]$ and $p \in \mathbb{N}$ the function $\rho\colon [0,\top]^2 \to [0,\top]$, $\rho(x_1,x_2) = (c_1x_1^p + c_2x_2^p)^{1/p}$. These functions are well-behaved, the Kantorovich-Rubinstein duality holds and the supremum [infimum] of the Kantorovich [Wasserstein] pseudometrics is always a maximum [minimum]. For the first function we obtain the ∞-product pseudometric $d_\infty((x_1,x_2),(y_1,y_2)) = \max(d_1(x_1,y_1), d_2(x_2,y_2))$ and for the other function the weighted p-product pseudometric $d_p((x_1,x_2),(y_1,y_2)) = (c_1d_1^p(x_1,y_1) + c_2d_2^p(x_2,y_2))^{1/p}$.

Note that the pseudometric space $(X_1 \times X_2, d_\infty)$ is the usual binary (category theoretic) product of (X_1,d_1) and (X_2,d_2). Similarly, we can also obtain the binary coproduct.

▶ **Example 5.2** (Coproduct Bifunctor). For the coproduct bifunctor $F\colon \mathbf{Set}^2 \to \mathbf{Set}$, where $F(X_1,X_2) = X_1 + X_2 = X_1 \times \{1\} \cup X_2 \times \{2\}$ and $F(f_1,f_2) = f_1 + f_2$ we take the evaluation function $ev_F\colon [0,\top] + [0,\top] \to [0,\top]$, $ev_F(x,i) = x$. This function is well-behaved, the Kantorovich-Rubinstein duality holds and the supremum of the Kantorovich pseudometric is always a maximum whereas the infimum of the Wasserstein pseudometric is a minimum if and only if any coupling of the two elements exists. We obtain the coproduct pseudometric d_+ where $d_+((x_1,i_1),(x_2,i_2))$ is equal to $d_i(x_1,x_2)$ if $i_1 = i_2 = i$ and equal to $\top$ otherwise.

6 Final Coalgebra and Coalgebraic Behavioral Pseudometrics

In this section we assume an arbitrary lifting $\overline{F}\colon \mathbf{PMet} \to \mathbf{PMet}$ of an endofunctor F on $\mathbf{Set}$. For any pseudometric space (X,d) we write d^F for the pseudometric obtained by applying $\overline{F}$ to (X,d). Such a lifting can be obtained as described earlier, but also by taking a lifted multifunctor and fixing all parameters apart from one, or by the composition of such functors. The following result ensures that if $\kappa\colon \Omega \to F\Omega$ is a final F-coalgebra, then there is also a final $\overline{F}$-coalgebra which is constructed by simply enriching Ω with a pseudometric d_Ω.

▶ **Theorem 6.1.** *Let $\overline{F}\colon \mathbf{PMet} \to \mathbf{PMet}$ be a lifting of a functor $F\colon \mathbf{Set} \to \mathbf{Set}$ which has a final coalgebra $\kappa\colon \Omega \to F\Omega$. For every ordinal i we construct a pseudometric $d_i\colon \Omega\times\Omega \to [0,\top]$ as follows: $d_0 := 0$ is the zero pseudometric, $d_{i+1} := d_i^F \circ (\kappa \times \kappa)$ for all ordinals i and $d_j = \sup_{i<j} d_i$ for all limit ordinals j. This sequence converges for some ordinal θ, i.e $d_\theta = d_\theta^F \circ (\kappa \times \kappa)$. Moreover $\kappa\colon (\Omega,d_\theta) \xrightarrow{1} (F\Omega, d_\theta^F)$ is the final $\overline{F}$-coalgebra.*

We noted that for any set X, the set of pseudometrics over X, with pointwise order, is a complete lattice. Moreover the lifting $\overline{F}$ induces a monotone function $\_^F$ which maps any pseudometric d on X to d^F on FX. If, additionally, such function is ω-continuous, i.e., it preserves the supremum of ω-chains, the construction in Theorem 6.1 will converge in at most ω steps, i.e., $d_\theta = d_\omega$. We show in [3] that the liftings induced by the finite powerset functor and the probability distribution functor with finite support are ω-continuous. The arguments used for convergence here suggests a connection with the work in [20], which provides fixed-point results for metric functors which are not locally contractive.

Beyond equivalences of states, in **PMet** we can measure the distance of behaviors in the final coalgebra. More precisely, the *behavioral distance* of two states $x, y \in X$ of some coalgebra $\alpha\colon X \to FX$ is defined via the pseudometric $bd(x,y) = d_\theta(\llbracket x \rrbracket, \llbracket y \rrbracket)$. Such distances can be computed analogously to d_θ above, replacing $\kappa\colon \Omega \to F\Omega$ by α. This way we do not need to explore the entire final coalgebra (which might be too large) but can restrict to the interesting part.

▶ **Theorem 6.2.** *Let the chain of the d_i converge in θ steps and $\overline{F}$ preserve isometries. Let furthermore $\alpha\colon X \to FX$ be an arbitrary coalgebra. For all ordinals i we define a pseudometric $e_i\colon X \times X \to [0,\top]$ as follows: e_0 is the zero pseudometric, $e_{i+1} = e_i^F \circ (\alpha \times \alpha)$ for all ordinals i and $e_j = \sup_{i<j} e_i$ for all limit ordinals j. Then we reach a fixed point after $\zeta \le \theta$ steps, i.e. $e_\zeta = e_\zeta^F \circ (\alpha \times \alpha)$, such that $bd = e_\zeta$.*

Since d_θ is a pseudometric, we have that if $\llbracket x \rrbracket = \llbracket y \rrbracket$ then $bd(x,y) = 0$. The other direction does not hold in general: for this d_θ has to be a proper metric. Theorem 6.5 at the end of this section provides sufficient conditions guaranteeing this property.

To this aim, we proceed by recalling the final coalgebra construction via the final chain which was first presented in the dual setting (free/initial algebra).

▶ **Definition 6.3** (Final Coalgebra Construction [1])**.** Let **C** be a category with terminal object **1** and limits of ordinal-indexed cochains. For any functor $F\colon \mathbf{C} \to \mathbf{C}$ the *final chain* consists of objects W_i for all ordinals i and *connection morphisms* $p_{i,j}\colon W_j \to W_i$ for all ordinals $i \le j$. The objects are defined as $W_0 := \mathbf{1}$, $W_{i+1} := FW_i$ for all ordinals i, and $W_j := \lim_{i<j} W_i$ for all limit ordinals j. The morphisms are determined by $p_{0,i} := {!}\colon W_i \to \mathbf{1}$, $p_{i,i} = \mathrm{id}_{W_i}$ for all ordinals i, $p_{i+1,j+1} := Fp_{i,j}$ for all ordinals $i < j$ and if j is a limit ordinal the $p_{i,j}$ are the morphisms of the limit cone. They satisfy $p_{i,k} = p_{i,j} \circ p_{j,k}$ for all ordinals $i \le j \le k$. We say that the chain *converges* in λ steps if $p_{\lambda,\lambda+1}\colon W_{\lambda+1} \to W_\lambda$ is an iso.

This construction does not necessarily converge, but if it does, we get a final coalgebra.

▶ **Proposition 6.4** ([1])**.** *Let **C** be a category with terminal object **1** and limits of ordinal-indexed cochains. If the final chain of a functor $F\colon \mathbf{C} \to \mathbf{C}$ converges in λ steps then $p_{\lambda,\lambda+1}^{-1}\colon W_\lambda \to FW_\lambda$ is the final coalgebra.*

We now show under which circumstances d_θ is a metric and how our construction relates to the construction of the final chain.

▶ **Theorem 6.5.** *Let $\overline{F}\colon \mathbf{PMet} \to \mathbf{PMet}$ be a lifting of a functor $F\colon \mathbf{Set} \to \mathbf{Set}$ which has a final coalgebra $\kappa\colon \Omega \to F\Omega$. Assume that $\overline{F}$ preserves isometries and metrics, that the final chain for F converges and the chain of the d_i converges in θ steps. Then d_θ is a metric, i.e. for $x, y \in \Omega$ we have $d_\theta(x,y) = 0 \iff x = y$.*

We will now get back to the examples studied at the beginning of the paper (Example 2.5 and Example 2.6) and discuss in which sense they are instances of our framework.

▶ **Example 6.6** (Probabilistic Transition System, revisited)**.** To model the behavioral distance from Example 2.5 in our framework, we set $\top = 1$ and proceed to lift the following three functors: we first consider the identity functor Id with evaluation map $ev_{\mathrm{Id}}\colon [0,1] \to [0,1]$, $ev_{\mathrm{Id}}(z) = c \cdot z$ in order to integrate the discount (Example 4.11). Then, we take the coproduct with the singleton metric space (Example 5.2). The combination of the two functors yields the discrete version of the refusal functor of [23], namely $\overline{R}(X,d) = (X + \mathbf{1}, \hat{d})$ where $\hat{d}$ is taken from Example 2.5. Finally, we lift the probability distribution functor $\mathcal{D}$ to obtain $\overline{\mathcal{D}}$ (Example 3.3). All functors satisfy the Kantorovich-Rubinstein duality and preserve metrics.

It is readily seen that $\overline{\mathcal{D}}(\overline{R}(X,d)) = (\mathcal{D}(X+1), \overline{d})$, where $\overline{d}$ is defined as in Example 2.5). Then, the least solution of $d(x,y) = \overline{d}(\alpha(x), \alpha(y))$ can be computed as in Theorem 6.2.

▶ **Example 6.7** (Metric Transition Systems, revisited)**.** To obtain propositional distances in metric transition systems we set $\top = \infty$. We also define, for the auxiliary functor G, an evaluation function $ev_G \colon G([0,\infty], \ldots, [0,\infty]) \to [0,\infty]$ with $ev_G(u) = \max_{r \in \Sigma} u(r)$. Let $\overline{G}$ be the corresponding lifted functor. It can be shown, similarly to Example 5.1, that the Kantorovich-Rubinstein duality holds and metrics are preserved. We instantiate the given pseudometric spaces (M_{r_i}, d_{r_i}) as parameters and obtain the functor $\overline{F}(X,d) = \overline{G}((M_{r_1}, d_{r_1}), \ldots, (M_{r_n}, d_{r_n})) \times \overline{\mathcal{P}_{fin}}(X,d)$ (for the lifting of the powerset functor see Example 4.13). Then, via Theorem 6.2, we obtain exactly the least solution of (1) in Example 2.6.

7 Related and Future Work

The ideas for our framework are heavily influenced by work on quantitative variants of (bisimulation) equivalence of probabilistic systems. In that context at first Giacalone et al. [12] observed that probabilistic bisimulation [16] is too strong and therefore introduced a metric based on the notion of ε-bisimulations.

Using a logical characterization of bisimulation for labelled Markov processes (LMP) [8], Desharnais et al. defined a family of metrics between these LMPs [9] via functional expressions: if evaluated on a state of an LMP, such a functional expression measures the extent to which a formula is satisfied in that state. A different, coalgebraic approach, which inspired ours, is used by van Breugel et al. [23]. As presented in more detail in the examples above, they define a pseudometric on probabilistic systems via the Kantorovich pseudometric for probability measures. Moreover, they show in [22] that this metric is related to the logical pseudometric by Desharnais et al.

Our framework provides a toolbox to determine behavioral distances for different types of transition systems modeled as coalgebras. Moreover, the liftings introduced in this paper pave the way to extend several coalgebraic methods to reason about quantitative properties of systems. For instance the bisimulation proof principle, which allows to check behavioral equivalence, assumes a specific meaning in **PMet**: every coalgebra $\alpha \colon (X,d) \to \overline{F}(X,d)$ *coinductively proves* that the behavioral distance bd of the underlying F-coalgebra on **Set** is smaller or equal than d. Indeed, since $[\![\_]\!]$ is nonexpansive, $d \geq d_\theta([\![\_]\!], [\![\_]\!]) = bd$. This principle, which has already been stated in different formulations (see e.g. [7, 10, 21]), can now be enhanced via *up-to techniques* by exploiting the liftings introduced in this paper and the coalgebraic understanding of such enhancements given in [4].

Since up-to techniques can exponentially improve algorithms for equivalence-checking, we hope that they could also optimize some of the algorithms for computing (or approximating) behavioral distances [23, 21, 5, 2]. At this point, it is worth recalling that the Kantorovich-Rubinstein duality has been exploited in [23] for defining one of these algorithms: the characterization given by the Wasserstein metric allows to reduce to linear programming.

Another line of research potentially stemming from our work concerns the so-called *abstract GSOS* [15] which provides abstract coalgebraic conditions ensuring compositionality of behavioral equivalence (with respect to some operators). By taking our lifting to **PMet**, abstract GSOS guarantees the nonexpansiveness of behavioral distance, a property that has captured the interest of several researchers [9, 11]. The main technical challenge would be to lift to **PMet** not only functors, but also distributive laws. Lifting of distributive laws would also be needed for defining *linear behavioral distances*, exploiting the coalgebraic account of trace semantics based on Kleisli categories [14].

We finally observe that the chains of Theorems 6.1 and 6.2 can be understood in terms of fibrations along the lines of [13]. A detailed comparison with [13] can be found in [3].

Acknowledgements. The authors are grateful to Franck van Breugel, Neil Ghani and Daniela Petrişan for several precious suggestions and inspiring discussions. The second author acknowledges the support by project `ANR 12IS0 2001 PACE`.

References

1 Jiří Adámek. Free algebras and automata realizations in the language of categories. *Commentationes Mathematicae Universitatis Carolinae*, 015(4):589–602, 1974.

2 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. On-the-fly exact computation of bisimilarity distances. In *Proc. of TACAS'13*, pages 1–15, 2013.

3 Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Behavioral metrics via functor lifting. Extended version with proofs, see `arXiv:1410.3385`, 2014.

4 Filippo Bonchi, Daniela Petrisan, Damien Pous, and Jurriaan Rot. Coinduction up to in a fibrational setting. In *Proc. of CSL-LICS'14*, 2014.

5 Di Chen, Franck van Breugel, and James Worrell. On the complexity of computing probabilistic bisimilarity. In *Proc. of FoSSaCS'12*, volume 7213 of *LNCS*, pages 437–451, 2012.

6 Luca de Alfaro, Marco Faella, and Mariëlle Stoelinga. Linear and branching system metrics. *IEEE Transactions on Software Engineering*, 25(2), 2009.

7 Yuxin Deng, Tom Chothia, Catuscia Palamidessi, and Jun Pang. Metrics for action-labelled quantitative transition systems. *ENTCS*, 153(2):79–96, 2006.

8 Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. *Information and Computation*, 179(2):163–193, 2002.

9 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *TCS*, 318(3):323–354, 2004.

10 Joseé Desharnais, Radha Jagadeesan, Vineet Gupta, and Prakash Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proc. of LICS'02*, pages 413–422. IEEE, 2002.

11 Daniel Gebler and Simone Tini. Compositionality of approximate bisimulation for probabilistic systems. In *Proc. of EXPRESS/SOS'13*, pages 32–46, 2013.

12 Alessandro Giacalone, Chi-Chang Jou, and Scott A. Smolka. Algebraic reasoning for probabilistic concurrent systems. In *Proc. IFIP TC2 Working Conference on Programming Concepts and Methods*, pages 443–458. North-Holland, 1990.

13 Ichiro Hasuo, Kenta Cho, Toshiki Kataoka, and Bart Jacobs. Coinductive predicates and final sequences in a fibration. *ENTCS*, 298:197–214, 2013.

14 Ichiro Hasuo, Bart Jacobs, and Ana Sokolova. Generic trace semantics via coinduction. *LMCS*, 3 (4:11):1–36, November 2007.

15 Bartek Klin. Bialgebras for structural operational semantics: An introduction. *TCS*, 412(38):5043–5069, 2011.

16 Kim G. Larsen and Arne Skou. Bisimulation through probabilistic testing. In *Proc. of POPL'89*, pages 344–352, 1989.

17 J. J. M. M. Rutten. Universal coalgebra: a theory of systems. *TCS*, 249:3–80, 2000.

18 Lutz Schröder. Expressivity of coalgebraic modal logic: The limits and beyond. *TCS*, 390(2–3):230–247, 2008.

19 Ana Sokolova. Probabilistic systems coalgebraically: A survey. *TCS*, 412(38):5095–5110, 2011. CMCS Tenth Anniversary Meeting.

20 Franck van Breugel, Claudio Hermida, Michael Makkai, and James Worrell. Recursively defined metric spaces without contraction. *TCS*, 380(1–2):143–163, July 2007.

21 Franck van Breugel, Babita Sharma, and James Worrell. Approximating a behavioural pseudometric without discount for probabilistic systems. *LMCS*, 4(2), 2008.

22 Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *TCS*, 331:115–142, 2005.

23 Franck van Breugel and James Worrell. Approximating and computing behavioural distances in probabilistic transition systems. *TCS*, 360(1):373–385, 2006.

24 Cédric Villani. *Optimal Transport – Old and New*, volume 338 of *A Series of Comprehensive Studies in Mathematics*. Springer, 2009.