Behavioural Metrics: Compositionality of the Kantorovich Lifting and an Application to Up-To Techniques

Keri D'Angelo
Cornell University, USA

Sebastian Gurke
Universität Duisburg-Essen, Germany

Johanna Maria Kirss
University of Copenhagen, Denmark

Barbara König
Universität Duisburg-Essen, Germany

Matina Najafi
Amirkabir University of Technology, Iran

Wojciech Różowski
University College London, UK

Paul Wild
FAU Erlangen-Nürnberg, Germany

Abstract

Behavioural distances of transition systems modelled via coalgebras for endofunctors generalize traditional notions of behavioural equivalence to a quantitative setting, in which states are equipped with a measure of how (dis)similar they are. Endowing transition systems with such distances essentially relies on the ability to lift functors describing the one-step behavior of the transition systems to the category of pseudometric spaces. We consider the category theoretic generalization of the Kantorovich lifting from transportation theory to the case of lifting functors to quantale-valued relations, which subsumes equivalences, preorders and (directed) metrics. We use tools from fibred category theory, which allow one to see the Kantorovich lifting as arising from an appropriate fibred adjunction. Our main contributions are compositionality results for the Kantorovich lifting, where we show that that the lifting of a composed functor coincides with the composition of the liftings. In addition, we describe how to lift distributive laws in the case where one of the two functors is polynomial (with finite coproducts). These results are essential ingredients for adapting up-to-techniques to the case of quantale-valued behavioural distances. Up-to techniques are a well-known coinductive technique for efficiently showing lower bounds for behavioural distances. We illustrate the results of our paper in two case studies.

2012 ACM Subject Classification Theory of computation → Models of computation

Keywords and phrases behavioural metrics, coalgebra, Galois connections, quantales, Kantorovich lifting, up-to techniques

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2024.20

Related Version Proofs and extra material can be found in the full version.
Full Version: `https://arxiv.org/abs/2404.19632` [10]

Funding *Sebastian Gurke, Barbara König and Paul Wild:* were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 434050016 (SpeQt).
Johanna Maria Kirss: was supported by NSF grant DMS-2216871.
Wojciech Różowski: partially supported by ERC grant Autoprobe (no. 101002697).

Acknowledgements We would like to thank Avi Craimer for helpful discussions and the organizers of the ACT Adjoint School for enabling us to meet and write this paper together!

35th International Conference on Concurrency Theory (CONCUR 2024).
Editors: Rupak Majumdar and Alexandra Silva; Article No. 20; pp. 20:1–20:19
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

In concurrency theory, behavioural equivalences are a fundamental concept: they explain whether two states exhibit the same behaviour and there are efficient techniques for checking behavioural equivalence [25]. More recently, this idea has been generalized to behavioural metrics that measure the behavioural distance of two states [28, 11, 12]. This is in particular interesting for quantitative systems where the behaviour of two states might be similar but not identical.

There are two dimensions along which notions of behavioural distances for (quantitative) transition systems can be generalized: first, instead of only considering metrics that return a real-valued distance, one can use an arbitrary quantale, yielding the notion of quantale-valued relations or conformances that subsume equivalences, preorders and (directed) metrics. Second, one can abstract from the branching type of the transition system under consideration, using a functor specifying various forms of branching (non-deterministic, probabilistic, weighted, etc.). This leads us to a coalgebraic framework [24] that provides several techniques for studying and analyzing systems and has also been adapted to behavioural metrics [2]. A coalgebra is a function $c\colon X \to FX$ where X is a set of states and F is a (set) functor.

For defining (and computing) behavioural conformances in coalgebraic generality, a fundamental notion is the lifting of a functor F to $\overline{F}$, acting on conformances. In the case when F is a distribution functor and the conformances are metrics, there is a well-known way of obtaining $\overline{F}$ through results from transportation theory using either Kantorovich or Wasserstein liftings, which are known to coincide through the so-called Kantorovich-Rubinstein duality [29]. Recent work [2] generalized these two approaches to lifting functors to the category of pseudometric spaces, as well as to more general quantale-valued relations [3, 5]. It turns out that at this level of generality, the analogue of the Kantorovich-Rubinstein duality does not hold in general anymore. In both the Kantorovich and Wasserstein approaches, given a set X and a conformance d (preorder, equivalence, metric, ...) on X, the lifted functor $\overline{F}$ canonically determines a conformance on FX, based on (a set of) evaluation maps. Intuitively, these maps provide a way of testing the one-step behaviour of the system and generalize the calculation of the expected value taking place in the definitions of the Kantorovich/Wasserstein liftings. Combining the lifting with a subsequent reindexing with the coalgebra c results in a function whose (greatest) fixpoint is the desired behavioural conformance. In this paper we focus on directed conformances such as preorders or directed metrics (also called hemimetrics).

The aim of this paper is twofold: we consider the so-called Kantorovich lifting of functors [2] that – as opposed to the Wasserstein lifting [5] – offers some extra flexibility, since it allows the use of a set of evaluation maps and places fewer restrictions on both the functor and these predicate liftings. We study compositionality results for the Kantorovich lifting, answering the question under which conditions the lifting is compositional in the sense that $\overline{FG} = \overline{F}\,\overline{G}$. This compositionality result is then an essential ingredient in adapting up-to techniques to the setting of behavioural metrics based on the Kantorovich lifting, inspired by the results of [5], which were developed for the Wasserstein lifting. Up-to techniques are coinductive techniques that allow small witnesses for lower bounds of behavioural distances by exploiting an algebraic structure on the state space.

We first set up a framework based on Galois connections and fibred adjunctions, extending [4]. This sets the stage for the definition of the Kantorovich lifting based on this adjunction. We next answer the question of compositionality positively for the case where F is polynomial with finite coproducts, and show several negative results for combinations of powerset and distribution functor.

The positive compositionality result for the case where F is polynomial with finite coproducts opens the door to developing up-to techniques for trace-like behavioural conformances that are computed on determinized transition systems, or – more generally – determinized coalgebras. More concretely, we consider coalgebras of type $c\colon X \to FTX$ where F is a finite coproduct polynomial functor, providing the explicit branching type of the coalgebra, and T is a monad, describing the implicit branching. For instance, in the case of a non-deterministic automaton we would use $FX = 2 \times X^A$ and $T = \mathcal{P}$ (powerset functor). There is a well-known determinization construction [17, 26] that transforms such a coalgebra into $c^\#\colon TX \to FTX$ via a distributive law $\zeta\colon TF \Rightarrow FT$. This yields a determinized system $c^\#$ acting on the state space TX, which has an algebraic structure given by the multiplication of the monad. A behavioural conformance is then computed on TX and passed back to X using the unit of the monad.

As TX might be very large (e.g., in the case of $T = \mathcal{P}$) or even infinite (e.g., in the case of $T = \mathcal{D}$, distribution functor), it can be hard to compute conformances for $c^\#$. However, the algebraic structure on TX can be fruitfully employed using up-to techniques [23, 6] that allow to consider post-fixpoints up to the algebraic structure, making it much easier to display a witness proving a (lower) bound on the distance of two states. The validity of the up-to technique rests on a compatibility property that is ensured whenever the distributive law ζ can be lifted, i.e., if it is non-expansive wrt. the lifted conformances. We show that this holds, where an essential step in the proof relies on the compositionality results.

In this sense, we complement the up-to techniques in [5] that provide a similar result for the Wasserstein lifting. This enables us to use the up-to technique for Kantorovich liftings, which are more versatile and allow more control over the distance values, in particular in the presence of products and coproducts. Indeed, as Wasserstein liftings are based on couplings, which in general need not exist [2], they often produce trivial distance values. We will see later in the paper that several of our case studies can only be treated appropriately in the Kantorovich case. Furthermore, we can show the lifting of the distributive law for an entire class of functors (polynomial functors with finite coproducts), while [5] contained generic results for a different class of canonical liftings. In the non-canonical case it was necessary to prove a complex condition ensuring compositionality in [5].

We apply our technique to several examples, such as trace metrics for probabilistic automata and trace semantics for systems with exceptions. We give concrete instances where up-to techniques help and show how witnesses yielding upper bounds can be reduced in size or even become finite (as opposed to infinite).

2 Preliminaries

We begin by recalling some relevant definitions and fix some notation.

As outlined in the introduction, we use quantales as the domain for behavioural conformances. A *quantale* is a complete lattice $(\mathcal{V}, \sqsubseteq)$ that is equipped with a commutative monoid structure $(\mathcal{V}, \otimes, k)$ (where k is the unit of $\otimes$) such that $\otimes$ is *join-continuous*, that is, $a \otimes \bigsqcup b_i = \bigsqcup a \otimes b_i$ for each $a \in \mathcal{V}$ and each family b_i in $\mathcal{V}$, where $\bigsqcup$ denotes least upper bounds. Join-continuity of $\otimes$ implies that the operation $a \otimes -$ has a right adjoint, which we denote by $d_\mathcal{V}(a, -)$. This means that we have $a \otimes b \sqsubseteq c \iff b \sqsubseteq d_\mathcal{V}(a, c)$ for all $a, b, c \in \mathcal{V}$. The operation $d_\mathcal{V}$ is called the *residuation* or *internal hom* of the quantale.

We work with three main examples of quantales. The first is the *Boolean quantale* $2_\sqcap$, consisting of two elements $\bot \sqsubseteq \top$ and with binary meet $\sqcap$ as the monoid structure. In this quantale, the unit k is $\top$, and residuation is Boolean implication: $d_\mathcal{V}(a, b) = a \to b = \neg a \sqcup b$.

The second main example is the *unit interval quantale* $[0,1]_\oplus$, where the underlying lattice is the unit interval $[0,1]$ under the reversed order (that is, $\sqsubseteq\, = \,\geq$), and with *truncated addition* $a \oplus b = \min(a+b,1)$ as the monoid structure. In this case, the unit of $\oplus$ is 0, and its residuation is *truncated subtraction*, which is given by $d_\mathcal{V}(a,b) = b \ominus a = \max(b-a,0)$. The third main example is the quantale $[0,\infty]_+$ of *extended positive reals*, with structure analogous to $[0,1]_\oplus$, i.e. with reversed lattice order and using the extended addition (with ∞) as the monoid operation.

▶ **Remark 1.** As many of our examples use the real-valued quantales $[0,1]_\oplus$ and $[0,\infty]_+$, where the order is reversed, we reserve the use of the symbols $\geq$ and $\leq$ for the usual order in the reals, and instead use $\sqsubseteq$ and $\sqsupseteq$ when working with general quantales. Similarly, we use $\bigsqcup$ and $\bigsqcap$ for joins and meets in the quantalic order, but switch to inf and sup when working in $[0,1]_\oplus$ or $[0,\infty]_+$.

We consider several types of conformances based on a quantale $\mathcal{V}$. First, given a set X, we may consider $\mathcal{V}$-valued endorelations on X, that is, maps of type $d\colon X \times X \to \mathcal{V}$. We call these structures $\mathcal{V}$*-graphs* [20], and write $\mathcal{V}\text{-}\mathsf{Graph}_X$ for the set of $\mathcal{V}$-graphs with underlying set X. Each set $\mathcal{V}\text{-}\mathsf{Graph}_X$ is a complete lattice where both the order and joins are pointwise, that is $d \sqsubseteq d'$ if $d(x,y) \sqsubseteq d'(x,y)$ for all $x,y \in X$. Given two $\mathcal{V}$-graphs $d_X \in \mathcal{V}\text{-}\mathsf{Graph}_X$ and $d_Y \in \mathcal{V}\text{-}\mathsf{Graph}_Y$ we say that a map $f\colon X \to Y$ is *non-expansive* or a $\mathcal{V}$*-functor* if $d_X \sqsubseteq d_Y \circ (f \times f)$ in $\mathcal{V}\text{-}\mathsf{Graph}_X$ and in this case we write $f\colon (X,d_X) \to (Y,d_Y)$. $\mathcal{V}$-graphs and non-expansive maps form a category $\mathcal{V}\text{-}\mathsf{Graph}$.

▶ **Remark 2.** In some parts of the literature, the category $\mathcal{V}\text{-}\mathsf{Graph}$ is denoted by $\mathcal{V}\text{-}\mathsf{Rel}$ instead [5]. We opt for $\mathcal{V}\text{-}\mathsf{Graph}$ as in [20], as $\mathcal{V}\text{-}\mathsf{Rel}$ more often denotes the category with sets as objects and $\mathcal{V}$-valued relations between them as morphisms [13].

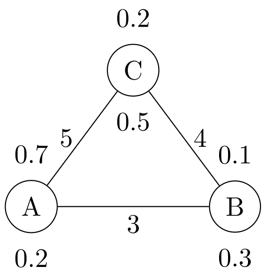
Second, within $\mathcal{V}\text{-}\mathsf{Graph}$ we consider the subcategory consisting of those $\mathcal{V}$-graphs that satisfy additional axioms, corresponding to a generalized notion of a metric space, or, equivalently, (small) categories enriched over $\mathcal{V}$ [20]. A $\mathcal{V}$*-category* is an object of $\mathcal{V}\text{-}\mathsf{Graph}_X$ for some set X where we additionally have $k \sqsubseteq d(x,x)$ and $d(x,y) \otimes d(y,z) \sqsubseteq d(x,z)$ for all $x,y,z \in X$. Instantiated to the quantales $2_\sqcap$ and $[0,1]_\oplus$, $\mathcal{V}$-categories correspond precisely to *preorders* and 1*-bounded hemimetric spaces*, respectively. The quantale $\mathcal{V}$ itself becomes a $\mathcal{V}$-category when equipped with the residuation $d_\mathcal{V}$. The class of all $\mathcal{V}$-categories is denoted by $\mathcal{V}\text{-}\mathsf{Cat}$, and the set of $\mathcal{V}$-categories based on a set X is denoted by $\mathcal{V}\text{-}\mathsf{Cat}_X$.

In this paper, we use a coalgebraic framework. Recall that a *coalgebra* is a pair (X,c), where X is the *state space* and $c\colon X \to FX$ is the transition structure parametric on an endofunctor $F\colon \mathsf{Set} \to \mathsf{Set}$. We also utilize two specific functors for (counter)examples below. The *powerset functor* is defined as $\mathcal{P}(X) = \{U \mid U \subseteq X\}$ on sets and $\mathcal{P}(f)(U) = \{f(x) \mid x \in U\}$ on functions. The *countably supported distribution functor* is defined as $\mathcal{D}(X) = \left\{p : X \to [0,1] \mid \sum_{x \in X} p(x) = 1, \mathsf{supp}(p) \text{ is countable}\right\}$ on sets, where $\mathsf{supp}(p) = \{x \mid p(x) \neq 0\}$, and as $\mathcal{D}(f)(p)(y) = \sum\{p(x) \mid x \in X, f(x) = y\}$ on functions.

Given $f_i\colon X \to Y_i$, $i \in \{1,2\}$, we denote by $\langle g_1, g_2\rangle\colon X \to Y_1 \times Y_2$ the mediating morphism of the product, namely $\langle g_1, g_2\rangle(x) = (g_1(x), g_2(x))$. Given $g_i\colon X_i \to Y$, $i \in \{1,2\}$, $[g_1,g_2]\colon X_1 + X_2 \to Y$ is the mediating morphism of the coproduct, namely $[g_1,g_2](x) = g_i(x)$ if $x \in X_i$.

3 Motivation from Transportation Theory

In this section, we give a brief description of the original Kantorovich distance on probability distributions, before we introduce its category theoretic generalization. Motivated by the transportation problem [29], the Kantorovich distance on probability distributions aims to maximize the transport by finding the optimal flow of commodities that satisfies demand from supplies and minimizes the flow cost. In fact, it is based on the dual version of this problem that asks for the optimal price function. For the sake of the example, consider a metric space defined on a three element set $X = \{A, B, C\}$ with a distance function $d\colon X \times X \to [0, \infty]$, such that $d(A, A) = d(B, B) = d(C, C) = 0$, $d(A, B) = d(B, A) = 3$, $d(A, C) = d(C, A) = 5$ and $d(B, C) = d(C, B) = 4$. Based on this distance we now want to define a distance on probability distributions on the set X, which is a function $d^{\uparrow}\colon \mathcal{D}(X) \times \mathcal{D}(X) \to [0, \infty]$. As a concrete example, let us take the distributions P and Q, such that $P(A) = 0.7$, $P(B) = 0.1$ and $P(C) = 0.2$, while $Q(A) = 0.2$, $Q(B) = 0.3$ and $Q(C) = 0.5$.



In order to define their distance, we can interpret the three elements A, B, C as places where a certain product is produced and consumed (imagine the places are maple syrup farms, each with an adjacent café where one can eat the pancakes with the aforementioned syrup). The geographical distance between the maple syrup farms is given by the distance function d, while the above distributions model the supply (Q) and demand (P) of the product in proportion to the total supply or demand. We assume that the total value of supply is the same as the total value of demand. As the owners of the farms, we are interested in transporting the product in a way to avoid excess supply and meet all demands. We can deal with this issue in two ways: do the transport on our own or find a logistics firm which will do it for us. The Kantorovich lifting relies on the latter perspective. We assume that for each place it sets up a price for the logistics company at which it will buy a unit of our product (at farms with overproduction) or sell it (at cafés with excess demand). This is equivalent to giving a function $f\colon X \to [0, \infty]$. We require that prices are *competitive*, that is for all $x, y \in X$, we have that $|f(x) - f(y)| \le d(x, y)$, which is equivalent to saying that f is a non-expansive function from d into the Euclidean distance d_e on $[0, \infty]$. Given a pricing f, the profit made by the transportation company is given by $c_f = \sum_{x\in X} f(x)P(x) - \sum_{x\in X} f(x)Q(X) = \sum_{x\in X} f(x)\,(P(x) - Q(x))$. Because the transportation company wants to make the most profit, it is aiming to pick a pricing f maximizing the formula given above. Combining all the moving parts together we are left with formula $d^{\uparrow}(P, Q) = \max\{\sum_{x\in X} f(x)\,(P(X) - Q(X)) \mid f\colon (X, d_X) \to ([0, \infty], d_e)\}$ defining the Kantorovich distance between probability distributions P and Q.

It is not straightforward to see, but in this example an optimal pricing function is $f(A) = 0$, $f(B) = 3$, $f(C) = 5$, which can be easily seen to be non-expansive and yields a distance of 2.1.

More abstractly, the formula above can be dissected into three pieces:

1. Taking all pricing plans f, which are non-expansive with respect to the Euclidean distance on $[0, \infty]$.
2. Evaluating each of the pricing plans, by calculating the expected value of f given a distribution on the set X.
3. Picking a pricing plan which maximizes the difference between the expected values.

In the following, we will concentrate on the directed case, where distance functions can be asymmetric and d_e is the directed Euclidean distance, that is, $d_e(x, y) = y \ominus x$.

In the category-theoretic generalization [2, 3] the calculation of the expected value (step 2) is replaced with (the set of) evaluation functions, intuitively scoring or testing the observable behaviour. At the same time, the steps of taking all non-expansive plans for a given metric and of generating a metric that maximizes the difference between expected values (steps 1 and 3) generalize to the setting of quantales and their residuation. In the next section, we show that the generalizations of those two steps form a fibred adjunction.

4 Setting Up a Fibred Adjunction

One of the key aspects of this paper is equipping sets of states of coalgebras with an extra structure of conformances (in particular preorders or hemimetrics). The very idea of "extra structure" can be phrased formally through the lenses of fibrations and fibred category theory, extending the ideas of [4]. In particular, we show that those results can be strengthened to the setting of fibred adjunctions.

The category $\mathcal{V}$-SPred. The adjunction-based framework from [4], besides working with $\mathcal{V}$-graphs, makes use of sets of $\mathcal{V}$-valued predicates, i.e., maps of the form $p\colon X \to \mathcal{V}$. We will use $\mathcal{V}\text{-}\mathsf{SPred}_X$ to denote the collection of sets of $\mathcal{V}$-valued predicates over some set X. A morphism between sets $S \subseteq \mathcal{V}^X$ and $T \subseteq \mathcal{V}^Y$ is a function $f\colon X \to Y$ satisfying $f^\bullet(T) := \{q \circ f \mid q \in T\} \subseteq S$, where $f^\bullet$ describes reindexing. We obtain a category $\mathcal{V}$-SPred with objects being pairs $(X, S \subseteq \mathcal{V}^X)$ and arrows are defined as above.

Grothendieck completion. It turns out that both $\mathcal{V}$-Graph/$\mathcal{V}$-Cat and $\mathcal{V}$-SPred can be viewed as fibred categories [20, 5]. We here only consider fibred categories arising from the Grothendieck construction, which can be viewed equivalently as a special kind of split indexed categories, that in our case are functors $A : \mathsf{Set}^{\mathrm{op}} \to \mathsf{Pos}$. Intuitively, such functors assign to each set a poset of "extra structure" and take functions $f : X \to Y$ to monotone maps canonically reindexing the "extra structure" on set Y by f.

The *Grothendieck completion* [14] of a functor $A\colon \mathsf{Set}^{\mathrm{op}} \to \mathsf{Pos}$ is a category denoted $\int A$, whose objects are pairs (X, i) where $X \in \mathsf{Set}$ and $i \in A(X)$. The arrows $(X, i) \to (Y, j)$ are maps $f\colon X \to Y$ such that $i \sqsubseteq A(f)(j)$, where $\sqsubseteq$ is the partial order on $A(X)$. The corresponding fibration is given by the forgetful functor $U\colon \int A \to \mathsf{Set}$ taking each pair (X, i) to its underlying set X. The fibre of each set X corresponds to the poset $A(X)$.

$\mathcal{V}$-Graph/$\mathcal{V}$-Cat and $\mathcal{V}$-SPred as Grothendieck completions. The category $\mathcal{V}$-Graph arises as the Grothendieck completion of the functor $\Phi : \mathsf{Set}^{\mathrm{op}} \to \mathsf{Pos}$ that takes each set X to $\Phi(X) = (\mathcal{V}\text{-}\mathsf{Graph}_X, \sqsubseteq)$, the lattice of $\mathcal{V}$-valued relations equipped with the pointwise order $\sqsubseteq$. Given a function $f\colon X \to Y$, we define $\Phi(f) = f^*$ by reindexing, where $f^*(d_Y) = d_Y \circ (f \times f)$. Analogously for $\mathcal{V}$-Cat.

Similarly, to obtain $\mathcal{V}$-SPred, we define a functor $\Psi : \mathsf{Set}^{\mathrm{op}} \to \mathsf{Pos}$ that maps each set X to $\Psi(X) = (\mathcal{V}\text{-SPred}_X, \supseteq)$, the lattice of collections of $\mathcal{V}$-valued predicates on X ordered by reverse inclusion. Furthermore $\Psi(f) = f^\bullet$.

Galois connection on the fibres. In the adjunction-based framework from [4], the Kantorovich lifting of a functor is phrased through the means of a contravariant Galois connection between the fibres of $\mathcal{V}$-Cat and $\mathcal{V}$-SPred and we here generalize from $\mathcal{V}$-Cat to $\mathcal{V}$-Graph. For each set X, we define a map $\alpha_X : \mathcal{V}\text{-SPred}_X \to \mathcal{V}\text{-Graph}_X$ given by:

$$\alpha_X(S)(x_1, x_2) = \bigsqcap_{f\in S} d_{\mathcal{V}}(f(x_1), f(x_2)) \qquad (S \subseteq \mathcal{V}^X)$$

Intuitively, α_X takes a collection S of $\mathcal{V}$-valued predicates on X and generates the greatest conformance on X that turns all predicates into non-expansive maps. For the other part of the Galois connection, we have a map $\gamma_X : \mathcal{V}\text{-Graph}_X \to \mathcal{V}\text{-SPred}_X$ defined by the following:

$$\gamma_X(d_X) = \{f : X \to \mathcal{V} \mid d_{\mathcal{V}} \circ (f \times f) \sqsupseteq d_X\}$$

Given a conformance $d_X : X \times X \to \mathcal{V}$, γ_X generates a set of $\mathcal{V}$-valued predicates on X which are non-expansive maps from d_X to the residuation distance. As mentioned before, we can instantiate the previous result [4, Theorem 7] and obtain the following:

▶ **Theorem 3** ([4]). *Let X be an arbitrary set, $d_X : X \times X \to \mathcal{V}$ a $\mathcal{V}$-graph and $S \subseteq \mathcal{V}^X$ a collection of $\mathcal{V}$-valued predicates. Then α_X and γ_X are both antitone (in $\subseteq$, $\sqsubseteq$) and form a contravariant Galois connection:*

$$d_X \sqsubseteq \alpha_X(S) \iff S \subseteq \gamma_X(d_X).$$

▶ **Example 4.** In the setting of Section 3, γ corresponds to Step 1 and α to Step 3.

Fibred Adjunction. Theorem 3 is a "local" property that only holds fiberwise. However, it turns out that we can argue something stronger, namely that we have a fibred adjunction situation. This is a "global" property, as fibred functors additionally preserve the notion of reindexing. Note that every natural transformation between functors of type $\mathsf{Set}^{\mathrm{op}} \to \mathsf{Pos}$ (a so-called morphism of split indexed categories) [14, Definition 1.10.5] induces a fibred functor between the corresponding Grothendieck completions.

One can quite easily verify that α is natural on $\mathcal{V}$-Graph, while γ is only laxly natural on $\mathcal{V}$-Graph and natural on $\mathcal{V}$-Cat. For the latter result, we rely on a quantalic version of the McShane-Whitney extension theorem [22, 30], mentioned also in [20] for the real-valued case.

▶ **Theorem 5.** *Let $d_X \in \mathcal{V}\text{-Cat}_X$ and $d_Y \in \mathcal{V}\text{-Cat}_Y$ be elements of $\mathcal{V}$-Cat. If $i : (Y, d_Y) \to (X, d_X)$ is an isometry, then for any non-expansive map $f : (Y, d_Y) \to (\mathcal{V}, d_{\mathcal{V}})$ there exists a non-expansive map $g : (X, d_X) \to (\mathcal{V}, d_{\mathcal{V}})$ such that $f = g \circ i$.*

▶ **Proposition 6.** *We have that $\alpha : \Psi \Rightarrow \Phi$ is natural and $\gamma : \Phi \Rightarrow \Psi$ is laxly natural, that is for all functions $f : X \to Y$ and all $\mathcal{V}$-valued relations d_Y on the set Y we have that $(f^\bullet \circ \gamma_Y)(d_Y) \subseteq (\gamma_X \circ f^*)(d_Y)$. When restricted to $\mathcal{V}$-Cat, $\gamma : \Phi \Rightarrow \Psi$ is natural.*

Note that we can safely make this restriction, while still keeping α to be well-defined, as its image always lies within $\mathcal{V}$-Cat.

▶ **Proposition 7.** *For all sets X and $S \subseteq \mathcal{V}^X$, $\alpha_X(S)$ is a $\mathcal{V}$-category, i.e., an object of $\mathcal{V}$-Cat. The co-closure $\alpha_X \circ \gamma_X$ is the identity when restricted to $\mathcal{V}$-Cat. Combined, this implies that for $d \in \mathcal{V}\text{-Graph}_X$, $\alpha_X(\gamma_X(d))$ is the metric closure of d, i.e. the least element of $\mathcal{V}\text{-Cat}_X$ above d.*

Because of the Grothendieck construction ([14, Theorem 1.10.7]), α and γ respectively correspond to fibred functors $\alpha\colon \mathcal{V}\text{-SPred} \to \mathcal{V}\text{-Cat}$ and $\gamma : \mathcal{V}\text{-Cat} \to \mathcal{V}\text{-SPred}$. Both functors keep morphisms unchanged and act on objects by respectively applying the appropriate components of α and γ. Now, we can state the strengthened version of Theorem 3, by showing that α and γ form a fibred adjunction. Note that due to the choice of the orderings, γ becomes the left and α the right adjoint (cf. [19]).

▶ **Theorem 8.** *There is an adjunction $\gamma \dashv \alpha$.*

5 The Coalgebraic Kantorovich Lifting

5.1 Definition of the Coalgebraic Kantorovich Lifting

The coalgebraic Kantorovich lifting [2] (originally defined for the real-valued case and for a single evaluation map) – extended to codensity liftings in [18] – is parametric in a set of *evaluation functions* for a set functor F. Evaluation functions are maps of type $F\mathcal{V} \to \mathcal{V}$ (generalizing the expected value computation in the traditional Kantorovich lifting) and as such can be used to lift $\mathcal{V}$-valued predicates on a set X to $\mathcal{V}$-valued predicates on the set FX. More precisely, given an evaluation function $ev\colon F\mathcal{V} \to \mathcal{V}$ and a predicate $f\colon X \to \mathcal{V}$ we obtain the predicate $ev \circ Ff\colon FX \to \mathcal{V}$. This operation extends to sets of evaluation functions and sets of $\mathcal{V}$-valued predicates, where a set Λ^F of evaluation functions for F induces the fibred functor $\Lambda^F\colon \mathcal{V}\text{-SPred} \to \mathcal{V}\text{-SPred}$, defined on $S \subseteq \mathcal{V}^X$ as follows:

$$\Lambda^F_X(S) = \{ev \circ Ff \mid ev \in \Lambda^F, f \in S\} \subseteq \mathcal{V}^{FX}.$$

The Kantorovich lifting can now be restated via the fibred adjunction introduced previously. Given F and Λ^F as above, we can define its Kantorovich lifting K_{Λ^F} as follows:

$$K_{\Lambda^F} = \alpha F \circ \Lambda^F \circ \gamma$$

or more concretely, for an object d_X of $\mathcal{V}$-Graph and $s, t \in FX$:

$$K_{\Lambda^F}(d_X)(s,t) = \bigsqcap_{ev\in\Lambda^F} \bigsqcap_{f\in\gamma_X(d_X)} d_{\mathcal{V}}(ev(Ff(s)), ev(Ff(t))).$$

If the set Λ^F is clear from the context we sometimes write $\overline{F}$ instead of K_{Λ^F}.

▶ **Lemma 9.** *The Kantorovich lifting of a functor F : $\mathsf{Set} \to \mathsf{Set}$ is a functor $\overline{F} = K_{\Lambda^F}\colon \mathcal{V}\text{-Graph} \to \mathcal{V}\text{-Graph}$, and fibred when restricted to $\mathcal{V}$-Cat.*

▶ **Remark 10.** Instantiating the construction above with the distribution functor $\mathcal{D}$ and a single evaluation function $\mathbb{E}$ taking the expected values yields the usual Kantorovich lifting, while in the case of powerset functor $\mathcal{P}$ and a single evaluation function sup, one obtains the (directed) Hausdorff metric. Despite those two instantiations corresponding to well-known constructions, there is no well-defined notion of *canonical lifting* and there are often different possibilities for a given functor. For example, the usual symmetric Hausdorff distance arises by additionally considering the dual evaluation function inf [32]. We will later also see that constant factors admit more than one choice of evaluation functions.

5.2 Compositionality of the Kantorovich Lifting

When an endofunctor is given as the composition of two or more individual set functors, it is natural to ask under which conditions its Kantorovich lifting is also the composition of Kantorovich liftings of the respective functors. Specifically, our aim in this section is to identify situations where this composition happens already at the level of the underlying sets of evaluation maps. If ev_F is an evaluation function for F and ev_G is an evaluation function for G, then an evaluation function for FG is given by $ev_F * ev_G := ev_F \circ F ev_G$. Extending this to sets of evaluation functions, we put $\Lambda^F * \Lambda^G = \{ev_F * ev_G \mid ev_F \in \Lambda^F, ev_G \in \Lambda^G\}$ for sets Λ^F and Λ^G of evaluation functions for functors F and G, respectively.

▶ **Definition 11** (Compositionality). Given two functors F and G and sets Λ^F and Λ^G of evaluation functions, we say that we have *compositionality* if $K_{\Lambda^F} \circ K_{\Lambda^G} = K_{\Lambda^F * \Lambda^G}$.

Expanding definitions, compositionality amounts to showing that

$$K_{\Lambda^F} \circ K_{\Lambda^G} = \alpha FG \circ \Lambda^F \circ \gamma G \circ \alpha G \circ \Lambda^G \circ \gamma = \alpha FG \circ \Lambda^F \circ \Lambda^G \circ \gamma = K_{\Lambda^F * \Lambda^G}. \tag{1}$$

One inequality ("$\sqsubseteq$") always holds: As α and γ form a Galois connection [4], we have $\mathrm{id}_{\mathcal{V}\text{-}\mathsf{SPred}_X} \subseteq \gamma_{GX} \circ \alpha_{GX}$, and thus we may use antitonicity of α to deduce "$\sqsubseteq$" in (1). Baldan et al. [2, Lemma 7.5] prove this for the special case of pseudometric liftings.

The other inequality, "$\sqsupseteq$", does not hold in general, and requires more work. Still, one notices that a sufficient condition is that $\Lambda^F \circ \gamma \circ \alpha \subseteq \gamma F \circ \alpha F \circ \Lambda^F$:

$$\begin{aligned} K_{\Lambda^F} \circ K_{\Lambda^G} &= \alpha FG \circ \Lambda^F \circ \gamma G \circ \alpha G \circ \Lambda^G \circ \gamma \\ &\sqsupseteq \alpha FG \circ \gamma FG \circ \alpha FG \circ \Lambda^F \circ \Lambda^G \circ \gamma = \alpha FG \circ \Lambda^{FG} \circ \gamma = K_{\Lambda^F * \Lambda^G}, \end{aligned}$$

using that $\alpha \circ \gamma \circ \alpha = \alpha$ for every Galois connection. Note that it is enough to prove the sufficient condition on non-empty sets, since γ always yields a non-empty set.

Before discussing the problem of systematically constructing sets of evaluation functions such that the sufficient condition (lax commutativity of Λ^F with the closure induced by the Galois connection) holds, we consider a few examples where compositionality fails:

▶ **Example 12.** Consider the powerset functor $\mathcal{P}$ with the predicate lifting sup ($\Lambda^{\mathcal{P}} = \{\sup\}$), and the discrete distribution functor $\mathcal{D}$ with the predicate lifting $\mathbb{E}$ that takes expected values ($\Lambda^{\mathcal{D}} = \{\mathbb{E}\}$). With just these predicate liftings, compositionality fails for all four combinations $\mathcal{PP}$, $\mathcal{PD}$, $\mathcal{DP}$, $\mathcal{DD}$. We show this for the case of $\mathcal{PD}$ and discuss the others in [10]. Let X be the two-element set $\{x, y\}$, equipped with the discrete metric d (that is, $d(x,y) = d(y,x) = 1$), so that in particular all maps $g\colon X \to [0,1]$ are non-expansive. We also consider the non-expansive function $f_{\mathcal{D}}\colon (\mathcal{D}X, K_{\{\mathbb{E}\}}d) \to ([0,1], d_{[0,1]_\oplus})$ given by $f_{\mathcal{D}}(p \cdot x + (1-p) \cdot y) = \min(p, 1-p)$. Put $U = \{1 \cdot x, 1 \cdot y\}$ and $V = \{1 \cdot x, 1/2 \cdot x + 1/2 \cdot y, 1 \cdot y\}$. Then $\sup f_{\mathcal{D}}[U] = \max(0,0) = 0$ and $\sup f_{\mathcal{D}}[V] = \max(0, 1/2, 0) = 1/2$, so that $K_{\{\sup\}}(K_{\{\mathbb{E}\}}d)(U,V) \geq 1/2$. For every $g\colon X \to [0,1]$ one finds that

$$(\sup * \mathbb{E})(g)(U) = \max(g(x), g(y)) = \max(g(x), g(x)+g(y)/2, g(y)) = (\sup * \mathbb{E})(g)(V),$$

implying that $K_{\{\sup * \mathbb{E}\}}d(U,V) = 0$.

5.3 Finite Coproduct Polynomial Functors

We now assume that the first functor (F with the lifting $\overline{F} = K_{\Lambda_F}$) is in fact a polynomial functor (with finite coproducts) and we show that in this case compositionality holds automatically for certain sets of predefined evaluation maps. This will later allow us to use compositionality to define up-to techniques for large classes of coalgebras that are based on such functors.

Consider the set of polynomial functors (with finite coproducts) given by

$$F ::= C_B \mid \mathrm{Id} \mid \textstyle\prod_{i \in I} F_i \mid F_1 + F_2$$

where C_B is the constant functor mapping to some set B and Id is the identity functor. We support products over arbitrary index sets, but we restrict to finitary coproducts for simplicity.

For such polynomial functors we can obtain compositionality in a structured manner, by constructing suitable sets of predicate liftings alongside with the functors themselves. We recursively define a set Λ^F of evaluation functions for each polynomial functor F as follows:

constant functors: $F = C_B$. Here we choose Λ^F to be any set of maps of type $B \to \mathcal{V}$. (For instance, when $B = \mathcal{V}$ we can put $\Lambda^F = \{\mathrm{id}_{\mathcal{V}}\}$.)

identity functor: $F = \mathrm{Id}$. We put $\Lambda^F = \{\mathrm{id}_{\mathcal{V}}\}$.

product functors: $F = \prod_{i \in I} F_i$. Put $\Lambda^F = \{ev_i \circ \pi'_i \mid i \in I, ev_i \in \Lambda^{F_i}\}$ where $\pi'_i \colon \prod_{i \in I} F_i\mathcal{V} \to F_i\mathcal{V}$ are the projections.

coproduct functors: $F = F_1 + F_2$. We put $\Lambda^F = \{[ev_1, \top] \mid ev_1 \in \Lambda^{F_1}\} \cup \{[\bot, ev_2] \mid ev_2 \in \Lambda^{F_2}\} \cup \{[\bot, \top]\}$, where $\top$ and $\bot$ denote constant maps into $\mathcal{V}$.

▶ **Remark 13.** We note that the construction for coproduct functors is associative, that is, for functors F_1, F_2 and F_3 the sets $\Lambda^{(F_1+F_2)+F_3}$ and $\Lambda^{F_1+(F_2+F_3)}$ coincide up to isomorphism.

Note also that the exponentiation $F^A X = (FX)^A$ is special case of the product where $I = A$ and $F_i = F$ for all $i \in I$.

▶ **Remark 14.** Throughout the paper, we restrict our attention to finite coproducts for the sake of simplicity, but we would like to note that our construction could be generalized to infinite sets. In general, since our lifting for the coproduct will be based on prioritization, we need to compare the sets in order of preference, i.e. have the additional structure of a well-order. This immediately works for countable sets.

The choice of evaluation maps above induces the following liftings, leading to the natural expected distances in the directed case.

▶ **Proposition 15.** *Given a polynomial functor F and a set of evaluation maps Λ^F as defined above, the corresponding lifting $\overline{F} = K_{\Lambda^F}$ is defined as follows on objects of $\mathcal{V}$-*Graph*: $\overline{F}(d_X) = d_X^F \colon FX \times FX \to \mathcal{V}$ where*

constant functors: $d_X^F \colon B \times B \to \mathcal{V}$, $d_X^F(b, c) = \bigsqcap_{ev \in \Lambda^F} d_{\mathcal{V}}(ev(b), ev(c))$.

identity functor: $d_X^F \colon X \times X \to \mathcal{V}$ *with* $d_X^F = \alpha_X(\gamma_X(d))$.

product functors: $d_X^F \colon \prod_{i \in I} F_iX \times \prod_{i \in I} F_iX \to \mathcal{V}$, $d_X^F(s, t) = \bigsqcap_{i \in I} d_X^{F_i}(\pi_i(s), \pi_i(t))$ *where* $\pi_i \colon \prod_{i \in I} F_iX \to F_iX$ *are the projections.*

coproduct functors: $d_X^F \colon (F_1X + F_2X) \times (F_1X + F_2X) \to \mathcal{V}$, *where*

$$d_X^F(s,t) = \begin{cases} d^{F_i}(s,t) & \text{if } s,t \in F_iX \text{ for } i \in \{1,2\} \\ \top & \text{if } s \in F_1X,\ t \in F_2X \\ \bot & \text{if } s \in F_2X,\ t \in F_1X \end{cases}$$

Under this choice of evaluation functions we can show the following, which implies compositionality (cf. Section 5.2):

▶ **Proposition 16.** *For every polynomial functor F and the corresponding set Λ^F of evaluation maps (as above) we have $\Lambda^F \circ \gamma \circ \alpha \subseteq \gamma F \circ \alpha F \circ \Lambda^F$ on non-empty sets of predicates.*

Using the arguments of Section 5.2, we infer:

▶ **Corollary 17.** *Let F and G be functors, and Λ^F and Λ^G be sets of predicate liftings for them. If F and λ^F are as in Proposition 16, then $K_{\Lambda^F} \circ K_{\Lambda^G} = K_{\Lambda^F * \Lambda^G}$.*

▶ **Example 18.** We consider a running example specifying standard directed trace metrics for probabilistic automata as introduced in [17]. We take the polynomial functor $F = [0,1] \times \_^A$ ("machine functor"), monad $T = \mathcal{D}$ and quantale $\mathcal{V} = [0,1]_\oplus$. Furthermore we use expectation ($\mathbb{E}$) as evaluation map for T and as evaluation maps for the functor F we take ev_* mapping to the first component and ev_a (for each $a \in A$) with $ev_a(r,g) = g(a)$. These evaluation maps are of the type described in this section and hence we have compositionality.

Note that this example is not directly realizable in the Wasserstein approach [5]: the issue with the Wasserstein lifting is that whenever no coupling of two elements exists, the distance is automatically the bottom element in the quantale. This can be seen for the functor F where $(r_1,x),(r_2,x) \in FX$ have no coupling whenever $r_1 \neq r_2$. Hence it is harder to parameterize and would not work here.

Using a set of evaluation maps Λ^F as opposed to a single evaluation map gives us additional flexibility.

6 Application: Up-To Techniques

We now adapt results from [5] on up-to techniques from Wasserstein to Kantorovich liftings. In particular, we instantiate the fibrational approach to coinductive proof techniques from [6] that allows to prove lower bounds for greatest fixpoints, using post-fixpoints up-to as witnesses. As shown in the running example and in Section 7 this can greatly help to reduce the size of such witnesses, even allowing finitary witnesses which would be infinite otherwise.

6.1 Introduction to Up-To Techniques

We first recall the notion of a bialgebra [15], a coalgebra with a compatible algebra structure.

▶ **Definition 19.** Consider two functors F, T and a natural transformation $\zeta\colon TF \Rightarrow FT$. An *$F$-$T$-bialgebra* for ζ is a tuple (Y,a,c) such that $a\colon TY \to Y$ is a T-algebra and $c\colon Y \to FY$ is an F-coalgebra so that the diagram below commutes.

$$\begin{array}{ccccc} TY & \xrightarrow{a} & Y & \xrightarrow{c} & FY \\ \downarrow{\scriptstyle Tc} & & & & \uparrow{\scriptstyle Fa} \\ TFY & & \xrightarrow{\zeta_Y} & & FTY \end{array}$$

In order to construct such bialgebras, distributive laws exchanging functors and monads are helpful.

▶ **Definition 20.** A *distributive law* or *EM-law* of a monad $T : \mathbf{C} \to \mathbf{C}$ with unit $\eta : \mathrm{Id} \Rightarrow T$ and multiplication $\mu : TT \Rightarrow T$ over a functor $F : \mathbf{C} \to \mathbf{C}$ is a natural transformation $\zeta : TF \Rightarrow FT$ such that the following diagrams commute:

$$\begin{array}{ccc} FX & & \\ {\scriptstyle \eta_{FX}}\downarrow & \searrow{\scriptstyle F\eta_X} & \\ TFX & \xrightarrow[\zeta_X]{} & FTX \end{array} \qquad \begin{array}{ccccc} T^2FX & \xrightarrow{T\zeta_X} & TFTX & \xrightarrow{\zeta_{TX}} & FT^2X \\ {\scriptstyle \mu_{FX}}\downarrow & & & & \downarrow{\scriptstyle F\mu_X} \\ TFX & & \xrightarrow[\zeta_X]{} & & FTX \end{array}$$

Whenever T is a monad and ζ is an EM-law, then an F-T-bialgebra can be obtained by determinizing a coalgebra $c\colon X \to FTX$. More concretely, we obtain $c^\# : Y \to FY$ where $Y = TX$ and $c^\# = F\mu_X \circ \zeta_{TX} \circ Tc$. The algebra map is $a = \mu_X : TY \to Y$.

We now assume a bialgebra (Y, a, c) and Kantorovich liftings $\overline{T} = K_{\Lambda^T}$, $\overline{F} = K_{\Lambda^F}$ of T, F. Based on this we can define a *behaviour function* beh via

$$\mathcal{V}\text{-}\mathsf{Graph}_Y \xrightarrow{\overline{F}} \mathcal{V}\text{-}\mathsf{Graph}_{FY} \xrightarrow{c^*} \mathcal{V}\text{-}\mathsf{Graph}_Y$$

Remember that c^* denotes reindexing via c. The greatest fixpoint of beh corresponds to a behavioural conformance (e.g., behavioural equivalence or bisimulation metric).

▶ **Example 21.** We continue with Example 18. We use the standard distributive law $\zeta\colon TF \Rightarrow FT$ given by the following components where $\mathbb{E}_\mu \pi_1 = \mathbb{E}(\mathcal{D}\pi_1(\mu))$:

$$\begin{array}{rcl} \zeta_X\colon \mathcal{D}([0,1] \times X^A) & \to & [0,1] \times \mathcal{D}X^A \\ \zeta_X(\mu) & = & (\mathbb{E}_\mu \pi_1, a \mapsto \mathcal{D}(\mathsf{eval}_a \circ \pi_2)(\mu)) \end{array}$$

where $\mathsf{eval}_a(f) = f(a)$. Given an Eilenberg-Moore coalgebra $c\colon X \to FTX$ (more concretely: $c\colon X \to [0,1] \times \mathcal{D}X^A$) and its determinization $c^\#\colon TX \to FTX$, the behavioural distance on TX arises as the greatest fixpoint (in the quantale order) of the map beh $= (c^\#)^* \circ \bar{F}$ defined above.

By unravelling the fixpoint equation one can see that it coincides with the directed trace metric on probability distributions that is defined as follows: for each state $x \in X$ let $tr_x\colon A^* \to [0,1]$ be a map that assigns to each word (trace) $w \in A^*$ the expected payoff for this word when read from x, where the payoff of a state x' is $\pi_1(c(x'))$. Then

$$\nu\text{beh}(p,q) = \sup_{w \in A^*} \big(\sum\nolimits_{x\in X} tr_x(w) \cdot q(x) \ominus \sum\nolimits_{x \in X} tr_x(w) \cdot p(x)\big)$$

If p, q are Dirac distributions δ_x, δ_y, we have: $\nu\text{beh}(\delta_x, \delta_y) = \sup_{w\in A^*}(tr_y(w) \ominus tr_x(w))$.

One can typically avoid computing the full fixpoint νbeh when checking the behavioural distance of two states; this is facilitated through the use of an *up-to function* u defined via

$$\mathcal{V}\text{-}\mathsf{Graph}_Y \xrightarrow{\overline{T}} \mathcal{V}\text{-}\mathsf{Graph}_{TY} \xrightarrow{\Sigma_a} \mathcal{V}\text{-}\mathsf{Graph}_Y$$

where $\Sigma_f\colon \mathcal{V}\text{-}\mathsf{Graph}_X \to \mathcal{V}\text{-}\mathsf{Graph}_Y$ is defined as $\Sigma_f(d)(y_1, y_2) = \bigsqcup_{f(x_i)=y_i} d(x_1, x_2)$ for $f\colon X \to Y$ (direct image).

Both functions (beh, u) are monotone functions on a complete lattice. Hence we can use the Knaster-Tarski theorem [27] and the theory of up-to techniques [23]. In particular, given a monotone function $f\colon L \to L$ over a complete lattice $(L, \sqsubseteq)$, we have the guarantee that $\ell \sqsubseteq f(\ell)$ for $\ell \in L$ guarantees $\ell \sqsubseteq \nu f$, i.e., a post-fixpoint of f is always a lower bound for the greatest fixpoint νf, an essential proof rule in coinductive reasoning. Even more widely applicable are proof rules based on up-to functions. An up-to function is a monotone function $u\colon L \to L$ that is f-compatible (i.e., $u \circ f \sqsubseteq f \circ u$). Then we can infer that $\ell \sqsubseteq f(u(\ell))$ (i.e., ℓ is a post-fixpoint up-to u) implies $\ell \sqsubseteq \nu f$ (ℓ is a lower bound for the greatest fixpoint). Typically u is extensive ($\ell \sqsubseteq u(\ell)$) and hence it is "easier" to find a post-fixpoint up-to rather than a post-fixpoint.

From [6] we obtain the following result that ensures compatibility:

▶ **Proposition 22** ([6]). *Whenever the EM-law $\zeta\colon TF \Rightarrow FT$ lifts to $\zeta\colon \overline{T}\,\overline{F} \Rightarrow \overline{F}\,\overline{T}$, we have that $u \circ$ beh $\sqsubseteq$ beh $\circ\, u$ (for u,* beh *as defined above), i.e., u is* beh*-compatible.*

Hence, we can deduce that every post-fixpoint up-to witnesses a lower bound of the greatest fixpoint. More concretely: $d_Y \sqsubseteq \text{beh}(u(d_Y))$ implies $d_Y \sqsubseteq \nu\,\text{beh}$ (where $d_Y \in \mathcal{V}\text{-}\mathsf{Graph}_Y$) (coinduction up-to proof principle).

6.2 Lifting Distributive Laws

To use the proof technique laid out in the previous section, we have to show that ζ lifts accordingly. We start by defining distributive laws and lifting them to $\mathcal{V}$-Graph.

Let F be a polynomial functor (cf. Section 5.3) and (T, μ, η) a monad over Set. Following [16, Exercise 5.4.4], EM-laws $\zeta : TF \Rightarrow FT$ can then be constructed inductively over the structure of F, i.e., for F being an identity, constant, product and coproduct functor. In the coproduct case we extend [16] by weakening the requirement that T preserves coproducts.

In the following, we inductively construct an EM-law $\zeta : TF \Rightarrow FT$ and first lift it to $\zeta : \overline{TF} \Rightarrow \overline{FT}$ (and then to $\zeta : \overline{T}\,\overline{F} \Rightarrow \overline{F}\,\overline{T}$). For the evaluation maps we assume that Λ^F is defined as in Section 5.3 and that $\Lambda^T = \{ev_T\}$, where $ev_T : T\mathcal{V} \to \mathcal{V}$ is a T-algebra.

constant functors: For $F = C_B$ we have that $TFX = TB$ and $FTX = B$, and so define the EM-law as $\zeta : TC_B \Rightarrow C_B$, where the (unique) component $\zeta_X : TB \to B$ is an arbitrary T-algebra on B. (From now on, we assume that evaluation maps $ev : B \to \mathcal{V}$ for constant functors are T-algebra homomorphisms between $\zeta : TB \to B$ and $ev_T : T\mathcal{V} \to \mathcal{V}$.)

identity functor: For $F = \mathrm{Id}$, we let the EM-law be the identity map $\mathrm{id} : T \Rightarrow T$.

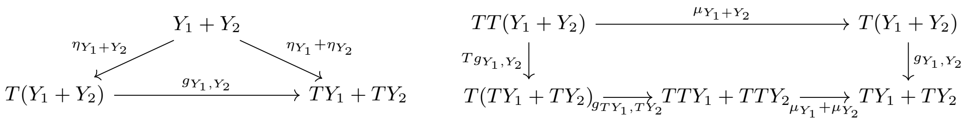
product functors: For $F = \prod_{i \in I} F_i$, assuming we have distributive laws $\zeta^i : TF_i \Rightarrow F_i T$, the EM-law is

$$\langle \zeta^i \circ T\pi_i \rangle : T \textstyle\prod_{i \in I} F_i \Rightarrow \prod_{i \in I} F_i T.$$

coproduct functors: For $F = F_1 + F_2$, assume that we have distributive laws $\zeta^i : TF_i \Rightarrow F_i T$ and a natural transformation $g : T((-) + (-)) \Rightarrow T + T$ between bifunctors. The EM-law is given by

$$(\zeta^1 + \zeta^2) \circ g_{F_1,F_2} : T(F_1 + F_2) \Rightarrow TF_1 + TF_2 \Rightarrow F_1 T + F_2 T$$

▶ **Definition 23.** Let T be a monad and let $g : T((-) + (-)) \Rightarrow T + T$ be a natural transformation as above. We say that g is *compatible with the unit* η of the monad if for all sets Y_1, Y_2 the left diagram below commutes. Analogously, g is *compatible with the multiplication* μ of the monad if the right diagram commutes for all sets Y_1, Y_2.



▶ **Proposition 24.** *Assume that the natural transformation g is compatible with unit and multiplication of T. Then the transformation ζ as defined above is an EM-law of T over F.*

In order to lift natural transformations (respectively distributive laws), we will use the following result:

▶ **Proposition 25.** *Let F, G be functors on* Set *and let the sets of evaluation maps of F and G be denoted by Λ^F and Λ^G. Let $\zeta : F \Rightarrow G$ be a natural transformation. If*

$$\Lambda^G \circ \zeta_{\mathcal{V}} := \{ev_G \circ \zeta_{\mathcal{V}} \mid ev_G \in \Lambda^G\} \subseteq \Lambda^F, \tag{2}$$

*then ζ lifts to $\zeta : \overline{F} \Rightarrow \overline{G}$ in $\mathcal{V}$-*Graph*, where $\overline{F} = K_{\Lambda^F}$, $\overline{G} = K_{\Lambda^G}$.*

We can show that the inclusion (2) (even equality) holds under some conditions.

▶ **Definition 26.** Let $g : T((-)+(-)) \Rightarrow T+T$ be a natural transformation as introduced above and let $ev_T : T\mathcal{V} \to \mathcal{V}$ be the evaluation map of the monad. We say that g is *well-behaved* wrt. ev_T if the following diagrams commute for $f_i \colon X_i \to \mathcal{V}$, where $\bot, \top$ are constant maps of the appropriate type.

$$\begin{array}{ccc} T(X_1+X_2) & \xrightarrow{T[f_1,\top_{X_2}]} & T\mathcal{V} \\ \downarrow{\scriptstyle g_{X_1,X_2}} & & \downarrow{\scriptstyle ev_T} \\ TX_1+TX_2 & \xrightarrow{[ev_T \circ Tf_1, \top_{TX_2}]} & \mathcal{V} \end{array} \qquad \begin{array}{ccc} T(X_1+X_2) & \xrightarrow{T[\bot_{X_1},f_2]} & T\mathcal{V} \\ \downarrow{\scriptstyle g_{X_1,X_2}} & & \downarrow{\scriptstyle ev_T} \\ TX_1+TX_2 & \xrightarrow{[\bot_{TX_1}, ev_T \circ Tf_2]} & \mathcal{V} \end{array} \qquad \begin{array}{ccc} T(X_1+X_2) & \xrightarrow{T[\bot_{X_1},\top_{X_2}]} & T\mathcal{V} \\ \downarrow{\scriptstyle g_{X_1,X_2}} & & \downarrow{\scriptstyle ev_T} \\ TX_1+TX_2 & \xrightarrow{[\bot_{TX_1}, \top_{TX_2}]} & \mathcal{V} \end{array}$$

▶ **Lemma 27.** *Let F be a polynomial functor and T a monad with $\Lambda^T = \{ev_T\}$.*

For distributive laws ζ as described above where the component g is well-behaved wrt. ev_T and evaluation maps as defined in Section 5.3, we have that

$$(\Lambda^F * \Lambda^T) \circ \zeta_{\mathcal{V}} = \Lambda^T * \Lambda^F.$$

Then, when we have a coalgebra of the form $Y \to FTY$ for F polynomial and T a monad as above, and we determinize it to get a coalgebra $X \to FX$ for $X = TY$, we obtain a bialgebra with the algebra structure given by the monad multiplication $\mu_Y : TX \to X$. The EM-law obtained then also forms a distributive law for the bialgebra. By Proposition 25 and Lemma 27 we know that the distributive law ζ lifts to $\mathcal{V}$-Graph, i.e., $\zeta \colon \overline{FT} \Rightarrow \overline{TF}$ where $\overline{FT} = K_{\Lambda^F * \Lambda^T}$ and $\overline{TF} = K_{\Lambda^T * \Lambda^F}$.

We now show that natural transformations g as required above do exist for the powerset and subdistribution monad for suitable quantales. Note that they are "asymmetric" and prioritize one of the two sets over the other.

▶ **Proposition 28.** *Let $T = \mathcal{P}$ be the powerset monad with evaluation map $ev_T = \sup$ for $\mathcal{V} = [0,1]_{\oplus}$. Then g_{X_1,X_2} below is a natural transformation that is compatible with unit and multiplication of T and is well-behaved.*

$$g_{X_1,X_2} \colon \mathcal{P}(X_1+X_2) \to \mathcal{P}X_1 + \mathcal{P}X_2 \qquad g_{X_1,X_2}(X') = \begin{cases} X' \cap X_1 & \textit{if } X' \cap X_1 \neq \emptyset \\ X' & \textit{otherwise} \end{cases}$$

▶ **Proposition 29.** *Let $T = \mathcal{S}$ be the subdistribution monad where $\mathcal{S}(X) = \{p \colon X \to [0,1] \mid \sum_{x \in X} p(x) \le 1\}$. Assume that its evaluation map is $ev_T = \mathbb{E}$ for the quantale $\mathcal{V} = [0,\infty]_+$ (where we assume that $p \cdot \infty = \infty$ if $p > 0$ and 0 otherwise). Then g_{X_1,X_2} below is a natural transformation that is compatible with unit and multiplication of T and is well-behaved.*

$$g_{X_1,X_2} \colon \mathcal{S}(X_1+X_2) \to \mathcal{S}X_1 + \mathcal{S}X_2 \qquad g_{X_1,X_2}(p) = \begin{cases} p|_{X_1} & \textit{if } supp(p) \cap X_1 \neq \emptyset \\ p|_{X_2} & \textit{otherwise} \end{cases}$$

It is left to show that the EM-law $\zeta \colon FT \Rightarrow TF$ lifts to $\zeta \colon \overline{F}\,\overline{T} \Rightarrow \overline{T}\,\overline{F}$, where $\overline{F} = K_{\Lambda^F}$, $\overline{T} = K_{\Lambda^T}$ where $\Lambda^T = \{ev_T\}$, and where the evaluation maps for F are obtained as described earlier. We namely prove that its components are all non-expansive, i.e., $\mathcal{V}$-Graph-morphisms, commutativity is already known. Using the previous results we obtain:

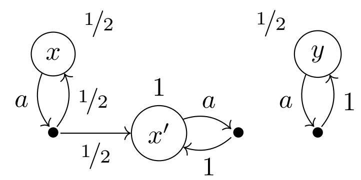
$$\overline{F}\,\overline{T} \overset{(1)}{\sqsubseteq} \overline{TF} \overset{(2)}{\Rightarrow} \overline{FT} \overset{(3)}{=} \overline{F}\,\overline{T}$$

(1) follows from the inequality in Section 5.2. This implies that the identity $\mathrm{id}_{TFX} \colon \overline{T}\,\overline{F}X \to \overline{TF}X$ is non-expansive (a $\mathcal{V}$-Graph-morphism), resulting in the natural transformation $\overline{T}\,\overline{F} \overset{\mathrm{id}}{\Rightarrow} \overline{TF}$.

(2) is implied by the results of this section (Lemma 27, Proposition 25)
(3) is guaranteed by the fact that $\overline{FT} = \overline{F}\,\overline{T}$ if F is polynomial (and we have suitable evaluation maps), hence compositionality holds (Proposition 16)

▶ **Example 30.** We continue Example 21 and first observe that the distributive law given there is obtained by the inductive construction given above.

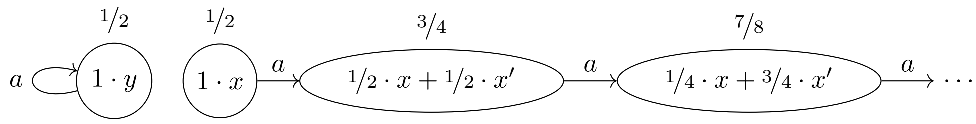
For this concrete example fix the label set as a singleton: $A = \{a\}$. Consider the coalgebra with states $X = \{x, x', y\}$ drawn below on the left. The payoff is written next to each state.



	ϵ	a	aa	aaa	...
x	$1/2$	$3/4$	$7/8$	$15/16$	...
y	$1/2$	$1/2$	$1/2$	$1/2$	...

Here, the trace map has the values for states x, y given in the table above on the right, leading to the behavioural distance $\nu\text{beh}(\delta_x, \delta_y) = 1/2$ (the supremum of all differences).

Determinizing the probabilistic automaton above leads to an automaton with infinite state space, even if we only consider the reachable states (which are probability distributions):



Our aim is now to define a witness distance of type $d\colon TX \times TX \to [0, 1]$ that is a post-fixpoint up-to in the quantale order and a pre-fixpoint for the order on the reals ($d \geq \text{beh}(u(d))$). From this we can infer that $\nu\text{beh} \leq d$, obtaining an upper bound. We set $d(1 \cdot x, 1 \cdot y) = 1/2$, $d(1 \cdot x', 1 \cdot y) = 1/2$ and $d(p, q) = 1$ for all other pairs of probability distributions p, q. We now show that d is a pre-fixpoint of beh in the order on the reals, i.e., our aim is to prove for all p, q that $d(p, q) \geq \text{beh}(u(d))(p, q)$. This is obvious for the cases where $d(p, q) = 1$, hence only two cases remain:

- If $p = 1 \cdot x$, $q = 1 \cdot y$, we have:

$$\begin{aligned}\text{beh}(u(d))(1 \cdot x, 1 \cdot y) &= \max\{u(d)(1/2 \cdot x + 1/2 \cdot x', y), |1/2 - 1/2|\} \\ &= u(d)(1/2 \cdot x + 1/2 \cdot x', y) \\ &\leq 1/2 \cdot d(1 \cdot x, 1 \cdot y) + 1/2 \cdot d(1 \cdot x', 1 \cdot y) \\ &= 1/2 \cdot 1/2 + 1/2 \cdot 1/2 = 1/2 = d(1 \cdot x, 1 \cdot y)\end{aligned}$$

- The case $p = 1 \cdot x'$, $q = 1 \cdot y$ can be shown analogously.

We are using the following inequalities: (i) $u(d)(p, q) \leq d(p, q)$ (follows from the definition of u); (ii) $u(d)(r_1 \cdot p_1 + r_2 \cdot p_2, r_1 \cdot q_1 + r_2 \cdot q_2) \leq r_1 \cdot d(p_1, q_1) + r_2 \cdot d(p_2, q_2)$ (metric congruence, see [10] for more details.). This concludes the argument.

Note that here up-to techniques in fact allow us to consider finitary witnesses for bounds for behavioural distances, even when the determinized coalgebra has an infinite state space.

7 Case Study: Transition Systems with Exceptions

In addition to the running example treated in the paper we now consider a second case study on transition systems with exceptions that helps to concretely show upper bounds (lower bounds in the quantalic order) for behavioural distances via appropriate witnesses.

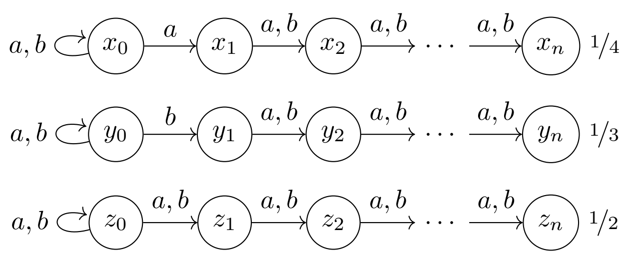
We consider a case study involving the coproduct, in particular the polynomial functor $F = [0,1] + (-)^A$ and the monad $T = \mathcal{P}$ with evaluation map sup. In a coalgebra of type $c\colon X \to FTX$, a state can either perform transitions or terminate with some output value taken from the interval $[0,1]$; in applications this value could for example be considered as the severity of the error encountered upon terminating a computation. Hence the (directed) distance of two states x, y can intuitively be interpreted as measuring how much worse the errors reached from a state y are compared to the errors from x.

Note that a state can also terminate without an exception by transitioning to the empty set. Due to the asymmetry in the distributive law for the coproduct, a state $X_0 \in \mathcal{P}(X)$ in the determinized automaton $c^\#$ will throw an exception as long as one of the elements $x \in X_0$ throws an exception ($c[X_0] \cap [0,1] \neq \emptyset$). In this case, $c^\#(X_0) = \sup(c[X_0] \cap [0,1])$ and X_0 performs a transition if all states in X_0 do so in the original coalgebra.

The behavioural distance on TX obtained as the greatest fixpoint (in the quantale order) of beh can be characterized as follows: for a set of states $X_0 \subseteq X$ and a word $w \in A^*$ let $ec(X_0, w)$ be the length of the least prefix which causes an exception when starting in X_0 (undefined if there are no exceptions) and let $E(X_0, w) \subseteq [0,1]$ be the set of exception values reached by that prefix. We define a distance $d_w^E\colon \mathcal{P}X \times \mathcal{P}X \to [0,1]$ as $d_w^E(X_1, X_2) = 0$ if $ec(X_2, w)$ is undefined, $d_w^E(X_1, X_2) = 1$ if $ec(X_1, w)$ is undefined and $ec(X_2, w)$ defined. In the case where both are defined we set $d_w^E(X_1, X_2) = \sup E(X_2, w) \ominus \sup E(X_1, w)$ if $ec(X_1, w) = ec(X_2, w)$, $d_w^E(X_1, X_2) = 1$ if $ec(X_1, w) > ec(X_2, w)$ and $d_w^E(X_1, X_2) = 0$ if $ec(X_1, w) < ec(X_2, w)$. Then it can be shown that for $X_1, X_2 \subseteq X$:

$$\nu\text{beh}(X_1, X_2) = \sup\nolimits_{w \in A^*} d_w^E(X_1, X_2)$$

As a concrete example we take the label set $A = \{a, b\}$ and the coalgebra c given below, which is inspired by a similar example in [7]:



Here, the outputs of the final states are $c(x_n) = 1/4$, $c(y_n) = 1/3$ and $c(z_n) = 1/2$, as indicated. It holds that $\nu\text{beh}(\{x_0, y_0\}, \{z_0\}) = 1/4$. Intuitively this is true, since the largest distance is achieved if we follow a path from x_0 where a is the n-last letter, yielding exception value $1/4$, while the same path results in the value $1/2$ from z_0, hence we obtain distance $1/2 \ominus 1/4 = 1/4$.

Note that the determinization of the transition system above is of exponential size and the same holds for a representation of a post-fixpoint for witnessing an upper bound for behavioural distance. We construct a $\mathcal{V}$-valued relation d of linear size witnessing that $\nu\text{beh}(\{x_0, y_0\}, \{z_0\}) \leq 1/4$ via up-to techniques. To this end let d be defined by

$$d(\{x_0, y_0\}, \{z_0\}) = 1/4 \qquad d(\{x_i\}, \{z_i\}) = 1/4 \qquad d(\{y_i\}, \{z_i\}) = 1/6$$

and distance 1 for all other arguments.

It suffices to show that d is a pre-fixed point of beh up-to u (wrt. $\leq$). We can use the property that $u(d)(X_1 \cup X_2, Y_1 \cup Y_2) \leq \max\{d(X_1, Y_1), d(X_2, Y_2)\}$ (see [10] for more details). Now the claim follows from unfolding the fixpoint equation by considering a- and b-transitions:

$$\begin{aligned}\mathrm{beh}(u(d))(\{x_0, y_0\}, \{z_0\}) &= \max\{u(d)(\{x_0, x_1, y_0\}, \{z_0, z_1\}), u(d)(\{x_0, y_0, y_1\}, \{z_0, z_1\})\} \\ &\leq \max\{d(\{x_0, y_0\}, \{z_0\}), d(\{x_1\}, \{z_1\}), d(\{y_1\}, \{z_1\})\} = 1/4\end{aligned}$$

The arguments for $d(\{x_i\}, \{z_i\})$, $d(\{y_i\}, \{z_i\})$ are similar, concluding the proof.

8 Conclusion

Related work. While the notion of Kantorovich distance on probability distributions is much older, Kantorovich liftings in a categorical framework have first been introduced in [2] and have since been generalized, for instance to codensity liftings [19] or to lifting fuzzy lax extensions [31].

A coalgebraic theory of up-to techniques was presented in [6] and has been instantiated to the setting of coalgebraic behavioural metrics in [5]. The latter paper concentrated on Wasserstein liftings, which leads to a significantly different underlying theory. Furthermore, Wasserstein liftings are somewhat restricted, since they rely only on a single evaluation map and on couplings (which sometimes do not exist, making it difficult to define more fine-grained metrics). We are not aware of a way to handle the two case studies directly in the Wasserstein approach.

Setting up the fibred adjunction (Section 4) and the definition of the Kantorovich lifting (Section 5.1) has some overlap to [4] and the recent [19]. Our focus is primarily on showing fibredness (naturality) via a quantalic version of the extension theorem.

The aim of [19] is on combining behavioural conformance via algebraic operations, which is different than our notion of compositionality via functor liftings. There is some similarity in the aims of both papers, namely the lifting of distributive laws and the motivation to study up-to techniques. From our point of view, the obtained results are largely orthogonal: while the focus of [19] is on providing n-ary operations for composing conformances and games and it provides a more general high-level account, we focus concretely on compositionality of functor liftings (studying counterexamples and treating the special case of polynomial functors), giving concrete recipes for lifting distributive laws and applying the results to proving upper bounds via up-to techniques.

Our up-to techniques are a form of up-to convex contextual closure: here they arise as specific instances of a general construction, but they have already been investigated in the Wasserstein setting [5] and earlier in [9]. Similar constructions are studied in [21].

Future work. Our aim is to better understand the metric congruence results employed in the case studies, comparing them with the similar proof rules in [21]. Compositionality of functor liftings fails in important cases (cf. Example 12), which could be fixed by using different sets of evaluation maps such as the Moss liftings in [31]. We also plan to study case studies involving the convex powerset functor [8]. Finally, we want to develop witness generation methods, by constructing suitable post-fixpoints up-to on-the-fly, similar to [1].

References

1 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. On-the-fly exact computation of bisimilarity distances. *Logical Methods in Computer Science*, 13(2:13):1–25, 2017. doi:10.23638/LMCS-13(2:13)2017.

2 Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Coalgebraic behavioral metrics. *Logical Methods in Computer Science*, 14(3), 2018. Selected Papers of the 6th Conference on Algebra and Coalgebra in Computer Science (CALCO 2015). doi:10.23638/LMCS-14(3:20)2018.

3 Harsh Beohar, Sebastian Gurke, Barbara König, and Karla Messing. Hennessy-Milner theorems via Galois connections. In *Proc. of CSL '23*, volume 252 of *LIPIcs*, pages 12:1–12:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.CSL.2023.12.

4 Harsh Beohar, Sebastian Gurke, Barbara König, Karla Messing, Jonas Forster, Lutz Schröder, and Paul Wild. Expressive quantale-valued logics for coalgebras: an adjunction-based approach. In *Proc. of STACS '24*, volume 289 of *LIPIcs*, pages 10:1–10:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024. doi:10.4230/LIPIcs.STACS.2024.10.

5 Filippo Bonchi, Barbara König, and Daniela Petrişan. Up-to techniques for behavioural metrics via fibrations. *Mathematical Structures in Computer Science*, 33(4–5):182–221, 2023. doi:10.1017/s0960129523000166.

6 Filippo Bonchi, Daniela Petrisan, Damien Pous, and Jurriaan Rot. A general account of coinduction up-to. *Acta Informatica*, 54(2):127–190, 2017. doi:10.1007/s00236-016-0271-4.

7 Filippo Bonchi and Damien Pous. Checking NFA equivalence with bisimulations up to congruence. In *Proc. of POPL '13*, pages 457–468. ACM, 2013. doi:10.1145/2429069.2429124.

8 Filippo Bonchi, Alexandra Silva, and Ana Sokolova. The power of convex algebras. In *Proc. of CONCUR '17*, volume 85 of *LIPIcs*, pages 23:1–23:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.CONCUR.2017.23.

9 Konstantinos Chatzikokolakis, Daniel Gebler, Catuscia Palamidessi, and Lili Xu. Generalized bisimulation metrics. In *Proc. of CONCUR '14*. Springer, 2014. LNCS/ARCoSS 8704. doi:10.1007/978-3-662-44584-6_4.

10 Keri D'Angelo, Sebastian Gurke, Johanna Maria Kirss, Barbara König, Matina Najafi, Wojciech Różowski, and Paul Wild. Behavioural metrics: Compositionality of the Kantorovich lifting and an application to up-to techniques, 2024. arXiv:2404.19632. doi:10.48550/arXiv.2404.19632.

11 Luca de Alfaro, Marco Faella, and Mariëlle Stoelinga. Linear and branching system metrics. *IEEE Transactions on Software Engineering*, 35(2):258–273, 2009. doi:10.1109/TSE.2008.106.

12 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318:323–354, 2004. doi:10.1016/j.tcs.2003.09.013.

13 Dirk Hofmann. Topological theories and closed objects. *Adv. Math.*, 215(2):789–824, 2007. doi:10.1016/j.aim.2007.04.013.

14 Bart Jacobs. *Categorical Logic and Type Theory*, volume 141 of *Studies in Logic and the Foundation of Mathematics*. Elsevier, 1999.

15 Bart Jacobs. A bialgebraic review of deterministic automata, regular expressions and languages. In *Essays Dedicated to Joseph A. Goguen*, pages 375–404. Springer, 2006. LNCS 4060. doi:10.1007/11780274_20.

16 Bart Jacobs. *Introduction to Coalgebra. Towards Mathematics of States and Observations*. Cambridge University Press, December 2016. doi:10.1017/CBO9781316823187.

17 Bart Jacobs, Alexandra Silva, and Ana Sokolova. Trace semantics via determinization. *Journal of Computer and System Sciences*, 81(5):859–879, 2015. doi:10.1016/j.jcss.2014.12.005.

18 Shin-ya Katsumata and Tetsuya Sato. Codensity liftings of monads. In *Proc. of CALCO '15*, volume 35 of *LIPIcs*, pages 156–170. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015. doi:10.4230/LIPIcs.CALCO.2015.156.

19 Mayuko Kori, Kazuki Watanabe, Jurriaan Rot, and Shin ya Katsumata. Composing codensity bisimulations. In *Proc. of LICS '24*, pages 52:1–52:13. ACM, 2024. doi:10.1145/3661814.3662139.

20 Francis William Lawvere. Metric spaces, generalized logic, and closed categories. *Rendiconti del seminario matématico e fisico di Milano*, 43(1):135–166, 1973. doi:10.1007/BF02924844.

21 Radu Mardare, Prakash Panangaden, and Gordon Plotkin. Quantitative algebraic reasoning. In *Proc. of LICS '16*, pages 700–709. ACM, 2016. doi:10.1145/2933575.2934518.

22 E. J. McShane. Extension of range of functions. *Bull. Amer. Math. Soc.*, 40(12):837–842, 1934.

23 Damien Pous. Complete lattices and up-to techniques. In *Proc. of APLAS '07*, pages 351–366. Springer, 2007. LNCS 4807. doi:10.1007/978-3-540-76637-7_24.

24 J.J.M.M. Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249:3–80, 2000. doi:10.1016/S0304-3975(00)00056-6.

25 Davide Sangiorgi. *Introduction to Bisimulation and Coinduction*. Cambridge University Press, 2011.

26 Alexandra Silva, Filippo Bonchi, Marcello M. Bonsangue, and Jan J. M. M. Rutten. Generalizing determinization from automata to coalgebras. *Logical Methods in Computer Science*, 9(1:09), 2013. doi:10.2168/LMCS-9(1:9)2013.

27 Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5:285–309, 1955.

28 Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331:115–142, 2005. doi:10.1016/j.tcs.2004.09.035.

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

30 Hassler Whitney. Analytic extensions of differentiable functions defined in closed sets. *Transactions of the American Mathematical Society*, 36(1):63–89, 1934.

31 Paul Wild and Lutz Schröder. Characteristic logics for behavioural metrics via fuzzy lax extensions. In *Proc. of CONCUR '20*, volume 171 of *LIPIcs*, pages 27:1–27:23. Schloss Dagstuhl, 2020. doi:10.4230/LIPIcs.CONCUR.2020.27.

32 Paul Wild and Lutz Schröder. Characteristic logics for behavioural hemimetrics via fuzzy lax extensions. *Log. Methods Comput. Sci.*, 18(2), 2022. doi:10.46298/lmcs-18(2:19)2022.