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
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Abstract

Behavioural distances of transition systems modelled as coalgebras for endofunctors generalize the 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 Kantorovich lifting of a functor on 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. These results are essential ingredients for adopting 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 conclude by illustrating the results of our paper in two case studies.

Funding *Sebastian Gurke*: as well as 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)

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 [22]. More recently, this idea has been generalized to behavioural metrics that measure the behavioural distance of two states [25, 8, 9]. 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 [21] 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 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. 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. 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 [2] – has the advantage of 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 [2]. Then 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 [4], 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 [3]. 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, and show several negative results for combinations of powerset and distribution functor.

The positive compositionality result for the case where F is polynomial 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 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 [14, 23] 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. On the other hand, 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), making it hard to compute conformances for $c^{\#}$.

However, the algebraic structure on TX can be fruitfully employed using up-to techniques [20, 5] 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 [4] that provide a similar result for the Wasserstein lifting. This enables us to use the up-to technique for Kantorovich liftings that allow more flexibility. The reason for this extra flexibility is that they are based on a set of evaluation maps, rather than a single evaluation map in the Wasserstein case, which leads to less fine-grained distances, in particular in the presence of products and coproducts. Indeed, as can be seen later in the paper, several 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), while [4] 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 [4].

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 $[0, \infty]_{\oplus}$ of *extended positive reals*, with structure analogous to $[0, 1]_{\oplus}$, i.e. reversed lattice order and using truncated addition as the monoid.

▶ **Remark 1.** As many of our examples use the real-valued quantales $[0, 1]_{\oplus}$ and $[0, \infty]_{\oplus}$, 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 $[0, \infty]_{\oplus}$.

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*, 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 $\mathcal{V}$-Graph is denoted by $\mathcal{V}$-Rel instead [4]. We opt for $\mathcal{V}$-Graph as in [17], as $\mathcal{V}$-Rel more often denotes the category with sets as objects and $\mathcal{V}$-valued relations between them as morphisms [10].

Second, within $\mathcal{V}$-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 $\mathcal{V}$ [17]. A *$\mathcal{V}$-category* is an object of $\mathcal{V}\text{-Graph}_X$ 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 $\mathcal{V}$-Cat, and the set of $\mathcal{V}$-categories based on a set X is denoted by $\mathcal{V}\text{-Cat}_X$.

In this paper, we use a coalgebraic framework. Recall 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 $\mathcal{D}(X) = \big\{p : X \to [0,1] \mid \sum_{x\in X} p(x) = 1, \mathsf{supp}(p) \text{ is countable}\big\}$ 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 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 [3]. 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 [3], 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 $\mathcal{V}\text{-SPred}_X$ 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 [17, 4]. 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* [11] 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 is 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 a pointwise order $\sqsubseteq$. Given a function $f\colon X \to Y$, we define $\Phi(f) = f^*$ by reindexing, where $f^*(d_Y) = d_Y \circ (f \times f)$.

Similarly, to obtain $\mathcal{V}$-SPred, we define a functor $\Psi : \mathsf{Set}^{\mathrm{op}} \to \mathsf{Pos}$ that maps each set X to $\Psi(X) = (\mathcal{V}\text{-}\mathsf{SPred}_X, \supseteq)$, 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 [3], 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 $\alpha_X\colon \mathcal{V}\text{-}\mathsf{SPred}_X \to \mathcal{V}\text{-}\mathsf{Graph}_X$ given by:

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

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 $\gamma_X\colon \mathcal{V}\text{-}\mathsf{Graph}_X \to \mathcal{V}\text{-}\mathsf{SPred}_X$ defined by the following:

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

Given a conformance $d_X\colon X \times X \to \mathcal{V}$, 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 [3, Theorem 7] and obtain the following:

▶ **Theorem 3** ([3]). *Let X be an arbitrary set, $d_X\colon X \times X \to \mathcal{V}$ 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:*

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

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 morphisms of split indexed categories) [11, 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 [19, 27], mentioned also in [17] for the real-valued case.

▶ **Theorem 4.** *Let $d_X \in \mathcal{V}\text{-}\mathsf{Cat}_X$ and $d_Y \in \mathcal{V}\text{-}\mathsf{Cat}_Y$ be elements of $\mathcal{V}$-Cat. If $i\colon (Y, d_Y) \to (X, d_X)$ is an isometry, then for any non-expansive map $f\colon (Y, d_Y) \to (\mathcal{V}, d_{\mathcal{V}})$ there exists a non-expansive map $g\colon (X, d_X) \to (\mathcal{V}, d_{\mathcal{V}})$ such that $f = g \circ i$.*

▶ **Proposition 5.** *We have that $\alpha\colon \Psi \Rightarrow \Phi$ is natural and $\gamma\colon \Phi \Rightarrow \Psi$ is laxly natural, that is for all functions $f\colon X \to Y$ 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, $\gamma\colon \Phi \Rightarrow \Psi$ 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 6.** *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 ([11, 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. [16]).

▶ **Theorem 7.** *There is an adjunction $\gamma \dashv \alpha$.*

4 The Kantorovich Lifting

4.1 Definition of Kantorovich Lifting

Motivated by the transportation problem [26], the Kantorovich distance on probability distributions aims to maximize the relative logistics profit 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.

The coalgebraic Kantorovich lifting [2] – also known as codensity lifting [15] – is parametric in a set of *evaluation functions* for a set functor F. Evaluation functions are 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 $s, t \in FX$:

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

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

▶ **Lemma 8.** *The Kantorovich lifting of a functor $F : \mathsf{Set} \to \mathsf{Set}$ is a functor $\overline{F} = K_{\Lambda^F} : \mathcal{V}\text{-Graph} \to \mathcal{V}\text{-Graph}$, and fibred when restricted to $\mathcal{V}$-Cat.*

4.2 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 9** (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 [3], 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 $\Lambda^F \circ \gamma \circ \alpha \subseteq \gamma F \circ \alpha F \circ \Lambda^F$:

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

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 $\Lambda^F \circ c \subseteq c \circ \Lambda^F$ holds, we consider a few examples where compositionality fails:

▶ **Example 10.** 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 the appendix (Example 34). 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 $K_{\{\sup * \mathbb{E}\}}d(U,V) = 0$.

4.3 Polynomial Functors

We now assume that the first functor (F with the lifting $\overline{F} = K_{\Lambda_F}$) is in fact a polynomial functor 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 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 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: $F = \prod_{i \in I} F_i$. Let $\Lambda^F = \{ev_i \circ \pi'_i \mid i \in I, ev_i \in \Lambda^{F_i}\}$ 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 11.** 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 $i \in I$.

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

▶ **Proposition 12.** *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 4.2):

▶ **Proposition 13.** *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.*

From this we can infer that $K_{\Lambda^F} \circ K_{\Lambda^G} = K_{\Lambda^F * \Lambda^G}$ using the arguments of Section 4.2, hence we have compositionality.

5 Application: Up-To Techniques

We now adapt results from [4] on up-to techniques from Wasserstein to Kantorovich liftings. In particular we instantiate the fibrational approach to coinductive proof techniques from [5] that allows to prove lower bounds for greatest fixpoints, using post-fixpoints up-to as witnesses. As shown in Section 6 this can greatly help to reduce the size of such witnesses, even allowing finitary witnesses which would be infinite otherwise.

5.1 Introduction to Up-To Techniques

We first recall the notion of a bialgebra [12], i.e., a coalgebra with a compatible algebra structure.

▶ **Definition 14.** Consider two functors F, T and a natural transformation $\zeta\colon TF \Rightarrow FT$. A *bialgebra* for ζ is a tuple (Y, a, c) such that $a\colon TY \to Y$ is a T-algebra, $c\colon Y \to FY$ is an F-coalgebra so that the diagram below commutes.

$$\begin{tikzcd} TY \arrow[r, "a"] \arrow[d, "Tc"] & Y \arrow[r, "c"] & FY \\ TFY \arrow[rr, "\zeta_Y"] & & FTY \arrow[u, "Fa"'] \end{tikzcd}$$

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

▶ **Definition 15.** 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:

$$\begin{tikzcd} FX \arrow[d, "\eta_{FX}"'] \arrow[dr, "F\eta_X"] & \\ TFX \arrow[r, "\zeta_X"'] & FTX \end{tikzcd} \qquad \begin{tikzcd} T^2FX \arrow[r, "T\zeta_X"] \arrow[d, "\mu_{FX}"'] & TFTX \arrow[r, "\zeta_{TX}"] & FT^2X \arrow[d, "F\mu_X"] \\ TFX \arrow[rr, "\zeta_X"'] & & FTX \end{tikzcd}$$

Whenever T is a monad and ζ is an EM-law, then a bialgebra can be obtained by determinizing a coalgebra $c\colon X \to FTX$. More concretely, we obtain $c^\#\colon Y \to FY$ where $Y = TX$ and $c^\# = F\mu_X \circ \zeta_{TX} \circ Tc$. The algebra map is $a = \mu_X\colon TY \to Y$. Examples involving determinization will be given in Section 6.

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

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

The greatest fixpoint of beh corresponds to a behavioural conformance (e.g., behavioural equivalence or bisimulation metric). Furthermore we can define an *up-to function* u 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 [24] and the theory of up-to techniques [20]. In particular, given a monotone function $f\colon L \to L$ over a complete lattice $(L, \sqsubseteq)$, we know that its greatest fixpoint νf coincides with its greatest post-fixpoint, where a post-fixpoint is an element $\ell \in L$ with $\ell \sqsubseteq f(\ell)$. Furthermore, given a monotone function $u\colon L \to L$ that is f-compatible (i.e., $u \circ f \sqsubseteq f \circ u$), we can infer that $\ell \sqsubseteq f(u(\ell))$ (i.e., ℓ is a post-fixpoint up-to u) implies $\ell \sqsubseteq \nu f$.

From [4] we obtain the following result that ensures compatibility:

▶ **Proposition 16** ([4]). *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).*

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

To use this proof technique, we have to show that ζ lifts accordingly to obtain an up-to technique usable for Kantorovich liftings. We start by defining distributive laws and lifting them to $\mathcal{V}$-Graph.

5.2 Lifting Distributive Laws

Let F be a polynomial functor (cf. Section 4.3) and (T, μ, η) a monad over Set. Following [13, 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 [13] 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 4.3 and that $\Lambda^T = \{ev_T\}$, where $ev_T \colon T\mathcal{V} \to \mathcal{V}$ is a T-algebra.

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

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

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

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

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

▶ **Definition 17.** 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.

$$\begin{tikzcd} & Y_1 + Y_2 \arrow[dl, "\eta_{Y_1+Y_2}"'] \arrow[dr, "\eta_{Y_1}+\eta_{Y_2}"] & \\ T(Y_1+Y_2) \arrow[rr, "g_{Y_1,Y_2}"] & & TY_1 + TY_2 \end{tikzcd}
\qquad
\begin{tikzcd} TT(Y_1+Y_2) \arrow[rr, "\mu_{Y_1+Y_2}"] \arrow[d, "Tg_{Y_1,Y_2}"'] & & T(Y_1+Y_2) \arrow[d, "g_{Y_1,Y_2}"] \\ T(TY_1+TY_2) \arrow[r, "g_{TY_1,TY_2}"'] & TTY_1 + TTY_2 \arrow[r, "\mu_{Y_1}+\mu_{Y_2}"'] & TY_1 + TY_2 \end{tikzcd}$$

▶ **Proposition 18.** *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 19.** *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 20.** 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 $f_i : X_i \to \mathcal{V}$, where $\bot, \top$ are constant maps of the appropriate type.

$$\begin{tikzcd} T(X_1+X_2) \arrow[r, "{T[f_1,\top_{X_2}]}"] \arrow[d, "g_{X_1,X_2}"'] & T\mathcal{V} \arrow[d, "ev_T"] \\ TX_1+TX_2 \arrow[r, "{[ev_T\circ Tf_1,\top_{TX_2}]}"'] & \mathcal{V} \end{tikzcd} \qquad \begin{tikzcd} T(X_1+X_2) \arrow[r, "{T[\bot_{X_1},f_2]}"] \arrow[d, "g_{X_1,X_2}"'] & T\mathcal{V} \arrow[d, "ev_T"] \\ TX_1+TX_2 \arrow[r, "{[\bot_{TX_1},ev_T\circ Tf_2]}"'] & \mathcal{V} \end{tikzcd} \qquad \begin{tikzcd} T(X_1+X_2) \arrow[r, "{T[\bot_{X_1},\top_{X_2}]}"] \arrow[d, "g_{X_1,X_2}"'] & T\mathcal{V} \arrow[d, "ev_T"] \\ TX_1+TX_2 \arrow[r, "{[\bot_{TX_1},\top_{TX_2}]}"'] & \mathcal{V} \end{tikzcd}$$

▶ **Lemma 21.** *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 defined in Section 4.3, 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 19 and Lemma 21 we know that the distributive law ζ lifts to $\mathcal{V}$-Graph, i.e., $\zeta : \overline{FT} \Rightarrow \overline{TF}$ 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 22.** *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.*

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

▶ **Proposition 23.** *Let $T = \mathcal{S}$ be the subdistribution monad where $\mathcal{S}(X) = \{p : X \to [0,1] \mid \sum_{x\in X} p(x) \leq 1\}$. Assume that its evaluation map is $ev_T = \mathbb{E}$ for quantale $\mathcal{V} = [0,\infty]_\oplus$ (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.*

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

It is left to show that the EM-law $\zeta : FT \Rightarrow FT$ lifts to $\zeta : \overline{F}\,\overline{T} \Rightarrow \overline{T}\,\overline{F}$, where $\overline{F} = K_{\Lambda^F}$, $\overline{T} = K_{\Lambda^T}$ where $\Lambda^T = \{ev_T\}$, where the evaluation maps 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 4.2. This implies that the identity $\mathrm{id}_{TFX} : \overline{T}\,\overline{F}X \to \overline{TF}X$ 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 21, Proposition 19)
(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 13)

6 Case Studies

We now consider two case studies and show how up-to techniques can help concretely to show upper bounds for behavioural distances via appropriate witnesses. Note that both case studies rely on a set of evaluation maps Λ^F as opposed to a singleton evaluation map and would hence not be directly realizable in the Wasserstein approach [4].

6.1 Probabilistic Automata

We consider a standard directed trace metrics for probabilistic automata as introduced in [14]. We take the 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 Section 4.3.

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 in Section 5.1.

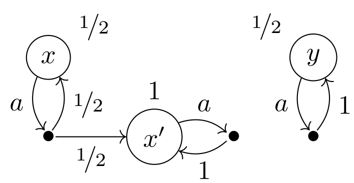
By unravelling the fixpoint equation one can see that it coincides with the directed trace metric on probability distributions that is defined as follows: for each state $x \in X$ let $tr_x\colon A^* \to [0,1]$ be a map that assigns to each word (trace) $w \in A^*$ the expected payoff for this word when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

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

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

The proof techniques elaborated in Section 5.1 apply and we explain how they can be exploited via an example.

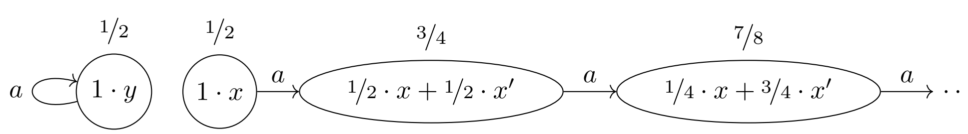
▶ **Example 24.** For this concrete example we fix the label set as a singleton: $A = \{a\}$. Consider the coalgebra with states $X = \{x, x', y\}$ drawn below on the left. The payoff is written next to each state.



	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 $\nu\mathrm{beh}(\delta_x, \delta_y) = 1/2$ (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 ($d \geq \mathrm{beh}(u(d))$). From this we can infer that $\nu\mathrm{beh} \leq d$, 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 $d(p,q) \geq \mathrm{beh}(u(d))(p,q)$. 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:

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

- 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, Corollary 36 in the appendix). 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.

6.2 Transition Systems with Exceptions

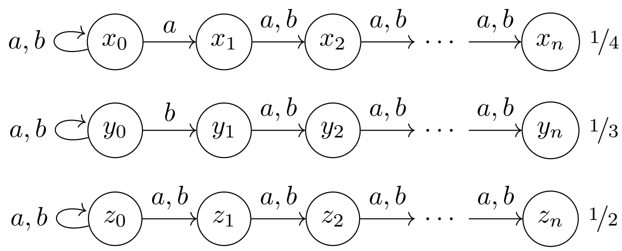
Next, we consider a case study involving the coproduct, in particular the polynomial functor $F = [0,1] + (-)^A$ and the monad $T = \mathcal{P}$ with evaluation map sup. In a coalgebra of type $c\colon X \to FTX$ a state can either perform transitions or terminate with some output value taken from the 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 it 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 $X_1, X_2 \subseteq X$:

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

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 [6]:



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 $\nu\mathrm{beh}(\{x_0, y_0\}, \{z_0\}) = 1/4$. Intuitively this is true, since the largest distance is achieved if we follow a path from x_0 with a a 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 $\nu\mathrm{beh}(\{x_0, y_0\}, \{z_0\}) \le 1/4$ 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. $\le$). We can use the property that $u(d)(X_1 \cup X_2, Y_1 \cup Y_2) \le \max\{d(X_1, Y_1), d(X_2, Y_2)\}$ (cf. Corollary 36 in the appendix). Now the claim follows from unfolding the fixpoint equation by considering a- and b-transitions:

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

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

7 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 [16] or to lifting fuzzy lax extensions [28].

A coalgebraic theory of up-to techniques was presented in [5] and has been instantiated to the setting of coalgebraic behavioural metrics in [4]. 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 metric). We are not aware of a way to handle the examples of Section 6 directly in the Wasserstein approach.

Setting up the fibred adjunction (Section 3) and the definition of the Kantorovich lifting (Section 4.1) has some overlap to [3] and the very recent [16]. Our focus is primarily on showing fibredness (naturality) via a quantalic version of the extension theorem.

The aim of [16] 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 [16] 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.

Future work. Our aim is to better understand the metric congruence results employed in Section 6, comparing them with the similar proof rules in [18]. Compositionality of functor liftings fails in important cases (cf. Example 10), which could be fixed by using different sets of evaluation maps as in the Moss liftings in [28]. We also plan to study case studies involving the convex powerset functor [7]. Finally, we want to develop witness generation methods, by constructing suitable post-fixpoint up-to on-the-fly, similar to [1].

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

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.
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).
3. 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 Center for Informatics, 2024.
4. Filippo Bonchi, Barbara König, and Daniela Petrişan. Up-to techniques for behavioural metrics via fibrations. *Mathematical Structures in Computer Science*, 2023. to appear.
5. Filippo Bonchi, Daniela Petrisan, Damien Pous, and Jurriaan Rot. A general account of coinduction up-to. *Acta Informatica*, 54(2):127–190, 2017.
6. Filippo Bonchi and Damien Pous. Checking NFA equivalence with bisimulations up to congruence. In *Proc. of POPL '13*, pages 457–468. ACM, 2013.
7. 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 fuer Informatik, 2017.
8. Luca de Alfaro, Marco Faella, and Mariëlle Stoelinga. Linear and branching system metrics. *IEEE Transactions on Software Engineering*, 35(2):258–273, 2009.
9. Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318:323–354, 2004.
10. Dirk Hofmann. Topological theories and closed objects. *Adv. Math.*, 215(2):789 – 824, 2007.
11. Bart Jacobs. *Categorical Logic and Type Theory*, volume 141 of *Studies in Logic and the Foundation of Mathematics*. Elsevier, 1999.
12. 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.
13. Bart Jacobs. *Introduction to Coalgebra. Towards Mathematics of States and Observations*. Cambridge University Press, December 2016.
14. Bart Jacobs, Alexandra Silva, and Ana Sokolova. Trace semantics via determinization. *Journal of Computer and System Sciences*, 81(5):859–879, 2015.

15 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 Center for Informatics, 2015.

16 Mayuko Kori, Kazuki Watanabe, Jurriaan Rot, and Shin ya Katsumata. Composing codensity bisimulations, 2024. arXiv:2404.08308.

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

18 Radu Mardare, Prakash Panangaden, and Gordon Plotkin. Quantitative algebraic reasoning. In *Proc. of LICS '16*, pages 700–709. ACM, 2016.

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

20 Damien Pous. Complete lattices and up-to techniques. In *Proc. of APLAS '07*, pages 351–366. Springer, 2007. LNCS 4807.

21 J.J.M.M. Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249:3–80, 2000.

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

23 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.

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

25 Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331:115–142, 2005.

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

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

28 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.

A Proofs for Section 2 (Preliminaries)

▶ **Lemma 25.** *Residuation (internal hom) in a quantale satisfies the following properties for $u, v, w, a, b, c, a_i, b_i \in \mathcal{V}$:*

1. $d_\mathcal{V}(v, w) = \bigsqcup\{u \in \mathcal{V} \mid u \otimes v \sqsubseteq w\}$
2. $d_\mathcal{V}(v, v) \sqsupseteq k$ *(reflexivity).*
3. $d_\mathcal{V}(u, v) \otimes d_\mathcal{V}(v, w) \sqsubseteq d_\mathcal{V}(u, w)$ *(triangle inequality).*
4. $d_\mathcal{V}(k, w) = w$.
5. $d_\mathcal{V}(\bot, w) = \top$.
6. $d_\mathcal{V}(v, \top) = \top$.
7. $d_\mathcal{V}(\top, \bot) = \bot$.
8. $d_\mathcal{V}(a, c) \sqsubseteq d_\mathcal{V}(d_\mathcal{V}(b, a), d_\mathcal{V}(b, c))$
9. $d_\mathcal{V}(a, c) \sqsubseteq d_\mathcal{V}(d_\mathcal{V}(c, b), d_\mathcal{V}(a, b))$
10. $\bigsqcap_{i \in I} d_\mathcal{V}(a_i, b_i) \sqsubseteq d_\mathcal{V}(\bigsqcap_{i \in I} a_i, \bigsqcap_{i \in I} b_i)$

Proof.

1. Residuation is the right adjoint of the associative operation $\otimes$.
2. $d_\mathcal{V}(v, v) = \bigsqcup\{u \in \mathcal{V} \mid u \otimes v \sqsubseteq v\} \sqsupseteq k$, since $k \otimes v = v \sqsubseteq v$.
3. From $d_\mathcal{V}(v, w) \sqsubseteq d_\mathcal{V}(v, w)$ and the adjunction property, we obtain $v \oplus d_\mathcal{V}(v, w) \sqsubseteq w$ and analogously $u \oplus d_\mathcal{V}(u, v) \sqsubseteq v$. Hence $u \otimes d_\mathcal{V}(u, v) \otimes d_\mathcal{V}(v, w) \sqsubseteq v \otimes \otimes d_\mathcal{V}(v, w) \sqsubseteq w$. This implies $d_\mathcal{V}(u, v) \otimes d_\mathcal{V}(v, w) \sqsubseteq d_\mathcal{V}(u, w)$.
4. $d_\mathcal{V}(k, w) = \bigsqcup\{u \in \mathcal{V} \mid u \otimes k \sqsubseteq w\} = \bigsqcup\{u \in \mathcal{V} \mid u \sqsubseteq w\} = w$.
5. $d_\mathcal{V}(\bot, w) = \bigsqcup\{u \in \mathcal{V} \mid u \otimes \bot \sqsubseteq w\} = \bigsqcup\{u \in \mathcal{V} \mid \bot \sqsubseteq w\} = \bigsqcup \mathcal{V} = \top$. This holds, since $\bot$ is the empty join and $\otimes$ preserves joins, hence $u \otimes \bot = \bot$.
6. $d_\mathcal{V}(v, \top) = \bigsqcup\{u \in \mathcal{V} \mid u \otimes v \sqsubseteq \top\} \bigsqcup \mathcal{V} = \top$.
7. $d_\mathcal{V}(\top, \bot) = \bigsqcup\{u \in \mathcal{V} \mid u \otimes \top \sqsubseteq \bot\} \sqsubseteq \bigsqcup\{u \in \mathcal{V} \mid u \otimes k \sqsubseteq \bot\} = \{u \in \mathcal{V} \mid u \sqsubseteq \bot\} = \bot$, where the inequality holds since $k \sqsubseteq \top$.
8. Recall that the because of the definition of residuation, we have the following property:

   $$d_\mathcal{V}(a, c) \sqsubseteq d_\mathcal{V}\left(d_\mathcal{V}(b, a), d_\mathcal{V}(b, c)\right) \iff d_\mathcal{V}(a, c) \otimes d_\mathcal{V}(b, a) \sqsubseteq d_\mathcal{V}(b, c)$$

   Hence, it suffices to show the second inequality:

   $$\begin{aligned} d_\mathcal{V}(a, c) \otimes d_\mathcal{V}(b, a) &= d_\mathcal{V}(b, a) \otimes d_\mathcal{V}(a, c) \\ &\sqsubseteq d_\mathcal{V}(b, c) && \text{(Triangle inequality, Lemma 25(3))} \end{aligned}$$

9. Symmetric to (8)
10. Because of the definition of the residuation, for all $i \in I$, we have that:

    $$d_\mathcal{V}(a_i, b_i) \sqsubseteq d_\mathcal{V}(a_i, b_i) \iff d_\mathcal{V}(a_i, b_i) \otimes a_i \sqsubseteq b_i$$

    Therefore, we have that:

    $$\begin{aligned} \bigsqcap_{i \in I} b_i &\sqsupseteq \bigsqcap_{i \in I} d_\mathcal{V}(a_i, b_i) \otimes a_i \\ &\sqsupseteq \bigsqcap_{i \in I} d_\mathcal{V}(a_i, b_i) \otimes \bigsqcap_{i \in I} a_i \end{aligned}$$

    Again, by the definition of residuation the above is equivalent to $\bigsqcap_{i \in I} d_\mathcal{V}(a_i, b_i) \sqsubseteq d_\mathcal{V}(\bigsqcap_{i \in I} a_i, \bigsqcap_{i \in I} b_i)$ as desired, which completes the proof.

◀

B Proofs for Section 3 (Setting Up a Fibred Adjunction)

▶ **Lemma 26.** $\mathcal{V}$-Graph $= \int \Phi$ *and* $\mathcal{V}$-SPred $= \int \Psi$.

Proof. The objects of $\int \Phi$ are pairs (X, d_X), where X is a set and d_X is an element of the lattice $\mathcal{V}\text{-Graph}_X$ of $\mathcal{V}$-valued relations. The arrows $f\colon (X, d_X) \to (Y, d_Y)$ are functions $f\colon X \to Y$, which satisfy $d_X \sqsubseteq f^*(d_Y)$, which is the same as $d_X \sqsubseteq d_Y \circ (f \times f)$. This precisely corresponds to the definition of the $\mathcal{V}$-Graph category.

Similarly, the objects of $\int \Psi$ are pairs (X, S), where X is a set and $S \in \mathcal{V}\text{-SPred}_X$ (and hence $S \subseteq \mathcal{V}^X$) is the element of the lattice of subsets of $\mathcal{V}^X$ ordered by reverse inclusion. The arrows $(X, S) \to (Y, T)$ are functions $f\colon X \to Y$, such that $S \supseteq f^\bullet(T)$, which is the same as $\{q \circ f \mid q \in T\} \subseteq S$. Again this precisely corresponds to the definition of $\mathcal{V}$-SPred category. ◀

Additionally, observe that if we were to restrict Φ to send each set X to the lattice $(\mathcal{V}\text{-Cat}_X, \sqsubseteq)$ of $\mathcal{V}$-categories on X, which is a sublattice of $(\mathcal{V}\text{-Graph}_X, \sqsubseteq)$ of all $\mathcal{V}$-graphs, then the Grothendieck completion would yield $\mathcal{V}$-Cat instead of $\mathcal{V}$-Graph.

▶ **Lemma 27.** $\alpha\colon \Psi \Rightarrow \Phi$ *is natural.*

Proof. Let $f\colon X \to Y$ be an arbitrary function. We need to check the commutativity of the following diagram in Pos:

$$\begin{tikzcd}
\Psi(Y) \arrow[r, "f^\bullet"] \arrow[d, "\alpha_Y"] & \Psi(X) \arrow[d, "\alpha_X"] \\
\Phi(Y) \arrow[r, "f^*"] & \Phi(X)
\end{tikzcd}$$

Note that as a consequence of Theorem 3, we have that for all sets X and all $S, S' \in \mathcal{V}\text{-SPred}_X$, we have antitonicity of α_X and hence α_X is a monotone map from the lattice of sets of predicates with inverse inclusion order to the lattice of $\mathcal{V}$-valued relations. It suffices to check $\alpha_X \circ f^\bullet = f^* \circ \alpha_Y$. Let $T \subseteq \mathcal{V}^Y$ and $x_1, x_2 \in X$ and consider the following:

$$\begin{aligned}
(\alpha_X \circ f^\bullet)(T)(x_1, x_2) &= \bigsqcap_{p \in T} d_{\mathcal{V}}((p \circ f)(x_1), (p \circ f)(x_2)) \\
&= \alpha_Y(T)(f(x_1), f(x_2)) \\
&= (f^* \circ \alpha_Y(T))(x_1, x_2)
\end{aligned}$$

◀

Unfortunately, in the case of γ, without any extra assumptions, we can only prove a weaker statement, namely that it is laxly natural.

▶ **Lemma 28.** $\gamma\colon \Phi \Rightarrow \Psi$ *is laxly natural, that is for all functions $f\colon X \to Y$ and all $\mathcal{V}$-valued relations d_Y on the set Y we have that $(f^\bullet \circ \gamma_Y)(d_Y) \subseteq (\gamma_X \circ f^*)(d_Y)$.*

Proof. Observing that γ_X is antitone, we can show that for all sets X, γ_X is a monotone function from the lattice of $\mathcal{V}$-valued relations to the lattice of collections of $\mathcal{V}$-valued predicates (ordered by reverse inclusion order). The only thing we need to show is that for all functions $f\colon X \to Y$, we have that $(f^\bullet \circ \gamma_Y)(d_Y) \subseteq (\gamma_X \circ f^*)(d_Y)$ for all $\mathcal{V}$-valued relations d_Y on the set Y.

Assume that $p\colon X \to \mathcal{V}$ belongs to $(f^\bullet \circ \gamma_Y)(d_Y)$. In other words, $p = q \circ f$ for some $q\colon Y \to \mathcal{V}$, such that $d_{\mathcal{V}} \circ (q \times q) \sqsupseteq d_Y$. Observe that pre-composing $f \times f$ yields the inequality

$d_{\mathcal{V}} \circ ((q \circ f) \times (q \circ f)) \sqsupseteq d_Y \circ (f \times f)$ and therefore $d_{\mathcal{V}} \circ (p \times p) \sqsupseteq d_Y \circ (f \times f) = f^*(d_Y)$. Hence, $p \in (\gamma_X \circ f^*)(d_X)$, which completes the proof. ◀

To obtain the converse inclusion, we need to be able to factorise any $\mathcal{V}$-valued predicate on X, which is non-expansive with respect to $f^*(d_Y)$ as the composition of f and a $\mathcal{V}$-valued predicate on Y which is non-expansive with respect to d_Y. Since f can be seen as isometry between $(X, f^*(d_Y))$ and (Y, d_Y), we are essentially focusing on extending a non-expansive map into a residuation metric through an isometry, which in the case of pseudometric spaces can be proved through the McShane-Whitney extension Theorem. It turns out that those results can be generalised to the setting of $\mathcal{V}$-categories.

▶ **Theorem 29** (Quantalic McShane-Whitney Extension Theorem). *Let $\mathcal{V}$ be a quantale and (X, d) be an object of $\mathcal{V}$-*Cat*. For any subset $Y \subseteq X$ and any non-expansive map $f \colon (Y, d) \to (\mathcal{V}, d_{\mathcal{V}})$ there is a non-expansive extension $\bar{f} \colon (X, d) \to (\mathcal{V}, d_{\mathcal{V}})$ of f.*

Proof. Define for every element $y \in Y$ a function $g_y \colon X \to \mathcal{V}$ by

$$g_y(x) = d_{\mathcal{V}}(d(x, y), f(y))$$

Note that all the functions g_y are non-expansive, since for $x_1, x_2 \in X$ we have $d(x_1, x_2) \otimes d(x_2, y) \sqsubseteq d(x_1, y)$ (triangle inequality) and hence

$$\begin{aligned} d(x_1, x_2) &\sqsubseteq d_{\mathcal{V}}(d(x_2, y), d(x_1, y)) \\ &\sqsubseteq d_{\mathcal{V}}(d_{\mathcal{V}}(d(x_1, y), f(y)), d_{\mathcal{V}}(d(x_2, y), f(y))) && \text{(Lemma 25(9))} \\ &= d_{\mathcal{V}}(g_y(x_1), g_y(x_2)). \end{aligned}$$

Now define $\bar{f} \colon X \to \mathcal{V}$ by

$$\bar{f}(x) = \bigsqcap_{y \in Y} g_y(x) = \bigsqcap_{y \in Y} d_{\mathcal{V}}(d(x, y), f(y))$$

Once can easily verify that $\bar{f}$ is also non-expansive, because

$$\begin{aligned} d_{\mathcal{V}}(\bar{f}(x_1), \bar{f}(x_2)) &= d_{\mathcal{V}}\Big(\bigsqcap_{y \in Y} g_y(x_1), \bigsqcap_{y \in Y} g_y(x_2)\Big) \\ &\sqsupseteq \bigsqcap_{y \in Y} d_{\mathcal{V}}(g_y(x_1), g_y(x_2)) && \text{(Lemma 25(9))} \\ &\sqsupseteq \bigsqcap_{y \in Y} d(x_1, x_2) = d(x_1, x_2) \end{aligned}$$

It remains to show that $\bar{f}(y) = f(y)$ for all $y \in Y$. Consider $z \in Y$, then from the antitonicity of $d_{\mathcal{V}}$ in its first argument and Lemma 25(4):

$$\bar{f}(z) \sqsubseteq d_{\mathcal{V}}(d(z, z), f(z)) \sqsubseteq d_{\mathcal{V}}(k, f(z)) = f(z)$$

On the other hand, by the nonexpansivity of f, for all $y \in Y$ we have by commutativity:

$$\begin{aligned} d_{\mathcal{V}}(f(z), f(y)) \sqsupseteq d(z, y) &\iff d(z, y) \otimes f(z) \sqsubseteq f(y) \\ &\iff d_{\mathcal{V}}(d(z, y), f(y)) \sqsupseteq f(z) \end{aligned}$$

And therefore

$$\bar{f}(z) = \bigsqcap_{y \in Y} d_{\mathcal{V}}(d(z, y), f(y)) \sqsupseteq f(z)$$

◀

▶ **Remark 30.** Note that the function $\bar{f}$ constructed in the proof is the largest non-expansive extension of f. That means for all non-expansive extensions $h\colon (X,d) \to (\mathcal{V}, d_{\mathcal{V}})$ of f we have $h \sqsubseteq \bar{f}$. For if h is another non-expansive extension of f, then for all $x \in X$ and all $y \in Y$ we have

$$\begin{aligned} d_{\mathcal{V}}(h(x), f(y)) = d_{\mathcal{V}}(h(x), h(y)) \sqsupseteq d(x,y) &\iff h(x) \otimes d(x,y) \sqsubseteq f(y) \\ &\iff h(x) \sqsubseteq d_{\mathcal{V}}(d(x,y), f(y)) \end{aligned}$$

and therefore

$$h(x) \sqsubseteq \bigsqcap_{y \in Y} d_{\mathcal{V}}(d(x,y), f(y)) = \bar{f}(x).$$

Similarly, the function $g\colon X \to \mathcal{V}$ defined by

$$g(x) = \bigsqcup_{y \in Y} f(y) \otimes d(y,x)$$

is the smallest non-expansive function extending f.

As a corollary, we obtain the desired property needed to show the converse inclusion.

▶ **Theorem 4.** *Let $d_X \in \mathcal{V}\text{-}\mathsf{Cat}_X$ and $d_Y \in \mathcal{V}\text{-}\mathsf{Cat}_Y$ be elements of $\mathcal{V}\text{-}\mathsf{Cat}$. If $i\colon (Y, d_Y) \to (X, d_X)$ is an isometry, then for any non-expansive map $f\colon (Y, d_Y) \to (\mathcal{V}, d_{\mathcal{V}})$ there exists a non-expansive map $g\colon (X, d_X) \to (\mathcal{V}, d_{\mathcal{V}})$ such that $f = g \circ i$.*

Proof. The function $i\colon Y \to X$ can be factorised as the composition of $e\colon Y \to i[Y]$, where $i[Y]$ denotes the image of Y under i and $m\colon i[Y] \to X$, where m is the canonical inclusion function of $i[Y]$ into X. We can equip $i[Y]$ with a distance $d_{i[Y]}\colon i[Y] \times i[Y] \to \mathcal{V}$ defined to be the restriction of d_X to the domain $i[Y] \times i[Y]$. One can observe that $(i[Y], d_{i[Y]})$ is also an element of $\mathcal{V}\text{-}\mathsf{Cat}$ and that $m\colon (Y, d_Y) \to (i[Y], d_{i[Y]})$ is an isometry as a consequence of i being an isometry.

We define a function $h\colon i[Y] \to \mathcal{V}$ to be given by $h(e(y)) = f(y)$. We now argue that h is well-defined. Let $y_1, y_2 \in Y$, such that $e(y_1) = e(y_2)$. Since $f\colon (Y, d_Y) \to (\mathcal{V}, d_{\mathcal{V}})$ is non-expansive, we have that:

$$d_{\mathcal{V}}(f(y_1), f(y_2)) \sqsupseteq d_Y(y_1, y_2) = d_{i[Y]}(e(y_1), e(y_2)) \sqsupseteq k$$

The second to last step relies on the fact that $d_{i[Y]}$ is reflexive. By the definition of $d_{\mathcal{V}}$ we have that the above inequality is equivalent to

$$k \otimes f(y_1) \sqsubseteq f(y_2) \iff f(y_1) \sqsubseteq f(y_2)$$

The second step above relies on the fact that k is the unit of $\otimes$. The argument that $f(y_2) \sqsubseteq f(y_1)$ is symmetric and therefore omitted.

We now argue that $h\colon (i[Y], d_{i[Y]}) \to (\mathcal{V}, d_{\mathcal{V}})$ is non-expansive. Let $x_1, x_2 \in i[Y]$. Since e is surjective, there must exist y_1, y_2 with $e(y_1) = x_1$ and $e(y_2) = x_2$.

$$\begin{aligned} d_{\mathcal{V}}(h(x_1), h(x_2)) &= d_{\mathcal{V}}(h(e(y_1)), h(e(y_2))) && (e \text{ is surjective}) \\ &= d_{\mathcal{V}}(f(y_1), f(y_2)) && (\text{Def. of } h) \\ &\sqsupseteq d_Y(y_1, y_2) && (f \text{ is non-expansive}) \\ &= d_{i[Y]}(e(y_1), e(y_2)) && (e \text{ is an isometry}) \\ &= d_{i[Y]}(x_1, x_2) \end{aligned}$$

We can sum up the situation as the commuting diagram in $\mathcal{V}$-Cat.

$$\begin{tikzcd}
(Y, d_Y) \arrow[d, "e"] \arrow[dr, "f"] & \\
(i[Y], d_{i[Y]}) \arrow[r, "h"] & (\mathcal{V}, d_{\mathcal{V}}) \\
(X, d_X) \arrow[u, hook, "m"'] \arrow[ur, "g"'] &
\end{tikzcd}$$

The upper triangle commutes by the definition of h, while g exists and makes the bottom one commute because of Theorem 29. Since $i = m \circ e$, we have that there exists non-expansive g, such that $g \circ i = f$, which completes the proof. ◀

Note, that the property above does not hold for arbitrary $\mathcal{V}$-graphs and only works for $\mathcal{V}$-categories. To make use of that, we restrict $\Phi(X)$ to be the lattice $(\mathcal{V}\text{-Cat}_X, \sqsubseteq)$ of $\mathcal{V}$-categories on X.

▶ **Lemma 31.** *When restricted to $\mathcal{V}$-Cat, $\gamma\colon \Phi \to \Psi$ is natural.*

Proof. It suffices to show the converse inclusion to the one argued in Lemma 28. Let $p\colon X \to \mathcal{V}$ be an arbitrary $\mathcal{V}$-valued predicate on X, such that $d_Y \circ (f \times f) \sqsubseteq d_{\mathcal{V}} \circ (p \times p)$. We want to show that there exists a non-expansive map $q\colon (Y, d_Y) \to (\mathcal{V}, d_{\mathcal{V}})$, such that $p = q \circ f$. Observe that $p\colon X \to \mathcal{V}$ is a non-expansive map from $(X, d_Y \circ (f \times f))$ to $(\mathcal{V}, d_{\mathcal{V}})$ and $f\colon (Y, d_Y \circ (f \times f)) \to (Y, d_Y)$ is an isometry.

$$\begin{tikzcd}
(X, d_Y \circ (f \times f)) \arrow[d, "f"] \arrow[dr, "p"] & \\
(Y, d_Y) \arrow[r, "\exists q"] & (\mathcal{V}, d_{\mathcal{V}})
\end{tikzcd}$$

We can use Theorem 4 and conclude that there exists a non-expansive function $q\colon (Y, d_Y) \to (\mathcal{V}, d_{\mathcal{V}})$ such that $p = q \circ f$. Hence, $p \in (f^\bullet \circ \gamma_Y)(d_Y)$, which completes the proof. ◀

▶ **Proposition 5.** *We have that $\alpha\colon \Psi \Rightarrow \Phi$ is natural and $\gamma\colon \Phi \Rightarrow \Psi$ is laxly natural, that is for all functions $f\colon X \to Y$ 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, $\gamma\colon \Phi \Rightarrow \Psi$ is natural.*

Proof. Follows from Lemmas 27, 28 and 31. ◀

▶ **Lemma 32.** *For all sets X and collections of predicates $S \subseteq \mathcal{V}^X$, $\alpha_X(S)$ is a $\mathcal{V}$-category.*

Proof. Let $x \in X$. For reflexivity, we have that for all $p \in S$, $d_{\mathcal{V}}(p(x), p(x)) \sqsupseteq k$ and hence

$$\alpha_X(S)(x, x) = \bigsqcap_{p \in S} d_{\mathcal{V}}(p(x), p(x)) \sqsupseteq k$$

For transitivity, let $x_1, x_2 \in X$. We have to show the following for all $z \in X$:

$$\alpha_X(S)(x_1, z) \otimes \alpha_X(S)(z, x_2) \sqsubseteq \alpha_X(S)(x_1, x_2) = \bigsqcap_{p \in S} d_{\mathcal{V}}(p(x_1), p(x_2))$$

To show the above, it is enough to argue that for all $p \in S$, we have that $\alpha_X(S)(x_1, z) \otimes \alpha_X(z, x_2) \sqsubseteq d_{\mathcal{V}}(p(x_1), p(x_2))$. We have the following:

$$\begin{aligned}
\alpha_X(S)(x_1, z) \otimes \alpha_X(S)(z, x_2) &= \left(\bigsqcap_{r \in S} d_{\mathcal{V}}(r(x_1), r(z))\right) \otimes \left(\bigsqcap_{r \in S} d_{\mathcal{V}}(r(z), r(x_2))\right) \\
&\sqsubseteq \bigsqcap_{r \in S} d_{\mathcal{V}}(r(x_1), r(z)) \otimes d_{\mathcal{V}}(r(z), r(x_2)) \\
&\qquad\qquad ((\textstyle\bigsqcap_{i \in I} a_i) \otimes (\bigsqcap_{i \in I} b_i) \sqsubseteq \bigsqcap_{i \in I} a_i \otimes b_i) \\
&\sqsubseteq \bigsqcap_{r \in S} d_{\mathcal{V}}(r(x_1), r(x_2)) \qquad (d_{\mathcal{V}} \text{ is transitive, Lemma 25(3)}) \\
&\sqsubseteq d_{\mathcal{V}}(p(x_1), p(x_2)) \qquad (p \in S)
\end{aligned}$$

which completes the proof. ◀

Additionally, when restricting our attention to $\mathcal{V}$-Cat (instead of all $\mathcal{V}$-Graph), it turns out that the co-closure operator of the Galois connection (Theorem 3) becomes an identity.

▶ **Lemma 33.** *The co-closure $\alpha_X \circ \gamma_X$ is the identity, when restricted to $\mathcal{V}$-Cat.*

Proof. Let (X, d_X) be an object of $\mathcal{V}$-Cat. The inequality $d_X \sqsubseteq (\alpha_X \circ \gamma_X)(X, d_X)$ holds immediately because of Theorem 3. For the other direction, we will have to show that for all $x_1, x_2 \in X$, we have that $d_X(x_1, x_2) \sqsupseteq (\alpha_X \circ \gamma_X)(X, d_X)(x_1, x_2)$. To observe that, we define a function $p\colon X \to \mathcal{V}$ given by $p(x) = d_X(x_1, x)$. First, recall that for all $x_1, x_2 \in X$

$$\alpha_X(\gamma_X(d_X))(x_1, x_2) = \bigsqcap_{\substack{p\colon X \to \mathcal{V} \\ d_X \sqsubseteq d_{\mathcal{V}} \circ (p \times p)}} d_{\mathcal{V}}(p(x_1), p(x_2))$$

We show that it is non-expansive with respect to $d_{\mathcal{V}}$

$$\begin{aligned}
d_{\mathcal{V}}(p(x_1), p(x_2)) &= d_{\mathcal{V}}(d_X(x_1, x_1), d_X(x_1, x_2)) \\
&\sqsubseteq d_{\mathcal{V}}(k, d_X(x_1, x_2)) \qquad (\text{Anttonicity of } d_{\mathcal{V}}(-,-) \text{ in the first argument}) \\
&= d_X(x_1, x_2) \qquad (x = d_{\mathcal{V}}(k, x), \text{ Lemma 25(5)})
\end{aligned}$$

Hence

$$\alpha_X(\gamma_X(d_X))(x_1, x_2) \sqsubseteq d_{\mathcal{V}}(p(x_1), p(x_2)) \sqsubseteq d_X(x_1, x_2)$$

◀

▶ **Proposition 6.** *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 $d \in \mathcal{V}$-Graph$_X$, $\alpha_X(\gamma_X(d))$ is the metric closure of d, i.e., the least element of $\mathcal{V}$-Cat$_X$ above d.*

Proof. The first two statements follow from Lemmas 33 and 32.

Now let $d \in \mathcal{V}$-Graph$_X$, $d' \in \mathcal{V}$-Cat$_X$ with $d \sqsubseteq d'$. By antitonicity of α, γ it holds that $\alpha_X(\gamma_X(d)) \sqsubseteq \alpha_X(\gamma_X(d')) = d'$. Furthermore $\alpha_X(\gamma_X(d))$ is a $\mathcal{V}$-category, which implies the statement. ◀

▶ **Theorem 7.** *There is an adjunction $\gamma \dashv \alpha$.*

Proof. Let $d_X \in \mathcal{V}\text{-}\mathsf{Cat}_x$ be an arbitrary reflexive and transive $\mathcal{V}$-valued relation on X, $T \subseteq \mathcal{V}^Y$ and $f\colon X \to Y$ a non-expansive map from (X, d_X) to $(Y, \alpha_Y(T))$. We will define a non-expansive map $\eta_{d_X}\colon (X, d_X) \to (X, \alpha_X(\gamma_X(d_X)))$ and show that there exists a unique arrow $g\colon \gamma_X(d_X) \to T$ which makes the following diagram commute.

$$\begin{tikzcd}
(X, d_X) \arrow[r, "\eta_{d_X}"] \arrow[d, "f"] & (X, \alpha_X(\gamma_X(d_X))) \arrow[dl, dashed, "\alpha(g)"] \\
(Y, \alpha_Y(T)) &
\end{tikzcd}$$

First, we set $\eta_{d_X} = \mathrm{id}$ to be the identity map. In order to show that it is non-expansive, we need to argue that $d_X \sqsubseteq \alpha_X(\gamma_X(d_X)) = \alpha_X(\gamma_X(d_X))$. This immediately follows from Theorem 3, where we have shown that α_X and γ_X form a contravariant Galois connection.

Recall that $\alpha(g) = g$ and hence the diagram above would uniquely commute if we set $g = f$. The only remaining thing is to argue that f is a morphism in $\mathcal{V}\text{-}\mathsf{SPred}$ between $(X, \gamma_X(d_X))$ and (Y, T). In other words, we need to prove that:

$$f^\bullet(T) = \{q \circ f \mid q \in T\} \subseteq \gamma_X(d_X) = \{p\colon X \to \mathcal{V} \mid d_X \sqsubseteq d_{\mathcal{V}} \circ (p \times p)\}$$

Because $f\colon (X, d_X) \to \alpha_X(T)$ is non-expansive, we have that for all $x_1, x_2 \in X$:

$$d_X(x_1, x_2) \sqsubseteq \bigsqcap_{q \in T} d_{\mathcal{V}}(q(f(x_1)), q(f(x_2)))$$

and therefore for any $q \in T$, we have that:

$$d_X(x_1, x_2) \sqsubseteq d_{\mathcal{V}}((q \circ f)(x_1), (q \circ f)(x_2))$$

Observe that because of the inequality above every element of $f^\bullet(T)$ also belongs to $\gamma_X(d_X)$, which completes the proof. ◀

C Proofs for Section 4 (The Kantorovich Lifting)

▶ **Lemma 8.** *The Kantorovich lifting of a functor F : $\mathsf{Set} \to \mathsf{Set}$ is a functor $\overline{F} = K_{\Lambda^F}\colon \mathcal{V}\text{-}\mathsf{Graph} \to \mathcal{V}\text{-}\mathsf{Graph}$, and fibred when restricted to $\mathcal{V}\text{-}\mathsf{Cat}$.*

Proof. We need to show that for an arbitrary non-expansive function $f\colon (X, d_X) \to (Y, d_X)$, Ff becomes non-expansive function between $\overline{F}(X, d_X)$ and $\overline{F}(Y, d_Y)$. In other words, we require that:

$$(\alpha_{FX} \circ \Lambda^F_X \circ \gamma_X)(d_X) \sqsubseteq (\alpha_{FY} \circ \Lambda^F_Y \circ \gamma_Y)(d_Y) \circ (Ff \times Ff)$$

First, we simplify the right-hand side:

$$\begin{aligned}
&\big((\alpha_{FY} \circ \Lambda^F_Y \circ \gamma_Y)(d_Y) \circ (Ff \times Ff)\big) \\
&= \big((\alpha_{FY} \circ \Lambda^F_Y \circ \gamma_Y(d_Y)) \circ (Ff \times Ff)\big) \\
&= \big(Ff^* \circ \alpha_{FY} \circ \Lambda^F_Y \circ \gamma_Y(d_Y)\big) && \text{(Definition of } (-)^*\text{)} \\
&= \big(\alpha_{FX} \circ (Ff)^\bullet \circ \Lambda^F_Y \circ \gamma_Y(d_Y)\big) && (\alpha \text{ is natural}) \\
&= \big(\alpha_{FX} \circ \Lambda^F_Z \circ f^\bullet \circ \gamma_Y(d_Y)\big)
\end{aligned}$$

Hence, we have the following:

$$(\alpha_{FX} \circ \Lambda_X^F \circ \gamma_X)(d_X) \sqsubseteq \big(\alpha_{FX} \circ \Lambda_X^F \circ f^\bullet \circ \gamma_Y(d_Y)\big)$$

Since α_{FX} is antitone (Theorem 3), it suffices to argue the following:

$$(\Lambda_X^F \circ \gamma_X)(d_X) \supseteq (\Lambda_X^F \circ f^\bullet \circ \gamma_Y)(d_Y)$$

Because of the lax naturality of γ (Lemma 28), we simplify the above even further:

$$(\Lambda_X^F \circ \gamma_X)(d_X) \supseteq (\Lambda_X^F \circ \gamma_X \circ f^*)(d_Y)$$

Since it is not hard to see that Λ_X^F is monotone, we are left with proving:

$$\gamma_Y(d_Y \circ (f \times f)) \subseteq \gamma_X(d_X)$$

Let $p \in \gamma_X(d_Y \circ (f \times f))$. In other words, for all $x_1, x_2 \in X$ we have that

$$d_{\mathcal{V}}(p(x_1), p(x_2)) \sqsupseteq d_Y(f(x_1), f(x_2))$$

Combining the above with the fact that f is non-expansive function from (X, d_X) to (Y, d_Y), we obtain that $d_{\mathcal{V}}(p(x_1), p(x_2)) \sqsupseteq d_X(x_1, x_2)$ and therefore $p \in \gamma_X(d_X)$, thus showing functoriality.

For fibredness, consider a function $f : X \to Y$ in Set, and an object (Y, d_Y) in $\mathcal{V}$-Cat. We wish to see that for all $s, t \in FX$

$$K_{\Lambda^F}(d_Y \circ (f \times f))(s,t) = K_{\Lambda^F}(d_Y) \circ (Ff \times Ff)(s,t)$$

This is equivalent to

$$\begin{aligned} &\bigsqcap_{ev \in \Lambda^F} \bigsqcap_{p \in \gamma_X(d_Y \circ (f \times f))} d_{\mathcal{V}}(ev(Fp(s)), ev(Fp(t))) \\ = &\bigsqcap_{ev \in \Lambda^F} \bigsqcap_{q \in \gamma_Y(d_Y)} d_{\mathcal{V}}(ev(Fq(Ff(s))), ev(Fq(Ff(t)))) \end{aligned}$$

and since in the category $\mathcal{V}$-Cat, we have that $(f^\bullet \circ \gamma_Y)(d_Y) = (\gamma_X \circ f^*)(d_Y)$, then the above equality holds, completing the proof.

◀

▶ **Example 34.** Consider the powerset functor $\mathcal{P}$ with the predicate lifting sup, and the discrete distribution functor $\mathcal{D}$ with the predicate lifting $\mathbb{E}$ that takes expected values. With just these predicate liftings, compositionality fails for all four combinations $\mathcal{PP}$, $\mathcal{PD}$, $\mathcal{DP}$, $\mathcal{DD}$. In all cases, 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 following two non-expansive functions on $\mathcal{P}X$ respectively $\mathcal{D}X$:

$$f_{\mathcal{P}}\colon (\mathcal{P}X, K_{\{\sup\}}d) \to ([0,1], d_{[0,1]_\oplus}) \qquad f_{\mathcal{D}}\colon (\mathcal{D}X, K_{\{\mathbb{E}\}}d) \to ([0,1], d_{[0,1]_\oplus})$$

$$f_{\mathcal{P}}(A) = \begin{cases} 1, & \text{if } A = \{x,y\} \\ 0, & \text{otherwise} \end{cases} \qquad f_{\mathcal{D}}(p \cdot x + (1-p) \cdot y) = \min(p, 1-p)$$

To see that $f_{\mathcal{P}}$ is indeed non-expansive, note that for $B \subseteq A$ we have $f_{\mathcal{P}}(B) \ominus f_{\mathcal{P}}(A) = 0 = K_{\{\sup\}}d(A,B)$ by monotonocity of $f_{\mathcal{P}}$ and sup, and for $B \nsubseteq A$ we have $K_{\{\sup\}}d(A,B) = 1$ as we may pick g to be the characteristic function of some $z \in B \setminus A$. In $\mathcal{D}X$, the distance of $\mu = p \cdot x + (1-p) \cdot y$ and $\nu = q \cdot x + (1-q) \cdot y$ can be computed as

$$K_{\{\mathbb{E}\}}d(\mu, \nu) = \sup\{(q \cdot a + (1-q) \cdot b) - (p \cdot a + (1-p) \cdot b) \mid a, b \in [0,1]\} = |q - p|,$$

so that $f_{\mathcal{D}}$ is easily seen to be non-expansive as well.

1. Put $U = \{\{x\},\{y\}\}$ and $V = \{\{x\},\{y\},\{x,y\}\}$. Then

$$\sup f_{\mathcal{P}}[U] = 0 \quad \text{and} \quad \sup f_{\mathcal{P}}[V] = 1,$$

so that $K_{\{\sup\}}(K_{\{\sup\}}d)(U,V) = 1$. For every $g\colon X \to [0,1]$ one finds that

$$(\sup * \sup)(g)(U) = \max(g(x), g(y)) = (\sup * \sup)(g)(V),$$

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

2. 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

$$\sup f_{\mathcal{D}}[U] = \max(0,0) = 0 \quad \text{and} \quad \sup f_{\mathcal{D}}[V] = \max(0, 1/2, 0) = 1/2,$$

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

3. Put $\mu = 1/2\cdot\{x\} + 1/2\cdot\{y\}$ and $\nu = 1\cdot\{x,y\}$. Then

$$\mathbb{E}_{\mu}(f_{\mathcal{P}}) = 1/2\cdot 0 + 1/2\cdot 0 = 0 \quad \text{and} \quad \mathbb{E}_{\nu}(f_{\mathcal{P}}) = 1\cdot 1 = 1,$$

so that $K_{\{\mathbb{E}\}}(K_{\{\sup\}}d)(U,V) = 1$. For every $g\colon X \to [0,1]$ one finds that

$$(\mathbb{E} * \sup)(g)(\mu) = g(x)+g(y)/2 \quad \text{and} \quad (\mathbb{E} * \sup)(g)(\nu) = \max(g(x), g(y)),$$

implying that $K_{\{\mathbb{E} * \sup\}}(d)(\mu,\nu) \leq 1/2$.

4. Put $\mu = 1/2\cdot(1\cdot x) + 1/2\cdot(1\cdot y)$ and $\nu = 1\cdot(1/2\cdot x + 1/2\cdot y)$. Then

$$\mathbb{E}_{\mu}(f_{\mathcal{D}}) = 0 \quad \text{and} \quad \mathbb{E}_{\nu}(f_{\mathcal{D}}) = 1/2,$$

so that $K_{\{\mathbb{E}\}}(K_{\{\mathbb{E}\}}d)(U,V) \geq 1/2$. For every $g\colon X \to [0,1]$ one finds that

$$(\mathbb{E} * \mathbb{E})(g)(\mu) = g(x)+g(y)/2 = (\mathbb{E} * \mathbb{E})(g)(\nu),$$

implying that $K_{\{\mathbb{E} * \mathbb{E}\}}(d)(\mu,\nu) = 0$.

▶ **Proposition 12.** *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*

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

Proof. **constant functors:** let $b, c \in B$. We have, since Ff is the identity:

$$\begin{aligned} d_X^F(b,c) &= \bigsqcap_{ev\in\Lambda^F} \bigsqcap_{f\in\gamma(d_X)} d_{\mathcal{V}}(ev(Ff(b)), ev(Ff(c))) \\ &= \bigsqcap_{ev\in\Lambda^F} d_{\mathcal{V}}(ev(b), ev(c)) \end{aligned}$$

identity functor: let $x, y \in X$. We have, since $Ff = f$ and ev is the identity:

$$\begin{aligned} d_X^F(x,y) &= \bigsqcap_{ev \in \Lambda^F} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev(Ff(x)), ev(Ff(y))) \\ &= \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(f(x)), f(y))) = \alpha_X(\gamma_X(d_X))(x,y) \end{aligned}$$

product functors: Let $s, t \in \prod_{i \in I} F_i X$. Let $\pi_i \colon \prod_{i \in I} F_i X \to F_i X$, $\pi_i' \colon \prod_{i \in I} F_i \mathcal{V} \to F_i \mathcal{V}$. Then

$$\begin{aligned} d_X^F(s,t) &= \bigsqcap_{ev \in \Lambda^F} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev(Ff(s)), ev(Ff(t))) \\ &= \bigsqcap_{i \in I} \bigsqcap_{ev_i \in \Lambda^{F_i}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev_i(\pi_i'(Ff(s))), ev_i(\pi_i'(Ff(t)))) \\ &= \bigsqcap_{i \in I} \bigsqcap_{ev_i \in \Lambda^{F_i}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev_i(F_i f(\pi_i(s))), ev_i(F_i f(\pi_i(t)))) \\ &= \bigsqcap_{i \in I} d_X^{F_i}(\pi_i(Ff(s)), \pi_i(Ff(t))) \end{aligned}$$

using the fact that $\pi_i' \circ Ff = F_i f \circ \pi_i$.

coproduct functors: We use the properties of $d_{\mathcal{V}}$ shown in Lemma 25. Whenever $s, t \in F_1 X$, we have that

$$\begin{aligned} d_X^F(s,t) &= \bigsqcap_{ev_1 \in \Lambda^{F_1}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev_1(Ff(s)), ev_1(Ff(t))) \sqcap d_{\mathcal{V}}(\bot, \bot) \\ &= d^{F_1}(s,t) \sqcap \top = d^{F_1}(s,t) \end{aligned}$$

If instead $s, t \in F_2 X$, we obtain

$$\begin{aligned} d_X^F(s,t) &= \bigsqcap_{ev_2 \in \Lambda^{F_2}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev_2(Ff(s)), ev_2(Ff(t))) \sqcap d_{\mathcal{V}}(\top, \top) \\ &= d^{F_2}(s,t) \sqcap \top = d^{F_2}(s,t) \end{aligned}$$

Now assume that $s \in F_1 X$, $t \in F_2 X$. Then:

$$\begin{aligned} d_X^F(s,t) &= \bigsqcap_{ev_1 \in \Lambda^{F_1}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev_1(Ff(s)), \top) \sqcap \\ &\qquad \bigsqcap_{ev_2 \in \Lambda^{F_2}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(\bot, ev_1(Ff(t))) \sqcap d_{\mathcal{V}}(\bot, \top) \\ &= \top \sqcap \top \sqcap \top = \top \end{aligned}$$

Now assume that $s \in F_2 X$, $t \in F_1 X$. Then, since $d_{\mathcal{V}}(\bot, \top) = \bot$:

$$\begin{aligned} d_X^F(s,t) &= \bigsqcap_{ev_1 \in \Lambda^{F_1}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(\top, ev_1(Ff(t))) \sqcap \\ &\qquad \bigsqcap_{ev_2 \in \Lambda^{F_2}} \bigsqcap_{f \in \gamma(d_X)} d_{\mathcal{V}}(ev_2(Ff(s)), \bot)) \sqcap d_{\mathcal{V}}(\top, \bot) = \bot \end{aligned}$$

◀

▶ **Proposition 13.** *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.*

Proof. Given F and Λ^F, we need to show that for every set X, every $S \subseteq \mathcal{V}^X$, every $f \colon (X, \alpha_X(S)) \to (\mathcal{V}, d_{\mathcal{V}})$, every $ev \in \Lambda^F$ and every $t_1, t_2 \in FX$ we have

$$\alpha_X(\Lambda_X^F(S))(t_1, t_2) \sqsubseteq d_{\mathcal{V}}(ev \circ Ff(t_1), ev \circ Ff(t_2)). \tag{3}$$

We proceed by induction over F.

- For $F = C_B$, the set Λ^F consists of some maps of type $B \to \mathcal{V}$. Then for (3) to hold we need that for every $b_1, b_2 \in B$ and every $ev \in \Lambda^F$ we have

$$\bigsqcap_{ev' \in \Lambda^F} d_{\mathcal{V}}(ev'(b_1), ev'(b_2)) \sqsubseteq d_{\mathcal{V}}(ev(b_1), ev(b_2)).$$

  This holds automatically, regardless of the specific choice of Λ^F, as the term on the right is one of the terms in the meet on the left.
- For $F = \mathrm{Id}$, the equation (3) reduces to the statement that f is non-expansive wrt. $\alpha_X(S)$, which is true by assumption.
- For $F = \prod_{i \in I} F_i$, let $s = (s_i)_{i \in I}, t = (t_i)_{i \in I} \in FX$ and fix some $j \in I$. Then we have for each $ev \in \Lambda^{F_j}$ and each $g \colon X \to \mathcal{V}$ that $ev \circ \pi_j \circ Fg(s) = ev \circ F_j g(s_j)$ and $ev \circ \pi_j \circ Fg(t) = ev \circ F_j g(t_j)$, and thus

$$\begin{aligned}
&\alpha_X(\Lambda^F_X(S))(s,t) \\
&= \bigsqcap_{i \in I} \bigsqcap_{ev \in \Lambda^{F_i}} \bigsqcap_{g \in S} d_{\mathcal{V}}(ev \circ \pi_i \circ Fg(s), ev \circ \pi_i \circ Fg(t)) \\
&\sqsubseteq \bigsqcap_{ev \in \Lambda^{F_j}} \bigsqcap_{g \in S} d_{\mathcal{V}}(ev \circ \pi_j \circ Fg(s), ev \circ \pi_j \circ Fg(t)) \\
&= \bigsqcap_{ev \in \Lambda^{F_j}} \bigsqcap_{g \in S} d_{\mathcal{V}}(ev \circ F_j g(s_j), ev \circ F_j g(t_j)) \\
&= \alpha_X(\Lambda^{F_j}_X(S))(s_j, t_j) \\
&\sqsubseteq d_{\mathcal{V}}(ev \circ F_j f(s_j), ev \circ F_j f(t_j)) && \text{(IH)} \\
&= d_{\mathcal{V}}(ev \circ \pi_j \circ Ff(s), ev \circ \pi_j \circ Ff(t))
\end{aligned}$$

  for each $f \colon (X, \alpha_X(S)) \to (\mathcal{V}, d_{\mathcal{V}})$ and each $ev \in \Lambda^{F_j}$.
- For $F = F_1 + F_2$, let $s \in F_i X \subseteq FX$ and $t \in F_j X \subseteq FX$, and let $ev \in \Lambda^F$.
  - If $i = 1$ and ev is of the form $[\bot, \top]$ or $[\bot, ev_2]$, then the right hand side of (3) evaluates to $d_{\mathcal{V}}(\bot, v)$ for some $v \in \mathcal{V}$, which equals $\top$.
  - If $j = 2$ and ev is of the form $[\bot, \top]$ or $[ev_1, \top]$, then the right hand side of (3) evaluates to $d_{\mathcal{V}}(v, \top)$ for some $v \in \mathcal{V}$, which equals $\top$.
  - If $i = j = 1$, and ev is of the form $[ev_1, \top]$, then we show (3) as follows.

$$\begin{aligned}
&\alpha_X(\Lambda^F_X(S))(s,t) \\
&\sqsubseteq \bigsqcap_{ev_1 \in \Lambda^{F_1}} \bigsqcap_{g \in S} d_{\mathcal{V}}([ev_1, \top] \circ Fg(s), [ev_1, \top] \circ Fg(t)) \\
&= \bigsqcap_{ev_1 \in \Lambda^{F_1}} \bigsqcap_{g \in S} d_{\mathcal{V}}(ev_1 \circ F_1 g(s), ev_1 \circ F_1 g(t)) \\
&= \alpha_X(\Lambda^{F_1}_X(S))(s,t) \\
&\sqsubseteq d_{\mathcal{V}}(ev_1 \circ F_1 f(s), ev_1 \circ F_1 f(t)) && \text{(IH)} \\
&= d_{\mathcal{V}}([ev_1, \top] \circ Ff(s), [ev_1, \top] \circ Ff(t))
\end{aligned}$$

  - If $i = j = 2$, and ev is of the form $[\bot, ev_2]$, the proof is very similar to the previous item.
  - The only remaining case is that where $i = 2$ and $j = 1$. In that case, we can pick $[\bot, \top] \in \Lambda^F$ and some arbitrary $g \in S$ (here it is important that S is non-empty) to show that the left hand side of (3) evaluates to $\bot$:

$$\alpha_X(\Lambda^F_X(S))(s,t) \sqsubseteq d_{\mathcal{V}}([\bot, \top] \circ Fg(s), [\bot, \top] \circ Fg(t)) = d_{\mathcal{V}}(\top, \bot) = \bot.$$

◀

D Proofs for Section 5 (Application: Up-To Techniques)

▶ **Proposition 18.** *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.*

Proof. We will proceed by induction over the shape of F. We wish to show that they $\zeta\colon TF \Rightarrow FT$ is a natural transformations and that the two diagrams in Definition 15 commute.

constant functors: If $F = C_B$, the naturality is clear and the distributive law diagrams become

$$\begin{tikzcd} B \arrow[d, "\eta_X"'] \arrow[dr, "\mathrm{Id}_B"] & \\ TB \arrow[r, "\zeta"'] & B \end{tikzcd} \qquad \begin{tikzcd} T^2B \arrow[r, "T\zeta"] \arrow[d, "\mu_B"'] & TB \arrow[r, "\zeta"] & B \arrow[d, "\mathrm{Id}_B"] \\ TB \arrow[rr, "\zeta"'] & & B, \end{tikzcd}$$

which are precisely the conditions for the map $\zeta : TB \to B$ being an algebra over T.

identity functor: For $F = \mathrm{Id}$, it is also clear that $\mathrm{id} : T \Rightarrow T$ is a natural transformation and a distributive law.

product functors: For $F = \prod_{i\in I} F_i$, we note that the definition of the distributive law ζ in terms of the ζ^i amounts to the set of equalities $\zeta^i \circ T\pi_i = \pi_i \circ \zeta$ for each $i \in I$. Therefore we can verify the triangle and pentagon diagrams via the following computations:

$$\begin{aligned} F\eta &= \langle F_i\eta \circ \pi_i\rangle \\ &= \langle \zeta^i \circ \eta_{F_i} \circ \pi_i\rangle && \text{(IH)} \\ &= \langle \zeta^i \circ T\pi_i \circ \eta_F\rangle && (\eta\text{ natural}) \\ &= \langle \pi_i \circ \zeta \circ \eta_F\rangle = \zeta \circ \eta_F \\ F\mu \circ \zeta \circ T\zeta &= \langle F_i\mu \circ \pi_i \circ \zeta \circ T\zeta\rangle \\ &= \langle F_i\mu \circ \zeta^i \circ T\pi_i \circ T\zeta\rangle \\ &= \langle F_i\mu \circ \zeta^i \circ T\zeta^i \circ TT\pi_i\rangle \\ &= \langle \zeta^i \circ \mu_{F_i} \circ TT\pi_i\rangle && \text{(IH)} \\ &= \langle \zeta^i \circ T\pi_i \circ \mu_F\rangle && (\mu\text{ natural}) \\ &= \langle \pi_i \circ \zeta \circ \mu_F\rangle = \zeta \circ \mu_F \end{aligned}$$

coproduct functors: For $F = F_1 + F_2$, we first check the commutativity of the triangular diagram. In the diagram below, the left inner triangle commutes because of assumption on g, while the right inner triangle commutes by induction hypothesis.

$$\begin{tikzcd} & F_1X + F_2X \arrow[dl, "\eta_{F_1X+F_2X}"'] \arrow[d, "\eta_{F_1X}+\eta_{F_2X}"] \arrow[dr, "F_1\eta_X+F_2\eta_X"] & \\ T(F_1X + F_2X) \arrow[r, "g_{F_1X,F_2X}"'] & TF_1X + TF_2X \arrow[r, "\zeta^1_X+\zeta^2_X"'] & F_1TX + F_2TX \end{tikzcd}$$

Hence, the outer diagram commutes, establishing the required condition. Moving on to the pentagonal diagram, we have the following:

$$\begin{tikzcd}
TT(F_1X+F_2X) \arrow[r, "Tg_{F_1X,F_2X}"] \arrow[ddd, "\mu_{F_1X+F_2X}"] & T(TF_1X+TF_2X) \arrow[r, "T(\zeta^1_X+\zeta^2_X)"] \arrow[d, "g_{TF_1X,TF_2X}"] & T(F_1TX+F_2TX) \arrow[d, "g_{F_1TX,F_2TX}"] \\
 & TTF_1X+TTF_2X \arrow[r, "T\zeta^1_X+T\zeta^2_X"] \arrow[dd, "\mu_{F_1X}+\mu_{F_2X}"] & TF_1TX+TF_2TX \arrow[d, "\zeta^1_{TX}+\zeta^2_{TX}"] \\
 & & F_1TTX+F_2TTX \arrow[d, "F_1\mu_X+F_2\mu_X"] \\
T(F_1X+F_2X) \arrow[r, "g_{F_1X,F_2X}"] & TF_1X+TF_2X \arrow[r, "\zeta^1_X+\zeta^2_X"] & F_1TX+F_2TX
\end{tikzcd}$$

The top right square commutes by naturality of g, while the bottom right diagram commutes by induction hypothesis. Finally, the left diagram commutes because of the assumption we make on g. The commutativity of the outer diagram yields the desired result.

◀

▶ **Proposition 19.** *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 $\overline{F} = K_{\Lambda^F}$, $\overline{G} = K_{\Lambda^G}$.*

Proof. To prove this, we need to show that the transformation ζ is non-expansive in each component, i.e. that for every $(X, d) \in \mathcal{V}$-Graph,

$$d^F \sqsubseteq d^G \circ (\zeta_X \times \zeta_X).$$

Unpacking the proof of the Kantorovich lifting, we wish to show that for every $t_1, t_2 \in FX$

$$\begin{aligned}
d^F(t_1,t_2) &= \bigsqcap \{d_\mathcal{V}((ev_F \circ Ff)(t_1), (ev_F \circ Ff)(t_2)) \mid ev_F \in \Lambda^F, \\
&\qquad\qquad d \sqsubseteq d_\mathcal{V} \circ (f \times f),\ f : X \to \mathcal{V}\} \\
&\sqsubseteq \bigsqcap \{d_\mathcal{V}((ev_G \circ \zeta_\mathcal{V} \circ Ff)(t_1), (ev_G \circ \zeta_\mathcal{V} \circ Ff)(t_2)) \mid ev_G \in \Lambda^G, \\
&\qquad\qquad d \sqsubseteq d_\mathcal{V} \circ (f \times f),\ f : X \to \mathcal{V}\} \\
&= \bigsqcap \{d_\mathcal{V}((ev_G \circ Gf \circ \zeta_X)(t_1), (ev_G \circ Gf \circ \zeta_X)(t_2)) \mid ev_G \in \Lambda^G, \\
&\qquad\qquad d \sqsubseteq d_\mathcal{V} \circ (f \times f),\ f : X \to \mathcal{V}\} \\
&= d^G(\zeta_X(t_1), \zeta_X(t_2)).
\end{aligned}$$

by the assumption we know that the set of functions of the form $ev_G \circ \zeta_\mathcal{V}$ is a superset of the set of functions of the form ev_F, and so the first infimum is smaller or equal than the second. And by the naturality of ζ, we have that $\zeta_\mathcal{V} \circ Ff = Gf \circ \zeta_X$. ◀

▶ **Lemma 21.** *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 defined in Section 4.3, we have that

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

Proof. We will prove this via the inductive definition of F. In particular we will show that for each $ev \in \Lambda^T * \Lambda^F$ there exists $ev' \in \Lambda^F * \Lambda^T$ such that $ev = ev' \circ \zeta_V$. And for each $ev' \in \Lambda^F * \Lambda^T$ there exists $ev \in \Lambda^T * \Lambda^F$ such that $ev = ev' \circ \zeta_V$.

For the case where $F = C_B$ we obtain:

$$\begin{aligned}\Lambda^F * \Lambda^T &= \{ev_F \circ F ev_T \mid ev_F \in \Lambda^F\} = \{ev_F \circ \mathrm{id}_B \mid ev_F \in \Lambda^F\} = \Lambda^F \\ \Lambda^T * \Lambda^F &= \{ev_T \circ T ev_F \mid ev_F \in \Lambda^F\} = \{ev_F \circ \zeta_{\mathcal{V}} \mid ev_F \in \Lambda^F\}\end{aligned}$$

where the last equality holds since each map ev_F is an algebra map from $\zeta_{\mathcal{V}}$ to ev_T by assumption. In this case we clearly have the correspondence stated above.

For $F = \mathrm{Id}$ we have remember that $ev_F \in \Lambda^F$ and $\zeta_{\mathcal{V}}$ are both the identity. Hence:

$$\begin{aligned}\Lambda^F * \Lambda^T &= \{ev_F \circ F ev_T \mid ev_F \in \Lambda^F\} = \{ev_T\} = \Lambda^T \\ \Lambda^T * \Lambda^F &= \{ev_T \circ \mathrm{id}_{T\mathcal{V}} \mid ev_F \in \Lambda^F\} = \Lambda^T\end{aligned}$$

Since both sets are the same and $\zeta_{\mathcal{V}}$ is the identity, the correspondence clearly holds.

Now, consider the case for the product functor $F = \prod_{i\in I} F_i$. Then

$$\begin{aligned}\Lambda^F * \Lambda^T &= \{ev_i' \circ \pi_i \circ F ev_T \mid i \in I, ev_i' \in \Lambda^{F_i}, ev_T \in \Lambda^T\} \\ \Lambda^T * \Lambda^F &= \{ev_T \circ T ev_i \circ T\pi_i \mid i \in I, ev_i \in \Lambda^{F_i}, ev_T \in \Lambda^T\}\end{aligned}$$

Here $\zeta_{\mathcal{V}} : T\prod_{i\in I} F_i\mathcal{V} \to \prod_{i\in I} F_i T\mathcal{V}$ is given as the universal map $\langle \zeta_{\mathcal{V}}^i \circ T\pi_i\rangle_{i\in I}$, i.e. the map such that $\pi_i' \circ \zeta_{\mathcal{V}} = \zeta_{\mathcal{V}}^i \circ T\pi_i$, where $\pi_i' \colon \prod_{i\in I} F_i T\mathcal{V} \to F_i T\mathcal{V}$.

Fix $i \in I$, $ev_i \in \Lambda^{F_i}$. Then, $ev_T \circ T ev_i \in \Lambda^T * \Lambda^{F_i}$ and by the induction hypothesis there exists $ev_i' \circ F_i ev_T \in \Lambda^{F_i} * \Lambda^T$ with $ev_i' \in \Lambda^{F_i}$ such that $ev_T \circ T ev_i = ev_i' \circ F_i ev_T \circ \zeta_{\mathcal{V}}^i$.

Now for $ev_T \circ T ev_i \circ T\pi_i \in \Lambda^T * \Lambda^F$ we pick $ev_i' \circ \pi_i \circ F ev_T \in \Lambda^F * \Lambda^T$ as corresponding map and we have:

$$\begin{aligned}ev_i' \circ \pi_i \circ F ev_T \circ \zeta_{\mathcal{V}} &= ev_i' \circ \pi_i \circ F ev_T \circ \langle \zeta_{\mathcal{V}}^i \circ T\pi_i\rangle_i \\ &= ev_i' \circ \pi_i \circ \langle F_i ev_T \circ \zeta_{\mathcal{V}}^i \circ T\pi_i\rangle \\ &= ev_i' \circ F_i ev_T \circ \zeta_{\mathcal{V}}^i \circ T\pi_i \\ &= ev_T \circ T ev_i \circ T\pi_i\end{aligned}$$

The proof is analogous for the other direction.

Finally we address the case of the coproduct functor. Then

$$\begin{aligned}\Lambda^F * \Lambda^T &= \{[ev_1 \circ F ev_T, \top] \mid ev_1 \in \Lambda^{F_1}\} \cup \{[\bot, ev_2 \circ F ev_T] \mid ev_2 \in \Lambda^{F_2}\} \cup \{[\bot, \top]\} \\ \Lambda^T * \Lambda^F &= \{ev_T \circ T[ev_1, \top] \mid ev_1 \in \Lambda^{F_1}\} \cup \{ev_T \circ T[\bot, ev_2] \mid ev_2 \in \Lambda^{F_2}\} \cup \{ev_T \circ T[\bot, \top]\}\end{aligned}$$

If we fix $ev_i \in \Lambda^{F_i}$, then $ev_T \circ T ev_i \in \Lambda^T * \Lambda^{F_i}$ and by the induction hypothesis there exists $ev_i' \circ F_i ev_T \in \Lambda^{F_i} * \Lambda^T$ with $ev_i' \in \Lambda^{F_i}$ such that $ev_T \circ T ev_i = ev_i' \circ F_i ev_T \circ \zeta_{\mathcal{V}}^1$.

We now consider three cases, according to the three subsets above.

For $ev_T \circ T[ev_1, \top] \in \Lambda^T * \Lambda^F$ we pick $[ev_1' \circ F ev_T, \top] \in \Lambda^F * \Lambda^T$ as corresponding map and we have, using well-behavedness of g:

$$\begin{aligned}[ev_1' \circ F ev_T, \top] \circ \zeta_{\mathcal{V}} &= [ev_1' \circ F ev_T, \top] \circ (\zeta_{\mathcal{V}}^1 + \zeta_{\mathcal{V}}^2) \circ g_{F_1X, F_2X} \\ &= [ev_1' \circ F ev_T \circ \zeta_{\mathcal{V}}^1, \top \circ \zeta_{\mathcal{V}}^2] \circ g_{F_1X, F_2X} \\ &= [ev_1' \circ F ev_T \circ \zeta_{\mathcal{V}}^1, \top] \circ g_{F_1X, F_2X} \\ &= [ev_T \circ T ev_1, \top] \circ g_{F_1X, F_2X} \\ &= ev_T \circ T[ev_1, \top]\end{aligned}$$

For $ev_T \circ T[\bot, ev_2] \in \Lambda^T * \Lambda^F$ we pick $[\bot, ev_2' \circ Fev_T] \in \Lambda^F * \Lambda^T$ as corresponding map and we have, using well-behavedness of g:

$$\begin{aligned}
[\bot, ev_2' \circ Fev_T] \circ \zeta_\mathcal{V} &= [\bot, ev_2' \circ Fev_T] \circ (\zeta_\mathcal{V}^1 + \zeta_\mathcal{V}^2) \circ g_{F_1X, F_2X} \\
&= [\bot \circ \zeta_\mathcal{V}^1, ev_2' \circ Fev_T \circ \zeta_\mathcal{V}^2] \circ g_{F_1X, F_2X} \\
&= [\bot, ev_2' \circ Fev_T \circ \zeta_\mathcal{V}^2] \circ g_{F_1X, F_2X} \\
&= [\bot, ev_T \circ Tev_2] \circ g_{F_1X, F_2X} \\
&= ev_T \circ T[\bot, ev_2]
\end{aligned}$$

For $ev_T \circ T[\bot, \top] \in \Lambda^T * \Lambda^F$ we pick $[\bot, \top] \in \Lambda^F * \Lambda^T$ as corresponding map and we have, using well-behavedness of g:

$$\begin{aligned}
[\bot, \top] \circ \zeta_\mathcal{V} &= [\bot, \top] \circ (\zeta_\mathcal{V}^1 + \zeta_\mathcal{V}^2) \circ g_{F_1X, F_2X} \\
&= [\bot \circ \zeta_\mathcal{V}^1, \top circ \zeta_\mathcal{V}^2] \circ g_{F_1X, F_2X} \\
&= [\bot, \top] \circ g_{F_1X, F_2X} \\
&= ev_T \circ T[\bot, ev_2]
\end{aligned}$$

The other directions can be shown analogously. This concludes the proof. ◀

▶ **Proposition 22.** *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.*

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

Proof. We have to prove the following properties of g:

naturality: let $f_i \colon X_i \to Y_i$, $X' \subseteq X_1 + X_1$. We distinguish the following cases: if $X' \cap X_1 \neq \emptyset$, then

$$\begin{aligned}
(\mathcal{P}(f_1) + \mathcal{P}(f_2)) \circ g_{X_1,X_2}(X') &= (\mathcal{P}(f_1) + \mathcal{P}(f_2))(X' \cap X_1) \\
&= \mathcal{P}(f_1)(X' \cap X_1) \\
&= f_1[X' \cap X_1] \\
&= g_{Y_1,Y_2}(f_1[X' \cap X_1] + f_2[X' \cap X_2]) \\
&= g_{Y_1,Y_2}((f_1 + f_2)[X']) \\
&= g_{Y_1,Y_2} \circ \mathcal{P}(f_1 + f_2)(X')
\end{aligned}$$

Note that the equality on the third line holds since $X' \cap X_1$ is non-empty and the same is true for $f_1[X' \cap X_1]$.

If $X' \cap X_1 = \emptyset$, hence $X' \subseteq X_2$, then

$$\begin{aligned}
(\mathcal{P}(f_1) + \mathcal{P}(f_2)) \circ g_{X_1,X_2}(X') &= (\mathcal{P}(f_1) + \mathcal{P}(f_2))(X') \\
&= \mathcal{P}(f_2)(X') \\
&= f_2[X'] \\
&= g_{X_1,X_2}(f_2[X']) \\
&= g_{X_1,X_2}((f_1 + f_2)[X']) \\
&= g_{X_1,X_2} \circ \mathcal{P}(f_1 + f_2)(X')
\end{aligned}$$

compatibility with η: $g_{X_1,X_2} \circ \eta_{X_1+X_2} = \eta_{X_1} + \eta_{X_2}$
Let $x \in X_1 + X_2$. Then, since g_{X_1,X_2} is the identity on singletons:

$$g_{X_1,X_2} \circ \eta_{X_1+X_2}(x) = g_{X_1,X_2}(\{x\}) = \{x\}$$

compatibility with μ: $g_{X_1,X_2} \circ \mu_{X_1+X_2} = (\mu_{X_1} + \mu_{X_2}) \circ g_{TX_1,TX_2} \circ Tg_{X_1,X_2}$
Let $\mathcal{X} \subseteq \mathcal{P}(X_1 + X_2)$. We can split $\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2$, where $\mathcal{X}_1 = \{X' \in \mathcal{X} \mid X' \cap X_1 \neq \emptyset\}$ and $\mathcal{X}_2 = \{X' \in \mathcal{X} \mid X' \cap X_1 = \emptyset\} \subseteq \mathcal{P}(X_2)$. Note that $\mathcal{X}_1 \cap X_1 = \{X' \cap X_1 \mid X' \in \mathcal{X}_1\} \subseteq \mathcal{P}(X_1)$. We consider two cases: if $\mathcal{X}_1 \neq \emptyset$, we have that $\bigcup \mathcal{X}$ contains at least one element from X_1 and hence:

$$\begin{aligned}
g_{X_1,X_2} \circ \mu_{X_1+X_2}(\mathcal{X}) &= g_{X_1,X_2}(\bigcup \mathcal{X}) \\
&= \left(\bigcup \mathcal{X}\right) \cap X_1 \\
&= \bigcup (\mathcal{X} \cap X_1) \\
&= \bigcup (\mathcal{X}_1 \cap X_1) \\
&= (\mu_{X_1} + \mu_{X_2})(\mathcal{X}_1 \cap X_1) \\
&= (\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{P}(X_1),\mathcal{P}(X_2)}((\mathcal{X}_1 \cap X_1) + \mathcal{X}_2) \\
&= (\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{P}(X_1),\mathcal{P}(X_2)} \circ g_{X_1,X_2}[\mathcal{X}_1 + \mathcal{X}_2] \\
&= (\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{P}(X_1),\mathcal{P}(X_2)} \circ \mathcal{P}(g_{X_1,X_2})(\mathcal{X})
\end{aligned}$$

If instead $\mathcal{X}_1 = \emptyset$, we have that $\bigcup \mathcal{X}$ contains no element from X_1 and $\mathcal{X} = \mathcal{X}_2$. Then:

$$\begin{aligned}
g_{X_1,X_2} \circ \mu_{X_1+X_2}(\mathcal{X}) &= g_{X_1,X_2} \circ \mu_{X_1+X_2}(\mathcal{X}_2) \\
&= g_{X_1,X_2}(\bigcup \mathcal{X}_2) \\
&= \bigcup \mathcal{X}_2 \\
&= (\mu_{X_1} + \mu_{X_2})(\mathcal{X}_2) \\
&= (\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{P}(X_1),\mathcal{P}(X_2)}(\mathcal{X}_2) \\
&= (\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{P}(X_1),\mathcal{P}(X_2)} \circ g_{X_1,X_2}[\mathcal{X}_2] \\
&= (\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{P}(X_1),\mathcal{P}(X_2)} \circ \mathcal{P}(g_{X_1,X_2})(\mathcal{X})
\end{aligned}$$

well-behavedness: First, note that here $\top = 0$ and $\bot = 1$.

- $[\sup \circ \mathcal{P} f_1, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2} = \sup \circ \mathcal{P}[f_1, \top_{X_2}]$:
  Let $X' \subseteq X_1 + X_2$. We distinguish three subcases.
  - If $X' \cap X_2 \neq \emptyset$, then

$$\begin{aligned}
([\sup \circ \mathcal{P} f_1, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2})(X') &= (\sup \circ \mathcal{P} f_1)(X' \cap X_1) \\
&= \sup f_1[X' \cap X_1] \\
&= \sup\{f_1[X' \cap X_1] \cup \top_{X_2}[X' \cap X_2]\} \\
&= \sup\left(\mathcal{P}[f_1, \top_{X_2}](X')\right)
\end{aligned}$$

  - If $X' = \emptyset$, then

$$\begin{aligned}
([\sup \circ \mathcal{P} f_1, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2})(X') &= \top_{\mathcal{P}X_2}(\emptyset) \\
&= \top \\
&= \sup(\emptyset) \\
&= \sup([\mathcal{P} f_1, \top_{X_2}](\emptyset)) \\
&= (\sup \circ [\mathcal{P} f_1, \top_{X_2}])(X')
\end{aligned}$$

- $X' \cap X_1 = \emptyset$ and $X' \cap X_1 \neq \emptyset$, hence $X' \subseteq X_2$. We can assume that $X' \neq \emptyset$.

$$\begin{aligned}([\sup \circ \mathcal{P} f_1, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2})(X') &= \top_{\mathcal{P}X_2}(X' \cap X_2)\\ &= \top\\ &= \sup(\top_{X_2}(X' \cap X_2))\\ &= \sup([\mathcal{P} f_1, \top_{X_2}](X' \cap X_2))\\ &= (\sup \circ [\mathcal{P} f_1, \top_{X_2}])(X')\end{aligned}$$

- $[\perp_{\mathcal{P}X_1}, \sup \circ \mathcal{P} f_2] \circ g_{X_1,X_2} = \sup \circ \mathcal{P}[\perp_{X_1}, f_2]$:
  Let $X' \subseteq X_1 + X_2$. We distinguish two subcases.
  - If $X' \cap X_1 \neq \emptyset$, then:

$$\begin{aligned}([\perp_{\mathcal{P}X_1}, \sup \circ \mathcal{P} f_2] \circ g_{X_1,X_2})(X') &= \perp_{\mathcal{P}X_1}(X' \cap X_2)\\ &= \perp\\ &= \sup(\perp_{X_1}[X' \cap X_1] \cup f_2[X' \cap X_2])\\ &= (\sup \circ \mathcal{P}[\perp_{X_1}, f_1])(X')\end{aligned}$$

  - If $X' \cup X_1 = \emptyset$, hence $X' \subseteq X_2$

$$\begin{aligned}([\perp_{\mathcal{P}X_1}, \sup \circ \mathcal{P} f_2] \circ g_{X_1,X_2})(X') &= (\sup \circ \mathcal{P} f_2)(X')\\ &= \sup(f_2[X'])\\ &= \sup(\perp_{X_1}[X' \cap X] \cup f_2[X' \cap X_2])\\ &= (\sup \circ \mathcal{P}[\perp_{X_1}, f_2])(X')\end{aligned}$$

- $[\perp_{\mathcal{P}X_1}, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2} = \sup \circ \mathcal{P}[\perp_{X_1}, \top_{X_2}]$: Let $X' \subseteq X_1 + X_2$. We distinguish three subcases:
  - If $X' \cap X_1 \neq \emptyset$, then

$$\begin{aligned}([\perp_{\mathcal{P}X_1}, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2})(X') &= \perp_{\mathcal{P}X_1}(X' \cap X_1)\\ &= \perp\\ &= \sup(\perp_{X_1}[X' \cap X_1])\\ &= \sup(\perp_{X_1}[X' \cap X_1] \cup \top_{X_2}[X' \cap X_2])\\ &= (\sup \circ \mathcal{P}[\perp_{X_1}, \top_{X_2}])(X')\end{aligned}$$

  - If $X' = \emptyset$, then

$$\begin{aligned}([\perp_{\mathcal{P}X_1}, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2})(X') &= \top_{\mathcal{P}X_2}(\emptyset)\\ &= \top\\ &= \sup(\emptyset)\\ &= (\sup \circ \mathcal{P}[\perp_{X_1}, \top_{X_2}])(\emptyset)\\ &= (\sup \circ \mathcal{P}[\perp_{X_1}, \top_{X_2}])(X')\end{aligned}$$

- If $X' \cap X_2 \neq \emptyset$, hence $X' \subseteq X_2$. We can assume that $X' \neq \emptyset$.

$$\begin{aligned}([\bot_{\mathcal{P}X_1}, \top_{\mathcal{P}X_2}] \circ g_{X_1,X_2})(X') &= \top_{\mathcal{P}X_2}(X' \cap X_2)\\ &= \top\\ &= \sup(\top_{X_2}(X' \cap X_2))\\ &= \sup(\bot_{X_1}[X' \cap X_1] \cup \top_{X_2}[X' \cap X_2])\\ &= (\sup \circ \mathcal{P}[\bot_{X_1}, \top_{X_2}])(X')\end{aligned}$$

◀

▶ **Proposition 23.** *Let $T = \mathcal{S}$ be the subdistribution monad where $\mathcal{S}(X) = \{p\colon X \to [0,1] \mid \sum_{x\in X} p(x) \leq 1\}$. Assume that its evaluation map is $ev_T = \mathbb{E}$ for quantale $\mathcal{V} = [0,\infty]_\oplus$ (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.*

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

Proof. We have to prove the following properties of g:

naturality: let $f_i\colon X_i \to Y_i$, $p \in \mathcal{S}(X_1 + X_1)$. We distinguish the following cases: if $\mathrm{supp}(p) \cap X_1 \neq \emptyset$, then for all $y \in Y_1$ we have that

$$\begin{aligned}((\mathcal{S}(f_1) + \mathcal{S}(f_2)) \circ g_{X_1,X_2}(p))\,(y) &= ((\mathcal{S}(f_1) + \mathcal{S}(f_2))(p|_{X_1}))\,(y)\\ &= (\mathcal{S}(f_1)(p|_{X_1}))\,(y)\\ &= \sum_{y=f_1(x)} p|_{X_1}(x)\\ &= \sum_{y=(f_1+f_2)(x)} p|_{X_1}(x)\\ &= (\mathcal{S}(f_1 + f_2)(p|_{X_1}))\,(y)\\ &= (\mathcal{S}(f_1 + f_2)(p))\,(y)\\ &= (\mathcal{S}(f_1 + f_2)(p))|_{Y_1}(y)\\ &= (g_{Y_1,Y_2} \circ \mathcal{S}(f_1 + f_2)(p))\,(y)\end{aligned}$$

For the remaining case, we have that for all $y \in Y_2$

$$\begin{aligned}((\mathcal{S}(f_1) + \mathcal{S}(f_2)) \circ g_{X_1,X_2}(p))\,(y) &= \mathcal{S}(f_2)(p|_{X_2})(y)\\ &= \sum_{y=f_2(x)} p|_{X_2}(x)\\ &= \sum_{y=(f_1+f_2)(x)} p_{X_2}(x)\\ &= (\mathcal{S}(f_1 + f_2)(p|_{X_2}))\,(y)\\ &= (\mathcal{S}(f_1 + f_2)(p))\,(y)\\ &= (\mathcal{S}(f_1 + f_2)(p))|_{Y_2}(y)\\ &= (g_{Y_1,Y_2} \circ \mathcal{S}(f_1 + f_2)(p))\,(y)\end{aligned}$$

compatibility with η: $g_{X_1,X_2} \circ \eta_{X_1+X_2} = \eta_{X_1} + \eta_{X_2}$. Since g_{X_1,X_2} is an identity when applied to Dirac distributions, we have that:

$$g_{X_1,X_2} \circ \eta_{X_1+X_2}(x) = g_{X_1,X_2}(\delta_x) = \delta_x = \eta_{X_1}(x)$$

compatibility with μ: $g_{X_1,X_2} \circ \mu_{X_1+X_2} = (\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{S}X_1,\mathcal{S}X_2} \circ \mathcal{S}g_{X_1,X_2}$. Let $p \in \mathcal{S}\mathcal{S}(X_1 + X_2)$. Note that we can partition $\operatorname{supp}(p) = \mathcal{X}_1 \cup \mathcal{X}_2$ in a way that $\mathcal{X}_1 = \{\nu \in \mathcal{S}(X_1+X_2) \mid \operatorname{supp}(\nu) \cap X_1 \neq \emptyset\}$ and $\mathcal{X}_2 = \{\nu \in \mathcal{S}(X_1 + X_2) \mid \operatorname{supp}(\nu) \cap X_1 = \emptyset\}$. We consider two cases: if $\mathcal{X}_1 \neq \emptyset$, then $\operatorname{supp}(\mu_{X_1+X_2}(p)) \cap X_1 \neq \emptyset$ and hence for all $x \in X_1$, we have that:

$$\begin{aligned}
((g_{X_1,X_2} \circ \mu_{X_1+X_2})(p))\,(x) &= (\mu_{X_1+X_2}(p))|_{X_1}(x) \\
&= \sum_{\nu \in \mathcal{S}(X_1+X_2)} p(\nu)\nu|_{X_1}(x) \\
&= \sum_{\nu \in \mathcal{X}_1} p(\nu)\nu|_{X_1}(x) \\
&= \sum_{\nu \in \mathcal{X}_1} \left(\mathcal{S}(g_{X_1,X_2})(p)\right)(\nu)\nu(x) \\
&= \sum_{\nu \in \mathcal{S}(X_1+X_2)} \left(\mathcal{S}g_{X_1,X_2}(p)|_{\mathcal{S}X_1}\right)(\nu)\nu(x) \\
&= \left((\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{S}X_1,\mathcal{S}X_2} \circ \mathcal{S}g_{X_1,X_2}(p)\right)(x)
\end{aligned}$$

Now, consider the case when $\operatorname{supp}(\mu_{X_1+X_2}(p)) \cap X_1 = \emptyset$. For all $x \in X_2$ we have that:

$$\begin{aligned}
((g_{X_1,X_2} \circ \mu_{X_1+X_2})(p))\,(x) &= (\mu_{X_1+X_2}(p))|_{X_2}(x) \\
&= \sum_{\nu \in \mathcal{S}(X_1+X_2)} p(\nu)\nu|_{X_2}(x) \\
&= \sum_{\nu \in \mathcal{X}_2} p(\nu)\nu|_{X_2}(x) \\
&= \sum_{\nu \in \mathcal{X}_2} p(\nu)\nu(x) \\
&= \sum_{\nu \in \mathcal{X}_2} \left(\mathcal{S}(g_{X_1,X_2})(p)\right)(\nu)\nu(x) \\
&= \sum_{\nu \in \mathcal{S}(X_1+X_2)} \left(\mathcal{S}(g_{X_1,X_2})(p)|_{\mathcal{S}X_2}\right)(\nu)\nu(x) \\
&= \sum_{\nu \in \mathcal{S}(X_1+X_2)} \left(\mathcal{S}(g_{X_1,X_2})(p)|_{\mathcal{S}X_2}\right)(\nu)\nu(x) \\
&= \left((\mu_{X_1} + \mu_{X_2}) \circ g_{\mathcal{S}X_1,\mathcal{S}X_2} \circ \mathcal{S}g_{X_1,X_2}(p)\right)(x)
\end{aligned}$$

well-behavedness: First, note that here $\top = 0$ and $\bot = \infty$.

- $[\mathbb{E} \circ \mathcal{S}f_1, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2} = \mathbb{E} \circ \mathcal{S}[f_1, \top_{X_2}]$: Let $p \in S(X_1 + X_2)$. We first consider the case when $\mathrm{supp}(p) \cap X_1 \neq \emptyset$. We have that:

$$\begin{aligned}
[\mathbb{E} \circ \mathcal{S}f_1, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2}(p) &= [\mathbb{E} \circ \mathcal{S}f_1, \top_{\mathcal{S}X_2}](p|_{X_1}) \\
&= \mathbb{E} \circ \mathcal{S}f_1(p|_{X_1}) \\
&= \sum_{y \in \mathcal{V}} y \left(\sum_{y = f_1(x)} p(x) \right) \\
&= \left(\sum_{y \in \mathcal{V}} y \left(\sum_{y = f_1(x)} p(x) \right) \right) + 0 \left(\sum_{0 = \top_{X_2}(x)} p(x) \right) \\
&= \sum_{y \in \mathcal{V}} y \left(\sum_{y = [f_1, \top_{X_2}](x)} p(x) \right) \\
&= \mathbb{E} \circ \mathcal{S}[f_1, \top_{X_2}]
\end{aligned}$$

Now, consider the case when $\mathrm{supp}(p) = \emptyset$. We have that:

$$\begin{aligned}
[\mathbb{E} \circ \mathcal{S}f_1, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2}(p) &= [\mathbb{E} \circ \mathcal{S}f_1, \top_{\mathcal{S}X_2}](p|_{X_2}) \\
&= \top \\
&= 0 \\
&= \mathbb{E} \circ \mathcal{S}[f_1, \top_{X_2}](p)
\end{aligned}$$

Finally, consider the case when $\mathrm{supp}(p) \cap X_1 = \emptyset$, but $\mathrm{supp}(p) \neq \emptyset$. We have that:

$$\begin{aligned}
[\mathbb{E} \circ \mathcal{S}f_1, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2}(p) &= [\mathbb{E} \circ \mathcal{S}f_1, \top_{\mathcal{S}X_2}](p|_{X_2}) \\
&= \top \\
&= 0 \\
&= \sum_{y \in \mathcal{V}} y \left(\sum_{y = \top_{X_2}(x)} p(x) \right) \\
&= \mathbb{E} \circ \mathcal{S}(\top_{X_2})(p) \\
&= \mathbb{E} \circ \mathcal{S}[f_1, \top_{X_2}](p)
\end{aligned}$$

- $[\perp_{\mathcal{S}X_1}, \mathbb{E} \circ \mathcal{S}f_2] \circ g_{X_1,X_2} = \mathbb{E} \circ \mathcal{S}[\perp_{X_1}, f_2]$: Let $p \in S(X_1 + X_2)$. Firstly, consider the case when $\mathrm{supp}(p) \cap X_1 \neq \emptyset$. In such a case, we have

$$\begin{aligned}
[\perp_{\mathcal{S}X_1}, \mathbb{E} \circ \mathcal{S}f_2] \circ g_{X_1,X_2}(p) &= [\perp_{\mathcal{S}X_1}, \mathbb{E} \circ \mathcal{S}f_2](p|_{X_1}) \\
&= \perp \\
&= \infty \\
&= \infty + \left(\sum_{y \in \mathbb{R}} y \left(\sum_{y = [\perp_{X_1}, f_2](x)} p(x) \right) \right) \\
&= \left(\sum_{\infty = [\perp_{X_1}, f_2](x)} p(x) \right) + \left(\sum_{y \in \mathbb{R}} y \left(\sum_{y = [\perp_{X_1}, f_2](x)} p(x) \right) \right) \\
&= \sum_{y \in [0, +\infty]} y \left(\sum_{y = [\perp_{X_1}, f_2](x)} p(x) \right) \\
&= \mathbb{E} \circ \mathcal{S}[\perp_{X_1}, f_2](p)
\end{aligned}$$

If $\mathrm{supp}(p) = \emptyset$, then we have that

$$\begin{aligned}
[\bot_{\mathcal{S}X_1}, \mathbb{E} \circ \mathcal{S}f_2] \circ g_{X_1,X_2}(p) &= [\bot_{\mathcal{S}X_1}, \mathbb{E} \circ \mathcal{S}f_2](p|_{X_2}) \\
&= \mathbb{E} \circ \mathcal{S}f_2(p|_{X_2}) \\
&= 0 \\
&= \mathbb{E} \circ \mathcal{S}[\bot_{X_1}, f_2](p)
\end{aligned}$$

Finally, consider the case when $\mathrm{supp}(p) \cap X_1 = \emptyset$, but $\mathrm{supp}(p) \neq \emptyset$. Then, we have that

$$\begin{aligned}
[\bot_{\mathcal{S}X_1}, \mathbb{E} \circ \mathcal{S}f_2] \circ g_{X_1,X_2}(p) &= [\bot_{\mathcal{S}X_1}, \mathbb{E} \circ \mathcal{S}f_2](p|_{X_2}) \\
&= \mathbb{E} \circ \mathcal{S}f_2(p|_{X_2}) \\
&= \sum_{y \in \mathcal{V}} y \left(\sum_{y = f_2(x)} p|_{X_2}(x) \right) \\
&= \sum_{y \in \mathcal{V}} y \left(\sum_{y = [\bot_{X_1}, f_2](x)} p(x) \right) \\
&= \mathbb{E} \circ \mathcal{S}[\bot_{X_1}, f_2](p)
\end{aligned}$$

- $[\bot_{\mathcal{S}X_1}, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2} = \mathbb{E} \circ \mathcal{S}[\bot_{X_1}, \top_{X_2}]$: Let $p \in \mathcal{S}(X_1 + X_2)$. First, consider the case when $\mathrm{supp}(p) \cap X_1 \neq \emptyset$. We have that

$$\begin{aligned}
[\bot_{\mathcal{S}X_1}, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2}(p) &= [\bot_{\mathcal{S}X_1}, \top_{\mathcal{S}X_2}](p|_{X_1}) \\
&= \bot_{\mathcal{S}X_1}(p|_{X_1}) \\
&= \infty \\
&= \infty + 0 \\
&= \left(\sum_{\infty = [\bot_{X_1}, \top_{X_2}](x)} p(x) \right) + \left(\sum_{0 = [\bot_{X_1}, \top_{X_2}](x)} p(x) \right) \\
&= \sum_{y \in [0, +\infty]} y \left(\sum_{y = [\bot_{X_1}, \top_{X_2}](x)} p(x) \right) \\
&= \mathbb{E} \circ \mathcal{S}[\bot_{X_1}, \top_{X_2}](p)
\end{aligned}$$

Then, consider a case when $\mathrm{supp}(p) = \emptyset$.

$$\begin{aligned}
[\bot_{\mathcal{S}X_1}, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2}(p) &= [\bot_{\mathcal{S}X_1}, \top_{\mathcal{S}X_2}](p|_{X_2}) \\
&= \top_{\mathcal{S}X_2}(p|_{X_2}) \\
&= 0 \\
&= \mathbb{E} \circ \mathcal{S}[\bot_{X_1}, \top_{X_2}](p)
\end{aligned}$$

Finally, consider a case when $\mathrm{supp}(p) \cap X_1 = \emptyset$ and $\mathrm{supp}(p) \cap X_2 \neq \emptyset$. We have that

$$\begin{aligned}[\bot_{\mathcal{S}X_1}, \top_{\mathcal{S}X_2}] \circ g_{X_1,X_2}(p) &= [\bot_{\mathcal{S}X_1}, \top_{\mathcal{S}X_2}](p|_{X_2}) \\ &= \top_{\mathcal{S}X_2}(p|_{X_2}) \\ &= 0 \\ &= \left(\sum_{y=[\bot_{X_1},\top_{X_2}](x)} p(x) \right) \\ &= \mathbb{E} \circ \mathcal{S}[\bot_{X_1}, \top_{X_2}](p)\end{aligned}$$

◀

E Proofs for Section 6 (Case Studies)

We will now show a property of the up-to function u that is closely related to the fact that under certain conditions the Kantorovich lifting is always bounded by the Wasserstein lifting, as shown in [2]. The requirement is part of the notion of a well-behaved evaluation function introduced there.

▶ **Proposition 35.** *Let (Y, a, c) be a bialgebra for F, T, ev_T an evaluation map for T and let u be the up-to function (as in Section 5.1). Furthermore assume that $: Y \times Y \to \mathcal{V}$ is an object of $\mathcal{V}$-*Graph *and $t \in T(Y \times Y)$. Then:*

$$u(d)(a(T\pi_1(t), a(T\pi_2(t))) \sqsupseteq ev_T \circ Td(t)$$

whenever

$$d_{\mathcal{V}} \circ (ev_T \times ev_T) \circ \langle T\pi_1', T\pi_2' \rangle \sqsupseteq ev_T \circ Td_{\mathcal{V}}. \tag{4}$$

where $\pi_i \colon Y \times Y \to Y$ and $\pi_i' \colon \mathcal{V} \times \mathcal{V} \to \mathcal{V}$.

Proof.

$$\begin{aligned}u(d)(a(T\pi_1(t), a(T\pi_2(t))) &= \bigsqcup_{a(t_i)=a(T\pi_i(t))} \overline{T}(d)(t_1, t_2) \\ &\sqsupseteq \overline{T}(d)(T\pi_1(t), T\pi_1(t)) \\ &= \bigsqcap_{f \in \gamma_Y(d)} d_{\mathcal{V}}(ev_T \circ Tf \circ T\pi_1(t), ev_T \circ Tf \circ T\pi_2(t)) \\ &\sqsupseteq ev_T \circ Td(t)\end{aligned}$$

The last inequality follows directly from [2, Lemma 5.25] whenever Condition (4) holds. Note that in [2] the quantale is $\mathcal{V}$ and $d_{\mathcal{V}}$ (called d_e there) is assumed to be symmetric, but neither assumption is used in the proof. ◀

▶ **Corollary 36.** *In the setting of Proposition 35 assume that $Y = TX$, $a = \mu_X$ and $\mathcal{V} = [0, 1]$. Furthermore consider the monad $T = \mathcal{P}$ ($T = \mathcal{D}$). Then the evaluation map $ev_T = \sup$ ($ev_T = \mathbb{E}$) satisfies Condition (4) and hence we obtain:*

- *$T = \mathcal{P}$:*

$$u(d)(\bigcup_{i\in I} X_i, \bigcup_{i\in I} Y_i) \leq \sup_{i\in I} d(X_i, Y_i)$$

whenever $X_i, Y_i \subseteq X$, $i \in I$.

- $T = \mathcal{D}$:

$$u(d)(\sum_{i\in I} r_i \cdot p_i, \sum_{i\in I} r_i \cdot q_i) \leq \sum_{i\in I} r_i \cdot d(p_i, q_i)$$

whenever $p_i, q_i \in \mathcal{D}(X)$, $i \in I$.

Proof. We first prove Condition (4) in both cases ($T = \mathcal{P}$, $T = \mathcal{D}$):

- $T = \mathcal{P}$: Let $t \in T(\mathcal{V} \times \mathcal{V})$, i.e., $t \subseteq [0,1] \times [0,1]$ respectively $t = \{(s_i, r_i) \mid i \in I, r_i, s_i \in [0,1]\}$. Then

$$\begin{aligned} d_{\mathcal{V}} \circ (ev_T \times ev_T) \circ \langle T\pi_1', T\pi_2' \rangle(t) &= d_{\mathcal{V}}(ev_T(T\pi_1'(t)), ev_T(T\pi_2'(t))) \\ &= \sup_{i\in I} r_i \ominus \sup_{i\in I} s_i \\ &\leq \sup_{i\in I}(r_i \ominus s_i) \\ &= ev_T \circ T d_{\mathcal{V}}(t). \end{aligned}$$

- $T = \mathcal{D}$: Let $t \in T(\mathcal{V} \times \mathcal{V})$, i.e., $t\colon [0,1] \times [0,1] \to [0,1]$ is a probability distribution respectively $t = \sum_{i\in I} p_i \cdot (s_i, r_i)$, written as a formal sum where $p_i, r_i, s_i \in [0,1]$ and $\sum_{i\in I} p_i = 1$. Then

$$\begin{aligned} d_{\mathcal{V}} \circ (ev_T \times ev_T) \circ \langle T\pi_1', T\pi_2' \rangle(t) &= d_{\mathcal{V}}(ev_T(T\pi_1'(t)), ev_T(T\pi_2'(t))) \\ &= \sum_{i\in I} p_i \cdot r_i \ominus \sum_{i\in I} p_i \cdot s_i \\ &\leq \sum_{i\in I} p_i \cdot (r_i \ominus s_i) \\ &= ev_T \circ T d_{\mathcal{V}}(t). \end{aligned}$$

Finally, we spell out the inequalities:

- $T = \mathcal{P}$: Let $t \in T(TX \times TX)$, i.e., $t \subseteq \mathcal{P}(X) \times \mathcal{P}(X)$ respectively $t = \{(X_i, Y_i) \mid i \in I, X_i, Y_i \subseteq X\}$. Then

$$\begin{aligned} u(d)(\bigcup_{i\in I} X_i, \bigcup_{i\in I} Y_i) &= u(d)(\mu_X(T\pi_1(t)), \mu_X(T\pi_2(t))) \\ &\leq ev_T \circ Td(t) \\ &= \sup_{i\in I} d(X_i, Y_i) \end{aligned}$$

- $T = \mathcal{D}$: Let $t \in T(TX \times TX)$, i.e., $t\colon \mathcal{D}(X) \times \mathcal{D}(X) \to [0,1]$ is a probability distribution respectively $t = \sum_{i\in I} r_i \cdot (p_i, q_i)$, written as a formal sum where $r_i \in [0,1]$ with $\sum_{i\in I} r_i = 1$, $p_i, q_i \in \mathcal{D}(X)$. Then

$$\begin{aligned} u(d)(\sum_{i\in I} r_i \cdot p_i, \sum_{i\in I} r_i \cdot q_i) &= u(d)(\mu_X(T\pi_1(t)), \mu_X(T\pi_2(t))) \\ &\leq ev_T \circ Td(t) \\ &= \sup_{i\in I} r_i \cdot d(p_i, q_i) \end{aligned}$$

◀