


PAPER

# Up-to techniques for behavioural metrics via fibrations

Filippo Bonchi<sup>1</sup>, Barbara König<sup>2</sup>  and Daniela Petrişan<sup>3</sup>

<sup>1</sup>Università di Pisa, Pisa, Italy, <sup>2</sup>Universität Duisburg-Essen, Duisburg, Germany, <sup>3</sup>CNRS, IRIF, Université Paris Diderot, Paris, France

**Corresponding author:** Barbara König; Email: [barbara\\_koenig@uni-due.de](mailto:barbara_koenig@uni-due.de)

(Received 18 October 2021; revised 16 May 2023; accepted 18 May 2023; first published online 10 July 2023)

## Abstract

Up-to techniques are a well-known method for enhancing coinductive proofs of behavioural equivalences. We introduce up-to techniques for behavioural metrics between systems modelled as coalgebras, and we provide abstract results to prove their soundness in a compositional way.

In order to obtain a general framework, we need a systematic way to lift functors: we show that the Wasserstein lifting of a functor, introduced in a previous work, corresponds to a change of base in a fibrational sense. This observation enables us to reuse existing results about soundness of up-to techniques in a fibrational setting. We focus on the fibrations of predicates and relations valued in a quantale. To illustrate our approach, we provide an example on distances between regular languages.

**Keywords:** behavioural metrics; bisimilarity; up-to techniques; coalgebras; fibrations

## 1. Introduction

Checking whether two systems have an equivalent (or similar) behaviour is a crucial problem in computer science. In concurrency theory, one standard methodology for establishing behavioural equivalence of two systems is constructing a bisimulation relation between them. When the systems display a quantitative behaviour, the notion of behavioural equivalence is replaced with the more robust notion of behavioural metric (de Alfaro et al. 2004; Desharnais et al. 2004; van Breugel and Worrell 2005).

Due to the sheer complexity of state-based systems, computing their behavioural equivalences and metrics can be very costly; therefore, optimization techniques – the so-called *up-to techniques* (Pous and Sangiorgi 2019) – have been developed to render these computations more efficient. These techniques are related to Galois connections (Baldan et al. 2020; Bonchi et al. 2018a) and can be found applications in various domains such as checking algorithms (Bonchi et al. 2017a; Bonchi and Pous 2013), abstract interpretation (Bonchi et al. 2018a) and proof assistants (Blanchette et al. 2017; Danielsson 2017). In the qualitative setting and in particular in concurrency, the theory of up-to techniques for bisimulations and various other coinductive predicates has been thoroughly studied (Bacci et al. 2021; Hur et al. 2013; Milner and Sangiorgi 1992; Pous and Sangiorgi 2011). On the other hand, in the quantitative setting, so far, there are only few papers that investigate up-to techniques for behavioural metrics, in particular (Chatzikokolakis et al. 2016). There, the notion of up-to techniques and the accompanying theory of soundness are specific for probabilistic automata and are not derived as instances of the standard lattice-theoretic framework, which we will briefly recall next.

Suppose we want to verify whether two states in a system behave in the same way (e.g. whether two states of an NFA accept the same language). The starting observation is that the relation of interest (e.g. behavioural equivalence or language equivalence) can be expressed as the greatest fixed point  $\nu b$  of a monotone function  $b: \text{Rel}_Q \rightarrow \text{Rel}_Q$  on the complete lattice  $\text{Rel}_Q$  of relations on the state space  $Q$  of the system. Hence, in order to prove that two states  $x$  and  $y$  are behaviourally equivalent, that is,  $(x, y) \in \nu b$ , it suffices to find a witness relation  $r$  which on one hand is a post-fixpoint of  $b$ , that is,  $r \subseteq b(r)$  and on the other hand contains the pair  $(x, y)$ . This is simply the coinduction proof principle. However, exhibiting such a witness relation  $r$  can be sometimes computationally expensive. In many situations, this computation can be significantly optimized, if instead of computing a post-fixpoint of  $b$  one exhibits a relaxed invariant, that is a relation  $r$  such that  $r \subseteq b(f(r))$  for a suitable function  $f$ . The function  $f$  is called a *sound up-to* technique when the proof principle

$$\frac{(x, y) \in r \quad r \subseteq b(f(r))}{(x, y) \in \nu b}$$

is valid. Establishing the soundness of up-to techniques on a case-by-case basis can be a tedious and sometimes delicate problem, see for example Milner (1989). For this reason, several works (Hur et al. 2013; Parrow and Weber 2016; Pous 2007, 2016; Pous and Sangiorgi 2011; Sangiorgi 1998) have established a lattice-theoretic framework for proving soundness results in a modular fashion. The key notion is compatibility: for arbitrary monotone maps  $b$  and  $f$  on a complete lattice  $(C, \leq)$ , the up-to technique  $f$  is *b-compatible* iff  $f \circ b \leq b \circ f$ . Compatible techniques are sound and, most importantly, can be combined in several useful ways.

In this paper, we develop a generic theory of up-to techniques for behavioural metrics applicable to different kinds of systems and metrics, which reuses established methodology. To achieve this, we exploit the theory developed in Bonchi et al. (2014) by modelling systems as *coalgebras* (Jacobs 2016; Rutten 2000) and behavioural metrics as coinductive predicates in a *fibration* (Hermida and Jacobs 1998). In order to provide general soundness results, we need a principled way to lift functors from Set to metric spaces, a problem that has been studied in Hofmann (2007), Hofmann and Nora (2020), Baldan et al. (2014) and Wild and Schröder (2022). Our key observation is that these liftings arise from a change-of-base situation between  $\mathcal{V}$ -Rel and  $\mathcal{V}$ -Pred, namely the fibrations of relations, respectively predicates, valued over a quantale  $\mathcal{V}$  (see Sections 4 and 5).

In Section 6, we provide sufficient conditions ensuring the compatibility of basic quantitative up-to techniques, as well as proper ways to compose them. Interestingly enough, the conditions ensuring compatibility of the quantitative analogue of up-to reflexivity and up-to transitivity are subsumed by those used in Hofmann (2007) to extend monads to a bicategory of many-valued relations and generalize those in Baldan et al. (2014) (see the discussion after Theorem 27).

When the state space of a system is equipped with an algebraic structure, for example in process algebras, one can usually exploit this structure by reasoning up-to context. Assuming that the system forms a *bialgebra* (Klin 2011; Turi and Plotkin 1997), that is, that the algebraic structure distributes over the coalgebraic behaviour as in GSOS specifications, we give sufficient conditions ensuring the compatibility of the quantitative version of contextual closure (Theorem 42).

In the qualitative setting, the sufficient conditions for compatibility are automatically met when taking as lifting the *canonical* relational one, see Bonchi et al. (2014). We show that the situation is similar in the quantitative setting for a certain notion of *quantitative canonical* lifting. In particular, up-to context is compatible for the canonical lifting under mild assumptions (Theorem 46). As an immediate corollary, we have that, in a bialgebra, syntactic contexts are *non-expansive* with respect to the behavioural metric induced by the canonical lifting. This property and weaker variants of it (such as non-extensiveness or uniform continuity), considered to be the quantitative analogue of behavioural equivalence being a congruence, have recently received considerable attention, see for example Desharnais et al. (2004), Bacci et al. (2013), Tini et al. (2017).

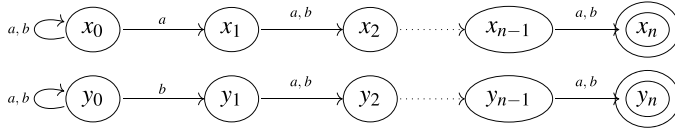


Figure 1. Example automaton.

To fix the intuition, Section 2 provides a motivating example (formally treated in Section 7). Section 3 contains the preliminaries, while quantale-valued predicates and relations are introduced and studied in Section 4. In the following, Section 5 discusses the lifting of functors to such generalized predicates and relations, which paves the way for quantitative up-to techniques, see Section 6. As mentioned above, Section 7 formally works out the example and Section 8 provides a detailed comparison with Chatzikokolakis et al. (2016). We conclude with a comparison to related work and a discussion of open problems in Section 9.

Additional proofs are provided in the appendix.

**2. Motivating Example: Distances Between Regular Languages**

Computing various distances (such as the edit-distance or Cantor metric) between strings, and more generally between regular languages or string distributions, has found various practical applications in various areas such as speech and handwriting recognition or computational biology. In this section, we focus on a simple distance between regular languages, which we will call *shortest-distinguishing-word-distance* and is defined as  $d_{sdw}(L, K) = c^{|w|}$  – where  $w$  is the shortest word which belongs to exactly one of the languages  $L, K$  and  $c$  is a constant such that  $0 < c < 1$ .

As a running example, which will be formally explained in Section 7, we consider the non-deterministic finite automaton in Fig. 1 and the languages accepted by the states  $x_0$ , respectively  $y_0$ . We can similarly define a distance on the states of an automaton as the aforementioned distance between the languages accepted by the two states. The inequality

$$d_{sdw}(x_0, y_0) \leq c^n \quad (\text{even } d_{sdw}(x_0, y_0) = c^n) \tag{1}$$

holds in this example since no word of length smaller than  $n$  is accepted by either state. Note that computing this distance is PSPACE-hard since the language equivalence problem for non-deterministic automata can be reduced to it.

One way to show the bound is to determinize the automaton in Fig. 1 and to use the fact that for deterministic automata the shortest-distinguishing-word-distance can be expressed as the greatest fixpoint for a monotone function. Indeed, for a finite deterministic automaton  $(Q, (\delta_a : Q \rightarrow Q)_{a \in A}, F \subseteq Q)$  over a finite alphabet  $A$ , we have that  $d_{sdw} : Q \times Q \rightarrow [0, 1]$  is the greatest fixpoint of a function  $b$  defined on the complete lattice  $[0, 1]^{Q \times Q}$  of functions ordered with the *reversed* pointwise order  $\leq$  and given by

$$b(d)(q_1, q_2) = \begin{cases} 1, & \text{if only one of } q_1, q_2 \text{ is in } F \\ \max_{a \in A} c \cdot \{d(\delta_a(q_1), \delta_a(q_2))\}, & \text{otherwise} \end{cases} \tag{2}$$

Notice that we use the reversed order on  $[0, 1]$ , for technical reasons (see Section 4).

In order to prove (1) we can define a witness distance  $\bar{d}$  on the states of the determinized automaton such that  $\bar{d}(\{x_0\}, \{y_0\}) \leq c^n$  and which is a post-fixpoint for  $b$ , that is,  $\bar{d} \leq b(\bar{d})$ . Notice that this would entail  $\bar{d} \leq d_{sdw}$  and hence  $d_{sdw}(\{x_0\}, \{y_0\}) \leq \bar{d}(\{x_0\}, \{y_0\}) \leq c^n$ .

This approach is problematic since the determinization of the automaton is of exponential size, so we have to define  $\bar{d}$  for exponentially many pairs of sets of states. In order to mitigate the state

space explosion, we will use an up-to technique, which, just as up-to congruence in Bonchi and Pous (2013), exploits the join-semilattice structure of the state set  $\mathcal{P}Q$  of the determinization of an NFA with state set  $Q$ . The crucial observation is the fact that given the states  $Q_1, Q_2, Q'_1, Q'_2 \in \mathcal{P}Q$  in the determinization of an NFA, the following inference rule holds

$$\frac{d_{\text{sdw}}(Q_1, Q_2) \leq r \quad d_{\text{sdw}}(Q'_1, Q'_2) \leq r}{d_{\text{sdw}}(Q_1 \cup Q'_1, Q_2 \cup Q'_2) \leq r}$$

Based on this, we can define a monotone function  $f$  on  $[0, 1]^{\mathcal{P}Q \times \mathcal{P}Q}$  that closes a function  $d$  according to such proof rules (the formal definition of  $f$  is given in Section 7). The general theory developed in this paper allows us to show in Section 7 that  $f$  is a sound up-to technique, that is, it is sufficient to prove  $\bar{d} \leq b(f(\bar{d}))$  (which is easier than  $\bar{d} \leq b(\bar{d})$ , since  $f$  is extensive) in order to establish  $\bar{d} \leq d_{\text{sdw}}$ .

Using this technique, it suffices to consider a quadratic number of pairs of sets of states in the example. In particular we define a function  $\bar{d}: \mathcal{P}Q \times \mathcal{P}Q \rightarrow [0, 1]$  as follows:

$$\bar{d}(\{x_i\}, \{y_j\}) = c^{n - \max\{i, j\}}$$

and  $\bar{d}(X_1, X_2) = 1$  for all other values. Note that this function is not a metric but rather, what we will call in Section 4, a relation valued in  $[0, 1]$ .

It holds that  $\bar{d}(\{x_0\}, \{y_0\}) = c^n$ . It remains to show that  $\bar{d} \leq b(f(\bar{d}))$ . For this, it suffices to prove that

$$b(f(\bar{d}))(\{x_i\}, \{y_j\}) \leq \bar{d}(\{x_i\}, \{y_j\}).$$

For instance, when  $i = j = 0$  we compute the sets of  $a$ -successors, which are  $\{x_0, x_1\}, \{y_0\}$ . We have that  $\bar{d}(\{x_0\}, \{y_0\}) = c^n \leq c^{n-1}$ ,  $\bar{d}(\{x_0\}, \{y_1\}) = c^{n-1}$ , and using the up-to proof rule introduced above, we obtain that  $f(\bar{d})(\{x_0, x_1\}, \{y_0\}) \leq c^{n-1}$ . The same holds for the sets of  $b$ -successors, and since  $x_0$  and  $y_0$  are both non-final, we infer  $b(f(\bar{d}))(\{x_0\}, \{y_0\}) \leq c \cdot c^{n-1} = c^n = \bar{d}(\{x_0\}, \{y_0\})$ . The remaining cases (when  $i \neq 0$  or  $j \neq 0$ ) are analogous.

Our aim is to introduce such proof techniques for behavioural metrics, to make this kind of reasoning precise, not only for this specific example but also for coalgebras in general. Furthermore, we will not limit ourselves to metrics and distances, but we will consider more general relations valued in arbitrary quantales, of which the interval  $[0, 1]$  is an example.

### 3. Preliminaries

We recall here formal definitions for notions such as coalgebras, bialgebras or fibrations.

**Definition 1.** A coalgebra for a functor  $F: \mathcal{C} \rightarrow \mathcal{C}$  or an  $F$ -coalgebra is a morphism  $\gamma: X \rightarrow FX$  for some object  $X$  of  $\mathcal{C}$ , referred to as the carrier of the coalgebra  $\gamma$ . A morphism between two coalgebras  $\gamma: X \rightarrow FX$  and  $\xi: Y \rightarrow FY$  is a morphism  $f: X \rightarrow Y$  such that  $\xi \circ f = Ff \circ \gamma$ . Algebras for the functor  $F$ , or  $F$ -algebras, are defined dually as morphisms of the form  $\alpha: FX \rightarrow X$ .

**Definition 2.** Consider two functors  $F, T$  and a natural transformation  $\zeta: TF \Rightarrow FT$ . A bialgebra for  $\zeta$  is a tuple  $(X, \alpha, \gamma)$  such that  $\alpha: TX \rightarrow X$  is a  $T$ -algebra,  $\gamma: X \rightarrow FX$  is an  $F$ -coalgebra so that the diagram on the left commutes. We call  $\zeta$  the distributive law of the bialgebra  $(X, \alpha, \gamma)$ , even when  $T$  is not a monad.

$$\begin{array}{ccc}
 TX & \xrightarrow{\alpha} & X & \xrightarrow{\gamma} & FX \\
 \downarrow T\gamma & & & & \uparrow F\alpha \\
 TFX & \xrightarrow{\zeta_X} & & & FTX
 \end{array}$$

**Example 3.** The determinization of an NFA can be seen as a bialgebra with  $X = \mathcal{P}Q$ , the algebra  $\mu_Q: \mathcal{P}\mathcal{P}Q \rightarrow \mathcal{P}Q$  given by the multiplication of the powerset monad, a coalgebra for the functor  $F(X) = 2 \times X^A$ , and a distributive law  $\zeta: \mathcal{P}F \rightarrow F\mathcal{P}$  defined for  $M \subseteq 2 \times X^A$  by  $\zeta_X(M) = (\bigvee_{(b,f) \in M} b, [a \mapsto \{f(a) \mid (b,f) \in M\}])$ . See Silva et al. (2013), Jacobs et al. (2015) for more details.

We now introduce the notions of fibration and bifibration.

**Definition 4.** A functor  $p: \mathcal{E} \rightarrow \mathcal{B}$  is called a fibration when for every morphism  $f: X \rightarrow Y$  in  $\mathcal{B}$  and every  $R$  in  $\mathcal{E}$  with  $p(R) = Y$  there exists a map  $\tilde{f}_R: f^*(R) \rightarrow R$  such that  $p(\tilde{f}_R) = f$ , satisfying the following universal property: For all maps  $g: Z \rightarrow X$  in  $\mathcal{B}$  and  $u: Q \rightarrow R$  in  $\mathcal{E}$  sitting above  $fg$  (i.e.  $p(u) = fg$ ), there is a unique map  $v: Q \rightarrow f^*(R)$  such that  $u = \tilde{f}_R v$  and  $p(v) = g$ .

$$\begin{array}{ccc}
 Q & \xrightarrow{\forall u} & R \\
 \exists! v \swarrow & & \downarrow \tilde{f}_R \\
 f^*(R) & \xrightarrow{\tilde{f}_R} & R \\
 \\ 
 Z & \xrightarrow{fg} & Y \\
 g \searrow & & \downarrow f \\
 X & \xrightarrow{f} & Y
 \end{array}$$

For  $X$  in  $\mathcal{B}$ , we denote by  $\mathcal{E}_X$  the fibre above  $X$ , that is, the subcategory of  $\mathcal{E}$  with objects mapped by  $p$  to  $X$  and arrows sitting above the identity on  $X$ .

A map  $\tilde{f}_R$  as above is called a *Cartesian lifting* of  $f$  and is unique up to isomorphism. If we make a choice of Cartesian liftings, the association  $R \mapsto f^*(R)$  gives rise to the so-called *reindexing functor*  $f^*: \mathcal{E}_Y \rightarrow \mathcal{E}_X$ . In what follows we will only consider *split fibrations*, that is, the Cartesian liftings are chosen such that we have  $(fg)^* = g^*f^*$  and  $id^* = id$ .

A functor  $p: \mathcal{E} \rightarrow \mathcal{B}$  is called a *bifibration* if both  $p: \mathcal{E} \rightarrow \mathcal{B}$  and  $p^{op}: \mathcal{E}^{op} \rightarrow \mathcal{B}^{op}$  are fibrations. Interestingly, a fibration is a bifibration if and only if each reindexing functor  $f^*: \mathcal{E}_Y \rightarrow \mathcal{E}_X$  has a left adjoint  $\Sigma_f \dashv f^*$ , see Jacobs (1999, Lemma 9.1.2). We will call the functors  $\Sigma_f$  *direct images along  $f$* .

Two important examples of bifibrations are those of relations over sets,  $p: \text{Rel} \rightarrow \text{Set}$ , and of predicates over sets,  $p: \text{Pred} \rightarrow \text{Set}$ , which played a crucial role in Bonchi et al. (2014). We do not recall their exact definitions here, as they arise as instances of the more general bifibrations of quantale-valued relations and predicates described in detail in the next section.

$$\begin{array}{ccc}
 \mathcal{E} & \xrightarrow{\widehat{F}} & \mathcal{E}' \\
 p \downarrow & & \downarrow p' \\
 \mathcal{B} & \xrightarrow{F} & \mathcal{B}'
 \end{array}$$

Given fibrations  $p: \mathcal{E} \rightarrow \mathcal{B}$  and  $p': \mathcal{E}' \rightarrow \mathcal{B}'$  and a functor on the base categories  $F: \mathcal{B} \rightarrow \mathcal{B}'$ , we call  $\widehat{F}: \mathcal{E} \rightarrow \mathcal{E}'$  a *lifting* of  $F$  when  $p'\widehat{F} = Fp$ . Notice that a lifting  $\widehat{F}$  restricts to a functor between the fibres  $\widehat{F}_X: \mathcal{E}_X \rightarrow \mathcal{E}'_{F(X)}$ . We omit the subscript  $X$  when it is clear from the context.

Consider an arbitrary lifting  $\widehat{F}$  of  $F$  and a morphism  $f: X \rightarrow Y$  in  $\mathcal{B}$ . For any  $R \in \mathcal{E}_Y$ , the maps  $\widetilde{Ff}_{\widehat{F}R}: (Ff)^*(\widehat{F}R) \rightarrow \widehat{F}R$  and  $\widehat{F}(\tilde{f}_R): \widehat{F}(f^*R) \rightarrow \widehat{F}R$  sit above  $Ff$ . Using the universal property in Definition 4, we obtain a canonical morphism

$$\widehat{F} \circ f^*(R) \rightarrow (Ff)^* \circ \widehat{F}(R). \tag{3}$$

A lifting  $\widehat{F}$  is called a *fibred lifting* when the natural transformation in (3) is an isomorphism.

We will use the *Beck-Chevalley* condition for fibrations  $p : \mathcal{E} \rightarrow \mathcal{B}$ , which will be needed in some of the proofs. Assume we have a commuting square:

$$\begin{array}{ccc}
 A & \xrightarrow{f} & B \\
 u \downarrow & & \downarrow v \\
 C & \xrightarrow{g} & D
 \end{array} \tag{4}$$

Since the fibration is split, we have a commuting diagram

$$\begin{array}{ccc}
 \mathcal{E}_A & \xleftarrow{f^*} & \mathcal{E}_B \\
 u^* \uparrow & & \uparrow v^* \\
 \mathcal{E}_C & \xleftarrow{g^*} & \mathcal{E}_D
 \end{array}$$

Using the adjunctions  $\Sigma_f \dashv f^*$  and  $\Sigma_g \dashv g^*$ , we obtain the so-called mate of the above square

$$\begin{array}{ccc}
 \mathcal{E}_A & \xrightarrow{\Sigma_f} & \mathcal{E}_B \\
 u^* \uparrow & \searrow & \uparrow v^* \\
 \mathcal{E}_C & \xrightarrow{\Sigma_g} & \mathcal{E}_D
 \end{array} \tag{5}$$

obtained using the unit and the counit of the above adjunctions, as the composite

$$\Sigma_f u^* \xrightarrow{\Sigma_f u^* \eta} \Sigma_f u^* g^* \Sigma_g \Rightarrow \Sigma_f f^* v^* \Sigma_g \xrightarrow{\varepsilon v^* \Sigma_g} v^* \Sigma_g$$

**Definition 5.** We say that the square (4) has the *Beck-Chevalley condition* if the mate (5) is an isomorphism.

### 4. Moving Towards a Quantitative Setting

We start by introducing two fibrations which are the foundations for our quantitative reasoning: predicates and relations valued in a quantale.

**Definition 6.** A *quantale*  $\mathcal{V}$  is a complete lattice equipped with an associative operation  $\otimes : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$  which is distributive on both sides over arbitrary joins  $\bigvee$ .

This implies that for every  $y \in \mathcal{V}$ , the functor  $- \otimes y$  has a right adjoint  $[y, -]$ . Similarly, for every  $x \in \mathcal{V}$ , the functor  $x \otimes -$  has a right adjoint, denoted by  $[[x, -]]$ . Thus, for every  $x, y, z \in \mathcal{V}$ , we have:  $x \otimes y \leq z \iff x \leq [y, z] \iff y \leq [[x, z]]$ .

If  $\otimes$  has an identity element or *unit* 1 for  $\otimes$ , the quantale is called *unital*. If  $x \otimes y = y \otimes x$  for every  $x, y \in \mathcal{V}$  the quantale is called *commutative* and we have  $[x, -] = [[x, -]]$ . Hereafter, we only work with unital, commutative quantales.

**Example 7.** The Boolean algebra  $2$  with  $\otimes = \wedge$  is a unital and commutative quantale: the unit is 1 and  $[y, z] = y \rightarrow z$ . The complete lattice  $[0, \infty]$  ordered by the reversed order of the reals, that is,  $\leq = \geq_{\mathbb{R}}$  and with  $\otimes = +$  is a unital commutative quantale: the unit is 0 and for every  $y, z \in [0, \infty]$  we have  $[y, z] = z - y$  (truncated subtraction). Also  $[0, 1]$  is a unital quantale where  $r \otimes s = \min(r + s, 1)$  (truncated addition). To avoid confusion, we use  $\vee, \wedge$  in the quantale and  $\inf, \sup$  in the reals.

**Definition 8.** Given a set  $X$  and a quantale  $\mathcal{V}$ , a  $\mathcal{V}$ -valued predicate on  $X$  is a map  $p : X \rightarrow \mathcal{V}$ . A  $\mathcal{V}$ -valued relation on  $X$  is a map  $r : X \times X \rightarrow \mathcal{V}$ .

Given two  $\mathcal{V}$ -valued predicates  $p, q : X \rightarrow \mathcal{V}$ , we say that  $p \leq q \iff \forall x \in X. p(x) \leq q(x)$ .

**Definition 9.** A morphism between  $\mathcal{V}$ -valued predicates  $p : X \rightarrow \mathcal{V}$  and  $q : Y \rightarrow \mathcal{V}$  is a map  $f : X \rightarrow Y$  such that  $p \leq q \circ f$ . We consider the category  $\mathcal{V}\text{-Pred}$  whose objects are  $\mathcal{V}$ -valued predicates and arrows are as above.

**Definition 10.** A morphism between  $\mathcal{V}$ -valued relations  $r : X \times X \rightarrow \mathcal{V}$  and  $q : Y \times Y \rightarrow \mathcal{V}$  is a map  $f : X \rightarrow Y$  such that  $p \leq q \circ (f \times f)$ . We consider the category  $\mathcal{V}\text{-Rel}$  whose objects are  $\mathcal{V}$ -valued relations and arrows are as above.

**The bifibration of  $\mathcal{V}$ -valued predicates.** The forgetful functor  $\mathcal{V}\text{-Pred} \rightarrow \text{Set}$  mapping a predicate  $p : X \rightarrow \mathcal{V}$  to  $X$  is a bifibration. The fibre  $\mathcal{V}\text{-Pred}_X$  is the lattice of  $\mathcal{V}$ -valued predicates on  $X$ . For  $f : X \rightarrow Y$  in  $\text{Set}$  the reindexing and direct image functors on a predicate  $p \in \mathcal{V}\text{-Pred}_Y$  are given by

$$f^*(p) = p \circ f \quad \text{and} \quad \Sigma_f(p)(y) = \bigvee \{p(x) \mid x \in f^{-1}(y)\}.$$

The bifibration  $\mathcal{V}\text{-Pred} \rightarrow \text{Set}$  has the Beck-Chevalley condition for weak pullback squares in  $\text{Set}$ . Essentially, we have to show that if (4) is a weak pullback, then for every  $p \in \mathcal{V}\text{-Pred}_C$  and  $b \in B$  we have

$$\bigvee_{a \in f^{-1}(b)} p(u(a)) = \bigvee_{c \in g^{-1}(v(b))} p(c) \tag{6}$$

Proving  $\leq$  is easy (we just use that the square commutes), but for  $\geq$  we need that (4) is a weak pullback.

**The bifibration of  $\mathcal{V}$ -valued relations.** Notice that we have the following pullback in  $\text{Cat}$ , where  $\Delta X = X \times X$ . This is a change-of-base situation and thus the functor  $\mathcal{V}\text{-Rel} \rightarrow \text{Set}$  mapping each  $\mathcal{V}$ -valued relation to its underlying set is also a bifibration.

We denote by  $\mathcal{V}\text{-Rel}_X$  the fibre above a set  $X$ . For each set  $X$ , the functor  $\iota$  restricts to an isomorphism  $\iota_X : \mathcal{V}\text{-Rel}_X \rightarrow \mathcal{V}\text{-Pred}_{X \times X}$ .

$$\begin{array}{ccc} \mathcal{V}\text{-Rel} & \xrightarrow{\iota} & \mathcal{V}\text{-Pred} \\ \downarrow \lrcorner & & \downarrow \\ \text{Set} & \xrightarrow{\Delta} & \text{Set} \end{array}$$

For  $f : X \rightarrow Y$  in  $\text{Set}$  the reindexing and direct image on  $p \in \mathcal{V}\text{-Rel}_Y$  are given by

$$f^*(p) = p \circ (f \times f) \quad \text{and} \quad \Sigma_f(p)(y) = \bigvee \{p(x, x') \mid (x, x') \in (f \times f)^{-1}(y, y')\}.$$

For two relations  $p, q \in \mathcal{V}\text{-Rel}_X$ , we define their *composition*  $p \cdot q : X \times X \rightarrow \mathcal{V}$  by  $p \cdot q(x, y) = \bigvee \{p(x, z) \otimes q(z, y) \mid z \in X\}$ . We define the *diagonal relation*  $\text{diag}_X \in \mathcal{V}\text{-Rel}_X$  by  $\text{diag}_X(x, y) = 1$  if  $x = y$  and  $\perp$  otherwise.

**Definition 11.** We say that a  $\mathcal{V}$ -valued relation  $r : X \times X \rightarrow \mathcal{V}$  is

- reflexive if for all  $x \in X$  we have  $r(x, x) \geq 1$ , (i.e.  $r \geq \text{diag}_X$ );
- transitive if  $r \cdot r \leq r$ ;
- symmetric if  $r = r \circ \text{sym}_X$ , where  $\text{sym}_X : X \times X \rightarrow X \times X$  is the symmetry isomorphism.

We denote by  $\mathcal{V}\text{-Cat}$  the full subcategory of  $\mathcal{V}\text{-Rel}$  consisting of reflexive, transitive relations and by  $\mathcal{V}\text{-Cat}_{\text{sym}}$  the full subcategory of  $\mathcal{V}\text{-Rel}$  that are additionally symmetric.

Note that  $\mathcal{V}$ -Cat is the category of small categories enriched over the  $\mathcal{V}$  in the sense of Kelly (1982).

**Example 12.** For  $\mathcal{V} = 2$ ,  $\mathcal{V}$ -valued relations are just relations. Reflexivity, transitivity and symmetry coincide with the standard notions, so  $\mathcal{V}$ -Cat is the category of preorders, while  $\mathcal{V}$ -Cat<sub>sym</sub> is the category of equivalence relations.

For  $\mathcal{V} = [0, \infty]$ ,  $\mathcal{V}$ -Cat is the category of generalized metric spaces à la Lawvere (Lawvere 2002) (i.e. directed pseudo-metrics and non-expansive maps), while  $\mathcal{V}$ -Cat<sub>sym</sub> is the one of pseudo-metrics.

Similar categories of quantale-valued relations have been used in the context of lambda calculus and thus higher-order languages (Pistone 2021).

### 5. Lifting Functors to $\mathcal{V}$ -Pred and $\mathcal{V}$ -Rel

In the previous section, we have introduced the fibrations of interest for quantitative reasoning. In order to deal with coinductive predicates in this setting, it is convenient to have a structured way to lift Set-functors to  $\mathcal{V}$ -valued predicates and relations and eventually to  $\mathcal{V}$ -enriched categories. Our strategy is to first lift functors to  $\mathcal{V}$ -Pred and then, by exploiting the change of base, move these liftings to  $\mathcal{V}$ -Rel. A comparison with the extensions of Set-monads to the bicategory of  $\mathcal{V}$ -matrices (Hofmann 2007) is provided in Section 9.

#### 5.1 $\mathcal{V}$ -predicate liftings

Liftings of Set-functors to the category Pred (for  $\mathcal{V} = 2$ ) of predicates have been widely studied in the context of coalgebraic modal logic, as they correspond to modal operators (see e.g. Schröder 2008). For  $\mathcal{V}$ -Pred, we proceed in a similar way. Let us analyse what it means to have a fibred lifting  $\widehat{F}$  to  $\mathcal{V}$ -Pred of an endofunctor  $F$  on Set. First, recall that the fibre  $\mathcal{V}$ -Pred<sub>X</sub> is just the poset  $\mathcal{V}^X$ . So the restriction  $\widehat{F}_X$  to such a fibre corresponds to a monotone map  $\mathcal{V}^X \rightarrow \mathcal{V}^{FX}$ . The fact that  $\widehat{F}$  is a fibred lifting essentially means that the maps  $(\mathcal{V}^X \rightarrow \mathcal{V}^{FX})_X$  form a natural transformation between the contravariant functors  $\mathcal{V}^-$  and  $\mathcal{V}^{F-}$ . Furthermore, by Yoneda lemma we know that natural transformations  $\mathcal{V}^- \Rightarrow \mathcal{V}^{F-}$  are in one-to-one correspondence with maps  $ev : F\mathcal{V} \rightarrow \mathcal{V}$ , which we will call hereafter *evaluation maps*.

One can characterize the evaluation maps which correspond to the monotone natural transformations. In Proposition 14, we show that these are the monotone evaluation maps  $ev : (F\mathcal{V}, \ll) \rightarrow (\mathcal{V}, \leq)$  with respect to the usual order  $\leq$  on  $\mathcal{V}$  and an order  $\ll$  on  $F\mathcal{V}$  defined below and obtained by applying the standard canonical relation lifting of  $F$  – in the sense of Barr (1970) – to the relation  $\leq$ . Explicitly, we apply the functor  $F$  to the relation  $\leq$  seen as the span below in order to obtain a relation on  $F\mathcal{V}$ . Note that  $[\leq] = \{(v_1, v_2) \in \mathcal{V} \times \mathcal{V} \mid v_1 \leq v_2\}$  and  $o$  is the embedding of  $[\leq]$  in  $\mathcal{V} \times \mathcal{V}$ .

$$\begin{array}{ccccc}
 & & [\leq] & & \\
 & & \downarrow o & & \\
 \mathcal{V} & \xleftarrow{\pi_1} & \mathcal{V} \times \mathcal{V} & \xrightarrow{\pi_2} & \mathcal{V}
 \end{array}$$

**Definition 13** (Relation  $\ll$  on  $F\mathcal{V}$ ). We define a relation  $\ll$  on  $F\mathcal{V}$ : let  $v_1, v_2 \in F\mathcal{V}$ . We define  $v_1 \ll v_2$  whenever

$$\exists r \in F[\leq] \text{ s.t. } F(\pi_1 \circ o)(r) = v_1 \text{ and } F(\pi_2 \circ o)(r) = v_2$$

The relation  $\ll$  will also be denoted by  $\leq^F$  (i.e. the order  $\leq$  lifted under  $F$  via the standard relation lifting Barr 1970).



According to Balan and Kurz (2011) relation lifting transforms preorders into preorders whenever  $F$  preserves weak pullbacks (but not necessarily orders into orders).

**Proposition 14.** *There is a bijective correspondence between*

- fibred liftings  $\widehat{F}$  of  $F$  to  $\mathcal{V}$ -Pred,
- monotone natural transformations  $\mathcal{V}^- \Rightarrow \mathcal{V}^{F^-}$ ,
- monotone evaluation maps  $ev: F\mathcal{V} \rightarrow \mathcal{V}$ .

*Proof.* The equivalence of the first two bullets is well-known in coalgebraic modal logic for  $\mathcal{V} = 2$ . For the sake of completeness, we include here full details.

$\widehat{F}$  is a lifting of  $F$  to  $\mathcal{V}$ -Pred if and only if the following two conditions are met for all sets  $X$  and functions  $f: X \rightarrow Y$ :

- (1)  $\widehat{F}_X: \mathcal{V}\text{-Pred}_X \rightarrow \mathcal{V}\text{-Pred}_{FX}$  is monotone, and,
- (2) the inequality  $\widehat{F}_X \circ f^*(R) \leq (Ff)^* \circ \widehat{F}_Y$  holds.

These two conditions alone are equivalent to the laxness of the following square

$$\begin{array}{ccc}
 \mathcal{V}^Y & \xrightarrow{f} & \mathcal{V}^X \\
 \widehat{F}_Y \downarrow & \wr & \downarrow \widehat{F}_X \\
 \mathcal{V}^{FY} & \xrightarrow{Ff} & \mathcal{V}^{FX}
 \end{array}$$

However,  $\widehat{F}$  is a fibred lifting of  $F$  if and only if item 1 holds and the inequality in item 2 above is in fact an equality. Hence,  $\widehat{F}$  is a fibred lifting if and only if the above square is actually commutative, which amounts to the existence of a natural transformation  $\gamma: \mathcal{V}^- \rightarrow \mathcal{V}^{F^-}$  with each component  $\gamma_X$  being *monotone*.

We have thus proved the equivalence of the first two conditions. Now, let us turn to the equivalence between the first and third one. By Yoneda lemma, we know that natural transformations  $\mathcal{V}^- \rightarrow \mathcal{V}^{F^-}$  are in one-to-one correspondence with evaluation maps  $ev: F(\mathcal{V}) \rightarrow \mathcal{V}$ . It remains to characterize the monotonicity condition. We show that this is equivalent to requiring that  $ev_F: F\mathcal{V} \rightarrow \mathcal{V}$  is monotone for the order  $\ll$  on  $F\mathcal{V}$  and  $\leq$  on  $\mathcal{V}$ .

$\Leftarrow$  Assume that  $ev_F$  is monotone and take  $f_1, f_2: X \rightarrow \mathcal{V}$  such that  $f_1 \leq f_2$ . This means that  $\langle f_1, f_2 \rangle$  factors through  $o$  as depicted below, where  $u: X \rightarrow [\leq]$  is defined as  $u(x) = (f_1(x), f_2(x))$ .

$$\begin{array}{ccc}
 & \langle f_1, f_2 \rangle & \\
 & \curvearrowright & \\
 X & \xrightarrow{u} [\leq] \xrightarrow{o} & \mathcal{V} \times \mathcal{V}
 \end{array}$$

If we apply  $F$  to the diagram above and post-compose with  $F\pi_1, F\pi_2, ev_F$ , we obtain the following diagram.

$$\begin{array}{ccccccc}
 FX & \xrightarrow{Fu} & F[\leq] & \xrightarrow{Fo} & F(\mathcal{V} \times \mathcal{V}) & \xrightarrow{F\pi_1} & F\mathcal{V} \xrightarrow{ev_F} \mathcal{V} \\
 & & & & & \xrightarrow{F\pi_2} & \\
 & & & & & & 
 \end{array}$$

Let  $t \in FX$ . Our aim is to show  $\widehat{F}f_1(t) \leq \widehat{F}f_2(t)$ , which implies  $\widehat{F}f_1 \leq \widehat{F}f_2$ .

First, define  $r = Fu(t) \in F[\leq]$ . Now observe that  $F(\pi_1 \circ o)(r) = F(\pi_1 \circ o \circ u)(t) = F(\pi_1 \circ \langle f_1, f_2 \rangle)(t) = Ff_1(t)$ . Analogously,  $F(\pi_2 \circ o)(r) = Ff_2(t)$ . Hence  $Ff_1(t) \ll Ff_2(t)$ .

Using the monotonicity of  $ev_F$ , we can conclude that

$$\widehat{F}f_1(t) = ev_F(Ff_1(t)) \leq ev_F(Ff_2(t)) = \widehat{F}f_2(t).$$

‘ $\Rightarrow$ ’ Assume that  $\widehat{F}$  is monotone. In order to show monotonicity of  $ev_F$ , take  $v_1, v_2 \in F\mathcal{V}$  such that  $v_1 \ll v_2$ . This means that there exists  $r \in F[\leq]$  such that  $F(\pi_1 \circ o)(r) = v_1, F(\pi_2 \circ o)(r) = v_2$ .

Now consider  $\pi_1 \circ o, \pi_2 \circ o: [\leq] \rightarrow \mathcal{V}$ . It holds that  $\pi_1 \circ o \leq \pi_2 \circ o$  and with monotonicity of  $\widehat{F}$  we can conclude  $\widehat{F}(\pi_1 \circ o) \leq \widehat{F}(\pi_2 \circ o)$ . Hence,

$$\begin{aligned} ev_F(v_1) &= ev_F(F(\pi_1 \circ o)(r)) = \widehat{F}(\pi_1 \circ o)(r) \\ &\leq \widehat{F}(\pi_2 \circ o)(r) = ev_F(F(\pi_2 \circ o)(r)) = ev_F(v_2) \end{aligned}$$

that is, we have shown that  $ev_F$  is monotone. □

Notice that the correspondence between fibred liftings and monotone evaluation maps is given in one direction by  $ev = \widehat{F}(id_{\mathcal{V}})$ , and conversely, by  $\widehat{F}(p: X \rightarrow \mathcal{V}) = ev \circ F(p)$ .

**Evaluation maps as Eilenberg-Moore algebras.** Evaluation maps have also been extensively considered in the coalgebraic approach to modal logics (Schröder 2008). A special kind of evaluation map arises when the truth values  $\mathcal{V}$  have an algebraic structure for a given monad  $(T, \mu, \eta)$ , that is, we have  $\mathcal{V} = T\Omega$  for some object  $\Omega$  and the evaluation map  $T\mathcal{V} \rightarrow \mathcal{V}$  is an Eilenberg-Moore algebra for  $T$ . This notion of monadic modality has been studied in Hasuo (2015) where the category of free algebras for  $T$  was assumed to be order enriched. In Lemma 15 below we show that under reasonable assumptions, the evaluation map obtained as the free Eilenberg-Moore algebra on  $\Omega$  (i.e.  $ev: T\mathcal{V} \rightarrow \mathcal{V}$  is just  $\mu_{\Omega}: T^2\Omega \rightarrow T\Omega$ ) is a monotone evaluation map, and hence gives rise to a fibred lifting of  $T$ .

**Lemma 15.** *Assume that  $T$  is a monad and  $\mathcal{V} = T\Omega$  a quantale as detailed above. Assume that there is a partial order  $\sqsubseteq$  on  $\Omega$  such that the lattice order  $\leq$  of the quantale is obtained by lifting  $\sqsubseteq$  under  $T$ , that is,  $\leq = \sqsubseteq^T$  (as in Definition 13). Then  $ev = \mu_{\Omega}: (T\mathcal{V}, \leq^T) \rightarrow (\mathcal{V}, \leq)$  is monotone, and consequently corresponds to a fibred lifting  $\widehat{T}$  of  $T$ .*

*Proof.* Let  $t'_1, t'_2 \in T\mathcal{V}$  such that  $t'_1 \ll t'_2$ , that is,  $t'_1 \leq^T t'_2$ . We have to show that  $\mu_{\Omega}(t'_1) \leq \mu_{\Omega}(t'_2)$ .

Since  $\leq$  is obtained by lifting  $\sqsubseteq$  under  $T$  we can infer that there exists a witness function  $w: \leq \rightarrow T(\sqsubseteq)$  that assigns to every pair of elements  $t_1, t_2 \in \mathcal{V}$  with  $t_1 \leq t_2$  a witness  $w(t_1, t_2) \in T(\sqsubseteq)$  with  $T\pi_i(w(t_1, t_2)) = t_i$ . Hence  $T\pi_i \circ w = \pi'_i$ , where  $\pi_i: \sqsubseteq \rightarrow \Omega$  and  $\pi'_i: \leq \rightarrow \mathcal{V}$  are the usual projections.

Since  $t'_1 \leq^T t'_2$ , there exists a witness  $t' \in T(\leq)$  with  $T\pi'_i(t') = t'_i$ .

We show that  $t = \mu_{\sqsubseteq}(Tw(t'))$  is a witness for  $\mu_{\Omega}(t'_1) \leq \mu_{\Omega}(t'_2)$ . It holds that  $T\pi_i \circ \mu_{\sqsubseteq} \circ Tw = \mu_{\Omega} \circ TT\pi_i \circ Tw = \mu_{\Omega} \circ T(T\pi_i \circ w) = \mu_{\Omega} \circ T\pi'_i$ , where the first equality holds since  $\mu$  is a natural transformation. This implies  $T\pi_i(t) = (T\pi_i \circ \mu_{\sqsubseteq} \circ Tw)(t') = (\mu_{\Omega} \circ T\pi'_i)(t') = \mu_{\Omega}(t'_i)$ . □

We provide next several examples of monotone evaluation maps which arise in this fashion.

**Example 16.** *When  $T$  is the powerset monad  $\mathcal{P}$  and  $\Omega = 1$ , we obtain  $\mathcal{V} = 2$  and  $\mu_1: \mathcal{P}2 \rightarrow 2$  corresponds to the  $\diamond$  modality, that is, to an existential predicate transformer, see Hasuo (2015).*

**Example 17.** *When  $T$  is the probability distribution functor  $\mathcal{D}$  on  $\text{Set}$  and  $\Omega = 2 = \{0, 1\}$  equipped with the order  $1 \sqsubseteq 0$ , we obtain  $\mathcal{V} = \mathcal{D}\{0, 1\} \cong [0, 1]$  with the reversed order of the reals, that is,  $\leq = \geq_{\mathbb{R}}$ . In this case  $ev_{\mathcal{D}}(f) = \sum_{r \in [0,1]} r \cdot f(r)$  for  $f: [0, 1] \rightarrow [0, 1]$  a probability distribution (expectation of the identity random variable).*

**The canonical evaluation map.** In the case  $\mathcal{V} = 2$ , there exists a simple way of lifting a functor  $F: \text{Set} \rightarrow \text{Set}$ : given a predicate  $p: U \rightarrow X$ , one defines the canonical predicate lifting  $\widehat{F}_{\text{can}}(U)$  of  $F$  as the epi-mono factorization of  $Fp: FU \rightarrow FX$ . This lifting corresponds to a canonical evaluation map  $\text{true}: 1 \rightarrow 2$  which maps the unique element of 1 into the element 1 of the quantale 2. For  $\mathcal{V}$ -relations, a generalized notion of canonical evaluation map was introduced in Hofmann (2007). For  $r \in \mathcal{V}$ , consider the subset  $\uparrow r = \{v \in \mathcal{V} \mid v \geq r\}$  and write  $\text{true}_r: \uparrow r \hookrightarrow \mathcal{V}$  for the

inclusion. Given  $u \in F\mathcal{V}$ , we write  $u \in F(\uparrow r)$  when  $u$  is in the image of the injective function  $F(\text{true}_r)$ . Following Hofmann (2007), we define  $ev_{\text{can}} : F\mathcal{V} \rightarrow \mathcal{V}$  as follows:

$$ev_{\text{can}}(u) = \bigvee \{r \mid u \in F(\uparrow r)\}.$$

**Example 18.** Assume  $F$  is the powerset functor  $\mathcal{P}$  and let  $u \in \mathcal{P}(\mathcal{V})$ . We obtain that

$$ev_{\text{can}}(u) = \bigvee \{r \mid u \subseteq \uparrow r\}, \text{ or equivalently, } ev_{\text{can}}(u) = \bigwedge u.$$

When  $\mathcal{V} = 2$ , we obtain  $ev_{\text{can}} : \mathcal{P}2 \rightarrow 2$  given by  $ev_{\text{can}}(u) = 1$  iff  $u = \emptyset$  or  $u = \{1\}$ . This corresponds to the  $\square$  operator from modal logic. If  $\mathcal{V} = [0, \infty]$  we have  $ev_{\text{can}}(u) = \sup u$ .

**Example 19.** The canonical evaluation map for the distribution monad  $\mathcal{D}$  and  $\mathcal{V} = [0, 1]$  is  $ev_{\text{can}}(f) = \sup_{r \in [0,1]} f(r)$ , which is not the monad multiplication.

The canonical evaluation map  $ev_{\text{can}}$  is monotone whenever the functor  $F$  preserves weak pull-backs (see Lemma 52 in Appendix A). For such functors, by Proposition 14, the map  $ev_{\text{can}}$  induces a fibred lifting  $\widehat{F}_{\text{can}}$  of  $F$ , called the canonical  $\mathcal{V}$ -Pred-lifting of  $F$  and defined by

$$\widehat{F}_{\text{can}}(p)(u) = \bigvee \{r \mid F(p)(u) \in F(\uparrow r)\} \text{ for } p \in \mathcal{V}\text{-Pred}_X \text{ and } u \in FX.$$

**Example 20.** Consider a signature  $\Sigma$  and the usual notion of  $\Sigma$ -term. We write  $\text{Var}(t)$  for the set of variables occurring in a  $\Sigma$ -term  $t$  and  $T_\Sigma X$  for the set of  $\Sigma$ -terms  $t$  such that  $\text{Var}(t) \subseteq X$ . The assignment  $X \mapsto T_\Sigma X$  extends to a functor  $T_\Sigma : \text{Set} \rightarrow \text{Set}$ , which is also known as term monad.

The canonical  $\mathcal{V}$ -Pred-lifting of  $T_\Sigma$  is defined for all  $p \in \mathcal{V}\text{-Pred}$  and  $t \in T_\Sigma X$  by

$$\widehat{T}_{\Sigma\text{can}}(p)(t) = \bigwedge_{x \in \text{Var}(t)} p(x).$$

**5.2 From predicates to relations via Wasserstein**

We describe next how functor liftings to  $\mathcal{V}$ -Rel can be systematically obtained using the change-of-base situation described above. In particular, we will show how the Wasserstein metric between probability distributions (defined in terms of couplings of distributions) can be naturally modelled in the fibrational setting.

Consider a  $\mathcal{V}$ -predicate lifting  $\widehat{F}$  of a Set-functor  $F$ . A natural way to lift  $F$  to  $\mathcal{V}$ -relations using  $\widehat{F}$  is to regard a  $\mathcal{V}$ -relation  $r : X \times X \rightarrow \mathcal{V}$  as a  $\mathcal{V}$ -predicate on the product  $X \times X$ . Formally, we will use the isomorphism  $\iota_X$  described in Section 4. We can apply the functor  $\widehat{F}$  to the predicate  $\iota_X(r)$  in order to obtain the predicate  $\widehat{F} \circ \iota_X(r)$  on the set  $F(X \times X)$ . Ideally, we would want to transform this predicate into a relation on  $FX$ . So first, we have to transform it into a predicate on  $FX \times FX$ . To this end, we use the natural transformation

$$\lambda^F : F \circ \Delta \Rightarrow \Delta \circ F \text{ defined by } \lambda_X^F = \langle F\pi_1, F\pi_2 \rangle : F(X \times X) \rightarrow FX \times FX. \tag{7}$$

We drop the superscript and simply write  $\lambda$  when the functor  $F$  is clear from the context. Additionally, the bifibrational structure of  $\mathcal{V}$ -Rel plays a crucial role, as we can use the direct image functor  $\Sigma_{\lambda_X}$  to transform  $\widehat{F} \circ \iota_X(r)$  into a predicate on  $FX \times FX$ . Putting all the pieces together, we define a lifting of  $F$  on the fibre  $\mathcal{V}\text{-Rel}_X$  as the composite  $W(\widehat{F})_X$  given by:

$$W(\widehat{F})_X : \mathcal{V}\text{-Rel}_X \xrightarrow{\iota_X} \mathcal{V}\text{-Pred}_{\Delta X} \xrightarrow{\widehat{F}_{\Delta X}} \mathcal{V}\text{-Pred}_{F\Delta X} \xrightarrow{\Sigma_{\lambda_X}} \mathcal{V}\text{-Pred}_{\Delta FX} \xrightarrow{\iota_{FX}^{-1}} \mathcal{V}\text{-Rel}_{FX} \tag{8}$$

**Remark 21.** The notation  $W(\widehat{F})$  emphasizes the dependence on the given  $\mathcal{V}$ -predicate lifting  $\widehat{F}$ . However, for ease of notation and if no confusion can arise, we will in the following often abbreviate  $W(\widehat{F})$  by  $\overline{F}$ , leaving the underlying predicate lifting implicit. In particular,  $\overline{F}_{\text{can}}$  is based on the canonical evaluation map, that is  $\overline{F}_{\text{can}} = W(\widehat{F}_{\text{can}})$ .

We still have to verify that  $W(\widehat{F})$ , as explained in Remark 21 hereafter denoted by  $\overline{F}$ , is indeed a lifting of  $F$  to  $\mathcal{V}$ -Rel. The above construction provides the definition of  $\overline{F}$  on the fibres and, in particular, on the objects of  $\mathcal{V}$ -Rel. For a morphism between  $\mathcal{V}$ -relations  $p \in \mathcal{V}\text{-Rel}_X$  and  $q \in \mathcal{V}\text{-Rel}_Y$ , that is, a map  $f : X \rightarrow Y$  such that  $p \leq f^*(q)$ , we define  $\overline{F}(f)$  as the map  $Ff : FX \rightarrow FY$ . To see that this is well defined, it remains to show that  $\overline{F}p \leq (Ff)^*(\overline{F}q)$ . This is the first part of the next proposition.

**Proposition 22.** The functor  $\overline{F} = W(\widehat{F})$  defined above is a well-defined lifting of  $F$  to  $\mathcal{V}$ -Rel. Furthermore, when  $F$  preserves weak pullbacks and  $\widehat{F}$  is a fibred lifting of  $F$  to  $\mathcal{V}$ -Pred, then  $\overline{F}$  is a fibred lifting of  $F$  to  $\mathcal{V}$ -Rel.

*Proof.* To prove that  $\overline{F}$  is a well-defined functor on  $\mathcal{V}$ -Rel, it remains to show that  $\overline{F}p \leq (Ff)^*(\overline{F}q)$  whenever  $p \leq f^*q$  (for  $f : X \rightarrow Y$ ). From the definition of  $\overline{F}$  as given in (8), we know that on each fibre  $\overline{F}$  is monotone; hence,  $\overline{F}p \leq \overline{F}(f^*(q))$ . Hence, it suffices to show that  $\overline{F}(f^*(q)) \leq (Ff)^* \circ \overline{F}(q)$ .

This follows from the sequence of (in)equalities (9)–(14), where on each line we underlined the sub-expression that was rewritten and which we will explain in turn.

$$\overline{F} \circ f^*(q) = \iota_{FX}^{-1} \circ \Sigma_{\lambda_X} \circ \widehat{F}_{\Delta X} \circ \iota_X \circ f^*(q) \tag{9}$$

$$= \iota_{FX}^{-1} \circ \Sigma_{\lambda_X} \circ \widehat{F}_{\Delta X} \circ (\Delta f)^* \circ \iota_Y(q) \tag{10}$$

$$\leq \iota_{FX}^{-1} \circ \Sigma_{\lambda_X} \circ (F\Delta f)^* \circ \widehat{F}_{\Delta Y} \circ \iota_Y(q) \tag{11}$$

$$\leq \iota_{FX}^{-1} \circ (\Delta Ff)^* \circ \Sigma_{\lambda_Y} \circ \widehat{F}_{\Delta Y} \circ \iota_Y(q) \tag{12}$$

$$= (Ff)^* \circ \iota_{FY}^{-1} \circ \Sigma_{\lambda_Y} \circ \widehat{F}_{\Delta Y} \circ \iota_Y(q) \tag{13}$$

$$= (Ff)^* \circ \overline{F}(q) \tag{14}$$

We obtained (9) and (14) using the definition of  $\overline{F}$ . To derive the equalities in (10) and (13), we used the fact that  $\iota$  is a fibred lifting of  $\Delta$ . The inequality (11) follows from the fact that  $\widehat{F}$  is a lifting of  $F$ , and hence, we have the inequality

$$\widehat{F}_{\Delta X} \circ (\Delta f)^* \leq (F\Delta f)^* \circ \widehat{F}_{\Delta Y} . \tag{15}$$

Finally, the inequality (12) follows from the commutativity of the naturality squares of  $\lambda$  as an instance of (5).

$$\Sigma_{\lambda_X} \circ (F\Delta f)^* \leq (\Delta Ff)^* \circ \Sigma_{\lambda_Y} . \tag{16}$$

Now let us focus on the second part of the proof. Since  $\widehat{F}$  is a fibred lifting by assumption, then the inequality (15) becomes an equality. When the functor  $F$  preserves weak pullbacks, then by Lemma 51 in Appendix A we know that the naturality squares of  $\lambda$  are weak pullbacks. Hence, since the fibration  $\mathcal{V}$ -Rel has the Beck-Chevalley property for weak pullback squares, it follows that (16) is also an equality. We obtain that all the inequalities (9)–(14) are in fact equalities. This amounts to the fact that  $\overline{F}$  is a fibred lifting.  $\square$

Spelling out the concrete description of the direct image functor and of  $\lambda_X$ , we obtain for a relation  $p \in \mathcal{V}\text{-Rel}_X$  and  $t_1, t_2 \in FX$ , that

$$\overline{F}(p)(t_1, t_2) = \bigvee \{ \widehat{F}(p)(t) \mid t \in F(X \times X), F\pi_i(t) = t_i \} \tag{17}$$

Unravelling the definition of  $\widehat{F}(p)(t) = ev \circ F(p)$ , we obtain for  $\overline{F}(p)$  the same formula as for the extension of  $F$  on  $\mathcal{V}$ -matrices, as given in Hofmann (2007, Definition 3.4). This definition

in Hofmann (2007) is obtained by a direct generalization of the Barr extensions of Set-functors to the bicategory of relations. In contrast, we obtained (17) by exploiting the fibrational change-of-base situation and by first considering a  $\mathcal{V}$ -Pred-lifting.

We call a lifting of the form  $\bar{F}$  the Wasserstein lifting of  $F$  corresponding to  $\widehat{F}$ . This terminology is motivated by the next example.

**Example 23.** When  $F = \mathcal{D}$  (the distribution functor),  $\mathcal{V} = [0, 1]$  and  $ev_F$  is as in Example 17, then  $\bar{F}$  is the original Wasserstein metric from transportation theory (Villani 2009), which – by the Kantorovich-Rubinstein duality – is the same as the Kantorovich metric. Here we compare two probability distributions  $t_1, t_2 \in \mathcal{D}X$  and obtain as a result the coupling  $t \in \mathcal{D}(X \times X)$  with marginal distributions  $t_1, t_2$ , giving us the optimal plan to transport the ‘supply’  $t_1$  to the ‘demand’  $t_2$ . More concretely, given a metric  $d: X \times X \rightarrow \mathcal{V}$ , the (discrete) Wasserstein metric is defined as

$$d^W(t_1, t_2) = \inf\left\{ \sum_{x,y \in X} d(x, y) \cdot t(x, y) \mid \sum_y t(x, y) = t_1(x), \sum_x t(x, y) = t_2(y) \right\}.$$

On the other hand, when  $ev_F$  is the canonical evaluation map of Example 19 the corresponding  $\mathcal{V}$ -Rel-lifting  $\bar{F}_{\text{can}}$  minimizes the longest distance (and hence the required time) rather than the total cost of transport.

**Example 24.** Let us spell out the definition when  $F = \mathcal{P}$  (powerset functor),  $\mathcal{V} = [0, 1]$  and  $ev_F: \mathcal{P}[0, 1] \rightarrow [0, 1]$  corresponds to  $\text{sup}$ , which is clearly monotone and is the canonical evaluation map as in Example 18.

Then, given a metric  $d: X \times X \rightarrow [0, 1]$  and  $X_1, X_2 \subseteq X$ , the lifted metric is defined as follows (remember that the order is reversed on  $[0, 1]$ ):

$$\bar{F}(d)(X_1, X_2) = \inf\{\text{sup } d[Y] \mid Y \subseteq X \times X, \pi_i[Y] = X_i\}$$

As explained in Baldan et al. (2018), Mémoli (2011), this is the same as the Hausdorff metric  $d^H$  defined by:

$$d^H(X_1, X_2) = \sup\left\{ \sup_{x_1 \in X_1} \inf_{x_2 \in X_2} d(x_1, x_2), \sup_{x_2 \in X_2} \inf_{x_1 \in X_1} d(x_1, x_2) \right\}$$

**Example 25.** Recall the  $\widehat{T}_{\Sigma \text{can}}$ , the canonical  $\mathcal{V}$ -Pred-lifting of the term monad  $T_{\Sigma}$ , from Example 20. We now illustrate  $\bar{T}_{\Sigma \text{can}}$ , the Wasserstein lifting corresponding to  $\widehat{T}_{\Sigma \text{can}}$ . By (17) we have that for all  $d \in \mathcal{V}$ -Rel and  $t_1, t_2 \in T_{\Sigma}X$ , it holds that

$$\bar{T}_{\Sigma \text{can}}(d)(t_1, t_2) = \bigvee \{ \widehat{T}_{\Sigma \text{can}}(d)(t) \mid t \in T_{\Sigma}(X \times X), T_{\Sigma}\pi_i(t) = t_i \}$$

Assume that for  $t_1, t_2 \in T_{\Sigma}X$ , there exists a  $\Sigma$ -context  $C(-_1, \dots, -_n)$  such that  $t_1 = C(x_1, \dots, x_n)$  and  $t_2 = C(y_1, \dots, y_n)$  for  $j \in \{1, \dots, n\}$  and variables  $x_j, y_j \in X$ .<sup>1</sup> Then, such a context  $C$  is unique and the above set contains exactly one  $t \in T_{\Sigma}(X \times X)$  that is  $t = C((x_1, y_1), \dots, (x_n, y_n))$ . Thus,  $\bar{T}_{\Sigma \text{can}}(d)(t_1, t_2) = \widehat{T}_{\Sigma \text{can}}(d)(C((x_1, y_1), \dots, (x_n, y_n)))$  that, by definition of  $\widehat{T}_{\Sigma \text{can}}$ , is  $\bigwedge_{j \in \{1, \dots, n\}} d(x_j, y_j)$ . Instead, if such a context  $C$  does not exist, then the above set is empty and  $\bar{T}_{\Sigma \text{can}}(d)(t_1, t_2) = \perp$ . In a nutshell,

$$\bar{T}_{\Sigma \text{can}}(d)(t_1, t_2) = \begin{cases} \bigwedge_{j \in \{1, \dots, n\}} d(x_j, y_j) & \exists C(-_1, \dots, -_n) \text{ such that} \\ & t_1 = C(x_1, \dots, x_n) \text{ and } t_2 = C(y_1, \dots, y_n) \\ \perp & \text{otherwise} \end{cases} \quad (18)$$

As an example take as terms  $t_1 = f(g(x), h(y))$  and  $t_2 = f(g(z), h(x))$ , then the context is  $C(-_1, -_2) = f(g(-_1), h(-_2))$  and the term  $t \in T_{\Sigma}(X \times X)$  is  $t = f(g((x, z), h((y, x)))$ . Thus,

$\overline{T}_{\Sigma_{\text{can}}}(d)(t_1, t_2) = d(x, z) \wedge d(y, x)$ . Instead, if one takes  $t_1 = f(g(x), h(y))$  and  $t_2 = f(g(z), y)$ , then  $\overline{T}_{\Sigma_{\text{can}}}(d)(t_1, t_2) = \perp$ .

The next lemma establishes that this construction is functorial: liftings of natural transformations to  $\mathcal{V}$ -Pred can be converted into liftings of natural transformations between the corresponding Wasserstein liftings on  $\mathcal{V}$ -Rel.

We now prove the following lemma:

**Lemma 26.** *If there exists a lifting  $\widehat{\zeta} : \widehat{F} \Rightarrow \widehat{G}$  of a natural transformation  $\zeta : F \Rightarrow G$ , then there exists a lifting  $\overline{\zeta} : \overline{F} \Rightarrow \overline{G}$  between the corresponding Wasserstein liftings (where  $\overline{F} = W(\widehat{F})$ ,  $\overline{G} = W(\widehat{G})$ ). Furthermore, when  $\widehat{F}$  and  $\widehat{G}$  correspond to monotone evaluation maps  $ev_F$  and  $ev_G$ , then the lifting  $\widehat{\zeta}$  exists and is unique if and only if  $ev_F \leq ev_G \circ \zeta_{\mathcal{V}}$ .*

*Proof.* The existence (and in this case uniqueness) of the lifting  $\widehat{\zeta}$  is equivalent to the fact that  $\widehat{F}_X \leq (\zeta_X)^* \circ \widehat{G}_X$  for all  $X$ . This is fairly standard, but we include here an explanation for the sake of completeness. If  $\widehat{\zeta}$  exists, then for all  $p \in \mathcal{V}\text{-Pred}_X$  we have the next diagram, where the dashed arrow exists and is unique by the universal property in Definition 4.

$$\begin{array}{ccc} \widehat{F}(p) & \xrightarrow{\widehat{\zeta}_p} & \widehat{G}(p) \\ \downarrow \exists! & \nearrow \widetilde{\zeta}_{X\widehat{G}(p)} & \\ (\zeta_X)^*(\widehat{G}(p)) & & \end{array}$$

$$X \xrightarrow{\zeta_X} Y$$

Since the fibres in  $\mathcal{V}$ -Pred are posets, this means that  $\widehat{F}(p) \leq (\zeta_X)^* \circ \widehat{G}(p)$ , since there is a unique morphism in the fibre from  $\widehat{F}(p)$  to  $\widehat{G}(p)$ . For the same reason, any two liftings of  $\zeta$  must coincide. Conversely, if the inequality  $\widehat{F}(p) \leq (\zeta_X)^* \circ \widehat{G}(p)$  holds, we compose with  $\widetilde{\zeta}_{X\widehat{G}(p)}$  in order to obtain  $\widehat{\zeta}_p$ .

We have to show that  $\overline{F}_X \leq (\zeta_X)^* \circ \overline{G}_X$ . We obtain:

$$\begin{aligned} (\zeta_X)^* \circ \overline{G} &= (\zeta_X)^* \circ \iota_{GX}^{-1} \circ \Sigma_{\lambda_X^G} \circ \widehat{G} \circ \iota_X \\ &= \iota_{FX}^{-1} \circ (\zeta_X \times \zeta_X)^* \circ \Sigma_{\lambda_X^G} \circ \widehat{G} \circ \iota_X \\ &\geq \iota_{FX}^{-1} \circ \Sigma_{\lambda_X^F} \circ (\zeta_{X \times X})^* \circ \widehat{G} \circ \iota_X \\ &\geq \iota_{FX}^{-1} \circ \Sigma_{\lambda_X^F} \circ \widehat{F} \circ \iota_X \\ &= \overline{F} \end{aligned}$$

To show the inequality on the third line, we notice that  $\Sigma_{\lambda_X^F} \circ (\zeta_{X \times X})^* \leq (\zeta_X \times \zeta_X)^* \circ \Sigma_{\lambda_X^G}$  is the mate, as in (5), of the following commutative square, which in turn commutes by the naturality of  $\zeta$  and the uniqueness of mediating morphisms into the product.

$$\begin{array}{ccc} F(X \times X) & \xrightarrow{\lambda_X^F} & FX \times FX \\ \zeta_{X \times X} \downarrow & & \downarrow \zeta_X \times \zeta_X \\ G(X \times X) & \xrightarrow{\lambda_X^G} & GX \times GX \end{array}$$

To summarize, the proof of the first part of the lemma follows from the next lax diagram, by composing with the isomorphisms  $\iota_X$  and  $\iota_{FX}^{-1}$ .

$$\begin{array}{ccccc}
 \mathcal{V}\text{-Pred}_{\Delta X} & \xrightarrow{\widehat{F}} & \mathcal{V}\text{-Pred}_{F\Delta X} & \xrightarrow{\Sigma_{\lambda_X^F}} & \mathcal{V}\text{-Pred}_{\Delta FX} \\
 & \searrow^{\widehat{G}} & \uparrow^{(\zeta_{\Delta X})^*} & \wedge & \uparrow^{(\Delta\zeta_X)^*} \\
 & & \mathcal{V}\text{-Pred}_{G\Delta X} & \xrightarrow{\Sigma_{\lambda_X^G}} & \mathcal{V}\text{-Pred}_{\Delta GX}
 \end{array}$$

It remains to prove that  $\widehat{F} \leq \zeta_X^* \circ \widehat{G}$  is equivalent to  $ev_F \leq ev_G \circ \zeta_V$ . The implication from left to right is obtained by setting  $X = \mathcal{V}$  and applying the functors on both sides to  $id_{\mathcal{V}}$ . We get the other direction by taking  $p: X \rightarrow \mathcal{V}$  and computing  $((\zeta_X)^* \circ \widehat{G})(p) = ev_G \circ Gp \circ \zeta_X = ev_G \circ \zeta_V \circ Fp \geq ev_F \circ Fp = \widehat{F}(p)$ . Note that this uses the naturality of  $\zeta$ .  $\square$

**5.3 Preservation of reflexivity, symmetry and transitivity**

For  $\mathcal{V} = [0, \infty]$ , one is also interested in lifting functors to the category of (generalized) pseudometric spaces, not just of  $[0, \infty]$ -valued relations. This motivates the next question: when does the lifting  $\widehat{F}$  restrict to a functor on  $\mathcal{V}\text{-Cat}$  and  $\mathcal{V}\text{-Cat}_{\text{sym}}$ ? We have the following characterization theorem, where  $\kappa_X: X \rightarrow \mathcal{V}$  is the constant function  $x \mapsto 1$  and  $u \otimes v: X \rightarrow \mathcal{V}$  denotes the pointwise tensor of two predicates  $u, v: X \rightarrow \mathcal{V}$ , that is,  $(u \otimes v)(x) = u(x) \otimes v(x)$ .

Furthermore, let  $\delta_X: X \rightarrow X \times X$  be the diagonal function on a set  $X$ . A relation  $r: X \times X \rightarrow \mathcal{V}$  is reflexive if and only if

$$\delta_X^* \circ \iota_X(r) \geq \kappa_X. \tag{19}$$

**Theorem 27.** Assume  $\widehat{F}$  is a lifting of  $F$  to  $\mathcal{V}\text{-Pred}$  and  $\overline{F}$  is the corresponding  $\mathcal{V}\text{-Rel}$  Wasserstein lifting, that is,  $\overline{F} = W(\widehat{F})$ . Then

- If  $\widehat{F}(\kappa_X) \geq \kappa_{FX}$ , then  $\overline{F}(\text{diag}_X) \geq \text{diag}_{FX}$ ; hence,  $\overline{F}$  preserves reflexive relations;
- If  $\widehat{F}$  is a fibred lifting,  $F$  preserves weak pullbacks and  $\widehat{F}(p \otimes q) \geq \widehat{F}(p) \otimes \widehat{F}(q)$ ; then,  $\overline{F}(p \cdot q) \geq \overline{F}(p) \cdot \overline{F}(q)$ ; hence,  $\overline{F}$  preserves transitive relations;
- $\overline{F}$  preserves symmetric relations.

Consequently, when all the above hypotheses are satisfied, then the corresponding  $\mathcal{V}\text{-Rel}$  Wasserstein lifting  $\overline{F}$  restricts to a lifting of  $F$  to both  $\mathcal{V}\text{-Cat}$  and  $\mathcal{V}\text{-Cat}_{\text{sym}}$ .

For  $\mathcal{V} = [0, \infty]$ , the first condition of Theorem 27 is a relaxed version of a condition in Baldan et al. (2018, Definition 5.14) used to guarantee reflexivity. The second condition (for transitivity) is equivalent to a non-symmetric variant of a condition in Baldan et al. (2018) (see Lemma 54 in Appendix A).

The proof is immediate from Lemmas 28, 30 and 31 which we prove next.

**Lemma 28.** Assume  $\widehat{F}$  is a lifting of  $F$  such that

$$\widehat{F}(\kappa_X) \geq \kappa_{FX}.$$

Then,  $\overline{F}(\text{diag}_X) \geq \text{diag}_{FX}$ ; hence,  $\overline{F}$  preserves reflexive relations.

*Proof.* Notice that

$$\text{diag}_X = \iota_X^{-1} \Sigma_{\delta_X}(\kappa_X).$$

Using this observation, we obtain that

$$\bar{F}(diag_X) = \iota_{FX}^{-1} \circ \Sigma_{\lambda_X} \circ \widehat{F} \circ \Sigma_{\delta_X}(\kappa_X) \tag{20}$$

$$\geq \iota_{FX}^{-1} \circ \Sigma_{\lambda_X} \circ \Sigma_{F\delta_X} \circ \widehat{F}(\kappa_X) \tag{21}$$

$$\geq \iota_{FX}^{-1} \circ \Sigma_{\lambda_X} \circ \Sigma_{F\delta_X}(\kappa_{FX}) \tag{22}$$

$$= \iota_{FX}^{-1} \circ \Sigma_{\delta_{FX}}(\kappa_{FX}) \tag{23}$$

$$= diag_{FX} \tag{24}$$

In (20), we used the definition of  $\bar{F}$ . For the inequality (21), we used that  $\widehat{F}$  is a lifting of  $F$  and the mate of (3), that is,  $\widehat{F} \circ \Sigma_{\delta_X} \geq \Sigma_{F\delta_X} \circ \widehat{F}$ . The inequality (22) is the hypothesis, while in (23) we used that  $\lambda_X \circ F\delta_X = \delta_{FX}$ .

Preservation of reflexive relations is now immediate. For  $r \in \mathcal{V}\text{-Rel}_X$  is reflexive when  $r \geq diag_X$ . Hence,  $\bar{F}(r) \geq \bar{F}(diag_X) \geq diag_{FX}$ , which entails that  $\bar{F}$  is reflexive.  $\square$

We now turn our attention to the preservation of composition of relations and of the transitivity property.

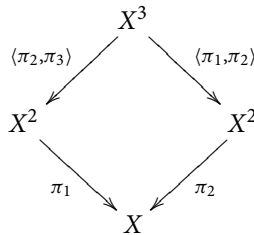
We will use the notations  $\pi_i : X \times X \times X \rightarrow X$  to denote the  $i$ th projection on  $X^3$  and  $\tau_i : FX \times FX \rightarrow FX$  to denote the  $i$ th projection on  $(FX)^3$ .

We will use the fact that the composition  $p \cdot q$  of two relations  $p, q : X \times X \rightarrow \mathcal{V}$  can be written as the composite

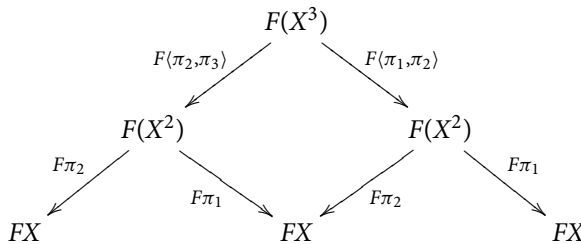
$$p \cdot q = \iota_X^{-1} \Sigma_{(\pi_1, \pi_3)}(\langle \pi_2, \pi_3 \rangle^*(\iota_X q) \otimes \langle \pi_1, \pi_2 \rangle^*(\iota_X p)) \tag{25}$$

**Lemma 29.** Assume  $F$  preserves weak pullbacks and  $u, w : F(X \times X) \rightarrow \mathcal{V}$  are in  $\mathcal{V}\text{-Pred}_{F(X \times X)}$ . We denote by  $v_X : F(X^3) \rightarrow (FX)^3$  the map defined as  $v_X = \langle F\pi_1, F\pi_2, F\pi_3 \rangle$ . Then we have:  $\Sigma_{v_X}((F\langle \pi_2, \pi_3 \rangle)^*(u) \otimes (F\langle \pi_1, \pi_2 \rangle)^*(w)) = \langle \tau_2, \tau_3 \rangle^* \Sigma_{\lambda_X}(u) \otimes \langle \tau_1, \tau_2 \rangle^* \Sigma_{\lambda_X}(w)$ .

*Proof.* It is easy to verify that the square below is a pullback.



By applying  $F$  to the diagram, we obtain the diagram below where the square is a weak pullback (since  $F$  preserves weak pullbacks).



Using this diagram, we can show that the square below is a weak pullback as well. Assume that  $t_1, t_2 \in F(X^2)$ ,  $(s_1, s_2, s_3) \in (FX)^3$  are given such that  $\lambda_X(t_1) = (s_2, s_3)$  (which means  $F\pi_1(t_1) = s_2$ ,  $F\pi_2(t_1) = s_3$ ) and  $\lambda_X(t_2) = (s_1, s_2)$  (which means  $F\pi_1(t_2) = s_1$ ,  $F\pi_2(t_2) = s_2$ ). That is,  $t_1, t_2$  live on the middle level and  $s_3, s_2, s_1$  on the lower level (in that order) in the diagram above. Since the square is a weak pullback, there exists  $t \in F(X^3)$  such that  $F\langle \pi_2, \pi_3 \rangle(t) = t_1$  and  $F\langle \pi_1, \pi_2 \rangle(t) = t_2$ .



It remains to verify that  $\nu_X(t) = (s_1, s_2, s_3)$ : for instance  $F\pi_1(t) = (F\pi_1 \circ F\langle\pi_1, \pi_2\rangle)(t) = F\pi_1(t_2) = s_1$ . (Analogously for  $s_2, s_3$ .)

$$\begin{array}{ccc}
 F(X^3) & \xrightarrow{\nu_X} & (FX)^3 \\
 \downarrow \langle F\langle\pi_2, \pi_3\rangle, F\langle\pi_1, \pi_2\rangle \rangle & & \downarrow \langle \langle\tau_2, \tau_3\rangle, \langle\tau_1, \tau_2\rangle \rangle \\
 F(X^2) \times F(X^2) & \xrightarrow{\lambda_X \times \lambda_X} & (FX)^2 \times (FX)^2
 \end{array}$$

Since the Beck-Chevalley condition holds, we obtain

$$\Sigma_{\nu_X} \langle F\langle\pi_2, \pi_3\rangle, \rangle F\langle\pi_1, \pi_2\rangle^* = \langle \langle\tau_2, \tau_3\rangle, \langle\tau_1, \tau_2\rangle \rangle^* \Sigma_{\lambda_X \times \lambda_X}.$$

Then, we will apply this to a predicate of the form  $\otimes \circ (u \times w)$  and using the facts

- $\langle h_1, h_2 \rangle^* (\otimes \circ (u \times w)) = h_1^*(u) \otimes h_2^*(w)$ .
- $\Sigma_{f \times f} (\otimes \circ (u \times w)) = \otimes \circ (\Sigma_f(u) \times \Sigma_f(w))$ .

we derive the desired equality.

While the first item above is straightforward, the second has to be further explained. Whenever  $f : X \rightarrow Y, p, p' : X \rightarrow \mathcal{V}, y, y' \in Y$ , we have (using distributivity):

$$\begin{aligned}
 & \Sigma_{f \times f} (\otimes \circ (p \times p'))(y, y') \\
 &= \bigvee \{ p(x) \otimes p'(x') \mid f(x) = y, f(x') = y' \} \\
 &= \left( \bigvee_{f(x)=y} p(x) \right) \otimes \left( \bigvee_{f(x')=y'} p'(x') \right) \\
 &= \Sigma_f(p)(y) \otimes \Sigma_f(p')(y') \\
 &= \otimes \circ (\Sigma_f(p) \times \Sigma_f(p'))(y, y') \quad \square
 \end{aligned}$$

**Lemma 30.** Assume  $F$  preserves weak pullbacks and  $\widehat{F}$  is a fibred lifting of  $F$  such that

$$\widehat{F}(u \otimes v) \geq \widehat{F}(u) \otimes \widehat{F}(v) \tag{26}$$

Then,  $\overline{F}(p \cdot q) \geq \overline{F}(p) \cdot \overline{F}(q)$ ; hence,  $\overline{F}$  preserves transitive relations.

*Proof.* We denote by  $\nu_X : F(X^3) \rightarrow (FX)^3$  the map defined as  $\nu_X = \langle F\pi_1, F\pi_2, F\pi_3 \rangle$ .

$$\overline{F}(p \cdot q) = \iota_X^{-1} \Sigma_{\lambda_X} \widehat{F} \Sigma_{\langle\pi_1, \pi_3\rangle} (\langle\pi_2, \pi_3\rangle^* (\iota_X q) \otimes \langle\pi_1, \pi_2\rangle^* (\iota_X p)) \tag{27}$$

$$\geq \iota_X^{-1} \Sigma_{\lambda_X} \Sigma_{F\langle\pi_1, \pi_3\rangle} \widehat{F} (\langle\pi_2, \pi_3\rangle^* (\iota_X q) \otimes \langle\pi_1, \pi_2\rangle^* (\iota_X p)) \tag{28}$$

$$\geq \iota_X^{-1} \Sigma_{\lambda_X} \Sigma_{F\langle\pi_1, \pi_3\rangle} \widehat{F} (\langle\pi_2, \pi_3\rangle^* (\iota_X q)) \otimes \widehat{F} (\langle\pi_1, \pi_2\rangle^* (\iota_X p)) \tag{29}$$

$$= \iota_X^{-1} \Sigma_{\lambda_X} \Sigma_{F\langle\pi_1, \pi_3\rangle} (F\langle\pi_2, \pi_3\rangle^* \widehat{F} (\iota_X q) \otimes F\langle\pi_1, \pi_2\rangle^* \widehat{F} (\iota_X p)) \tag{30}$$

$$= \iota_X^{-1} \Sigma_{\langle\tau_1, \tau_3\rangle} \Sigma_{\nu_X} (F\langle\pi_2, \pi_3\rangle^* \widehat{F} (\iota_X q) \otimes F\langle\pi_1, \pi_2\rangle^* \widehat{F} (\iota_X p)) \tag{31}$$

$$= \iota_X^{-1} \Sigma_{\langle\tau_1, \tau_3\rangle} (\langle\tau_2, \tau_3\rangle^* \Sigma_{\lambda_X} (\widehat{F} (\iota_X q)) \otimes \langle\tau_1, \tau_2\rangle^* \Sigma_{\lambda_X} (\widehat{F} (\iota_X p))) \tag{32}$$

$$= \iota_X^{-1} \Sigma_{\langle\tau_1, \tau_3\rangle} (\langle\tau_2, \tau_3\rangle^* (\iota_X \overline{F} q) \otimes \langle\tau_1, \tau_2\rangle^* (\iota_X \overline{F} p)) \tag{33}$$

$$= \overline{F} p \cdot \overline{F} q \tag{34}$$

The equalities (27), (33) and (34) follow by unravelling the definition of  $\overline{F}$  and from (25). The inequality (28) follows using by the mate of (3). The inequality (29) follows from the hypothesis on  $\widehat{F}$ . The equality (30) is obtained using the  $\overline{F}$  is a fibred lifting. To prove the equality in (31), we use that  $\lambda_X \circ F\langle\pi_1, \pi_3\rangle = \langle\tau_1, \tau_3\rangle \circ \nu_X$ . Finally, (32) follows from Lemma 29.

Assume  $r \in \mathcal{V}\text{-Rel}_X$  is transitive, that is,  $r \cdot r \leq r$ . Then, we have  $\bar{F}r \cdot \bar{F}r \leq \bar{F}(r \cdot r) \leq \bar{F}r$ , hence  $\bar{F}r$  is transitive.  $\square$

**Lemma 31.** *The Wasserstein lifting preserves symmetric  $\mathcal{V}$ -valued relations.*

*Proof.* We first observe that the square below commutes.

$$\begin{array}{ccc} F(X \times X) & \xrightarrow{\lambda_X} & FX \times FX \\ F\text{sym}_X \downarrow & & \downarrow \text{sym}_{FX} \\ F(X \times X) & \xrightarrow{\lambda_X} & FX \times FX \end{array}$$

Knowing that  $\lambda_X = \langle F\pi_1^X, F\pi_2^X \rangle$  and that  $\text{sym}_X = \langle \pi_2, \pi_1 \rangle$ , where  $\pi_i^X : X \times X \rightarrow X$ , we can easily show that the square commutes:

$$\begin{aligned} & \text{sym}_{FX} \circ \lambda_X \\ &= \langle \pi_2^{FX}, \pi_1^{FX} \rangle \circ \langle F\pi_1^X, F\pi_2^X \rangle \\ &= \langle \pi_2^{FX} \circ \langle F\pi_1^X, F\pi_2^X \rangle, \pi_1^{FX} \circ \langle F\pi_1^X, F\pi_2^X \rangle \rangle \\ &= \langle F\pi_2^X, F\pi_1^X \rangle \\ &= \langle F(\pi_1^X \circ \langle \pi_2^X, \pi_1^X \rangle), F(\pi_2^X \circ \langle \pi_2^X, \pi_1^X \rangle) \rangle \\ &= \langle F(\pi_1^X \circ \text{sym}_X), F(\pi_2^X \circ \text{sym}_X) \rangle \\ &= \langle F\pi_1^X, F\pi_2^X \rangle \circ F\text{sym}_X \\ &= \lambda_X \circ F\text{sym}_X \end{aligned}$$

Recall that  $p \in \mathcal{V}\text{-Rel}_Y$  is symmetric when  $p = p \circ \text{sym}_Y$ . We cannot perform a reindexing along  $\text{sym}_Y$  in the fibration  $\mathcal{V}\text{-Rel}$ , since  $\text{sym}_Y$  is not a morphism on  $Y$ , but on  $Y \times Y$ . Instead, we have that  $p$  is symmetric if and only if

$$\iota_Y p = (\text{sym}_Y)^*(\iota_Y p)$$

in  $\mathcal{V}\text{-Pred}$ . Hence, we want to show that for any  $r \in \mathcal{V}\text{-Rel}_X$  the implication holds

$$\iota_X r = (\text{sym}_X)^*(\iota_X r) \Rightarrow \iota_{FX} \bar{F}r = (\text{sym}_{FX})^*(\iota_{FX} \bar{F}r)$$

We have the following inequalities:

$$\begin{aligned} \iota_{FX} \bar{F}r &= \Sigma_{\lambda_X} \circ \widehat{F}_{\Delta X} \circ \iota_X(r) \\ &= \Sigma_{\lambda_X} \circ \widehat{F}_{\Delta X} \circ (\text{sym}_X)^*(\iota_X r) \\ &\leq \Sigma_{\lambda_X} \circ (F\text{sym}_X)^* \circ \widehat{F}_{\Delta X}(\iota_X r) \\ &\leq (\text{sym}_{FX})^* \circ \Sigma_{\lambda_X} \circ \widehat{F}_{\Delta X}(\iota_X r) \\ &= (\text{sym}_{FX})^*(\iota_{FX} \bar{F}r) \end{aligned}$$

However, using the idempotency of  $\text{sym}_{FX}$  and the monotonicity of  $(\text{sym}_{FX})^*$  from the inequality

$$\iota_{FX} \bar{F}r \leq (\text{sym}_{FX})^*(\iota_{FX} \bar{F}r)$$

that we have just proved above we can infer that the equality also holds.  $\square$

We can establish generic sufficient conditions on a monotone evaluation map  $ev$  so that the corresponding  $\mathcal{V}\text{-Pred}$ -lifting  $\widehat{F}$  satisfies the conditions of Theorem 27. In Proposition 53 in Appendix A, we show that  $\widehat{F}(p \otimes q) \geq \widehat{F}(p) \otimes \widehat{F}(q)$  holds whenever the map  $\otimes : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$  is the carrier of a lax morphism in the category of  $F$ -algebras between  $(\mathcal{V}, ev)^2 \rightarrow (\mathcal{V}, ev)$ , that is,  $\otimes \circ (ev \times ev) \circ \lambda_{\mathcal{V}} \leq ev \circ F(\otimes)$ . Furthermore,  $\widehat{F}(\kappa_X) \geq \kappa_X$  holds whenever the map  $\kappa_{\mathbb{1}} : \mathbb{1} \rightarrow \mathcal{V}$  is the carrier of a lax morphism from the one-element  $F$ -algebra  $! : F\mathbb{1} \rightarrow \mathbb{1}$  to  $(\mathcal{V}, ev)$ , that is,

$\kappa_{\perp} \circ ! \leq ev \circ F\kappa_{\perp}$ . These two requirements correspond to the conditions  $(Q_{\otimes})$ , respectively  $(Q_k)$  satisfied by a topological theory in the sense of Hofmann (2007, Definition 3.1). Since these two are satisfied by the canonical evaluation map  $ev_{can}$ ,<sup>2</sup> we immediately obtain

**Proposition 32.** *Whenever  $F$  preserves weak pullbacks the canonical lifting  $\widehat{F}_{can}$  satisfies the conditions in Theorem 27:*

- (1)  $\widehat{F}_{can}(p \otimes q) \geq \widehat{F}_{can}(p) \otimes \widehat{F}_{can}(q)$ , for all  $p, q \in \mathcal{V}\text{-Pred}_X$ ,
- (2)  $\widehat{F}_{can}(\kappa_X) \geq \kappa_X$ .

*Proof.* (1) Given  $t \in FX$ , we have on one hand that

$$\begin{aligned} \widehat{F}_{can}(p \otimes q)(t) &= ev_{can}(F(p \otimes q)(t)) \\ &= \bigvee \{r \mid F(p \otimes q)(t) \in F(\uparrow r)\}, \end{aligned}$$

and on the other, that

$$\begin{aligned} &(\widehat{F}_{can}(p) \otimes \widehat{F}_{can}(q))(t) \\ &= ev_{can}(Fp(t)) \otimes ev_{can}(Fq(t)) \\ &= \bigvee \{r \mid Fp(t) \in F(\uparrow r)\} \otimes \bigvee \{s \mid Fq(t) \in F(\uparrow s)\} \\ &= \bigvee \{r \otimes s \mid Fp(t) \in F(\uparrow r), Fq(t) \in F(\uparrow s)\}. \end{aligned}$$

Hence, in order to show the desired inequality it is sufficient to show that

$$Fp(t) \in F(\uparrow r), Fq(t) \in F(\uparrow s) \text{ imply } F(p \otimes q)(t) \in F(\uparrow (r \otimes s)).$$

Let  $r, s \in \mathcal{V}$  so that  $Fp(t) \in F(\uparrow r)$  and  $Fq(t) \in F(\uparrow s)$ . Note that  $p \otimes q: X \rightarrow \mathcal{V}$  is the composite:

$$\begin{array}{ccccc} X & \xrightarrow{\delta_X} & X \times X & \xrightarrow{p \times q} & \mathcal{V} \times \mathcal{V} & \xrightarrow{\otimes} & \mathcal{V} \\ & & & & \searrow & \nearrow & \\ & & & & & p \otimes q & \end{array}$$

Hence,  $F(p \otimes q)$  is the composite of the arrows on the top line of the diagram below:

$$\begin{array}{ccccccc} FX & \xrightarrow{F\delta_X} & F(X \times X) & \xrightarrow{F(p \times q)} & F(\mathcal{V} \times \mathcal{V}) & \xrightarrow{F(\otimes)} & F\mathcal{V} \\ & \searrow & \downarrow \lambda_X & & \downarrow \lambda_{\mathcal{V}} & & \\ & \delta_{FX} & FX \times FX & \xrightarrow{Fp \times Fq} & F\mathcal{V} \times F\mathcal{V} & & \end{array}$$

Note that the triangle and the square above are commutative. Using the abbreviation  $\theta = F((p \otimes q) \circ \delta_X)(t)$ , we have that:

$$F(p \otimes q)(t) = F(\otimes)(\theta) \tag{35}$$

$$((Fp)(t), (Fq)(t)) = \lambda_{\mathcal{V}}(\theta) \tag{36}$$

From Lemma 51 in Appendix A, we know that the square in the diagram below is a weak pullback.

$$\begin{array}{ccc} F((\uparrow r) \times (\uparrow s)) & \xrightarrow{F(\text{true}_r \times \text{true}_s)} & F(\mathcal{V} \times \mathcal{V}) \\ \downarrow (F\pi_1, F\pi_2) & & \downarrow \lambda_{\mathcal{V}} \\ F(\uparrow r) \times F(\uparrow s) & \xrightarrow{F\text{true}_r \times F\text{true}_s} & F\mathcal{V} \times F\mathcal{V} \end{array} \tag{37}$$

By hypothesis, we know that there exists  $u \in F(\uparrow r)$  and  $v \in F(\uparrow s)$  such that  $Fp(t) = F\text{true}_r(u)$  and  $Fq(t) = F\text{true}_s(v)$ . Hence,

$$(F\text{true}_r \times F\text{true}_s)(u, v) = \lambda_{\mathcal{V}}(\theta).$$

Using the fact that the square (37) is a weak pullback, there exists  $w \in F((\uparrow r) \times (\uparrow s))$  such that  $F(\text{true}_r \times \text{true}_s)(w) = \theta$ ,  $F\pi_1(w) = u$  and  $F\pi_2(w) = v$ .

Thus far, we have shown that

$$F(p \otimes q)(t) = F(\otimes)(\theta) = F(\otimes) \circ F(\text{true}_r \times \text{true}_s)(w)$$

for some  $w \in F((\uparrow r) \times (\uparrow s))$ . To finish the proof of the first item, we will prove that  $F(\otimes) \circ F(\text{true}_r \times \text{true}_s)$  factors through  $F\text{true}_{r \otimes s}: F(\uparrow(r \otimes s)) \rightarrow F\mathcal{V}$ .

To this end, notice that due to monotonicity of the tensor product, we know that  $(\uparrow r) \otimes (\uparrow s) \subseteq \uparrow(r \otimes s)$ . Hence,  $\otimes: \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$  restricts to a function  $\otimes|_{\uparrow r, \uparrow s}$  on  $\uparrow r \times \uparrow s$  so that the square below commutes.

$$\begin{array}{ccc} \uparrow r \times \uparrow s & \xrightarrow{\text{true}_r \times \text{true}_s} & \mathcal{V} \times \mathcal{V} \\ \downarrow \otimes|_{\uparrow r, \uparrow s} & & \downarrow \otimes \\ \uparrow(r \otimes s) & \xrightarrow{\text{true}_{r \otimes s}} & \mathcal{V} \end{array}$$

Now we put  $z := F(\otimes|_{\uparrow r, \uparrow s})(w)$  and observe that

$$\begin{aligned} F(p \otimes q)(t) &= F(\otimes) \circ F(\text{true}_r \times \text{true}_s)(w) \\ &= F(\text{true}_{r \otimes s}) \circ F(\otimes|_{\uparrow r, \uparrow s})(w) \\ &= F(\text{true}_{r \otimes s})(z). \end{aligned}$$

We conclude that  $F(p \otimes q)(t) \in F(\uparrow(r \otimes s))$ .

(2) Now let us prove the second item. Given  $t \in FX$ , we know that

$$\begin{aligned} \widehat{F}_{\text{can}}\kappa_X(t) &= \text{ev}_{\text{can}}(F\kappa_X(t)) \\ &= \bigvee \{r \mid F\kappa_X(t) \in F(\uparrow r)\}. \end{aligned}$$

In order to show that  $\widehat{F}_{\text{can}}\kappa_X(t) \geq 1$  it is sufficient to prove that  $F\kappa_X(t) \in F(\uparrow 1)$ .

Let  $e: X \rightarrow \uparrow 1$  a constant mapping with  $e(x) = 1$ . Then, the diagram to the left below commutes and by applying the functor  $F$  we obtain the diagram below.

$$\begin{array}{ccc} X & & FX \\ e \swarrow & & \swarrow Fe \\ \uparrow 1 & \xrightarrow{\text{true}_1} & \mathcal{V} \end{array} \quad \begin{array}{ccc} & & FX \\ & & \downarrow F\kappa_X \\ & & F\mathcal{V} \\ & \swarrow Fe & \\ & F(\uparrow 1) & \xrightarrow{F\text{true}_1} \end{array}$$

Now  $F\kappa_X(t) = F\text{true}_1(Fe(t))$ , hence  $F\kappa_X(t) \in F(\uparrow 1)$ . □

An immediate consequence of Proposition 32 and of Theorem 27 is that the Wasserstein lifting  $\overline{F}_{\text{can}}$  that corresponds to  $\widehat{F}_{\text{can}}$  restricts to a lifting of  $F$  to both  $\mathcal{V}$ -Cat and  $\mathcal{V}$ -Cat<sub>sym</sub>.

### 6. Quantitative up-to Techniques

The fibrational constructions of the previous section provide a convenient setting to develop an abstract theory of quantitative up-to techniques. We rely on Bonchi et al. (2014) and use the generic results for showing compatibility of the various up-to functions.

The coinductive object of interest is the greatest fixpoint of a monotone map  $b$  on  $\mathcal{V}$ -Rel, hereafter denoted by  $\nu b$ . Recall that an up-to technique, namely a monotone map  $f$  on  $\mathcal{V}$ -Rel, is sound whenever  $d \leq b(f(d))$  implies  $d \leq \nu b$ , for all  $d \in \mathcal{V}\text{-Rel}_X$ ; it is compatible if  $f \circ b \leq b \circ f$  in the pointwise order. It is well-known that compatibility entails soundness. Another useful property is:

$$\text{if } f \text{ is compatible, then } f(\nu b) \leq \nu b. \tag{38}$$

Following Bonchi et al. (2014), we assume hereafter that  $b$  can be seen as the composite

$$b: \mathcal{V}\text{-Rel}_X \xrightarrow{\tilde{F}} \mathcal{V}\text{-Rel}_{FX} \xrightarrow{\xi^*} \mathcal{V}\text{-Rel}_X. \tag{39}$$

where  $\xi: X \rightarrow FX$  is some coalgebra for  $F: \text{Set} \rightarrow \text{Set}$  and  $\tilde{F}: \mathcal{V}\text{-Rel} \rightarrow \mathcal{V}\text{-Rel}$  is an arbitrary lifting of  $F$ .

We now consider several up-to functions  $f$  that can be combined with such a monotone map  $b$ .

When  $F$  admits a final coalgebra  $\omega: \Omega \rightarrow F\Omega$ , the unique morphism  $!: X \rightarrow \Omega$  induces the behavioural closure up-to technique

$$bhv: \mathcal{V}\text{-Rel}_X \xrightarrow{\Sigma_!} \mathcal{V}\text{-Rel}_\Omega \xrightarrow{!^*} \mathcal{V}\text{-Rel}_X \tag{40}$$

where  $bhv(p)(x, y) = \bigvee \{p(x', y') \mid !(x) = !(x') \text{ and } !(y) = !(y')\}$ . For  $\mathcal{V} = 2$ , behavioural closure corresponds to the usual up-to behavioural equivalence (bisimilarity).

Other immediate generalizations are the up-to reflexivity (*ref*), up-to transitivity (*trn*) and up-to symmetry (*sym*) techniques. Inductively, take  $(-)^0 = id: \mathcal{V}\text{-Rel}_X \rightarrow \mathcal{V}\text{-Rel}_X$  and  $(-)^{n+1} = id \cdot (-)^n$ . Call *diag*:  $\mathcal{V}\text{-Rel}_X \rightarrow \mathcal{V}\text{-Rel}_X$  the constant function to  $diag_X$  and *inv*:  $\mathcal{V}\text{-Rel}_X \rightarrow \mathcal{V}\text{-Rel}_X$  be the inversion function mapping  $d$  into  $d \circ sym_X$ . Then *ref*, *trn* and *sym* are defined as follows.

$$ref = id \vee diag \qquad trn = \bigvee_i (-)^i \qquad sym = id \vee inv \tag{41}$$

The last up-to techniques we consider in this paper are quantitative generalizations of the up-to contextual closure. For this, one needs to assume that the coalgebra  $\xi: X \rightarrow FX$  carries a [bialgebra]bialgebras  $(X, \alpha: TX \rightarrow X, \xi: X \rightarrow FX)$  for some natural transformation  $\zeta: T \circ F \Rightarrow F \circ T$ . For an arbitrary lifting  $\tilde{T}: \mathcal{V}\text{-Rel} \rightarrow \mathcal{V}\text{-Rel}$  of  $T: \text{Set} \rightarrow \text{Set}$ , up-to contextual closure is seen as the composite:

$$f: \mathcal{V}\text{-Rel}_X \xrightarrow{\tilde{T}} \mathcal{V}\text{-Rel}_{TX} \xrightarrow{\Sigma_\alpha} \mathcal{V}\text{-Rel}_X. \tag{42}$$

**Example 33.** The functor  $T_\Sigma: \text{Set} \rightarrow \text{Set}$  from Example 20 carries a monad where the multiplication  $\mu_X: T_\Sigma T_\Sigma X \rightarrow T_\Sigma X$  is just term composition. Given a context  $C$ , hereafter we will write  $C[t_1, \dots, t_n]$  for an arbitrary element of  $T_\Sigma T_\Sigma X$  (a term of terms) and  $C(t_1, \dots, t_n)$  for the term  $\mu_X(C[t_1, \dots, t_n]) \in T_\Sigma X$  (term substitution, as before). We will use  $T_1$  and  $T_2$  for elements of  $T_\Sigma T_\Sigma X$  and  $t_1, t_2, s_j^1$  and  $s_j^2$  for elements of  $T_\Sigma X$ .

In this example, we consider the following up-to technique

$$ctx: \mathcal{V}\text{-Rel}_{T_\Sigma X} \xrightarrow{\overline{T_\Sigma \text{can}}} \mathcal{V}\text{-Rel}_{T_\Sigma T_\Sigma X} \xrightarrow{\Sigma_{\mu_X}} \mathcal{V}\text{-Rel}_{T_\Sigma X}.$$

where  $\overline{T_\Sigma \text{can}}$  is defined as in (18) and  $\Sigma_{\mu_X}$  is the direct image of the multiplication  $\mu_X$ . In order to understand why this is a proper generalization of the usual notion of contextual closure of a relation is convenient to spell out the definition. For  $d \in \mathcal{V}\text{-Rel}_{T_\Sigma X}$  and  $t_1, t_2 \in T_\Sigma X$ , we have that

$$\begin{aligned} ctx(d)(t_1, t_2) &= \Sigma_{\mu_X}(\overline{T_\Sigma \text{can}}(d))(t_1, t_2) \\ &= \bigvee \{\overline{T_\Sigma \text{can}}(d)(T_1, T_2) \mid \mu_X(T_i) = t_i\} \\ &= \bigvee \{\overline{T_\Sigma \text{can}}(d)(T_1, T_2) \mid \mu_X(T_i) = t_i \text{ and } \exists C[-0, \dots, -n] \text{ s.t. } T_i = C[s_1^i, \dots, s_n^i]\} \\ &= \bigvee_C \{\bigwedge_j d(s_j^1, s_j^2) \mid t_i = C(s_0^i, \dots, s_n^i)\} \end{aligned}$$

Notice that for  $\mathcal{V} = 2$ , this boils down to the usual notion of contextual closure of a relation.

**Example 34.** Let  $\mathcal{V} = [0, 1]$ . In Chatzikokolakis et al. (2016), the convex closure of  $d \in \mathcal{V}\text{-Rel}_{\mathcal{D}(X)}$  is defined for  $\Delta, \Theta \in \mathcal{D}(X)$  as

$$\text{cvx}(d)(\Delta, \Theta) = \inf\left\{ \sum_i p_i \cdot d(\Delta_i, \Theta_i) \mid \Delta = \sum_i p_i \cdot \Delta_i, \Theta = \sum_i p_i \cdot \Theta_i \right\}$$

where  $\Delta_i, \Theta_i \in \mathcal{D}(X)$ ,  $p_i \in [0, 1]$ . This can be obtained as in (42) by taking the Wasserstein lifting  $\overline{\mathcal{D}}$  of  $\mathcal{D}$  from Example 23 corresponding to the evaluation map of Example 17, and the free algebra structure on  $\mathcal{D}X$  given by the monad multiplication  $\mu_X: \mathcal{D}\mathcal{D}X \rightarrow \mathcal{D}X$ . This results in:

$$f: \mathcal{V}\text{-Rel}_{\mathcal{D}X} \xrightarrow{\overline{\mathcal{D}}} \mathcal{V}\text{-Rel}_{\mathcal{D}\mathcal{D}X} \xrightarrow{\Sigma_{\mu_X}} \mathcal{V}\text{-Rel}_{\mathcal{D}X}.$$

Let  $\Delta, \Theta \in \mathcal{D}X$  and  $d: \mathcal{D}X \times \mathcal{D}X \rightarrow [0, 1]$ . Then, by expanding the definitions of the direct image and of the Wasserstein lifting we obtain:

$$\begin{aligned} f(d)(\Delta, \Theta) &= \Sigma_{\mu_X}(\overline{\mathcal{D}}(d))(\Delta, \theta) \\ &= \inf\{\overline{\mathcal{D}}(d)(\tilde{\Delta}, \tilde{\Theta}) \mid \tilde{\Delta}, \tilde{\Theta} \in \mathcal{D}\mathcal{D}X, \mu_X(\tilde{\Delta}) = \Delta, \mu_X(\tilde{\Theta}) = \Theta\} \\ &= \inf\{\inf\{\widehat{\mathcal{D}}(d)(\Gamma) \mid \Gamma \in \mathcal{D}(\mathcal{D}X \times \mathcal{D}X), \mathcal{D}\pi_1(\Gamma) = \tilde{\Delta}, \mathcal{D}\pi_2(\Gamma) = \tilde{\Theta}\} \mid \\ &\quad \tilde{\Delta}, \tilde{\Theta} \in \mathcal{D}\mathcal{D}X, \mu_X(\tilde{\Delta}) = \Delta, \mu_X(\tilde{\Theta}) = \Theta\} \\ &= \inf\{\widehat{\mathcal{D}}(d)(\Gamma) \mid \Gamma \in \mathcal{D}(\mathcal{D}X \times \mathcal{D}X), \mu_X(\mathcal{D}\pi_1(\Gamma)) = \Delta, \mu_X(\mathcal{D}\pi_2(\Gamma)) = \Theta\} \end{aligned}$$

Now observe that  $\Gamma \in \mathcal{D}(\mathcal{D}X \times \mathcal{D}X)$  can be written as a formal sum  $\Gamma = \sum_i p_i \cdot (\Delta_i, \Theta_i)$  where  $\Delta_i, \Theta_i \in \mathcal{D}X$  and  $p_i = \Gamma(\Delta_i, \Theta_i)$ . Then,

$$\begin{aligned} \widehat{\mathcal{D}}(d)(\Gamma) &= \text{ev}(\mathcal{D}(d)(\Gamma)) = \sum_{r \in [0,1]} r \cdot \sum_{\Gamma(\Delta', \Theta')=r} d(\Delta', \Theta') \\ &= \sum_{\Delta', \Theta'} d(\Delta', \Theta') \cdot \Gamma(\Delta', \Theta') = \sum_i p_i \cdot d(\Delta_i, \Theta_i) \end{aligned}$$

In addition  $\mu_X(\mathcal{D}\pi_1(\Gamma)) = \Delta$  means  $\sum_i p_i \cdot \Delta_i = \Delta$  and similarly  $\mu_X(\mathcal{D}\pi_2(\Gamma)) = \Theta$  means  $\sum_i p_i \cdot \Theta_i = \Theta$ . Therefore,  $f(d)(\Delta, \Theta) = \text{cvx}(d)(\Delta, \Theta)$ .

**Remark 35.** In the following, we will use the convention that  $b$  is obtained as in equation (39) and  $f$  is as in (42). In addition  $\tilde{F}, \tilde{T}$  are arbitrary liftings of  $F, T$  to  $\mathcal{V}\text{-Rel}$ . Later, we will restrict to Wasserstein liftings.

### 6.1 Compatibility for arbitrary liftings

We now study compatibility of the up-to techniques illustrated so far with respect to  $b$  as in (39).

With the definitions of  $bhv, ref, trn, sym$  and two results in Bonchi et al. (2017b), it is immediate to prove the following.

**Proposition 36.** Let  $\tilde{F}: \mathcal{V}\text{-Rel} \rightarrow \mathcal{V}\text{-Rel}$  be an arbitrary lifting of  $F: \text{Set} \rightarrow \text{Set}$ .

- If  $\tilde{F}(\text{diag}_X) \geq \text{diag}_{FX}$ , then  $ref$  is  $b$ -compatible.
- If  $\tilde{F}(p \cdot q) \geq \tilde{F}(p) \cdot \tilde{F}(q)$  for all  $p, q \in \mathcal{V}\text{-Rel}_X$ , then  $trn$  is  $b$ -compatible.
- If  $\tilde{F}(d) \circ \text{sym}_X \leq \tilde{F}(d \circ \text{sym}_X)$ , then  $sym$  is  $b$ -compatible.
- If  $\tilde{F}$  is a fibred lifting, then  $bhv$  is  $b$ -compatible.

*Proof.* Observe that

- $diag$  is compatible by the hypothesis and Bonchi et al. (2017b, Proposition 6.3). Then,  $ref$  is compatible since  $id$  is compatible and the join of compatible functions is compatible.

- For all  $i, (-)^i$  is compatible (the proof goes by induction: for the base case,  $id$  is compatible; for the inductive case, we use Proposition 38). Then,  $trn$  is compatible (following the same argument as above).
- $inv$  is compatible by the hypothesis and Bonchi et al. (2017b, Proposition 6.3). Then  $sym$  is compatible.
- Theorem 6.1 in Bonchi et al. (2017b) entails that  $bhv$  is compatible. □

For up-to context, it is enough to use (Bonchi et al. 2014, Theorem 2) stating the following.

**Proposition 37.** *If there exists a lifting  $\bar{\zeta} : \tilde{T} \circ \tilde{F} \Rightarrow \tilde{F} \circ \tilde{T}$  of  $\zeta$ , then  $f$  is  $b$ -compatible.*

*Proof.* Follows immediately from Theorem 2 in Bonchi et al. (2014). □

As usual, compatible techniques can be combined either by function composition ( $\circ$ ) or by arbitrary joins ( $\bigvee$ ). For instance, compatibility of *up-to metric closure*, defined as the composite  $mtr = trn \circ sym \circ ref$ , follows immediately from compatibility of  $trn$ ,  $sym$  and  $ref$ . In  $\mathcal{V}$ -Rel there is yet another useful way to combine up-to techniques – called *chaining* in Chatzikokolakis et al. (2016) – and defined as the composition ( $\cdot$ ) of relations (for a definition see Section 4)

**Proposition 38.** *Let  $g_1, g_2 : \mathcal{V}\text{-Rel}_X \rightarrow \mathcal{V}\text{-Rel}_X$  be arbitrary up-to functions which are compatible with respect to  $b$ :  $\mathcal{V}\text{-Rel}_X \rightarrow \mathcal{V}\text{-Rel}_X$ . If  $\tilde{F}(p \cdot q) \geq \tilde{F}(p) \cdot \tilde{F}(q)$  for all  $p, q \in \mathcal{V}\text{-Rel}_X$ , then  $g_1 \cdot g_2$  is  $b$ -compatible.*

*Proof.* Follows immediately from Proposition 6.3 in Bonchi et al. (2017b). □

### 6.2 Compatibility for Wasserstein lifting

By moving from arbitrary liftings to Wasserstein liftings, it is possible to state more interesting results. First, observe that whenever  $\bar{F}$  is the Wasserstein lifting corresponding to some  $\mathcal{V}$ -Pred-lifting  $\hat{F}$  (i.e.  $\bar{F} = W(\hat{F})$ ) which satisfies the conditions of Theorem 27, then the hypotheses of Proposition 36 immediately hold when replacing  $\tilde{F}$  by  $\bar{F}$ , so  $bhv$ ,  $ref$ ,  $trn$ ,  $sym$  are compatible.

For up-to context, more work is needed. By Proposition 37, to prove compatibility it is enough to find a lifting

$$\bar{\zeta} : \bar{T} \circ \bar{F} \Rightarrow \bar{F} \circ \bar{T}$$

of  $\zeta$ . In order to prove a theorem that establishes sufficient conditions for the existence of such lifting, we first need the following two lemmas.

**Remark 39.** *In the statement of the next results, we use the following notation: let  $F, T$  be functors and  $ev_F, ev_T$  their evaluation maps, leading to  $\mathcal{V}$ -predicate liftings  $\hat{F}, \hat{T}$ . Notice that  $\widehat{T \circ F} := \hat{T} \circ \hat{F}$  and  $\widehat{F \circ T} := \hat{F} \circ \hat{T}$  are liftings of the composite functors  $T \circ F$ , respectively  $F \circ T$ . As before we are using the convention  $\bar{F} = W(\hat{F})$ ,  $\bar{T} = W(\hat{T})$ . Additionally, we will denote by  $\overline{T \circ F}$  and  $\overline{F \circ T}$  the Wasserstein liftings obtained from  $\widehat{T \circ F}$ , respectively,  $\widehat{F \circ T}$  as in Section 5. That is,  $\overline{T \circ F} := W(\widehat{T \circ F})$ ,  $\overline{F \circ T} := W(\widehat{F \circ T})$ .*

**Lemma 40.** *Whenever  $\hat{T} \circ \Sigma_{\lambda_X^F} \leq \Sigma_{T\lambda_X^F} \circ \hat{T}$ , the identity natural transformation  $TF \Rightarrow TF$  lifts to  $\overline{T \circ F} \Rightarrow \overline{T \circ F}$ .*

*Proof.* We have to show that the (identity) maps underlying the natural transformation are non-expansive, in particular  $\overline{T \circ F}(r) \leq \overline{T \circ F}(r)$  for every  $r$  in  $\mathcal{V}\text{-Rel}_X$ .

Let  $\lambda_X^F : F(X \times X) \rightarrow FX \times FX$ ,  $\lambda_{FX}^T : T(FX \times FX) \rightarrow TFX \times TFX$ ,  $\lambda_X^{TF} : TF(X \times X) \rightarrow TFX \times TFX$  be the natural transformations on which the three liftings are based. From the uniqueness

of the mediating morphism of the product, we obtain  $\lambda_X^{TF} = \lambda_{FX}^T \circ T\lambda_X^F$ . This yields that  $\Sigma_{\lambda_{FX}^{TF}} = \Sigma_{\lambda_{FX}^T} \circ \Sigma_{T\lambda_X^F}$  and allows us to prove:

$$\begin{aligned} \overline{T} \circ \overline{F} &= \iota_{TFX}^{-1} \circ \Sigma_{\lambda_{FX}^T} \circ \widehat{T} \circ \iota_{FX} \circ \iota_{FX}^{-1} \circ \Sigma_{\lambda_X^F} \circ \widehat{F} \circ \iota_X \\ &= \iota_{TFX}^{-1} \circ \Sigma_{\lambda_{FX}^T} \circ \widehat{T} \circ \Sigma_{\lambda_X^F} \circ \widehat{F} \circ \iota_X \\ &\leq \iota_{TFX}^{-1} \circ \Sigma_{\lambda_{FX}^T} \circ \Sigma_{T\lambda_X^F} \circ \widehat{T} \circ \widehat{F} \circ \iota_X \\ &= \iota_{TFX}^{-1} \circ \Sigma_{\lambda_{FX}^{TF}} \circ \widehat{TF} \circ \iota_X = \overline{T \circ F}. \end{aligned} \quad \square$$

**Lemma 41.** *The identity natural transformation  $F \circ T \Rightarrow F \circ T$  lifts to  $\overline{F \circ T} \Rightarrow \overline{F} \circ \overline{T}$ .*

*Proof.* It always holds that  $\Sigma_{Ff} \circ \widehat{F} \leq \widehat{F} \circ \Sigma_f$  for all  $f : X \rightarrow Y$ . Indeed, by (3),  $\widehat{F} \circ f^* \leq (Ff)^* \circ \widehat{F}$  holds in  $\mathcal{V}$ -Rel. Then, using the fact that  $\Sigma_f, f^*$  (respectively  $\Sigma_{Ff}, (Ff)^*$ ) are adjoint, we obtain the desired inequality.

The rest of the proof is analogous to Lemma 40. In particular, the inequality in the computation turns into an equality. □

In fact, if we combine Lemmas 40 and 41 (with the roles of  $T, F$  switched), we obtain compositionality of the Wasserstein lifting under the additional requirement of Lemma 40. In other words,  $W(\widehat{T}) \circ W(\widehat{F}) = \overline{T \circ F} = \overline{T} \circ \overline{F} = W(\widehat{T \circ F})$ .

**Theorem 42.** *Assume the natural transformation  $\zeta : T \circ F \Rightarrow F \circ T$  lifts to a natural transformation  $\widehat{\zeta} : \widehat{T} \circ \widehat{F} \Rightarrow \widehat{F} \circ \widehat{T}$  and that we have  $\widehat{T} \circ \Sigma_{\lambda_X^F} \leq \Sigma_{T\lambda_X^F} \circ \widehat{T}$ . Then,  $\zeta$  lifts to a distributive law  $\overline{\zeta} : \overline{T} \circ \overline{F} \Rightarrow \overline{F} \circ \overline{T}$ .*

*Proof.* We split the proof obligation into three parts:

$$\overline{T} \circ \overline{F} \xrightarrow{(1)} \overline{T \circ F} \xrightarrow{(2) \widehat{\zeta}} \overline{F \circ T} \xrightarrow{(3)} \overline{F} \circ \overline{T}.$$

- (1) lifts the identity natural transformation on  $T \circ F$ . Its existence is proved using the hypothesis  $\widehat{T} \circ \Sigma_{\lambda_X^F} \leq \Sigma_{T\lambda_X^F} \circ \widehat{T}$ , see Lemma 40.
- (2) is obtained by applying Lemma 26 to  $\widehat{\zeta}$ . Such liftings have already been studied in Baldan et al. (2015).
- (3) lifts the identity natural transformation on  $F \circ T$ . It exists by Lemma 41. □

The first requirement of the previous theorem holds for the canonical  $\mathcal{V}$ -Pred-liftings under mild assumptions on  $F$  and  $T$ .

**Proposition 43.** *Assume that  $\zeta : T \circ F \Rightarrow F \circ T$  is a natural transformation and that, furthermore,  $T$  preserves weak pullbacks and  $F$  preserves intersections. Then,  $\zeta$  lifts to a natural transformation  $\widehat{\zeta} : \widehat{T}_{\text{can}} \circ \widehat{F}_{\text{can}} \Rightarrow \widehat{F}_{\text{can}} \circ \widehat{T}_{\text{can}}$ .*

*Proof.* We use the following notations:

$$ev_{TF} := ev_{\text{can}}^T \circ T(ev_{\text{can}}^F)$$

and

$$ev_{FT} := ev_{\text{can}}^F \circ F(ev_{\text{can}}^T).$$

Notice that  $ev_{TF}$  and  $ev_{FT}$  are exactly the evaluation maps corresponding to the liftings  $\widehat{T}_{\text{can}} \circ \widehat{F}_{\text{can}}$ , respectively  $\widehat{F}_{\text{can}} \circ \widehat{T}_{\text{can}}$ . Using the second part of Lemma 26, it suffices to show  $ev_{TF} \leq ev_{FT} \circ \zeta_{\mathcal{V}}$ .

We will consider the inclusion maps  $\text{true}_r : \uparrow r \rightarrow \mathcal{V}$  and write  $t \in F(\uparrow r)$  for  $t \in \text{true}_r(\uparrow r)$ .



We first consider the diagram below. We will show that the dotted arrow exists and that the resulting square is a weak pullback.

$$\begin{array}{ccc}
 F(\uparrow r) & \xrightarrow{F\text{true}_r} & F\mathcal{V} \\
 \downarrow \text{ev}_{\text{can}}^F|_{\uparrow r} & & \downarrow \text{ev}_{\text{can}}^F \\
 \uparrow r & \xrightarrow{\text{true}_r} & \mathcal{V}
 \end{array}$$

Let  $t \in F(\uparrow r)$ . This means that  $\text{ev}_{\text{can}}^F(t) = \bigvee \{s \mid t \in F(\uparrow s)\} \geq r$ , since the set contains  $r$  itself. Hence,  $\text{ev}_{\text{can}}^F$  restricts to  $\text{ev}_{\text{can}}^F|_{\uparrow r}$ .

In order to show that the square is a weak pullback take  $t \in F\mathcal{V}$  such that  $\text{ev}_{\text{can}}^F(t) = \bigvee \{s \mid t \in F(\uparrow s)\} = \bar{s} \geq r$ , hence  $\bar{s} \geq r$ . We have to show that  $t \in F(\uparrow r)$ , that is, that  $r$  is contained in the set, which we will do by showing that  $\{s \mid t \in F(\uparrow s)\}$  is downward-closed and contains its supremum. The set  $\{s \mid t \in F(\uparrow s)\}$  is downward-closed since  $F$  as a Set-functor preserves injections with non-empty domains, and hence,  $s' \leq s$  implies  $\uparrow s \subseteq \uparrow s'$ , and thus,  $F(\uparrow s) \subseteq F(\uparrow s')$ . It follows that  $t \in F(\uparrow s)$  implies  $t \in F(\uparrow s')$ . If  $F$  preserves intersections, the set contains its supremum:  $\uparrow \bar{s} = \bigcap \{\uparrow s \mid t \in F(\uparrow s)\}$ , and hence,  $F(\uparrow \bar{s}) = \bigcap \{F(\uparrow s) \mid t \in F(\uparrow s)\} \ni t$ . This implies  $t \in F(\uparrow \bar{s}) \subseteq F(\uparrow r)$ , as desired.

Similarly one obtains such a commuting square (not necessarily a weak pullback) for  $T$  and  $\text{ev}_{\text{can}}^T$ . This results in the following diagram where the right-hand square and the upper ‘square’ commute and the left-hand square is a weak pullback (using pullback preservation of  $T$ ).

$$\begin{array}{ccccccc}
 & & & \zeta_{\uparrow r} & & & \\
 & & & \curvearrowright & & & \\
 TF(\uparrow r) & \xrightarrow{TF\text{true}_r} & TF\mathcal{V} & \xrightarrow{\zeta_{\mathcal{V}}} & FT\mathcal{V} & \xleftarrow{FT\text{true}_r} & FT(\uparrow r) \\
 \downarrow T(\text{ev}_{\text{can}}^F|_{\uparrow r}) & & \downarrow T\text{ev}_{\text{can}}^F & & \downarrow F\text{ev}_{\text{can}}^T & & \downarrow F(\text{ev}_{\text{can}}^T|_{\uparrow r}) \\
 T(\uparrow r) & \xrightarrow{T\text{true}_r} & T\mathcal{V} & & F\mathcal{V} & \xleftarrow{F\text{true}_r} & F(\uparrow r)
 \end{array}$$

In order to prove that

$$\text{ev}_{TF}^T = \text{ev}_{\text{can}}^T \circ T\text{ev}_{\text{can}}^F \leq \text{ev}_{\text{can}}^F \circ F\text{ev}_{\text{can}}^T \circ \zeta_{\mathcal{V}} = \text{ev}_{FT}^T \circ \zeta_{\mathcal{V}},$$

let  $t \in TF\mathcal{V}$ . Since

$$\text{ev}_{\text{can}}^T(T\text{ev}_{\text{can}}^F(t)) = \bigvee \{r \mid T\text{ev}_{\text{can}}^F(t) \in T(\uparrow r)\}$$

and

$$\text{ev}_{\text{can}}^F(F\text{ev}_{\text{can}}^T(\zeta_{\mathcal{V}}(t))) = \bigvee \{r \mid F\text{ev}_{\text{can}}^T(\zeta_{\mathcal{V}}(t)) \in F(\uparrow r)\}$$

it suffices to show that

$$T\text{ev}_{\text{can}}^F(t) \in T(\uparrow r) \text{ implies } F\text{ev}_{\text{can}}^T(\zeta_{\mathcal{V}}(t)) \in F(\uparrow r).$$

So let  $T\text{ev}_{\text{can}}^F(t) \in T(\uparrow r)$  and the fact that the left-hand square is a weak pullback implies that there exists  $t' \in TF(\uparrow r)$  with  $TF\text{true}_r(t') = t$ .

Then, using naturality of  $\zeta$ , we obtain

$$\begin{aligned}
 F\text{ev}_{\text{can}}^T(\zeta_{\mathcal{V}}(t)) &= F\text{ev}_{\text{can}}^T(\zeta_{\mathcal{V}}(TF\text{true}_r(t'))) \\
 &= F\text{ev}_{\text{can}}^T(FT\text{true}_r(\zeta_{\uparrow r}(t'))) \\
 &= F\text{true}_r(F(\text{ev}_{\text{can}}^T|_{\uparrow r})(\zeta_{\uparrow r}(t'))) \\
 &\in F(\uparrow r)
 \end{aligned}$$

□

The next proposition establishes sufficient conditions for the second hypothesis of Theorem 42. We require a property on  $\mathcal{V}$  that holds for the quantales in Example 7 and was also assumed in Hofmann (2007). Given  $u, v \in \mathcal{V}$  we write  $u \lll v$  ( $u$  is totally below  $v$ ) if for every  $W \subseteq \mathcal{V}$ ,  $v \leq \bigvee W$  implies that there exists  $w \in W$  with  $u \leq w$ . The quantale  $\mathcal{V}$  is *constructively completely distributive* iff for all  $v \in \mathcal{V}$  it holds that  $v = \bigvee \{u \in \mathcal{V} \mid u \lll v\}$ .

**Example 44.** In the real interval  $[0, 1]$ , the order  $\lll$  coincides with  $>_{\mathbb{R}}$ , whereas in a powerset lattice  $\mathcal{P}M$  we have that  $M_1 \lll M_2$  for  $M_1, M_2 \subseteq M$  whenever  $M_1 \subseteq M_2$  and  $M_1$  contains at most one element. Both lattices are constructively completely distributive.

We now show a result for the canonical predicate lifting. In Appendix B.2, we prove a more general statement (Proposition 56) for arbitrary predicate liftings of  $T$  that is useful to guarantee the result for interesting liftings, such as the one in Example 23.

**Proposition 45.** Assume that  $T$  preserves weak pullbacks and that  $\mathcal{V}$  is constructively completely distributive. Then  $\widehat{T}_{\text{can}} \circ \Sigma_f \leq \Sigma_{Tf} \circ \widehat{T}_{\text{can}}$ .

*Proof.* In Appendix B.2, we prove a more general result (Proposition 56). Proposition 45 follows thus from Proposition 56, whose conditions are shown to be satisfied by the canonical lifting in Lemmas 57 and 58. □

Combining Theorem 42 and Propositions 37, 43 and 45, we conclude:

**Theorem 46.** Let  $(X, \alpha : TX \rightarrow X, \xi : X \rightarrow FX)$  be a bialgebra for a natural transformation  $\zeta : T \circ F \Rightarrow F \circ T$ . If  $\mathcal{V}$  is constructively completely distributive,  $T$  preserves weak pullbacks and  $F$  preserves intersections, then  $f = \overline{T}_{\text{can}} \circ \Sigma_{\alpha}$  is compatible with respect to  $b = \overline{F}_{\text{can}} \circ \xi^*$ .

When  $\alpha$  is the free algebra for a signature  $\mu_X : T_{\Sigma} T_{\Sigma} X \rightarrow T_{\Sigma} X$  (as in Example 33) and  $F$  is a suitable functor, the above theorem guarantees that up-to contextual closure is compatible with respect to  $b$ . Then, by (38), it holds that for all terms  $t_1, t_2$  and unary contexts  $C$  that

$$vb(t_1, t_2) \leq ctx(vb)(C(t_1), C(t_2)) \leq vb(C(t_1), C(t_2)),$$

where the first inequality follows from the definition of  $ctx$  (Example 33) and the second from (38), since  $ctx$  is  $b$ -compatible.

For  $\mathcal{V} = 2$ , since the canonical quantitative lifting coincides with the relational lifting of the canonical predicate lifting, then  $vb$  is exactly the standard coalgebraic notion of behavioural equivalence (Hermida and Jacobs 1998). Therefore, the above inequality just means that behavioural equivalence is a congruence.

For  $\mathcal{V} = [0, \infty]$  instead, this property boils down to *non-expansiveness* of contexts with respect to the behavioural metric. It is worth to mention that this property often fails in probabilistic process algebras (see e.g. Gebler et al. 2015) when taking the standard Wasserstein lifting which, as shown in Example 23, is *not* based on the canonical one. We leave as future work to explore the implications of this insight.

### 7. Example: Distance Between Regular Languages

We will now work out the quantitative version of the up-to congruence technique for non-deterministic automata. We consider the shortest-distinguishing-word-distance  $d_{\text{sdw}}$ , proposed in Section 2. As explained, we will assume an on-the-fly determinization of the non-deterministic automaton, that is, formally we will work with a coalgebra that corresponds to a deterministic automaton on which we have a join-semilattice structure.

We explain next the various ingredients of the example:

**Coalgebra and algebra.** As outlined in Section 2 and Example 3, the determinization of an NFA with state space  $Q$  is a bialgebra  $(X, \alpha, \xi)$  for the distributive law  $\zeta_X : \mathcal{P}(2 \times X^A) \rightarrow 2 \times (\mathcal{P}X)^A$ ,

where  $X = \mathcal{P}Q$ ,  $\alpha : \mathcal{P}X \rightarrow X$  is given by union and  $\xi : X \rightarrow 2 \times X^A$  specifies the DFA structure of the determinization. Hence, we instantiate the generic results in the previous section with  $TX = \mathcal{P}X$ ,  $FX = 2 \times X^A$  and  $\zeta$  as defined in Example 3.

**Lifting the functors.** We take the quantale  $\mathcal{V} = [0, 1]$  (Example 7) and consider the Wasserstein liftings of the endofunctors  $F$  and  $T$  to  $\mathcal{V}$ -Rel corresponding to the following evaluation maps:

- $ev_F(b, f) := c \cdot \max_{a \in A} f(a)$ , where  $b \in \{0, 1\}$ ,  $f : A \rightarrow [0, 1]$  and  $c$  is the constant used in  $d_{sdw}$ , and,
- $ev_T := ev_{can}^{\mathcal{P}} = \sup$ , the canonical evaluation map as in Example 18.

These are monotone evaluation maps that satisfy the hypothesis of Theorem 27:

**Lemma 47.** *The evaluation maps  $ev_F$  and  $ev_T$  defined in Section 7 induce liftings which satisfy the requirements of Theorem 27.*

*Proof.* These evaluation maps satisfy the required properties:  $ev_T$  is the canonical evaluation map (see Section 5), thus the statement follows from Proposition 32. For  $ev_F$ , notice that this is of the form  $g \circ ev$ , where  $ev$  satisfies the requirements of Proposition 53 in Appendix A (since it is canonical) and  $g : \mathcal{V} \rightarrow \mathcal{V}$  with  $g(r) = c \cdot r$  is monotone,  $g(a \otimes b) \geq g(a) \otimes g(b)$  and  $g(1) \geq 1$ . It is thus straightforward to see that  $ev_F$  fulfils the conditions of Proposition 53 and thus the corresponding lifting, those of Theorem 27.  $\square$

Hence, the corresponding Wasserstein liftings restrict to  $\mathcal{V}$ -Cat. We computed the Wasserstein lifting of  $T = \mathcal{P}$  in Example 24: applying the lifted functor  $\bar{T}$  to a map  $d : X \times X \rightarrow [0, 1]$  gives us the Hausdorff distance, that is,  $\bar{T}(d)(X_1, X_2) = d^H(X_1, X_2)$ , where  $X_1, X_2 \subseteq X$  and  $d^H$  denotes the Hausdorff metric based on  $d$ . On the other hand, the Wasserstein lifting of  $F$  corresponding to  $ev_F$  associates to a metric  $d : X \times X \rightarrow [0, 1]$  the metric  $\bar{F}(d) : FX \times FX \rightarrow [0, 1]$  given by

$$((b_1, f_1), (b_2, f_2)) \mapsto \begin{cases} 1 & \text{if } b_1 \neq b_2 \\ \max_{a \in A} c \cdot \{d(f_1(a), f_2(a))\} & \text{otherwise} \end{cases}$$

**Fixpoint equation.** The map  $b$  for the fixpoint equation was defined in Section 6 as the composite  $\xi^* \circ \bar{F}$ . Using the above lifting  $\bar{F}$ , this computation yields exactly the map  $b$  defined in (2), whose largest fixpoint (smallest with respect to the natural order on the reals) is the shortest-distinguishing-word-distance introduced in Section 2.

**Up-to technique.** The next step is to determine the map  $f$  introduced in Section 6 for the up-to technique and defined as the composite  $\Sigma_\alpha \circ \bar{T}$  on  $\mathcal{V}$ -Rel. Combining the definition of the direct image functors on  $\mathcal{V}$ -Rel with the lifting  $\bar{T}$ , we obtain that

$$f(d)(x_1, x_2) = \inf\{d^H(X_1, X_2) \mid X_1, X_2 \subseteq X, \alpha(X_i) = x_i\}$$

for all maps  $d : X \times X \rightarrow [0, 1]$ . To show that  $f(d)(Q_1, Q_2) \leq_{\mathbb{R}} r$  for two sets  $Q_1, Q_2 \subseteq Q$  (i.e.  $Q_1, Q_2 \in X$ ) and a constant  $r$ , we use the following rules:

$$f(d)(\emptyset, \emptyset) \leq_{\mathbb{R}} r \quad \frac{d(Q_1, Q_2) \leq_{\mathbb{R}} r}{f(d)(Q_1, Q_2) \leq_{\mathbb{R}} r} \quad \frac{f(d)(Q_1, Q_2) \leq_{\mathbb{R}} r \quad f(d)(Q'_1, Q'_2) \leq_{\mathbb{R}} r}{f(d)(Q_1 \cup Q'_1, Q_2 \cup Q'_2) \leq_{\mathbb{R}} r}$$

**Lifting of distributive law.** In order to prove that the distributive law lifts to  $\mathcal{V}$ -Rel and hence that the up-to technique is sound by virtue of Proposition 37, we can prove that the two conditions of Theorem 42 are met by the  $\mathcal{V}$ -Pred liftings of  $F$  and  $T$  corresponding to the evaluation maps  $ev_F$  and  $ev_T$ :

**Lemma 48.** Assume  $\widehat{F}$  and  $\widehat{T}$  are the  $[0, 1]$ -Pred liftings of  $FX = 2 \times X^A$  and  $TX = \mathcal{P}X$  which correspond to the evaluation maps  $ev_F$  and  $ev_T$  defined in the example in Section 7. Then, we have:

- (1)  $\widehat{T} \circ \Sigma_{\lambda_X^F} \leq \Sigma_{T\lambda_X^F} \circ \widehat{T}$ , and,
- (2)  $\zeta : T \circ F \Rightarrow F \circ T$  lifts to a natural transformation  $\widehat{\zeta} : \widehat{T} \circ \widehat{F} \Rightarrow \widehat{F} \circ \widehat{T}$

*Proof.* Recall that on the quantale  $[0, 1]$  the quantale order is the reversed order on the reals, so in order to avoid confusion we use  $\leq, \vee, \wedge$  in the quantale and  $\geq_{\mathbb{R}}, \inf, \sup$  in the reals.

(1) To prove the first item, we can rely on Proposition 45, since  $\widehat{T}_{\text{can}}$  is the canonical lifting and we are working in the quantale  $\mathcal{V} = [0, 1]$ , which is constructively completely distributive.

(2) Recall that  $\widehat{T} \circ \widehat{F}$  is a lifting of  $T \circ F$  which corresponds to the evaluation map  $ev_{TF} = ev_T \circ T(ev_F)$ . Similarly,  $\widehat{F} \circ \widehat{T}$  corresponds to the evaluation map  $ev_{FT} = ev_F \circ F(ev_T)$ . The existence of  $\widehat{\zeta}$  is then equivalent to the inequality

$$ev_{TF} \geq_{\mathbb{R}} ev_{FT} \circ \zeta_{\mathcal{V}} . \tag{43}$$

Here we are almost in the setting of canonical liftings treated in Proposition 43, apart from the fact that  $ev_F = g \circ ev_{\text{can}}^F$ , where the function  $g$  is given by  $g(r) = c \cdot r$ . Recall  $ev_T = ev_{\text{can}}^T$ . Furthermore,  $T$  preserves weak pullbacks and  $F$  preserves intersections, hence by (the proof of) Proposition 43, we know that

$$ev_{\text{can}}^T \circ T(ev_{\text{can}}^F) \geq_{\mathbb{R}} ev_{\text{can}}^F \circ F(ev_{\text{can}}^T) \circ \zeta_{\mathcal{V}} .$$

In order to obtain the desired lifting of natural transformations, we first notice that  $ev_{\text{can}}^T \circ Tg = g \circ ev_{\text{can}}^T$ . Indeed, for all  $R \subseteq [0, 1]$  we have

$$ev_{\text{can}}^T(Tg(R)) = \sup c \cdot R = c \cdot \sup R = g(ev_{\text{can}}^T(R)) .$$

To conclude, we use the above equalities and the monotonicity of  $g$ :

$$\begin{aligned} ev_{TF} &= ev_T \circ T(ev_F) \\ &= ev_{\text{can}}^T \circ T(g) \circ T(ev_F) \\ &\geq_{\mathbb{R}} g \circ ev_{\text{can}}^T \circ T(ev_F) \\ &\geq_{\mathbb{R}} g \circ ev_{\text{can}}^F \circ F(ev_{\text{can}}^T) \circ \zeta_{\mathcal{V}} \\ &= ev_F \circ F(ev_{\text{can}}^T) \circ \zeta_{\mathcal{V}} \\ &= ev_{FT} \circ \zeta_{\mathcal{V}} . \end{aligned} \quad \square$$

Everything combined, we obtain a sound up-to technique, which implies that the reasoning in Section 2 is valid. Furthermore, as the example shows, the up-to technique can significantly simplify behavioural distance arguments and speed up computations.

### 8. A Detailed Comparison with Chatzikokolakis et al. (2016)

In this appendix, we discuss in details the relationship between our work and Chatzikokolakis et al. (2016) where a general framework of up-to techniques for behavioural metric is introduced.

The systems of interest in Chatzikokolakis et al. (2016) are probabilistic automata which are known (Bartels et al. 2004) to be coalgebras for the functor  $\mathcal{P}(A \times \mathcal{D}(-))$ . The behavioural metrics under consideration are defined as the greatest fixed points of

$$b : \mathcal{V}\text{-Rel}_X \xrightarrow{K} \mathcal{V}\text{-Rel}_{\mathcal{D}(X)} \xrightarrow{\overline{\mathcal{P}(A \times \overline{\phantom{x}})}} \mathcal{V}\text{-Rel}_{\mathcal{D}(\mathcal{D}(X))^A} \xrightarrow{\xi^*} \mathcal{V}\text{-Rel}_X \tag{44}$$

where  $\xi : X \rightarrow \mathcal{P}(A \times \mathcal{D}(X))$  is a probabilistic automaton,  $\overline{\mathcal{P}(A \times -)}$  is the canonical lifting of  $\mathcal{P}(A \times -)$  (based on the Hausdorff distance, Example 24) and  $K$  is some lifting of  $\mathcal{D}$ . Please note that the quantale  $\mathcal{V}$  in Chatzikokolakis et al. (2016) is  $[0, \infty]$  (Example 7) so the ordering used in this paper and the one in Chatzikokolakis et al. (2016) are always inverted.

Observe that the definition of  $b$  as in (44) is an instance of (39) by taking  $\tilde{F} = K \circ \overline{\mathcal{P}(A \times -)}$ . It is worth to mention that  $K$  in Chatzikokolakis et al. (2016) is not arbitrary, but it is supposed to be an instance of a parametric construction called *generalized Kantorovic metric*. For a certain value of the parameter, this coincides (via the well-known duality) with the Wasserstein metric from transportation theory (Example 23).

The authors of Chatzikokolakis et al. (2016) introduced several basic techniques – which can be easily defined in our framework, for example, metric closure (Section 6), convex closure (Example 34) or contextual closure (Example 33) – and combine them via composition ( $\circ$ ), supremum ( $\bigvee$ ) and chaining ( $\cdot$ ). In Proposition 38, we have provided sufficient conditions ensuring that  $\cdot$  preserves compatibility. The same result for  $\circ$  and  $\bigvee$  follows immediately from the standard theory of compatible up-to techniques (Pous and Sangiorgi 2011). This is not the case for Chatzikokolakis et al. (2016), where these results need novel proofs since the basic notions of up-to techniques and compatibility (or respectfulness) do not fit within the standard lattice-theoretic framework.

Indeed in Chatzikokolakis et al. (2016), an up-to technique is defined to be some map  $f$  of type  $\mathcal{V}\text{-Rel}_{\mathcal{D}(X)} \rightarrow \mathcal{V}\text{-Rel}_{\mathcal{D}(X)}$  and a bisimulation up-to  $f$  to be some  $d \in \mathcal{V}\text{-Rel}_X$  such that  $d \leq (\xi^* \circ \overline{\mathcal{P}}) \circ f \circ K(d)$ .<sup>3</sup> Soundness is defined in the expected way. The notion to prove soundness (Definition 5 in Chatzikokolakis et al. 2016) amounts to the following, modulo the usual difference between compatibility and respectfulness (that is well-known and discussed in several papers Bonchi et al. 2017b; Pous 2016)

**Definition 49.** *A monotone map  $f : \mathcal{V}\text{-Rel}_{\mathcal{D}(X)} \rightarrow \mathcal{V}\text{-Rel}_{\mathcal{D}(X)}$  is a well-behaved up-to technique iff there exists an  $f' : \mathcal{V}\text{-Rel}_X \rightarrow \mathcal{V}\text{-Rel}_X$  such that*

- (1)  $f \circ K \leq K \circ f'$  and
- (2)  $f' \circ b \leq b \circ f'$ .

Observe that whenever  $f$  is well-behaved, a bisimulation up-to  $f$  in the sense of Chatzikokolakis et al. (2016) can be transformed into a bisimulation up-to  $f'$  in our sense by means of the first item:

$$d \leq (\xi^* \circ \overline{\mathcal{P}(A \times -)}) \circ f \circ K(d) \leq (\xi^* \circ \overline{\mathcal{P}(A \times -)}) \circ K \circ f'(d) = b(f'(d)). \tag{45}$$

Moreover, thanks to the second item,  $f'$  is compatible w.r.t.  $b$ .

This observation shows that the techniques in Chatzikokolakis et al. (2016) can be reformulated within the standard theory of Pous and Sangiorgi (2011) and thus proved compatible by means of our framework.

**Lemma 50.** *Consider a probabilistic automaton and let  $K$  denote a convex (in the sense of Chatzikokolakis et al. 2016) lifting of the probability distribution functor. Then, a bisimulation metric up-to convex closure in the sense of Chatzikokolakis et al. (2016) is exactly a bisimulation metric, that is, a post-fixpoint of  $b$  in (44).*

*Proof.* As above, let  $\xi : X \rightarrow \mathcal{P}(A \times \mathcal{D}(X))$  denote the coalgebra structure corresponding to the probabilistic automaton. The up-to convex closure is defined as in Example 34. Recall that a bisimulation metric up-to convex closure in the sense of Chatzikokolakis et al. (2016) is a bisimulation metric  $d$  such that  $d$  progresses to  $cvx \circ K(d)$ , written using the notation in Chatzikokolakis et al. (2016, Definition 2) as  $d \rightsquigarrow cvx \circ K(d)$ . Spelling out that definition, we obtain that, in the quantale order (i.e. the reversed of the order on the reals used Chatzikokolakis et al. 2016), we have

$$d \leq \xi^* \circ \overline{\mathcal{P}(A \times -)} \circ cvx \circ K(d). \tag{46}$$

On the other hand, the respectfulness of  $cvx$  – established via Chatzikokolakis et al. (2016, Theorem 11) – uses the fact that for all  $d \in \mathcal{V}\text{-Rel}_X$  we have that  $K(d)$  is convex; hence, the  $f'$  used above is simply the identity function on  $\mathcal{V}\text{-Rel}_{\mathcal{D}X}$ . In other words, we have

$$cvx \circ K(d) \leq K(d) \tag{47}$$

Combining (46) and (47), we obtain that

$$d \leq \xi^* \circ \overline{\mathcal{P}(A \times -)} \circ K(d),$$

or equivalently, that  $d$  is simply a bisimulation metric. □

### 9. Related and Future Work

Up-to techniques for behavioural metrics in a probabilistic setting have been considered in Chatzikokolakis et al. (2016) using a generalization of the Kantorovich lifting (Chatzikokolakis et al. 2014). In Section 6, we have shown that the basic techniques introduced in Chatzikokolakis et al. (2016) (e.g. metric closure, convex closure and contextual closure) as well as the ways to combine them (composition, join and chaining) naturally fit within our framework. The main difference with our approach – beyond the fact that we consider arbitrary coalgebras while the results in Chatzikokolakis et al. (2016) just cover coalgebras for a fixed functor – is that the definition of up-to techniques and the criteria to prove their soundness do not fit within the standard framework of Pous and Sangiorgi (2011). Nevertheless, as illustrated by a detailed comparison in Section 8, the techniques of Chatzikokolakis et al. (2016) can be reformulated within the standard theory and thus proved sound by means of our framework. An important observation brought to light by compositional methodology inherent to the fibrational approach is that for probabilistic automata, a bisimulation metric up-to convexity in the sense of Chatzikokolakis et al. (2016) is exactly a bisimulation metric, see Lemma 50. Nevertheless, the up-to convex closure technique can find meaningful applications in linear, trace-based behavioural metrics (see Baldan et al. 2015).

The Wasserstein (respectively Kantorovich) lifting of the distribution functor involving couplings was first used for defining behavioural pseudo-metrics using final coalgebras in van Breugel and Worrell (2001). Our work is based instead on liftings for arbitrary functors, a problem that has been considered in several works (see e.g. Balan et al. 2015; Baldan et al. 2018; Hofmann 2007; Katsumata and Sato 2015), despite with different shades. The closest to our approach are Sprunger et al. (2018, 2021), Hofmann (2007) and Baldan et al. (2018) that we discuss next.

In Sprunger et al. (2018, 2021), the authors introduce a generalization of the Wasserstein lifting for  $CLat_{\wedge}$ -fibrations, inspired by our original conference publication (Bonchi et al. 2018b). This construction uses the fact that the functor  $\Delta$  has a left adjoint  $2 \times \_$ , resulting in an adjunction that can be lifted to an adjunction between  $\mathcal{V}\text{-Pred}$  and  $\mathcal{V}\text{-Rel}$ . Then,  $\bar{F}$  can be obtained as a pushforward (direct image) along the natural transformation  $\theta: LF\Delta \Rightarrow F$  and correctness is ensured by Hermida’s adjunction lifting theorem.

In Hofmann (2007), Hofmann introduces a generalization of the Barr extension (of Set-functors to Rel), namely he defines extensions of Set-monads to the bicategory of  $\mathcal{V}$ -matrices, in which 0-cells are sets and the  $\mathcal{V}$ -relations are 1-cells. Some of the definitions and techniques do overlap between the developments in Hofmann (2007), and the results we presented in Section 5. However, there are also some (subtle) differences which would not allow us to use off the shelf his results.

First, in order to reuse the results in Bonchi et al. (2014), we need to recast the theory in a fibrational setting, rather than the bicategorical setting of Hofmann (2007). The definition of *topological theory* (Hofmann 2007, Definition 3.1) comprises what we call an evaluation map, but which additionally has to satisfy various conditions. An important difference with what we do is that the condition  $(Q_{\vee})$  in the aforementioned definition entails that the predicate lifting one would obtain from such an evaluation map would be an *opfibred lifting*, rather than a fibred lifting as in our setting. Indeed, the condition  $(Q_{\vee})$  can be equivalently expressed in terms of a natural

transformation involving the *covariant* functor  $P_{\mathcal{V}}$ , as opposed to the *contravariant* one  $\mathcal{V}^-$  that we used in Section 5.1. Lastly, in our framework we need to work with arbitrary functors, not necessarily carrying a monad structure.

In Baldan *et al.* (2018), we provided a generic construction for the Wasserstein lifting of a functor to the category of pseudo-metric spaces, rather than on arbitrary quantale-valued relations. The realization that this construction is an instance as a change-of-base situation between  $\mathcal{V}$ -Rel and  $\mathcal{V}$ -Pred allows us to exploit the theory in Bonchi *et al.* (2014) for up-to techniques and, as a side result, provides simpler (and cleaner) conditions for the restriction to  $\mathcal{V}$ -Cat (Theorem 27).

We leave for future work several open problems. What is a universal property for the canonical Wasserstein lifting? Secondly, can the Wasserstein liftings presented here be captured in the framework of Balan *et al.* (2015) or Katsumata and Sato (2015)? In particular, the relation to codensity liftings (Katsumata and Sato 2015; Komorida *et al.* 2021, 2019), a generalized notion of Kantorovich lifting, has to be studied in more detail. A result which shows under which conditions our notion of Wasserstein lifting coincides with the codensity lifting, that is a generalized Kantorovich-Rubinstein result (Villani 2009), is non-trivial to obtain and would be of great interest.

We also leave for future work the concrete development of up-to techniques for other quantales than 2 and  $[0, 1]$ . We are particularly interested in weighted automata (Droste *et al.* 2009) over quantales and in conditional transition systems, a variant of featured transition systems.

Another interesting question is whether the rule-based reasoning that we employed in Section 8 can be formally seen as an instance of quantitative algebraic reasoning (Mardare *et al.* 2016; Mio and Vignudelli 2020, Mio *et al.* 2021).

Finally, we plan to study the setting of Bonchi *et al.* (2019), which provides an algebraic theory for specifying systems with nondeterminism, probability and termination. We believe that up-to techniques based on such algebraic theories can be defined in a way similar to the contextual closure studied in Example 33. This would enable us to exhibit proof rules to simplify checking behavioural trace distance in such a context.

**Acknowledgements.** The authors would like to thank the anonymous reviewers for their valuable feedback. Furthermore, we are grateful to Shin-ya Katsumata, Henning Kerstan, Damien Pous and Paolo Baldan for precious suggestions and inspiring discussions and to Thomas Colcombet for the knowledge package used for creating the hyperlinks in this paper.

The first author was supported by the Ministero dell'Universit'a e della Ricerca of Italy under Grant No. 201784YSZ5, PRIN2017 – ASPRA (Analysis of Program Analyses) and partially supported by Universit'a di Pisa Project PRA-2022-2023 FM4HD. The second author acknowledges support from DFG project SpeQt (project number 434050016).

**Competing interests.** The author(s) declare none.

## Notes

- 1 For a context  $C(-_1, \dots, -_n)$  we denote by  $C(t_1, \dots, t_n)$  the term obtained by substituting the term  $t_i$  in the hole  $-_i$ .
- 2 The same observation is present in Hofmann (2007, Theorem 3.3(b)) but in a slightly different setting.
- 3 Note that  $d \leq \xi^* \circ \mathcal{P}(A \times -)(d')$  means, in the language of Chatzikokolakis *et al.* (2016), that  $d$  progresses to  $d'$ .

## References

- Bacci, G., Bacci, G., Larsen, K. and Mardare, R. (2013). Computing behavioral distances, compositionally. In: *MFCS, LNCS*, vol. 8087, Springer, 74–85.
- Bacci, G., Bacci, G., Larsen, K. G., Tribastone, M., Tschaikowski, M. and Vandin, A. (2021). Efficient local computation of differential bisimulations via coupling and up-to methods. In: *LICS, IEEE*, 1–14.
- Balan, A. and Kurz, A. (2011). Finitary functors: From Set to Preord and Poset. In: *CALCO, LNCS*, vol. 6859, Springer, 85–99.
- Balan, A., Kurz, A. and Velebil, J. (2015). Extensions of functors from Set to  $\mathcal{V}$ -cat. In: *CALCO, LIPIcs*, vol. 35, 17–34.
- Baldan, P., Bonchi, F., Kerstan, H. and König, B. (2014). Behavioral metrics via functor lifting. In: *FSTTCS, LIPIcs*, vol. 29, 403–415.
- Baldan, P., Bonchi, F., Kerstan, H. and König, B. (2015). Towards trace metrics via functor lifting. In: *CALCO, LIPIcs*, vol. 35, 35–49.

- Baldan, P., Bonchi, F., Kerstan, H. and König, B. (2018). Coalgebraic behavioral metrics. *Logical Methods in Computer Science* **14** (3). Selected Papers of the 6th Conference on Algebra and Coalgebra in Computer Science (CALCO 2015).
- Baldan, P., König, B. and Padoan, T. (2020). Abstraction, up-to techniques and games for systems of fixpoint equations. In: *CONCUR, LIPIcs*, vol. 171, Schloss Dagstuhl – Leibniz Center for Informatics, 25:1–25:20.
- Barr, M. (1970). Relational algebras. In: *Reports of the Midwest Category Seminar IV*, Springer, 39–55.
- Bartels, F., Sokolova, A. and de Vink, E. (2004). A hierarchy of probabilistic system types. *Theoretical Computer Science* **327** (1–2) 3–22.
- Blanchette, J. C., Bouzy, A., Lochbihler, A., Popescu, A. and Traytel, D. (2017). Friends with benefits – implementing corecursion in foundational proof assistants. In: *ESOP, LNCS*, vol. 10201, Springer, 111–140. .
- Bonchi, F., Ganty, P., Giacobazzi, R. and Pavlovic, D. (2018a). Sound up-to techniques and complete abstract domains. In: *LICS, ACM*, 175–184.
- Bonchi, F., König, B. and Küpper, S. (2017a). Up-to techniques for weighted systems. In: *TACAS, LNCS*, vol. 10205, Springer, 535–552. .
- Bonchi, F., König, B. and Petrisan, D. (2018b). Up-to techniques for behavioural metrics via fibrations. arXiv:1806.11064.
- Bonchi, F., Petrisan, D., Pous, D. and Rot, J. (2014). Coinduction up-to in a fibrational setting. In: *CSL-LICS, ACM*, 1–9. Paper No. 20.
- Bonchi, F., Petrisan, D., Pous, D. and Rot, J. (2017b). A general account of coinduction up-to. *Acta Informatica* **54** (2) 127–190.
- Bonchi, F. and Pous, D. (2013). Checking NFA equivalence with bisimulations up to congruence. In: *POPL, ACM*, 457–468.
- Bonchi, F., Sokolova, A. and Vignudelli, V. (2019). The theory of traces for systems with nondeterminism and probability. In: *LICS, IEEE*, 1–14.
- Chatzikokolakis, K., Gebler, D., Palamidessi, C. and Xu, L. (2014). Generalized bisimulation metrics. In: *CONCUR, LNCS/ARCoSS*, vol. 8704, Springer.
- Chatzikokolakis, K., Palamidessi, C. and Vignudelli, V. (2016). Up-to techniques for generalized bisimulation metrics. In: *CONCUR, LIPIcs*, vol. 59, 35:1–35:14.
- Danielsson, N. (2017). Up-to techniques using sized types. In: *POPL*, vol. 2, ACM, 43:1–43:28.
- de Alfaro, L., Faella, M. and Stoelinga, M. (2004). Linear and branching metrics for quantitative transition systems. In: *ICALP, LNCS*, vol. 3142, Springer, 97–109.
- Desharnais, J., Gupta, V., Jagadeesan, R. and Panangaden, P. (2004). Metrics for labelled Markov processes. *Theoretical Computer Science* **318** (3) 323–354.
- Droste, M., Kuich, W. and Vogler, H. (2009). *Handbook of Weighted Automata*, Monographs in Theoretical Computer Science, Springer, Berlin, Heidelberg.
- Gebler, D., Larsen, K. G. and Tini, S. (2015). Compositional metric reasoning with probabilistic process calculi. In: *FOSSACS, LNCS*, vol. 9034, Springer, 230–245.
- Hasuo, I. (2015). Generic weakest precondition semantics from monads enriched with order. *Theoretical Computer Science* **604** 2–29.
- Hermida, C. and Jacobs, B. (1998). Structural induction and coinduction in a fibrational setting. *Information and Computation* **145** 107–152.
- Hofmann, D. (2007). Topological theories and closed objects. *Advances in Mathematics* **215** (2) 789–824.
- Hofmann, D. and Nora, P. (2020). Hausdorff coalgebras. *Applied Categorical Structures* **28** (5) 773–806.
- Hur, C.-K., Neis, G., Dreyer, D. and Vafeiadis, V. (2013). The power of parameterization in coinductive proof. In: *POPL, ACM*, 193–206.
- Jacobs, B. (1999). *Categorical Logic and Type Theory*, Elsevier, Amsterdam.
- Jacobs, B. (2016). *Introduction to Coalgebra. Towards Mathematics of States and Observations*, Cambridge University Press, Cambridge.
- Jacobs, B., Silva, A. and Sokolova, A. (2015). Trace semantics via determinization. *Journal of Computer and System Sciences* **81** (5) 859–879.
- Katsumata, S. and Sato, T. (2015). Codensity liftings of monads. In: *CALCO, LIPIcs*, vol. 35, 156–170.
- Kelly, M. (1982). *Basic Concepts of Enriched Category Theory*, Lecture Notes in Mathematics, vol. 64, Cambridge University Press, Cambridge.
- Klin, B. (2011). Bialgebras for structural operational semantics: An introduction. *Theoretical Computer Science* **412** (38) 5043–5069.
- Komorida, Y., Katsumata, S., Kupke, C., Rot, J. and Hasuo, I. (2021). Expressivity of quantitative modal logics: Categorical foundations via codensity and approximation. In: *LICS, IEEE*, 1–14.
- Komorida, Y., Katsumata, S.-Y., Hu, N., Klin, B. and Hasuo, I. (2019). Codensity games for bisimilarity. In: *LICS, ACM*, 1–13.
- Lawvere, F. (2002). Metric spaces, generalized logic, and closed categories. *Reprints in Theory and Applications of Categories* **1** 1–37.
- Mardare, R., Panangaden, P. and Plotkin, G. (2016). Quantitative algebraic reasoning. In: *LICS, ACM*, 700–709.



Mémoli, F. (2011). Gromov-Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics* **11** (4) 417–487.

Milner, R. (1989). *Communication and Concurrency*. Prentice Hall, Upper Saddle River, NJ, USA.

Milner, R. and Sangiorgi, D. (1992). Techniques of weak bisimulation up-to. In: *CONCUR*, LNCS, vol. 630, Springer-Verlag.

Mio, M., Sarkis, R. and Vignudelli, V. (2021). Combining nondeterminism, probability, and termination: Equational and metric reasoning. In: *LICS*, IEEE, 1–14.

Mio, M. and Vignudelli, V. (2020). Monads and quantitative equational theories for nondeterminism and probability. In: *CONCUR*, LIPIcs, vol. 171, 28:1–28:18.

Parrow, J. and Weber, T. (2016). The largest respectful function. arXiv:1605.04136.

Pistone, P. (2021). On generalized metric spaces for the simply typed lambda-calculus. In: *LICS*, IEEE, 1–14.

Pous, D. (2007). Complete lattices and up-to techniques. In: *APLAS*, LNCS, vol. 4807, Springer, 351–366.

Pous, D. (2016). Coinduction all the way up. In: *LICS*, ACM, 307–316.

Pous, D. and Sangiorgi, D. (2011). Enhancements of the coinductive proof method. In: Sangiorgi, D. and Rutten, J. (eds.) *Advanced Topics in Bisimulation and Coinduction*, Cambridge University Press.

Pous, D. and Sangiorgi, D. (2019). Bisimulation and coinduction enhancements: A historical perspective. *Formal Aspects of Computing* **31** (6) 733–749.

Rutten, J. (2000). Universal coalgebra: A theory of systems. *Theoretical Computer Science* **249** (1) 3–80.

Sangiorgi, D. (1998). On the bisimulation proof method. *Mathematical Structures in Computer Science* **8** (5) 447–479.

Schröder, L. (2008). Expressivity of coalgebraic modal logic: The limits and beyond. *Theoretical Computer Science* **390** 230–247.

Silva, A., Bonchi, F., Bonsangue, M. M. and Rutten, J. (2013). Generalizing determinization from automata to coalgebras. *Logical Methods in Computer Science* **9** (1) 1–23.

Sprunger, D., Katsumata, S.-y., Dubut, J. and Hasuo, I. (2021). Fibrational bisimulations and quantitative reasoning: Extended version. *Journal of Logic and Computation* **31** (6) 1526–1559.

Sprunger, D., ya Katsumata, S., Dubut, J. and Hasuo, I. (2018). Fibrational bisimulations and quantitative reasoning. In: *CMCS*, LNCS, vol. 11202, Springer, 190–213.

Tini, S., Larsen, K. and Gebler, D. (2017). Compositional bisimulation metric reasoning with probabilistic process calculi. *Logical Methods in Computer Science* **12** (4) 1–38.

Turi, D. and Plotkin, G. (1997). Towards a mathematical operational semantics. In: *LICS*, IEEE, 280–291.

van Breugel, F. and Worrell, J. (2001). Towards quantitative verification of probabilistic transition systems. In: *ICALP*, LNCS, vol. 2076, Springer, 421–432.

van Breugel, F. and Worrell, J. (2005). A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science* **331** 115–142.

Villani, C. (2009). *Optimal Transport – Old and New*, A Series of Comprehensive Studies in Mathematics, vol. 338, Springer, Berlin, Heidelberg.

Wild, P. and Schröder, L. (2022). Characteristic logics for behavioural hemimetrics via fuzzy lax extensions. *Logical Methods in Computer Science* **18** (2) 1–35.

**Appendix A. Proofs and additional material for Section 5**

We now provide several lemmas that are used to prove the results in Section 5. These proofs have been moved to the appendix in order to obtain a more compact structure of the main body of the paper.

**Lemma 51** (Corollary 2.7 in Hofmann 2007). *If  $F : \text{Set} \rightarrow \text{Set}$  is weak pullback-preserving, then the naturality squares of the binatural transformation  $\langle F\pi_1 d, F\pi_2 \rangle : F(X \times Y) \rightarrow FX \times FY$  are weak pullbacks, where  $\pi_1 : X \times Y \rightarrow X$  and  $\pi_2 : X \times Y \rightarrow Y$  denote the projections. In particular, the naturality squares of the natural transformation  $\lambda$  are weak pullbacks.*

*Proof.* Consider morphisms  $f : X \rightarrow X'$  and  $g : Y \rightarrow Y'$ . We want to prove that the square

$$\begin{array}{ccc}
 F(X \times Y) & \xrightarrow{\langle F\pi_1, F\pi_2 \rangle} & F(X) \times F(Y) \\
 \downarrow F(f \times g) & & \downarrow Ff \times Fg \\
 F(X' \times Y') & \xrightarrow{\langle F\pi_1, F\pi_2 \rangle} & F(X') \times F(Y')
 \end{array} \tag{A1}$$

is a weak pullback. To this end, we will consider the following diagram:

$$\begin{array}{ccccc}
 & & F(X \times Y) & & \\
 & & \swarrow F(X \times g) & \searrow F(f \times Y) & \\
 & F(X \times Y') & & & F(X' \times Y) \\
 & \swarrow F\pi_1 & \searrow F(f \times Y') & \swarrow F(X' \times g) & \searrow F\pi_2 \\
 F(X) & & F(X' \times Y') & & F(Y) \\
 & \searrow Ff & \swarrow F\pi_1 & \swarrow F\pi_2 & \searrow Fg \\
 & & F(X') & & F(Y')
 \end{array} \tag{A2}$$

The three squares above are obtained by applying the functor  $F$  to weak pullbacks; hence, by the assumption on  $F$ , they are also weak pullbacks.

Assume  $s' \in F(X')$  and  $t' \in F(Y')$  are such that there exist  $s \in F(X)$ ,  $t \in F(Y)$  and  $u \in F(X' \times Y')$  satisfying  $Ff(s) = s'$ ,  $Fg(t) = t'$  and  $\langle F\pi_1, F\pi_2 \rangle(u) = (s', t')$ . Proving that (A1) is a weak pullback amounts to showing the existence of  $v \in F(X \times Y)$  so that  $F\pi_1(v) = s$ ,  $F\pi_2(v) = t$  and  $F(f \times g)(v) = u$ .

From the fact that the lower left square in (A2) is a weak pullback, we infer the existence of  $u_1 \in F(X \times Y')$  such that  $F\pi_1(u_1) = s$  and  $F(f \times Y')(u_1) = u$ .

Analogously, using that the lower right square in (A2) is a weak pullback we obtain the existence of  $u_2 \in F(X' \times Y)$  such that  $F\pi_2(u_2) = t$  and  $F(X' \times g)(u_2) = u$ .

Since the upper square is also a weak pullback, we deduce the existence of  $v \in F(X \times Y)$  satisfying  $F(X \times g)(v) = u_1$  and  $F(f \times Y)(v) = u_2$ . Upon noticing that  $F\pi_1 \circ F(X \times g) = F(\pi_1)$ ,  $F\pi_2 \circ F(f \times Y) = F(\pi_2)$ , and  $F(f \times Y) \circ F(X \times g) = F(f \times g)$ , we conclude that  $v$  is the element we were looking for in  $F(X \times Y)$ .  $\square$

**Lemma 52.** Assume the functor  $F$  preserves weak pullbacks. The map  $ev_{can}: F\mathcal{V} \rightarrow \mathcal{V}$  is a monotone evaluation map, that is, it is monotone with respect to the order  $\ll$  on  $F\mathcal{V}$  defined in 13 and the order  $\leq$  on  $\mathcal{V}$ .

*Proof.* It is sufficient to show that  $ev_{can}: (F\mathcal{V}, \ll) \rightarrow (\mathcal{V}, \leq)$  is monotone.

Hence, let  $u'_1, u'_2 \in F\mathcal{V}$  such that  $u'_1 \ll u'_2$ , which implies that there exists  $u' \in F(\mathcal{V} \times \mathcal{V})$  with  $F\pi_i(u') = u'_i$ , where  $\pi_i: \leq \rightarrow \mathcal{V}$  are the projections.

We have to show that  $ev_{can}(u'_1) \leq ev_{can}(u'_2)$ . It is sufficient to show that  $u'_1 \in F(\uparrow r)$  implies  $u'_2 \in F(\uparrow r)$ . Assume  $r \in \mathcal{V}$  is such that  $u'_1 \in F(\uparrow r)$ . Then, there exists  $u_1 \in \uparrow r$  such that  $Ftrue_r(u_1) = u'_1$ .

Now consider the diagram on the left in (A3), where

$$e_r: \leq_{\uparrow r} \rightarrow \leq$$

embeds  $\leq$  restricted to  $\uparrow r$  into the full relation. Furthermore, the functions  $\bar{\pi}_i$  are the projections for  $\leq_{\uparrow r}$ . This diagram commutes for  $i = 1, 2$  and is a weak pullback for  $i = 1$ . Hence, the diagram on the right in (A3) is also a weak pullback.

$$\begin{array}{ccc}
 \leq_{\uparrow r} & \xrightarrow{e_r} & \leq \\
 \bar{\pi}_i \downarrow & & \downarrow \pi_i \\
 \uparrow r & \xrightarrow{true_r} & \mathcal{V}
 \end{array}
 \qquad
 \begin{array}{ccc}
 F(\leq_{\uparrow r}) & \xrightarrow{Fe_r} & F(\leq) \\
 F\bar{\pi}_1 \downarrow & & \downarrow F\pi_1 \\
 F(\uparrow r) & \xrightarrow{Ftrue_r} & F\mathcal{V}
 \end{array} \tag{A3}$$

We have  $u_1 \in F(\uparrow r)$  and  $u' \in F(\leq)$  with  $Ftrue_r(u_1) = u'_1 = F\pi_1(u')$ . Hence, there must be an element  $u \in F(\leq_{\uparrow r})$  with  $F\bar{\pi}_1(u) = u_1$  and  $Fe_r(u) = u'$ .

We set  $u_2 = F\bar{\pi}_2(u) \in F(\uparrow r)$  and observe that  $F\text{true}_r(u_2) = F(\text{true}_r \circ \bar{\pi}_2)(u) = F(\pi_2 \circ e_r)(u) = F\pi_2(u') = u'_2$ . This means that  $u'_2 \in F(\uparrow r)$  as required.  $\square$

**Proposition 53.** Assume  $ev: F\mathcal{V} \rightarrow \mathcal{V}$  is monotone evaluation map and let  $\widehat{F}$  be the corresponding fibred lifting of  $F$ . Then we have:

- (1)  $\widehat{F}(p \otimes q) \geq \widehat{F}(p) \otimes \widehat{F}(q)$  holds whenever the map  $\otimes: \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$  is a lax  $F$ -algebra morphism, in the sense that we have a lax diagram:

$$\begin{array}{ccc}
 F(\mathcal{V} \times \mathcal{V}) & \xrightarrow{\lambda_{\mathcal{V}}} & F\mathcal{V} \times F\mathcal{V} & \xrightarrow{ev \times ev} & \mathcal{V} \times \mathcal{V} \\
 F(\otimes) \downarrow & & \geq & & \downarrow \otimes \\
 F\mathcal{V} & \xrightarrow{\quad\quad\quad} & & \xrightarrow{ev} & \mathcal{V}
 \end{array}$$

- (2)  $\widehat{F}(\kappa_X) \geq \kappa_X$  holds whenever the map  $\kappa_{\mathbb{1}}: \mathbb{1} \rightarrow \mathcal{V}$  is a lax algebra morphism, that is, we have the lax diagram

$$\begin{array}{ccc}
 F\mathbb{1} & \xrightarrow{!} & \mathbb{1} \\
 F\kappa_{\mathbb{1}} \downarrow & \geq & \downarrow \kappa_{\mathbb{1}} \\
 F\mathcal{V} & \xrightarrow{ev} & \mathcal{V}
 \end{array}$$

*Proof.* (1) We start with the observation that the predicate  $p \otimes q$  is computed as the composite

$$X \xrightarrow{\delta_X} X \times X \xrightarrow{p \times q} \mathcal{V} \times \mathcal{V} \xrightarrow{\otimes} \mathcal{V}$$

Upon recalling that  $\widehat{F}(p) = ev \circ F(p)$ , we notice that the leftmost path from  $FX$  to  $\mathcal{V}$  in the next diagram evaluates to  $\widehat{F}(p \otimes q)$ . Similarly, the rightmost path from  $FX$  to  $\mathcal{V}$  evaluates to  $\widehat{F}(p) \otimes \widehat{F}(q)$ . Now, the desired inequality  $\widehat{F}(p \otimes q) \geq \widehat{F}(p) \otimes \widehat{F}(q)$  follows using the fact that the upper triangle commutes, the naturality of  $\lambda$  and the lax diagram from the hypothesis.

$$\begin{array}{ccccc}
 & & FX & & \\
 & F\delta_X \swarrow & & \searrow \delta_{FX} & \\
 F(X \times X) & \xrightarrow{\lambda_X} & & FX \times FX & \\
 F(p \times q) \downarrow & & & \downarrow Fp \times Fq & \\
 F(\mathcal{V} \times \mathcal{V}) & \xrightarrow{\lambda_{\mathcal{V}}} & & F\mathcal{V} \times F\mathcal{V} & \\
 F(\otimes) \downarrow & & \geq & \downarrow ev \times ev & \\
 F\mathcal{V} & & & \mathcal{V} \times \mathcal{V} & \\
 & ev \searrow & & \swarrow \otimes & \\
 & & \mathcal{V} & & 
 \end{array}$$

- (2) We consider the following diagram

$$\begin{array}{ccccc}
 & & F\mathbb{1} & \xrightarrow{!} & \mathbb{1} \\
 & F! \curvearrowright & \downarrow F\kappa_{\mathbb{1}} & \geq & \downarrow \kappa_{\mathbb{1}} \\
 FX & \xrightarrow{F\kappa_X} & F\mathcal{V} & \xrightarrow{ev} & \mathcal{V}
 \end{array}$$

Since  $\kappa_X = \kappa_{\mathbb{1}} \circ !$  the left triangle commutes. Hence, we obtain

$$ev \circ F\kappa_X \geq \kappa_{\mathbb{1}} \circ ! \circ F!$$

or, equivalently,

$$\widehat{F}(\kappa_X) \geq \kappa_X.$$

□

We show how one of the conditions for well-behavedness that we required for the Wasserstein lifting in Baldan et al. (2018, Definition 5.14) for the quantale  $\mathcal{V} = [0, \infty]$  is related to the conditions in Proposition 27. Our original condition was  $d_e \circ (ev_F \times ev_F) \circ \langle F\pi_1, F\pi_2 \rangle \leq \widehat{F}(d_e)$  where  $d_e(-1, -2) = [-1, -2] \wedge [-2, -1]$  (which evaluates to  $d_e(r, s) = |r - s|$  on the reals). This clearly implies the non-symmetric variant stated in the lemma below.

**Lemma 54.**  $\widehat{F}(p \otimes q) \geq \widehat{F}(p) \otimes \widehat{F}(q)$  holds for all  $p, q: X \rightarrow \mathcal{V}$  if and only if  $[\pi_1, \pi_2](ev_F \times ev_F)\langle F\pi_1, F\pi_2 \rangle \geq \widehat{F}[\pi_1, \pi_2]$ .

*Proof.*

$$[\pi_1, \pi_2](ev_F \times ev_F)\langle F\pi_1, F\pi_2 \rangle \geq \widehat{F}[\pi_1, \pi_2] \tag{A4}$$

$$\iff [\pi_1, \pi_2](\widehat{F}\pi_1, \widehat{F}\pi_2) \geq \widehat{F}[\pi_1, \pi_2] \tag{A5}$$

$$\iff [\widehat{F}\pi_1, \widehat{F}\pi_2] \geq \widehat{F}[\pi_1, \pi_2] \tag{A6}$$

$$\iff \widehat{F}\pi_2 \geq \widehat{F}[\pi_1, \pi_2] \otimes \widehat{F}\pi_1 \tag{A7}$$

using

- For the equivalences (A4)  $\iff$  (A5)  $\iff$  (A6) just simple rewriting along with  $\widehat{F} = ev_F \circ F$ .
- For the equivalence (A6)  $\iff$  (A7) the tensor property  $x \otimes y \leq z \iff x \leq [y, z]$ .

Using this, we now aim to show that (A7)  $\iff$  (26).

- Showing (26)  $\implies$  (A7) is straightforward: from  $[\pi_1, \pi_2] \leq [\pi_1, \pi_2]$ , we can infer  $[\pi_1, \pi_2] \otimes \pi_1 \leq \pi_2$ . Using this, the monotonicity of  $\widehat{F}$  and (26) (by taking  $u = [\pi_1, \pi_2]: X \rightarrow \mathcal{V}$  and  $v = \pi_1: X \rightarrow \mathcal{V}$ ) we obtain inequality (A7) as follows:

$$\widehat{F}\pi_2 \geq \widehat{F}([\pi_1, \pi_2] \otimes \pi_1) \geq \widehat{F}[\pi_1, \pi_2] \otimes \widehat{F}\pi_1$$

- The implication (A7)  $\implies$  (26) can be shown by rewriting  $u \otimes v = \pi_2 \circ \langle v, u \otimes v \rangle$  and then using (A7) as follows

$$\begin{aligned} & \widehat{F}(u \otimes v) \\ &= \widehat{F}(\pi_2 \circ \langle v, u \otimes v \rangle) = \widehat{F}\pi_2 \circ F\langle v, u \otimes v \rangle \\ &\geq (\widehat{F}[\pi_1, \pi_2] \otimes \widehat{F}\pi_1) \circ F\langle v, u \otimes v \rangle \\ &= \widehat{F}([\pi_1, \pi_2] \circ \langle v, u \otimes v \rangle) \otimes \widehat{F}(\pi_1 \circ \langle v, u \otimes v \rangle) \\ &= \widehat{F}[v, u \otimes v] \otimes \widehat{F}v \geq \widehat{F}u \otimes \widehat{F}v \end{aligned}$$

where the last inequality follows again from monotonicity of  $\widehat{F}$  and the definitions of  $\otimes$  and  $[-, -]$ . In particular,  $u \otimes v \leq u \otimes v$ , and hence,  $u \leq [v, u \otimes v]$ . □

**Appendix B. Proofs and additional material for Section 6**

**B.1 Lifting of the distributive law**

We include an alternative proof of Theorem 42 which brings the pieces of proof in one large diagram.

*Alternative proof of Theorem 42.* The existence of the lifting  $\widehat{\zeta}$  is equivalent to

$$\widehat{T} \circ \widehat{F} \leq (\zeta_X)^* \circ \widehat{F} \circ \widehat{T},$$

while the existence of the lifting  $\overline{\zeta}$  is equivalent to

$$\overline{T} \circ \overline{F} \leq (\zeta_X)^* \circ \overline{F} \circ \overline{T}.$$

The latter inequality, which we have to prove, is in turn equivalent to the inequality obtained by using the isomorphism  $\iota$ .

$$\begin{aligned} \iota_{TFX} \circ \overline{T} \circ \overline{F} \circ \iota_X^{-1} &\leq \iota_{TFX} \circ (\zeta_X)^* \circ \overline{F} \circ \overline{T} \circ \iota_X^{-1} \\ &= (\Delta \zeta_X)^* \circ \iota_{FTX} \circ \overline{F} \circ \overline{T} \circ \iota_X^{-1} \end{aligned}$$

The left-hand side of the above inequality rewrites using the definitions of the Wasserstein liftings as the composite of the outermost right-then-down path  $\mathcal{V}\text{-Pred}_{\Delta X}$  to  $\mathcal{V}\text{-Pred}_{\Delta TFX}$  in the next diagram. The right-hand side similarly evaluates to the outermost down-then-right path in the diagram. So it suffices to establish the inequality between these two paths. We do this by decomposing the diagram into smaller pieces (see Fig. B1) and explaining each inequality in turn.

The two inequalities in the top pentagon and top-right square follow from the hypothesis. The two triangles at the bottom are equalities that follow from the fact that

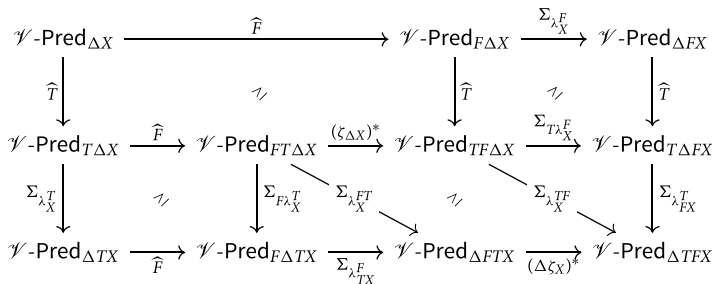
$$\lambda_X^{FT} = \lambda_{TX}^F \circ F \lambda_X^T \quad \text{and} \quad \lambda_X^{TF} = \lambda_{FX}^T \circ T \lambda_X^F$$

The inequality in the left-down square holds since  $\widehat{F}$  is a lifting and is obtained via adjoint transposes from (3).

Using the naturality of  $\zeta$ , one can show that the next square commutes

$$\begin{array}{ccc} TF\Delta X & \xrightarrow{\lambda_X^{TF}} & \Delta TFX \\ \zeta_{\Delta X} \downarrow & & \downarrow \Delta \zeta_X \\ FT\Delta X & \xrightarrow{\lambda_X^{FT}} & \Delta FTX. \end{array}$$

and hence, the inequality in the bottom rhombus can be derived as an instance of a generic result for bifibrations, see (5). □



**Figure B1.** Existence of the lifting  $\widehat{\zeta}$ .

**B.2 Details on constructively completely distributive quantales**

In this appendix, we provide a result (Proposition 56 below) for proving  $\widehat{T} \circ \Sigma_f \leq \Sigma_{Tf} \circ \widehat{T}$  that is more general than Proposition 45. This is useful, for instance to prove such property for liftings different than the canonical one.

For this more general result, we need some additional properties, in particular the lifting  $\widehat{T}$  must preserve a special type of supremum of predicates (even stronger than uniform convergence).

**Definition 55.** Let  $(p_i: X \rightarrow \mathcal{V})_{i \in I}$  be a family of predicates. We call its sup constructively convergent if for every predicate  $q: X \rightarrow \mathcal{V}$  with  $q \lll \bigvee_{i \in I} p_i$  (pointwise), there exists  $i \in I$  with  $q \leq p_i$ .

**Proposition 56.** Assume  $\mathcal{V}$  is a constructively completely distributive quantale and assume  $\widehat{T}$  is a lifting of a Set-functor  $T$ . Then, we have that  $\widehat{T} \circ \Sigma_f \leq \Sigma_{Tf} \circ \widehat{T}$  whenever either of the conditions below is met

- $f$  is surjective and  $\widehat{T}$  preserves constructively convergent sups.
- $f$  is injective,  $T$  preserves weak pullbacks,  $\widehat{T}$  is a fibred lifting corresponding to an evaluation map  $ev$  such that for every  $t \in T\mathcal{V}$ ,  $ev(t) \neq \perp$  implies  $t \in T(\mathcal{V} \setminus \{\perp\})$ . (In other words:  $ev^{-1}(\mathcal{V} \setminus \{\perp\}) \subseteq T(\mathcal{V} \setminus \{\perp\})$ .)
- $f$  is an arbitrary function and all the above properties are satisfied.

*Proof.* Let  $f = m \circ e$  be the epi-mono factorization of  $f$ , that is,  $e: X \rightarrow Z$  is surjective and  $m: Z \rightarrow Y$  is injective. We will show the inequality separately for  $m, e$ , from which we can straightforwardly derive the inequality for  $f$ .

- $\widehat{T} \circ \Sigma_e \leq \Sigma_{Te} \circ \widehat{T}$ :  
Let  $p: X \rightarrow \mathcal{V}, z \in Z$ . Observe that

$$\Sigma_e(p)(z) = \bigvee \{p(x) \mid e(x) = z\} = \bigvee \{(p \circ g)(z) \mid g \in G\}$$

where  $G = \{g: Z \rightarrow X \mid e \circ g = id_Z\}$  is the set of all choice functions. Note that the last equality in the displayed equation above requires surjectivity of  $e$ , because otherwise no choice functions exist.

So  $\Sigma_e(p) = \bigvee_{g \in G} p \circ g = \bigvee_{g \in G} g^*(p)$  and we show that this sup is constructively convergent. Let  $q: Z \rightarrow \mathcal{V}$  with  $q \lll \Sigma_e(p)$ . Now for a  $z \in Z$ , we observe that  $q(z) \lll \bigvee_{e(x)=z} p(x)$ , and hence, since we are working in a cccd-lattice, there exists an  $x_z \in X$  with  $e(x_z) = z$  and  $p(x_z) \geq q(z)$ . On  $z$ , we define the choice function  $g$  as  $g(z) = x_z$ . We have  $e \circ g = id_Z$ , and furthermore for all,  $z \in Z$  we have  $q(z) \leq p(x_z) = (p \circ g)(z)$ , hence  $q \leq p \circ g$  as desired.

According to our assumption  $\widehat{T}$  preserves such suprema and we get:

$$\widehat{T}(\Sigma_e(p)) = \widehat{T}(\bigvee_{g \in G} g^*(p)) = \bigvee_{g \in G} \widehat{T}(g^*(p)) \leq \bigvee_{g \in G} (Tg)^*(\widehat{T}p)$$

We will now show  $(Tg)^* \leq \Sigma_{Te}$  as an intermediate result: Let  $p: TX \rightarrow \mathcal{V}$  and  $t \in TZ$ . Then,

$$(Tg)^*(p)(t) = (p \circ Tg)(t) = p(Tg(t)) \leq \bigvee_{Te(s)=t} p(s) = \Sigma_{Te}(p)(t)$$

since  $s = Tg(t)$  satisfies  $Te(s) = Te(Tg(t)) = Tid_Z(t) = t$ . This implies

$$\bigvee_{g \in G} (Tg)^*(\widehat{T}p) \leq \bigvee_{g \in G} \Sigma_{Te}(\widehat{T}p) = (\Sigma_{Te} \circ \widehat{T})(p)$$

By combining everything, we obtain the desired result.

- $\widehat{T} \circ \Sigma_m \leq \Sigma_{Tm} \circ \widehat{T}$  :

Let  $p: Z \rightarrow \mathcal{Y}$ ,  $t \in FY$ , we have to show that  $\widehat{T}(\Sigma_m(p))(t) \leq \Sigma_{Tm}(\widehat{T}p)(t)$ .

We consider the following two cases:

- $t$  is in the image of  $Tm$ : in this case, there exists  $s \in TX$  with  $Tm(s) = t$ . Since  $m$  is injective, we have that for any  $y \in Y$  in the image of  $m$   $\Sigma_m(p)(y) = p(z)$ , where  $z \in Z$  is the unique preimage of  $y$ . Hence,  $\Sigma_m(p) \circ m = p$ .

Using the fact that we have a fibred lifting (Proposition 14), this means that

$$\begin{aligned} \widehat{T}(\Sigma_m(p))(t) &= \widehat{T}(\Sigma_m(p))(Tm(s)) = \widehat{T}(\Sigma_m(p) \circ m)(s) = \widehat{T}p(s) \\ &\leq \bigvee_{Tm(s)=t} \widehat{T}p(s) = \Sigma_{Tm}(\widehat{T}p)(t) \end{aligned}$$

- $t$  is not in the image of  $Tm$ : we show that in this case  $\widehat{T}(\Sigma_m(p))(t) = \perp$ . (The right-hand side of the inequality is also  $\perp$ , due to the empty supremum.) Note that  $\Sigma_m(p)(y) = \perp$  for all  $y \in Y$  which are not in the image of  $m$ .

Now assume that  $\widehat{T}(\Sigma_m(p))(t) \neq \perp$ . Take the pullback on the left below and observe that  $Y' = \{y \in Y \mid \Sigma_m(p)(y) \neq \perp\} \subseteq m[X]$ .

$$\begin{array}{ccc} \mathcal{Y} \setminus \{\perp\} & \longrightarrow & \mathcal{Y} \\ \uparrow & \Sigma_m(p) \uparrow & \uparrow \\ Y' & \xrightarrow{m} & Y \end{array} \qquad \begin{array}{ccc} T(\mathcal{Y} \setminus \{\perp\}) & \longrightarrow & T\mathcal{Y} \xrightarrow{ev} \mathcal{Y} \\ \uparrow & T(\Sigma_m(p)) \uparrow & \nearrow \widehat{T}(\Sigma_m(p)) \\ T(Y') & \xrightarrow{Tm} & TY \end{array}$$

Since  $T$  is weak pullback-preserving, the square on the right above is a weak pullback. By assumption, we have  $ev(T(\Sigma_m(p))(t)) = \widehat{T}(\Sigma_m(p))(t) \neq \perp$ . This implies that  $T(\Sigma_m(p))(t) \in T(\mathcal{Y} \setminus \{\perp\})$ .

Since the square to the right above is a weak pullback, this means that  $t \in T(Y')$ , hence  $t \in T(m[X])$ , which is a contradiction, since  $t$  is not in the image of  $Tm$ .  $\square$

**Lemma 57.** Assume that  $T$  preserves weak pullbacks and  $\mathcal{Y}$  is a constructively completely distributive quantale. Then, the canonical predicate lifting  $\widehat{T}_{can}$  preserve2s constructively convergent sups, that is,  $\widehat{T}_{can}(\bigvee_{i \in I} p_i) = \bigvee_{i \in I} \widehat{T}_{can}(p_i)$ .

*Proof.* First, we obviously have  $\widehat{T}_{can}(\bigvee_{i \in I} p_i) \geq \bigvee_{i \in I} \widehat{T}_{can}(p_i)$  due to monotonicity.

We now show  $\widehat{T}_{can}(\bigvee_{i \in I} p_i) \leq \bigvee_{i \in I} \widehat{T}_{can}(p_i)$ : first denote  $\bigvee_{i \in I} p_i$  by  $p$ . Let  $t \in TX$ , hence the inequality spells out to

$$\bigvee \{r \mid Tp(t) \in T(\uparrow r)\} \leq \bigvee_{i \in I} \bigvee \{s \mid Tp_i(t) \in T(\uparrow s)\}$$

Now let  $r$  be such that  $Tp(t) \in T(\uparrow r)$  and take the pullback on the left below. Note that  $X_r = \{x \in X \mid p(x) \geq r\}$ .

$$\begin{array}{ccc} \uparrow r & \longrightarrow & \mathcal{Y} \\ \uparrow & & \uparrow p \\ X_r & \longrightarrow & X \end{array} \qquad \begin{array}{ccc} T(\uparrow r) & \longrightarrow & T\mathcal{Y} \\ \uparrow & & \uparrow Tp \\ T(X_r) & \longrightarrow & TX \end{array}$$

Due to weak pullback preservation, the square above on the right is a weak pullback. This means that  $t \in TX$ , which satisfies  $t \in T(\uparrow r)$ , is also contained in  $T(X_r)$ .

Now let  $u \lll r$ . We define a predicate  $p' : X \rightarrow \mathcal{V}$  with

$$p'(x) = \begin{cases} u & \text{if } x \in X_r \\ \perp & \text{otherwise} \end{cases}$$

Note that  $p'$  satisfies  $p' \lll p$ . This is also true whenever  $p'(x) = \perp$ , since  $\perp$  is totally below everything. Then, due to constructively convergence of the sup there exists an index  $i \in I$  with  $p' \leq p_i$ .

Now obtain the set  $X_u^i$  with the following pullback on the left, where  $X_u^i = \{x \in X \mid p_i(x) \geq u\}$ .

$$\begin{array}{ccc} \uparrow u & \longrightarrow & \mathcal{V} \\ \uparrow & & \uparrow p_i \\ X_u^i & \longrightarrow & X \end{array} \qquad \begin{array}{ccc} T(\uparrow u) & \longrightarrow & T\mathcal{V} \\ \uparrow & & \uparrow Tp_i \\ T(X_u^i) & \longrightarrow & TX \end{array}$$

We can observe that  $X_r \subseteq X_u^i$ : let  $x \in X_r$ , then  $u = p'(x) \leq p_i(x)$ , hence  $x \in X_u^i$ .

This means that  $t \in T(X_u^i)$  and since the square on the right above commutes, this gives us  $Tp_i(t) \in T(\uparrow u)$ . From this, we infer

$$\bigvee_{i \in I} \bigvee \{s \mid Tp_i(t) \in T(\uparrow s)\} \geq u$$

Since this holds for all  $u \lll r$ , we have

$$r = \bigvee_{u \lll r} u \leq \bigvee_{i \in I} \bigvee \{s \mid Tp_i(t) \in T(\uparrow s)\}$$

which entails the required inequality. □

**Lemma 58.** *For a constructively completely distributive quantale, it holds for the canonical predicate lifting  $\widehat{T}_{\text{can}}$  that  $ev_{\text{can}}(t) \neq \perp$  implies  $t \in T(\mathcal{V} \setminus \{\perp\})$  for  $t \in T\mathcal{V}$ .*

*Proof.* Assume that

$$ev_{\text{can}}(t) = \bigvee \{r \mid t \in T(\uparrow r)\} \neq \perp$$

Hence, there is at least one  $r \neq \perp$  with  $t \in T(\uparrow r) \subseteq T(\mathcal{V} \setminus \{\perp\})$ ; otherwise, the supremum would equal  $\perp$ . □

Whenever  $T = \mathcal{D}$  (the probability distribution functor), we have that  $\mathcal{D}$  preserves constructively convergent sups (since it preserves uniform convergence). We now consider the evaluation map as in Example 17, namely the expectation  $ev(t) = \sum_{r \in [0,1]} r \cdot t(r)$  with  $t \in \mathcal{D}[0, 1]$ . Note that in the quantale  $[0, 1]$  we have  $\perp = 1$ .

However, the property  $ev(t) \neq 1 \Rightarrow t \in \mathcal{D}[0, 1)$  for all  $t \in \mathcal{D}[0, 1]$  does not hold in this case. If  $t(0) = 1, t(r) = 0$  for all  $r \neq 0$ , we have  $ev(t) = 0 \neq 1$ , but  $t \notin \mathcal{D}[0, 1)$ .

Nevertheless, the corresponding  $\mathcal{V}$ -Pred-lifting of  $\mathcal{D}$  still satisfies the property  $\widehat{\mathcal{D}} \circ \Sigma_\lambda \leq \Sigma_{\mathcal{D}\lambda} \circ \widehat{\mathcal{D}}$ , required in Theorem 42. Here we can rely on the fact that for  $\mathcal{D}$  the components  $\lambda_X : \mathcal{D}(X \times X) \rightarrow \mathcal{D}X \times \mathcal{D}X$  are *surjective*, so we can apply the second item of Proposition 56.