Correspondences Between Codensity and Coupling-Based Liftings, a Practical Approach

Samuel Humeau ✉ 🔗
ENS de Lyon, CNRS, LIP, UMR 5668, 69342, Lyon cedex 07, France

Daniela Petrisan ✉ 🔗
CNRS, IRIF, Université Paris Diderot, Paris, France

Jurriaan Rot ✉ 🔗
Institute for Computing and Information Sciences, Radboud University, Nijmegen, The Netherlands

—— **Abstract** ——

The Kantorovich distance is a widely used metric between probability distributions. The Kantorovich-Rubinstein duality states that it can be defined in two equivalent ways: as a supremum, based on non-expansive functions into $[0, 1]$, and as an infimum, based on probabilistic couplings.

Orthogonally, there are categorical generalisations of both presentations proposed in the literature, in the form of *codensity liftings* and what we refer to as *coupling-based liftings*. Both lift endofunctors on the category **Set** of sets and functions to that of pseudometric spaces, and both are parameterised by modalities from coalgebraic modal logic.

A generalisation of the Kantorovich-Rubinstein duality has been more nebulous – it is known not to work in some cases. In this paper we propose a compositional approach for obtaining such generalised dualities for a class of functors, which is closed under coproducts and products. Our approach is based on an explicit construction of modalities and also applies to and extends known cases such as that of the powerset functor.

1 Introduction

The *Kantorovich* (or *Wasserstein*, or *Monge-Kantorovich*) distance [17] is a standard and widely used metric between probability distributions, studied amongst others in transportation theory [27]. In concurrency theory, the Kantorovich distance forms the basis of so-called *behavioural metrics*, which are quantitative generalisations of bisimilarity. They allow a more fine-grained and robust comparison of system behaviours than classical Boolean-valued behavioural equivalences [8, 9, 26].

In its discrete version the Kantorovich distance takes as argument a (pseudo-)metric on a set X, and lifts it to a (pseudo-)metric on the set of (finitely supported) distributions $\mathcal{D}(X)$. The celebrated *Kantorovich-Rubinstein duality* [18] states that this distance can be computed in two ways, yielding the same result: as an infimum indexed by probabilistic couplings, and as a supremum indexed by non-expansive functions into the $[0, 1]$ interval with the Euclidean

distance. This fundamental result is useful for analysis and computation of these distances (e.g., [1, 7, 15, 25, 27]). A detailed proof in a broader context than just finitely supported distributions can be found in [27].

Orthogonally to this duality, in the last years there have been several proposals to generalise the Kantorovich distance from distributions to general endofunctors on the category **Set** of sets and functions. The problem then becomes to lift such functors to the category of (pseudo-)metric spaces. This is particularly useful in the context of a coalgebraic presentation of systems, where the type of the system at hand is parametric in the given **Set** endofunctor. In particular, this allows a uniform presentation of various types of probabilistic systems, but also, for instance, metrics on deterministic automata [6].

There are categorical generalisations of both presentations of the Kantorovich distance: the *coupling-based* approach and the one based on non-expansive maps [2, 5, 6, 11]. The latter has recently been established as an instance of the so-called *codensity liftings*[1] [24]. Both approaches are parametric in (sets of) *modalities* or *evaluation maps*, that allow a degree of freedom in the choice of liftings.

We aim to relate the two approaches, by studying generalisations of the Kantorovich-Rubinstein duality to a wide class of functors beyond $\mathcal{D}$. This problem was first proposed and studied in [2], where it is shown to hold in some concrete cases but also to fail in other basic instances, even for very elementary functors such as the diagonal functor Δ mapping a set X to $X \times X$. In the latter article the authors restrict the study to liftings parametric in exactly one modality they call an *evaluation map*, which is assumed to be the same for both codensity and coupling-based liftings. We depart from this by allowing modalities to differ on both sides. This approach can already be found in [10]. There, it is shown that every coupling-based lifting can be presented as a codensity lifting; but the proof is non-constructive and yields a large collection of modalities.

In the current paper, we approach the problem of generalised Kantorovich-Rubinstein dualities from a concrete perspective, with an emphasis on compositionality aspects. Given a modality as parameter for a coupling-based lifting, we aim to explicitly translate it to modalities for a codensity lifting, in such a way that the two correspond. More explicitly, we show that the class of such correspondences between coupling-based and codensity liftings is closed under coproducts and products (and conversely, that every correspondence for a coproduct of functors can be recovered from correspondences on its constituents). We also investigate correspondences for the identity functor, where there is flexibility in the choice of modalities; and for the powerset functor, extending earlier correspondence results [2, 12].

These correspondence results then allow us to define a concrete *grammar* of functors for which we obtain a correspondence between coupling-based and codensity liftings. In fact, we obtain several grammars based on different assumptions on the underlying poset of truth values (assumed to be a quantale). Base cases include the constant functors, distribution functor, powerset functor and the identity functor; recursive constructions the product and coproduct. These results allow us to obtain or recover induced codensity and coupling-based presentations for a range of examples of behavioural metrics, including metrics on streams, labelled Markov chains, deterministic automata, and non-deterministic automata.

---

[1] *Kantorovich metric* is used interchangeably in the literature to refer to both presentations. We therefore avoid this terminology by consistently referring to the two presentations as *coupling-based* and *codensity* liftings respectively.

In the last part of the paper, we investigate the limitations of our approach through the concrete example of *conditional transition systems* [4, 3]. In contrast to the earlier examples, here, our grammar does give us a metric (and a correspondence result), but it is not the one considered earlier in the literature. We prove that, in fact, the metric from the literature can be expressed with a codensity lifting but not with a coupling-based lifting.

2 Codensity and coupling-based liftings: correspondences by example

In this section we motivate and describe the problem of correspondences between codensity and coupling-based liftings at the general level of functors, and our approach in this paper, by means of two examples: the Kantorovich-Rubinstein duality for distributions, and a similar correspondence for the shortest-distinguishing-word-distance on deterministic automata [6].

Distributions. We start by recalling the classical Kantorovich distance, in the discrete case. In this paper we focus on pseudometrics (defined like metrics except that different elements can have distance 0) as is common in the use of these types of distances in concurrency theory. For a pseudometric $d\colon X \times X \to [0,1]$, the Kantorovich distance is a pseudometric on the set $\mathcal{D}(X)$ of distributions on X, defined, for $\mu,\nu \in \mathcal{D}(X)$ by:

$$\mathcal{D}^{\downarrow}(d)(\mu,\nu) = \inf_{\sigma \in \Omega(\mu,\nu)} \sum_{x,y \in X} d(x,y) \cdot \sigma(x,y) \tag{1}$$

where $\Omega(\mu,\nu)$ is the set of couplings between μ and ν, i.e., probability distributions on $X \times X$ whose marginals are μ and ν respectively. It can equivalently be computed as:

$$\mathcal{D}^{\uparrow}(d)(\mu,\nu) = \sup_{f\colon X \to [0,1] \text{ n.e.}} \left| \sum_{x \in X} f(x) \cdot \mu(x) - \sum_{x \in X} f(x) \cdot \nu(x) \right| \tag{2}$$

where the supremum ranges over *non-expansive functions* f into $[0,1]$ equipped with the Euclidean distance, i.e., such that $|f(x) - f(y)| \leq d(x,y)$ for all $x,y \in X$. The equality $\mathcal{D}^{\uparrow}(d) = \mathcal{D}^{\downarrow}(d)$ is an instance (see [27, Particular case 5.16]) of a general result known as the Kantorovich-Rubinstein duality (see [27, Theorem 5.10]).

A different example. The Kantorovich distance can be seen as a *lifting* of the distribution functor on the category **Set** of sets and functions to the category **PMet** of pseudometric spaces and non-expansive functions between them. We proceed with a quite different example of a similar phenomenon: a lifting of a functor from **Set** to $\mathbf{PMet}_{\leq 1}$ the category of pseudometric spaces bounded by 1 presented in two ways, as an infimum over a variant of couplings and a supremum over non-expansive functions. This example describes a distance between *deterministic finite automata (DFA)*. Empty infima and suprema are defined w.r.t. the interval $[0,1]$ where the pseudometrics take their values: $\sup \emptyset = 0$ and $\inf \emptyset = 1$.[2]

We view DFA over an alphabet A as coalgebras for the functor $F\colon \mathbf{Set} \to \mathbf{Set}$, $F(X) = 2 \times X^A$ where $2 = \{0,1\}$. Coalgebras for this functor are of the form $\langle l,\delta \rangle\colon X \to 2 \times X^A$, where X is the set of states, the output function $l\colon X \to 2$ describes which states are accepting ($l(x) = 1$), and $\delta\colon X \to X^A$ is the transition function. As usual, a word $\omega \in A^*$ is *accepted* in a state x when, after reading ω starting on x we end up with an accepting state.

---

[2] The *knowledge* package is used throughout the paper. Most of the introduced vocabulary and notations are clickable and the associated links brings the reader to their definitions.

Given $c \in [0,1)$, the *shortest-distinguishing-word-distance* $d_{\mathrm{sdw}}(x,y)$ between states x,y is 0 if they recognise the same language and $c^{|\omega|}$ for ω a shortest word that belongs exactly to one of the two languages recognised by x and y. As shown in [6], this distance can be computed recursively as a fixpoint of the map Φ on pseudometrics $X \times X \to [0,1]$ given by

$$\Phi(d)(x,y) = \begin{cases} 1 \text{ if } l(x) \neq l(y) \\ c \cdot \max_{a \in A}\{d(\delta(x)(a), \delta(y)(a))\} \end{cases}$$

The shortest-distinguishing-word-distance can be obtained via a *lifting* $\overline{F} \colon \mathbf{PMet} \to \mathbf{PMet}$ of the **Set** functor F, mapping a pseudometric space (X,d) to a pseudometric space $(FX, \overline{F}d)$, that we will recall below. The operator Φ above factors through the pseudometric $\overline{F}d$. Explicitly, it decomposes as $\Phi = \overline{F}d \circ \langle l, \delta \rangle$.

In fact, just as in the Kantorovich-Rubinstein duality, the lifting $\overline{F}$ can be obtained in two different ways. Given a pseudometric $d \colon X \times X \to [0,1]$ and $(l_1, \delta_1), (l_2, \delta_2) \in 2 \times X^A$, it arises as a coupling-based lifting $\overline{F}^{\downarrow}d$ by:

$$\overline{F}^{\downarrow}(d)((l_1,\delta_1),(l_2,\delta_2)) = \inf_{(l,\delta) \in F(X \times X), \ F\pi_i(l,\delta) = (l_i, \delta_i)} \left(\sup_{a \in A} c \cdot d(\delta(a)) \right) \tag{3}$$

Here elements of $2 \times (X \times X)^A$ are viewed as a variant of couplings.

Secondly we obtain $\overline{F}$ as a so-called codensity lifting by:

$$\overline{F}^{\uparrow}(d)((l_1,\delta_1),(l_2,\delta_2)) = \sup_{f \colon X \to [0,1] \text{ n.e.}} \{|c \cdot f(\delta_1(a)) - c \cdot f(\delta_2(a))| \mid a \in A\} \cup \{|l_1 - l_2|\} \tag{4}$$

where, as above, the supremum again ranges over non-expansive functions f. Just as for the Kantorovich distance, we again obtain an equality of functors on **PMet**: $\overline{F}^{\uparrow} = \overline{F}^{\downarrow}$.

Modalities as parameters. The abstract coupling-based and codensity liftings take as input a single *modality* (coupling-based), or a *set of modalities* (codensity), known from the semantics of coalgebraic modal logic. For the Kantorovich lifting, this modality is the expected value function $\mathbb{E} \colon \mathcal{D}([0,1]) \to [0,1]$, which appears implicitly in both presentations.

In the case of deterministic automata, while a single modality suffices for the coupling-based lifting, for the codensity lifting we actually use a *set* of modalities (one for each letter, plus a modality for acceptance of states). This observation is important for the investigation in this paper. Instead of aiming for a one-to-one matching of modalities, which we refer to as a *duality*, we allow a coupling-based lifting specified by a single modality to be matched by a codensity lifting specified by a set of modalities; we refer to this as a *correspondence*. This allows us to cover examples such as DFA and way more, and circumvent the problem already observed in [2]: even for the product functor with the modality $\max \colon [0,1] \times [0,1] \to [0,1]$ there is no duality. However, there is a correspondence, that is, multiple modalities are needed for the codensity lifting to match the coupling-based one.

The idea of allowing multiple modalities for codensity liftings is not new and can be found already at the origins of codensity liftings [19, 24, 21]. Here we show how it can be used to generalise Kantorovich-Rubinstein dualities in a very concrete manner; the aim is to define classes of functors and modalities for which we can explicitly describe correspondences between associated coupling-based and codensity liftings.

Outline. We work at the general level of pseudometrics valued in a quantale; we recall the definitions in Section 3. We then recall the abstract notions of coupling-based and codensity liftings of functors along modalities, and formulate the problem of correspondences (Section 4).

In Section 5 we prove our main results on correspondences: we show that the class of functors for which we have correspondences is closed under products and coproducts, and includes constant and identity functors, yielding a family of correspondences for simple polynomial functors including DFA. In Section 6 we revisit the duality results for the finite powerset and finite probability distribution functors. As a consequence we are able to study in Section 7 a class of functors for which we have correspondences, generated by a grammar. The case of conditional transition systems, for which we prove that there are no correspondences possible, is treated in Section 8.

3 Preliminaries: quantales and pseudometrics

We use quantales to model a general notion of truth object, including both Booleans and real number intervals. In this section we recall the necessary preliminaries on quantales, and the associated general notion of pseudometrics valued in a quantale; instances include the standard notion of pseudometrics on real numbers as well as equivalence relations.

▶ **Definition 1.** *A* quantale $\mathcal{V}$ *is a complete lattice with an associative "and" operation* $\otimes\colon \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ *which is distributive over arbitrary joins. We only consider quantales that are* commutative*, i.e., the operation $\otimes$ is commutative,* unital*, meaning $\otimes$ admits a unit element, and* affine*, which means that the top element $\top$ is the unit of $\otimes$: $x \otimes \top = x = \top \otimes x$ for all $x \in \mathcal{V}$.*

Given a quantale $\mathcal{V}$, there is an operation $[-,-]\colon \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ that is characterised by $x \otimes y \leq z \Leftrightarrow x \leq [y, z]$. It is simply defined by $[y, z] = \bigvee\{x \in \mathcal{V} \mid x \otimes y \leq z\}$.

▶ **Example 2.**
- Any complete Boolean algebra is a quantale with $\otimes = \wedge$; in particular, the usual Boolean algebra $2 = \{\bot, \top\}$ with $\bot \leq \top$. In this case, $[x, y] = \top$ iff $x \leq y$.
- Any interval $[0, M]$ with $M \in (0, \infty]$, given with *reversed order* and $x \otimes y = \min(x + y, M)$ the *truncated sum* is a quantale. Here the top element is given by $\top = 0$, and the bottom element by $\bot = M$. For this quantale, we have $[x, y] = \max\{y - x, 0\}$. We write $\overline{\mathbb{R}}$ for the case that $M = \infty$, i.e., the non-negative real numbers extended with a top element.

We use quantales as an abstract notion of truth object; accordingly, we define a $\mathcal{V}$-*predicate* on a set X to be a map $p\colon X \to \mathcal{V}$. Of particular interest are $\mathcal{V}$-pseudometrics. These are predicates on $X \times X$ that are reflexive, symmetric and transitive in a suitable sense that can be expressed at the general level of quantales.

▶ **Definition 3.** *A $\mathcal{V}$-pseudometric on a set X is a map $d\colon X \times X \to \mathcal{V}$ which is:*
- reflexive*: $d(x, x) = \top$ for all $x \in X$,*
- symmetric*: $d(x, y) = d(y, x)$ for all $x, y \in X$,*
- transitive*: $\bigvee_{z \in X} d(x, z) \otimes d(z, y) \leq d(x, y)$ for all $x, y \in X$.*

▶ **Example 4.**
- 2-pseudometrics are equivalence relations.
- $\overline{\mathbb{R}}$-pseudometrics are the "usual" pseudometrics, that is, maps $d\colon X \times X \to \overline{\mathbb{R}}$ such that $d(x, x) = 0$, $d(x, y) = d(y, x)$ and $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in \overline{\mathbb{R}}$. To see why we obtain the triangle inequality from transitivity, recall that the order on the quantale is reversed w.r.t. the usual order on real numbers, and that $\otimes = +$.

The order on a quantale $\mathcal{V}$ is extended pointwise to functions:

$$\forall f, g \colon X \to \mathcal{V}, \; f \leq g \Leftrightarrow (\forall x \in X, \; f(x) \leq g(x))$$

Given two $\mathcal{V}$-pseudometrics d_X and d_Y on sets X and Y respectively, a map $f \colon X \to Y$ is a *($\mathcal{V}$-pseudometric) morphism* from d_X to d_Y if $d_X \leq d_Y \circ (f \times f)$. We write $\mathcal{V}$-**PMet** for the category whose objects are $\mathcal{V}$-pseudometrics and arrows morphisms between them.

▶ Remark 5. Because of the reversal of the order on $\overline{\mathbb{R}}$, morphisms of $\overline{\mathbb{R}}$-pseudometrics are non-expansive maps. To avoid confusion with the quantale definition, we refrain from using the word *non-expansive* altogether, and replace it by *($\mathcal{V}$-pseudometric) morphism* instead.

The following canonical $\mathcal{V}$-pseudometric is essential for the notion of codensity lifting.

▶ **Definition 6.** *The* Euclidean pseudometric $d_e \colon \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ *is defined as follows:*

$$d_e(x, y) = [x, y] \wedge [y, x]$$

▶ **Example 7.**
- For the quantale 2, the Euclidean pseudometric $d_e \colon 2 \times 2 \to 2$ sends equal elements to $\top$ and different ones to $\bot$.
- For the quantale $\overline{\mathbb{R}}$, the Euclidean pseudometric $d_e \colon \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ instantiates to the usual Euclidean distance, i.e., $d_e(x, y) = |y - x|$.

4   Liftings, dualities, and correspondences

In this section we recall the definitions of coupling-based and codensity liftings from the literature (Section 4.2) and define the problem of their correspondence. This is followed by a few technical tools that we use in the proofs of correspondence (Section 4.3). We start with the notion of (well-behaved) modality (Section 4.1), used in these liftings.

4.1   Well-behaved modalities

Given a **Set** endofunctor F, a *modality* for F is a function $\tau \colon F\mathcal{V} \to \mathcal{V}$. Modalities are standard in the semantics of coalgebraic modal logic [23]. Note that here we assume $\mathcal{V}$ to be a quantale, to have a suitable notion of pseudometrics.

In order for coupling-based liftings to be well-defined some particular conditions on the functor and the associated modalities are needed. The underlying functor F is assumed to preserve weak pullbacks. In this context, we say τ is *well-behaved* [2] when:
- it is *monotone*, meaning that for all $\mathcal{V}$-predicates $p \leq q$ we have $\tau \circ F(p) \leq \tau \circ F(q)$,
- for all $\mathcal{V}$-predicates p and q, $\tau \circ F(p \otimes q) \geq (\tau \circ Fp) \otimes (\tau \circ Fq)$,
- with $i \colon \{\top\} \hookrightarrow \mathcal{V}$ the inclusion map, $Fi(F\{\top\}) = \tau^{-1}(\top)$.

▶ **Example 8.**
- On the identity functor, with the quantale $\overline{\mathbb{R}}$, a modality is just a map $\tau \colon \mathcal{V} \to \mathcal{V}$. It is well-behaved if and only if it is monotone, subadditive (i.e., $\tau(r + s) \leq \tau(r) + \tau(s)$), and satisfies $\tau(0) = 0$.
- For the *finite powerset functor* $\mathcal{P}$ mapping a set X to the set of its finite subsets $\mathcal{P}(X)$, with the same quantale $\overline{\mathbb{R}}$, the maximum function is well-behaved.
- For the *finite distribution functor* $\mathcal{D}$ mapping a set X to the set of its finitely supported probability distributions $\mathcal{D}(X)$, with the quantale $[0, 1]$, the map $\mathbb{E} \colon \mathcal{D}([0, 1]) \to [0, 1]$ giving the expected value of a probability distribution is well-behaved.

In the rest of this paper we consider constant, identity, finite powerset, finite probability distribution functors, and combine them using products and coproducts. All these functors preserve weak pullbacks (see [14, Propositions 4.2.6 and 4.2.10] for example), ensuring we can consider well-behaved modalities for them.

Well-behaved modalities can be constructed from existing ones. The case of composition is a generalisation to quantales of a particular case of [2, Theorem 7.2].

▶ **Proposition 9.** *Given three arbitrary well-behaved modalities* $\tau, \tau' \colon F\mathcal{V} \to \mathcal{V}$ *and* $\tau_{\mathrm{Id}} \colon \mathcal{V} \to \mathcal{V}$, *the following modalities are again well-behaved:* $\tau_{\mathrm{Id}} \circ \tau$, $\tau \otimes \tau'$ *and* $\tau \wedge \tau'$.

4.2 Liftings and correspondences

Given a functor $P \colon \mathbf{C} \to \mathbf{D}$, a *lifting* of a $\mathbf{D}$-functor $F \colon \mathbf{D} \to \mathbf{D}$ from $\mathbf{D}$ to $\mathbf{C}$ is a functor $\overline{F} \colon \mathbf{C} \to \mathbf{C}$ such that $P \circ \overline{F} = F \circ P$. Here we only consider the case when $\mathbf{D} = \mathbf{Set}$ and $\mathbf{C} = \mathcal{V}\text{-}\mathbf{PMet}$, with P the forgetful functor sending a $\mathcal{V}$-pseudometric of type $X \times X \to \mathcal{V}$ to its underlying set X and acting as the identity on arrows. We define coupling-based liftings, codensity liftings, and the associated notion of correspondence that we study here.

Given $t_1, t_2 \in FX$, a *coupling* of t_1 and t_2 is an element $t \in F(X \times X)$ such that $F\pi_i(t) = t_i$. The set of couplings of t_1 and t_2 is denoted by $\Omega(t_1, t_2)$. In particular, if $F = \mathcal{D}$ is the distribution functor on $\mathbf{Set}$, these are precisely probabilistic couplings: joint distributions on $X \times X$ whose marginals coincide with given distributions t_1, t_2. The following lifting arises from [11]. See also [6]. It is not referred to in the literature as "coupling-based"; we use this terminology to distinguish it from the codensity lifting.

▶ **Definition 10.** *Let* $F \colon \mathbf{Set} \to \mathbf{Set}$ *be a functor which preserves weak pullbacks, and let* τ *be a well-behaved modality for* F. *The* coupling-based lifting *of* F *to* $\mathcal{V}\text{-}\mathbf{PMet}$ *is given by, with* $d \in \mathcal{V}\text{-}\mathbf{PMet}$ *and* $t_1, t_2 \in FX$:

$$F_\tau^\downarrow(d)(t_1, t_2) = \bigvee_{t \in \Omega(t_1, t_2)} \tau \circ Fd(t)$$

▶ **Example 11.** In Section 2, we studied a coupling-based lifting for DFA, which yields the shortest-distinguishing-word-distance. To see this as an instance of Definition 10, the quantale is $\mathcal{V} = [0, 1]$ as introduced in Section 3, the functor F is the one mapping a set X to $2 \times X^A$. The well-behaved modality takes as argument an element $(l, \delta) \in 2 \times \mathcal{V}^A$ and returns $\bigwedge_{a \in A} c \cdot \delta(a)$ for some $c \in [0, 1)$. The resulting lifting is precisely the one given in Equation (3).

The *codensity lifting* is defined in a very general setting for monads in [19]. Here, we use an instance, which appears in [24, 21] and, for a single modality, in [2].

▶ **Definition 12.** *Let* $F \colon \mathbf{Set} \to \mathbf{Set}$, *and let* Γ *be a family of modalities for* F. *The* codensity lifting *of* F *along* Γ *to* $\mathcal{V}\text{-}\mathbf{PMet}$ *is defined by, with* $d \in \mathcal{V}\text{-}\mathbf{PMet}$ *and* $t_1, t_2 \in FX$:

$$F_\Gamma^\uparrow(d)(t_1, t_2) = \bigwedge_{\tau \in \Gamma, \ f \colon d \to \mathrm{d}_e} \mathrm{d}_e(\tau \circ Ff(t_1), \tau \circ Ff(t_2))$$

Note that the maps $f \colon d \to \mathrm{d}_e$ in the definition of the codensity lifting are $\mathcal{V}$-pseudometric morphisms, so that, for instance, in $\overline{\mathbb{R}}$, they correspond to non-expansive maps (cf. Remark 5).

▶ **Example 13.** In the example of Section 2 the codensity lifting has one modality mapping $(l, \delta) \in 2 \times \mathcal{V}^A$ to l, as well as one modality τ_a for each letter $a \in A$ given by $\tau_a(l, \delta) = c \cdot \delta(a)$. The resulting codensity lifting coincides with the lifting given in Equation (4).

Note that the coupling-based liftings are only allowed to have one modality, whereas the codensity ones may have multiple modalities. Informally, one of the reasons coupling-based liftings do not require multiple modalities can be seen by looking again at the functor for DFA. Recall that it maps a set X to $2 \times X^A$ for $2 = \{0, 1\}$ and A an alphabet. Considering couplings forces comparisons using d to be done separately on each letter. Codensity liftings instead use one modality per letter to ensure a similar condition. On a higher level of abstraction, the multiple modalities of codensity liftings relate to the expressivity of some modal logics (see, e.g., [20]). On the other hand, the unique modality for coupling-based liftings has often been called an evaluation function (see [2]). Having only one way of evaluating something seems natural in a given evaluation context.

In the rest of this paper we study possible equalities between the coupling-based and codensity liftings. In general, we allow a well-behaved modality for the coupling-based liftings to be matched by a set of (different) modalities for the codensity lifting. In case the modalities are the same, we call this Kantorovich-Rubinstein duality.

▶ **Definition 14.** *Let $F \colon \mathbf{Set} \to \mathbf{Set}$ be a weak-pullback preserving functor, $\tau \colon F\mathcal{V} \to \mathcal{V}$ a well-behaved modality, and Γ a set of modalities for F. A* correspondence *is a triple of the form (F, τ, Γ) such that the associated coupling-based and codensity liftings coincide, as in:*

$$F_\tau^\downarrow = F_\Gamma^\uparrow$$

If $\Gamma = \{\tau\}$ then we refer to this as Kantorovich-Rubinstein duality*, or simply duality.*

Two correspondences (F, τ_1, Γ_1) and (F, τ_2, Γ_2) for the same functor are called *equivalent* when they yield the same liftings:

$$F_{\tau_1}^\downarrow = F_{\Gamma_1}^\uparrow = F_{\Gamma_2}^\uparrow = F_{\tau_2}^\downarrow$$

This is denoted by $(F, \tau_1, \Gamma_1) \sim (F, \tau_2, \Gamma_2)$.

When the associated liftings are not equal but one *inequality* holds nonetheless such as

$$F_{\tau_1}^\downarrow = F_{\Gamma_1}^\uparrow \leq F_{\Gamma_2}^\uparrow = F_{\tau_2}^\downarrow$$

we write $(F, \tau_1, \Gamma_1) \leq (F, \tau_2, \Gamma_2)$.

4.3   Kantorovich-Rubinstein dualities and tools to get them

The focus of this paper is on general correspondences between coupling-based and codensity liftings, but in some cases there is the stronger property of Kantorovich-Rubinstein duality, meaning correspondences like $(F, \tau, \{\tau\})$. In the remainder of this section we consider a few technical definitions and results that are useful in obtaining such duality results.

First, coupling-based liftings are always smaller than codensity liftings. This is a known result; it is a direct consequence of codensity liftings being initial in a well-chosen category as shown in [10]. It can also be found in [2, Theorem 5.27] in the particular case of the quantale $[0, M]$ as presented in Section 3.

▶ **Proposition 15.** *Given a $\mathbf{Set}$ endofunctor F and an associated well-behaved modality, we have $F_\tau^\downarrow \leq F_{\{\tau\}}^\uparrow$.*

Hence to get a Kantorovich-Rubinstein duality, we only need the other inequality. Two tools are introduced for that: optimal couplings to know the value of coupling-based liftings and optimal functions to know that of codensity liftings.

Given a functor $F \colon \mathbf{Set} \to \mathbf{Set}$, an associated modality τ, a $\mathcal{V}$-pseudometric $d \colon X \times X \to \mathcal{V}$, and $t_1, t_2 \in FX$, an *optimal coupling* is a coupling $t \in F(X \times X)$ of t_1 and t_2 such that

$$F_\tau^\downarrow d(t_1, t_2) = \tau \circ Fd(t)$$

When for all d, t_1, and t_2 either an optimal coupling exists or no couplings of t_1 and t_2 exist we say that *F has all optimal couplings*. Optimal couplings are well-known in the context of transport theory [27] and in existing proofs of duality for the finite powerset functor [12, 2]. We note that, whenever all elements in a class of functors $\mathcal{F}$ have all optimal couplings, then functors in the closure of $\mathcal{F}$ by coproducts and products do too.

An *optimal function* for a functor F, an associated modality τ, a $\mathcal{V}$-pseudometric $d\colon X \times X \to \mathcal{V}$, and $t_1, t_2 \in FX$ is a morphism in $\mathcal{V}\text{-}\mathbf{PMet}$ $f\colon d \to \mathrm{d}_e$ such that:

$$F^{\uparrow}_{\{\tau\}}d(t_1, t_2) = \mathrm{d}_e(\tau \circ Ff(t_1), \tau \circ Ff(t_2))$$

In the context of the general continuous Kantorovich-Rubinstein duality from optimal transport theory such optimal functions do exist (see [27, proof of Theorem 5.10]) but are not explicitly defined. As we will see they exist for all the functors we consider.

The following lemma is useful to design optimal functions. It is well-known in the context of quantale-enriched categories. Instead of defining some pseudometric morphism $f\colon d \to \mathrm{d}_e$ on X as a whole, it can be defined on $Y \subseteq X$ only, first as some morphism $g\colon d \circ i \to \mathrm{d}_e$ and then extended using the lemma. This is useful when only parts of X are of interest, say some subset of it for the powerset functor, or the support of some probability distribution for the distribution functor. The extension of g to f from Y to X is done by giving the greatest values to f on $X \backslash Y$ while ensuring it is a $\mathcal{V}$-pseudometric morphism.

▶ **Lemma 16.** *Let $d\colon X \times X \to \mathcal{V}$ be a $\mathcal{V}$-pseudometric, and $Y \overset{i}{\subseteq} X$. For all $\mathcal{V}$-pseudometric morphisms $g\colon d \circ (i \times i) \to \mathrm{d}_e$ there exists $f\colon d \to \mathrm{d}_e$ s.t. $g = f \circ i$ and f is the least such morphism.*

5 Correspondences through coproducts and products

In this section we obtain new correspondences between coupling-based and codensity liftings, for polynomial functors. More specifically, given a set of functors $\mathcal{F}$ and associated correspondences (F, τ_F, Γ_F), these are extended to correspondences for the coproduct (Section 5.1) and product (Section 5.2) of the underlying functors. In Section 5.3 we study correspondences for constant and identity functors. Finally in Section 5.4 we combine these results to describe several collections of functors for which we have correspondences. Throughout this section we fix a quantale $\mathcal{V}$.

5.1 Coproduct functors

We start by showing that correspondences are closed under coproducts. Given sets S_i we write κ_j for the jth *coprojection* $\kappa_j\colon S_j \to \coprod S_i$.

▶ **Proposition 17.** *Given correspondences $(F_i, \tau_i\colon F_i\mathcal{V} \to \mathcal{V}, \Gamma_i)$ there is a correspondence*

$$\left(\coprod F_i, [\tau_i], \bigcup \overline{\Gamma_i} \right) \qquad \text{where } \overline{\Gamma_i} = \{\overline{\tau} \mid \tau \in \Gamma_i\} \cup \{\tau_{\top,i}\};$$

the map $[\tau_i]\colon \coprod F_i\mathcal{V} \to \mathcal{V}$ is the cotupling of the individual modalities; for $\tau \in \Gamma_i$,

$$\overline{\tau}(x) = \begin{cases} \tau(y) & \text{when there exists } y \in F_i\mathcal{V} \text{ and } x = \kappa_i y \\ \top & \text{otherwise} \end{cases}$$

and finally, $\tau_{\top,i}(x) = \begin{cases} \top & \text{when there exists } y \in F_i\mathcal{V} \text{ and } x = \kappa_i y \\ \bot & \text{otherwise} \end{cases}$.

When there are no couplings for two given elements, as it happens for instance for elements in different components of a coproduct, the coupling-based lifting always gives the value $\bot$. The modalities $\tau_{\top,i}$ are there to ensure that the codensity lifting also gives the value $\bot$ in this case. The other modalities are just the extensions of the modalities from the components of the coproduct to the coproduct itself.

▶ **Remark 18.** Proposition 17 gives a correspondence but not a duality, in the sense of Definition 14, that is, the correspondence relates a single modality for the coupling-based lifting to a set of modalities for the codensity lifting. This is the case even if the correspondences of the component functors F_i are dualities.

As it turns out, the inverse process of going from a correspondence at the level of the coproduct to correspondences for its components is also possible:

▶ **Proposition 19.** *If there is a correspondence* $(\coprod F_i, \tau, \Gamma)$, *then there are correspondences* $(F_i, \tau \circ \kappa_i, \Gamma \circ \kappa_i)$ *for each i.*

Going back and forth between the components of a coproduct and the coproduct itself using the results above gives back the same correspondences:

▶ **Proposition 20.** *Whenever one of the correspondences on the left exists we also have the corresponding equivalence:*

$$(F_i, \tau_i, \Gamma_i) \quad \sim \quad \left(F_i, [\tau_j] \circ \kappa_i, \bigcup \overline{\Gamma_j} \circ \kappa_i\right) \ \ and$$

$$\left(\coprod F_i, \tau, \Gamma\right) \quad \sim \quad \left(\coprod F_i, [\tau \circ \kappa_i], \bigcup \overline{\Gamma \circ \kappa_i}\right)$$

This shows that all correspondences on the coproduct arises as in Proposition 17.

5.2 Product functors

We turn to the construction of correspondences for products from ones on their components. For this, we ask for some kind of distributivity condition on the quantale. In fact, the case of products is more involved than that of coproducts. Each of the three propositions for coproducts above has a version for products, but each gets their own specific restrictions in the form of conditions on the quantale and modalities or even on the statement itself.

There are several possibilities depending on the number of couplings and on whether we want finite or arbitrary products. Below, we say that a functor F has *finite couplings* if, for any $t_1, t_2 \in FX$, the set $\Omega(t_1, t_2) \subseteq F(X \times X)$ of couplings is finite.

▶ **Proposition 21.** *Let (F, τ_F, Γ_F) and (G, τ_G, Γ_G) be correspondences. If one of the following conditions holds:*

1. *F and G have finite couplings and $\mathcal{V}$ is distributive, meaning for all $x, y, z \in \mathcal{V}$, we have that $(x \vee z) \wedge (y \vee z) = (x \wedge y) \vee z$; or,*
2. *$\mathcal{V}$ is join-infinite distributive: for all $x \in \mathcal{V}$ and $V \subseteq \mathcal{V}$, $x \wedge \bigvee V = \bigvee \{x \wedge v \mid v \in V\}$;*

then we have a correspondence $(F \times G, (\tau_F \circ \pi_1) \wedge (\tau_G \circ \pi_2), (\Gamma_F \circ \pi_1) \cup (\Gamma_G \circ \pi_2))$.

Under stronger distributivity conditions, this can be extended to infinite products.

▶ **Proposition 22.** *Let $(F_i, \tau_{F_i}, \Gamma_{F_i})$ be correspondences. If one of the following holds:*

1. *F_i has finite couplings and $\mathcal{V}$ is meet-infinite distributive, meaning that for all $x \in \mathcal{V}$ and $V \subseteq \mathcal{V}$, $x \vee \bigwedge V = \bigwedge \{x \vee v \mid v \in V\}$; or,*
2. *$\mathcal{V}$ is completely distributive, meaning that for all sets $K \subseteq I \times J$ such that K projects onto I, and any subset $\{x_{ij} \mid (i,j) \in K\} \subseteq \mathcal{V}$, $\bigwedge_{i \in I} \left(\bigvee_{j \in K(i)} x_{ij} \right) = \bigvee_{f \in A} \left(\bigwedge_{i \in I} x_{if(i)} \right)$ where $A = \{f \colon I \to J \mid \forall i \in I, \ f(i) \in K(i)\}$;*

then we have a correspondence $(\prod F_i, \bigwedge(\tau_{F_i} \circ \pi_i), \bigcup(\Gamma_{F_i} \circ \pi_i))$.

▶ **Example 23.** The quantales 2 and $[0, M]$ from Example 2 are both completely distributive.

Replacing the infinite distributivity conditions on the quantale by countably infinite versions directly gives similar statements for countable products.

In order for the inverse process from products to components to be uniquely defined we first need to ensure that the functor has been decomposed as much as possible using coproducts following Propositions 17, 19, and 20. The following well-known lemma is useful for this (see [22, Proposition 1.3.12]):

▶ **Lemma 24.** *Let $F \colon \mathbf{Set} \to \mathbf{Set}$ be a functor. There is a family of functors $\{F_i\}_{i \in F\{\top\}}$ indexed by $F\{\top\}$ such that $F \simeq \coprod F_i$, and if $F_i = \coprod G_j$ then all G_j but one are the empty functor sending any set to the empty set.*

Having characterised how functors decompose along coproducts and how correspondences work with regard to such decompositions, we can now assume functors not to be writable as non-trivial coproducts meaning that if a functor can be written as a coproduct, all but one of the components must be the empty functor. Furthermore we assume for convenience that functors are not empty functors. Those two hypotheses can be formalised as follows: by the previous lemma, functors send $\{\top\}$ and by extension any singleton to a singleton. Hence, given τ well-behaved, with $i \colon \{\top\} \hookrightarrow \mathcal{V}$, $Fi(F\{\top\}) = \tau^{-1}(\top)$ must be a singleton.

▶ **Proposition 25.** *Given a correspondence $\left(\prod_{i \in I} F_i, \tau, \Gamma \right)$, we again have correspondences $\left(F_i, \tau_{|i}, \Gamma_{|i} \right)$ where $\tau_{|i}(x) = \tau(\top_1, \ldots, \top_{i-1}, x, \top_{i+1}, \ldots, \top_I)$, with $(\top_j)_j = \tau^{-1}(\top)$ and similarly for modalities in $\Gamma_{|i}$.*

In a similar fashion as for the coproduct, we would like that going back and forth between products and their components, following Proposition 21 or 22 and Proposition 25, preserves correspondences. We can only do that partially:

▶ **Proposition 26.** *Whenever the correspondences on the left exist and we have the right distributivity conditions from Proposition 21 or 22 for the correspondence on the right to exist, we have the following equivalence and inequality:*

$$
\begin{aligned}
(F_i, \tau_i, \Gamma_i) &\sim \left(F_i, \left(\bigwedge(\tau_j \circ \pi_j) \right)_{|i}, \left(\bigcup(\Gamma_j \circ \pi_j) \right)_{|i} \right) \\
\left(\prod F_i, \tau, \Gamma \right) &\leq \left(\prod F_i, \bigwedge(\tau_{|i} \circ \pi_i), \bigcup(\Gamma_{|i} \circ \pi_i) \right)
\end{aligned}
$$

The inequality comes from the following aspect of the proof. When using the correspondences on the components of a product of functors to retrieve a correspondence on the product itself, a choice is made to build modalities for the product. The natural choice of using a meet ensures the inequality while there is no obvious way to get an equivalence of correspondences.

5.3 Constant and identity functors

In this subsection we give correspondence results for constant functors, and for the identity functor. For constant functors, the essence is given by the constant-to-**1** functor.

▶ **Proposition 27.** *The triple* $(\mathbf{1}, \tau, \Gamma)$ *with the functor* $\mathbf{1}$ *sending any set to a fixed singleton is a correspondence if and only if* $\Gamma = \{\tau\}$ *and* τ *is the constant to* $\top$ *modality.*

Proof. We write $\mathbf{1} = \{*\}$.

$\Rightarrow$: as modalities $\tau \colon \{*\} \to \mathcal{V}$ are assumed well-behaved, given the inclusion $i \colon \{\top\} \hookrightarrow \mathcal{V}$ we must have $\mathbf{1}i[\mathbf{1}\{\top\}] = \tau^{-1}(\top)$ which translates to $\tau^{-1}(\top) = \{*\}$.

$\Leftarrow$: both liftings lift all pseudometrics to the one constant to $\top$. ◄

Next, we extend this result to arbitrary constant functors. There is only one correspondence result in this case.

▶ **Corollary 28.** *For A a set, denoting by A the associated* constant functor *mapping all sets to A, there is a correspondence* $(A, \tau_\top, \{\tau_a \mid a \in A\})$ *with* $\tau_\top$ *the constant to* $\top$ *modality and*

$$\tau_a(b) = \begin{cases} \top \text{ when } b = a \\ \bot \text{ otherwise} \end{cases}$$

Furthermore, any other correspondence (A, τ, Γ) *is equivalent to* $(A, \tau_\top, \{\tau_a \mid a \in A\})$.

Proof. This is a direct consequence of Propositions 17, 19, 20 and 27 viewing the functor constant and equal to A as a coproduct:

$$A \simeq \coprod_{a \in A} \mathbf{1}$$

◄

We give a general duality result for the *identity functor* Id acting as the identity on both sets and functions, generalising [2, Example 5.29]. An instance of well-behaved modalities for the identity functor was discussed in Example 8.

▶ **Proposition 29.** *For any well-behaved modality* $\tau \colon \mathcal{V} \to \mathcal{V}$ *for the identity functor, there is a correspondence* $(\mathrm{Id}, \tau, \{\tau\})$.

Proof. By Proposition 15 we need only prove

$$\mathrm{Id}_\tau^\downarrow \geq \mathrm{Id}_\tau^\uparrow$$

Let X be a set, $d \colon X \times X \to \mathcal{V}$ a $\mathcal{V}$-pseudometric, and $x, y \in X$. There is exactly one coupling of x with y given by (x, y). In particular it is optimal:

$$\mathrm{Id}_\tau^\downarrow d(x, y) = \tau(d(x, y))$$

Furthermore,

$$\mathrm{Id}_\tau^\uparrow d(x, y) = \bigwedge_{f \colon d \to \mathrm{d}_e} \mathrm{d}_e(\tau \circ f(x), \tau \circ f(y))$$

Using Lemma 16, we define $f \colon d_{|\{x\}} \to \mathrm{d}_e$ by $f(x) = \top$ and extend it to a morphism $f \colon d \to \mathrm{d}_e$. This gives $f(y) = d(x, y)$ so that $\mathrm{d}_e(\tau \circ f(x), \tau \circ f(y)) = \tau(d(x, y))$. Hence

$$\mathrm{Id}_\tau^\uparrow d(x, y) \leq \tau \circ d(x, y) = \mathrm{Id}_\tau^\downarrow d(x, y)$$

ending the proof. Note that f is *a fortiori* an optimal function. ◄

5.4 Putting it together: correspondences for polynomial functors

Having seen correspondences for products, coproducts, the identity functor, and constant functors, we combine these results to obtain a "grammar" of functors with correspondences.

We start without products. Using Propositions 17 and 29 as well as Corollary 28 gives correspondences for functors in the following grammar of **Set** endofunctors:

$$F ::= A \mid \mathrm{Id}_\tau \mid \coprod F_i$$

where τ indicates choices of well-behaved modalities in the case of the identity functor. Furthermore by Propositions 20 and 29 and Corollary 28, any coupling-based lifting of a functor in this grammar is equivalent to a correspondence given by our construction.

Observing that for a set A, $\coprod_{a \in A} F \simeq A \times F$, the grammar can equivalently be defined as

$$F ::= A \mid \mathrm{Id}_\tau \mid A \times F \mid \coprod F_i$$

where τ is a well-behaved modality for the identity functor. We can also note that these functors are exactly those isomorphic to $A + B \times \mathrm{Id}$ for some sets A and B.

▶ **Example 30** (Discounting on streams). Consider the stream functor $X \mapsto A \times X$ for some alphabet A. The associated final coalgebra is the set of streams A^ω. The results above tell us that to get a correspondence we need one well-behaved modality $\tau : \mathcal{V} \to \mathcal{V}$ per letter $a \in A$ for the codensity lifting, through the isomorphism $A \times \mathrm{Id} \simeq \coprod_{a \in A} \mathrm{Id}$: we have a way of computing distance between streams in a specific way for each letter in A. When $\mathcal{V} = ([0, M], \geq, +)$ for some real number $M \in (0, \infty]$, choosing $\tau_a(x) = c \cdot x$, the constant c is then a discount factor allowing one to give more value to the first letters of a stream. We can choose a different constant c_a for each letter a giving different values to different letters. We can also go further: if we do not want a linear discount we can use arbitrary well-behaved modalities for the identity functor. On the quantale $([0, M], \geq, +)$, those are exactly sub-additive monotone maps mapping 0 to 0.

We also get correspondences for the following grammar of *simple polynomial functors*:

$$F ::= A \mid \mathrm{Id}_\tau \mid A \times F \mid \coprod F_i \mid \prod F_i$$

These functors all have finite couplings. Following Propositions 21 and 22 we can either get finite or arbitrary products in the grammar above by assuming $\mathcal{V}$ to be distributive or meet-infinite distributive respectively.

▶ **Example 31.** We retrieve the correspondence of Section 2 as the particular case of the quantale $([0, M], \geq, +)$, the functor $2 \times \mathrm{Id}^A$, and indexed modalities τ for the identity functors always mapping $x \in [0, M]$ to $c \cdot x$ for some $c \in [0, 1)$. Replacing $[0, M]$ by the usual Boolean quantale $2 = \{\top, \bot\}$ and indexed modalities by the identity functions gives a correspondence equivalent to the usual relation lifting [14] associated to behavioural equivalence of DFA.

6 Dualities for the powerset and probability distribution functors

We give duality results for the powerset functor mapping a set X to the set of its subsets $\mathcal{P}X$, recovering a result from the literature [2, 12] reproduced here as Corollary 35, and complementing it with a duality result where $\mathcal{V}$ is a total order (Corollary 34).

▶ **Remark 32.** The non-empty powerset functor $\mathcal{P}_{\geq 1}$ mapping a set X to the set of its non-empty subsets might have been chosen instead of $\mathcal{P}$. Note that $\mathcal{P} = \mathbf{1} + \mathcal{P}_{\geq 1}$. Up-to equivalence, as a consequence of Propositions 17, 19 and 27 a correspondence for $\mathcal{P}$ is exactly given by one for $\mathcal{P}_{\geq 1}$ and conversely.

Given $T_1, T_2 \in \mathcal{P}X$, $t_1 \in T_1$, and a $\mathcal{V}$-pseudometric $d\colon X \times X \to \mathcal{V}$, we write M_{t_1} for the *set of maximal elements* of $\{d(t_1, t_2) \mid t_2 \in T_2\}$, and similarly for elements $t_2 \in T_2$. Expressing a coupling-based lifting as a codensity one in a correspondence implies turning a join in a meet. The following proposition gives a condition allowing one to do that in the case of the powerset functor by making a modality absorb a meet.

▶ **Proposition 33.** *Let $\tau\colon \mathcal{PV} \to \mathcal{V}$ be a well-behaved modality. If for every $\mathcal{V}$-pseudometric $d\colon X \times X \to \mathcal{V}$ and sets $T_1, T_2 \in \mathcal{P}X$ the following equality holds*

$$\tau\left\{\bigvee M_{t_1} \mid t_1 \in T_1\right\} \bigwedge \tau\left\{\bigvee M_{t_2} \mid t_2 \in T_2\right\} = \tau\left(\left(\bigcup_{t_1 \in T_1} M_{t_1}\right) \cup \left(\bigcup_{t_2 \in T_2} M_{t_2}\right)\right)$$

then there is a duality $(\mathcal{P}, \tau, \{\tau\})$.

▶ **Corollary 34.** *Let $\tau\colon \mathcal{PV} \to \mathcal{V}$ be a well-behaved modality. If the order on $\mathcal{V}$ is total then there is a duality $(\mathcal{P}, \tau, \{\tau\})$.*

▶ **Corollary 35.** *We assume $\mathcal{V}$ to be completely distributive. We have duality for the coupling-based and codensity liftings of the powerset functor, both along the meet modality which maps a subset of $\mathcal{V}$ to the meet of its elements.*

For the probability distribution functor we fix a constant $M \in (0, \infty]$ and use the quantale $([0, M], \geq, +)$. The following (Kantorovich-Rubinstein) duality result is well-known (see, e.g., [27, Theorem 5.10] for a general proof in the continuous case and [16, Appendix A] for the discrete finite case):

▶ **Proposition 36.** *There is a duality $(\mathcal{D}, \mathbb{E}, \{\mathbb{E}\})$ with $\mathcal{D}$ the finite probability distribution functor and $\mathbb{E}$ giving the expectation of probability distributions.*

We extend this result slightly to allow post-composition by well-behaved modalities:

▶ **Proposition 37.** *Let $\tau\colon [0, M] \to [0, M]$ be a well-behaved modality. If τ is additive, meaning that for all $x, y \in [0, M]$, $\tau(x + y) = \tau(x) + \tau(y)$, then there is a duality $(\mathcal{D}, \tau \circ \mathbb{E}, \{\tau \circ \mathbb{E}\})$.*

7 Correspondences for grammars of functors

In this section we combine the results from the previous two sections, allowing the construction of correspondences for certain classes of functors including powerset, distribution, identity and constant functors, as well as products and coproducts thereof.

Whenever $\mathcal{V}$ is totally ordered and completely distributive, using results from Sections 5 and 6 we can construct correspondences for the following grammar of functors:

$$F ::= A \mid \mathrm{Id}_\tau \mid \mathcal{P}_\tau \mid \prod F_i \mid \coprod F_i$$

where the indices τ are well-behaved modalities for either the identity or powerset functors.

Noting that $\mathcal{P}\left(\coprod_{i \in I} F_i\right) \cong \prod_{i \in I} \mathcal{P}(F_i)$, this grammar can be reformulated as follows:

$$G ::= A \mid \mathrm{Id}_\tau \mid A \times G \mid \coprod G$$

$$F ::= A \mid \mathrm{Id}_\tau \mid \mathcal{P} \circ G \mid \prod F_i \mid \coprod F_i$$

where the indices of the form τ indicate a choice of possible well-behaved modalities for the identity functor when appearing in F and for the powerset functor when appearing in G.

▶ **Example 38.** In a similar fashion as in Example 31, considering the completely distributive quantale $([0, M], \geq, +)$, the functor $2 \times \mathcal{P}^A$ which is associated to *non-deterministic automata*, and indexed modalities for the powerset functors to be all mapping finite subsets $V \subset [0, M]$ to $c \cdot \max(V)$ for a fixed $c \in [0, 1)$, we get a lifting associated to a shortest-distinguishing-word-distance for non-deterministic systems as is detailed in Section 2 for DFA.

▶ **Remark 39.** This grammar with finite products only defines a fragment of finitary Kripke polynomial functors as defined in [14] which correspond to replacing $\mathcal{P}(G)$ by $\mathcal{P}(F)$, and having only finite products but including arbitrary exponents.

▶ **Remark 40.** Using the identity modality for the identity functors and the meet modality for the powerset functors, the modalities given by this construction for the coupling-based liftings are exactly the *canonical evaluation maps* as defined in [6].

In the particular case of $\mathcal{V} = [0, M]$ we can extend the above grammar with the finite probability distribution functor and get correspondence results for the following:

$$F ::= A \mid \mathrm{Id}_\tau \mid \mathcal{P}_\tau \mid \mathcal{D}_\tau \mid \prod F_i \mid \coprod F_i$$

where the indices τ indicate choices of well-behaved modalities for either the identity, powerset, or distribution functors, furthermore of a modality of the form $\tau \circ \mathbb{E}$ for the latter as described in Proposition 37.

Observe that $\mathcal{D}\left(\coprod_{i \in I} F_i\right) \cong \mathcal{D}(I) \times \prod_{i \in I} \mathcal{D}(F_i)$ so that the grammar above can be expressed as follows:

$$G ::= A \mid \mathrm{Id}_\tau \mid A \times G \mid \coprod G$$
$$F ::= A \mid \mathrm{Id}_\tau \mid \mathcal{P} \circ G \mid \mathcal{D} \circ G \mid \prod F_i \mid \coprod F_i$$

where the indices τ indicate choices of possible well-behaved modalities for the identity functor in Id_τ and for the powerset or distribution functors when appearing in G for $\mathcal{P} \circ G$ and $\mathcal{D} \circ G$ respectively.

▶ **Example 41.** We fix $M = 1$ so that $\mathcal{V} = [0, 1]$. Consider the functor $\mathcal{D}(\mathbf{1} + \mathrm{Id})^A$ for a set of labels A and $\mathbf{1}$ a singleton as in Proposition 27. It is associated to *labelled Markov processes*. To get a correspondence we need one modality $\tau_a \colon \mathcal{D}[0, 1] \to [0, 1]$ per label $a \in A$. By Proposition 37 we can take them all to be $\tau_a(\mu) = c \cdot \mathbb{E}(\mu)$ with $c \in [0, 1)$ a fixed constant. The resulting correspondence is associated to the usual metrics for labelled Markov processes as introduced in [8]. This is a direct consequence of [26, Proposition 29] which expresses the associated lifting for one label in a "codensity-like" manner.

8 A counter-example: conditional transition systems

We now highlight an example where a correspondence can *not* be provided by our construction: some codensity lifting may not be obtained as a coupling-based lifting.

A *conditional transition system* (CTS) over an alphabet A and a partially ordered set of conditions $(\mathcal{L}, \leq)$ is a tuple $(X, A, \mathcal{L}, \delta)$ where X is a set of states and the transition map $\delta \colon X \times A \to ((\mathcal{L}, \leq) \to (\mathcal{P}(X), \supseteq))$ associates to each pair $(x, a) \in X \times A$ a monotone function $\delta_{x,a}$ from the poset of conditions to $\mathcal{P}(X)$. For the associated functor to live in **Set** we restrict to the CTSs for which $\mathcal{L}$ is trivially ordered by $\forall x, y \in \mathcal{L}, x \leq y$. Hence any map $\delta_{x,a} \colon \mathcal{L} \to \mathcal{P}(X)$ is monotone, and CTSs are exactly coalgebras for the functor $F = \mathcal{P}(-)^{A \times \mathcal{L}}$ (see [4]).

The associated bisimulations, not detailed here, are related to a lifting $\overline{F}\colon \mathcal{P}(\mathcal{L})\textbf{-PMet} \to \mathcal{P}(\mathcal{L})\textbf{-PMet}$ of F defined by

$$\overline{F}r(T_1, T_2) = \{l \in \mathcal{L} \mid \forall a \in A, \forall (x, y) \in T_1(l, a) \times T_2(l, a),$$
$$\exists (x', y') \in T_1(l, a) \times T_2(l, a),$$
$$(l \in r(x, y')) \text{ and } (l \in r(x', y))\}$$

on objects and $\overline{F}f = Ff$ on maps.

▶ **Proposition 42.** *For $a \in A$ consider a modality $\tau_a \colon \mathcal{P}(\mathcal{P}(\mathcal{L}))^{\mathcal{L} \times A} \to \mathcal{P}(\mathcal{L})$, defined by:*

$$\forall f \colon \mathcal{L} \times A \to \mathcal{P}(\mathcal{P}(\mathcal{L})), \tau_a(f) = \{l \in \mathcal{L} \mid l \in \cap f(l, a)\}$$

The codensity lifting defined by the family $(\tau_a)_{a \in A}$ of modalities is equal to $\overline{F}$.

▶ **Proposition 43.** *If A is non-empty and $|\mathcal{L}| \geq 2$ then there is no well-behaved modality $\tau \colon \mathcal{P}(\mathcal{P}(\mathcal{L}))^{\mathcal{L} \times A} \to \mathcal{P}(\mathcal{L})$ such that the resulting coupling-based lifting is equal to $\overline{F}$.*

Proof. Let $a \in A$ be some letter and $c_1, c_2 \in \mathcal{L}$ be two distinct conditions. Consider $d \colon X \times X \to \mathcal{P}(\mathcal{L})$ constant to $\top$, that is, $\mathcal{L}$. It is obviously a $\mathcal{P}(\mathcal{L})$-pseudometric. Finally we consider $T_1, T_2 \in \mathcal{P}(X)^{\mathcal{L} \times A}$ such that $T_1(c_1, a) = \emptyset$ and $T_1(c_2, a) = X$, and $T_2(c_1, a) = T_2(c_2, a) = X$. Directly, because $T_1(c_1, a) = \emptyset$ but $T_2(c_1, a) \neq \emptyset$ there are no couplings of T_1 and T_2, and for all well-behaved modalities, $F_\tau^\downarrow d(T_1, T_2) = \emptyset$. On the other hand,

$$\forall (x, y) \in T_1(c_2, a) \times T_2(c_2, a), \exists (x', y') \in T_1(c_2, a) \times T_2(c_2, a), (l \in d(x, y')) \text{ and } (l \in d(x', y))$$

Hence $c_2 \in \overline{F}d(T_1, T_2) \neq \emptyset$. ◀

Hence, conditional transition systems give a limitation to the correspondence results provided above. Note however that the associated functor has a correspondence induced by a grammar above. The reason it is not equivalent to the lifting $\overline{F}$ is that it considers the set $\mathcal{L} \times A$ as being a set of actions, while we want conditions in $\mathcal{L}$ to be independent of one another and compared separately. Indeed, conditions are fixed throughout the execution of a CTS whereas actions may change at each step.

9 Conclusions and future work

We have studied correspondences between coupling-based and codensity liftings, moving from the classical Kantorovich-Rubinstein duality for distributions to different types of endofunctors on **Set**. In particular, we have shown that such types of correspondences are closed under coproducts and products and used that to provide explicit correspondences for several grammars of functors, including polynomial functors with the possibility of extending them using the powerset and the probability distribution functors. This instantiates to usual liftings of functors associated to (non)deterministic finite automata, or labelled Markov processes, both with discount.

In [10] the authors have shown that on an abstract level all coupling-based liftings are in fact codensity liftings, implying that all coupling-based liftings have some associated correspondences. Section 8 shows that the converse does not hold by providing an example of lifting that arises as a codensity lifting but not as a coupling-based one. Our work proves that correspondences for coproducts of functors are characterised by ones for the components of the coproducts, but gives no such result for products of functors. For example we could consider

the coupling-based lifting of the *diagonal functor* $\Delta\colon X \mapsto X \times X$ along the well-behaved modality $\otimes\colon \mathcal{V} \times \mathcal{V} \to \mathcal{V}$. The question of whether it arises from coupling-based liftings of the identity functors, and the problem of relating it to a codensity lifting in a correspondence remain open. Note however that if coupling-based liftings require their modalities to be well-behaved, codensity liftings, as defined in [10] need no such assumption and can be defined along sets of any modalities. Hence it is possible that to see the coupling-based lifting of Δ along $\otimes$ as a codensity lifting it is needed to consider general modalities for the latter.

In Section 8 we have seen that a certain lifting for conditional transition systems can not be obtained as a coupling-based lifting. This leads to the question of whether the definition of coupling-based liftings could be extended somehow to encompass this example and obtain a correspondence with the existing codensity lifting.

References

1 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. On-the-fly computation of bisimilarity distances. *Log. Methods Comput. Sci.*, 13(2), 2017. `doi:10.23638/LMCS-13(2:13)2017`.

2 Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Coalgebraic behavioral metrics. *Log. Methods Comput. Sci.*, 14(3), 2018. `doi:10.23638/LMCS-14(3:20)2018`.

3 Harsh Beohar, Barbara König, Sebastian Küpper, and Alexandra Silva. Conditional transition systems with upgrades. *Sci. Comput. Program.*, 186, 2020. `doi:10.1016/J.SCICO.2019.102320`.

4 Harsh Beohar, Barbara König, Sebastian Küpper, Alexandra Silva, and Thorsten Wißmann. A coalgebraic treatment of conditional transition systems with upgrades. *Log. Methods Comput. Sci.*, 14(1), 2018. `doi:10.23638/LMCS-14(1:19)2018`.

5 Filippo Bonchi, Barbara König, and Daniela Petrisan. Up-to techniques for behavioural metrics via fibrations. In Sven Schewe and Lijun Zhang, editors, *29th International Conference on Concurrency Theory, CONCUR 2018, September 4-7, 2018, Beijing, China*, volume 118 of *LIPIcs*, pages 17:1–17:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. `doi:10.4230/LIPICS.CONCUR.2018.17`.

6 Filippo Bonchi, Barbara König, and Daniela Petrisan. Up-to techniques for behavioural metrics via fibrations. *Math. Struct. Comput. Sci.*, 33(4-5):182–221, 2023. `doi:10.1017/S0960129523000166`.

7 Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, R. Mohayaee, and A. Sobolevskiĭ. Reconstruction of the early Universe as a convex optimization problem. *Monthly Notices of the Royal Astronomical Society*, 346(2):501–524, December 2003. `doi:10.1046/j.1365-2966.2003.07106.x`.

8 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled markov processes. *Theor. Comput. Sci.*, 318(3):323–354, 2004. `doi:10.1016/J.TCS.2003.09.013`.

9 Alessandro Giacalone, Chi-Chang Jou, and Scott A. Smolka. Algebraic reasoning for probabilistic concurrent systems. In Manfred Broy and Cliff B. Jones, editors, *Programming concepts and methods: Proceedings of the IFIP Working Group 2.2, 2.3 Working Conference on Programming Concepts and Methods, Sea of Galilee, Israel, 2-5 April, 1990*, pages 443–458. North-Holland, 1990.

10 Sergey Goncharov, Dirk Hofmann, Pedro Nora, Lutz Schröder, and Paul Wild. Kantorovich functors and characteristic logics for behavioural distances. In Orna Kupferman and Pawel Sobocinski, editors, *Foundations of Software Science and Computation Structures - 26th International Conference, FoSSaCS 2023, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2023, Paris, France, April 22-27, 2023, Proceedings*, volume 13992 of *Lecture Notes in Computer Science*, pages 46–67. Springer, 2023. `doi:10.1007/978-3-031-30829-1_3`.

11   Dirk Hofmann. Topological theories and closed objects. *Advances in Mathematics*, 215(2):789–824, 2007. `doi:10.1016/j.aim.2007.04.013`.

12   Dirk Hofmann and Pedro Nora. Hausdorff coalgebras. *Appl. Categorical Struct.*, 28(5):773–806, 2020. `doi:10.1007/S10485-020-09597-8`.

13   Samuel Humeau, Daniela Petrisan, and Jurriaan Rot. Correspondences between codensity and coupling-based liftings, a practical approach, 2024. Version of this paper with the appendix. URL: `https://hal.science/hal-04789352`.

14   Bart Jacobs. *Introduction to Coalgebra: Towards Mathematics of States and Observation*, volume 59 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2016. `doi:10.1017/CBO9781316823187`.

15   Bart Jacobs. Drawing from an urn is isometric. In Naoki Kobayashi and James Worrell, editors, *Foundations of Software Science and Computation Structures - 27th International Conference, FoSSaCS 2024, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2024, Luxembourg City, Luxembourg, April 6-11, 2024, Proceedings, Part I*, volume 14574 of *Lecture Notes in Computer Science*, pages 101–120. Springer, 2024. `doi:10.1007/978-3-031-57228-9_6`.

16   Bart Jacobs. Drawing with distance, 2024. `arXiv:2405.18182`, `doi:10.48550/arXiv.2405.18182`.

17   Leonid Kantorovich. On the translocation of masses (in Russian). *Doklady Akademii Nauk*, 37(7–8):227–229, 1942. Translated to English in Management Science, 5, 1 (1958).

18   Leonid Kantorovich and Gennady S. Rubinstein. On a space of totally additive functions (in Russian). *Vestnik Leningrad. Univ*, 13(7):52–59, 1958.

19   Shin-ya Katsumata, Tetsuya Sato, and Tarmo Uustalu. Codensity lifting of monads and its dual. *Log. Methods Comput. Sci.*, 14(4), 2018. `doi:10.23638/LMCS-14(4:6)2018`.

20   Yuichi Komorida, Shin-ya Katsumata, Clemens Kupke, Jurriaan Rot, and Ichiro Hasuo. Expressivity of quantitative modal logics : Categorical foundations via codensity and approximation. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*, pages 1–14. IEEE, 2021. `doi:10.1109/LICS52264.2021.9470656`.

21   Barbara König and Christina Mika-Michalski. (metric) bisimulation games and real-valued modal logics for coalgebras. In Sven Schewe and Lijun Zhang, editors, *29th International Conference on Concurrency Theory, CONCUR 2018, September 4-7, 2018, Beijing, China*, volume 118 of *LIPIcs*, pages 37:1–37:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. `doi:10.4230/LIPICS.CONCUR.2018.37`.

22   Robert Paré. Taut functors and the difference operator, 2024. `arXiv:2407.21129`.

23   Lutz Schröder. Expressivity of coalgebraic modal logic: The limits and beyond. In Vladimiro Sassone, editor, *Foundations of Software Science and Computational Structures, 8th International Conference, FOSSACS 2005, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2005, Edinburgh, UK, April 4-8, 2005, Proceedings*, volume 3441 of *Lecture Notes in Computer Science*, pages 440–454. Springer, 2005. `doi:10.1007/978-3-540-31982-5_28`.

24   David Sprunger, Shin-ya Katsumata, Jérémy Dubut, and Ichiro Hasuo. Fibrational bisimulations and quantitative reasoning: Extended version. *J. Log. Comput.*, 31(6):1526–1559, 2021. `doi:10.1093/LOGCOM/EXAB051`.

25   Franck van Breugel. Probabilistic bisimilarity distances. *ACM SIGLOG News*, 4(4):33–51, 2017. `doi:10.1145/3157831.3157837`.

26   Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theor. Comput. Sci.*, 331(1):115–142, 2005. `doi:10.1016/J.TCS.2004.09.035`.

27   Cédric Villani et al. *Optimal transport: old and new*, volume 338. Springer, 2009. `doi:10.1007/978-3-540-71050-9`.