Threshold-Based Behavioural Distances

Jonas Forster
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Lutz Schröder
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Paul Wild
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Barbara König
Universität Duisburg-Essen, Germany

Pedro Nora
Universität Duisburg-Essen, Germany

Abstract

Behavioural distances generally offer more fine-grained means of comparing quantitative systems than two-valued behavioural equivalences. They often relate to quantitative modal logics that characterize a given behavioural distance in terms of the induced logical distance. We develop a unified framework for behavioural distances and logics induced by a special type of modalities that lift two-valued predicates to quantitative predicates. A typical example is the probability operator, which maps a two-valued predicate A to a quantitative predicate on probability distributions assigning to each distribution the respective probability of A. Correspondingly, the prototypical example of our framework is ε-bisimulation distance of Markov chains, which has recently been shown to coincide with the behavioural distance induced by the popular Lévy-Prokhorov distance on distributions. Other examples include behavioural distance on metric transition systems and Hausdorff behavioural distance on fuzzy transition systems. We establish a number of general results in this framework, including existence and polynomial-time computation of distinguishing formulae in two characteristic modal logics: A two-valued logic with a notion of satisfaction up to ε, and a quantitative logic. These general results instantiate to new results in many of the mentioned examples. Notably, we obtain polynomial-time computation of distinguishing formulae for ε-bisimulation distance of Markov chains in a quantitative logic featuring a "generally" modality used in probabilistic knowledge representation.

2012 ACM Subject Classification Theory of computation → Concurrency; Theory of computation → Modal and temporal logics

Keywords and phrases Behavioural distance, modal logic, quantitative logic, coalgebra, Sugeno integration

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2026.35

Related Version *Full Version*: `https://arxiv.org/abs/2602.12084` [20]

Funding Research funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project numbers 434050016 (Jonas Forster, Barbara König, Pedro Nora, Lutz Schröder), 531706730 (Lutz Schröder) and 419850228 (Paul Wild).

1 Introduction

In systems carrying quantitative information, behavioural distances often provide a more stable and fine-grained means of comparing the behaviour of processes than two-valued equivalences [22]. Behavioural distances can be treated parametrically over the system type (non-deterministic, weighted, probabilistic etc.) in the framework of universal coalgebra [34], which encapsulates the system type in the choice of a functor on a suitable base category. For instance, Markov chains are coalgebras for the discrete distribution functor on the category of sets. Coalgebraic treatments of behavioural distances are based either on liftings of the given functor to the category of metric spaces [5] or on so-called (quantitative) *lax extensions* of the functor, which extend the functor to act on quantitative relations [41]. Both lax extensions and functor liftings can be induced from a choice of quantitative modalities, interpreted coalgebraically as quantitative predicate liftings [41, 5]. A well-known example is the standard Kantorovich distance on probability distributions, which is induced by the expectation modality; that is, given a metric on the outcome space, the distance of two probability distributions is defined as the supremum of the deviations between the respective expected values, taken over all non-expansive predicates on the outcome space.

In the present work, we develop a coalgebraic framework for behavioural distances that are based on fixing threshold values for allowed deviations in long-term behaviour. One central example is the behavioural distance on labelled Markov chains induced by *ε-bisimilarity* [13]: One defines a notion of ε-bisimulation, which, roughly speaking, works like the two-valued notion of probabilistic bisimulation but allows probabilities to deviate by up to ε; then, the induced behavioural distance between states x, y is the infimum over all ε such that x, y are ε-bisimilar. It was recently shown that this behavioural distance, originally called *ε-distance*, is induced by the Lévy-Prokhorov metric on distributions [14]; we hence refer to it as *Lévy-Prokhorov behavioural distance*. Due to its favourable statistical robustness properties, the Lévy-Prokhorov distance is popular in machine learning tasks such as conformal prediction [3] and corruption resistance [9]; other related distances have been employed, for instance, in mobile security [15] and in differential privacy [8].

Our coalgebraic treatment of similar distances, which we term *threshold-based*, starts from the choice of a set of modalities of particular type: They are induced by predicate liftings that turn 2-valued predicates on the base set into $\mathcal{V}$-valued predicates on the functorial image of this set, where we generally write $\mathcal{V}$ for the unit interval $[0, 1]$; we refer to such predicate liftings as *2-to-$\mathcal{V}$*. For instance, Lévy-Prokhorov behavioural distance is induced by the single 2-to-$\mathcal{V}$ predicate lifting that just evaluates probability distributions on 2-valued predicates. Beyond this, the framework subsumes many known examples of coalgebraic behavioural distance, such as Hausdorff behavioural distances on metric [12] or fuzzy [41] transition systems; the most notable non-example are Kantorovich-type distances on probabilistic systems.

Throughout, we cover both symmetric distances, i.e. behavioural pseudometrics, and asymmetric ones, i.e. behavioural hemimetrics such as simulation distance on metric transition systems [16]. Correspondingly, we work with *ε-simulations*, which take on the nature of bisimulations if the underlying set of modalities is closed under duals. We proceed to develop the meta-theory and algorithmics of this framework, obtaining the following main results.

1. We show that the behavioural distance induced from a notion of ε-(bi-)similarity coincides with the one induced by a *threshold-based lax extension* induced by the given modalities (Theorem 3.11). In the probabilistic case, this lax extension is the Lévy-Prokhorov extension, so one instance of this result is the mentioned characterization of ε-distance as Lévy-Prokhorov behavioural distance [14].

2. From the underlying 2-to-$\mathcal{V}$ predicate liftings, we construct quantitative predicate liftings [35, 41], which lift $\mathcal{V}$-valued predicates to $\mathcal{V}$-valued predicates, by Sugeno integration [37] (Section 6). We show that the lax extension induced by these quantitative *Sugeno modalities* coincides with the threshold-based lax extension from Item (1) (Theorem 6.13). It then follows from general results [27, 41] that the quantitative modal logic of the Sugeno modalities is *characteristic* for threshold-based behavioural distance on finitely branching systems in the sense that logical distance coincides with behavioural distance. (Another consequence is instantiation of a quantitative van-Benthem theorem [40].) In the probabilistic case, the Sugeno modality is precisely the *generally* modality that has been used in probabilistic knowledge representation [35] to model the notion "with high probability"; thus, one instance of our generic result is a recent result showing that the *generally* modality induces the Lévy-Prokhorov extension [43].

3. We introduce a two-valued modal logic equipped with a notion of satisfaction up to ε, which we show characterizes threshold-based behavioural distance on finitely branching systems (Section 5). The probabilistic instance yields Desharnais et al.'s characteristic logic for Lévy-Prokhorov behavioural distance [13], restricted to finitely branching systems. All other instances are new; these include a two-valued characteristic logic for behavioural distance on metric transition systems [12].

4. For both the two-valued and the quantitative modal logic, we provide algorithms that extract distinguishing formulae witnessing lower bounds for behavioural distance from Spoiler strategies in a codensity-style [26] Spoiler-Duplicator game. The extracted formulae are of polynomial dag size (the tree size of distinguishing formulae is worst-case exponential even in standard two-valued Kripkean modal logics, e.g. [17]), and the algorithms run in polynomial time under mild assumptions (Theorems 5.5 and 6.11). In the quantitative case, this recovers a corresponding result for metric transition systems [12]. All other instances are new; notably, we obtain polynomial-time computation of characteristic formulae, both two-valued and quantitative, for Lévy-Prokhorov distance of Markov chains.

We use Lévy-Prokhorov behavioural distance as a running example for our main results, presented in Sections 3–6, and discuss a range of further examples in Section 7.

Omitted proofs and additional details are found in the full version [20].

Related Work. As mentioned above, Desharnais et al. [13] introduce ε-distance on labelled Markov chains and an associated two-valued characteristic modal logic, without however considering computational properties of the logic (they do provide an algorithm for the computation of behavioural distance). The fact that the quantitative modal logic of *generally* is characteristic for Lévy-Prokhorov behavioural distance follows from mentioned recent results [43] by existing coalgebraic quantitative Hennessy-Milner theorems [27, 41, 18], which however rely on fairly involved constructions that approximate non-expansive properties at every modal depth using Stone-Weierstraß-type density theorems, and thus do not imply our algorithmic results on computation of distinguishing formulae. Our generic algorithm for the computation of distinguishing formulae takes inspiration from a corresponding algorithm for the modal logic of metric transition systems [12], which as indicated above is at the same time the only previously known instance of our general result. The computation of distinguishing formulae for two-valued behavioural equivalences goes back to work on labelled transition systems [11]; by now, corresponding algorithms have been designed in coalgebraic generality [28, 45]. (The two-valued setting allows very stringent quasi-linear time bounds [45] that presumably do not generalize to the quantitative case.) As mentioned above, one case not covered by our framework of threshold-based behavioural distances are Kantorovich-type distances on Markov chains, for which an algorithm computing distinguishing formulae has been presented recently [33]. This algorithm works with a different specification: It computes an exact witness of the distance under the n-th iterate of the functional defining the behavioural distance as a fixpoint, while our algorithm finds a formula witnessing distance strictly above a given ε.

2 Preliminaries

We assume basic familiarity with category theory (e.g. [2]). We recall basic notions on distance functions and coalgebras, and fix some notation.

Distance functions and relations. We write $\mathcal{V}$ to denote the unit interval $[0,1] \subseteq \mathbb{R}$, with suprema and infima under the usual ordering of the reals written as $\vee$ and $\wedge$, and $\oplus$ and $\ominus$ used to denote truncated addition and subtraction on $\mathcal{V}$, respectively (i.e. $x \oplus y = \min(x+y, 1)$, $x \ominus y = \max(x-y, 0)$). We generally work with predicates $A \in 2^X$ where $2 = \{\top, \bot\}$, which we silently identify with subsets $A \subseteq X$ when convenient. Let X, Y be sets. For relations $R \subseteq X \times Y$, we write $R\colon X \nrightarrow Y$, and we write $R^\circ\colon Y \nrightarrow X$ for the converse of a relation $R\colon X \nrightarrow Y$. We denote the relational composite of relations $R\colon X \nrightarrow Y, S\colon Y \nrightarrow Z$ by $S \cdot R\colon X \nrightarrow Z$. The relational image $R[A] \in 2^Y$ of $A \in 2^X$ under R is given by $R[A] = \{y \mid \exists x \in A.\, x\ R\ y\}$.

A *$\mathcal{V}$-valued relation* $r\colon X \times Y \to \mathcal{V}$ will be denoted by $r\colon X \nrightarrow_{\mathcal{V}} Y$ and its *converse* by $r^\circ\colon Y \nrightarrow_{\mathcal{V}} X$, with $r^\circ(y,x) = r(x,y)$ for $x \in X$, $y \in Y$. We order $\mathcal{V}$-valued relations pointwise, i.e. by $r \le r'$ iff $r(x,y) \le r'(x,y)$ for all $x \in X$, $y \in Y$. A $\mathcal{V}$-valued relation $d\colon X \nrightarrow_{\mathcal{V}} X$ is *symmetric* if it coincides with its converse, and a *hemimetric* if $d(x,x) = 0$ for all $x \in X$ and $d(x,z) \le d(x,y) \oplus d(y,z)$ for all $x,y,z \in X$; a *pseudometric* is a symmetric hemimetric. Similarly to the qualitative case, the composite $s \cdot r\colon X \nrightarrow_{\mathcal{V}} Z$ of $\mathcal{V}$-valued relations $r\colon X \nrightarrow_{\mathcal{V}} Y, s\colon Y \nrightarrow_{\mathcal{V}} Z$ is given by

$$(s \cdot r)(x,z) = \bigwedge\nolimits_{y \in Y} r(x,y) \oplus s(y,z).$$

Universal coalgebra. Given an endofunctor $F\colon \mathbf{Set} \to \mathbf{Set}$, an *$F$-coalgebra* is a pair (X, ξ), where X is a set of *states* and $\xi\colon X \to FX$ is the *transition* map. Coalgebras generalize the notion of state-based systems: The transition map defines how individual states can transition between each other, with the transition dynamic (nondeterministic, probabilistic, etc.) being specified by the endofunctor F. For instance, coalgebras for the *subdistribution functor* $\mathcal{S}\colon \mathbf{Set} \to \mathbf{Set}$, where $\mathcal{S}X = \{\mu\colon X \to \mathcal{V} \mid \{x \mid \mu(x) \neq 0\} \text{ finite}, \sum_{x \in X} \mu(x) \le 1\}$ and $\mathcal{S}f(\mu)(y) = \sum_{x \in f^{-1}(y)} \mu(x)$, are Markov chains that optionally terminate with some probability in each state. The *distribution functor* is the subfunctor $\mathcal{D}$ of $\mathcal{S}$ where we additionally require that $\sum_{x \in X} \mu(x) = 1$.

A functor element $a \in FX$ is *finitely supported* if there exist a finite subset $X_0 \subseteq X$ and an element $a_0 \in FX_0$ such that $Fi(a_0) = a$ where i denotes the inclusion map $X_0 \hookrightarrow X$. We may identify FX_0 with a subset of FX [6], and thus just write $a \in FX_0$ in this case. A coalgebra (X, ξ) is *finitely branching* if for each $x \in X$, $\xi(x)$ is finitely supported. A functor F is *finitary* if all its elements are finitely supported, so all its coalgebras are finitely branching. Both $\mathcal{S}$ and $\mathcal{D}$ are finitary, as we restrict to finitely supported (sub-)distributions. A *homomorphism* between two F-coalgebras $(X,\xi), (Y,\zeta)$ is a function $h\colon X \to Y$ such that $\zeta \cdot h = Fh \cdot \xi$. States $x \in X, y \in Y$ in coalgebras $(X,\xi), (Y,\zeta)$ are *behaviourally equivalent* if there is a third coalgebra (Z,δ) and coalgebra homomorphisms $g\colon (X,\xi) \to (Z,\delta)$ and $h\colon (Y,\zeta) \to (Z,\delta)$ such that $g(x) = h(y)$; this captures standard notions of (e.g., probabilistic) bisimilarity [1, 7, 25].

3 Threshold-Based Behavioural Distances

We proceed to introduce the central notion of *threshold-based behavioural distances*. These distances generalize a distance on (labelled) Markov chains that has been termed *ε-distance* [13, 14]; we will see in Section 7 that our general notion in fact subsumes many other well-known behavioural distances on various system types. As indicated above, ε-distance is based on first introducing a notion of ε-(bi-)simulation in which probabilities under successor distributions are allowed to deviate by at most ε, and then defining the induced (symmetric or asymmetric) distance between states x, y as the infimum over all ε such that x, y are ε-(bi-)similar. Our generalization works as follows.

We parametrize the overall framework over a set functor F encapsulating the system type in a coalgebraic modelling as recalled in Section 2. In this setting, both existing and our new notions of behavioural distances and coalgebraic modal logics are closely tied to predicate liftings, which provide a means of lifting predicates on the state space X of a coalgebra to the space FX or successor structures. We work with both two-valued and quantitative predicate liftings, as well as with a new type of heterogeneous predicate lifting:

▶ **Definition 3.1.** A 2*-to-*$\mathcal{V}$ (respectively 2*-to-*2, $\mathcal{V}$*-to-*$\mathcal{V}$) *predicate lifting* for F is a collection of maps $\lambda_X : 2^X \to \mathcal{V}^{FX}$ (respectively $\lambda_X : 2^X \to 2^{FX}$, $\lambda_X : \mathcal{V}^X \to \mathcal{V}^{FX}$) subject to the *naturality* condition $\lambda_X(g \cdot f) = \lambda_Y(g) \cdot Ff$ for all $f : X \to Y$ and all predicates g of the appropriate type ($g \in 2^Y$ or $g \in \mathcal{V}^Y$).

The naturality condition equivalently states that predicate liftings are natural transformations involving the contravariant two-valued and $\mathcal{V}$-valued powerset functors 2^- and $\mathcal{V}^-$. Heterogeneous predicate liftings of type 2-to-$\mathcal{V}$ can interact both with two-valued and quantitative concepts by respectively casting them as predicate liftings of type 2-to-2 and $\mathcal{V}$-to-$\mathcal{V}$. For the former this is achieved by introducing a threshold value and then treating values above that threshold as $\top$ and values below it as $\bot$, while for the latter this is achieved through a process called Sugeno integration (see Section 6). These ideas respectively form the technical cores of Sections 5 and 6.

▶ **Definition 3.2.** A 2-to-$\mathcal{V}$ predicate lifting λ is *monotone* if whenever $A, B \in 2^X$ and $A \subseteq B$, then $\lambda_X(A)(a) \leq \lambda_X(B)(a)$ for all $a \in FX$. The *dual* $\overline{\lambda}$ of λ is given by $\overline{\lambda}_X(A)(a) = 1 - \lambda_X(X \setminus A)(a)$ for $A \in 2^X$, $a \in FX$. Given a set Λ of 2-to-$\mathcal{V}$ predicate liftings, we write $\overline{\Lambda}$ for the closure of Λ under duals, i.e. $\overline{\Lambda} = \Lambda \cup \{\overline{\lambda} \mid \lambda \in \Lambda\}$.

We will generally use the terms *predicate lifting* and *modality* interchangeably. We restrict to unary liftings purely in the interest of readability; the treatment of higher arities, i.e. predicate liftings of type $(2^-)^n \to \mathcal{V}^{F-}$, requires no more than additional indexing. One basic example of a monotone 2-to-$\mathcal{V}$ predicate lifting is *probability*, given by

$$P(A)(\mu) = \mu(A) \quad \text{for } A \in 2^X,\ \mu \in \mathcal{S}X.$$

In a Markov chain $\xi : X \to \mathcal{S}X$, $P(A)(\xi(x))$ gives the probability of transitioning from x into the set $A \in 2^X$, while $\overline{P}(A)(\xi(x))$ is the probability of either transitioning into A or terminating.

▶ **Convention 3.3.** For the remainder of the paper, we fix a functor F and a set Λ of monotone 2-to-$\mathcal{V}$ predicate liftings.

From the given choice of 2-to-$\mathcal{V}$ predicate liftings, we obtain a notion of (asymmetric) behavioural distance as follows:

▶ **Definition 3.4** (ε-(Bi-)simulation, threshold-based behavioural distance)**.** Let $\xi : X \to FX$, $\zeta : Y \to FY$ be F-coalgebras, and let $\varepsilon \in \mathcal{V}$. A (two-valued) relation $R : X \nrightarrow Y$ is an ε-Λ-*simulation* (from (X, ξ) to (Y, ζ)) if whenever $x \mathrel{R} y$, then

$$\lambda(R[A])(\zeta(y)) \geq \lambda(A)(\xi(x)) - \varepsilon \quad \text{for all } A \in 2^X,\ \lambda \in \Lambda. \tag{1}$$

A relation R as above is an ε-Λ-*bisimulation* if both R and its converse R° are ε-Λ-simulations. If states $x \in X$, $y \in Y$ are related by some ε-Λ-(bi-)simulation, then we say that y ε-Λ-*simulates* x or that x and y are ε-Λ-*bisimilar*, and write $x \preceq_{\varepsilon,\Lambda} y$ or $x \approx_{\varepsilon,\Lambda} y$, respectively, defining $\preceq_{\varepsilon,\Lambda}$ and $\approx_{\varepsilon,\Lambda}$ as relations between X and Y. We define the induced *(threshold-based)* Λ-*behavioural distance* d_Λ between states $x \in X$, $y \in Y$ by

$$d_\Lambda(x, y) = \bigwedge\{\varepsilon \mid x \preceq_{\varepsilon,\Lambda} y\}.$$

▶ Remark 3.5. Like in the two-valued case [23] and in the purely quantitative setting [41], it is easy to see that if Λ is closed under duals (i.e. $\Lambda = \overline{\Lambda}$), then every ε-Λ-simulation R is in fact an ε-Λ-bisimulation (in the notation of Definition 3.4, to see that R° is a Λ-simulation, show $\lambda(R^\circ[B])(\xi(x)) \geq \lambda(B)(\zeta(y)) - \varepsilon$ for $B \in 2^Y$ and $x\ R\ y$ by applying (1) to $\overline{\lambda}$ and $A = X \setminus R^\circ[B]$). It follows that in this case, the threshold-based distance d_Λ is symmetric, hence a pseudometric. We thus focus almost entirely on ε-Λ-simulations in the main line of the technical development, as we obtain corresponding results on bisimulations by just closing under duals.

▶ **Example 3.6.** The notion of ε-Λ-simulation generalizes Desharnais et al.'s ε-simulations on labelled Markov chains [13]. For the moment, we restrict to the single-action case, in which we just speak of Markov chains (with optional termination), and treat the general case in Example 7.1. As indicated in Section 2, a Markov chain is a coalgebra for the subdistribution functor $\mathcal{S}$. For $\Lambda = \{P\}$, ε-Λ-simulations on Markov chains are precisely ε-simulations in the sense of Desharnais et al. Explicitly, a relation $R\colon X \nrightarrow Y$ is an ε-Λ-simulation between Markov chains $\xi\colon X \to \mathcal{S}X$, $\zeta\colon Y \to \mathcal{S}Y$ if whenever $x\ R\ y$, then $\zeta(y)(R[A]) \geq \xi(x)(A) - \varepsilon$ for all $A \in 2^X$. We use this case as a running example, deferring the presentation of further instances to Section 7. By Remark 3.5, ε-$\overline{\Lambda}$-simulations are precisely ε-bisimulations in the sense of Desharnais et al. (again restricted to the case of a single action for the moment, with the general case covered in Example 7.1).

We note some basic properties of ε-Λ-similarity:

▶ **Lemma 3.7.** *Let $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ be F-coalgebras.*
1. *The relation $\preceq_{\varepsilon,\Lambda}$ is the greatest ε-Λ-simulation from (X, ξ) to (Y, ζ).*
2. *Threshold-quantitative Λ-behavioural distance is a (class-sized) hemimetric.*
3. *Let $\delta \geq \varepsilon \geq 0$, $x \in X$, $y \in Y$. If $x \preceq_{\varepsilon,\Lambda} y$, then $x \preceq_{\delta,\Lambda} y$.*
4. *Let $x \in X$, $y \in Y$ such that $d_\Lambda(x, y) < \varepsilon$. Then $x \preceq_{\varepsilon,\Lambda} y$.*
5. *Suppose that (Y, ζ) is finitely branching, and let $\varepsilon \geq 0$, $x \in X$, $y \in X$. If $x \preceq_{\varepsilon',\Lambda} y$ for every $\varepsilon' > \varepsilon$, then $x \preceq_{\varepsilon,\Lambda} y$. In particular, $d_\Lambda(x, y) \leq \varepsilon$ iff $x \preceq_{\varepsilon,\Lambda} y$.*

Without finite branching, Lemma 3.7.5 does not hold in general. Details are in the full version.

Relators and Lax Extensions. We capture the notion of ε-Λ-similarity more concisely using (two-valued) *relators* over F, which we just understand as assignments L mapping relations $R\colon X \nrightarrow Y$ to relations $LR\colon FX \nrightarrow FY$, subject only to monotonicity ($R \subseteq S$ implies $LR \subseteq LS$), following usage in earlier work [38, 32] (note that the term is sometimes used with a stricter meaning [4, 30]). In this terminology, we have relators $L_{\varepsilon,\Lambda}$ over F for all $\varepsilon \geq 0$, given by

$$a\ L_{\varepsilon,\Lambda}R\ b \iff \forall \lambda \in \Lambda, A \subseteq X.\, \lambda(R[A])(b) \geq \lambda(A)(a) - \varepsilon$$

for $R \subseteq X \times Y$, $a \in FX$, $b \in FY$. Then given F-coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$, a relation $R\colon X \nrightarrow Y$ is an ε-Λ-simulation iff R is an $L_{\varepsilon,\Lambda}$-*simulation*, i.e. $\xi(x)\ L_{\varepsilon,\Lambda}R\ \zeta(y)$ for all $(x,y) \in R$. We will see that this family of relators gives rise to a quantitative lax extension of F, in the following sense:

▶ **Definition 3.8.** A mapping L of $\mathcal{V}$-valued relations $r\colon X \nrightarrow_{\mathcal{V}} Y$ to $\mathcal{V}$-valued relations $Lr\colon FX \nrightarrow_{\mathcal{V}} FY$ is a (quantitative) *lax extension* of F if for all $\mathcal{V}$-valued relations $r, r'\colon X \nrightarrow_{\mathcal{V}} Y$, $s\colon Y \nrightarrow_{\mathcal{V}} Z$ and every function $f\colon X \to Y$, we have

1. $r \leq r' \implies Lr \leq Lr'$ (*monotonicity*)
2. $L(s \cdot r) \leq Ls \cdot Lr$ (*lax functoriality*)
3. $Lf \leq Ff$ and $L(f^\circ) \leq (Ff)^\circ$ (*extension*)

where we treat f as a $\mathcal{V}$-valued relation with $f(x,y) = 0$ if $f(x) = y$ and $f(x,y) = 1$ otherwise.

Using the family $(L_{\varepsilon,\Lambda})_{\varepsilon\in\mathcal{V}}$ of relators, we define a lax extension $\mathbf{L}_\Lambda$ by

$$\mathbf{L}_\Lambda r(a,b) = \bigwedge\{\varepsilon \in \mathcal{V} \mid a\ L_{\varepsilon,\Lambda} r_\varepsilon\ b\} \tag{2}$$

for $r\colon X \nrightarrow_{\mathcal{V}} Y$, $a \in FX$, $b \in FY$, where $r_\varepsilon = \{(x,y) \in X \times Y \mid r(x,y) \leq \varepsilon\}$. We defer the proof that $\mathbf{L}_\Lambda$ is in fact a lax extension – we show later that $\mathbf{L}_\Lambda$ is induced by a set of quantitative modalities via the so-called Kantorovich construction (Theorem 6.13), which by general results always yields a lax extension [41]. We frequently consider cases where Λ consists only of a single predicate lifting λ; in this case, we write $L_{\varepsilon,\lambda}$ and L_λ in place of $L_{\varepsilon,\{\lambda\}}$ and $L_{\{\lambda\}}$, respectively. Lax extensions restrict to liftings of pseudometrics (hemimetrics); that is, whenever d is a pseudometric (hemimetric), then Ld is also a pseudometric (hemimetric) [41].

▶ Remark 3.9. Every 2-to-2 predicate lifting induces a 2-to-$\mathcal{V}$ predicate lifting by postcomposing with the inclusion of the sets 2^{FX} into the sets $\mathcal{V}^{FX}$; we say that a 2-to-$\mathcal{V}$ predicate lifting λ is *two-valued* if it is induced in this way, and we then also use λ as a 2-to-2 predicate lifting. If all predicate liftings in Λ are two-valued, then the relators $L_{\varepsilon,\Lambda}$ all coincide for $\varepsilon < 1$ (indeed, they coincide with the two-valued lax extension induced by the corresponding 2-to-2 predicate liftings [31]). Nevertheless, $\mathbf{L}_\Lambda$ is typically properly quantitative, and in fact often coincides with familiar quantitative lax extensions. For instance, in case $F = \mathcal{P}$ and $\Lambda = \{\Diamond\}$ where $\Diamond\colon 2^- \to 2^{\mathcal{P}-}$ is the standard 2-valued diamond modality given by $\Diamond_X(A) = \{S \in \mathcal{P}X \mid A \cap S \neq \emptyset\}$ for $A \in 2^X$, $\mathbf{L}_\Lambda$ is precisely the one-sided Hausdorff lax extension given by $\mathbf{L}_\Lambda r(S,T) = \bigvee_{x\in S}\bigwedge_{y\in T} r(x,y)$ for $S \in \mathcal{P}X$, $T \in \mathcal{P}Y$, and $r\colon X \nrightarrow_{\mathcal{V}} Y$. By Remark 3.5, the closure $\overline{\Lambda}$ of Λ under duals, $\overline{\Lambda} = \{\Diamond, \Box\}$, induces the symmetric Hausdorff lax extension, which, when applied to a metric on X, yields the Hausdorff pseudometric on $\mathcal{P}X$, given by $L_{\overline{\Lambda}} r(S,T) = \mathbf{L}_\Lambda r(S,T) \vee \mathbf{L}_\Lambda r^\circ(T,S) = \bigvee_{x\in S}\bigwedge_{y\in T} r(x,y) \vee \bigvee_{y\in T}\bigwedge_{x\in S} r(x,y)$ for data as above.

▶ **Example 3.10.** Continuing Example 3.6, i.e. taking $F = \mathcal{S}$ and $\Lambda = \{P\}$, we describe $\mathbf{L}_P$ as a one-sided form of the Lévy-Prokhorov lifting [14], that is we have

$$\mathbf{L}_P r(\mu,\nu) = \bigwedge\{\varepsilon \mid \forall A \in 2^X.\, \nu(r_\varepsilon[A]) \geq \mu(A) - \varepsilon\}$$

for $\mu \in \mathcal{S}X$, $\nu \in \mathcal{S}Y$, and $r\colon X \nrightarrow_{\mathcal{V}} Y$. Correspondingly, $\mathbf{L}_{\overline{\Lambda}}$ is the usual two-sided Lévy-Prokhorov lifting, i.e.

$$\mathbf{L}_{\overline{\Lambda}} r(\mu,\nu) = \bigwedge\{\varepsilon \mid \forall A \in 2^X.\, \nu(r_\varepsilon[A]) \geq \mu(A) - \varepsilon \wedge \forall B \in 2^Y.\mu((r^\circ)_\varepsilon[B]) \geq \nu(B) - \varepsilon\}.$$

We note here that over probability distributions, i.e. subdistributions with total weight 1, P and $\overline{P}$ coincide, so that already $\mathbf{L}_P$ defines the two-sided Lévy-Prokhorov lifting.

Our version of the Lévy-Prokhorov lifting differs from that due to Desharnais and Sokolova [14] in two regards. First, their version acts on pseudometrics, while ours acts more generally on $\mathcal{V}$-valued relations and hence in particular needs to use the converse r°. Second, their version is defined with open instead of closed neighbourhoods. This choice does not make a difference, as we show in the full version.

Quantitative lax extensions yield notions of quantitative simulations for F-coalgebras: If L is a quantitative lax extension, and (X,ξ) and (Y,ζ) are F-coalgebras, then a $\mathcal{V}$-valued relation $s\colon X \nrightarrow_{\mathcal{V}} Y$ is an *L-simulation* if $s \geq \zeta^\circ \cdot Ls \cdot \xi$, i.e. if

$$s(x,y) \geq Ls(\xi(x),\zeta(y)) \quad \text{for all } x \in X \text{ and } y \in Y.$$

As L is monotone, by the Knaster-Tarski fixed point theorem it follows that there is a smallest L-simulation between (X,ξ) and (Y,ζ), which we call *L-behavioural distance* (from (X,ξ) to (Y,ζ)), and often denote by d_L.

We show next that $\mathbf{L}_\Lambda$-behavioural distance in this sense coincides with Λ-behavioural distance in the sense of Definition 3.4 (generalizing recent results for the case of labelled Markov chains [14]), so that $\mathbf{L}_\Lambda$ captures ε-Λ-similarity:

▶ **Theorem 3.11.** *Let (X,ξ) and (Y,ζ) be F-coalgebras. For states $x \in X$, $y \in Y$, we have*

$$d_\Lambda(x,y) = d_{\mathbf{L}_\Lambda}(x,y).$$

▶ **Example 3.12.** As already indicated above, by Example 3.10, one instance of Theorem 3.11 yields the characterization of ε-distance on labelled Markov chains [13] via fixpoints of the Lévy-Prokhorov lifting established recently by Desharnais and Sokolova [14] (restricted for the moment to the case with only one label, with the general case discussed in Example 7.1); indeed we obtain this both for the two-sided (symmetric) and for the one-sided (asymmetric) variant, respectively applying Theorem 3.11 to $\{P,\overline{P}\}$ or to $\{P\}$. Correspondingly, we occasionally refer to this behavioural distance as *Lévy-Prokhorov behavioural distance*, adding the qualification *one-sided* or *two-sided* when needed.

4 Threshold-Quantitative Codensity Games

With a view to using Spoiler strategies in the computation of distinguishing formulae, we next introduce a Spoiler-Duplicator game for ε-similarity.

▶ **Definition 4.1.** Let Λ be a set of 2-to-$\mathcal{V}$ predicate liftings, and let $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ be F-coalgebras. The *codensity game up to ε* on ξ and ζ, played by Spoiler (S) and Duplicator (D), is given by the following table detailing positions, ownership, and moves:

Position	Owner	Moves
$(x,y) \in X \times Y$	S	$\{(\lambda,A,B) \in \Lambda \times 2^X \times 2^Y \mid \lambda(B)(\zeta(y)) < \lambda(A)(\xi(x)) - \varepsilon\}$
$(\lambda,A,B) \in \Lambda \times 2^X \times 2^Y$	D	$\{(x,y) \in X \times Y \mid x \in A, y \notin B\}$

As usual, any player who cannot move on their turn loses. Infinite plays are won by D.

Intuitively, we understand a position of the form (x,y) as D claiming that x and y are ε-Λ-similar, i.e. $x \preceq_{\varepsilon,\Lambda} y$. Spoiler can challenge such a claim by exhibiting predicates A and B such that $\lambda(B)(\zeta(y)) < \lambda(A)(\xi(x)) - \varepsilon$, which refutes D's claim provided that $\preceq^\varepsilon_{\varepsilon,\Lambda}[A] \subseteq B$ (here, $\preceq^\varepsilon_{\varepsilon,\Lambda}[A] \subseteq Y$ is the set of states that ε-Λ-simulate some state in A). Thus, D needs to challenge this inclusion by playing (x',y') where $x' \in A$ and $y' \notin B$, thereby claiming that $x' \preceq_{\varepsilon,\Lambda} y'$, i.e. that $y' \in \preceq_{\varepsilon,\Lambda}[A]$. The game is illustrated in Example 6.10 below.

▶ **Theorem 4.2.** *Let $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ be F-coalgebras. Duplicator wins the position (x, y) in the codensity game up to ε on ξ and ζ iff y ε-Λ-simulates x.*

In connection with Lemma 3.7.5, we obtain

▶ **Corollary 4.3** (Game correctness). *Let Λ be a set of 2-to-$\mathcal{V}$ predicate liftings, and let $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ be finitely branching F-coalgebras. Then Duplicator wins the position (x, y) in the codensity game up to ε on ξ and ζ iff $d_\Lambda(x, y) \leq \varepsilon$.*

▶ Remark 4.4. We expect that the codensity game up to ε can be cast as an instance of the very general notion of codensity games on fibrations [26], in this case on the fibration of binary relations over the base category $\mathbf{Set}^2$ of pairs of sets (that is, the fibre over a pair (X, Y) of sets consists of the relations of type $R\colon X \nrightarrow Y$). Desharnais et al. [13] present a game for the specific case of Lévy-Prokhorov distance, which however is quite different in spirit from the codensity game up to ε.

As we aim to use Spoiler strategies to extract distinguishing formulae, we are interested in computing them. Since the codensity game is a safety game, Spoiler's winning region is a least fixed point, and thus can be computed by Kleene iteration. The time estimate for this computation necessarily depends on computational properties of the functor, whose formulation requires a suitable finite representation of modalities $\lambda \in \Lambda$ and functor elements $a \in FX$ for finite X (e.g. [24]). Such representations are obvious in all of our main examples. In general, not all elements of FX will be representable; in particular, we regard only such elements of $\mathcal{S}X$ as representable that have only rational probabilities. A coalgebra $\xi\colon X \to FX$ is *representable* if $\xi(x)$ is representable for every $x \in X$. We generally assume that numbers are coded in binary. The requisite condition is then quite simple:

▶ **Definition 4.5.** A set Λ of 2-to-$\mathcal{V}$ predicate liftings is *polynomial-time solvable* if given $0 \leq \varepsilon \in \mathbb{Q}$, representable functor elements $a \in FX$, $b \in FY$ with X, Y finite, and a relation $S\colon X \nrightarrow Y$, one can decide in polynomial time (in the total size of the input a, b, S, ε) whether there exist $A \in 2^X$, $B \in 2^Y$, $\lambda \in \Lambda$ such that $\lambda(B)(b) < \lambda(A)(a) - \varepsilon$ and $A \times (Y \setminus B) \subseteq S$, and then compute these data in polynomial time.

▶ **Example 4.6.** One typically establishes polynomial-time solvability by forming the complement relation $R = (X \times Y) \setminus S$, and taking B to be $R[A]$. This ensures the condition $A \times (Y \setminus B) \subseteq S$, so that only the inequality needs to be checked. For instance, polynomial-time solvability of $\Lambda = \{P\}$ follows in this way from a construction by Desharnais et al. [13], who, given $\mu \in \mathcal{S}X$ and $\nu \in \mathcal{S}Y$, construct a flow network $\mathcal{N}(\mu, \nu, R)$ whose maximal flow is at least $\mu(X) - \varepsilon$ iff for all $A \in 2^X$, $P(R[A])(\nu) \geq P(A)(\mu) - \varepsilon$, i.e. $\nu(R[A]) \geq \mu(A) - \varepsilon$. In case the maximal flow (which can be computed in polynomial time by standard methods) is less than $\mu(X) - \varepsilon$, one obtains $A \in 2^X$, $B \in 2^Y$ such that $\nu(B) < \mu(A) - \varepsilon$ and $A \times (Y \setminus B) \subseteq S$ from a minimum cut of $\mathcal{N}(\mu, \nu, R)$ (which in turn is read off from a maximal flow in a standard manner); details are in the full version. Modifications for the two-sided case ($\Lambda = \{P, \overline{P}\}$) are straightforward.

▶ **Theorem 4.7** (Strategy computation). *Let $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ be representable finite coalgebras, and let $0 \leq \varepsilon \in \mathbb{Q}$. If Λ is polynomial-time solvable, then both the winning region of Spoiler in the codensity game up to ε on ξ and ζ and a history-free winning strategy on the winning region are computable in polynomial time.*

By the correctness of the game (Corollary 4.3), this entails:

▶ **Corollary 4.8.** *If Λ is polynomially solvable, then it is decidable in polynomial time whether two given states in representable finite coalgebras have distance at most ε.*

5 Two-Valued Distinguishing Formulae

We now introduce a two-valued coalgebraic modal logic equipped with a notion of satisfaction up to ε, with a view to certifying high behavioural distances by modal formulae (a quantitative logic for similar purposes is discussed in Section 6). To this end, we convert the given 2-to-$\mathcal{V}$ predicate liftings into 2-to-2 predicate lifting by means of threshold values; as a simple example, for the probability modality $\Lambda = \{P\}$ over subdistributions, one obtains the usual threshold modalities 'with probability at least q' known from two-valued probabilistic modal logics [29]. Satisfaction up to ε then means essentially that ε is subtracted from all thresholds. Generalizing Desharnais et al.'s logic for labelled Markov chains [13], we define the two-valued logic $\mathcal{L}_2(\Lambda)$ as follows. The set $\mathcal{F}_2(\Lambda)$ of *(two-valued) Λ-formulae* $\varphi, \psi, \dots$ is given by the grammar

$$\varphi, \psi ::= \bot \mid \top \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \lambda_q \varphi \qquad (\lambda \in \Lambda, q \in \mathcal{V}).$$

The *modal rank* of a formula φ is the maximal nesting depth of modalities λ_q in φ.

We interpret this logic in terms of satisfaction relations up-to-ε, denoted $\models_\varepsilon$, between states x in F-coalgebras $\xi\colon X \to FX$ and Λ-formulae. The relation $\models_\varepsilon$ is defined inductively, with the usual clauses for the Boolean connectives and with the modal case given by

$$x \models_\varepsilon \lambda_q \varphi \quad \text{iff} \quad \lambda(\llbracket \varphi \rrbracket_\varepsilon)(\xi(x)) \geq q \ominus \varepsilon$$

where $\llbracket \varphi \rrbracket_\varepsilon = \{z \in X \mid z \models_\varepsilon \varphi\}$. Intuitively, $x \models_\varepsilon \varphi$ means that state x might not satisfy φ exactly, but satisfies it at least up to ε; for $\varepsilon = 0$, we just have classical satisfaction.

► **Definition 5.1.** Given states $x \in X$, $y \in Y$ in F-coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$, respectively, we say that x is *logically included in y up to ε* if for all $\varphi \in \mathcal{F}_2(\Lambda)$, whenever $x \models_0 \varphi$, then $y \models_\varepsilon \varphi$.

(One can then in fact show that more generally, whenever $x \models_\delta \varphi$ for some $\delta \in \mathcal{V}$, then $y \models_{\delta+\varepsilon} \varphi$; details are in the full version.) Satisfaction is preserved up to behavioural distance (in generalization of the "if" direction of [13, Theorem 3]):

► **Lemma 5.2** (Preservation). *Let $x \in X$, $y \in Y$ be states in F-coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$, and let $\varepsilon \in \mathcal{V}$. If $d_\Lambda(x, y) \leq \varepsilon$, then x is logically included in y up to ε.*

Indeed, we can show that over finitely branching coalgebras, the threshold-based Λ-behavioural distance d_Λ is characterized by the logic $\mathcal{L}(\Lambda)$, a property often referred to as *expressiveness*:

► **Definition 5.3.** Let $x \in X$, $y \in Y$ be states in F-coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$, respectively. A formula $\varphi \in \mathcal{F}_2(\Lambda)$ is *ε-distinguishing* for (x, y) if $x \models_0 \varphi$ and $y \not\models_\varepsilon \varphi$.

► **Theorem 5.4** (Expressiveness, two-valued). *Let $x \in X$, $y \in Y$ be states in finitely branching coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$. Then $d_\Lambda(x, y) \leq \varepsilon$ iff x is logically included in y up to ε.*

That is, whenever $d_\Lambda(x, y) > \varepsilon$, then there exists an ε-distinguishing formula for (x, y) in $\mathcal{L}(\Lambda)$. We next focus on the computation of distinguishing formulae in case the involved coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ are finite. Let s be a memoryless winning strategy for Spoiler that wins all positions in the winning region of Spoiler in the codensity game up to ε on ξ and ζ as per Definition 4.1 (such a strategy exists because from Spoiler's perspective, the game is a reachability game). Since there are only $k = |X| \times |Y|$ Spoiler positions in the game, Spoiler wins any position in their winning region in at most k moves. We compute ε-distinguishing formulae for all pairs (x, y) of states in Spoiler's winning region in dag representation, i.e. sharing identical subformulae (cf. Remark 5.6), using dynamic programming, with stages indexed by the remaining number of Spoiler moves. As positions of the shape $(x, y) \in X \times Y$ belong to Spoiler, this number is always greater than 0. In stage $i > 0$, the algorithm proceeds as follows for each position (x_0, y_0) won by S in i moves:

▶ **Algorithm 1** (Extraction of two-valued distinguishing formulae)**.**

1. Put $(\lambda, A, B) = s(x_0, y_0)$ (Spoiler's winning move). By definition of the game, $\lambda(B)(\zeta(y)) < q - \varepsilon$ where $q = \lambda(A)(\xi(x))$.
2. For every possible reply by D, i.e. every pair $(x, y) \in X \times Y$ such that $x \in A$ and $y \notin B$, S wins (x, y) in at most $i - 1$ moves, so we have already computed an ε-distinguishing formula φ_{xy} for (x, y) in a previous stage.
3. Put

$$\varphi = \bigvee_{x \in A} \bigwedge_{y \in Y \setminus B} \varphi_{xy}. \tag{3}$$

The ε-distinguishing formula for (x_0, y_0) is $\varphi_{x_0 y_0} = \lambda_q \varphi$.

The algorithm is illustrated in Example 6.10 below. In the interest of abstractness, we have refrained from assuming representability of the input, which, of course, is a prerequisite for the actual execution of the algorithm; we discuss this point in Remark 5.7. Nevertheless, correctness of the algorithm and a size estimate for the constructed formulae can formulated already for the abstract version.

▶ **Theorem 5.5** (Extraction of polynomial-sized two-valued distinguishing formulae)**.** *Given finite F-coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ and a Spoiler strategy s for the codensity game up to ε on ξ and ζ that wins all positions in Spoiler's winning region, Algorithm 1 constructs two-valued ε-distinguishing formulae $\varphi_{xy} \in \mathcal{F}_2(\Lambda)$ for all pairs of states $(x, y) \in X \times Y$ such that $d_\Lambda(x, y) > \varepsilon$. The formulae $\varphi_{x,y}$ are of quadratic modal rank and polynomial, specifically quartic, dag size in $|X| + |Y|$.*

(For the quartic bound, just note that the algorithm constructs quadratically many formulae, each of which contributes quadratically many logical connectives.)

▶ Remark 5.6 (Dag representation). As indicated above, we measure formulae in *dag size*, i.e. the size of a representation where identical subformulae are shared, as opposed to the *tree size* measured without sharing. Even in the two-valued base case, i.e. in the standard modal logic of Kripke frames, the tree size of distinguishing formulae for non-bisimilar states is worst-case exponential [17], and existing algorithms computing distinguishing formulae for two-valued behavioural equivalences [11, 44] work with dag representations.

▶ Remark 5.7 (Constants, complexity). In the interest of succinctness, we have allowed real-valued thresholds q on modalities λ_q, following Desharnais et al. [13]. To restrict to rational thresholds, and to ensure executability, call a modality λ *polynomially rational* if for $A \in 2^X$ and representable $t \in FX$, $\lambda(A)(t)$ is rational and can be computed in polynomial time in t, in particular is of polynomial size. For instance, our running example, the modality P over Markov chains, is polynomially rational as it just sums over probabilities already given in the input distribution. It is apparent from Algorithm 1 that in case the input coalgebras are representable and all $\lambda \in \Lambda$ are polynomially rational, the extracted formulae contain only polynomial-sized rational thresholds. Moreover, the algorithm then becomes executable, and runs in polynomial time.

In combination with Theorem 4.7, we obtain:

▶ **Corollary 5.8** (Polynomial-time computation of two-valued distinguishing formulae)**.** *If Λ is polynomial-time solvable and all $\lambda \in \Lambda$ are polynomially rational, then ε-distinguishing formulae of states x, y of threshold-behavioural distance $d_\Lambda(x, y) > \varepsilon$ in representable finite coalgebras can be computed in polynomial time in the overall size of the input.*

▶ **Example 5.9.** Continuing Example 3.6, we obtain Desharnais et al.'s expressiveness result for ε-logic w.r.t. ε-simulation [13], restricted to finitely branching (and, for the moment, single-action) Markov chains, as an instance of Theorem 5.4 (we leave a generalization of Theorem 5.4 to countable branching as covered by Desharnais et al. to future work). As a new result, we obtain a polynomial size estimate for distinguishing formulae for ε-(bi-)simulation on finite Markov chains, and polynomial-time computation of such formulae on representable Markov chains. Additional discussion is in Example 7.1.

6 Quantitative Distinguishing Formulae

Having seen a characteristic two-valued modal logic for threshold-based behavioural distance d_Λ, we next develop a properly quantitative modal logic from the given 2-to-$\mathcal{V}$ predicate liftings. This logic will also be shown to be characteristic for d_Λ, now in the sense that behavioural distance equals the supremum over all deviations in truth values of quantitative modal formulae.

To this end, we now convert the given 2-to-$\mathcal{V}$ predicate liftings $\lambda \in \Lambda$ into $\mathcal{V}$-to-$\mathcal{V}$ predicate liftings S^λ (with components $\mathsf{S}^\lambda_X : \mathcal{V}^X \to \mathcal{V}^{FX}$) – the *Sugeno modalities* – by putting

$$\mathsf{S}^\lambda_X(f)(a) = \bigvee\nolimits_{\varepsilon \in \mathcal{V}} \varepsilon \wedge \lambda(f_\varepsilon)(a) \quad \text{for } f \in \mathcal{V}^X,\ a \in FX \tag{4}$$

where the 2-valued predicate $f_\varepsilon \in 2^X$ is given by $f_\varepsilon(x) = \top$ iff $f(x) \geq \varepsilon$. We put $\mathsf{S}[\Lambda] = \{\mathsf{S}^\lambda \mid \lambda \in \Lambda\}$. We combine these modalities with a propositional base used standardly in quantitative modal logics characterizing behavioural distances, which besides minimum and maximum operators notably includes constant shift operators [39, 42, 27, 41]. This propositional base is designed to ensure non-expansiveness of formula evaluation, and is correspondingly sometimes termed the *non-expansive propositional base* [21]. We define the set $\mathcal{F}_\mathcal{V}(\mathsf{S}[\Lambda])$ of *(quantitative)* Λ*-formulae* φ, ψ by the grammar

$$\mathcal{F}_\mathcal{V}(\mathsf{S}[\Lambda]) \ni \varphi, \psi ::= \top \mid \bot \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \oplus q \mid \varphi \ominus q \mid \mathsf{S}^\lambda \varphi \qquad (q \in \mathcal{V})$$

The *modal rank* of a formula φ is the maximal nesting depth of modalities S^λ in φ.

Given a coalgebra $\xi : X \to FX$, a formula φ is interpreted as a truth map $[\![\varphi]\!] \in \mathcal{V}^X$ given by the standard clauses for the propositional operators ($[\![\top]\!](x) = 1$, $[\![\bot]\!](x) = 0$, $[\![\varphi \wedge \psi]\!](x) = \min\{[\![\varphi]\!](x), [\![\psi]\!](x)\}$, $[\![\varphi \vee \psi]\!](x) = \max\{[\![\varphi]\!](x), [\![\psi]\!](x)\}$, $[\![\varphi \oplus q]\!](x) = [\![\varphi]\!](x) \oplus q$, $[\![\varphi \ominus q]\!](x) = [\![\varphi]\!](x) \ominus q$) and by

$$[\![\mathsf{S}^\lambda \varphi]\!](x) = \mathsf{S}^\lambda_X([\![\varphi]\!])(\xi(x)).$$

We refer to the arising logic $\mathcal{L}(\mathsf{S}[\Lambda])$ as *non-expansive coalgebraic modal logic*. We explicitly do not include negation (which in the quantitative setting is interpreted as $[\![\neg\varphi]\!](x) = 1 \ominus [\![\varphi]\!](x)$), as we also want to cover asymmetric distances. In case Λ is closed under duals, we can encode negation by taking negation normal forms (cf. Remark 6.5).

▶ **Example 6.1.** For $F = \mathcal{S}$ and $\lambda = P$, the Sugeno modality S^P, given by $S^P_X(f)(\mu) = \bigvee_{\varepsilon \in \mathcal{V}} \varepsilon \wedge \mu(f_\varepsilon)$, is precisely the *generally* modality used in probabilistic knowledge representation to model vague properties holding with high probability. As such, it serves as a computationally more tractable alternative to a modality *probably* that coincides with the expectation modality [35].

▶ Remark 6.2. We use the term *Sugeno modalities* and the designation S^λ in honour of the fact that (4) essentially defines S^λ using Sugeno integration [37]. We discuss this connection, as well as a categorical perspective of this construction, in the full version.

We have a number of alternative descriptions of S^λ:

▶ **Lemma 6.3.** *For $f \in \mathcal{V}^X$ and $a \in FX$, we have*

$$\mathsf{S}^\lambda_X(f)(a) = \bigvee\{\varepsilon \in \mathcal{V} \mid \lambda(f_\varepsilon)(a) \geq \varepsilon\} = \bigwedge\{\varepsilon \in \mathcal{V} \mid \lambda(f_\varepsilon)(a) \leq \varepsilon\} = \bigwedge\nolimits_{\varepsilon\in\mathcal{V}} \varepsilon \vee \lambda(f_\varepsilon)(a).$$

▶ Remark 6.4 (Strict inequalities). The values in Lemma 6.3 do not change when one instead uses strict inequalities in the second and third term or in the definition of f_ε.

▶ Remark 6.5. The Sugeno construction preserves duals, in the sense that taking the Sugeno modality of the dual of a 2-to-$\mathcal{V}$ predicate lifting is the same as taking the dual of its Sugeno modality. This is a consequence of Lemma 6.3; details are in the full version. This means that if Λ is closed under duals, then $\mathsf{S}[\Lambda]$ is also closed under duals, and we can therefore recursively encode negation, in particular putting $\neg\mathsf{S}^\lambda(\varphi) = \mathsf{S}^{\overline{\lambda}}(\neg\varphi)$ in the modal case.

▶ **Definition 6.6** (Sugeno logical distance)**.** Let $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ be F-coalgebras. The *(Λ-)-Sugeno logical distance* $d^{\mathsf{log}}{}_{\mathsf{S}[\Lambda]}(x, y)$ from a state $x \in X$ to a state $y \in Y$ is

$$d^{\mathsf{log}}{}_{\mathsf{S}[\Lambda]}(x, y) = \bigvee\nolimits_{\varphi\in\mathcal{F}_\mathcal{V}(\mathsf{S}[\Lambda])} [\![\varphi]\!](x) - [\![\varphi]\!](y).$$

Notice that like threshold-based Λ-behavioural distance, Λ-Sugeno distance is in general only a hemimetric, i.e. may fail to be symmetric; by Remark 6.5, it is symmetric, i.e. a pseudometric, if Λ is closed under duals. From Theorem 6.13 proved later, it follows by general results [27, 40, 18] that the non-expansive coalgebraic modal logic $\mathcal{L}(\mathsf{S}[\Lambda])$ is *expressive*, i.e. that Λ-Sugeno logical distance $d^{\mathsf{log}}{}_{\mathsf{S}[\Lambda]}$ coincides with threshold-based Λ-behavioural distance d_Λ. Still, we give a direct (and simpler) proof for the present case in the full version, obtaining

▶ **Theorem 6.7** (Sugeno expressiveness)**.** *Λ-Sugeno logical distance $d^{\mathsf{log}}{}_{\mathsf{S}[\Lambda]}$ and $\mathbf{L}_\Lambda$-behavioural distance d_Λ coincide on finitely branching F-coalgebras.*

In the following, we focus mainly on the algorithmic construction of distinguishing formulae, which now witness lower bounds for behavioural distance directly by formula evaluation:

▶ **Definition 6.8.** Given F-coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$, a quantitative formula $\varphi \in \mathcal{F}_\mathcal{V}(\mathsf{S}[\Lambda])$ is *ε-distinguishing* for a pair $(x, y) \in X \times Y$ of states if $[\![\varphi]\!](y) < [\![\varphi]\!](x) - \varepsilon$.

By Theorem 6.7, if ξ and ζ are finitely branching, then an ε-distinguishing formula for (x, y) exists whenever $d_\Lambda(x, y) > \varepsilon$. Similar to Section 5, we extract such a formula from a winning strategy for Spoiler in the codensity game. Given the same data as for Algorithm 1, our new algorithm works quite similarly, with main difference being that we no longer use thresholds, and instead use constant shifts in Step 2 and Sugeno modalities in Step 3.

▶ **Algorithm 2** (Extraction of quantitative distinguishing formulae)**.**

1. Take (λ, A, B) to be Spoiler's winning move $s(x_0, y_0)$. By definition of the game, $\lambda(B)(\zeta(y)) < q - \varepsilon$ where $q = \lambda(A)(\xi(x))$.
2. For every possible reply by D, i.e. every pair $(x, y) \in X \times Y$ such that $x \in A$ and $y \notin B$, S wins (x, y) in at most $i - 1$ moves, so we have already computed an ε-distinguishing formula φ_{xy} for (x, y) in previous stages. Using constant shifts, we normalize the $\varphi_{x,y}$ so that $[\![\varphi_{xy}]\!](x) = q$.
3. Put $\varphi = \bigvee_{x\in A} \bigwedge_{y\in Y\setminus B} \varphi_{xy}$. The ε-distinguishing formula for (x_0, y_0) is $\varphi_{x_0y_0} = \mathsf{S}^\lambda\varphi$.

▶ Remark 6.9 (Constants, complexity). Like in the two-valued case, we have admitted real-valued shifts in the grammar but note that the above algorithm uses only rational shifts provided that the input coalgebras are representable and the modalities are polynomially rational (cf. Remark 5.7). Moreover, the algorithm is then executable and runs in polynomial time.

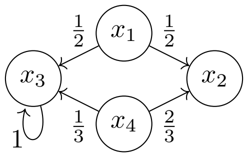

Figure 1 States of distance 1/6 in a Markov chain (states x_1, x_4).

▶ **Example 6.10.** Consider the Markov chain, i.e. coalgebra ξ for the subdistribution functor $\mathcal{S}$, depicted in Figure 1. The Lévy-Prokhorov behavioural distance ($\Lambda = \{P\}$) of x_1, x_4 is $\frac{1}{6}$. Indeed, for $\varepsilon < \frac{1}{6}$, Spoiler wins (x_1, x_4) in the codensity game up to ε (Definition 4.1) by playing $A = \{x_3\}$, $B = \{x_1, x_3, x_4\}$, which satisfies the inequality $\lambda(B)(\xi(x_4)) = \frac{1}{3} < \frac{1}{2} - \varepsilon = \lambda(A)(\xi(x_1)) - \varepsilon$. Now Duplicator must play (x_3, x_2). Spoiler then wins by playing $A = \{x_3\}$, $B = \{x_1, x_2, x_3, x_4\}$, to which Duplicator has no reply. We apply Algorithm 1 to extract a two-valued distinguishing formula: For states x_3, x_2, we construct $\varphi_{3,2} = P_1\top$ (where $\top = \bigvee_{x\in\{x_3\}} \bigwedge_{y\in\emptyset} \varphi_{x,y}$), which then results in $\varphi_{1,4} = P_{\frac{1}{2}} P_1 \top$ for states x_1, x_4. Note that $x_1 \models_0 \varphi_{1,4}$, while $x_4 \not\models_\varepsilon \varphi_{1,4}$. In the quantitative case (Algorithm 2), a shift is performed during the extraction: We first obtain $\psi_{3,2} = \mathsf{S}^P 1$, which witnesses distance 1 for states x_3, x_2. This is shifted by $q = \lambda(A)(\xi(x_1)) = \frac{1}{2}$, so the final formula is $\psi_{1,4} = \mathsf{S}^P(\mathsf{S}^P 1 \ominus \frac{1}{2})$, which gives us $[\![\psi_{1,4}]\!](x_1) = \frac{1}{2}$, $[\![\psi_{1,4}]\!](x_4) = \frac{1}{3}$, witnessing distance $\frac{1}{6} > \varepsilon$.

Again, we can have correctness and a size estimate on distinguishing formulae already for the abstract algorithm:

▶ **Theorem 6.11** (Extraction of polynomial-sized quantitative distinguishing formulae). *Given finite F-coalgebras $\xi\colon X \to FX$, $\zeta\colon Y \to FY$ and a Spoiler strategy s for the codensity game up to ε on ξ and ζ that wins all positions in Spoiler's winning region, Algorithm 2 constructs ε-distinguishing formulae φ_{xy} for all pairs of states $(x, y) \in X \times Y$ such that $d_\Lambda(x, y) > \varepsilon$. The resulting formulae are of quadratic modal rank and polynomial, specifically quartic, dag size in $|X| + |Y|$.*

The proof, found in the full version, is fairly similar to that of Theorem 5.5.

▶ **Corollary 6.12** (Polynomial-time computation of distinguishing quantitative formulae). *If Λ is polynomial-time solvable and all modalities in Λ are polynomially rational, then quantitative ε-distinguishing formulae of states x, y of threshold-behavioural distance $d_\Lambda(x, y) > \varepsilon$ can be computed in polynomial time in the overall input size.*

In particular, as a new result, we can compute quantitative ε-distinguishing formulae for the Lévy-Prokhorov distance on Markov chains in the logic of *generally* in polynomial time.

Kantorovich Characterization of Coalgebraic ε-Bisimulation. As our last main result, we show that the so-called Kantorovich quantitative (sometimes "fuzzy") lax extension [41] induced by the Sugeno modalities coincides with the quantitative lax extension $\mathbf{L}_\Lambda$ as per (2). This closes the open end from Section 3: it follows by general results [41] that $\mathbf{L}_\Lambda$ is indeed a quantitative lax extension. We recall the definition of the Kantorovich quantitative lax extension K_Θ of F induced by a set Θ of $\mathcal{V}$-valued predicate liftings for F. Given a $\mathcal{V}$-valued relation $r\colon X \nrightarrow_{\mathcal{V}} Y$, a pair $(f,g) \in \mathcal{V}^X \times \mathcal{V}^Y$ of predicates is *r-preserved* if for all $(x,y) \in X \times Y$, $g(y) \geq f(x) - r(x,y)$. For $a \in FX$, $b \in FY$, we then put

$$\mathbf{K}_\Theta r(a,b) = \bigvee \{\mu(f)(a) \ominus \mu(g)(b) \mid \mu \in \Theta, (f,g) \ r\text{-preserved}\}.$$

It has been shown that $\mathbf{K}_\Theta$ is always a quantitative lax extension [41] and that the quantitative modal logic induced by Θ is characteristic for the behavioural distance induced by $\mathbf{K}_\Theta$. The following theorem implies that these results apply to the Sugeno modalities and $\mathbf{L}_\Lambda$:

▶ **Theorem 6.13.** *For $\mathbf{L}_\Lambda$ as per (2) and $r\colon X \nrightarrow_{\mathcal{V}} Y$, we have $\mathbf{L}_\Lambda r = \mathbf{K}_{\mathsf{S}[\Lambda]} r$.*

Note here that the $\geq$-direction of the equality $\mathbf{L}_\Lambda r = \mathbf{K}_{\mathsf{S}[\Lambda]} r$ amounts to the Sugeno modalities preserving non-expansive predicates. Applying Theorem 6.13 to our running example of $F = \mathcal{S}$ and $\Lambda = \{P\}$, we obtain a Kantorovich characterization of Lévy-Prokhorov distance via the *generally* modality, reproving a recent specific result to this effect [43]. A general consequence of Theorem 6.13 is that when Λ is finite (as in all our examples), one obtains an instance of the coalgebraic quantitative van Benthem theorem [40], stating roughly that, up to approximation in bounded rank, quantitative coalgebraic modal logic is the behaviourally non-expansive fragment of quantitative coalgebraic first-order predicate logic.

7 Case Studies

We now discuss how our framework applies to a number of example functors and modalities. In all cases, we obtain the characterization of ε-(bi-)simulation distance both via a codensity game (Theorem 4.2) and via a lax extension $\mathbf{L}_\Lambda$ (Theorem 3.11), as well as a representation of the latter in Kantorovich form (Theorem 6.13). All of our examples are also polynomial-time solvable, so that we can compute distinguishing formulae both in a two-valued (Corollary 5.8) and in a quantitative (Corollary 6.12) logic in polynomial time. We thus recover known results (some of them very recent [14, 43]), and moreover establish several new results.

▶ **Example 7.1** (Labelled Markov chains)**.** Extending our running example of Markov chains, we consider the functor $F = \mathcal{S}^{\mathcal{A}}$, where $\mathcal{A}$ is a set of labels. Its coalgebras are *labelled Markov chains* as considered in previous work on Lévy-Prokhorov behavioural distance [13, 14], where in particular the characterization via a lax extension (Theorem 3.11) has already been established [14]. We put $\Lambda = \{\Diamond_\alpha \mid \alpha \in \mathcal{A}\}$, where $\Diamond_\alpha\colon 2^X \to \mathcal{V}^{FX}$ is given by $\Diamond_\alpha(A)(g) = g(\alpha)(A) = P(A)(g(\alpha))$ for $g \in (\mathcal{S}X)^{\mathcal{A}}$, $\alpha \in \mathcal{A}$, $A \in 2^X$, where P is the probability modality on $\mathcal{S}$ from our running example (we write $\Diamond_\alpha$ rather than P_α in the interest of conformity with previous work [13]). We can thus express the Sugeno modalities and the lax extension $\mathbf{L}_\Lambda$ in terms of their single-action counterparts:

$$\mathsf{S}^{\Diamond_\alpha}(f)(g) = \mathsf{S}^P(f)(g(\alpha)) \quad \text{and} \quad \mathbf{L}_\Lambda r(g,h) = \bigvee\nolimits_{\alpha \in \mathcal{A}} \mathbf{L}_P(g(\alpha), h(\alpha))$$

for $g \in (\mathcal{S}X)^{\mathcal{A}}$, $h \in (\mathcal{S}Y)^{\mathcal{A}}$, and $r\colon X \nrightarrow_{\mathcal{V}} Y$. The set Λ is polynomial-time solvable, essentially because we can reduce the computational problem to that for P discussed in Example 4.6. We therefore obtain polynomial-time computation of distinguishing formulae for ε-bisimilarity distance over probabilistic transition systems, both in the logic of the modalities $\Diamond_{\alpha,q}$ (indexed over labels α and thresholds q) and in the logic of the modalities $\mathsf{S}^{\Diamond_\alpha}$. While the existence of distinguishing formulae for the two-valued case is known [13], the computational results are new. We also obtain a Kantorovich characterization of the corresponding behavioural distance, generalizing a recent result for the case $|\mathcal{A}| = 1$ (and $\mathcal{D}$ in place of $\mathcal{S}$) [43].

▶ **Example 7.2** (Fuzzy transition systems). The *fuzzy powerset functor* $\mathcal{P}_\mathcal{V}$ sends a set X to the set of functions $g\colon X \to \mathcal{V}$, with $\mathcal{P}_\mathcal{V} f(g)(y) = \bigvee_{x \in f^{-1}(y)} g(x)$ for $f\colon X \to Y$. Its coalgebras are *fuzzy transition systems*, i.e. transition systems in which adjacency takes on values in the unit interval. We consider the 2-to-$\mathcal{V}$ predicate lifting $\Diamond$ given by $\Diamond(A)(g) = \bigvee_{x \in A} g(x)$ for $A \in 2^X$, $g \in \mathcal{P}_\mathcal{V}(X)$. The corresponding Sugeno modality $\mathsf{S}^\Diamond$ is precisely the *fuzzy diamond* modality $\mathsf{S}^\Diamond(f)(g) = \sup_{x \in X} g(x) \wedge f(x)$ from standard fuzzy modal logic (e.g. [36, 21]), which characterizes bisimulation distance on $\mathcal{V}$-valued relational models [42]. We newly obtain characterizations of this distance through the two-valued logic induced by the threshold modalities $\Diamond_q$ and through the lax extension $\mathbf{L}_\Diamond$. Moreover, $\Lambda = \{\Diamond\}$ is polynomial-time solvable, which means that as another new result we obtain polynomial-time computation of distinguishing formulae in both the two-valued and the quantitative logic of fuzzy diamond; the two-valued logic is in fact new in itself. The key observation for proving polynomial-time solvability is that instead of checking each subset $A \in 2^X$, it suffices to consider singleton sets, of which there are linearly many; details are in the full version.

▶ **Example 7.3** (Metric transition systems). We next consider the functor $F = \mathcal{P}(\mathcal{A} \times -)$, where $\mathcal{P}$ is the powerset functor and the set $\mathcal{A}$ carries a metric $d_\mathcal{A}$. Coalgebras for this functor are *metric transition systems* [12, 16]. We put $\Lambda = \{\Diamond_\alpha \mid \alpha \in \mathcal{A}\}$, where $\Diamond_\alpha$ is the 2-to-$\mathcal{V}$ predicate lifting given by $\Diamond_\alpha(A)(S) = \bigvee_{x \in A, (\beta, x) \in S} 1 - d_\mathcal{A}(\alpha, \beta)$ for $A \in 2^X$, $S \in FX$, saying roughly 'there is a transition into the set A whose label is close to α'. The Sugeno modalities are given by $\mathsf{S}^{\Diamond_\alpha}(f)(S) = \bigvee_{(\beta,x) \in S} (1 - d_\mathcal{A}(\alpha, \beta)) \wedge f(x)$ for $f \in \mathcal{V}^X$, $S \in FX$, and they are precisely the metric diamond modalities from the literature whose induced non-expansive modal logic characterizes behavioural metrics for metric transition systems [12, 10, 19].

The set Λ is polynomial-time solvable because, like in the case of fuzzy transition systems, the modalities are based on taking suprema and it therefore suffices to check singletons. We therefore obtain polynomial-time computation of distinguishing formulae in both logics; while this is known for the quantitative logic [12], the result for the two-valued logic and indeed the logic itself are new.

8 Conclusion

We have introduced a coalgebraic framework for behavioural distances that are *threshold-based*, i.e. determined by notions of ε-(bi-)simulation working with a fixed allowed deviation ε. Our framework is parametrized over a choice of *2-to-$\mathcal{V}$ predicate liftings* that lift two-valued predicates to quantitative predicates (taking truth values in the unit interval $\mathcal{V}$), in generalization of the probability operator. We have shown that such distances are equivalently induced by suitably constructed quantitative lax extensions, which in turn we have shown to be characterized by quantitative modalities induced from the given 2-to-$\mathcal{V}$ predicate liftings, the *Sugeno modalities*. Both for a two-valued modal logic determined by the original 2-to-$\mathcal{V}$ predicate liftings and for a quantitative modal logic determined by the Sugeno modalities, we have established quantitative Hennessy-Milner theorems stating that the respective notions of logical distance coincide with threshold-based behavioural distance on finitely branching systems, and we have shown that distinguishing formulae certifying high behavioural distance can be computed efficiently under mild conditions. Besides recovering both established [12, 13] and recent [14, 43] results on Markov chains and metric transition systems, we obtain a number of new results as instances of our generic results. Maybe most notably, we establish the polynomial-time computation of distinguishing formulae, both two-valued and quantitative, for Lévy-Prokhorov behavioural distance, also known as ε-distance, on (labelled) Markov chains [13].

An important direction for future work is to develop generic reasoning algorithms for both two-valued and quantitative characteristic modal logics induced by 2-to-$\mathcal{V}$ predicate liftings. A possible model for such algorithms is provided in recent work on reasoning in what has been termed *non-expansive fuzzy* $\mathcal{ALC}$ [21], effectively a multi-modal version of the modal logic of fuzzy transition systems that has featured as one of our case studies.

References

1. Peter Aczel and Nax Paul Mendler. A final coalgebra theorem. In David H. Pitt, David E. Rydeheard, Peter Dybjer, Andrew M. Pitts, and Axel Poigné, editors, *Category Theory and Computer Science, CTCS 1989*, LNCS, pages 357–365. Springer, 1989. doi:10.1007/BFB0018361.
2. Jiří Adámek, Horst Herrlich, and George E. Strecker. *Abstract and concrete categories: The joy of cats*. John Wiley & Sons Inc., 1990. Republished in: Reprints in Theory and Applications of Categories, No. 17 (2006) pp. 1–507. URL: http://tac.mta.ca/tac/reprints/articles/17/tr17abs.html.
3. Liviu Aolaritei, Oliver Wang, Julie Zhu, Michael Jordan, and Youssef Marzouk. Conformal prediction under Lévy-Prokhorov distribution shifts: Robustness to local and global perturbations. In *Neural Information Processing Systems, NeurIPS 2025*, 2025. To appear. Preprint available at https://arxiv.org/abs/2502.14105. doi:10.48550/arXiv.2502.14105.
4. Roland Carl Backhouse, Peter J. de Bruin, Paul F. Hoogendijk, Grant Malcolm, Ed Voermans, and Jaap van der Woude. Polynomial relators (extended abstract). In Maurice Nivat, Charles Rattray, Teodor Rus, and Giuseppe Scollo, editors, *Algebraic Methodology and Software Technology, AMAST 1991*, Workshops in Computing, pages 303–326. Springer, 1991.
5. 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.
6. Michael Barr. Terminal coalgebras in well-founded set theory. *Theor. Comput. Sci.*, 114(2):299–315, 1993. doi:10.1016/0304-3975(93)90076-6.
7. Falk Bartels, Ana Sokolova, and Erik P. de Vink. A hierarchy of probabilistic system types. *Theor. Comput. Sci.*, 327(1-2):3–22, 2004. doi:10.1016/J.TCS.2004.07.019.
8. Gilles Barthe, Boris Köpf, Federico Olmedo, and Santiago Zanella-Béguelin. Probabilistic relational reasoning for differential privacy. In John Field and Michael Hicks, editors, *Principles of Programming Languages, POPL 2012*, pages 97–110. ACM, 2012. doi:10.1145/2103656.2103670.
9. M. Amine Bennouna, Ryan Lucas, and Bart P. G. Van Parys. Certified robust neural networks: Generalization and corruption resistance. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *International Conference on Machine Learning, ICML 2023*, volume 202 of *Proc. Machine Learning Res.*, pages 2092–2112. PMLR, 2023. URL: https://proceedings.mlr.press/v202/bennouna23a.html.
10. Harsh Beohar, Sebastian Gurke, Barbara König, and Karla Messing. Hennessy-milner theorems via galois connections. In Bartek Klin and Elaine Pimentel, editors, *Computer Science Logic, CSL 2023*, LIPIcs, pages 12:1–12:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.CSL.2023.12.
11. Rance Cleaveland. On automatically explaining bisimulation inequivalence. In Edmund M. Clarke and Robert P. Kurshan, editors, *Computer Aided Verification, CAV 1990*, volume 531 of *LNCS*, pages 364–372. Springer, 1990. doi:10.1007/BFB0023750.
12. Luca de Alfaro, Marco Faella, and Mariëlle Stoelinga. Linear and branching system metrics. *IEEE Trans. Software Eng.*, 35(2):258–273, 2009. doi:10.1109/TSE.2008.106.
13. Josée Desharnais, François Laviolette, and Mathieu Tracol. Approximate analysis of probabilistic processes: Logic, simulation and games. In *Quantitative Evaluation of Systems, QEST 2008*, pages 264–273. IEEE Computer Society, 2008. doi:10.1109/QEST.2008.42.

14 Josée Desharnais and Ana Sokolova. ε-distance via Lévy-Prokhorov lifting. In Stefano Guerrini and Barbara König, editors, *Computer Science Logic, CSL 2026*, volume 363 of *LIPIcs*, pages 26:1–26:24. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026. `doi:10.4230/LIPIcs.CSL.2026.26`.

15 Gianluca Dini, Fabio Martinelli, Ilaria Matteucci, Andrea Saracino, and Daniele Sgandurra. Introducing probabilities in contract-based approaches for mobile application security. In Joaquín García-Alfaro, Georgios V. Lioudakis, Nora Cuppens-Boulahia, Simon N. Foley, and William M. Fitzgerald, editors, *Data Privacy Management and Autonomous Spontaneous Security, DPM-SETOP 2013*, volume 8247 of *LNCS*, pages 284–299. Springer, 2013. `doi:10.1007/978-3-642-54568-9_18`.

16 Uli Fahrenberg and Axel Legay. The quantitative linear-time-branching-time spectrum. *Theor. Comput. Sci.*, 538:54–69, 2014. `doi:10.1016/J.TCS.2013.07.030`.

17 Santiago Figueira and Daniel Gorín. On the size of shortest modal descriptions. In Lev D. Beklemishev, Valentin Goranko, and Valentin B. Shehtman, editors, *Advances in Modal Logic, AiML 2010*, pages 120–139. College Publications, 2010. URL: `http://www.aiml.net/volumes/volume8/Figueira-Gorin.pdf`.

18 Jonas Forster, Sergey Goncharov, Dirk Hofmann, Pedro Nora, Lutz Schröder, and Paul Wild. Quantitative Hennessy-Milner theorems via notions of density. In Bartek Klin and Elaine Pimentel, editors, *Computer Science Logic, CSL 2023*, volume 252 of *LIPIcs*, pages 22:1–22:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. `doi:10.4230/LIPIcs.CSL.2023.22`.

19 Jonas Forster, Lutz Schröder, Paul Wild, Harsh Beohar, Sebastian Gurke, Barbara König, and Karla Messing. Quantitative graded semantics and spectra of behavioural metrics. In *Computer Science Logic, CSL 2025*, volume 326 of *LIPIcs*, pages 33:1–33:21. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025. `doi:10.4230/LIPIcs.CSL.2025.33`.

20 Jonas Forster, Lutz Schröder, Paul Wild, Barbara König, and Pedro Nora. Computing distinguishing formulae for threshold-based behavioural distances. *CoRR*, abs/2602.12084, 2026. `doi:10.48550/arXiv.2602.12084`.

21 Stefan Gebhart, Lutz Schröder, and Paul Wild. Non-expansive fuzzy $\mathcal{ALC}$. In *International Joint Conference on Artificial Intelligence, IJCAI 2025*, pages 4509–4517. ijcai.org, 2025. `doi:10.24963/IJCAI.2025/502`.

22 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*, pages 443–458. North-Holland, 1990.

23 Daniel Gorín and Lutz Schröder. Simulations and bisimulations for coalgebraic modal logics. In Reiko Heckel and Stefan Milius, editors, *Algebra and Coalgebra in Computer Science, CALCO 2013*, volume 8089 of *LNCS*, pages 253–266. Springer, 2013. `doi:10.1007/978-3-642-40206-7_19`.

24 Daniel Hausmann and Lutz Schröder. Game-based local model checking for the coalgebraic mu-calculus. In Wan J. Fokkink and Rob van Glabbeek, editors, *Concurrency Theory, CONCUR 2019*, LIPIcs, pages 35:1–35:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. `doi:10.4230/LIPIcs.CONCUR.2019.35`.

25 Bartek Klin. Structural operational semantics for weighted transition systems. In Jens Palsberg, editor, *Semantics and Algebraic Specification, Essays Dedicated to Peter D. Mosses on the Occasion of His 60th Birthday*, LNCS, pages 121–139. Springer, 2009. `doi:10.1007/978-3-642-04164-8_7`.

26 Yuichi Komorida, Shin-ya Katsumata, Nick Hu, Bartek Klin, and Ichiro Hasuo. Codensity games for bisimilarity. In *Logic in Computer Science, LICS 2019*, pages 1–13. IEEE, 2019. `doi:10.1109/LICS.2019.8785691`.

27 Barbara König and Christina Mika-Michalski. (Metric) bisimulation games and real-valued modal logics for coalgebras. In Sven Schewe and Lijun Zhang, editors, *Concurrency Theory, CONCUR 2018*, volume 118 of *LIPIcs*, pages 37:1–37:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. `doi:10.4230/LIPIcs.CONCUR.2018.37`.

28 Barbara König, Christina Mika-Michalski, and Lutz Schröder. Explaining non-bisimilarity in a coalgebraic approach: Games and distinguishing formulas. In Daniela Petrisan and Jurriaan Rot, editors, *Coalgebraic Methods in Computer Science, CMCS 2020*, volume 12094 of *LNCS*, pages 133–154. Springer, 2020. doi:10.1007/978-3-030-57201-3_8.

29 Kim Guldstrand Larsen and Arne Skou. Bisimulation through probabilistic testing. *Inf. Comput.*, 94(1):1–28, 1991. doi:10.1016/0890-5401(91)90030-6.

30 Paul Blain Levy. Similarity quotients as final coalgebras. In Martin Hofmann, editor, *Foundations of Software Science and Computational Structures, FOSSACS 2011*, volume 6604 of *LNCS*, pages 27–41. Springer, 2011. doi:10.1007/978-3-642-19805-2_3.

31 Johannes Marti and Yde Venema. Lax extensions of coalgebra functors and their logic. *J. Comput. Syst. Sci.*, 81(5):880–900, 2015. doi:10.1016/J.JCSS.2014.12.006.

32 Pedro Nora, Jurriaan Rot, Lutz Schröder, and Paul Wild. Relational connectors and heterogeneous simulations. In Parosh Aziz Abdulla and Delia Kesner, editors, *Foundations of Software Science and Computation Structures, FoSSaCS 2025*, volume 15691 of *LNCS*, pages 111–132. Springer, 2025. doi:10.1007/978-3-031-90897-2_6.

33 Amgad Rady and Franck van Breugel. Explainability of probabilistic bisimilarity distances for labelled Markov chains. In Orna Kupferman and Pawel Sobocinski, editors, *Foundations of Software Science and Computation Structures, FoSSaCS 2023*, volume 13992 of *LNCS*, pages 285–307. Springer, 2023. doi:10.1007/978-3-031-30829-1_14.

34 Jan J. M. M. Rutten. Universal coalgebra: a theory of systems. *Theor. Comput. Sci.*, 249(1):3–80, 2000. doi:10.1016/S0304-3975(00)00056-6.

35 Lutz Schröder and Dirk Pattinson. Description logics and fuzzy probability. In Toby Walsh, editor, *International Joint Conference on Artificial Intelligence, IJCAI 2011*, pages 1075–1081. IJCAI/AAAI, 2011. doi:10.5591/978-1-57735-516-8/IJCAI11-184.

36 Umberto Straccia. A fuzzy description logic. In Jack Mostow and Chuck Rich, editors, *Artificial Intelligence/ Innovative Applications of Artificial Intelligence, AAAI 1998, IAAI 1998*, pages 594–599. AAAI Press / The MIT Press, 1998. URL: http://www.aaai.org/Library/AAAI/1998/aaai98-084.php.

37 Michio Sugeno. *Theory of fuzzy integrals and its applications.* PhD thesis, Tokyo Institute of Technology, 1974.

38 Albert Thijs. *Simulation and fixpoint semantics.* PhD thesis, University of Groningen, 1996.

39 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.

40 Paul Wild and Lutz Schröder. A quantified coalgebraic van Benthem theorem. In Stefan Kiefer and Christine Tasson, editors, *Foundations of Software Science and Computation Structures, FOSSACS 2021*, volume 12650 of *LNCS*, pages 551–571. Springer, 2021. doi:10.1007/978-3-030-71995-1_28.

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

42 Paul Wild, Lutz Schröder, Dirk Pattinson, and Barbara König. A van Benthem theorem for fuzzy modal logic. In Anuj Dawar and Erich Grädel, editors, *Logic in Computer Science, LICS 2018*, pages 909–918. ACM, 2018. doi:10.1145/3209108.3209180.

43 Paul Wild, Lutz Schröder, Karla Messing, Barbara König, and Jonas Forster. Generalized Kantorovich-Rubinstein duality beyond Hausdorff and Kantorovich. In Natalie Betrand and Stefan Milius, editors, *Foundations of Software Science and Computation Structures, FoSSaCS 2026*, LNCS, pages 636–657. Springer, 2026. doi:10.1007/978-3-032-22730-0_30.

44 Thorsten Wißmann, Stefan Milius, and Lutz Schröder. Explaining behavioural inequivalence generically in quasilinear time. In Serge Haddad and Daniele Varacca, editors, *Concurrency Theory, CONCUR 2021*, LIPIcs, pages 32:1–32:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.CONCUR.2021.32.

45 Thorsten Wißmann, Stefan Milius, and Lutz Schröder. Quasilinear-time computation of generic modal witnesses for behavioural inequivalence. *Log. Methods Comput. Sci.*, 18(4), 2022. doi:10.46298/LMCS-18(4:6)2022.