Expressivity of Bisimulation Pseudometrics over Analytic State Spaces

Daniel Luckhardt
UCL, University College London, UK

Harsh Beohar
University of Sheffield, UK

Clemens Kupke
University of Strathclyde, Glasgow, UK

Abstract

A Markov decision process (MDP) is a state-based dynamical system capable of describing probabilistic behaviour with rewards. In this paper, we view MDPs as coalgebras living in the category of analytic spaces, a very general class of measurable spaces. Note that analytic spaces were already studied in the literature on labelled Markov processes and bisimulation relations. Our results are twofold. First, we define bisimulation pseudometrics over such coalgebras using the framework of fibrations. Second, we develop a quantitative modal logic for such coalgebras and prove a quantitative form of Hennessy-Milner theorem in this new setting stating that the bisimulation pseudometric corresponds to the logical distance induced by modal formulae.

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

Keywords and phrases Markov decision process, quantitative Hennessy-Milner theorem

Digital Object Identifier 10.4230/LIPIcs.CALCO.2025.13

Related Version *Full Version*: `https://arxiv.org/abs/2505.23635` [2]

1 Introduction

Markov decision processes (MDPs) are a well known mathematical model for decision-theoretic planning [5] and reinforcement learning [37]. Informally, an MDP can be seen as a generalisation of an automaton, where the transition function (for each action in the alphabet) gives a probability distribution over the state space together with a reward function that for each state and action gives a real-valued number.

Inspired from the previous work on bisimulation pseudometrics on labelled Markov processes [8, 10] and probabilistic transition systems [39, 38], Ferns et al. [17, 18] defined (among other results) a notion of bisimulation pseudometric on the states of an MDP. Unlike [17] and the previous work on bisimulation equivalence for MDPs [23], the systems considered in [18] were MDPs with continuous state spaces. Conformances over continuous state MDPs have found applications in representation learning [21, 43] (a topic studied within the field of reinforcement learning).

In this paper, we propose a modal logic $\mathcal{L}$ (cf. Section 5) with quantitative semantics, i.e. the semantics of each formula is given by a real-valued function for MDPs with continuous state space. We then prove a quantitative version of the Hennessy-Milner theorem (a well known result [25] from concurrency theory), i.e. we show that the bisimulation pseudometrics on continuous state MDPs coincide with the logical distance in our logic. A major obstacle to overcome in the continuous setting is the definition of bisimulation pseudometrics itself. Moreover, the fundamental question "what is a distance on a measurable space" (besides the usual equations of a pseudometric, i.e. d is reflexive, symmetric, and satisfies the triangle inequality) needs addressing. In [18] the authors had to invoke an additional Polish structure

© Daniel Luckhardt, Harsh Beohar, and Clemens Kupke;
licensed under Creative Commons License CC-BY 4.0

inducing the σ-algebra as their methods forced them to work with lower semi-continuous distance functions d. In this sequel, we work in a purely measure theoretic set-up with a far more general class of distances, universally measurable distance functions, cf. Subsection 3.1 for details.

Although our approach is rooted in the theory of fibrations [26], the recent approaches [1, 29, 30] to obtain expressive modal logic for coalgebras do not apply. For instance, our fibration Pred of predicates[1] over a state space has only countable many meets; thus, it is not a complete lattice fibration as required in [1, 29, 30]. As a result, the codensity lifting used in [1, 29] to derive the Kantorovich lifting for the distribution endofunctor, cannot be used to derive the Kantorovich lifting for the Giry endofunctor over measurable spaces.

In the sequel, we recalibrate the fibration infrastructure in Subsection 2.1. Our inspiration is [4] which presented a coupling-based lifting for an endofunctor that – when instantiated to distribution endofunctor – gives rise to the well-known Wasserstein lifting on probability distributions. Analogously, we will show in Subsection 4.1 how to capture the Wasserstein lifting on probability measures. For our definition to work, we will restrict to a full subcategory of the category $(X, \mathcal{A})$ of measurable spaces – the category **Ana** of analytic spaces. Note that analytic spaces already appeared in the literature on labelled Markov processes (for instance, see [9]) to show that logical equivalence induced by a modal logic given in [9] coincide with probabilistic bisimilarity.

After having clarified our measure theoretic assumptions, we will define bisimulation metrics for MDPs as the least fixpoint of the following functional:

$$\mathsf{Pred}((X,\mathcal{A})\times(X,\mathcal{A})) \xrightarrow{\sigma_X} \mathsf{Pred}(B_{\mathsf{MDP}}(X,\mathcal{A})\times B_{\mathsf{MDP}}(X,\mathcal{A})) \xrightarrow{(\gamma\times\gamma)^*} \mathsf{Pred}((X,\mathcal{A})\times(X,\mathcal{A})),$$

where B_{MDP} is the endofunctor modelling MDPs as given in Section 3 and σ is a lifting of distance functions (or put simply, a distance lifting) for B_{MDP} as given in Section 4. The definition of our distance lifting σ is parameterised by a discount factor $c \in [0, 1]$. Furthermore, thanks to the Kantorovich-Rubinstein duality for measurable spaces [33, Theorem 5], the above functional corresponds to the functional given in [18, Theorem 3.12] whose least fixpoint is the bisimulation pseudometric on the state space of an MDP.

Moving on to our modal logic and comparing with the expressive modal logic for probabilistic systems studied in [7, 10, 39], the key distinguishing feature of our work is the semantics of our diamond modality $\diamond_a\varphi$. Intuitively, the $[\![\diamond_a\varphi]\!](x)$ (for a state x) gives the expected value of landing in an a-successor from x with some fixed probability $c \in [0, 1]$ combined with the reward for a when staying in the state x with probability $1 - c$. In other words, the semantics of $\diamond_a\varphi$ is a convex combination of expected value of moving to an a-successor and the reward for a at a state. Unlike the above references, we were unable (without breaking the proof of the adequacy result) to further decompose this modality into the traditional diamond modality and 0-ary reward modality as defined in [7, 10, 39].

This paper is organised as follows. In Section 2, we recall the preliminaries from measure theory and calibrate our fibration setup for measurable spaces. In Section 3, we give the concrete definition of behaviour endofunctors that model Markov reward processes (MRPs) and MDPs and establish a bifibration of predicates. The former can be seen as unlabelled version of an MDP. In Section 4, we capture the bisimulation pseudometrics for both MRPs and MDPs as least fixpoint of a functional as explained above. In Section 5, we define our modal logic and establish the adequacy and expressivity results. In Section 6, we end this paper by a discussion on related work and potential topics for future work. The proofs of all lemmas and theorems can be found in the clearly marked appendix.

[1] Note that a fibration of predicates is more fundamental than a fibration of conformances like pseudometrics and equivalence relations, since the latter can be derived from the former.

2 Preliminaries

2.1 Capturing behavioural conformances categorically

In this subsection, we refine the construction [4] of coupling-based lifting for an endofunctor on **Set** by working with two different fibrations of predicates (cf. Assumptions **A1** and **A2**). Moreover, our presentation works in a category $\mathbf{C}$ having products; unlike, in [4], where the coalgebras were living in **Set**. This will provide us a blueprint to define a bisimulation distance for both MRPs and MDPs when viewed as coalgebras in Section 3.

Throughout this section, let **Pos** be the category of posets and order preserving maps; and, let $B\colon \mathbf{C} \to \mathbf{C}$ be the functor modelling the branching type of systems of interest.

A1. There is an indexed category $\mathsf{Pred}\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Pos}$ such that $\mathsf{Pred}(X)$ (for $X \in \mathbf{C}$) is a poset and $\mathsf{Pred}(f)\colon \mathsf{Pred}(Y) \to \mathsf{Pred}(X)$ (for $f\colon X \to Y \in \mathbf{C}$) is order preserving.

Henceforth we write the reindexing f^* (instead of $\mathsf{Pred}(f)$) which is customary in the literature on fibrations [26]. In the sequel, we will view an element $p \in \mathsf{Pred}(X)$ intuitively as a predicate over an object $X \in \mathbf{C}$. The idea is to view Pred as a semantic universe in which we interpret the formulae of a modal logic. Thus, the operators of a modal logic (like negation, conjunction etcetera) must be operators definable over the fibre $\mathsf{Pred}(X)$.

Furthermore, the authors in [4] required that Pred is rather a bifibration, which is difficult to obtain in general for arbitrary measurable spaces (cf. Subsection 3.1). Our observation, which leads to **A2**, is that we can arrange both universally measurable predicates and lower semi-measurable predicates in such a way that the latter results in a bifibration structure and the former acts as a semantic universe to interpret our modal formulae.

A2. there is an indexed category IsPred such that IsPred is a subfunctor of Pred. Moreover, the indexed category IsPred has a bifibration structure, i.e. for every $f\colon X \to Y \in \mathbf{C}$ the reindexing functor f^* has a left adjoint $\exists_f\colon \mathsf{IsPred}(X) \to \mathsf{IsPred}(Y)$.

Now, following [4], one needs a predicate lifting to define a coupling-based lifting, which in our setting due to the presence of two fibrations of predicates takes the following shape.

A3. there is an indexed morphism $\sigma\colon \mathsf{Pred} \Rightarrow \mathsf{IsPred} \circ B^{\mathrm{op}}$, i.e. σ is a natural transformation of type $\mathsf{Pred} \Rightarrow \mathsf{IsPred} \circ B^{\mathrm{op}}$.

Thanks to the above three assumptions, every predicate lifting σ induces a lifting $\hat{\sigma}$, which simplifies to the composition given in [4, Eq. 5] when $\mathsf{IsPred} = \mathsf{Pred}$).

$$\mathsf{Pred}(X \times X) \xrightarrow{\sigma_{X\times X}} \mathsf{IsPred}(B(X \times X)) \xrightarrow{\exists_{\pi_X}} \mathsf{IsPred}(BX \times BX) \hookrightarrow \mathsf{Pred}(BX \times BX), \quad (1)$$

where $\pi_X\colon B(X \times X) \to BX \times BX$ is the unique map such that $\mathrm{pr}_i^{BX} \circ \pi_X = B(\mathrm{pr}_i^X)$ and $\mathrm{pr}_i\colon X \times X \to X$ are the obvious projection maps (for $i \in \{1, 2\}$). It is this lifting which will give us the usual Wasserstein lifting for a Giry functor B defined in Section 3. Now, for a given coalgebra $\gamma\colon X \to BX \in \mathbf{C}$, simply take the greatest fixpoint of the functional given below to define a *coupling-based lifting* for the endofunctor B.

$$\mathsf{Pred}(X \times X) \xrightarrow{\hat{\sigma}_X} \mathsf{Pred}(BX \times BX) \xrightarrow{(\gamma\times\gamma)^*} \mathsf{Pred}(X \times X). \quad (2)$$

To ensure this fixpoint exists, we require the following assumption

A4. the indexed category Pred has countable fibred limits, i.e. each fibre of Pred has countable meets and these countable meets are preserved by the reindexing operation.

▶ **Proposition 1.** *If the induced lifting $\hat{\sigma}$ defined in* (1) *is (Scott) cocontiuous, then the greatest fixpoint of the functional given in* (2) *exists.*

▶ Remark 2. It should be noted that, in the above proposition, we use the dual version to apply Kleene's fixpoint theorem on the lattice Pred. However, our concrete predicates in Section 3 will be ordered by pointwise lifting of the dual order (i.e. $\geq$) on the unit interval $[0,1]$, which then leads to the application of the usual Kleene's fixpoint theorem.

2.2 Measurable spaces

A measurable space is a pair $(X,\mathcal{A})$ consisting of a set X thought of as a space, e.g. state space of an MDP, and a σ-algebra $\mathcal{A}\subseteq\mathcal{P}(X)$ (whose elements are called *measurable* sets) of subsets of X stable under the complement operation $(_)^{\complement}$ and countable union (including empty union). In applications the σ-algebra often contains a given topology, i.e. the collection of open sets $\mathcal{T}$, on the state space. Often one considers the minimal such σ-algebra, the *Borel-σ-algebra*, denoted $\mathcal{B}_{\mathcal{T}}$. In this context, we use the notation $\sigma\langle\mathcal{P}\rangle$ to denote the minimal σ-algebra containing a given $\mathcal{P}\subseteq\mathcal{P}(X)$. The elements of $\mathcal{B}_{\mathcal{T}}=\sigma\langle\mathcal{T}\rangle$ are called *Borel* sets and $(\mathcal{T},\mathcal{B}_{\mathcal{T}})$ is called a *Borel space*. If a given measurable space $(X,\mathcal{A})$ stems from a Polish space, a completely metrisable separable topological space, then $(X,\mathcal{A})$ is called *standard*. The collection of all measurable spaces form a category **Meas** when endowed with maps that inversely preserve measurable sets, so-called *measurable maps*. The term *$\mathcal{A}$-$\mathcal{B}$-measurable* for a measurable morphism $(X,\mathcal{A})\to(Y,\mathcal{B})$ is also used to be specific about σ-algebras. The category **Meas** has arbitrary products; in particular, $(X,\mathcal{A})\times(Y,\mathcal{B})=(X\times Y,\mathcal{A}\otimes\mathcal{B})$ where $\mathcal{A}\otimes\mathcal{B}$ is the σ-algebra generated by the set $\{\,U\times V\mid U\in\mathcal{A}\wedge V\in\mathcal{B}\,\}$.

A *probability measure* $\mathfrak{m}$ on a measurable space $(X,\mathcal{A})$ is a function of type $\mathcal{A}\to[0,1]$ such that $\mathfrak{m}(X)=1$ and $\mathfrak{m}(\bigcup_{i\in\mathbb{N}}A_i)=\sum_{i\in\mathbb{N}}\mathfrak{m}(A_i)$ for any sequence of pairwise disjoint sets $(A_i\in\mathcal{A})_{i\in\mathbb{N}}$. We denote by $\mathcal{G}(X,\mathcal{A})$ the collection of all probability measures on $(X,\mathcal{A})$ endowed – going back to [22] – with the σ-algebra generated by all the evaluation maps $\mathrm{ev}_A\colon\mathcal{G}(X,\mathcal{A})\to[0,1]$ (one for each $A\in\mathcal{A}$) given by the mapping $\mathfrak{m}\mapsto\mathfrak{m}(A)$, i.e. the minimal σ-algebra making all maps $\mathrm{ev}_A\colon\mathcal{G}(X,\mathcal{A})\to([0,1],\mathcal{B}_{[0,1]})$ measurable.

We restrict the exposition of this theory to it bare minimum with some additional background given in Section A. We call a subset of a measurable space $(X,\mathcal{A})$ *Suslin*, if it is the image of an element in $\mathcal{B}_{\mathbb{N}^{\mathbb{N}}}\otimes\mathcal{A}$ along the projection $\mathbb{N}^{\mathbb{N}}\times X\to X$. Let $\mathfrak{S}\,\mathcal{A}$ denote the set of all Suslin subsets of $(X,\mathcal{A})$. A measurable space $(X,\mathcal{A})$ is called *analytic* if it is homeomorphic to a Suslin set of a standard space $(Y,\mathcal{B})$ endowed with the *restricted σ-algebra*, i.e. $\mathcal{B}|_X\coloneqq\{\,B\cap X\mid B\in\mathcal{B}\,\}$. Such constructions are also a common subject in descriptive set theory, cf. [27] and [20, Ch. 42]. By **Ana** we denote the full subcategory of **Meas** of analytic spaces, which admits countable products. The endofunctor $\mathcal{G}$ restricts to **Ana**. To provide full generality, cf. Remark 6, of our results, we also introduce the concept of a smooth space [16]: $(X,\mathcal{A})$ is *smooth*, cf. Subsection A.2, if for any other measurable space $(Y,\mathcal{B})$ any projection to Y of a Borel (or equivalently Suslin) subset of $(Y,\mathcal{B})\times(X,\mathcal{A})$ is Suslin. Nevertheless, it is perfectly fine to assume analytic spaces throughout at least for the first reading.

Given a measure space (i.e., a measurable space with a measure) $(X,\mathcal{A},\mathfrak{m})$ one may wish to extend $\mathcal{A}$. The $\mathfrak{m}$-completion of $(X,\mathcal{A})$, $\overline{\mathcal{A}}^{\mathfrak{m}}\supseteq\mathcal{A}$, is defined as the smallest σ-algebra containing $\mathcal{A}\cup\{B\subseteq X\mid\exists A\in\mathcal{A}.\ m(A)=0\text{ and }B\subseteq A\}$. A way to describe $\overline{(X,\mathcal{A})}^{\mathfrak{m}}$, when $\mathfrak{m}$ is a probability measure, is as the set of all A such that there are A_-,A_+ with $\mathfrak{m}(A_-)=\mathfrak{m}(A_+)$ and $A_-\subseteq A\subseteq A_+$. The measure $\mathfrak{m}$ uniquely extends to a measure on $\overline{\mathcal{A}}^{\mathfrak{m}}$. Given a measurable space $(X,\mathcal{A})$, the *universal completion* of $\mathcal{A}$ (denoted $\overline{\mathcal{A}}$) is the

intersection of all completions of $\mathcal{A}$ with respect to any (probability) measure on $(X, \mathcal{A})$. The universal completion is quite big; especially it contains $\mathfrak{S}\,\mathcal{A}$ (so every Suslin set is universally measurable).

3 Markov decision processes

In this section we are going to instantiate our categorical parameters (cf. Assumptions **A1** - **A4**) in the setting of measurable spaces. We begin by recalling the definition of a Markov decision processes from [18] and view them as coalgebras.

▶ **Definition 3.** *A (continuous)* Markov reward process *(MRP) is a coalgebra γ of type*

$$(X, \mathcal{A}) \to \mathcal{G}(X, \mathcal{A}) \times ([0,1], \mathcal{B}_{[0,1]}) \in \mathbf{Meas}.$$

In other words, γ is given by a pair (γ^0, γ^1) of maps satisfying the following properties:

- *$\gamma^0(x)$ (for each $x \in X$) is a probability measure;*
- *$\gamma^0(_)(U) \colon X \to [0,1]$ (for each measurable set $U \in \mathcal{A}$) is a measurable function; and*
- *γ^1 is a measurable function.*

Moreover, given a countable set Σ of actions, we define a Markov decision process *(MDP) [18] as a coalgebra γ of type*

$$(X, \mathcal{A}) \to \prod_{\Sigma} \left(\mathcal{G}(X, \mathcal{A}) \times ([0,1], \mathcal{B}_{[0,1]})\right) \in \mathbf{Meas}.$$

In other words, the map $\gamma(x)(a)$ (for each state $x \in X$ and action $a \in \Sigma$), is a Markov process. Henceforth we write $\gamma_{a,x}$ to denote $\gamma(x)(a)$; so, $\gamma^0_{a,x}$ corresponds to a probability measure and $\gamma^1_{a,x}$ corresponds to a "reward" at x for an action a.

Thus, the endofunctors of interest are the following:

- $B_{\mathsf{MRP}} = \mathcal{G} \times [0,1]$ whose coalgebras correspond to Markov reward processes
- $B_{\mathsf{MDP}} = \prod_{\Sigma} B_{\mathsf{MRP}}$ whose coalgebras correspond to Markov decision processes.

3.1 Fibrations induced universally/l.s.m. predicates

Having fixed the type of systems, we now look into the issue of endowing a bifibration structure on the space of all Boolean/quantitative predicates. Consider the indexed category $\mathsf{Pred}(X, \mathcal{A}) = \mathbf{Meas}((X, \mathcal{A}), (2, \mathcal{B}_2))$ of Boolean predicates, i.e. a predicate $p \in \mathsf{Pred}(X, \mathcal{A})$ is a measurable function of type $X \to 2$. The reindexing functor f^* (for a measurable function $f \colon (X, \mathcal{A}) \to (Y, \mathcal{B})$) is given by the inverse image operation (since inverse image of measurable sets is measurable). It is well known (originally due to Suslin [36]) that Borel measurable sets even on standard spaces are not closed under direct images; thus as a result the left adjoint to reindexing functor cannot exist in general. Nevertheless if we weaken measurable sets to Suslin sets (which are equivalently analytic sets for analytic spaces, cf. [20, 421K] and [27, 13.3iii)]), then Suslin sets of analytic spaces are preserved by direct image onto analytic spaces, cf. [6] for an in-depth discussion how to develop these concepts.

In lieu of the above discussion, we restrict our state spaces to be analytic, i.e. our working category for the remainder is the category **Ana** of analytic spaces and measurable functions as morphisms. This is, on the one hand, a bit more general than Polish spaces as required in [18] and on the other hand conceptually more elegant, as we are only working with measurable spaces and do not require an underlying topology. Moreover, we consider quantitative predicates on an analytic space $(X, \mathcal{A})$ to be *lower semimeasurable* (l.s.m.) functions (a real-valued generalisation of Suslin sets) of type $X \to [0,1]$.

▶ **Definition 4.** *Let $(X, \mathcal{A})$ be an analytic space. Then a function $p\colon X \to [0,1]$ is* lower semi-measurable, l.s.m. *for short, (resp.* universally measurable*) predicate iff the preimage of the interval $[0,r]$ (for every $r \in [0,1]$) under p is a Suslin set (resp. universally measurable set), i.e. for every $r \in [0,1]$, $p^{-1}([0,r]) \in \mathfrak{S}\,\mathcal{A}$ (resp. $p^{-1}([0,r]) \in \overline{\mathcal{A}}$).*

The term "lower semi-measurable" is chosen in parallel to the term "lower semi-continuous" in topology which refers to a real-valued function which is continuous with respect to the upper-interval topology $\{\,(r,\infty) \mid r \in \mathbb{R}\,\}$.

We can arrange l.s.m. predicates in an indexed category as follows. Consider the mapping $\mathsf{lsPred}\colon \mathbf{Ana}^{\mathrm{op}} \to \mathbf{Pos}$ such that $\mathsf{lsPred}(X,\mathcal{A})$ is the set of all l.s.m. predicates where the ordering relation is the pointwise lifting of the "greater-than-equality" relation on the unit interval[2]. The reindexing f^* (for an arrow $f\colon (X,\mathcal{A}) \to (Y,\mathcal{B}) \in \mathbf{Ana}$) is given by pre-composition, i.e.

$$f^*(q) = q \circ f \qquad (\text{for every } q \in \mathsf{lsPred}(Y,\mathcal{B})).$$

▶ **Lemma 5.** *The indexed category* lsPred *has countable fibred (co)limits, i.e. each fibre has countable meets and countable joins which are preserved by the reindexing functor. Moreover,* lsPred *has a bifibration structure, i.e., for every $f\colon (X,\mathcal{A}) \to (Y,\mathcal{B}) \in \mathbf{Ana}$, the reindexing functor has a left adjoint $\exists_f$ given by:*

$$\exists_f(p)(y) = \inf_{f(x)=y} p(x) \qquad (\textit{for every } p \in \mathsf{lsPred}(X,\mathcal{A}), y \in Y).$$

▶ Remark 6. It should be noted that the existence of a left adjoint can be stated in more general terms by requiring that $(X,\mathcal{A})$ is a smooth space, cf. Subsection A.2, and $(Y,\mathcal{B})$ is countably separated, i.e. there is a countable family of measurable sets distinguishing every pair of distinct points. Moreover, Axiom 2 could be weakened to maps of type $f\colon BX \to Y \in \mathbf{C}$:

A1'. the reindexing functor f^* has a left adjoint $\exists_f\colon \mathsf{lsPred}(BX) \to \mathsf{lsPred}(Y)$.

Any universally measurable subset of a standard space is countably separated as a measurable space, but also an analytic space. So our construction generalises to the full subcategory of **Meas** of measurable spaces expressible in this form.

To show that **A2** is satisfied, it remains to define an indexed category $\mathsf{Pred}\colon \mathbf{Ana}^{\mathrm{op}} \to \mathbf{Pos}$ such that each fibre $\mathsf{lsPred}(X,\mathcal{A})$ is contained in $\mathsf{Pred}(X,\mathcal{A})$. We disregard the trivial definition, i.e. $\mathsf{Pred} = \mathsf{lsPred}$, since Suslin sets are not closed under complementation. As a result, we cannot give semantics to the negation operator in our logic. Nonetheless, it is also known that Suslin sets are universally measurable sets [11], so we simply let $\mathsf{Pred}(X,\mathcal{A})$ be the set of universally measurable predicates on the analytic space $(X,\mathcal{A})$.

▶ **Proposition 7.** *Assumptions **A1** and **A2** are satisfied by* Pred *and* lsPred*, respectively.*

[2] In other words, we are viewing the unit interval as the Lawvere quantale $([0,1],\geq,+)$ where $+$ is the truncated addition. So, $p \leq q \iff \forall_x\, p(x) \geq q(x)$.

4 Bisimulation distance

The objective of this section is to define bisimulation pseudometrics (cf. Subsection 4.2) for Markov reward processes and MDPs as the least[3] fixpoint of a functional given in (2) on page 3 where $B = \{B_{\mathsf{MRP}}, B_{\mathsf{MDP}}\}$. In both cases, the definition of a pseudometric rests on a coupling-based lifting (1) for the Giry endofunctor $\mathcal{G}$ which we will work out in the following subsection.

4.1 Wasserstein lifting categorically

We begin by defining a predicate lifting for $\mathcal{G}$ (i.e. when $B = \mathcal{G}$ in Assumption **A3**). Consider the mapping $\sigma_{(X,\mathcal{A})}\colon \mathsf{Pred}(X,\mathcal{A}) \to \mathsf{IsPred}(\mathcal{G}(X,\mathcal{A}))$ given by

$$\sigma_{(X,\mathcal{A})}(p)(\mathfrak{m}) = \int p \,\mathrm{d}\mathfrak{m} \quad (\text{for every } p \in \mathsf{Pred}\,(X,\mathcal{A}), \mathfrak{m} \in \mathcal{G}(X,\mathcal{A})). \tag{3}$$

Henceforth, we drop the sigma-algebra notation from the subscript whenever it is clear from the context. Thus, $\sigma_X(p)(\mathfrak{m})$ is the expectation of random variable p under the measure $\mathfrak{m}$.

▶ **Theorem 8.** *The mapping σ defined in* (3) *is a natural transformation valued in* IsPred*; thus, an indexed category morphism. Moreover, σ preserves directed suprema which is a consequence of the monotone convergence theorem well known in measure theory.*

Note that predicate lifting improves universally measurable predicates even to Borel measurable predicates for analytic spaces; the proof of this fact can be found in [2].

Thus, **A3** is satisfied and invoking the $\hat{\sigma}$ given in (1) gives the usual Wasserstein lifting for the Giry endofunctor $\mathcal{G}$ as expected.

▶ **Definition 9.** *A predicate $d \in \mathsf{Pred}((X,\mathcal{A}) \times (X,\mathcal{A}))$ is a* pseudometric *on $(X,\mathcal{A}) \in \mathbf{Ana}$ iff d is reflexive, symmetric, and satisfies the triangle inequality.*

Moreover, a probability measure $\mathfrak{c} \in \mathcal{G}((X,\mathcal{A}) \times (X,\mathcal{A}))$ is a coupling *for two probability measures $\mathfrak{m}, \mathfrak{n}$ iff $\mathcal{G}(\mathrm{pr}_1)(\mathfrak{c}) = \mathfrak{m}$ and $\mathcal{G}(\mathrm{pr}_2)(\mathfrak{c}) = \mathfrak{n}$. We write $K(\mathfrak{m},\mathfrak{n})$ to denote the set of all couplings for the probability measures $\mathfrak{m}, \mathfrak{n}$.*

▶ **Proposition 10.** *Let d be a pseudometric on a space $(X,\mathcal{A}) \in \mathbf{Ana}$. Then, the lifting $\hat{\sigma}$ given in* (1) *evaluates to the following well known formula associated with Wasserstein lifting of probability measures. Moreover, $\hat{\sigma}(d)$ is a pseudometric on $\mathcal{G}(X,\mathcal{A})$.*

$$\hat{\sigma}(d)(\mathfrak{m},\mathfrak{n}) = \inf_{\mathfrak{c} \in K(\mathfrak{m},\mathfrak{n})} \int d \,\mathrm{d}\mathfrak{c}$$

4.2 Distance lifting for $B = \{B_{\mathsf{MRP}}, B_{\mathsf{MDP}}\}$

One way to define the distance lifting σ^{MRP} for B_{MRP}, i.e. a map of type

$$\sigma^{\mathsf{MRP}}_X\colon \mathsf{Pred}((X,\mathcal{A}) \times (X,\mathcal{A})) \Rightarrow \mathsf{Pred}(B_{\mathsf{MRP}}(X,\mathcal{A}) \times B_{\mathsf{MRP}}(X,\mathcal{A})),$$

is to first define a predicate lifting for B_{MRP} and then use the equation in (1) where $B = B_{\mathsf{MRP}}$. To this end one may follow [28, Subsection 5.6.2] in deriving a predicate lifting for B_{MRP} in a compositional manner. These results (though stated for the category of sets) can

[3] Recall the predicates are ordered by $\geq$, so the greatest fixpoint is actually least fixpoint under the usual order $\leq$.

be generalised to measurable spaces, but they are only applicable when the underlying endofunctor preserves weak pullbacks. In particular, it is known that the Giry functor (a composite functor in the case of B_{MRP}) does not preserve weak pullbacks in **Meas** [41].

So instead of compositionally deriving predicate liftings for B_{MRP} and then invoking (1), we derive the coupling based lifting for B_{MRP} in three stages:

- first, we view B_{MRP} as the composition of functors $B_I \circ \mathcal{G}$ where $B_I \colon \mathbf{Ana} \to \mathbf{Ana}$ maps every space to its product with the unit interval, i.e.

  $$B_I = \mathrm{Id} \times ([0,1], \mathcal{B}_{[0,1]}).$$

- second, for a fixed $c \in [0,1]$, we define

  $$\sigma^c_X \colon \mathsf{Pred}((X,\mathcal{A}) \times (X,\mathcal{A})) \Rightarrow \mathsf{Pred}(B_I(X,\mathcal{A}) \times B_I(X,\mathcal{A}))$$

  as $\sigma^c_X(d)((x,r),(y,s)) = cd(x,y) + (1-c)|r-s|$, for $x, y \in X$ and $r, s \in [0,1]$.
- third, recall $\hat{\sigma}$ from Proposition 10 and let σ^{MRP} be the composition:

$$\mathsf{Pred}((X,\mathcal{A})\times(X,\mathcal{A})) \xrightarrow{\hat{\sigma}_X} \mathsf{Pred}(\mathcal{G}(X,\mathcal{A})\times\mathcal{G}(X,\mathcal{A})) \xrightarrow{\sigma^c_{\mathcal{G}X}} \mathsf{Pred}(B_{\mathsf{MRP}}(X,\mathcal{A})\times B_{\mathsf{MRP}}(X,\mathcal{A})).$$

Note that c may take the extremal values 0 and 1. This is possible – in contrast to [18] – as the bisimulation distance is not obtained using a contraction-based fixpoint argument. However, in the extreme cases the bisimulation distance would not take into account either the transition or the reward part.

▶ **Lemma 11.** *The above mapping σ^c is well defined. Moreover, for any pseudometric $d \in \mathsf{Pred}((X,\mathcal{A}) \times (X,\mathcal{A}))$, the lift $\sigma^{\mathsf{MRP}}_X(d)$ is given by*

$$\sigma^{\mathsf{MRP}}_X(d)((\mathfrak{m},r),(\mathfrak{n},s)) = c\left(\inf_{\mathfrak{c}\in K(\mathfrak{m},\mathfrak{n})} \int d \,\mathrm{d}\mathfrak{c}\right) + (1-c)|r-s|$$

for $\mathfrak{m}, \mathfrak{n} \in \mathcal{G}(X,\mathcal{A})$, and $r, s \in [0,1]$ and is a pseudometric.

In a similar vein, we can now define a distance lifting σ^{MDP} for B_{MDP} (whose coalgebras model MDPs) by letting $B_{\mathsf{MDP}} = B_\Sigma \circ B_{\mathsf{MRP}}$, where $B_\Sigma = \prod_\Sigma \mathrm{Id}$. Now consider the distance lifting $\sigma^\Sigma_{(_)}$ for B_Σ as follows:

$$\sigma^\Sigma(d)(\vec{x},\vec{y}) = \sup_{a\in\Sigma} d(\vec{x}(a),\vec{y}(a)), \quad (\text{for } \vec{x},\vec{y} \in \prod_\Sigma X).$$

▶ **Lemma 12.** *The above mapping σ^Σ is well defined. Moreover, the mapping $\sigma^{\mathsf{MDP}} = \sigma^\Sigma_{B_{\mathsf{MRP}}(_)} \circ \sigma^{\mathsf{MRP}}$ for a distance $d \in \mathsf{Pred}((X,\mathcal{A}) \times (X,\mathcal{A}))$ evaluates to*

$$\sigma^{\mathsf{MDP}}(d)((\vec{\mathfrak{m}},\vec{r}),(\vec{\mathfrak{n}},\vec{s})) = \sup_{a\in\Sigma}\left[c\left(\inf_{\mathfrak{c}\in K(\vec{\mathfrak{m}}(a),\vec{\mathfrak{n}}(a))} \int d \,\mathrm{d}\mathfrak{c}\right) + (1-c)|\vec{r}(a) - \vec{s}(a)|\right].$$

Now composing the two distance liftings σ^B (for $B \in \{B_{\mathsf{MRP}}, B_{\mathsf{MDP}}\}$) with a B-coalgebra is the desired functional as given in (2). Clearly, **A4** is satisfied since Pred has countable suprema and they are preserved by reindexing functors. We end this subsection by showing that the least fixpoint exists for both of these functionals; thus, also paving a way to compute bisimulation pseudometrics for these systems. To this end we need a general result, whose proof is based on some classical results comprising a non-topological version of Riesz–Markov–Kakutani representation theorem [13, IV.5.1], Banach-Alaoglu theorem [13, V.4.2] and Sion's minimax theorem [35, Thm. 3].

▶ **Theorem 13.** *Let $\gamma\colon (X,\mathcal{A}) \to \mathcal{G}(X,\mathcal{A}) \in \mathbf{Ana}$. Then the functional $\gamma \circ \hat{\sigma}$ is ω-cpo continuous w.r.t. $\leq$, i.e. for any $\leq$-increasing sequence $d_i \in \mathsf{Pred}((X,\mathcal{A}) \times (X,\mathcal{A}))$ of pseudometrics with $i \in \mathbb{N}$, we have (for each $x, y \in X$):*

$$\inf_{\mathfrak{c} \in K(\gamma_x,\gamma_y)} \int \sup_{i\in\mathbb{N}} d_i \,\mathrm{d}\mathfrak{c} = \sup_{i\in\mathbb{N}} \inf_{\mathfrak{c} \in K(\gamma_x,\gamma_y)} \int d_i \,\mathrm{d}\mathfrak{c}.$$

▶ Remark 14. The above theorem can be stated for general measurable spaces as well, but by restricting the coalgebra map so that $\gamma(x)$ (for each $x \in X$) is a perfect measure (see Subsection A.1). Perfect measures were introduced by Kolmogorov [24, 22–23] and have many different equivalent definitions. For us a measure space $(X, \mathcal{A}, \mathfrak{m})$ is called *perfect*, if for any separable metrisable space $(Y, \mathcal{T})$ and every measurable map $f\colon X \to Y$ we have the following property: For every $A \in \mathcal{A}$ and $r < \mathfrak{m}(A)$ there is a compact set $K \subseteq \operatorname{im} f$ with $\mathfrak{m}(A \cap f^{-1}K) \geq r$, cf. [20, 451O(a)]. In this case, the measure $\mathfrak{m}$ is called *perfect*.

▶ Remark 15. Let $p \in [1,\infty)$ and recall the p-Wasserstein distance between probability measures $\mathfrak{m}, \mathfrak{n} \in \mathcal{G}(X,\mathcal{A})$. Below we argue how to capture this lifting in our setup.

$$W_p(d)(\mathfrak{m},\mathfrak{n}) = \sqrt[p]{\inf_{\mathfrak{c}\in K(\mathfrak{m},\mathfrak{n})} \int d^p \,\mathrm{d}\mathfrak{c}}, \qquad \text{for a pseudometric } d \in \mathsf{Pred}((X,\mathcal{A}) \times (X,\mathcal{A})).$$

Note that any monotonously increasing lower semicontinuous function $f\colon [0,1] \to [0,1]$ induces a ω-cpo-continuous map $f \circ _\colon \mathsf{Pred}(X,\mathcal{A}) \to \mathsf{Pred}(X,\mathcal{A})$. It is monotone by monotonicity of f and for any increasing sequence $(p_i \in \mathsf{Pred}(X,\mathcal{A}))_{i\in\mathbb{N}}$ we have

$$\sup_{i\in\mathbb{N}} f(p_i(x)) = f\left(\sup_{i\in\mathbb{N}} p_i(x)\right) \qquad \text{(for each } x \in X\text{)}.$$

Note that both $(_)^p$ and $\sqrt[p]{_}$ are monotonously increasing lower-semicontinuous functions. So $W_p = \sqrt[p]{_} \circ \hat{\sigma} \circ (_)^p$ is ω-cpo-continuous as a composition of ω-cpo-continuous functions.

▶ **Corollary 16.** *Let $\gamma_B\colon (X,\mathcal{A}) \to B(X,\mathcal{A}) \in \mathbf{Ana}$ be a coalgebra where $B \in \{B_{\mathsf{MRP}}, B_{\mathsf{MDP}}\}$. Then the least fixpoint for the functionals $\gamma_B \circ \sigma^B$ exists.*

Using the fact that any ω-cpo-continuous endofunction has a least fixpoint by Kleene's fixpoint theorem, we write $\mathbf{bd}_c^\gamma$ (or simply $\mathbf{bd}$ whenever the coalgebra structure is clear from the context) to denote the least fixpoint of the functionals in the above corollary.

Note that by using Kleene fixed point theorem we require weaker assumptions than [17, 3.12], who use the Banach fixed point theorem to define their bisimulation pseudometric restrict themselves to a set-up with contractions.

5 A quantitative modal logic and its expressivity

The signatures of the (logical) languages considered in this paper are parametrised by a set $\{ f_{iy} \mid y \in Y_i \}$ of ω-indexed families of Y_i-indexed function symbols of arity n_i as follows:

$$\top \mid \neg_ \mid _ \wedge _ \mid \diamond_a _, a \in A \mid f_{iy}, y \in Y_i, i \in \omega \tag{4}$$

In other words, the signature we are using extends the semi-lattice signature $(\top, \wedge)$ with negation $(\neg)$, modalities $\diamond_a$ (one for each action a) and additional function symbols f_{iy} (each f_{iy} could be viewed as n_i-ary predicate on the unit interval). We simply write $\mathcal{L}$ to denote the set of formulae generated by the above signature. The restriction to countably many families

of function symbols will become important when we construct a second-countable topology on $\mathcal{L}$. Note that we can also view the logical symbols, $\top, \neg_, _ \wedge _$, as (singleton index families of) function symbols. This will be very handy for proofs by structural induction over $\mathcal{L}$. Throughout this section, we let $\Omega = [0,1]$ and consider only MDPs (the modal logic for MRPs can be derive by letting the set Σ of actions to be a singleton set).

The reason to choose this general formulation with index sets Y_i are twofold. First, to endow a topology on $\mathcal{L}$ which is needed to prove both adequacy (i.e. the distance induced by formulae in $\mathcal{L}$ is below than the bisimulation distance **bd** for MDPs) and expressivity (which is the converse of adequacy). Second, it allows greater flexibility in applications, cf. (6), by accommodating additional operators for which adequacy and expressivity still hold.

5.1 Interpretation of modal formulae in $\mathcal{L}$

We give semantics to each formulae $\varphi \in \mathcal{L}$ by defining an *interpretation* $[\![\varphi]\!]$ as a predicate in $\mathsf{Pred}(X, \mathcal{A})$. This is done by structural induction over φ in the following way. The logical symbols, i.e. $\top$ (truth), $\neg_$ (negation), and $_ \wedge _$ (conjunction), respectively, are interpreted as the functions

$$1, \quad \lambda\, x.1 - x, \quad \lambda\, x\, y.\min\{x, y\} \quad \text{respectively.} \tag{5}$$

Next we consider function symbols f_{iy} of arity $n_i > 0$ (for some $y \in Y_i$) and a sequence $(\varphi_j)_{1 \le j \le n_i}$. Its interpretation is given as

$$[\![f_{iy}]\!](\varphi_1, \cdots, \varphi_{n_i}) = f_{iy}([\![\varphi_1]\!], \cdots, [\![\varphi_{n_i}]\!]) \text{ for some fixed } f_{iy} : [0,1]^{n_i} \to [0,1]$$

whilst, $[\![f_{iy}]\!] = f_{iy} \in [0,1]$ when $n_i = 0$ (i.e. f_{iy} is a constant). Besides the logical symbols we introduce the following *basic* families of function symbols indexed by $r, c \in [0,1]$ and $\alpha \in (0, \infty)$: *Scalar addition* $(_ + r)$, *scalar subtraction* $(_ - r)$, *scalar multiplication* $(r\,_)$, and *convex combination* $(_ +_c _)$ are interpreted as the function assigning to x (and y)

$$\min\{1, x + r\}, \quad \max\{0, x - r\}, \quad rx, \quad x +_c y, \tag{6}$$

respectively, throughout this paper. Note that only the first two basic operations are required for our main results.

Finally, for the modal operators we fix a parameter $c \in \Omega$ (which was also used to define bisimulation pseudometrics **bd** in the previous section) and a coalgebra $\gamma\colon (X, \mathcal{A}) \to B_{\mathsf{MDP}}(X, \mathcal{A}) \in \mathbf{Ana}$ to define the interpretation of $[\![\diamond_a \varphi]\!]$ as:

$$[\![\diamond_a \varphi]\!]_\gamma(x) = \int [\![\varphi]\!]\, \mathrm{d}\gamma^0_{a,x} \;+_c\; \gamma^1_{a,x}, \tag{7}$$

where $r +_c s \coloneqq cr + (1-c)s$ denotes convex combination of $r, s \in \Omega$. Intuitively, this means that the value of $\diamond_a \varphi$ is determined by a convex combination of the expected value of φ and the utility after performing a.

The defined interpretation function $[\![_]\!]$ gives rise a (quantitative) theory map $\mathrm{qTh}\colon X \to \Omega^{\mathcal{L}}$ defined by

$$\mathrm{qTh}(x)(\varphi) \coloneqq [\![\varphi]\!](x) \quad (\text{for every } \varphi \in \mathcal{L}). \tag{8}$$

The question whether the theory map lives in our working category **Ana** is an important step for expressivity (cf. Theorem 28). However, for adequacy (cf. Theorem 22), we only require that the function symbols f_{iy} in $\mathcal{L}$ are nonexpansive w.r.t. the suprema distance

$d_\infty(\vec{x},\vec{y}) = \max_{1\leq j\leq n_i}|\vec{x}(j),\vec{y}(j)|$. Nonetheless, before attempting these results we need the definition of a logical distance, which at this stage is purely a set-theoretic assignment. Later, in next subsection, we will show that the logical distance defined below is indeed a predicate over the product of a state space with itself (cf. Lemma 21).

▶ **Definition 17** (Logical distance). *Given an interpretation* $[\![_]\!]\colon \mathcal{L} \to \mathbf{Set}(X,[0,1])$*, we define the* logical distance *as follows (for every* $x,y \in X$*):*

$$\mathbf{d}_{\mathcal{L}}(x,y) \;=\; \sup_{\varphi\in\mathcal{L}} |[\![\varphi]\!](x) - [\![\varphi]\!](y)| \;=\; \sup_{\varphi\in\mathcal{L}} [\![\varphi]\!](x) - [\![\varphi]\!](y) \;. \tag{9}$$

Note that the absolute value is redundant since the negation operator $\neg$ *is in our logic* $\mathcal{L}$*.*

5.2 Endowing a topology on $\mathcal{L}$ through its shapes

As we are dealing with more than countably many function symbols, we will have to impose some structure on the set of function symbols in order to prove our expressivity theorem (cf. Theorem 28). Technically, we need a topology $\mathcal{T}_{\mathcal{L}}$ on $\mathcal{L}$ so that the theory map qTh (8) becomes *topologisable* in the following sense.

▶ **Definition 18.** *Let* $(X,\mathcal{A}) \in \mathbf{Ana}$*. Then the theory map* $\mathrm{qTh}\colon X \to \Omega^{\mathcal{L}}$ *is* topologisable *iff there is a topology* $\mathcal{T}_{\mathcal{L}}$ *on* $\mathcal{L}$ *such that the preimage of every open set* $U \subseteq \Omega^{\mathcal{L}}$ *(in the compact open topology on the function space* $\Omega^{\mathcal{L}}$*) is a universal measurable set, i.e.* $\mathrm{qTh}^{-1}(U) \in \overline{\mathcal{A}}$*.*

To be able to do this, we switch our focus from the uncountable language $\mathcal{L}$ to the countable language $\mathrm{Sh}(\mathcal{L})$ of *shapes* induced by the language $\mathcal{L}$. This language $\mathrm{Sh}(\mathcal{L})$ of shapes is constructed by collapsing all function symbols indexed by the same index set, i.e.

$$\top \mid \neg_ \mid _ \wedge _ \mid \diamond_a_, a \in A \mid f_i, i \in \omega. \tag{10}$$

To each formula φ on $\mathcal{L}$ one can assign a formula in $\mathrm{Sh}(\mathcal{L})$ by replacing each f_{iy} by f_i. This defines an equivalence relation on $\mathcal{L}$. For each $\psi \in \mathrm{Sh}(\mathcal{L})$ denote by $\widehat{\psi} \subseteq \mathcal{L}$ the corresponding equivalence class. For instance, for an n_i-ary function symbol f_i the set $\widehat{\psi}$ for $\psi = f_i(\top,\dots,\top)$ is in canonical bijection to Y_i, and the shape $r\neg\top \wedge r\top$ (when $\mathcal{L}$ allows for scalar multiplication) corresponds to $\Omega\times\Omega$, as each r ranges over $\Omega = [0,1]$. Through this equivalence relation we can subdivide $\mathcal{L}$ into countably many chunks, each of which associated to a finite product of Y_i's. If for a family $\{f_y\}_{y\in Y}$ of n-ary function symbols the set Y is endowed with a σ-algebra (resp. topology), we call $[\![_]\!]$ *jointly measurable* (resp. *jointly continuous*), if the function $[0,1]^n \times Y \to [0,1]$ is measurable (resp. continuous) with respect to the respective product σ-algebra (resp. product topology).

For the remainder of this subsection $\Omega^{(X,\mathcal{T})}$ will denote the set of continuous functions from $(X,\mathcal{T})$ to Ω, which will be endowed with the compact-open topology [14, 3.4] if not explicitly stated otherwise. Recall that a Hausdorff topological space $(X,\mathcal{T})$ is called *locally compact*, if each point admits a compact neighbourhood.

▶ **Lemma 19** ([14], 3.4.16). *Let* $(X,\mathcal{T})$ *be a locally compact second countable space. Moreover if* Ω *is second countable, then* $\Omega^{(X,\mathcal{T})}$ *is second countable w.r.t. the compact-open topology.*

▶ Remark 20. Lemma 19 will be used only once but at a crucial point in Theorem 28, which is our main result. One could ask if one can extend the classes of spaces for which Lemma 19 holds. For instance, does it hold for arbitrary second countable Hausdorff spaces? An ansatz would be to weaken the notion of compactness further. The promising notion is a k-space, the quotient space of some locally compact space, i.e. there is a quotient map $q\colon (Y,\mathcal{T}_Y) \to (X,\mathcal{T})$

for some locally compact space $(X, \mathcal{T})$ [14, p. 152]. The reason is that in this case there is a decomposition $\Omega^{(Y,\mathcal{T}_Y)} = \varprojlim_i \Omega^{K_i}$ for any directed system of of compact sets $K_i \subseteq X$ with $\bigcup_i K_i = X$ [14, 3.4.11]. Second countability follows as soon as i can be chosen to range over a countable set. Unfortunately, this already implies hemicompactness of $(Y, \mathcal{T}_Y)$ in case of regular spaces, cf. [40, 8.1(d)] and [14, 3.4.E(c)]. [40] also discusses other weakenings of compactness, but the mentioned Fact 8.1 therein do not provide a remedy.

We end this subsection by the main results of this section; namely that the logical distance $\mathbf{d}_\mathcal{L}$ is a predicate on $(X, \mathcal{A}) \times (X, \mathcal{A})$ and the logic $\mathcal{L}$ is adequate w.r.t. $\mathbf{bd}_c$.

▶ **Lemma 21.** *If* qTh *on* $(X, \mathcal{A})$ *is topologisable, then* $\mathbf{d}_\mathcal{L}$ *is measurable.*

▶ **Theorem 22.** *If the function symbols* f_{iy} *are interpreted by* $[\![_]\!]$ *as nonexpansive functions w.r.t.* d_∞*, then the logic* $\mathcal{L}$ *is adequate, i.e.* $\mathbf{bd} \geq \mathbf{d}_\mathcal{L}$.

Moreover, the language consisting of logical symbols, modalities, scalar addition, subtraction, multiplication, and convex combination is always adequate.

▶ Remark 23. With the aim to develop continuous version of first order logic, Yaacov and Berenstein studied metric structures and their model theory in [42]. Simply put, a metric structure consists of a complete bounded metric space (M, d) with a set of $\mathbb{R}$-valued predicates, a set of operators on the metric space (which are uniformly continuous of type $M^n \to M$ for some $n > 0$), and a set of distinguished elements of M. It is worthwhile to note that the operators defined by our signature (4) is a special case of a metric structure on the unit interval with an empty set of predicates.

5.3 A general expressivity theorem for $\mathcal{L}$

Our main expressivity result (cf. Theorem 28) is based on the well known Stone-Weierstraß theorem and Kantorovich-Rubinstein duality. This follows the tradition of expressivity results from recent papers [29, 1, 19] on coalgebraic modal logic; however, the proof of Theorem 28 does not follow from the abstract results established in the aforementioned articles. In contrast, we need an extra condition that the theory map qTh is topologisable (cf. Lemma 21).

▶ **Definition 24.** *A set* L *of functions* $X \to [0,1]$ approximates a function $f\colon X \to \mathbb{R}$ at a pair $x, y \in X$ up to ε *(for some* $\varepsilon > 0$*) if there is a* $g \in L$ *with* $|g(x) - f(x)|, |g(y) - f(y)| < \varepsilon$. *We further say that* L approximates f at x, y *if* L *approximates* f *at* x, y *for each* $\varepsilon > 0$.

Henceforth we write $[\![\mathcal{L}]\!] = \{\, [\![\varphi]\!] \mid \varphi \in \mathcal{L} \,\}$. It turns out that the operators (truth, conjunction, positive, and negative scaling) in our logic $\mathcal{L}$ approximate any predicate over a state space. The following lemma is taken from [7, Lemma 10].

The following lemma is a continuous (and thus simpler) version of [7, Lem. 10].

▶ **Lemma 25.** *Assume* $\top, _ \wedge _, _ + r, _ - r \in \mathcal{L}$ *for all* $r \in \Omega$ *and* $[\![_]\!]\colon \mathcal{L} \to \mathsf{Pred}(X, \mathcal{A})$ *an interpretation on predicates on some measurable space* $(X, \mathcal{A})$. *Let* $p \in \mathsf{Pred}(X, \mathcal{A})$, $x, y \in X$ *and* $\varepsilon > 0$. *Then*

$$\exists \varphi \in \mathcal{L}\colon 0 \leq p(x) - p(y) < [\![\varphi]\!](x) - [\![\varphi]\!](y) + \varepsilon \tag{11a}$$

$$\implies \big(\exists \psi \in \mathcal{L}\colon p(x) - [\![\psi]\!](x) \in [0, \varepsilon) \text{ and } p(y) - [\![\psi]\!](y) = 0\big) \tag{11b}$$

$$\implies [\![\mathcal{L}]\!] \text{ approximates } p \text{ at } x, y \text{ up to } \varepsilon. \tag{11c}$$

With the help of this lemma, we can approximate any short (aka nonexpansive) predicates w.r.t. logical distance $\mathbf{d}_{\mathcal{L}}$ by formulae in our logic $\mathcal{L}$. Given a measurable space $(X, \mathcal{A})$, we define the set of *short predicates with respect to logical distance* $\mathbf{d}_{\mathcal{L}}$ be

$$\mathsf{Pred}(X, \mathcal{A}, \mathbf{d}_{\mathcal{L}}) \coloneqq \{\, h \in \mathsf{Pred}(X, \mathcal{A}) \mid \forall x, y \in X \colon \mathbf{d}_{\mathcal{L}}(x, y) \geq h(x) - h(y) \,\}. \tag{12}$$

▶ **Corollary 26.** *As long as the scalar addition is allowed in the signature of $\mathcal{L}$ in (4), every short predicate can be approximated by the interpretation of modal formulae in $[\![\mathcal{L}]\!]$.*

Proof. Take $h \in \mathsf{Pred}(X, \mathcal{A}, \mathrm{d}_{[\![_]\!]})$. As $\neg_ \in \mathcal{L}$ we have for each $x, y \in X$ by (9) that $h(x) - h(y) \leq \sup_{f \in [\![\mathcal{L}]\!]} f(x) - f(y)$. This implies (11a) for any $\varepsilon > 0$. Thus $[\![\mathcal{L}]\!]$ approximates h at x, y for any $\varepsilon > 0$ by Lemma 25. Hence the claim follows. ◀

The next lemma is the well-known Kantorovich-Rubinstein duality extended to perfect measures. We provide a proof in the appendix, since the rather sketchy proof in [33, Thm. 5] considers only distances on their induced Borel-σ-algebra, while other known proofs, especially [12, 11.8.2&6], chose a topological instead of a purely measure theoretic set-up.

▶ **Lemma 27** (Kantorovich-Rubinstein theorem)**.** *Let $\mathfrak{m}, \mathfrak{n} \in \mathcal{G}(X, \mathcal{A})$ be perfect measures and $d \colon (X, \mathcal{A}) \times (X, \mathcal{A}) \to ([0, 1], \mathcal{B}_{[0,1]})$ a l.s.m. pseudo-metric such that $(X, \mathcal{A})$ is analytic (or smooth, or $\mathcal{T}_d$, the topology induced by d, is contained in $\overline{\mathcal{A}}$) and $d(_, x_0)$ is integrable for every $x_0 \in X$. Then*

$$\inf_{\mathfrak{c} \in K(\mathfrak{m}, \mathfrak{n})} \int d \,\mathrm{d}\mathfrak{c} = \sup_{h \in \mathsf{Pred}(X, \mathcal{A}, d)} \int h \,\mathrm{d}(\mathfrak{m} - \mathfrak{n}).$$

▶ **Theorem 28.** *Let $\mathcal{L}$ be a language with a coalgebra $\gamma \colon (X, \mathcal{A}) \to B_{\mathsf{MDP}}(X, \mathcal{A})$ so that $[\![_]\!]$ is well defined. Assume the following restrictions:*

1. *every measure $\gamma^0_{x,a}$ (for every $x \in X$, $a \in \Sigma$) is perfect,*
2. *the theory map* qTh *is topologisable by a locally compact and second countable topology, and*
3. *the scalar addition is in the signature of our language $\mathcal{L}$,*

then $\mathbf{d}_{\mathcal{L}}$ is a fixpoint of the functional $\sigma^{B_{\mathsf{MDP}}} \circ \gamma$. As a consequence, we have that the language $\mathbf{d}_{\mathcal{L}}$ is expressive w.r.t. $\mathbf{bd}$ (i.e., $\mathbf{bd} \leq \mathbf{d}_{\mathcal{L}}$).

Proof. Let $\mathfrak{m}_{x,a} = \gamma^0_{x,a}$ (for each $x \in X, a \in \Sigma$), let $r^a_x = \gamma^1_{a,x}$ and $r^a_{x,y} = |r^a_x - r^a_y|$. Recall the distance liftings $\hat{\sigma}$ and σ^{MRP} from Proposition 10 and Lemma 11, respectively. The claim $\mathbf{bd} \leq \mathbf{d}_{\mathcal{L}}$ is – using order preservation of $\sigma^{B_{\mathsf{MDP}}} \circ \gamma$, Theorem 13 – equivalent to $\forall \varepsilon > 0 \colon \sigma^{B_{\mathsf{MDP}}} \circ \gamma(x, y) \leq \mathbf{d}_{\mathcal{L}}(x, y) + \varepsilon$ for all $x, y \in X$. From the definition of $\sigma^{B_{\mathsf{MDP}}}$ this translates to the condition

$$\forall \varepsilon > 0 \colon \forall a \in \Sigma \colon \qquad \sigma^{\mathsf{MRP}}_X(\mathbf{d}_{\mathcal{L}}) \circ \gamma(x, y)((\mathfrak{m}_{x,a}, r^a_x), (\mathfrak{m}_{y,a}, r^a_y)) \leq \mathbf{d}_{\mathcal{L}} + \varepsilon. \tag{13}$$

To this end, we begin by expanding the left hand side of the above inequality:

$$\begin{aligned} &\sigma^{\mathsf{MRP}}_X(\mathbf{d}_{\mathcal{L}}) \circ \gamma(x, y)((\mathfrak{m}_{x,a}, r^a_x), (\mathfrak{m}_{y,a}, r^a_y)) \\ &\qquad = \inf_{\mathfrak{c} \in K(\mathfrak{m}_{x,a}, \mathfrak{m}_{y,a})} \int \mathbf{d}_{\mathcal{L}} \,\mathrm{d}\mathfrak{c} +_c r^a_{xy} \end{aligned} \tag{14}$$

we push this expression to $[0, 1]^{\mathcal{L}}$ by letting $\widetilde{\mathbf{d}_{\mathcal{L}}} \coloneqq \sup_{\varphi \in \mathrm{im}(\mathrm{qTh})}$ be a distance on $[0, 1]^{\mathcal{L}}$

$$= \inf_{\mathfrak{c} \in K(\mathrm{qTh}_*(\mathfrak{m}_{x,a}), \mathrm{qTh}(\mathfrak{m}_{y,a}))} \int \widetilde{\mathbf{d}_{\mathcal{L}}} \,\mathrm{d}\mathfrak{c} +_c r^a_{xy} \tag{15}$$

As $[0,1]^{\mathcal{L}}$ is second countable, cf. Lemma 19, we can (depending on $\mathrm{qTh}_*(\mathfrak{m}_{x,a})$ and $\mathrm{qTh}(\mathfrak{m}_{y,a})$) restrict the integral to $A \times A$ for some standard (and thus analytic) subspace $A \in \mathcal{B}_{[0,1]^{\mathcal{L}}}$ [15, Thm. 6] (using that $\mathrm{qTh}_*(\mathfrak{m}_{x,a})$ and $\mathrm{qTh}(\mathfrak{m}_{y,a})$ again are perfect). (Note that this argument actually requires only second countability of $\mathcal{L}$ as it can be refined using [14, 3.8.D], [3, 2.1.15] and [15, § 8 Remark].) As $\mathbf{d}_{\mathcal{L}}$ is bounded, so $\mathbf{d}_{\mathcal{L}}(_, x)$ is certainly integrable for any $x \in X$. Thus we can finally apply Kantorovic-Rubinstein duality Lemma 27, using Lemma 21 and Item 1

$$= \sup_{h \in \mathsf{Pred}([0,1]^{\mathcal{L}}, \mathcal{B}_{[0,1]^{\mathcal{L}}}, \widetilde{\mathbf{d}_{\mathcal{L}}})} \int h \,\mathrm{d}(\mathfrak{m}_{x,a} - \mathfrak{m}_{y,a}) +_c r^a_{xy} \tag{16}$$

$$\leq \sup_{h \in \mathsf{Pred}(X, \mathcal{A}, \mathbf{d}_{\mathcal{L}})} \int h \,\mathrm{d}(\mathrm{qTh}_*(\mathfrak{m}_{x,a}) - \mathrm{qTh}_*(\mathfrak{m}_{y,a})) +_c r^a_{xy} \tag{17}$$

By applying Corollary 26 we can approximate any short predicate h by the interpretation of formulae in our logic $[\![\mathcal{L}]\!]$.

Define $\mathcal{S} \subseteq \mathcal{A}$ to be $\mathcal{S} = \left\{ U, U^{\complement} \;\middle|\; U \in \mathrm{qTh}^{-1}[\mathcal{S}_{[\![_]\!]}] \cup h^{-1}[\mathcal{S}_\Omega] \right\}$. Let $\mathcal{T}$ denote the topology generated by $\mathcal{S}$ and $(X', \mathcal{T}')$ the Kolmogorov quotient, cf. Section C, of $(X, \mathcal{T})$ with unit map $\eta\colon (X, \mathcal{T}) \to (X', \mathcal{T}')$. Both $\mathfrak{m}^\flat_{z,a} \coloneqq \mathcal{G}(\eta)(\mathfrak{m}_{z,a})$ (for $z \in \{x, y\}$) are perfect since the push-forward measures of perfect measures [20, 451Ea] is perfect and using 1 we know $\mathfrak{m}_{z,a}$ is perfect. By [20, 451M] both measures are inner regular with respect to compact sets; thus we find compact sets $K_x, K_y \subseteq X'$ with $\mathfrak{m}^\flat_{x,a}((K_x)^{\complement}), \mathfrak{m}^\flat_{y,a}((K_y)^{\complement}) < \delta$. Thus $K_x \cup K_y$ is compact and so is $K \coloneqq \eta^{-1}(K_x \cup K_y)$ by Lemma 38. Moreover we have $\mathfrak{m}_{x,a}((K)^{\complement}), \mathfrak{m}_{y,a}((K_y)^{\complement}) < \delta$. As $\mathcal{S}$ is closed under complement, $(X', \mathcal{T}')$ is R_2.

So finally, the Stone-Weierstraß Theorem, Lemma 37, is applicable to $(K, \mathcal{T}|_K)$; thus, every nonexpansive predicate h can be approximated on K by a function from the family $\{\, [\![\varphi]\!]|_K \mid \varphi \in \mathcal{L} \,\}$. Let $\varphi_{h,\delta} \in \mathcal{L}$ denote a witness of a δ-approximation from $[\![\mathcal{L}]\!]|_K$ of $h|_K$.

Continuing at (17) we obtain (for all $\delta > 0$):

$$\begin{aligned}
&\sigma(\mathbf{d}_{\mathcal{L}})(\mathfrak{m}_{x,a}, \mathfrak{m}_{y,a}) +_c r^a_{x,y} \\
&\quad\leq \left(\sup_{h \in \mathsf{Pred}(X, \mathcal{A}, \mathbf{d}_{\mathcal{L}})} \int_K h \,\mathrm{d}(\mathfrak{m}_{x,a} - \mathfrak{m}_{y,a}) + \int_{K^{\complement}} h \,\mathrm{d}(\mathfrak{m}_{x,a} - \mathfrak{m}_{y,a}) \right) +_c r^a_{xy} \\
&\quad\leq \left(\sup_{h \in \mathsf{Pred}(X, \mathcal{A}, \mathbf{d}_{\mathcal{L}})} \int_K [\![\varphi_{h,\delta}]\!] \,\mathrm{d}(\mathfrak{m}_{x,a} - \mathfrak{m}_{y,a}) + \delta + \delta \cdot 1 \right) +_c r^a_{xy} \\
&\quad\leq \left(\sup_{h \in \mathsf{Pred}(X, \mathcal{A}, \mathbf{d}_{\mathcal{L}})} \int [\![\varphi_{h,\delta}]\!] \,\mathrm{d}(\mathfrak{m}_{x,a} - \mathfrak{m}_{y,a}) + 3\delta \right) +_c r^a_{xy} \\
&\quad\leq \left(\sup_{\varphi \in \mathcal{L}} \int [\![\varphi]\!] \,\mathrm{d}(\mathfrak{m}_{x,a} - \mathfrak{m}_{y,a}) + 3\delta \right) +_c r^a_{xy} \\
&\quad= \left(\sup_{\varphi \in \mathcal{L}} \left(\int [\![\varphi]\!] \,\mathrm{d}\mathfrak{m}_{x,a} +_c r^a_x \right) - \left(\int [\![\varphi]\!] \,\mathrm{d}\mathfrak{m}_{y,a} +_c r^a_y \right) \right) + 3c\delta \qquad (18) \\
&\quad= \sup_{\varphi \in \mathcal{L}} [\![\diamond_a \varphi]\!](x) - [\![\diamond_a \varphi]\!](y) + 3c\delta \\
&\quad\overset{(9)}{\leq} \mathbf{d}_{\mathcal{L}}(x, y) + 3c\delta.
\end{aligned}$$

Choosing $3c\delta < \varepsilon$, (13) follows – finishing the proof. ◀

It should be noted that the restrictions on perfect measures in the above theorem is redundant when the coalgebra map $\gamma \in \mathbf{Ana}$. Moreover, the second restriction from the previous theorem can also be discarded by imposing the following restrictions on the function symbols $(f_{iy})_{i \in \omega, y \in Y_i}$, which belong to the signature of our language $\mathcal{L}$.

▶ **Theorem 29.** *Assume that $\mathcal{L}$ given in* (4) *is such that each family Y_i is endowed with a second countable Hausdorff topology $\mathcal{T}_i$ and the interpretation of function symbols f_{iy} are jointly continuous with respect to $\mathcal{T}_i$, then* qTh *is topologisable by a second countable Hausdorff topology. This topology can be chosen to be locally compact, if each $\mathcal{T}_i$ has this property.*

Recalling that $[0,1]$ is compact, thus the assumptions of Theorem 29 are fulfilled.

▶ **Lemma 30.** *For a language $\mathcal{L}$ with the following signature in which the set Σ of actions is countable, the theory map* qTh *is topologisable.*

$$\mathcal{L} ::= \wedge \mid \neg \mid \top \mid _ + r, r \in [0,1] \mid _ - r, r \in [0,1] \mid \diamond_a, a \in \Sigma.$$

Now combining Lemma 30 and Theorems 22 and 28 we get that the modal language $\mathcal{L}$ defined in Lemma 30 is both adequate and expressive for bisimulation pseudo-metrics $\mathbf{bd}$ defined in Corollary 16 for MDPs.

▶ **Corollary 31.** *Let $\mathcal{L}$ be the language as given in Lemma 30 and let $\gamma\colon (X,\mathcal{A}) \to B_{\mathsf{MDP}}(X,\mathcal{A}) \in \mathbf{Ana}$ be an MDP. Then the bisimulation pseudometric (defined in Corollary 16) coincides with the logical distance $\mathbf{d}_{\mathcal{L}}$.*

One may anticipate, following [7], to decompose the semantics of our diamond modality into two modalities: one modelling the expectation modality $\diamond'_a\varphi$ and the 0-ary reward modality rew_a. The semantics of these two modalities given below in $\mathcal{L}'$ is taken from [7]. We argue next that this is unfortunately not possible without jeopardizing the adequacy result.

▶ Remark 32. Assume $c \in [0,1]$ and a language $\mathcal{L}'$ with signature

$$\top \mid \neg _ \mid _ \wedge _ \mid \mathrm{rew}_a, a \in \Sigma \mid \diamond'_a, a \in \Sigma \mid r\,_, r \in [0,1] \mid _ + _ \tag{19}$$

and an interpretation depending on $\gamma\colon (X,\mathcal{A}) \to B_{\mathsf{MDP}}(X,\mathcal{A}) \in \mathbf{Ana}$ defined by:

$$[\![\diamond_a\varphi]\!]_\gamma(x) = c\int [\![\varphi]\!]\,\mathrm{d}\gamma^0_{a,x} \qquad \text{and} \qquad [\![\mathrm{rew}_a]\!](x) = \gamma^1_{a,x}.$$

Then the language $\mathcal{L}'$ is expressive w.r.t. $\mathbf{bd}$ (i.e., $\mathbf{bd} \leq \mathbf{d}_{\mathcal{L}'}$). However, $\mathcal{L}'$ is not adequate since the binary (truncated) addition is not nonexpansive w.r.t. suprema distance.

6 Related work and concluding remarks

6.1 Related work

Our work is inspired by [18] and establishes a quantitative version of Hennessy-Milner theorem for the therein defined bisimulation pseudometric. To the best of our knowledge, such a generalisation is novel and has not been studied elsewhere in the literature. The key technical differences between the two works are as follows. First, our notion of conformance on continuous state MDPs is based on universal measurability; whilst, it is based on lower semi-continuity in [18]. Note that every lower semi-continuous function is universally measurable. Second, our MDPs are coalgebras living in **Ana** and the state space of an MDP is thus an analytic space in our paper; whilst, it is a Polish space in [18]. Third, the bisimulation pseudometric $\mathbf{bd}$ defined in this paper is based on Wasserstein lifting; whilst, the bisimulation pseudometric (denoted $\mathbf{bd}_{FPP}$) of Ferns et al. is based on Kantorovich lifting. Note that

the pseudometrics $\mathbf{bd}$ and $\mathbf{bd}_{FPP}$ are equivalent due to the Kantorovich-Rubinstein duality (Lemma 27). Finally, we employ the Kleene's fixpoint theorem to define $\mathbf{bd}$, whilst, Ferns et al. employed Banach fixed point theorem to define their bisimulation pseudometric $\mathbf{bd}_{FPP}$.

The recent work of Chen et al. [7] on continuous time Markov processes (i.e. a family of B_{MRP}-coalgebras indexed by non-negative real numbers) is also insightful, where a quantitative version of Hennessy-Milner theorem is also proven. The mathematical development followed in [7] is worth comparing, especially when this family is restricted to a singleton coalgebra (say, for instance, $\gamma\colon (X,\mathcal{A}) \to B_{\mathsf{MRP}}(X,\mathcal{A})$) and the σ-algebra $\mathcal{A}$ is generated by a Polish topology on X. The functional $\mathcal{F}_c$ (for some $0 < c < 1$) defined in [7] is not an endofunction in general on the lattice of lower semi-continuous functions on $(X,\mathcal{A})$. Using the notations of this paper, $\mathcal{F}_c$ can be rewritten as: $\mathcal{F}_c(d)(x,y) = c \cdot \hat{\sigma}(d)(\gamma^0(x),\gamma^0(y))$ for every $x, y \in X$.

Nevertheless, to capture their bisimulation pseudometric (denoted $\mathbf{bd}_{CCP}$) by a fixpoint argument, the authors had to work with continuous distance functions on X. The usual Knaster-Tarski fixpoint theorem is inapplicable and the authors constructed $\mathbf{bd}_{CCP}$ as the limit of following pseudometrics δ_i: $\delta_0 = (\gamma^1 \times \gamma^1)^*(d_E)$; $\delta_{i+1} = \mathcal{F}_c(\delta_i)$. As a result, the two bisimulation pseudometrics $\mathbf{bd}_{CCP}$ and $\mathbf{bd}$ (Lemma 11) are different.

The recent works [1, 30, 29, 19] on developing expressive modal logic for a behavioural conformance that are defined by codensity lifting (called Kantorovich lifting in [19]) can, unfortunately, not be directly applied to the current setting. This is due to the underlying assumption of behavioural conformances defined internally in a complete lattice fibration (or equivalently using the language of topological functors [19]). To this end, we adopted a coupling-based lifting approach (inspired from [4]) to define our bisimulation pseudometric. This adoption required significant effort in recasting old results from measure theory in our framework as outlined in Assumptions **A1**-**A4**.

6.2 Concluding remarks

To summarise, we model both MRPs and MDPs with continuous state spaces as coalgebras in **Ana** and define the notion of bisimulation pseudometric using the well known Kleene's fixpoint theorem. The latter was based on a given coalgebra $\gamma\colon (X,\mathcal{A}) \to \mathcal{G}(X,\mathcal{A}) \in \mathbf{Ana}$ and the fact that a functional $\gamma \circ \hat{\sigma}\colon \mathsf{Pred}(X,\mathcal{A}) \to \mathsf{Pred}(X,\mathcal{A})$ is ω-cpo-continuous (Theorem 13), whose proof was in turn based on classical results from functional analysis. In addition, we also presented a "quantitative" modal logic $\mathcal{L}$ whose formulae are interpreted as universally measurable predicates over the state space of an MDP and the logical distance $\mathbf{d}_{\mathcal{L}}$ generated by $\mathcal{L}$ coincides with the bisimulation distance $\mathbf{bd}$. To prove the expressivity result (Theorem 28) is, certainly, more involved than the adequacy result (Theorem 22); nonetheless, they both require that the theory map qTh (8) is topologisable.

For future work the fact that in the expressivity result a topological structure on the formulas instead of any requirement on the statement was the key assumption may stipulate new perspectives. A more concrete worthwhile enterprise would be to generalise the Stone-Weierstraß theorem for measurable spaces. This will help in directly invoking the argument to approximate a nonexpansive map h by logical formulae in the proof of Theorem 28; thus, avoiding the topological arguments used here.

References

1 Harsh Beohar, Sebastian Gurke, Barbara König, Karla Messing, Jonas Forster, Lutz Schröder, and Paul Wild. Expressive quantale-valued logics for coalgebras: An adjunction-based approach. In Olaf Beyersdorff, Mamadou Moustapha Kanté, Orna Kupferman, and Daniel Lokshtanov, editors, *41st International Symposium on Theoretical Aspects of Computer Science (STACS 2024)*, volume 289, pages 10:1–10:19, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.STACS.2024.10.

2 Harsh Beohar, Clemens Kupke, and Daniel Luckhardt. Henneysey-milner type theorem for measurable pseudometrics, 2025. arXiv:2505.23635 [cs.LO].

3 Vladimir I. Bogachev. Measures on topological spaces. *Journal of Mathematical Sciences*, 91(4):3033–3156, 1998.

4 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)*, volume 118 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 17:1–17:17, Dagstuhl, Germany, 2018. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CONCUR.2018.17.

5 Craig Boutilier, Thomas Dean, and Steve Hanks. Decision-theoretic planning: structural assumptions and computational leverage. *J. Artif. Int. Res.*, 11(1):1–94, July 1999. doi:10.1613/JAIR.575.

6 D. W. Bressler and M. Sion. The current theory of analytic sets. *Canadian Journal of Mathematics*, 16:207–230, 1964. doi:10.4153/CJM-1964-021-7.

7 Linan Chen, Florence Clerc, and Prakash Panangaden. A behavioural pseudometrics for continuous-time Markov processes. In *Foundations of Software Science and Computation Structures*, Lecture Notes in Computer Science, 2025. Accepted; arXiv:2501.13008 [cs.LO]; Event: FoSSaCS, Hamilton, Ontario, Canada, May 5–8, 2025. doi:10.48550/arXiv.2501.13008.

8 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labeled Markov systems. In Jos C. M. Baeten and Sjouke Mauw, editors, *CONCUR'99 Concurrency Theory*, pages 258–273, Berlin, Heidelberg, 1999. Springer. doi:10.1007/3-540-48320-9_19.

9 Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. *Information and Computation*, 179(2):163–193, 2002. doi:10.1006/inco.2001.2962.

10 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004. doi:10.1016/j.tcs.2003.09.013.

11 Ernst-Erich Doberkat. Measures and all that — a tutorial, 2014. arXiv:1409.2662 [math.FA], Version 3.

12 R. M. Dudley. *Real Analysis and Probability*. Number 74 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2002.

13 Nelson Dunford and Jacob Schwartz. *Linear Operators*, volume 3 of *Pure and Applied Mathematics*. Wiley-InterScience, 1958/1971.

14 Ryszard Engelking. *General topology*, volume 6 of *Sigma series in pure mathematics*. Heldermann Verlag, revised and completed edition edition, 1989.

15 Arnold M. Faden. The existence of regular conditional probabilities: Necessary and sufficient conditions. *The Annals of Probability*, 13(1):288–298, 1985.

16 Neil Falkner. Generalizations of analytic and standard measurable spaces. *Mathematica Scandinavica*, pages 283–301, 1981.

17 Norm Ferns, Prakash Panangaden, and Doina Precup. Metrics for finite Markov decision processes. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence (UAI 2004)*, pages 162–169, 2004. arXiv:1207.4114 [cs.AI]. URL: https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1&smnu=2&article_id=1103&proceeding_id=20.

18 Norm Ferns, Prakash Panangaden, and Doina Precup. Bisimulation metrics for continuous Markov decision processes. *SIAM Journal on Computing*, 40(6):1662–1714, 2011. doi:10.1137/10080484X.

19 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, *31st EACSL Annual Conference on Computer Science Logic (CSL 2023)*, volume 252, pages 22:1–22:20, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CSL.2023.22.

20 D.H. Fremlin. *Measure Theory*, volume 5. Torres Fremlin, 2000/2008.

21 Carles Gelada, Saurabh Kumar, Jacob Buckman, Ofir Nachum, and Marc G. Bellemare. DeepMDP: Learning continuous latent space models for representation learning. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 2170–2179. PMLR, June 2019. URL: https://proceedings.mlr.press/v97/gelada19a.html.

22 Michèle Giry. A categorical approach to probability. In B. Banaschewski, editor, *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer, 1982.

23 Robert Givan, Thomas Dean, and Matthew Greig. Equivalence notions and model minimization in Markov decision processes. *Artificial Intelligence*, 147(1):163–223, 2003. Planning with Uncertainty and Incomplete Information. doi:10.1016/S0004-3702(02)00376-4.

24 Boris Vladimirovich Gnedenko and Andrey Nikolaevich Kolmogoroff. *Predelnye raspredeleniya dlya summ nezavisimykh sluchaynykh velichin*. GITTL, 1949.

25 Matthew Hennessy and Robin Milner. On observing nondeterminism and concurrency. In Jaco de Bakker and Jan van Leeuwen, editors, *Automata, Languages and Programming*, pages 299–309, Berlin, Heidelberg, 1980. Springer Berlin Heidelberg. doi:10.1007/3-540-10003-2_79.

26 B. P. F. Jacobs. *Categorical Logic and Type Theory*. Number 141 in Studies in Logic and the Foundations of Mathematics. North Holland, Amsterdam, 1999.

27 Alexander Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer Science & Business Media, 2012.

28 Henning Kerstan. *Coalgebraic Behavior Analysis: From Qualitative To Quantitative Analyses*. PhD thesis, Universität Duisburg-Essen, May 2016. Submitted on 2016-05-09. URL: https://duepublico2.uni-due.de/receive/duepublico_mods_00041220.

29 Yuichi Komorida, Shin-ya Katsumata, Clemens Kupke, Jurriaan Rot, and Ichiro Hasuo. Expressivity of quantitative modal logics: Categorical foundations via codensity and approximation. In *Proceedings of the Thirty Sixth Annual IEEE Symposium on Logic in Computer Science (LICS 2021)*, pages 1–14, Rome, Italy, June 2021. IEEE Computer Society Press. doi:10.1109/LICS52264.2021.9470656.

30 Clemens Kupke and Jurriaan Rot. Expressive Logics for Coinductive Predicates. In Maribel Fernández and Anca Muscholl, editors, *28th EACSL Annual Conference on Computer Science Logic (CSL 2020)*, volume 152 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:18, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CSL.2020.26.

31 Jan K Pachl. Two classes of measures. *Colloquium Mathematicum*, 42(1):331–340, 1979. Erratum in Vol. 45.2 (1981), pp. 331–333.

32 D. Ramachandran and L. Rüschendorf. On the monge–kantorovich duality theorem. *Teoriya veroyatnostey i ee primeneniya*, 45(2), 2000.

33 Doraiswamy Ramachandran and Ludger Rüschendorf. A general duality theorem for marginal problems. *Probability Theory and Related Fields*, 101:311–319, 1995.

34 Czesław Ryll-Nardzewski. On quasi-compact measures. *Fundamenta Mathematicae*, 40:125–130, 1953.

35 Stephen Simons. *Minimax Theorems and Their Proofs*, chapter 1, pages 1–23. Number 5 in Nonconvex Optimization and Its Applications. Springer, 1995. doi:10.1007/978-1-4613-3557-3.

36 M. Ya. Suslin. Sur une définition des ensembles mesurables B sans nombres transfinis. *Comptes Rendus Mathématique. Académie des Sciences. Paris.*, 164(2):88–91, 1917.

37 Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, MA, USA, 2018.

38 Franck van Breugel and James Worrell. An algorithm for quantitative verification of probabilistic transition systems. In Kim G. Larsen and Mogens Nielsen, editors, *CONCUR 2001 — Concurrency Theory*, pages 336–350, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg. doi:10.1007/3-540-44685-0_23.

39 Franck van Breugel and James Worrell. Towards quantitative verification of probabilistic transition systems. In Fernando Orejas, Paul G. Spirakis, and Jan van Leeuwen, editors, *Automata, Languages and Programming*, pages 421–432, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg. doi:10.1007/3-540-48224-5_35.

40 Eric K. van Douwen. The integers and topology. In *Handbook of set-theoretic topology*, pages 111–167. Elsevier, 1984.

41 Ignacio D. Viglizzo. Final sequences and final coalgebras for measurable spaces. In *Algebra and Coalgebra in Computer Science*, pages 395–407, Berlin, Heidelberg, 2005. Springer. doi:10.1007/11548133_25.

42 Itaï Ben Yaacov, Alexander Berenstein, C. Ward Henson, and Alexander Usvyatsov. Model theory for metric structures. In Zoé Chatzidakis, Dugald Macpherson, Anand Pillay, and Alex Wilkie, editors, *Model Theory with Applications to Algebra and Analysis*, volume 350 of *London Mathematical Society Lecture Note Series*, pages 315–427. Cambridge University Press, Cambridge, 2008.

43 Amy Zhang, Rowan Thomas McAllister, Roberto Calandra, Yarin Gal, and Sergey Levine. Learning invariant representations for reinforcement learning without reconstruction. In *International Conference on Learning Representations*, 2021. URL: https://openreview.net/forum?id=-2FCwDKRREu.

A Notations and background

The power set of a set X is denoted by $\mathcal{P}(X)$. For a function $f\colon X \to Y$ we denote for a subset $A \subseteq X$ its direct image under f by $f_*(U)$ or $f[U]$ and for a $V \subseteq Y$ the inverse image by $f^{-1}(V)$.

A.1 Perfect measures

Perfect measures were introduced by Kolmogorov [24, 22–23]. The aim was to provide a convenient subclass of measures general enough for all applications. There are many different equivalent definitions. We choose the following one: A measure space $(X, \mathcal{A}, \mathfrak{m})$ is called *perfect*, if for any separable metrisable space $(Y, \mathcal{T})$ and every measurable map $f\colon X \to Y$ we have the following property: For every $A \in \mathcal{A}$ and $r < \mathfrak{m}(A)$ there is a compact set $K \subseteq \operatorname{im} f$ with $\mathfrak{m}(A \cap f^{-1}K) \geq r$, cf. [20, 451O(a)]. Also $\mathfrak{m}$ is called *perfect* in this case. As a direct consequence of the definition, perfectness is functorial, i.e. the push-forward of a perfect measure is perfect.

In typical real-world applications two points of a measurable space should only be distinguished by the σ-algebra if they can be distinguished by an observation. As only a finite amount of observations with limited precision can be made per unit of time, there should be a countable subset $\mathcal{S} \subseteq \mathcal{A}$ distinguishing as strong a $\mathcal{A}$ does ($\forall x, y \in X \forall A \in \mathcal{A}\colon x \in A \wedge y \notin A \implies (\exists S \in \mathcal{S}\colon \#S \cap \{x, y\} = 1)$). A measurable space enjoying this property is called *countably fibered*. Perfect countably fibered probability spaces are actually – despite being a very general notion, including, e.g. analytic spaces – quite close to standard spaces [15, § 8 Rem.]: Any such space is almost pre-standard with respect to some sub-σ-algebra $\mathcal{A}' \subseteq \mathcal{A}$. Almost pre-standard means that a space is standard when restricted to a Borel set of full measure and identifying all point not distinguished by $\mathcal{A}$ [15]. For countably generated spaces perfectness can even be characterised equivalently by being almost pre-standard [15, Thm. 6].

A.2 Suslin operation and smooth spaces

For a function $f_{(_)}\colon \omega \to X$ let $f_{(_)}|_{\leq i}$ denote the restriction to the first k indices. Further let $\omega^{<\omega}$ denote the set of all finite sequences in ω. The Suslin operation, cf. [27, 25.4] or [20, 421B], is denoted by $\mathfrak{S}$; we recall that it is defined by

$$\mathfrak{S}\,\mathcal{P} = \bigcup_{n_{(_)}\in\omega^\omega} \bigcap_{i\in\omega} A_{n_{(_)}|_{\leq i}} \tag{20a}$$

for a Suslin scheme $A_{(_)}\colon \omega^{<\omega}$. Further, we remind of its elementary properties: that for any paving $\mathcal{P}$ and countable family $\mathcal{S} \subseteq \mathfrak{S}\,\mathcal{P}$

$$\bigcup \mathcal{S} \in \mathfrak{S}\,\mathcal{P} \tag{20b}$$

$$\bigcap \mathcal{S} \in \mathfrak{S}\,\mathcal{P} \tag{20c}$$

[20, 421E], also for any function $f\colon X \to Y$ and paving $\mathcal{Q}$ on Y that

$$f^{-1}[\mathfrak{S}\,\mathcal{Q}] = \mathfrak{S}(f^{-1}[\mathcal{Q}]) \tag{20d}$$

[20, 421Cc], as well as monotonicity [20, 421Ca] and idempotence [20, 421D], that for any paving $\mathcal{P}$

$$\mathcal{P} \subseteq \mathfrak{S}\,\mathcal{P} \tag{20e}$$

$$\mathfrak{S}(\mathfrak{S}\,\mathcal{P}) = \mathfrak{S}\,\mathcal{P}. \tag{20f}$$

Denote the Giry monad by $\mathcal{G}(X,\mathcal{A}) = (\mathrm{M}_{(X,\mathcal{A})}, \mathrm{A}_{(X,\mathcal{A})})$.

We also recall a generalisation of analytic spaces: smooth spaces as introduced by Falkner [16]. It can be defined as follows [16, 1.3]:

▶ **Definition 33.** *A measurable space $(X,\mathcal{A})$ is called* smooth *if for any measurable space $(X,\mathcal{A})$ and any $A \in \mathfrak{S}(\mathcal{B}\otimes\mathcal{A})$ the projection on the first component $\mathrm{pr}_Y A$ is in $\mathfrak{S}\,\mathcal{B}$.*

Note the following fact [16, 1.3]:

▶ **Lemma 34.** *Every analytic space is smooth.*

B Proof of Theorem 13

The proof of Theorem 13 is based on Sion's minmax theorem [35, Thm. 3] which we recall first.

▶ **Lemma 35.** *Let U be a convex subset of a linear topological space, V a compact convex subset of a linear topological space, and $f\colon U\times V \to \mathbb{R}$ be upper semicontinuous on U and lower semicontinuous on V. Suppose that*

$$\forall y \in V, \lambda\in\mathbb{R}\colon \qquad \{\, x\in U \mid f(x,y) \geq \lambda \,\} \text{ is convex and} \tag{21a}$$

$$\forall x \in U, \lambda\in\mathbb{R}\colon \qquad \{\, y\in V \mid f(x,y) \leq \lambda \,\} \text{ is convex.} \tag{21b}$$

Then $\inf\limits_{y\in V}\sup\limits_{x\in U} f(x,y) = \sup\limits_{x\in U}\inf\limits_{y\in V} f(x,y)$.

In addition, we also need some important results from functional analysis; in particular, non-topological version of Riesz–Markov–Kakutani representation theorem [13, IV.5.1] and Banach-Alaoglu theorem [13, V.4.2].

The vector space $\mathrm{B}(X)$ of bounded functions $f\colon X \to \mathbb{R}$ is endowed with the uniform norm given by

$$\|f\|_{\mathrm{u}} = \sup\{\, |f(x)| \mid x \in X \,\}.$$

It is well-known that this space $\mathrm{B}(X)$ is complete [13, IV.5].

Let $\mathrm{B}(X,\mathcal{A})$ denote the Banach space [13, IV.5] consisting of limits of simple functions on $(X,\mathcal{A})$ with respect to the uniform norm $\|_\|_{\mathrm{u}}$. It is also functorial. Moreover, we recall that B^* denotes the dual of a Banach space B, i.e. all bounded linear functionals $B \to \mathbb{R}$. This dual can be either endowed with the uniform norm obtaining a Banach space or with the weak-* topology. The assignment $(_)^*$ is functorial with respect to both choices. We write the dual vector space of $\mathrm{B}(X,\mathcal{A})$ simply as $\mathrm{B}^*(X,\mathcal{A})$.

Recall that a *charge* is the same thing as a measure but only finitely additive. Denote by $\mathrm{Charge}(X,\mathcal{A})$ the set of all charges on $(X,\mathcal{A})$. We may also view Charge as a functor $\mathbf{Meas} \to \mathbf{Meas}$ by endowing $\mathrm{Charge}(X,\mathcal{A})$ with the same σ-algebra as $\mathcal{G}(X,\mathcal{A})$, i.e. the one making all ev_A with $A \in \mathcal{A}$ measurable. There is the following classical duality between charges and positive functionals given by the isometric (with respect to the uniform norm) isomorphism [13, IV.5.1]

$$\mathrm{Int}\colon \mathrm{Charge}(X,\mathcal{A}) \to \mathrm{B}^*(X,\mathcal{A}), \quad \mathfrak{m} \mapsto \lambda\, f.\, \int f \,\mathrm{d}\mathfrak{m}. \tag{22}$$

The above duality given by Int can be viewed as a non-topological version of the Riesz-Markov-Kakutani representation theorem. Note that probability charges $\mathfrak{m}$ are characterised by being positive, i.e. $\int f \,\mathrm{d}\mathfrak{m} = \mathrm{Int}(\mathfrak{m})(f) \geq 0$ for any $f \in \mathrm{B}(X,\mathcal{A})$ with $f \geq 0$, and normed, i.e. $\int 1 \,\mathrm{d}\mathfrak{m} = \mathrm{Int}(\mathfrak{m})(1) = 1$.

We are now ready to prove Theorem 13. Actually, we prove the following slight generalisation (recall that every probability measure on an analytic space is perfect).

▶ **Theorem 36.** *Let $\gamma\colon (X,\mathcal{A}) \to \mathcal{G}(X,\mathcal{A}) \in \mathfrak{m}$ such that $\gamma(x)$ is a perfect measure (for each $x \in X$). Then the $\hat{\sigma}$ is ω-cpo-continuous with respect to $\leq$.*

Proof. Assume an $\leq$-increasing sequence $(d_i)_{i\in\mathbb{N}}$ of pseudometrics over $(X,\mathcal{A})$, then we need to show that the following equation for every $x, y \in X$.

$$\inf_{\mathfrak{c}\in K(\gamma_x,\gamma_y)} \int \sup_{i\in\mathbb{N}} d_i \,\mathrm{d}\mathfrak{c} = \sup_{i\in\mathbb{N}} \inf_{\mathfrak{c}\in K(\gamma_x,\gamma_y)} \int d_i \,\mathrm{d}\mathfrak{c}. \tag{23}$$

Let $x, y \in X$ and let $\mathfrak{m}_1 = \gamma_x$ and $\mathfrak{m}_2 = \gamma_y$. Recall the notion of charge which is the same thing as a measure but only finitely additive and note that

$$\begin{aligned}&\{\, \mathfrak{c} \text{ measure on } \mathcal{A}\otimes\mathcal{A} \mid \mathcal{G}(\mathrm{pr}_1)\mathfrak{c} = \mathfrak{m}_1, \mathcal{G}(\mathrm{pr}_2)\mathfrak{c} = \mathfrak{m}_2 \,\}\\ &\quad = \{\, \mathfrak{c} \text{ charge on } \mathcal{A}\otimes\mathcal{A} \mid \mathcal{G}(\mathrm{pr}_1)\mathfrak{c} = \mathfrak{m}_1, \mathcal{G}(\mathrm{pr}_2)\mathfrak{c} = \mathfrak{m}_2 \,\}\end{aligned}$$

as $\mathfrak{m}_1, \mathfrak{m}_2$ are perfect [32, D5] (see also [31, Prop. 3], original result [34, Thm. VIII]). Now applying the duality (22) and by suppress $\mathcal{A}$ below in favour of readability (i.e. $X \times X$ as a shorthand for the measurable space $(X,\mathcal{A}) \times (X,\mathcal{A})$).

$$\begin{aligned}&\overset{\mathrm{Int}}{\cong} \left\{ c \in (\mathrm{B}^*(X\times X)) \;\middle|\; \begin{array}{l} c \text{ is positive}, c(1) = 1 \text{ and} \\ \quad \text{for } i = 1,2\colon (\mathrm{B}^*\,\mathrm{pr}_i)c = \mathrm{Int}(\mathfrak{m}_i) \end{array} \right\}\\ &= \bigcap \Big\{ {\mathrm{ev}_f}^{-1}([0,\infty)) \;\Big|\; f \in \mathrm{B}\big(X\times X\big), f \geq 0 \,\Big\} \cap {\mathrm{ev}_1}^{-1}(\{1\}) \\ &\qquad\qquad \cap \bigcap\nolimits_{i=1,2} (\mathrm{B}^*\,\mathrm{pr}_i)^{-1}(\{\mathrm{Int}\,\mathfrak{m}_i\})\end{aligned} \tag{24}$$

where $\mathrm{ev}_f : c \mapsto c(f)$ is the evaluation function.

$$=: V_{xx'}$$

From the last representation it is apparent that $V_{xx'}$ is an intersection of weak-*-closed subsets of $\mathrm{B}^*(X \times X)$ as $\mathrm{B}(\mathrm{pr}_1), \mathrm{B}(\mathrm{pr}_2) : \mathrm{B}\,X \to \mathrm{B}(X \times X)$ are continuous by functoriality. Hence $V_{xx'}$ is closed. Moreover the set $\{\, \phi \in \mathrm{B}^*(X \times X) \mid \phi(f) \in [-1,1] \text{ for any } f \text{ with } \|f\|_\mathrm{u} \leq 1 \,\}$ is compact by Banach-Alaoglu theorem [13, V.4.2]. This set also contains any positive $c \in \mathrm{B}^*(X \times X)$ with $c(1) = 1$ (for any f with $\|u\|_\mathrm{f} \leq 1$ note that $|c(f)| \underset{c \text{ linear}}{\leq} |c(|f|)| \underset{c \text{ positive}}{=} \int |f| \,\mathrm{d}c \underset{c \text{ positive,}}{\leq} \int 1 \,\mathrm{d}c = 1$). Thus $V_{xx'}$ is compact. [13, III.2.22, p. 119; III.1.5]

Set $U = [0,\infty)$ and define $\tilde{d}_x : U \to [0,1]$ by $\tilde{d}_x = d_i$ if $x \in [i, i+1)$. Note that $\tilde{d}_x$ is an increasing function in x. Further define $f : U \times V_{xx'} \to [0,1] \in \mathbf{Set}$ by $f(x,c) \coloneqq c(\tilde{d}_x) = \int \tilde{d}_x \,\mathrm{d}c$. For upper semi-continuity of f in its first argument fix a $c \in V_{xx'}$ and take any $r \in \mathbb{R}$. Observe that $f(_, c)^{-1}([0,r)) = \{\, x \in U \mid \exists i \in \omega : x < i+1 \wedge c(d_i) < r \,\}$ is open. As r was chosen arbitrarily, upper-semicontinuity is proven. For lower semi-continuity of f in its second argument fix an $x \in U$ and take again any $r \in [0,\infty)$. Let i be the element of ω with $x \in [i, i+1)$. Observe that $f(x, _)^{-1}((r,\infty)) = f(i, _)^{-1}((r,\infty)) = \{\, c \in V_{xx'} \mid c(d_i) > r \,\} = {\mathrm{ev}_{d_i}}^{-1}((r,\infty))$ is open by definition of weak-*-topology. Since x and r were chosen arbitrarily, lower-semicontinuity is proven. For each c and $\lambda \in \mathbb{R}$ the level set $\{\, x \in U \mid f(x,c) \geq \lambda \,\}$ is of form $[i,\infty) \subseteq \mathbb{R}$ and thus obviously convex. Convexity of $V_{xx'}$ is also quickly confirmed by noting that every operand in (24) is closed under convex combination. For each $x \in U$ and $\lambda \in \mathbb{R}$ the level set $\{\, c \in \mathrm{B}^*(X \times X) \mid c(\tilde{d}_x) \leq \lambda \,\}$ is convex by linearity of the c's. As the intersection of convex sets is convex, the level set $\{\, c \in V_{xx'} \mid c(\tilde{d}_x) \leq \lambda \,\}$ is convex.

Then by Lemma 35, $\inf_{c \in V_{xx'}} \sup_{x \in [0,\infty)} c(\tilde{d}_x) = \sup_{x \in [0,\infty)} \inf_{c \in V_{xx'}} c(\tilde{d}_x)$. By definition of $\tilde{d}$ this becomes $\inf_{c \in V_{xx'}} \sup_{i \in \mathbb{N}} c(d_i) = \sup_{i \in \mathbb{N}} \inf_{c \in V_{xx'}} c(d_i)$. By the canonical isomorphism and applying Levi's theorem (monotone convergence theorem) [20, 123A] the claim in (23) follows. ◀

C Topology

Recall that a topological space is R_2 if any pair of topologically distinct point (i.e. $\exists U \in \mathcal{T} : U \cap \{x,y\} \in \{\{x\},\{y\}\}$) are separated by disjoint open sets (i.e. $\exists U, V \in \mathcal{T} : x \in U \wedge y \in V \wedge U \cap V = \emptyset$). Also recall that a topological space is called T_2 if it is Hausdorff and T_0 if any pair of distinct point is separated by an open set. Obviously, a T_2-space is precisely an R_2-space that is T_0. Any topological space can be transformed canonically into a T_0 space by identifying all point that are not topologically distinct point, resulting in the so-called *Kolmogorov quotient* Kol. We now generalise a bit the well-known Stone-Weierstraß theorem using the fact that T_0-spaces form a reflective subcategory of topological spaces by the Kolmogorov quotient construction. Proofs of the following theorems are found in [2].

▶ **Lemma 37.** *Stone-Weierstraß Let $(X,\mathcal{T})$ a compact R_2-space. Let L be a set of continuous functions $X \to \mathbb{R}$ such that $\min\{f,g\}, \max\{f,g\} \in L$ for all $f,g \in L$. If some continuous function $f : X \to \mathbb{R}$ can be approximated at each pair of points by functions in L, then f itself can also be approximated by functions in L with respect to the uniform norm $\|_\|_\mathrm{u}$.*

▶ **Lemma 38.** *The unit of the Kolmogorov construction $\eta_X : X \to \mathrm{Kol}\,X$ is proper, i.e. preimages of compact sets are compact.*

D Proof of adequacy, Theorem 22

Proof. Set $[\![_]\!] = [\![_]\!]_\gamma$. We prove $\mathbf{bd} \geq [\![\varphi]\!]^\bullet \mathrm{d_E}$ for each $\varphi \in \mathcal{L}$ by structural induction over φ, where $[\![\varphi]\!]^\bullet = ([\![\varphi]\!] \times [\![\varphi]\!])^{-1}$. Recall again that all logical symbols are also function symbols, so the proof consists of two cases: Function symbols and modal operators.

Take a formula $\varphi = f(\varphi_1, \dots, \varphi_n)$ for an arbitrary n-ary function symbol f – interpreted by a function also denoted by $f\colon [0,1]^n \mapsto [0,1]$ – and formulas φ_i with $i = 1, \dots, n$ with $\mathbf{bd} \geq [\![\varphi_i]\!]^\bullet \mathrm{d_E}$ for each i. Given two states $x, y \in X$ we find

$$\begin{aligned}
[\![\varphi]\!]^\bullet \mathrm{d_E}(x,y) &\overset{\text{Subsection 5.1}}{=} |f([\![\varphi_1]\!](x), \dots, [\![\varphi_n]\!](x)) - f([\![\varphi_1]\!](y), \dots, [\![\varphi_n]\!](y))| \\
&\overset{f \text{ nonexpansive}}{\leq} \mathrm{d}_\infty\big(([\![\varphi_1]\!](x), \dots, [\![\varphi_n]\!](x)), [\![\varphi_1]\!](y), \dots, [\![\varphi_n]\!](y)\big) \\
&\overset{\substack{\forall i\colon \mathbf{bd} \geq [\![\varphi_i]\!]^\bullet \mathrm{d_E}, \\ \text{definition of } \mathrm{d}_\infty}}{\leq} \mathbf{bd}(x,y).
\end{aligned}$$

As x, y had been chosen arbitrarily, it follows that $\mathbf{bd} \geq [\![\varphi]\!]^\bullet \mathrm{d_E}$.

Turning in the final step to modal operators, take any $a \in \Sigma$ and assume that for a formula ψ we already know that $\mathbf{bd} \geq [\![\psi]\!]^\bullet \mathrm{d_E}$. Observe for any two states $x, y \in X$

$$\begin{aligned}
[\![\Diamond_i \psi]\!]^\bullet \mathrm{d_E}(x,y) &\overset{(7)}{=} \left| \int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,x} +_c rew_a(x) - \left(\int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,y} +_c rew_a(y) \right) \right| \\
&\overset{\substack{\text{definition of } +_c, \\ \text{linearity of integral}}}{=} \left| \left(\int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,x} - \int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,y} \right) +_c (rew_a(x) - rew_a(y)) \right| \\
&\overset{\substack{c, 1-c \geq 0; \text{ definition of } +_c; \\ \text{triangle inequality}}}{\leq} \left| \int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,x} - \int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,y} \right| +_c |rew_a(x) - rew_a(y)| \\
&= \left| \int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,x} - \int [\![\psi]\!] \,\mathrm{d}\mathfrak{m}_{a,y} \right| +_c |rew_a^\bullet(\mathrm{d_E})(x,y)|
\end{aligned}$$

let $\mathfrak{c}$ always range over all coupling $K(\mathfrak{m}_{a,x}, \mathfrak{m}_{a,y})$

$$\begin{aligned}
&\overset{\substack{\text{definition of coupling} \\ \text{Definition 9}}}{=} \left| \inf_{\mathfrak{c}} \int [\![\psi]\!](x) - [\![\psi]\!](y) \mathfrak{c}(\mathrm{d}x', \mathrm{d}y') \right| +_c rew_a^\bullet(\mathrm{d_E})(x,y) \\
&\overset{\text{Jensen's inequality}}{\leq} \inf_{\mathfrak{c}} \int |[\![\psi]\!](x) - [\![\psi]\!](y)| \,\mathfrak{c}(\mathrm{d}x', \mathrm{d}y') +_c rew_a^\bullet(\mathrm{d_E})(x,y) \\
&= \inf_{\mathfrak{c}} \int [\![\psi]\!]^\bullet \mathrm{d_E} \,\mathrm{d}\mathfrak{c} +_c rew_a^\bullet(\mathrm{d_E})(x,y) \\
&\leq \sup_{a \in \Sigma} \inf_{\mathfrak{c}} \int [\![\psi]\!]^\bullet \mathrm{d_E} \,\mathrm{d}\mathfrak{c} +_c rew_a^\bullet(\mathrm{d_E})(x,y) \\
&\overset{\text{definition of } \sigma^B}{=} \big(\gamma_B \circ \sigma^B([\![\psi]\!]^\bullet \mathrm{d_E})\big)(x,y) \\
&\overset{\substack{\text{monotonicity from} \\ \text{Theorem 13; } \mathbf{bd} \geq [\![\psi]\!]^\bullet \mathrm{d_E}}}{\leq} \big(\gamma_B \circ \sigma^B(\mathbf{bd})\big)(x,y) \\
&= \mathbf{bd}(x,y).
\end{aligned}$$

As x, y had been chosen arbitrarily, it follows that $\mathbf{bd} \geq [\![\Diamond_a \psi]\!]^\bullet \mathrm{d_E}$. ◀

Note that the usage of Jensen's inequality could be avoided using (9). But we chose to conduct the proof this way in view of possible further research into logics without negation.

E Remaining proofs

The remaining proofs in this paper and its appendices are found in [2].