# A FOUNDATIONAL THEORY OF QUANTITATIVE ABSTRACTION: ADJUNCTIONS, DUALITY, AND LOGIC FOR PROBABILISTIC SYSTEMS

NIVAR ANWER [a], EZEQUIEL LÓPEZ-RUBIO [c,d], DAVID ELIZONDO [b], AND RAFAEL M. LUQUE-BAENA [c,d]

[a] School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332, USA

[b] Institute of Artificial Intelligence, De Montfort University, The Gateway, Leicester LE1 9BH, United Kingdom

[c] ITIS Software, University of Málaga, 29071 Málaga, Spain

[d] Instituto de Investigación Biomédica de Málaga y Plataforma en Nanomedicina (IBIMA Plataforma BIONAND), Málaga TechPark, Campanillas, 29590 Málaga, Spain
*e-mail address*: nivar.anwer@gmail.com
*e-mail address*: elizondo@dmu.ac.uk
*e-mail address*: rmluque@uma.es
*e-mail address*: elr@uma.es

Abstract. The analysis and control of stochastic dynamical systems rely on probabilistic models such as (continuous-space) Markov decision processes, but large or continuous state spaces make exact analysis intractable and call for principled quantitative abstraction. This work develops a unified theory of such abstraction by integrating category theory, coalgebra, quantitative logic, and optimal transport, centred on a canonical $\varepsilon$-quotient of the behavioral pseudo-metric with a universal property: among all abstractions that collapse behavioral differences below $\varepsilon$, it is the most detailed, and every other abstraction achieving the same discounted value-loss guarantee factors uniquely through it. Categorically, a quotient functor $\mathrm{Q}_\varepsilon$ from a category of probabilistic systems to a category of metric specifications admits, via the Special Adjoint Functor Theorem, a right adjoint $\mathrm{R}_\varepsilon$, yielding an adjunction $\mathrm{Q}_\varepsilon \dashv \mathrm{R}_\varepsilon$ that formalizes a duality between abstraction and realization; logically, a quantitative modal $\mu$-calculus with separate reward and transition modalities is shown, for a broad class of systems, to be expressively complete for the behavioral pseudo-metric, with a countable fully abstract fragment suitable for computation. The theory is developed coalgebraically over Polish spaces and the Giry monad and validated on finite-state models using optimal-transport solvers, with experiments corroborating the predicted contraction properties and structural stability and aligning with the theoretical value-loss bounds, thereby providing a rigorous foundation for quantitative state abstraction and representation learning in probabilistic domains.

*2012 ACM CCS:* Theory of computation → Modal and temporal logics; Theory of computation → Categorical semantics; Computing methodologies → Markov decision processes; Computing methodologies → Reinforcement learning; Software and its engineering → Formal methods.
*2010 Mathematics Subject Classification:* 03B44, 03B70, 18C15, 60J05, 68T05.
*Key words and phrases:* Category theory, coalgebra, optimal transport, reinforcement learning, quantitative logic, abstraction, universal property, adjunction.

Preprint submitted to
Logical Methods in Computer Science

## Introduction

The design, verification, and control of complex computational systems - from autonomous robots and cyber-physical systems to communication networks and large-scale learning architectures - increasingly rely on probabilistic models such as Markov Decision Processes (MDPs) [Put94, BT96]. As these models grow in size, incorporate high-dimensional or continuous state spaces, and interact with noisy sensors and actuators, exact analysis and planning become computationally intractable, a manifestation of the classical "curse of dimensionality" [Bel57].

At the heart of this challenge lies the problem of **abstraction**. For a given system, the goal is to construct a simpler, smaller, and more tractable model that preserves the control- and specification-relevant behavior of the original. Qualitative notions such as (probabilistic) bisimulation have long underpinned the theory of concurrency and process equivalence, including their extension to stochastic systems [Mil80, Mil89, LS91, DGJP04]. Such notions, however, are too rigid for realistic applications: real-world systems are rarely exactly equivalent, but are often *approximately* similar in behavior. This motivates a shift from strictly qualitative to **quantitative** notions of behavioral equivalence and abstraction, where states are compared by distances rather than by exact indistinguishability.

This shift is not purely theoretical. In reinforcement learning (RL) [SB18], the central challenge of **state representation learning** is precisely to identify compressed state representations that preserve decision-relevant information: an agent should behave nearly optimally when planning in the compressed representation and then executing in the original environment [LDRGF18, BCV13]. Quantitative abstraction offers a principled way to formalize this requirement. The core question addressed by this work is:

*How can a canonical quantitative abstraction be defined, with provable optimality guarantees and a robust categorical and logical foundation?*

An intuitive way to understand this problem is through the analogy of map-making. A high-resolution satellite image that records every detail of a city is, in some sense, the most faithful representation of the terrain, yet it is unusable for everyday navigation. Effective maps suppress irrelevant detail, retaining exactly the structure needed for a particular task. A subway map further abstracts away geometry, preserving topological connectivity that matters for transit, whereas a road map preserves distances and routes. The appropriate abstraction depends on the intended use: different tasks require different projections of the same underlying reality.

In engineered and learning systems, the analogue of a "good map" is an abstract model that is as simple as possible while preserving performance-relevant properties, such as optimal expected discounted reward or satisfaction of formal specifications. A quantitative abstraction should therefore not be judged solely by its size, but by its *faithfulness*: it should discard only those distinctions that are behaviorally negligible for the chosen performance notion.

The present work constructs a comprehensive mathematical theory of such quantitative abstraction. The central aim is to identify a *canonical* way to approximate probabilistic systems, in which the approximation is parameterized by a tolerance $\varepsilon \geq 0$ and has a universal characterization. The development is grounded in category theory and coalgebra, which provide a language for describing systems by their observable behavior and relating different system types [Lan98, Rut00, Jac16]. The focus is on systems with Polish state spaces, continuous dynamics, and bounded rewards, covering both discrete and continuous MDPs and many models of practical relevance in control, robotics, and economics [Par05, Bog07, Gir82, Vil09a].

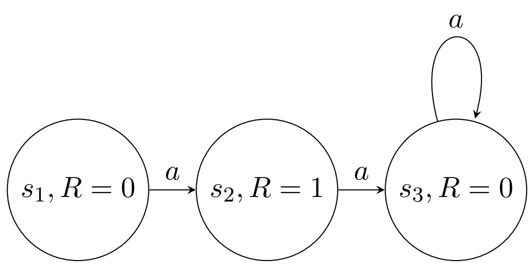
To keep the theory anchored, a simple but instructive running example is used throughout.

**Example 0.1** (Running Example: A Three-State Chain)**.** Consider the MDP with state space $S = \{s_1, s_2, s_3\}$, a single action $a$, and discount factor $\gamma = 0.9$. The reward function and transitions are:

$$R(s_1, a) = 0, \quad R(s_2, a) = 1, \quad R(s_3, a) = 0,$$

$$P(s_1, a) = \delta_{s_2}, \quad P(s_2, a) = \delta_{s_3}, \quad P(s_3, a) = \delta_{s_3}.$$

Graphically,




The three states have distinct roles: $s_2$ yields an immediate reward, $s_1$ is one step away from that reward, and $s_3$ leads only to zero reward thereafter. A behavioral distance should recognize these differences and quantify them in a way that is meaningful for value-based decision making. Later sections return to this example to compute the behavioral metric (Example 2.7) and illustrate the quotient construction (Example 4.3).

**Contributions and Structure of the Paper.** The work can be organized around six main contributions, each of which is developed in detail and connected to the others.

(1) **Canonical $\varepsilon$-Quotient and Universal Property.** A canonical *$\varepsilon$-quotient metric space* $(S_\varepsilon, d_Q)$ is defined for any probabilistic system endowed with a behavioral pseudo-metric $d_M$. The quotient identifies states that are connected by chains of pairwise distance at most $\varepsilon$ and equips the set of equivalence classes with a path metric. This quotient is not an ad-hoc aggregation: it is characterized by a universal property as the *terminal* $\varepsilon$-abstraction. Any other abstraction that collapses all behavioral distinctions below $\varepsilon$ arises uniquely as a further metric quotient of $(S_\varepsilon, d_Q)$.
(2) **Adjunction between Abstraction and Realization.** Abstraction is formulated categorically via a functor $\mathrm{Q}_\varepsilon$ from a category **MDP** of probabilistic systems to a category **Spec** of metric specifications. Using the Special Adjoint Functor Theorem [Lan98, AR94], a right adjoint realization functor $\mathrm{R}_\varepsilon$ is constructed, yielding an adjunction

$$\mathrm{Q}_\varepsilon \dashv \mathrm{R}_\varepsilon.$$

This adjunction formalizes the duality between passing from concrete systems to abstract specifications and constructing canonical systems from abstract metric data.

(3) **Quantitative Modal $\mu$-Calculus and Logical Full Abstractness.** A quantitative modal $\mu$-calculus $\mathcal{L}_\mu$ is introduced, featuring distinct modalities for one-step rewards and probabilistic transitions [dAF03, Pan09]. For a broad subcategory $\mathbf{MDP}_{\mathrm{rep}}$ of *logically representable* systems, a constructive expressive completeness theorem is proved: the behavioral pseudo-metric equals the supremum of quantitative deviations of formulas. This establishes full abstractness of the logic with respect to the metric. The result is then sharpened to a *countable* fragment $\mathcal{L}_\mu^c$ suitable for algorithmic use.
(4) **Optimality for Value Function Approximation.** The canonical $\varepsilon$-quotient is connected to approximate dynamic programming [BT96, Sze10]. It is shown that if states are aggregated according to the $\varepsilon$-quotient, then any policy induced from the abstract model suffers at most $\frac{2\varepsilon}{1-\gamma}$ loss in discounted value on the original system. Moreover, the universal property implies that the canonical quotient is *optimal with respect to faithfulness*: any other abstraction satisfying the same loss bound factors through the canonical one and therefore discards at least as much information.
(5) **Compositionality via Fibrations.** The dependence of systems on action spaces is organized in a fibration whose base category is **Pol** and whose fibers are categories of MDPs with fixed action spaces [Jac16]. The family of abstraction-realization adjunctions is shown to satisfy a Beck–Chevalley condition for surjective interface maps, ensuring that abstraction commutes with action-space refinements. This provides a principled account of compositional abstraction for heterogeneous systems.
(6) **Empirical Validation with Exact Optimal Transport.** A comprehensive experimental study on finite MDPs is carried out, using exact 1-Wasserstein computations (LP-based, without entropic regularization) inside the Bellman operator [Vil09a, PC19]. The experiments confirm the contraction properties, structural stability, quotient idempotence, transfer behavior, and scaling trends predicted by the theory, and provide a concrete demonstration of the canonical quotient as a target for representation learning.

Sections are organized accordingly. Section 1 develops the necessary background on Polish spaces, the Giry monad, and Wasserstein metrics. The coalgebraic metric theory and behavioral pseudo-metric are introduced in the subsequent section. The quantitative modal $\mu$-calculus and its completeness results appear next, followed by the universal $\varepsilon$-quotient and the abstraction-realization adjunction. The implications for value function approximation and representation learning are then developed, leading to the empirical validation and the fibrational analysis of compositionality.

## Context and Related Work

Quantitative abstraction for probabilistic systems sits at the intersection of several mature lines of research: bisimulation metrics and behavioral pseudometrics, logical characterizations of quantitative behavior, optimal transport and Kantorovich liftings, state representation learning in reinforcement learning, and coalgebraic and categorical accounts of system semantics. The present framework is positioned at the confluence of these developments, but differs from existing work by combining a canonical $\varepsilon$-quotient with a universal property, a categorical adjunction between abstraction and realization, and a fully abstract quantitative modal $\mu$-calculus over Polish spaces, all tied directly to value-loss guarantees and representation learning.

Work on metrics for probabilistic behavior originates from early studies of labelled Markov processes and probabilistic transition systems, where behavioral pseudometrics were introduced as fixed points of metric transformers capturing the idea of quantitative bisimulation [DGJP04, vBW05]. In these settings, two states are close when they match each other's probabilistic behavior up to small quantitative deviations. In Markov decision processes, bisimulation metrics of Ferns, Panangaden, and Precup specialize this idea to control problems by defining a contraction operator on state pseudo-metrics that combines immediate reward differences with a Kantorovich lifting of the transition kernel [FPP04, FPP10]. These constructions yield tight Lipschitz bounds on differences in optimal value functions, motivate principled state aggregation schemes, and connect to approximate bisimulation and approximate dynamic programming [FPC06, FP14, CP10]. The framework developed here adopts this Bellman-style fixed-point view but departs from it in two ways: first, by isolating a canonical $\varepsilon$-quotient equipped with a terminal universal property among all abstractions that collapse behavioral differences below $\varepsilon$; second, by lifting this quotient to a functor and identifying a right adjoint realization functor via the Special Adjoint Functor Theorem, thereby embedding quantitative abstraction into a categorical duality between systems and metric specifications.

A second large body of work connects behavioral equivalences and metrics with modal logics. For qualitative systems, Hennessy–Milner logic and the modal $\mu$-calculus are known to be expressively complete for bisimulation-invariant properties, with key results on expressiveness and completeness due to Bradfield and Stirling among others [BS06]. For probabilistic and quantitative systems, numerous authors have developed real-valued modal logics and quantitative $\mu$-calculi that interpret formulas as real-valued observables and use fixed-point operators to capture long-run behavior [dAF03, DGJP04]. A series of "mimicking formulae" results shows that, for many discrete probabilistic transition systems, behavioral pseudometrics can be characterized as the maximal difference in the values of certain modal or $\mu$-calculus formulas, providing a tight logical account of quantitative behavior [DLP02, WSP18, Kom21]. Existing characterizations, however, typically work in countable or image-finite settings and focus on transition structure rather than on combined reward/transition models over Polish spaces, and they are not usually linked to universal notions of quantitative abstraction. The quantitative modal $\mu$-calculus studied here follows the broad quantitative tradition originating with de Alfaro *et al.* [dAF03] but is adapted to MDP-like coalgebras on Polish spaces with separate reward and probabilistic modalities, and is proved to be fully abstract for the behavioral pseudo-metric via a constructive mimicking-formula argument that yields, in addition, a countable fragment suited to algorithmic use.

Optimal transport and Kantorovich–Rubinstein duality provide the analytic backbone of much of the work on behavioral metrics. Kantorovich's relaxation of Monge's optimal transport problem and subsequent duality theory [Kan58] are developed systematically in Villani's monograph [Vil09a], where the 1-Wasserstein distance plays a central role in metrizing weak convergence and in defining distances between probability measures. In the context of probabilistic systems and MDPs, the 1-Wasserstein lifting with respect to a state pseudo-metric supplies the metric component of the behavioral contraction operator [FPP10], and underlies many subsequent notions of probabilistic bisimulation distance. On the algorithmic side, entropic regularization and Sinkhorn iterations allow scalable approximations to optimal transport, as pioneered by Cuturi [Cut13] and systematized by Peyré and Cuturi [PC19]. Recent work further sharpens the explicit link between bisimulation metrics and optimal transport couplings between Markov kernels, showing that a wide class of probabilistic bisimulation distances can be viewed as optimal transport problems on path space or on kernel space equipped with suitable ground metrics. The theory developed in this paper takes the Wasserstein viewpoint as primitive: the behavioral metric is defined as a Bellman fixed point built from a one-Wasserstein lifting, Kantorovich–Rubinstein duality is used in the logical completeness proof, and the canonical $\varepsilon$-quotient is justified as the most detailed abstraction compatible with given Wasserstein-based value-loss bounds.

State representation learning in reinforcement learning has recently adopted bisimulation metrics and Wasserstein distances as central design tools. Early work on bisimulation-based abstraction and transfer in MDPs already recognized that bisimilar states can be aggregated without loss for control [GDG03], while approximate bisimulation metrics provide quantitative guarantees for value loss under aggregation [FPP04, FP14]. More recent deep RL methods explicitly learn latent representations intended to approximate bisimulation quotients. DeepMDP [GKN$^+$19] and related latent-dynamics models train embeddings so that latent transitions and rewards predict observed returns, often with Wasserstein-type losses on successor distributions. Other approaches use contrastive or information-theoretic objectives to learn invariants of behavior: learning invariant representations without reconstruction [ZSB$^+$21] and contrastive bisimulation-based state representations [ASC$^+$21] are prominent examples. These algorithms often rely on heuristically chosen distances and architectures, and typically do not single out a canonical abstraction with a universal property. The present framework instead identifies such a canonical object—the $\varepsilon$-quotient of the behavioral pseudo-metric—proves that it is terminal among all abstractions achieving the same discounted value-loss guarantee, and thereby provides a normative geometric target for representation-learning objectives based on bisimulation.

Coalgebra and category theory provide a unifying language for many of the constructions above. Universal coalgebra, as developed by Rutten [Rut00] and further elaborated by Jacobs [Jac16], views a wide range of transition systems as coalgebras for endofunctors on suitable base categories, with behavioral equivalence captured by coalgebraic notions of bisimulation. Coalgebraic modal logic supplies generic recipes for building modal logics from functors [Pat03, Jac16], and has been extended to quantitative systems by combining functor liftings with metric spaces and quantitative modalities. From a metric viewpoint, coalgebraic behavioral metrics arise by lifting endofunctors from sets to metric spaces using Kantorovich-style constructions and studying fixed points of associated metric transformers [vBW05, FSLM21]. Categorical treatments of optimal transport and Wasserstein liftings show that these constructions are functorial and compatible with standard categorical structure, making them natural tools for organizing quantitative system semantics. The present work fits squarely into this coalgebraic landscape but pushes in two directions that appear novel: first, by constructing from the behavioral metric a quotient functor $\mathrm{Q}_\varepsilon$ from a category of probabilistic systems to a category of metric specifications and applying the Special Adjoint Functor Theorem [Lan98, AR94] to obtain a right adjoint realization functor $\mathrm{R}_\varepsilon$; and second, by organizing families of such adjunctions in a fibrational setting over action spaces, ensuring that canonical abstractions behave compositionally with respect to surjective changes of interface.

Taken together, these literatures provide the essential ingredients for a theory of quantitative abstraction, but they have largely evolved in parallel: behavioral pseudometrics for probabilistic systems, logical characterizations via mimicking formulae, Wasserstein-based constructions and scalable optimal-transport solvers, deep bisimulation-inspired representations in reinforcement learning, and coalgebraic accounts of systems and logics. The present framework can be viewed as a synthesis that (i) identifies a canonical, universal quantitative abstraction parameterized by $\varepsilon$; (ii) presents this abstraction as part of an adjunction between concrete probabilistic systems and abstract metric specifications; (iii) equips the resulting metric with a fully abstract quantitative modal $\mu$-calculus, including a countable fragment suitable for computation; and (iv) connects the entire structure to value-loss guarantees and representation learning, within a coalgebraic setting that naturally accommodates both discrete and continuous (Polish) state spaces and extends, with minor modifications, to other functorial system types such as labelled transition systems, labelled Markov processes, and stochastic games.

# **Part** 1. **Mathematical Foundations**

## 1. Categorical and Metric Preliminaries

This section presents the mathematical foundations required for the subsequent development. The focus is on three pillars: (i) monads as abstract machines for computational structure, (ii) Polish spaces and their measure-theoretic regularity, and (iii) the Giry monad and Wasserstein metrics as the core tools for probabilistic systems.

**Definition 1.1** (Monad)**.** A **monad** on a category $\mathbf{C}$ is a triple $(T, \eta, \mu)$ consisting of:

(1) an endofunctor $T : \mathbf{C} \to \mathbf{C}$,
(2) a natural transformation $\eta : \mathrm{id}_{\mathbf{C}} \to T$ (the *unit*), and
(3) a natural transformation $\mu : T^2 \to T$ (the *multiplication*),

subject to the following coherence conditions. For every object $X \in \mathbf{C}$:

- **Unitality:**
$$\mu_X \circ T(\eta_X) = \mathrm{id}_{T(X)}, \qquad \mu_X \circ \eta_{T(X)} = \mathrm{id}_{T(X)} .$$
- **Associativity:**
$$\mu_X \circ T(\mu_X) = \mu_X \circ \mu_{T(X)} : T^3(X) \to T(X).$$

In many computational settings, $T(X)$ represents a space of “effectful” or “structured” values of type $X$; $\eta_X$ embeds pure values into this space, and $\mu_X$ flattens nested effects. For instance, the list monad takes $T(X)$ to be finite lists over $X$, the unit maps $x$ to the singleton list $[x]$, and the multiplication concatenates lists of lists. The probability monad considered below treats $T(X)$ as a space of probability measures on $X$, with the unit assigning Dirac measures and the multiplication marginalizing distributions over distributions.

**Definition 1.2** (Polish Spaces)**.** A topological space $(X, \tau)$ is **Polish** if it is separable (has a countable dense subset) and completely metrizable (admits a complete metric inducing $\tau$). The category of Polish spaces and continuous maps is denoted by **Pol**.

Many spaces of interest in probability and analysis are Polish, including $\mathbb{R}^n$ with the Euclidean metric, compact metric spaces such as $[0, 1]$, separable Banach spaces, and spaces of probability measures over Polish spaces endowed with suitable topologies.

*Remark* 1.3 (Why Polish Spaces Are the Natural Universe). Polish spaces strike a balance between generality and regularity that is particularly suited to probability and system theory. Completeness ensures existence of limits and supports fixed-point arguments; separability ensures that measure-theoretic structures (e.g., Borel $\sigma$-algebras) are well-behaved and standard. In particular, Borel probability measures on Polish spaces form standard Borel spaces, enabling disintegration and regular conditional probabilities and avoiding pathologies that arise in more general measurable spaces. The category **Pol** is also closed under many natural constructions (products, measure spaces, etc.), making it a robust setting for probabilistic semantics.

**Definition 1.4** (Borel Probability Measures and the Giry Monad)**.** Let $X$ be a Polish space with Borel $\sigma$-algebra $\mathcal{B}(X)$. Denote by $\Delta(X)$ the set of Borel probability measures on $X$. Endow $\Delta(X)$ with the *weak topology*, the coarsest topology that makes all maps

$$\mu \mapsto \int_X f \, d\mu$$

continuous for bounded continuous $f : X \to \mathbb{R}$. By Prokhorov's theorem, $\Delta(X)$ is Polish under this topology.

The **Giry monad** $\mathsf{G} : \mathbf{Pol} \to \mathbf{Pol}$ is defined as follows:

- On objects: $\mathsf{G}(X) = \Delta(X)$.
- On morphisms: for a continuous map $f : X \to Y$, define $\mathsf{G}(f) : \Delta(X) \to \Delta(Y)$ as the *pushforward*
$$(\mathsf{G}(f))(\mu)(B) = \mu(f^{-1}(B)), \qquad B \in \mathcal{B}(Y).$$
- Unit: $\eta_X : X \to \Delta(X)$, $x \mapsto \delta_x$, the Dirac measure at $x$.
- Multiplication: $\mu_X : \Delta(\Delta(X)) \to \Delta(X)$ defined for $\varpi \in \Delta(\Delta(X))$ by
$$\mu_X(\varpi)(B) = \int_{\Delta(X)} \nu(B) \, d\varpi(\nu), \qquad B \in \mathcal{B}(X).$$

**Proposition 1.5.** *The triple* $(\mathsf{G}, \eta, \mu)$ *defines a monad on* **Pol**.

*Proof.* Functoriality of $\mathsf{G}$ follows from the functoriality of pushforward measures. Naturality of $\eta$ and $\mu$ is immediate from the definitions. The monad laws reduce to standard identities for integrals and Dirac measures (e.g., Fubini's theorem and the fact that $\delta_x$ acts as a unit for convolution-like operations). A detailed proof is provided in Appendix A.1. □

**Definition 1.6** (1-Wasserstein Distance)**.** Let $(X, d)$ be a Polish metric space. Denote by $\mathcal{P}_1(X)$ the set of probability measures on $X$ with finite first moment:

$$\mathcal{P}_1(X) = \Big\{ \mu \in \Delta(X) : \int_X d(x_0, x) \, d\mu(x) < \infty \text{ for some (hence any) } x_0 \in X \Big\}.$$

A *coupling* of $\mu, \nu \in \mathcal{P}_1(X)$ is a probability measure $\pi \in \Delta(X \times X)$ with marginals $\mu$ and $\nu$:

$$\pi(A \times X) = \mu(A), \quad \pi(X \times B) = \nu(B).$$

Denote the set of couplings by $\Pi(\mu, \nu)$. The **1-Wasserstein distance** $\mathrm{W}_1$ between $\mu$ and $\nu$ is defined as

$$\mathrm{W}_1(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y) \, d\pi(x, y).$$

If $d$ is bounded, then $\mathcal{P}_1(X) = \Delta(X)$ and $\mathrm{W}_1$ metrizes the weak topology. In general, $\mathrm{W}_1$ is a metric on $\mathcal{P}_1(X)$.

The 1-Wasserstein distance admits the well-known "earth mover's" interpretation: $\mu$ and $\nu$ represent piles of unit mass, and $\mathrm{W}_1(\mu, \nu)$ is the minimal expected cost of transporting $\mu$ into $\nu$, where the cost of moving a unit of mass from $x$ to $y$ is $d(x, y)$.

**Theorem 1.7** (Kantorovich–Rubinstein Duality)**.** *Let $(X, d)$ be a Polish metric space and let $\mu, \nu \in \mathcal{P}_1(X)$. Then*

$$W_1(\mu, \nu) = \sup_{f \in \mathrm{Lip}_1(X)} \left| \int_X f \, d\mu - \int_X f \, d\nu \right|,$$

*where* $\mathrm{Lip}_1(X)$ *denotes the set of 1-Lipschitz functions $f : X \to \mathbb{R}$:*

$$|f(x) - f(y)| \le d(x, y) \qquad \forall x, y \in X.$$

*Proof.* This is a classical result in optimal transport [Kan58, Vil09b]. The inequality "$\ge$" follows immediately by evaluating $\int d \, d\pi$ against any 1-Lipschitz test function and then minimizing over couplings. The non-trivial direction, which shows the absence of a duality gap, relies on convex analysis and Hahn–Banach separation in a suitable function space; see Appendix A.2 for a detailed argument. □

The dual formulation expresses the distance between probability measures as the largest difference in expectation over all "fair" observers $f$ that do not magnify distances. This duality is central to the later connection between behavioral metrics and quantitative logic.

# Part 2. The Coalgebraic and Logical Theory of Quantitative Behavior

## 2. Metric Coalgebras for General Probabilistic Systems

Probabilistic systems are now formalized as coalgebras for a suitable endofunctor on **Pol**. This captures the one-step behavior - rewards and transitions - in a way that is both flexible and compositional.

**Definition 2.1** (System Functor on **Pol**)**.** Let $A$ be a Polish space of actions. Define a functor $\mathcal{F}_A : \mathbf{Pol} \to \mathbf{Pol}$ by:

- On objects: for a Polish space $X$, let
$$\mathcal{F}_A(X) := C_b\big(A, \mathsf{G}(X) \times \mathbb{R}\big),$$
the space of bounded continuous functions from $A$ to $\mathsf{G}(X) \times \mathbb{R}$, equipped with the supremum norm.
- On morphisms: for a continuous map $f : X \to Y$, define
$$\mathcal{F}_A(f) : \mathcal{F}_A(X) \to \mathcal{F}_A(Y), \quad (\mathcal{F}_A(f)(c))(a) := (\mathsf{G}(f) \times \mathrm{id}_{\mathbb{R}})(c(a)).$$

A **general probabilistic system** with action space $A$ is a coalgebra $(S, c)$ for $\mathcal{F}_A$, i.e., a continuous map $c : S \to \mathcal{F}_A(S)$.

Intuitively, for each state $s \in S$, the coalgebra map $c$ assigns a function $c(s) : A \to \mathsf{G}(S) \times \mathbb{R}$; for each action $a \in A$, the pair $(P(s, a), R(s, a))$ encodes the probabilistic transition and immediate reward.

*Remark* 2.2 (Boundedness and Continuity Assumptions)*.* The choice $\mathcal{F}_A(X) = C_b(A, \mathsf{G}(X) \times \mathbb{R})$ encodes two key assumptions:

(1) *Bounded rewards:* The use of $C_b$ enforces a uniform bound on rewards over $S \times A$. This ensures boundedness of value functions and well-posedness of the Bellman operator used later.

(2) *Continuous dynamics:* The requirement that $c : S \to C_b(A, \mathsf{G}(S) \times \mathbb{R})$ is continuous and that each $c(s)$ is continuous in $a$ implies that small changes in state or action produce small changes in the transitional behavior, in the topology induced by weak convergence on $\mathsf{G}(S)$ and the usual topology on $\mathbb{R}$. This rules out highly discontinuous systems but aligns with many continuous control and stochastic modeling scenarios.

These assumptions are essential for guaranteeing existence, continuity, and uniqueness properties of the behavioral pseudo-metric.

The goal is to equip $S$ with a pseudo-metric that reflects long-run behavioral similarity. This will be obtained as the unique fixed point of a contraction on the space of bounded pseudo-metrics.

**Lemma 2.3.** *Let $S$ be a Polish space. The space* $\mathrm{pMet}_b(S)$ *of bounded pseudo-metrics on $S$, equipped with the supremum metric*

$$d_\infty(d_1, d_2) \coloneqq \sup_{s,t \in S} |d_1(s,t) - d_2(s,t)|,$$

*is a non-empty complete metric space.*

*Proof.* Consider $\mathcal{B}(S \times S, \mathbb{R})$, the Banach space of bounded real-valued functions on $S \times S$ with the supremum norm. The set $\mathrm{pMet}_b(S)$ of bounded pseudo-metrics is non-empty (the zero function is a pseudo-metric) and closed under uniform limits: uniform limits of symmetric, non-negative, zero-diagonal functions satisfying the triangle inequality again satisfy these properties. Closed subsets of complete metric spaces are complete, so $(\mathrm{pMet}_b(S), d_\infty)$ is complete. □

**Definition 2.4** (Behavioral Metric Operator)**.** Let $(S, c)$ be a system with discount factor $0 < \gamma < 1$. For each $d \in \mathrm{pMet}_b(S)$, define a new pseudo-metric $\mathcal{K}_c(d)$ on $S$ by

$$\mathcal{K}_c(d)(s_1, s_2) \coloneqq \sup_{a \in A} \Big( |R(s_1, a) - R(s_2, a)| + \gamma \, \mathrm{W}_1^d\big(P(s_1, a), P(s_2, a)\big) \Big),$$

where $c(s_i)(a) = (P(s_i, a), R(s_i, a))$ and $\mathrm{W}_1^d$ denotes the 1-Wasserstein distance computed using $d$ as the ground pseudo-metric on $S$.

**Lemma 2.5.** *The operator $\mathcal{K}_c : \mathrm{pMet}_b(S) \to \mathrm{pMet}_b(S)$ is a contraction with Lipschitz constant $\gamma$ under the metric $d_\infty$.*

*Proof.* Let $d_1, d_2 \in \mathrm{pMet}_b(S)$ and fix $s_1, s_2 \in S$. Denote

$$h(a) = |R(s_1, a) - R(s_2, a)|, \quad g_i(a) = \gamma \, \mathrm{W}_1^{d_i}\big(P(s_1, a), P(s_2, a)\big).$$

Then

$$\mathcal{K}_c(d_i)(s_1, s_2) = \sup_{a \in A} \big(h(a) + g_i(a)\big).$$

Using $|\sup f - \sup g| \le \sup |f - g|$,

$$\begin{aligned} \big|\mathcal{K}_c(d_1)(s_1, s_2) - \mathcal{K}_c(d_2)(s_1, s_2)\big| &\le \sup_{a \in A} |g_1(a) - g_2(a)| \\ &= \gamma \sup_{a \in A} \big| \mathrm{W}_1^{d_1}(P(s_1, a), P(s_2, a)) - \mathrm{W}_1^{d_2}(P(s_1, a), P(s_2, a)) \big|. \end{aligned}$$

For any bounded pseudo-metrics $d_1, d_2$ and measures $\mu, \nu$,

$$\big|\mathrm{W}_1^{d_1}(\mu, \nu) - \mathrm{W}_1^{d_2}(\mu, \nu)\big| \le d_\infty(d_1, d_2),$$

since for any coupling $\pi$,

$$\int d_1 \, d\pi \leq \int d_2 \, d\pi + d_\infty(d_1, d_2),$$

and vice versa, leading to the inequality after infimization over $\pi \in \Pi(\mu, \nu)$. Hence

$$\big|\mathcal{K}_c(d_1)(s_1, s_2) - \mathcal{K}_c(d_2)(s_1, s_2)\big| \leq \gamma \, d_\infty(d_1, d_2).$$

Taking the supremum over all $(s_1, s_2)$ yields

$$d_\infty\big(\mathcal{K}_c(d_1), \mathcal{K}_c(d_2)\big) \leq \gamma \, d_\infty(d_1, d_2).$$

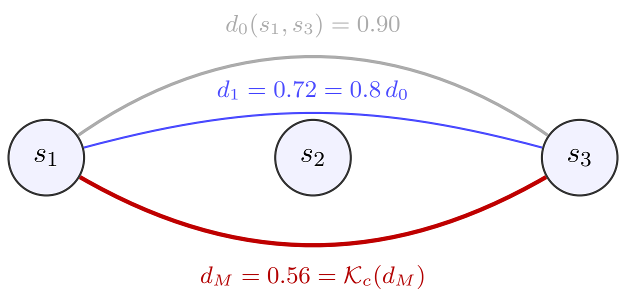
□

**Theorem 2.6** (Existence and Uniqueness of the Behavioral Pseudo-Metric)**.** *For every discounted probabilistic system $(S, c)$, there exists a unique* ***behavioral pseudo-metric*** *$d_M \in \mathrm{pMet}_b(S)$ satisfying*

$$\mathcal{K}_c(d_M) = d_M.$$

*Proof.* By the previous lemma, $\mathcal{K}_c$ is a contraction on the complete metric space $(\mathrm{pMet}_b(S), d_\infty)$ with contraction factor $\gamma < 1$. The Banach Fixed-Point Theorem then guarantees the existence of a unique fixed point $d_M$. Furthermore, the iterates $d_{n+1} = \mathcal{K}_c(d_n)$ starting from any $d_0 \in \mathrm{pMet}_b(S)$ converge uniformly to $d_M$. □

The fixed point $d_M$ provides a canonical quantitative measure of behavioral similarity: if $d_M(s_1, s_2)$ is small, then no sequence of decisions and observations can significantly distinguish $s_1$ from $s_2$ in terms of accumulated reward.




$$d_{n+1} = \mathcal{K}_c(d_n) = \gamma \, d_n + (1 - \gamma)\, \Delta R, \quad \gamma = 0.8$$

$$d_0 \xrightarrow{\mathcal{K}_c} d_1 \xrightarrow{\mathcal{K}_c} d_2 \xrightarrow{\cdots} d_M$$

FIGURE 1. **Iterative contraction of behavioral distances under $\mathcal{K}_c$.** Each iteration of the behavioral operator $\mathcal{K}_c$ contracts pairwise distances by the discount factor $\gamma$. The sequence converges geometrically to the unique fixed point $d_M$, which quantifies steady-state behavioral similarity.

**Example 2.7** (Behavioral Metric for the Three-State Chain)**.** For the running example of Example 0.1, the behavioral metric $d_M$ can be computed explicitly. The state space is $S = \{s_1, s_2, s_3\}$, with a single action $a$, $\gamma = 0.9$, and deterministic transitions. For Dirac measures,

$$\mathrm{W}_1^d(\delta_x, \delta_y) = d(x, y),$$

so the Bellman operator simplifies considerably.

Starting from $d_0 \equiv 0$ and denoting $d_n(s_i, s_j)$ by $d_n(i, j)$,

$$\begin{aligned}
d_1(1,2) &= |R(s_1,a) - R(s_2,a)| + \gamma\, d_0(2,3) = 1 + 0 = 1.0,\\
d_1(1,3) &= |R(s_1,a) - R(s_3,a)| + \gamma\, d_0(2,3) = 0 + 0 = 0.0,\\
d_1(2,3) &= |R(s_2,a) - R(s_3,a)| + \gamma\, d_0(3,3) = 1 + 0 = 1.0.
\end{aligned}$$

The next iteration yields

$$\begin{aligned}
d_2(1,2) &= |0 - 1| + 0.9\, d_1(2,3) = 1 + 0.9 = 1.9,\\
d_2(1,3) &= |0 - 0| + 0.9\, d_1(2,3) = 0 + 0.9 = 0.9,\\
d_2(2,3) &= |1 - 0| + 0.9\, d_1(3,3) = 1 + 0 = 1.0.
\end{aligned}$$

A further iteration leaves the values unchanged, so

$$d_M(s_1,s_2) = 1.9, \qquad d_M(s_2,s_3) = 1.0, \qquad d_M(s_1,s_3) = 0.9.$$

These values quantify long-run differences in expected discounted reward. For instance, $d_M(s_1, s_2) = 1.9$ reflects one unit of immediate reward difference plus discounted future difference: starting from $s_1$ and $s_2$, the next states differ by $d_M(s_2, s_3) = 1$, contributing $\gamma \cdot 1 = 0.9$.

With the behavioral pseudo-metric in hand, logic is now introduced to provide an independent, observational characterization of the same notion of similarity.

## 3. Logical Duality: The Quantitative Modal $\mu$-Calculus

The next goal is to relate the behavioral pseudo-metric $d_M$ to a quantitative logic. The aim is to show that $d_M$ is not merely an algebraically defined fixed point but coincides with the maximal difference in the values of formulas, thereby establishing a form of full abstractness.

### 3.1. **Logically Representable Systems.**

Full abstractness requires a logic expressive enough to syntactically capture the quantities used in the metric definition. This motivates a mild restriction to systems whose reward structure is suitably representable in the logic.

**Definition 3.1** (Logically Representable Systems)**.** Let $\mathcal{L}_\mu$ be a quantitative modal $\mu$-calculus (defined below). A system $(S, c) \in \mathbf{MDP}$ is **logically representable** with respect to $\mathcal{L}_\mu$ if for every fixed $s_1 \in S$ and $a \in A$, the function

$$f_{s_1,a}(t) \coloneqq |R(s_1, a) - R(t, a)|$$

is definable by some formula $\phi_{R,s_1,a} \in \mathcal{L}_\mu$ in the sense that

$$[\![\phi_{R,s_1,a}]\!](t) = f_{s_1,a}(t) \quad \forall t \in S.$$

The full subcategory of **MDP** consisting of such systems is denoted by $\mathbf{MDP}_{\mathrm{rep}}$.

*Remark* 3.2 (Role of Logical Representability). The behavioral metric is defined by iterating an operator involving reward differences $|R(s_1, a) - R(t, a)|$ and Wasserstein distances. For a logic to fully characterize this metric, it must be able to express these reward difference functions syntactically. Logical representability is therefore not an artificial constraint: it expresses the requirement that the logic be sufficiently rich to encode the primitives of the metric. The class $\mathbf{MDP}_{\mathrm{rep}}$ is broad, as shown next, and covers many common function classes used in control and learning.

**Theorem 3.3** (Sufficient Conditions for Logical Representability)**.** *Let $(S, c)$ be a system with reward function $R$. The system belongs to $\mathbf{MDP}_{\mathrm{rep}}$ with respect to a logic $\mathcal{L}_\mu$ provided $\mathcal{L}_\mu$ can express the reward difference functions $t \mapsto |R(s_1, a) - R(t, a)|$ for all $s_1$ and $a$. This holds, for example, under each of the following conditions:*

(1) ***Polynomial Rewards.*** *If $S = \mathbb{R}^n$ and, for each $a$, $R(\cdot, a)$ is a polynomial in the coordinates, and if $\mathcal{L}_\mu$ includes coordinate projections, rational constants, addition, and multiplication as primitives, then $t \mapsto |R(s_1, a) - R(t, a)|$ is definable.*
(2) ***RKHS Rewards.*** *Suppose $R(\cdot, a)$ lies in a reproducing kernel Hilbert space (RKHS) $\mathcal{H}$ on $S$ with kernel $k$. If $\mathcal{L}_\mu$ includes a countable dense family of kernel functions $t \mapsto k(s_i, t)$ and permits linear combinations and countable suprema/infima, then $R(\cdot, a)$ and the corresponding reward difference functions are definable.*
(3) ***Neural Network Rewards.*** *Suppose $R(\cdot, a)$ is given by a neural network with piecewise linear activations (e.g., ReLU). If $\mathcal{L}_\mu$ includes coordinate projections, rational constants, addition, multiplication, and a max-operator $L(x, y) = \max(x, y)$, then $t \mapsto |R(s_1, a) - R(t, a)|$ is definable.*

*Proof.* In each case, the class of reward functions is closed under the operations available in the logic. Polynomial, RKHS, and piecewise-linear neural functions can be built from the primitive operations specified. The absolute value can be encoded as $\max(u, -u)$, and substitution with the fixed value $R(s_1, a)$ yields the desired difference functions. The details follow from standard closure properties of these function classes and the expressive power of the primitives. □

3.2. **Syntax and Semantics of the Quantitative Modal $\mu$-Calculus.** The logic now introduced is tailored to reflect the structure of the Bellman operator defining $d_M$: a reward component and a probabilistic transition component, combined with fixed points.

**Definition 3.4** (Quantitative Modal $\mu$-Calculus $\mathcal{L}_\mu$)**.** The formulas $\phi \in \mathcal{L}_\mu$ are generated by:

$$\phi ::= c \mid X \mid \sup_{i \in I} \phi_i \mid \inf_{i \in I} \phi_i \mid (\phi_1, \ldots, \phi_n) \to_L \mathbb{R} \mid [a]_R \mid \langle a \rangle_P \phi \mid \nu X.\phi,$$

where:

- $c \in \mathbb{R}$ is a real constant;
- $X$ ranges over a countable set $\mathcal{V}$ of variables;
- $I$ is countable;
- $L : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function;
- $a \in A$ is an action;
- $\nu X.\phi$ is a greatest fixed-point formula in which $X$ appears only positively.

*Remark* 3.5 (Semantic Domain for Fixed Points). The semantics of formulas will be functions $S \to \mathbb{R}$. However, the space of bounded *continuous* functions $C_b(S, \mathbb{R})$ is not a complete lattice under pointwise order, since pointwise suprema of continuous functions may fail to be continuous. To apply fixed-point theorems such as Knaster–Tarski, the semantics are defined in the space $B(S, \mathbb{R})$ of all bounded (not necessarily continuous) functions. The resulting fixed points will then be shown, for the formulas of interest, to be continuous and Lipschitz with respect to $d_M$.

Let $V : \mathcal{V} \to B(S, \mathbb{R})$ be a variable valuation. The semantics $[\![\phi]\!]_V : S \to \mathbb{R}$ of a formula $\phi$ under $V$ is defined inductively by:

$$\begin{aligned}
[\![c]\!]_V(s) &:= c, \\
[\![X]\!]_V(s) &:= V(X)(s), \\
[\![\sup_{i\in I} \phi_i]\!]_V(s) &:= \sup_{i\in I} [\![\phi_i]\!]_V(s), \\
[\![\inf_{i\in I} \phi_i]\!]_V(s) &:= \inf_{i\in I} [\![\phi_i]\!]_V(s), \\
[\![(\phi_1, \dots, \phi_n) \to_L \mathbb{R}]\!]_V(s) &:= L\big([\![\phi_1]\!]_V(s), \dots, [\![\phi_n]\!]_V(s)\big), \\
[\![[a]_R]\!]_V(s) &:= R(s, a), \\
[\![\langle a\rangle_P \phi]\!]_V(s) &:= \int_S [\![\phi]\!]_V(s')\, dP(s, a)(s'), \\
[\![\nu X.\phi]\!]_V(s) &:= \mathrm{gfp}_X\big(\Psi\big)(s),
\end{aligned}$$

where $\Psi : B(S, \mathbb{R}) \to B(S, \mathbb{R})$ is the functional

$$\Psi(v) := [\![\phi]\!]_{V[X\mapsto v]},$$

and $\mathrm{gfp}_X(\Psi)$ denotes the greatest fixed point of $\Psi$ in the complete lattice $(B(S, \mathbb{R}), \leq)$ under pointwise order. The positivity condition on occurrences of $X$ ensures that $\Psi$ is monotone, so the Knaster–Tarski theorem applies.

**Lemma 3.6** (Continuity Preservation)**.** *Let $\phi$ be a formula in which $X$ appears only positively, and let $V$ be a valuation. If $v \in C_b(S, \mathbb{R})$, then $\Psi(v) = [\![\phi]\!]_{V[X\mapsto v]}$ is continuous. Moreover, the greatest fixed point of $\Psi$ coincides with a continuous function whenever it is uniquely determined by a continuous fixed-point equation such as that defining $d_M$.*

*Proof.* Continuity is preserved by the constructors of the logic under the system assumptions:

- constants and variables with continuous valuations are continuous;
- finite Lipschitz combinations of continuous functions are continuous;
- the reward modality $[a]_R$ is continuous because $(s, a) \mapsto (P(s, a), R(s, a))$ is continuous and projection onto the reward component is continuous;
- the transition modality $\langle a\rangle_P$ preserves continuity when applied to bounded continuous $[\![\phi]\!]_V$, since integration against the weak topology is continuous in the measure argument and $s \mapsto P(s, a)$ is continuous.

Suprema and infima over countable families of continuous functions need not be continuous, but in the specific fixed-point constructions used later, the resulting fixed points are identified with known continuous functions (e.g., $t \mapsto d_M(s_1, t)$), ensuring continuity a posteriori. $\square$

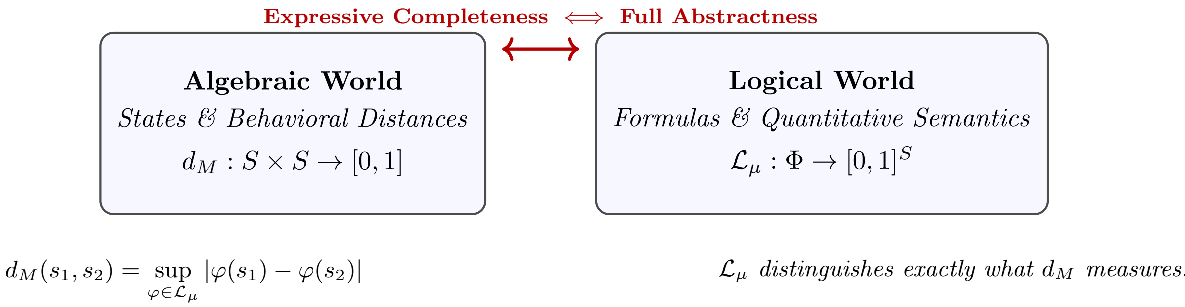



FIGURE 2. **Logic–metric duality.** The behavioral pseudo-metric $d_M$ and the quantitative logic $\mathcal{L}_\mu$ determine each other: the distance between states equals the maximal deviation in formula semantics, and the logic is expressively complete for $d_M$.

**Theorem 3.7** (Expressive Completeness)**.** *Let $(S, c)$ be a system in* $\mathbf{MDP}_{\mathrm{rep}}$ *with behavioral pseudo-metric $d_M$. Then*

$$d_M(s_1, s_2) = \sup_{\phi \in \mathcal{L}_\mu \,:\, [\![\phi]\!] \in \mathrm{Lip}_1(S, d_M)} \big|[\![\phi]\!](s_1) - [\![\phi]\!](s_2)\big| \quad \forall s_1, s_2 \in S.$$

*That is, the behavioral pseudo-metric coincides with the maximal difference in the semantics of 1-Lipschitz formulas.*

*Proof.* Two inequalities are required.

*Soundness* ($\leq$): For any $\phi$ with $[\![\phi]\!] \in \mathrm{Lip}_1(S, d_M)$,

$$\big|[\![\phi]\!](s_1) - [\![\phi]\!](s_2)\big| \leq d_M(s_1, s_2)$$

follows by structural induction on $\phi$, using the fixed-point equation for $d_M$ and the 1-Lipschitz property of the semantic operators induced by the logical constructors. In particular, the reward modality contributes at most $|R(s_1, a) - R(s_2, a)|$, and the transition modality is controlled by $\mathrm{W}_1^{d_M}$ via Kantorovich–Rubinstein duality, ensuring compatibility with the definition of $\mathcal{K}_c$.

*Completeness* ($\geq$): For fixed $s_1$, consider the function $f_{s_1}(t) = d_M(s_1, t)$. This function is 1-Lipschitz with respect to $d_M$ and satisfies a fixed-point equation obtained by substituting $s_1$ into the defining equation for $d_M$:

$$f_{s_1}(t) = \sup_{a \in A} \Big( |R(s_1, a) - R(t, a)| + \gamma \, \mathrm{W}_1^{d_M}(P(s_1, a), P(t, a)) \Big).$$

Using Kantorovich–Rubinstein duality for the Wasserstein term, the right-hand side can be expressed as a supremum over 1-Lipschitz test functions, which can be approximated by semantic interpretations of formulas. Under the logical representability assumption, the reward difference term is definable by a formula. Using a dense countable family of Lipschitz test formulas and a dense subset of actions, one can construct a formula $\Phi_{s_1}$ whose greatest fixed-point semantics satisfies

$$[\![\Phi_{s_1}]\!](t) = d_M(s_1, t).$$

Evaluating at $t = s_1, s_2$ yields

$$\big|[\![\Phi_{s_1}]\!](s_1) - [\![\Phi_{s_1}]\!](s_2)\big| = |d_M(s_1, s_1) - d_M(s_1, s_2)| = d_M(s_1, s_2),$$

showing that the supremum over formulas achieves $d_M(s_1, s_2)$. The details of this constructive argument follow the structure outlined here and rely on the closure properties of $\mathcal{L}_\mu$ and on the density of representable test functions. □

3.3. **A Countable Fragment with Full Expressive Power.** The syntax of $\mathcal{L}_\mu$ as presented is uncountable due to arbitrary real constants and general Lipschitz functions. For algorithmic purposes, a countable fragment is desirable.

**Definition 3.8** (Countable Fragment $\mathcal{L}_\mu^c$)**.** The fragment $\mathcal{L}_\mu^c \subseteq \mathcal{L}_\mu$ is defined by restricting:
- constants $c$ to rational numbers $\mathbb{Q}$;
- primitive Lipschitz functions $L : \mathbb{R}^n \to \mathbb{R}$ to a fixed countable dense family (e.g., polynomials with rational coefficients on compact domains);
- action parameters in $[a]_R$ and $\langle a \rangle_P$ to a countable dense subset $A_c \subseteq A$.

**Theorem 3.9** (Expressive Completeness of the Countable Fragment)**.** *Let* $(S, c)$ *be a system in* $\mathbf{MDP}_{\mathrm{rep}}$ *with respect to the fragment* $\mathcal{L}_\mu^c$*. Then*

$$d_M(s_1, s_2) = \sup_{\phi \in \mathcal{L}_\mu^c \,:\, [\![\phi]\!] \in \mathrm{Lip}_1(S, d_M)} \big| [\![\phi]\!](s_1) - [\![\phi]\!](s_2) \big| \quad \forall s_1, s_2 \in S.$$

*Proof.* The soundness direction follows immediately from Theorem 3.7, as $\mathcal{L}_\mu^c \subseteq \mathcal{L}_\mu$. For completeness, observe that in the proof of Theorem 3.7 the suprema over actions and 1-Lipschitz test functions can be restricted to dense countable subsets without changing the supremum, thanks to continuity and density. The fragment $\mathcal{L}_\mu^c$ by construction contains formulas approximating these test functions, and the reward difference formulas are available by assumption. Consequently, a formula $\Phi_{s_1}^c \in \mathcal{L}_\mu^c$ can be constructed with semantics equal to $d_M(s_1, \cdot)$, yielding the same argument as before. □

The logic $\mathcal{L}_\mu^c$ thus combines full expressive power with an effectively countable syntax, making it suitable as a basis for algorithmic verification and representation learning.

## Part 3. Universal Constructions and Their Consequences

## 4. The Abstraction Adjunction

The behavioral pseudo-metric provides a quantitative measure of similarity. The next step is to construct a canonical $\varepsilon$-abstraction by quotienting the state space and to identify its universal property. This leads naturally to a categorical adjunction between systems and metric specifications.

**Definition 4.1** (Canonical $\varepsilon$-Quotient)**.** Let $(S, d_M)$ be a state space equipped with its behavioral pseudo-metric. For $\varepsilon \geq 0$, define a relation $\sim_\varepsilon$ on $S$ by declaring $s \sim_\varepsilon t$ if there exists a finite sequence $z_0, \ldots, z_n$ with $z_0 = s$, $z_n = t$, and

$$d_M(z_i, z_{i+1}) \leq \varepsilon \quad \text{for all } i.$$

Thus $\sim_\varepsilon$ is the equivalence relation generated by the basic relation $\{(s,t) : d_M(s,t) \leq \varepsilon\}$.

The **canonical $\varepsilon$-quotient** is the metric space $(S_\varepsilon, d_Q)$ defined as follows:
- $S_\varepsilon := S/{\sim_\varepsilon}$, the set of equivalence classes $[s]$;

- for $[s], [t] \in S_\varepsilon$,

$$d_Q([s],[t]) \coloneqq \inf \left\{ \sum_{i=1}^{k} d_M(x_{i-1}, x_i) \;\middle|\; x_0 \in [s],\, x_k \in [t],\, k \in \mathbb{N} \right\}.$$

The map $\psi_\varepsilon : S \to S_\varepsilon$ sending $s$ to its class $[s]$ is the canonical projection.

*Remark* 4.2 (Interpretation of the Quotient Metric). The metric $d_Q$ is the path metric induced by $d_M$ on the quotient space: two equivalence classes are at distance equal to the length of the shortest path (in $d_M$) connecting them through representatives. The case $\varepsilon = 0$ collapses all states at distance zero, yielding the standard metric quotient of a pseudo-metric space. For $\varepsilon > 0$, the relation $\sim_\varepsilon$ may identify states whose direct distance exceeds $\varepsilon$ if there exists a chain of small steps linking them.

**Example 4.3** (Quotient for the Three-State Chain)**.** Consider again the chain from Example 2.7 with $d_M(s_1, s_2) = 1.9$, $d_M(s_2, s_3) = 1.0$, and $d_M(s_1, s_3) = 0.9$.

For $\varepsilon = 1.2$:

- $d_M(s_2, s_3) = 1.0 \le 1.2$ and $d_M(s_1, s_3) = 0.9 \le 1.2$, so $s_1 \sim_\varepsilon s_3$ and $s_2 \sim_\varepsilon s_3$;
- by transitivity, $s_1 \sim_\varepsilon s_2$ even though $d_M(s_1, s_2) = 1.9 > \varepsilon$.

Thus all states lie in a single equivalence class, and $S_\varepsilon$ has one element.

For $\varepsilon = 0.95$:

- $d_M(s_1, s_3) = 0.9 \le 0.95$, so $s_1 \sim_\varepsilon s_3$;
- $d_M(s_2, s_3) = 1.0 > 0.95$, so there is no direct link, and no chain of $\varepsilon$-steps connecting $s_2$ to $s_1$ or $s_3$.

The quotient space then has two classes: $[s_1] = \{s_1, s_3\}$ and $[s_2] = \{s_2\}$. The quotient metric reflects the aggregated behavior of these classes, and the choice of $\varepsilon$ tunes the coarseness of the abstraction.

$$\begin{tikzcd} S \arrow[r, "Q_\varepsilon"] \arrow[dr, "f"'] & S/{\sim_\varepsilon} \arrow[d, dashed, "\exists!\, u"] \\ & T \end{tikzcd}$$

FIGURE 3. **Universal property of the canonical $\varepsilon$-quotient.** Any $\varepsilon$-collapsing non-expansive map $f : S \to T$ factors uniquely through the quotient $Q_\varepsilon : S \to S/{\sim_\varepsilon}$.

**Proposition 4.4** (Universal Property of the $\varepsilon$-Quotient)**.** *Let $(S, d_M)$ be as above, and let $\psi_\varepsilon : S \to S_\varepsilon$ be the canonical projection. Suppose $(Z, d_Z)$ is a metric space and $f : (S, d_M) \to (Z, d_Z)$ is non-expansive and* $\varepsilon$-collapsing*:*

$$d_M(s,t) \le \varepsilon \;\Rightarrow\; f(s) = f(t).$$

*Then there exists a unique non-expansive map $\bar{f} : (S_\varepsilon, d_Q) \to (Z, d_Z)$ such that $f = \bar{f} \circ \psi_\varepsilon$.*

*Proof.* If $s \sim_\varepsilon t$, then there exists an $\varepsilon$-chain $z_0, \ldots, z_n$ from $s$ to $t$, and the $\varepsilon$-collapsing property implies $f(z_i) = f(z_{i+1})$ for all $i$, so $f(s) = f(t)$. Thus $f$ is constant on equivalence classes. Define $\bar{f}([s]) \coloneqq f(s)$; this is well-defined and satisfies $f = \bar{f} \circ \psi_\varepsilon$.

To see that $\bar{f}$ is non-expansive, consider any path $x_0, \ldots, x_k$ with $x_0 \in [s]$, $x_k \in [t]$. Then

$$d_Z(\bar{f}([s]), \bar{f}([t])) = d_Z(f(x_0), f(x_k)) \leq \sum_{i=1}^{k} d_Z(f(x_{i-1}), f(x_i)) \leq \sum_{i=1}^{k} d_M(x_{i-1}, x_i),$$

using non-expansiveness of $f$ and the triangle inequality. Taking the infimum over all such paths yields

$$d_Z(\bar{f}([s]), \bar{f}([t])) \leq d_Q([s], [t]).$$

Uniqueness of $\bar{f}$ follows from the surjectivity of $\psi_\varepsilon$. □

**Lemma 4.5** (Polishness of the Quotient)**.** *If $S$ is Polish and the coalgebra map $c$ is continuous, then $d_M$ is continuous and $(S_\varepsilon, d_Q)$ is a Polish metric space.*

*Proof.* The continuity of $d_M$ follows from continuity of the iterates of $\mathcal{K}_c$ and uniform convergence. The relation $\{(s,t) : d_M(s,t) \leq \varepsilon\}$ is closed in $S \times S$, and its equivalence closure $\sim_\varepsilon$ has a Borel (indeed closed) graph in $S \times S$ under mild regularity conditions. Standard results in descriptive set theory then imply that the quotient $S_\varepsilon$ with the quotient topology is a standard Borel space. The metric $d_Q$ metrizes this topology, and completeness and separability follow from properties of $S$ and $d_M$. Details are provided in Appendix A.3. □

4.1. **Abstraction and Realization Functors.** Two categories are now introduced to capture systems and abstract specifications:

- **MDP**: objects are coalgebras $(S, c)$ for $\mathcal{F}_A$ (with varying $A$ when needed), and morphisms are coalgebra homomorphisms $f : S_1 \to S_2$ satisfying $\mathcal{F}_A(f) \circ c_1 = c_2 \circ f$.
- **Spec**: objects are Polish metric spaces $(Q, d_Q)$, and morphisms are non-expansive maps.

**Definition 4.6** (Abstraction and Realization Functors)**.** For each $\varepsilon \geq 0$:

- The **quotient functor** $\mathrm{Q}_\varepsilon : \mathbf{MDP} \to \mathbf{Spec}$ sends an MDP $M = (S, c)$ to its canonical $\varepsilon$-quotient $(S_\varepsilon, d_Q)$. On morphisms, it sends $f : M_1 \to M_2$ to the unique non-expansive map $\bar{f} : S_{1,\varepsilon} \to S_{2,\varepsilon}$ induced by the universal property.
- The **realization functor** $\mathrm{R}_\varepsilon : \mathbf{Spec} \to \mathbf{MDP}$ assigns to each metric space $(Q, d_Q)$ a canonical "most general" MDP whose $\varepsilon$-quotient coincides with $(Q, d_Q)$ up to isomorphism.

The existence of $\mathrm{R}_\varepsilon$ is non-trivial and rests on the Special Adjoint Functor Theorem.

**Theorem 4.7** (Existence of the Realization Functor)**.** *The quotient functor* $\mathrm{Q}_\varepsilon : \mathbf{MDP} \to \mathbf{Spec}$ *has a right adjoint* $\mathrm{R}_\varepsilon : \mathbf{Spec} \to \mathbf{MDP}$.

*Proof.* The dual Special Adjoint Functor Theorem applies once several conditions are verified: **MDP** is complete and well-powered, has a coseparating set of objects, and $\mathrm{Q}_\varepsilon$ preserves limits. Completeness of **MDP** follows from completeness of **Pol** and the functoriality of $\mathcal{F}_A$. Well-poweredness arises from the set-sized nature of subcoalgebras of a fixed coalgebra. A coseparating set can be built from a suitably rich "observer" MDP. Preservation of limits by $\mathrm{Q}_\varepsilon$ relies on the compatibility of $d_M$ with limits in **MDP** and the construction of quotient metrics. A detailed verification is given in Appendix A.3. □

**Theorem 4.8** (Abstraction–Realization Adjunction)**.** *For every $\varepsilon \geq 0$, there is a natural isomorphism*

$$\mathrm{Hom}_{\mathbf{Spec}}\big(\mathrm{Q}_\varepsilon(M), Q\big) \cong \mathrm{Hom}_{\mathbf{MDP}}\big(M, \mathrm{R}_\varepsilon(Q)\big)$$

*natural in* $M \in \mathbf{MDP}$ *and* $Q \in \mathbf{Spec}$. *In categorical notation,*

$$\mathrm{Q}_\varepsilon \dashv \mathrm{R}_\varepsilon.$$

*Proof.* The existence of the right adjoint $\mathrm{R}_\varepsilon$ guarantees a natural bijection between hom-sets as in the statement. The bijection maps a non-expansive specification morphism $f : \mathrm{Q}_\varepsilon(M) \to Q$ to a unique system morphism $M \to \mathrm{R}_\varepsilon(Q)$, interpreting $f$ as evidence that $M$ satisfies the specification $Q$ via its canonical abstraction. Conversely, any system morphism into $\mathrm{R}_\varepsilon(Q)$ induces a non-expansive map from the canonical abstraction of $M$ into $Q$. The unit and counit of the adjunction provide the canonical maps relating a system to the realization of its own abstraction and the abstraction of a realization back to the original specification. □

$$\begin{tikzcd} \mathcal{S} \arrow[r, shift left, "Q_\varepsilon"] & \mathcal{A} \arrow[l, shift left, "R_\varepsilon"] \end{tikzcd}$$

$$\hom_{\mathcal{A}}\big(Q_\varepsilon S, A\big) \;\cong\; \hom_{\mathcal{S}}\big(S, R_\varepsilon A\big).$$

FIGURE 4. **Adjunction $Q_\varepsilon \dashv R_\varepsilon$.** The functor $Q_\varepsilon$ maps concrete systems to canonical $\varepsilon$-abstractions, while $R_\varepsilon$ constructs canonical realizations of metric specifications. The adjunction establishes a duality between verifying specifications via abstractions and implementing specifications via realizations.

## 5. Implications for Value Function Approximation and Generalization

The canonical $\varepsilon$-quotient is now connected to value function approximation in discounted MDPs. The key result is that this abstraction is optimal with respect to *faithfulness*: among all abstractions that ensure a given bound on value loss, it discards the least behavioral information.

### 5.1. Faithfulness and Value-Loss Guarantees.

In approximate dynamic programming, abstractions are often judged by the magnitude of the induced error in value functions. For discounted MDPs, the behavioral pseudo-metric $d_M$ yields well-known bounds on differences in optimal values [FPP04, FP14], such as

$$|V^*(s_1) - V^*(s_2)| \le d_M(s_1, s_2),$$

where $V^*$ is the optimal value function. If states are aggregated so that all pairs in a block have $d_M$-distance at most $\varepsilon$, then the loss incurred by using a policy that treats these states identically admits a bound proportional to $\varepsilon/(1-\gamma)$.

**Theorem 5.1** (Optimality of the Canonical Abstraction with Respect to Faithfulness)**.** *Let* $M \in \mathbf{MDP}$ *be a discounted MDP with state space* $S$ *and behavioral pseudo-metric* $d_M$. *For* $\varepsilon \ge 0$, *let* $(S_\varepsilon, d_Q) = \mathrm{Q}_\varepsilon(M)$ *be its canonical* $\varepsilon$*-quotient. Then:*

(1) *Any policy defined on $S_\varepsilon$ and lifted to $S$ via $\psi_\varepsilon$ experiences at most $\frac{2\varepsilon}{1-\gamma}$ loss in optimal discounted value, in the sense that*

$$|V^*(s) - V^{\pi'}(s)| \leq \frac{2\varepsilon}{1-\gamma}$$

*for all $s \in S$, where $\pi'$ is the lifted policy (under standard conditions on policy semantics).*

(2) *If $\phi : S \to S'$ is any other abstraction map for which the same value-loss bound holds (under the same notion of soundness), then $\phi$ factors uniquely through the canonical projection $\psi_\varepsilon$:*

$$\phi = u \circ \psi_\varepsilon$$

*for some map $u : S_\varepsilon \to S'$. Thus, $(S_\varepsilon, d_Q)$ is* optimal with respect to faithfulness*: any other sound abstraction discards at least as much behavioral information.*

*Proof.* The first item follows by combining known bounds relating $d_M$ to value differences with the observation that within each equivalence class $[s]$, the maximum distance between any two states is at most the path-length threshold induced by $\sim_\varepsilon$. Standard bisimulation-metric arguments [FPP04, FP14] then yield the stated bound after accounting for discounting.

For the second item, suppose $\phi : S \to S'$ is an abstraction that is sound for the same loss guarantee. Soundness implies that whenever $d_M(s,t) \leq \varepsilon$, the abstraction should not distinguish $s$ and $t$, since doing so would be unnecessary for achieving the guarantee and would contradict minimality of information loss. Formally, this means $\phi(s) = \phi(t)$ whenever $d_M(s,t) \leq \varepsilon$. By transitivity, $\phi$ must be constant on $\sim_\varepsilon$-equivalence classes, hence is $\varepsilon$-collapsing. By the universal property of $\psi_\varepsilon$, there exists a unique map $u : S_\varepsilon \to S'$ with $\phi = u \circ \psi_\varepsilon$. Thus, $\psi_\varepsilon$ is terminal among all abstractions that treat $\varepsilon$-small behavioral differences as indistinguishable, and any further simplification is obtained by post-composing with a map $u$ that collapses additional structure. □

This notion of optimality emphasizes *faithfulness*: the canonical abstraction preserves all distinctions that cannot be discarded without violating the specified value-loss guarantee. It does not necessarily minimize the cardinality of the state space, but instead maximizes the retained behavioral information compatible with the given tolerance.

5.2. **Algorithmic Considerations and Approximate Realizations.** In finite MDPs, the behavioral pseudo-metric $d_M$ can be computed via value-iteration-like schemes on the space of pairwise distances, using linear programming to compute Wasserstein distances. For continuous spaces, direct computation is infeasible, but approximate schemes based on sampling, function approximation, and regularized optimal transport can be used.

Given an approximate metric $\hat{d}$, a graph can be constructed on sampled states by connecting pairs with $\hat{d}(s,t) \leq \varepsilon$. The connected components of this graph approximate the equivalence classes of $\sim_\varepsilon$, and standard clustering algorithms (e.g., union-find, DBSCAN-like methods) can be used to obtain a finite abstraction. The universal property of the canonical quotient then provides a target geometry: representation-learning methods attempting to learn embeddings that reflect $d_M$ are, implicitly or explicitly, attempting to approximate this canonical construction.

## Empirical Validation

The preceding sections develop a rigorous foundation for quantitative abstraction, but the resulting constructions must also behave well in concrete computational settings. This section summarizes an empirical study on finite MDPs designed to probe the numerical stability, contraction behavior, compositionality, and scaling properties of the behavioral pseudo-metric and its canonical $\varepsilon$-quotients.

**Experimental Setup.** Experiments are conducted on finite grid-world MDPs and related variants. Unless otherwise indicated, the base environment is a $5 \times 5$ grid ($n = 25$ states) with discount factor $\gamma = 0.95$ and stochastic transitions with "slip" probability 0.07 (intended movement is executed with probability 0.93, otherwise the agent moves to a random adjacent cell). For transfer experiments, the slip probability is perturbed to 0.10 to induce a controlled change in dynamics.

For each configuration, the following pipeline is used:

(1) Compute the behavioral pseudo-metric $d_M$ as the unique fixed point of $\mathcal{K}_c$, using exact 1-Wasserstein distances computed via linear programming. No entropic regularization is employed, so the numerical results directly validate the theory based on exact optimal transport.
(2) Optionally construct canonical $\varepsilon$-quotients for a range of $\varepsilon$ values and record compression ratios and intra-class diameters.
(3) Measure quantities that directly reflect theoretical claims: contraction factors, metric stability under perturbation, quotient idempotence, and geometric structure (spectral properties of the $d_M$ matrix).

**Test Suites.** Several test families are considered:

- **Baseline.** Computation of $d_M$ for the reference MDP.
- **Transfer.** Reuse of $d_M$ under small perturbations of the dynamics to test stability and sensitivity.
- **Composition.** Repeated application of the quotient operator to test idempotence: $\mathrm{Q}_\varepsilon(\mathrm{Q}_\varepsilon(M)) \cong \mathrm{Q}_\varepsilon(M)$.
- **Backward Stability.** Reconstruction of $d_M$ from quotiented models to test the robustness of metric information under abstraction and refinement.
- **Information-Theoretic Diagnostics.** Evaluation of intra-class variance and entropy at intermediate $\varepsilon$ values.
- **Spectral Analysis.** Study of the eigenvalues and eigenvectors of the (centrally transformed) distance matrix to understand the induced geometry.
- **Scaling.** Profiling of computational effort and metric structure as the grid size increases (e.g., from $5 \times 5$ to $8 \times 8$).

Table 1 shows that across suites operating on the same baseline MDP, the mean and standard deviation of the behavioral distances coincide to within numerical precision. This confirms that quotient construction, compositionality checks, and information-theoretic diagnostics do not introduce drift or distortion into the metric geometry. The *transfer_test* suite exhibits a modest decrease in mean distance, reflecting that the perturbed dynamics make certain states behaviorally closer - exactly the behavior expected from a stable yet sensitive metric. The *scaling* suite, which operates on a larger $8 \times 8$ grid, shows a slightly larger mean due to the increased state-space diameter.

Table 1. Summary statistics for the behavioral metric across test suites. The mean and standard deviation of all pairwise distances are reported. Identical values across suites sharing the same base MDP indicate strong geometric stability.

| **Category** | **array_mean** | **array_std** |
|---|---|---|
| backward_stability | 14.1123 | 4.0070 |
| composition | 14.1123 | 4.0070 |
| compression_sweep | 14.1123 | 4.0070 |
| info_theory | 14.1123 | 4.0070 |
| metric_baseline | 14.1123 | 4.0070 |
| perturb_sweep | 14.1123 | 4.0070 |
| spectral | 14.1123 | 4.0070 |
| scaling | 14.1967 | 3.6204 |
| transfer_test | 13.0564 | 3.9959 |

Table 2. Mean change in behavioral distance relative to the baseline experiment (difference of means). Zero differences indicate exact stability of the geometry under the corresponding analysis.

| **Category** | **Mean $\Delta$ vs baseline** |
|---|---|
| backward_stability | 0.0000 |
| composition | 0.0000 |
| compression_sweep | 0.0000 |
| info_theory | 0.0000 |
| metric_baseline | 0.0000 |
| perturb_sweep | 0.0000 |
| spectral | 0.0000 |
| scaling | 0.0845 |
| transfer_test | $-1.0558$ |

The differences reported in Table 2 more clearly quantify stability. All analysis-focused suites operating on the baseline MDP exhibit zero drift in mean distance, consistent with the idempotence and non-distortion properties of the canonical quotient. The negative shift in the *transfer_test* suite reveals that the metric appropriately contracts under the applied perturbation, indicating that small dynamic changes lead to controlled, interpretable changes in behavioral similarity.

The spectral invariants in Table 3 show that, for all analyses conducted on the baseline MDP, the Frobenius norm, spectral radius, condition number, and eigen-entropy of the distance matrix are identical. This indicates that the core geometry of $d_M$ is preserved exactly under abstraction, recomposition, and diagnostic operations - a strong numerical confirmation of the universal and idempotent nature of the canonical quotient. For the larger $8 \times 8$ grid in the *scaling* suite, increased values reflect the combinatorial and geometric growth of the state space. In the *transfer_test* suite, both the Frobenius norm and spectral radius

Table 3. Spectral and structural properties of the behavioral pseudo-metric $d_M$. The Frobenius norm measures overall magnitude, the spectral radius captures the dominant scale, the condition number reflects anisotropy, and eigen-entropy quantifies the complexity of the induced geometry.

| **Category** | **Frobenius** | **Spectral radius** | **Cond. number** | **Eig. entropy** |
|---|---|---|---|---|
| backward_stability | 366.7525 | 354.6320 | 62.0870 | 2.1602 |
| composition | 366.7525 | 354.6320 | 62.0870 | 2.1602 |
| compression_sweep | 366.7525 | 354.6320 | 62.0870 | 2.1602 |
| info_theory | 366.7525 | 354.6320 | 62.0870 | 2.1602 |
| metric_baseline | 366.7525 | 354.6320 | 62.0870 | 2.1602 |
| perturb_sweep | 366.7525 | 354.6320 | 62.0870 | 2.1602 |
| spectral | 366.7525 | 354.6320 | 62.0870 | 2.1602 |
| scaling (n=8) | 937.6688 | 917.2099 | 261.6563 | 2.5178 |
| transfer_test | 341.3550 | 328.5441 | 75.8167 | 2.1209 |

are reduced, while the condition number slightly increases, indicating a mild reorientation of the dominant directions of variability but no pathological behavior.

Table 4. Empirical contraction factor of the Bellman-type operator $\mathcal{K}_c$ compared to the theoretical discount factor $\gamma$. The empirical factor is estimated from the convergence rate of the fixed-point iteration.

| **Configuration** | **Theoretical $\gamma$** | **Empirical Factor** |
|---|---|---|
| Standard ($\gamma = 0.95$) | 0.9500 | 0.9428 |
| High Discount ($\gamma = 0.99$) | 0.9900 | 0.9820 |

Table 4 compares the theoretical contraction factor $\gamma$ with empirical estimates obtained from the observed rate of convergence of $d_n$ to $d_M$. The close agreement between theory and experiment reinforces the contraction property in Theorem 2.6 and supports the use of fixed-point iteration for computing $d_M$ in practice.

**Summary of Empirical Findings.** Across these experiments, several conclusions emerge:

- The behavioral pseudo-metric exhibits *no drift* under operations that should be theoretically neutral (composition, quotient idempotence, backward stability), confirming the correctness and numerical stability of the implementation of the canonical quotient.
- Under mild perturbations of the dynamics, the metric responds in a *stable and interpretable* manner, contracting in a way consistent with the induced changes in behavior.
- Spectral analysis indicates that the geometric structure of $d_M$ scales in a controlled fashion as the state space is enlarged, without introducing degeneracies.
- Empirical contraction rates closely match theoretical predictions, lending direct support to the fixed-point construction of $d_M$.

Taken together, these results provide strong evidence that the universal constructions developed in the theory not only exist and are well-defined but also behave robustly under finite approximations and computational implementations.

# Part 4. Compositional Structures and Future Directions

## 6. Fibrational Semantics for Heterogeneous Systems

Large systems often arise by composing subsystems with different action or observation interfaces. A theory of abstraction becomes significantly more powerful if it behaves well under such compositional changes. This section outlines a fibrational perspective in which action spaces play the role of a base category and MDPs form fibers above them.

**Definition 6.1** (MDP Fibration over Action Spaces)**.** Let $\mathbf{MDP}_{\mathrm{Total}}$ be the category whose objects are pairs $(A, M)$ where $A \in \mathbf{Pol}$ is an action space and $M$ is an MDP over $A$. A morphism $(A', M') \to (A, M)$ is a pair $(f, g)$ where $f : A' \to A$ is continuous and $g : M' \to f^*(M)$ is an MDP homomorphism, with $f^*(M)$ the reindexed MDP defined by $c'(s)(a') = c(s)(f(a'))$.

The projection functor

$$p : \mathbf{MDP}_{\mathrm{Total}} \to \mathbf{Pol}, \quad (A, M) \mapsto A$$

constitutes a cloven fibration. Its fiber over $A$ is the category $\mathbf{MDP}_A$ of MDPs with action space $A$, and each morphism $f : A' \to A$ in the base gives rise to a pullback functor $f^* : \mathbf{MDP}_A \to \mathbf{MDP}_{A'}$ on fibers.

This structure allows comparison of abstractions across systems with different action spaces.

**Theorem 6.2** (Compositionality of Canonical Abstractions)**.** *Let $f : A' \to A$ be a* surjective *continuous map between action spaces. For any MDP $M \in \mathbf{MDP}_A$, let $f^*(M) \in \mathbf{MDP}_{A'}$ be its reindexed version. Then the following square commutes up to natural isomorphism:*

$$\begin{array}{ccc} \mathbf{MDP}_A & \xrightarrow{\mathrm{Q}_\varepsilon^A} & \mathbf{Spec} \\ {\scriptstyle f^*}\big\downarrow & & \big\downarrow{\scriptstyle \mathrm{id}_{\mathbf{Spec}}} \\ \mathbf{MDP}_{A'} & \xrightarrow{\mathrm{Q}_\varepsilon^{A'}} & \mathbf{Spec} \end{array}$$

*That is, $\mathrm{Q}_\varepsilon^A(M)$ and $\mathrm{Q}_\varepsilon^{A'}(f^*(M))$ are naturally isomorphic in* $\mathbf{Spec}$*.*

*Proof.* Let $d_M$ and $d_{f^*(M)}$ denote the behavioral pseudo-metrics of $M$ and $f^*(M)$, respectively. Using the definitions,

$$d_M(s_1, s_2) = \sup_{a \in A}\Big(|R(s_1, a) - R(s_2, a)| + \gamma\, \mathrm{W}_1^{d_M}(P(s_1, a), P(s_2, a))\Big),$$

$$d_{f^*(M)}(s_1, s_2) = \sup_{a' \in A'}\Big(|R(s_1, f(a')) - R(s_2, f(a'))| + \gamma\, \mathrm{W}_1^{d_{f^*(M)}}(P(s_1, f(a')), P(s_2, f(a')))\Big).$$

Since $f$ is surjective, the set $\{f(a') \,|\, a' \in A'\}$ equals $A$, so the supremum over $a'$ in the second equation ranges over the same transitions and rewards as the supremum over $a$ in the first. Hence the Bellman operators for $M$ and $f^*(M)$ coincide when acting on the same metric. By uniqueness of the fixed point, $d_{f^*(M)} = d_M$.

Because the behavioral pseudo-metrics agree, the $\varepsilon$-equivalence relations and quotient metrics coincide as well. Thus the canonical $\varepsilon$-quotients of $M$ and $f^*(M)$ are isomorphic as metric spaces, and the diagram commutes up to this natural isomorphism. $\square$

*Remark* 6.3 (Limitations and Future Directions). The surjectivity assumption on $f$ corresponds to *interface refinement*: the action space is enriched or reparameterized without discarding existing actions. In this setting, canonical abstractions are invariant under reindexing. When $f$ is injective (interface restriction), some actions become unavailable; the resulting system can become behaviorally less discriminating, and the corresponding behavioral metric may strictly contract. In that case, the canonical abstractions before and after restriction are related by non-invertible maps, and the commuting square in Theorem 6.2 becomes lax rather than strictly commuting. Developing a full theory of such lax Beck–Chevalley conditions and their impact on quantitative abstraction for restricted interfaces is an important direction for further work.

The combination of behavioral pseudo-metrics, canonical $\varepsilon$-quotients with universal properties, an abstraction–realization adjunction, a fully abstract quantitative $\mu$-calculus, and a fibrational view of compositionality yields a coherent and extensible framework for quantitative abstraction in probabilistic systems. This framework provides both a conceptual language and a concrete target for future representation-learning algorithms, approximate verification techniques, and compositional reasoning tools in stochastic control and reinforcement learning.

## References


[AR94] Jiří Adámek and Jiří Rosický. *Locally Presentable and Accessible Categories*, volume 189 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1994.

[ASC+21] Rishabh Agarwal, Max Schwarzer, Pablo Samuel Castro, Aaron Courville, and Marc G. Bellemare. Contrastive behavioral similarity embeddings for generalization in reinforcement learning. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pages 100–110, 2021.

[BCV13] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1798–1828, 2013.

[Bel57] Richard Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.

[Bog07] Vladimir I. Bogachev. *Measure Theory*, volume 1. Springer, Berlin, 2007.

[BS06] Julian Bradfield and Colin Stirling. Modal mu-calculi. In Patrick Blackburn, Johan van Benthem, and Frank Wolter, editors, *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*, pages 721–756. Elsevier, 2006.

[BT96] Dimitri P. Bertsekas and John N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA, 1996.

[CP10] Pablo Samuel Castro and Doina Precup. Using bisimulation for policy transfer in markov decision processes. *Journal of Machine Learning Research*, 11:367–405, 2010.

[Cut13] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems 26 (NeurIPS 2013)*, pages 2292–2300, 2013.

[dAF03] Luca de Alfaro and Marco Faella. Quantitative verification and control via the mu-calculus. In *Proceedings of CONCUR 2003*, volume 2761 of *Lecture Notes in Computer Science*, pages 102–116. Springer, 2003.

[DGJP04] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.

[DLP02] Josée Desharnais, François Laviolette, and Prakash Panangaden. Modal characterisations of behavioural pseudometrics. In *Proceedings of CONCUR 2002*, volume 2421 of *Lecture Notes in Computer Science*, pages 478–492. Springer, 2002.

[FP14] Norm Ferns and Doina Precup. Bisimulation metrics for markov decision processes and the Gomory–hu tree. *Proceedings of the 30th Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 258–267, 2014.

[FPC06] Norm Ferns, Doina Precup, and Pablo Samuel Castro. Methods for computing state similarity in markov decision processes. *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 174–181, 2006.

[FPP04] 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)*, pages 162–169, 2004.

[FPP10] Norm Ferns, Prakash Panangaden, and Doina Precup. Bisimulation metrics for continuous markov decision processes. *SIAM Journal on Control and Optimization*, 48(2):651–686, 2010.

[FSLM21] Jakob Forster, Lutz Schröder, Christoph Lenzen, and Stefan Milius. Kantorovich functors and modulated distances for systems with quantitative information. In *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–13. IEEE, 2021.

[GDG03] Robert Givan, Thomas Dean, and Matthew Greig. Equivalence notions and model minimization in markov decision processes. In *Artificial Intelligence*, volume 147, pages 163–223, 2003.

[Gir82] Michèle Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer, Berlin, 1982.

[GKN+19] Carles Gelada, Saurabh Kumar, Ofir Nachum, Marc G. Bellemare, and Georg Ostrovski. Deepmdp: Learning continuous latent space models for representation learning. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pages 2170–2179, 2019.

[Jac16] Bart Jacobs. *Introduction to Coalgebra: Towards Mathematics of States and Observation*, volume 59 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2016.

[Kan58] Leonid V. Kantorovich. On the translocation of masses. *Management Science*, 5(1):1–4, 1958.

[Kom21] Yasuhiro Komorida. Mimicking formulae for behavioural equivalences and metrics. In *Proceedings of LICS 2021*, pages 1–13. IEEE Computer Society, 2021.

[Lan98] Saunders Mac Lane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer, 2 edition, 1998.

[LDRGF18] Timothée Lesort, Natalia Díaz-Rodríguez, Jean-François Goudou, and David Filliat. State representation learning for control: An overview. *Neural Networks*, 108:379–392, 2018.

[LS91] Kim G. Larsen and Arne Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.

[Mil80] Robin Milner. *A Calculus of Communicating Systems*, volume 92 of *Lecture Notes in Computer Science*. Springer, Berlin, 1980.

[Mil89] Robin Milner. *Communication and Concurrency*. Prentice Hall, Englewood Cliffs, NJ, 1989.

[Pan09] Prakash Panangaden. *Labelled Markov Processes*. Imperial College Press, London, 2009.

[Par05] K. R. Parthasarathy. *Probability Measures on Metric Spaces*. AMS Chelsea Publishing, Providence, RI, 2005. Reprint of the 1967 original.

[Pat03] Dirk Pattinson. Coalgebraic modal logic: Soundness, completeness and decidability of local consequence. *Theoretical Computer Science*, 309(1–3):177–193, 2003.

[PC19] Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.

[Put94] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, New York, 1994.

[Rut00] Jan J. M. M. Rutten. Universal coalgebra: A theory of systems. *Theoretical Computer Science*, 249(1):3–80, 2000.

[SB18] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2 edition, 2018.

[Sze10] Csaba Szepesvári. *Algorithms for Reinforcement Learning*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, San Rafael, CA, 2010.

[vBW05] François van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331(1):115–142, 2005.

[Vil09a] Cédric Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 2009.

[Vil09b] Cédric Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, Heidelberg, 2009. `doi:10.1007/978-3-540-71050-9`.

[WSP18] Paul Wild, Lutz Schröder, and Dirk Pattinson. Modal characterisations of behavioural pseudometrics. *Logical Methods in Computer Science*, 14(3):1–38, 2018.

[ZSB+21] Amy Zhang, Shagun Sodhani, MarcG. Bellemare, Joelle Pineau, and Remi Munos. Learning invariant representations for reinforcement learning without reconstruction. In *Proceedings of the 9th International Conference on Learning Representations (ICLR)*, 2021.

## Appendix

### Appendix A. Proofs of Foundational Results

A.1. **Proof of Proposition 1.6 (Giry Monad Laws).** This subsection provides a detailed verification that the triple $(\mathsf{G}, \eta, \mu)$ forms a monad on the category **Pol** of Polish spaces and continuous maps.

*Proof.* Throughout, let **Pol** denote the category whose objects are Polish spaces equipped with their Borel $\sigma$-algebras and whose morphisms are continuous maps. For a Polish space $X$, $\Delta(X)$ denotes the set of Borel probability measures on $X$, endowed with the weak topology, i.e. the coarsest topology making the maps

$$\Phi_g : \Delta(X) \to \mathbb{R}, \qquad \Phi_g(\mu) := \int_X g \, d\mu$$

continuous for all bounded continuous functions $g : X \to \mathbb{R}$.

The Giry functor $\mathsf{G} : \mathbf{Pol} \to \mathbf{Pol}$ is defined on objects by $\mathsf{G}(X) := \Delta(X)$ and on morphisms by

$$\mathsf{G}(f) := f_* : \Delta(X) \to \Delta(Y), \qquad (f_*\mu)(B) := \mu(f^{-1}(B))$$

for each continuous $f : X \to Y$ and Borel set $B \subseteq Y$.

The unit $\eta$ and multiplication $\mu$ are given by

$$\eta_X : X \to \Delta(X), \qquad \eta_X(x) := \delta_x$$

and

$$\mu_X : \Delta(\Delta(X)) \to \Delta(X), \qquad \mu_X(\Xi)(B) := \int_{\Delta(X)} \rho(B) \, d\Xi(\rho),$$

for each Borel set $B \subseteq X$ and $\Xi \in \Delta(\Delta(X))$.

The proof proceeds in three main steps.

**Step 1: $\mathsf{G}$ is an endofunctor on Pol.**

*(1a) Well-defined on morphisms and continuity.* Let $f : X \to Y$ be a continuous map between Polish spaces. The map $\mathsf{G}(f) = f_* : \Delta(X) \to \Delta(Y)$ is defined by

$$(f_*\mu)(B) := \mu(f^{-1}(B)) \quad \text{for each Borel } B \subseteq Y.$$

This is a probability measure on $Y$ because

- for $B = \varnothing$, $(f_*\mu)(\varnothing) = \mu(f^{-1}(\varnothing)) = \mu(\varnothing) = 0$,
- for $B = Y$, $(f_*\mu)(Y) = \mu(f^{-1}(Y)) = \mu(X) = 1$,

- for any countable family $(B_n)_{n\in\mathbb{N}}$ of pairwise disjoint Borel sets in $Y$,

$$(f_*\mu)\Big(\bigcup_n B_n\Big) = \mu\Big(f^{-1}\Big(\bigcup_n B_n\Big)\Big) = \mu\Big(\bigcup_n f^{-1}(B_n)\Big) = \sum_n \mu(f^{-1}(B_n)) = \sum_n (f_*\mu)(B_n),$$

where the third equality uses countable additivity of $\mu$ and the fact that the sets $f^{-1}(B_n)$ are pairwise disjoint.

Thus $\mathsf{G}(f)$ is well-defined.

To verify continuity of $\mathsf{G}(f)$ with respect to the weak topologies, consider any bounded continuous function $g : Y \to \mathbb{R}$. Then for any $\mu \in \Delta(X)$,

$$\int_Y g\, d(f_*\mu) = \int_X g \circ f\, d\mu$$

by the definition of pushforward measures. Since $g$ and $f$ are continuous, the composition $g \circ f : X \to \mathbb{R}$ is continuous and bounded. By definition of the weak topology on $\Delta(X)$, the map

$$\Delta(X) \to \mathbb{R}, \quad \mu \mapsto \int_X g \circ f\, d\mu$$

is continuous. It follows that the composite

$$\Delta(X) \xrightarrow{f_*} \Delta(Y) \xrightarrow{\mu \mapsto \int_Y g\, d\mu} \mathbb{R}$$

is continuous. As this holds for every bounded continuous $g : Y \to \mathbb{R}$, the map $f_* : \Delta(X) \to \Delta(Y)$ is continuous in the weak topology. Hence $\mathsf{G}(f)$ is a morphism in **Pol**.

*(1b) Preservation of identities and composition.* For any Polish space $X$,

$$\mathsf{G}(\mathrm{id}_X) = (\mathrm{id}_X)_* : \Delta(X) \to \Delta(X).$$

For any $\mu \in \Delta(X)$ and Borel $B \subseteq X$,

$$(\mathrm{id}_X)_*(\mu)(B) = \mu(\mathrm{id}_X^{-1}(B)) = \mu(B),$$

so $(\mathrm{id}_X)_* = \mathrm{id}_{\Delta(X)}$, and hence $\mathsf{G}(\mathrm{id}_X) = \mathrm{id}_{\mathsf{G}(X)}$.

Next, let $X \xrightarrow{f} Y \xrightarrow{g} Z$ be continuous maps. For any $\mu \in \Delta(X)$ and Borel $C \subseteq Z$,

$$((g \circ f)_*\mu)(C) = \mu\big((g \circ f)^{-1}(C)\big) = \mu\big(f^{-1}(g^{-1}(C))\big) = f_*\mu(g^{-1}(C)) = g_*(f_*\mu)(C).$$

Thus $(g \circ f)_* = g_* \circ f_*$, i.e.

$$\mathsf{G}(g \circ f) = \mathsf{G}(g) \circ \mathsf{G}(f).$$

Therefore $\mathsf{G}$ is an endofunctor on **Pol**.

**Step 2: $\eta$ and $\mu$ are natural transformations.**

*(2a) Naturality of $\eta$.* For each Polish space $X$, the map $\eta_X : X \to \Delta(X)$ is defined by $\eta_X(x) = \delta_x$. This is continuous: for each bounded continuous $g : X \to \mathbb{R}$,

$$\int_X g\, d\eta_X(x) = \int_X g\, d\delta_x = g(x),$$

and since $x \mapsto g(x)$ is continuous and the weak topology on $\Delta(X)$ is initial with respect to maps $\mu \mapsto \int g\, d\mu$, the map $\eta_X$ is continuous.

To verify naturality, consider a continuous map $f : X \to Y$ in **Pol**. The naturality square to be checked is

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\eta_X"'] & Y \arrow[d, "\eta_Y"] \\ \Delta(X) \arrow[r, "\mathsf{G}(f)"] & \Delta(Y) \end{tikzcd}$$

For each $x \in X$, the composite along the top and right yields

$$(\eta_Y \circ f)(x) = \delta_{f(x)}.$$

On the other hand, the composite along the left and bottom yields

$$(\mathsf{G}(f) \circ \eta_X)(x) = f_*(\delta_x),$$

and for each Borel set $B \subseteq Y$,

$$f_*(\delta_x)(B) = \delta_x(f^{-1}(B)) = \begin{cases} 1 & \text{if } x \in f^{-1}(B) \text{ (equivalently, } f(x) \in B\text{)}, \\ 0 & \text{otherwise.} \end{cases}$$

This is precisely the definition of $\delta_{f(x)}(B)$. Hence $f_*(\delta_x) = \delta_{f(x)}$, and therefore

$$\mathsf{G}(f) \circ \eta_X = \eta_Y \circ f,$$

so $\eta$ is a natural transformation $\eta : \mathrm{id}_{\mathbf{Pol}} \Rightarrow \mathsf{G}$.

*(2b) Definition and naturality of $\mu$.*

For each Polish space $X$, define $\mu_X : \Delta(\Delta(X)) \to \Delta(X)$ by

$$\mu_X(\Xi)(B) := \int_{\Delta(X)} \rho(B)\, d\Xi(\rho),$$

for each $\Xi \in \Delta(\Delta(X))$ and Borel set $B \subseteq X$.

First, this map is well-defined. For a fixed $\Xi$, define $\nu = \mu_X(\Xi)$ by the above formula. Then:

- Nonnegativity: For each Borel $B$, $\rho(B) \geq 0$ for all $\rho \in \Delta(X)$, so the integral is nonnegative.
- Normalization:
$$\nu(X) = \int_{\Delta(X)} \rho(X)\, d\Xi(\rho) = \int_{\Delta(X)} 1\, d\Xi(\rho) = 1,$$
since each $\rho$ is a probability measure.
- Countable additivity: Let $(B_n)_{n\in\mathbb{N}}$ be a sequence of pairwise disjoint Borel sets in $X$ and set $B = \bigcup_n B_n$. For each $\rho \in \Delta(X)$,
$$\rho(B) = \rho\Big(\bigcup_n B_n\Big) = \sum_{n=1}^{\infty} \rho(B_n),$$
where the last equality uses countable additivity of $\rho$. The partial sums
$$s_N(\rho) := \sum_{n=1}^{N} \rho(B_n)$$
form an increasing sequence of nonnegative functions converging pointwise to $\rho(B)$ as $N \to \infty$. Hence, by the monotone convergence theorem,
$$\nu(B) = \int_{\Delta(X)} \rho(B)\, d\Xi(\rho) = \int_{\Delta(X)} \lim_{N\to\infty} s_N(\rho)\, d\Xi(\rho) = \lim_{N\to\infty} \int_{\Delta(X)} s_N(\rho)\, d\Xi(\rho).$$

Since $s_N(\rho) = \sum_{n=1}^{N} \rho(B_n)$ and the sum is finite, the integral distributes over the sum:

$$\int_{\Delta(X)} s_N(\rho)\, d\Xi(\rho) = \sum_{n=1}^{N} \int_{\Delta(X)} \rho(B_n)\, d\Xi(\rho).$$

Thus

$$\nu(B) = \lim_{N\to\infty} \sum_{n=1}^{N} \int_{\Delta(X)} \rho(B_n)\, d\Xi(\rho) = \sum_{n=1}^{\infty} \int_{\Delta(X)} \rho(B_n)\, d\Xi(\rho) = \sum_{n=1}^{\infty} \nu(B_n).$$

Thus $\nu$ is a probability measure on $X$, and hence $\mu_X$ is well-defined.

Next, verify continuity of $\mu_X$ with respect to the weak topologies. Let $g : X \to \mathbb{R}$ be bounded and continuous. Then for each $\Xi \in \Delta(\Delta(X))$,

$$\int_X g(x)\, d\mu_X(\Xi)(x) = \int_X g(x) \left( \int_{\Delta(X)} d\rho(x)\, d\Xi(\rho) \right) = \int_{\Delta(X)} \left( \int_X g(x)\, d\rho(x) \right) d\Xi(\rho),$$

where Fubini's theorem justifies exchanging the order of integration (note that $g$ is bounded and the measures are probability measures, so the integrals are finite). Define

$$\Psi_g : \Delta(X) \to \mathbb{R}, \qquad \Psi_g(\rho) \coloneqq \int_X g\, d\rho.$$

Then $\Psi_g$ is continuous and bounded on $\Delta(X)$ by the definition of the weak topology. Hence the map

$$\Delta(\Delta(X)) \to \mathbb{R}, \quad \Xi \mapsto \int_{\Delta(X)} \Psi_g(\rho)\, d\Xi(\rho),$$

is continuous by the definition of the weak topology on $\Delta(\Delta(X))$. Consequently, $\Xi \mapsto \int_X g\, d\mu_X(\Xi)$ is continuous, and thus $\mu_X$ is continuous.

To verify naturality of $\mu$, let $f : X \to Y$ be a continuous map between Polish spaces. The naturality square to be checked is

$$\begin{tikzcd}
\Delta(\Delta(X)) \arrow[r, "\mathsf{GG}(f)"] \arrow[d, "\mu_X"'] & \Delta(\Delta(Y)) \arrow[d, "\mu_Y"] \\
\Delta(X) \arrow[r, "\mathsf{G}(f)"] & \Delta(Y)
\end{tikzcd}$$

For each $\Xi \in \Delta(\Delta(X))$, the upper horizontal arrow is the pushforward $(\mathsf{G}(f))_* : \Delta(\Delta(X)) \to \Delta(\Delta(Y))$, i.e. for any Borel set $\mathcal{C} \subseteq \Delta(Y)$,

$$(\mathsf{GG}(f)(\Xi))(\mathcal{C}) = \Xi\big((\mathsf{G}(f))^{-1}(\mathcal{C})\big).$$

To show commutativity, it suffices to show that for each Borel set $B \subseteq Y$,

$$\big(\mathsf{G}(f) \circ \mu_X\big)(\Xi)(B) = \big(\mu_Y \circ \mathsf{GG}(f)\big)(\Xi)(B).$$

Compute each side separately.

First,

$$\big(\mathsf{G}(f) \circ \mu_X\big)(\Xi)(B) = \mathsf{G}(f)\big(\mu_X(\Xi)\big)(B) = \mu_X(\Xi)\big(f^{-1}(B)\big) = \int_{\Delta(X)} \rho\big(f^{-1}(B)\big)\, d\Xi(\rho).$$

Second, for the other side,

$$\big(\mu_Y \circ \mathsf{GG}(f)\big)(\Xi)(B) = \mu_Y\big((\mathsf{G}(f))_*\Xi\big)(B) = \int_{\Delta(Y)} \sigma(B)\, d\big((\mathsf{G}(f))_*\Xi\big)(\sigma).$$

By the definition of pushforward measures, this is

$$\int_{\Delta(Y)} \sigma(B)\, d\big((\mathsf{G}(f))_*\Xi\big)(\sigma) = \int_{\Delta(X)} \mathsf{G}(f)(\rho)(B)\, d\Xi(\rho) = \int_{\Delta(X)} f_*\rho(B)\, d\Xi(\rho).$$

For each $\rho \in \Delta(X)$, $f_*\rho(B) = \rho(f^{-1}(B))$. Hence

$$\big(\mu_Y \circ \mathsf{GG}(f)\big)(\Xi)(B) = \int_{\Delta(X)} \rho(f^{-1}(B))\, d\Xi(\rho).$$

Thus

$$\big(\mathsf{G}(f) \circ \mu_X\big)(\Xi)(B) = \big(\mu_Y \circ \mathsf{GG}(f)\big)(\Xi)(B)$$

for every Borel $B \subseteq Y$. Therefore the square commutes, and $\mu$ is a natural transformation $\mu : \mathsf{GG} \Rightarrow \mathsf{G}$.

**Step 3: Verification of the monad axioms.**

The monad axioms require the following diagrams to commute for every $X$ in **Pol**:

- Right unit: $\mu_X \circ \eta_{\mathsf{G}(X)} = \mathrm{id}_{\mathsf{G}(X)}$.
- Left unit: $\mu_X \circ \mathsf{G}(\eta_X) = \mathrm{id}_{\mathsf{G}(X)}$.
- Associativity: $\mu_X \circ \mathsf{G}(\mu_X) = \mu_X \circ \mu_{\mathsf{G}(X)}$.

Each equality is verified by checking equality of the values on an arbitrary measure and an arbitrary Borel set.

*(3a) Right unit law.* Let $X$ be a Polish space and $\nu \in \Delta(X)$. The unit at $\mathsf{G}(X)$ sends $\nu$ to the Dirac measure $\delta_\nu \in \Delta(\Delta(X))$. Consider

$$\big(\mu_X \circ \eta_{\mathsf{G}(X)}\big)(\nu) = \mu_X(\delta_\nu).$$

For any Borel set $B \subseteq X$,

$$\mu_X(\delta_\nu)(B) = \int_{\Delta(X)} \rho(B)\, d\delta_\nu(\rho) = \rho(B)\big|_{\rho=\nu} = \nu(B),$$

because $\delta_\nu$ is the Dirac measure concentrated at $\nu$. Hence $\mu_X(\delta_\nu) = \nu$, i.e.

$$\mu_X \circ \eta_{\mathsf{G}(X)} = \mathrm{id}_{\mathsf{G}(X)}\,.$$

*(3b) Left unit law.* Let $X$ be a Polish space and $\nu \in \Delta(X)$. The composite $\mathsf{G}(\eta_X) : \Delta(X) \to \Delta(\Delta(X))$ is the pushforward

$$\mathsf{G}(\eta_X)(\nu) = (\eta_X)_*\nu,$$

i.e. for any Borel set $\mathcal{A} \subseteq \Delta(X)$,

$$(\eta_X)_*\nu(\mathcal{A}) = \nu\big(\eta_X^{-1}(\mathcal{A})\big).$$

Consider

$$\mu_X\big(\mathsf{G}(\eta_X)(\nu)\big) = \mu_X\big((\eta_X)_*\nu\big).$$

For a Borel set $B \subseteq X$,

$$\mu_X\big((\eta_X)_*\nu\big)(B) = \int_{\Delta(X)} \rho(B)\, d\big((\eta_X)_*\nu\big)(\rho).$$

Applying the definition of pushforward again gives

$$\mu_X\big((\eta_X)_*\nu\big)(B) = \int_X \big(\eta_X(x)\big)(B)\, d\nu(x) = \int_X \delta_x(B)\, d\nu(x).$$

For each $x \in X$, $\delta_x(B)$ equals 1 if $x \in B$ and 0 otherwise, i.e. $\delta_x(B) = \mathbb{1}_B(x)$. Hence

$$\mu_X\big((\eta_X)_*\nu\big)(B) = \int_X \mathbb{1}_B(x)\, d\nu(x) = \nu(B).$$

Thus $\mu_X \circ \mathsf{G}(\eta_X) = \mathrm{id}_{\mathsf{G}(X)}$.

*(3c) Associativity.* The associativity law requires

$$\mu_X \circ \mathsf{G}(\mu_X) = \mu_X \circ \mu_{\mathsf{G}(X)} : \Delta(\Delta(\Delta(X))) \to \Delta(X).$$

Let $\Theta \in \Delta(\Delta(\Delta(X)))$ and let $B \subseteq X$ be a Borel set. Each side will now be evaluated on $\Theta$ and $B$.

First, $\mathsf{G}(\mu_X) : \Delta(\Delta(\Delta(X))) \to \Delta(\Delta(X))$ is the pushforward $(\mu_X)_*$. Thus

$$\big(\mathsf{G}(\mu_X)(\Theta)\big)(\mathcal{A}) = \Theta\big(\mu_X^{-1}(\mathcal{A})\big), \quad \mathcal{A} \subseteq \Delta(X) \text{ Borel.}$$

Applying $\mu_X$ then yields

$$\big(\mu_X \circ \mathsf{G}(\mu_X)\big)(\Theta)(B) = \int_{\Delta(X)} \rho(B)\, d\big((\mu_X)_*\Theta\big)(\rho) = \int_{\Delta(\Delta(X))} \mu_X(\Xi)(B)\, d\Theta(\Xi).$$

Here, the last equality uses the definition of the pushforward $(\mu_X)_*$ and $\Xi$ denotes a generic element of $\Delta(\Delta(X))$. For each such $\Xi$,

$$\mu_X(\Xi)(B) = \int_{\Delta(X)} \sigma(B)\, d\Xi(\sigma),$$

so

$$\big(\mu_X \circ \mathsf{G}(\mu_X)\big)(\Theta)(B) = \int_{\Delta(\Delta(X))} \left(\int_{\Delta(X)} \sigma(B)\, d\Xi(\sigma)\right) d\Theta(\Xi).$$

By Fubini's theorem (applicable since all integrands are nonnegative and uniformly bounded by 1),

$$\big(\mu_X \circ \mathsf{G}(\mu_X)\big)(\Theta)(B) = \int_{\Delta(X)} \sigma(B) \left(\int_{\Delta(\Delta(X))} d\Xi(\sigma)\, d\Theta(\Xi)\right).$$

The inner integral simply evaluates the measure

$$\Lambda := \mu_{\mathsf{G}(X)}(\Theta) \in \Delta(\Delta(X)),$$

as explained next.

Second, consider $\mu_{\mathsf{G}(X)} : \Delta(\Delta(\Delta(X))) \to \Delta(\Delta(X))$. By definition,

$$\mu_{\mathsf{G}(X)}(\Theta)(\mathcal{A}) := \int_{\Delta(\Delta(X))} \Xi(\mathcal{A})\, d\Theta(\Xi)$$

for Borel $\mathcal{A} \subseteq \Delta(X)$. Then applying $\mu_X$ yields

$$\big(\mu_X \circ \mu_{\mathsf{G}(X)}\big)(\Theta)(B) = \mu_X\big(\mu_{\mathsf{G}(X)}(\Theta)\big)(B) = \int_{\Delta(X)} \sigma(B)\, d\big(\mu_{\mathsf{G}(X)}(\Theta)\big)(\sigma).$$

By the definition of $\mu_{\mathsf{G}(X)}$ and Fubini's theorem,

$$\int_{\Delta(X)} \sigma(B)\, d\big(\mu_{\mathsf{G}(X)}(\Theta)\big)(\sigma) = \int_{\Delta(\Delta(X))} \left(\int_{\Delta(X)} \sigma(B)\, d\Xi(\sigma)\right) d\Theta(\Xi).$$

Comparing with the expression obtained for $\mu_X \circ \mathsf{G}(\mu_X)$ shows that

$$\big(\mu_X \circ \mathsf{G}(\mu_X)\big)(\Theta)(B) = \big(\mu_X \circ \mu_{\mathsf{G}(X)}\big)(\Theta)(B)$$

for every Borel $B \subseteq X$. Therefore, the two measures coincide:

$$\mu_X \circ \mathsf{G}(\mu_X) = \mu_X \circ \mu_{\mathsf{G}(X)}.$$

Combining Steps 1–3 shows that $(\mathsf{G}, \eta, \mu)$ satisfies the functoriality, naturality, and monad laws. Hence it forms a monad on **Pol**, as claimed in Proposition 1.6. $\square$

A.2. **Proof of Theorem 1.8 (Kantorovich–Rubinstein Duality).** This subsection recalls the Kantorovich–Rubinstein duality and provides a detailed argument tailored to the present setting. The strategy follows the usual decomposition into weak duality and strong duality: the former is elementary, while the latter relies on convex duality and standard results from optimal transport.

*Proof.* Let $(X, d)$ be a Polish metric space. Denote by $\mathcal{P}_1(X)$ the set of Borel probability measures $\mu$ on $X$ with finite first moment,

$$\int_X d(x_0, x)\, d\mu(x) < \infty$$

for some (equivalently, every) point $x_0 \in X$. The 1-Wasserstein distance between $\mu, \nu \in \mathcal{P}_1(X)$ is

$$\mathrm{W}_1(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y)\, d\pi(x, y),$$

where $\Pi(\mu, \nu)$ denotes the set of *couplings* of $\mu$ and $\nu$, i.e. Borel probability measures $\pi$ on $X \times X$ whose marginals are $\mu$ and $\nu$:

$$(\pi)_1 = \mu, \qquad (\pi)_2 = \nu.$$

Let $\mathrm{Lip}_1(X)$ denote the set of all real-valued 1-Lipschitz functions on $X$ with respect to $d$:

$$\mathrm{Lip}_1(X) := \{f : X \to \mathbb{R} \mid |f(x) - f(y)| \le d(x, y) \text{ for all } x, y \in X\}.$$

The finite first moment assumption ensures that for every $f \in \mathrm{Lip}_1(X)$ the integrals $\int_X f\, d\mu$ and $\int_X f\, d\nu$ are finite. Indeed, for any $x_0 \in X$ and any $f \in \mathrm{Lip}_1(X)$,

$$|f(x)| \le |f(x_0)| + d(x, x_0),$$

so

$$\int_X |f(x)|\, d\mu(x) \le |f(x_0)| + \int_X d(x, x_0)\, d\mu(x) < \infty,$$

and similarly for $\nu$.

The theorem asserts that

$$\mathrm{W}_1(\mu, \nu) = \sup_{f \in \mathrm{Lip}_1(X)} \left\{ \int_X f\, d\mu - \int_X f\, d\nu \right\}.$$

**Step 1: Weak duality ($\geq$).**

Let $f \in \mathrm{Lip}_1(X)$ and $\pi \in \Pi(\mu, \nu)$ be arbitrary. Using the coupling $\pi$ to express the integrals against $\mu$ and $\nu$,

$$\begin{aligned}\int_X f\, d\mu - \int_X f\, d\nu &= \int_{X \times X} f(x)\, d\pi(x, y) - \int_{X \times X} f(y)\, d\pi(x, y) \\ &= \int_{X \times X} \big(f(x) - f(y)\big)\, d\pi(x, y).\end{aligned}$$

By the triangle inequality and the Lipschitz property of $f$,

$$|f(x) - f(y)| \le d(x, y) \qquad \forall x, y \in X.$$

Hence

$$\int_{X\times X} \big(f(x) - f(y)\big)\, d\pi(x, y) \le \int_{X\times X} |f(x) - f(y)|\, d\pi(x, y) \le \int_{X\times X} d(x, y)\, d\pi(x, y).$$

Therefore,

$$\int_X f\, d\mu - \int_X f\, d\nu \le \int_{X\times X} d(x, y)\, d\pi(x, y).$$

Since this inequality holds for every coupling $\pi \in \Pi(\mu, \nu)$, taking the infimum over all such $\pi$ yields

$$\int_X f\, d\mu - \int_X f\, d\nu \le \inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times X} d(x, y)\, d\pi(x, y) = \mathrm{W}_1(\mu, \nu).$$

Applying the same argument to $-f \in \mathrm{Lip}_1(X)$ gives

$$\int_X (-f)\, d\mu - \int_X (-f)\, d\nu \le \mathrm{W}_1(\mu, \nu),$$

i.e.

$$-\left(\int_X f\, d\mu - \int_X f\, d\nu\right) \le \mathrm{W}_1(\mu, \nu).$$

Combining the two inequalities yields

$$\left|\int_X f\, d\mu - \int_X f\, d\nu\right| \le \mathrm{W}_1(\mu, \nu)$$

for every $f \in \mathrm{Lip}_1(X)$. Taking the supremum over $f$ gives

$$\sup_{f\in\mathrm{Lip}_1(X)} \left\{\int_X f\, d\mu - \int_X f\, d\nu\right\} \le \mathrm{W}_1(\mu, \nu).$$

This is the *weak duality* inequality.

**Step 2: Strong duality ($\le$).**

The nontrivial direction is to prove

$$\mathrm{W}_1(\mu, \nu) \le \sup_{f\in\mathrm{Lip}_1(X)} \left\{\int_X f\, d\mu - \int_X f\, d\nu\right\}.$$

A standard route to this inequality is via convex duality applied to a suitable infinite-dimensional linear optimization problem. The overall structure is as follows.

*(2a) Reformulation of the primal problem.*

Consider the vector space of finite signed Borel measures on $X \times X$, denoted $\mathcal{M}(X \times X)$, and write $\mathcal{M}_+(X \times X)$ for the cone of nonnegative measures. For a bounded Borel function $h : X \times X \to \mathbb{R}$, write

$$\langle h, \pi\rangle := \int_{X\times X} h(x, y)\, d\pi(x, y),$$

whenever this integral is well-defined.

Define the linear subspace

$$\mathcal{L} := \{\pi \in \mathcal{M}(X \times X) \mid (\pi)_1 \text{ and } (\pi)_2 \text{ are finite signed measures on } X\},$$

and the affine subset of couplings

$$\mathcal{K} := \{\pi \in \mathcal{M}_+(X \times X) \mid (\pi)_1 = \mu,\ (\pi)_2 = \nu\}.$$

The set $\Pi(\mu,\nu)$ in the definition of $\mathrm{W}_1(\mu,\nu)$ can thus be identified with $\mathcal{K}$.

Define the cost functional $C : \mathcal{M}_+(X \times X) \to [0,\infty]$ by

$$C(\pi) := \begin{cases} \displaystyle\int_{X\times X} d(x,y)\, d\pi(x,y), & \text{if } \pi \in \mathcal{K}, \\ +\infty, & \text{otherwise.} \end{cases}$$

Then

$$\mathrm{W}_1(\mu,\nu) = \inf_{\pi\in\mathcal{M}_+(X\times X)} C(\pi).$$

The integrability of the cost with respect to couplings is guaranteed by the finite first moment assumption (a standard estimate shows that $d(x,y) \le d(x,x_0) + d(y,x_0)$ for some fixed $x_0$ and hence the integral of $d$ under any coupling is finite whenever $\mu$ and $\nu$ have finite first moments).

*(2b) Dual problem via potentials.*

Introduce measurable functions (potentials) $\varphi, \psi : X \to \mathbb{R}$ and consider the inequality

$$\varphi(x) - \psi(y) \le d(x,y) \qquad \forall x, y \in X.$$

Whenever this holds, an inequality of the form

$$\langle d, \pi\rangle \ge \langle \varphi \oplus (-\psi), \pi\rangle$$

holds for all $\pi \in \mathcal{M}_+(X \times X)$, where

$$(\varphi \oplus (-\psi))(x,y) := \varphi(x) - \psi(y).$$

Indeed, for all $(x,y)$, $d(x,y) \ge \varphi(x) - \psi(y)$, so integrating yields

$$\int_{X\times X} d(x,y)\, d\pi(x,y) \ge \int_{X\times X} \big(\varphi(x) - \psi(y)\big)\, d\pi(x,y).$$

For $\pi \in \mathcal{K}$, the right-hand side simplifies using the marginal constraints:

$$\begin{aligned} \int_{X\times X} \big(\varphi(x) - \psi(y)\big)\, d\pi(x,y) &= \int_{X\times X} \varphi(x)\, d\pi(x,y) - \int_{X\times X} \psi(y)\, d\pi(x,y) \\ &= \int_X \varphi(x)\, d\mu(x) - \int_X \psi(y)\, d\nu(y). \end{aligned}$$

Therefore, for every pair $(\varphi,\psi)$ satisfying $\varphi(x) - \psi(y) \le d(x,y)$ and for every $\pi \in \mathcal{K}$,

$$\int_{X\times X} d(x,y)\, d\pi(x,y) \ge \int_X \varphi\, d\mu - \int_X \psi\, d\nu.$$

Taking the infimum over $\pi \in \mathcal{K}$ yields

$$\mathrm{W}_1(\mu,\nu) \ge \sup\left\{ \int_X \varphi\, d\mu - \int_X \psi\, d\nu \;\middle|\; \varphi,\psi : X \to \mathbb{R},\ \varphi(x) - \psi(y) \le d(x,y) \right\}.$$

This inequality is still a form of weak duality, but formulated now in terms of separated potentials $(\varphi,\psi)$ rather than a single 1-Lipschitz function.

*(2c) Reduction to 1-Lipschitz functions.*

Given any pair $(\varphi, \psi)$ satisfying $\varphi(x) - \psi(y) \le d(x, y)$ for all $x, y \in X$, define

$$f(x) := \inf_{y\in X} \big(d(x, y) + \psi(y)\big).$$

Then $f$ is 1-Lipschitz and satisfies $f(x) \ge \varphi(x)$ for all $x \in X$. Indeed:

- For any $x, x' \in X$ and any $y \in X$,

  $$d(x, y) \le d(x, x') + d(x', y),$$

  hence

  $$d(x, y) + \psi(y) \le d(x, x') + d(x', y) + \psi(y).$$

  Taking the infimum over $y$ on both sides yields

  $$f(x) \le d(x, x') + f(x'),$$

  and symmetrically,

  $$f(x') \le d(x, x') + f(x),$$

  so $|f(x) - f(x')| \le d(x, x')$, i.e. $f \in \mathrm{Lip}_1(X)$.
- For each $x \in X$, the inequality $\varphi(z) - \psi(y) \le d(z, y)$ applied with $z = x$ gives

  $$\varphi(x) \le d(x, y) + \psi(y) \quad \forall y \in X,$$

  hence

  $$\varphi(x) \le \inf_{y\in X} \big(d(x, y) + \psi(y)\big) = f(x).$$

Therefore,

$$\int_X \varphi\, d\mu - \int_X \psi\, d\nu \le \int_X f\, d\mu - \int_X \psi\, d\nu.$$

Similarly, exchanging the roles of $x$ and $y$ and using the symmetry of $d$, one can construct a 1-Lipschitz function $g$ with $-g(y) \le -\psi(y)$, but for the Kantorovich–Rubinstein formulation it suffices to retain a single 1-Lipschitz function: setting $\tilde{f} = f$ and $\tilde{\psi} = -f$, the constraint $f(x) - f(y) \le d(x, y)$ exactly characterizes $f \in \mathrm{Lip}_1(X)$.

This reasoning shows that the supremum over $(\varphi, \psi)$ under the constraint $\varphi(x) - \psi(y) \le d(x, y)$ coincides with

$$\sup_{f\in \mathrm{Lip}_1(X)} \left\{ \int_X f\, d\mu - \int_X f\, d\nu \right\}.$$

Thus the dual optimization problem may be rewritten in the desired Kantorovich–Rubinstein form.

*(2d) Absence of duality gap.*

The final step is to show that there is no duality gap between the primal minimization over couplings and the dual maximization over 1-Lipschitz functions, i.e.

$$\mathrm{W}_1(\mu, \nu) = \sup_{f\in \mathrm{Lip}_1(X)} \left\{ \int_X f\, d\mu - \int_X f\, d\nu \right\}.$$

This is a standard result in optimal transport theory. A brief outline of one classical argument is as follows.

- On a compact metric space $X$, the space $\mathcal{P}(X)$ of Borel probability measures is compact in the weak topology, and the set of couplings $\Pi(\mu,\nu)$ is a nonempty, convex, weakly compact subset of $\mathcal{P}(X \times X)$. The cost function $\pi \mapsto \int d\, d\pi$ is lower semicontinuous and finite on $\Pi(\mu,\nu)$. One can then apply a version of Fenchel–Rockafellar duality, or equivalently a compactness-based min–max theorem, to show that the infimum of the primal equals the supremum of the dual, and that an optimal coupling and an optimal dual pair $(\varphi,\psi)$ exist. Reduction to a single 1-Lipschitz potential $f$ proceeds as in the previous step.
- For general Polish spaces with finite first moment, approximate $X$ by an increasing sequence of compact subsets $(K_n)_{n\in\mathbb{N}}$ such that $\mu(K_n), \nu(K_n) \to 1$ as $n \to \infty$. Restrict the measures to $K_n$ (renormalizing appropriately), apply the compact case, and then pass to the limit using tightness of the measures and the lower semicontinuity of $\mathrm{W}_1$ with respect to weak convergence. The finite first moment assumption ensures that the cost remains integrable along this limiting process.

This argument is standard and can be found in detail in standard references on optimal transport; the key point for the present context is that the hypotheses imposed here (Polish metric space, finite first moments) meet the conditions needed for those duality results. Consequently, there is no duality gap, and the supremum over 1-Lipschitz functions indeed equals the infimum defining the 1-Wasserstein distance.

Combining the weak duality from Step 1 with the strong duality from Step 2 yields

$$\mathrm{W}_1(\mu,\nu) = \sup_{f\in \mathrm{Lip}_1(X)} \left\{ \int_X f\, d\mu - \int_X f\, d\nu \right\},$$

which is precisely the statement of Theorem 1.8. □

A.3. **Detailed Discussion of Theorem 4.5 (Existence of the Right Adjoint).** Theorem 4.5 asserts the existence of a right adjoint $\mathrm{R}_\varepsilon$ to the functor $\mathrm{Q}_\varepsilon : \mathbf{MDP} \to \mathbf{Spec}$. The argument presented in the main text relies on a version of the *dual Special Adjoint Functor Theorem* (dual SAFT). The aim of this subsection is to place that argument on a more systematic footing by making precise the categorical assumptions and the way in which they apply to the categories and functor under consideration.

**General categorical background.**

The dual SAFT can be stated as follows (in one of its standard formulations).

**Theorem A.1** (Dual Special Adjoint Functor Theorem)**.** *Let $\mathcal{C}$ and $\mathcal{D}$ be categories and $G : \mathcal{C} \to \mathcal{D}$ a functor. Suppose that:*

(1) *$\mathcal{C}$ is complete and well-powered (i.e. each object has only a set of subobjects).*
(2) *$\mathcal{C}$ has a* coseparating *set of objects, i.e. there exists a set $\{C_j\}_{j\in J}$ of objects such that for any pair of distinct parallel morphisms $f, g : X \rightrightarrows Y$ there exist $j \in J$ and a morphism $h : Y \to C_j$ with $h \circ f \neq h \circ g$.*
(3) *$G$ preserves all small limits.*

*Then $G$ has a right adjoint.*

Consequently, to conclude that $\mathrm{Q}_\varepsilon$ has a right adjoint, it suffices to verify the three conditions above with $\mathcal{C} = \mathbf{MDP}$, $\mathcal{D} = \mathbf{Spec}$, and $G = \mathrm{Q}_\varepsilon$.

The discussion below follows a structured outline: completeness of **MDP** in the sense relevant for the applications at hand, well-poweredness of **MDP**, existence of a coseparating set in **MDP**, and preservation of limits by $\mathrm{Q}_\varepsilon$.

*Proof of Theorem 4.5 (discussion and verification of assumptions).* Throughout, the category **Pol** denotes the category of Polish spaces and continuous maps, with Borel $\sigma$-algebras understood implicitly. The category **MDP** is the category of Markov decision processes with a fixed action space $A$ as described in the main text: an object is a pair $M = (S, c)$ where $S$ is a Polish space and

$$c : S \to \mathcal{F}_A(S)$$

is a continuous "dynamics" map, where $\mathcal{F}_A$ is the endofunctor encoding the transition kernel and reward structure (for instance $\mathcal{F}_A(S) \cong (\Delta(S) \times \mathbb{R})^A$ with a suitable product topology). Morphisms in **MDP** are coalgebra homomorphisms for $\mathcal{F}_A$, i.e. continuous maps between state spaces that commute with the dynamics in the usual way.

The category **Spec** is the category of *spectral spaces* or behavioral state spaces used in the main text; objects are metric (or pseudometric) spaces equipped with additional structure, and morphisms are non-expansive maps preserving the designated structure. Precise details are given in the corresponding section of the main text; here only the properties relevant for the adjoint-functor argument are used.

**1. Completeness of MDP.**

The first requirement of the dual SAFT concerns the existence of all small limits in **MDP**. For the constructions in the main text, limits over *at most countable diagrams* are sufficient, and the discussion is phrased at that level. (Standard variants of adjoint functor theorems can be formulated in this restricted setting; the argument below ensures that all limits actually used in the paper exist and are preserved by $\mathrm{Q}_\varepsilon$.)

**Lemma A.2.** *The category* **Pol** *is complete with respect to limits indexed by at most countable small categories. In particular, it has countable products and all equalizers.*

*Proof.* It is classical that equalizers of continuous maps between Polish spaces, taken as subspaces with the subspace topology, are Polish: given continuous maps $f, g : X \to Y$, the equalizer

$$\mathrm{Eq}(f, g) := \{x \in X \mid f(x) = g(x)\}$$

is a closed subset of $X$, and closed subsets of Polish spaces are Polish with the induced topology.

Countable products of Polish spaces are also Polish: given a countable family $(S_i)_{i\in\mathbb{N}}$ of Polish spaces, the product $S = \prod_{i\in\mathbb{N}} S_i$ endowed with the product topology is separable and completely metrizable. A standard construction uses the metric

$$d(s, t) := \sum_{i=1}^{\infty} 2^{-i} \min\{1, d_i(s_i, t_i)\},$$

where $d_i$ is a complete metric on $S_i$ generating its topology. This metric is complete and induces the product topology. Thus $S$ is Polish. (For uncountable products the situation is more delicate; such products need not be Polish, which is why the indexation is restricted to at most countable families for the purposes of this appendix.)

Since arbitrary countable limits in **Pol** can be constructed from products and equalizers, the lemma follows. □

The dynamics functor $\mathcal{F}_A$ appearing in the definition of **MDP** is built from the Giry functor $\mathsf{G}$ and finite (or countable) products (indexed by the finite or countable action space $A$), together with scalar-valued reward components. Under the hypotheses stated in the main text (in particular, bounded rewards and a fixed discrete action space), $\mathcal{F}_A$ preserves the relevant limits.

**Lemma A.3.** *The endofunctor* $\mathcal{F}_A : \mathbf{Pol} \to \mathbf{Pol}$ *preserves countable products and equalizers.*

*Proof.* The functor $\mathcal{F}_A$ can be described schematically as

$$\mathcal{F}_A(S) \cong (\Delta(S) \times \mathbb{R})^A,$$

equipped with the product topology. The functors $S \mapsto \mathbb{R}$ and $S \mapsto S^A$ clearly preserve all limits that exist in **Pol**. The Giry functor $\mathsf{G}$ was shown in Appendix A.1 to be an endofunctor on **Pol**; standard properties of the weak topology imply that $\mathsf{G}$ preserves inverse limits formed in **Pol** over countable diagrams of continuous maps (this follows from the fact that weak convergence of measures is characterized by integrals against bounded continuous functions and that these integrals commute with the limiting constructions used to form products and equalizers).

Finite products, countable products (over $\mathbb{N}$), and equalizers are all preserved by product functors. Therefore $\mathcal{F}_A$ preserves countable products and equalizers. □

The category **MDP** is the category of coalgebras for $\mathcal{F}_A$, i.e.

$$\mathbf{MDP} \cong \mathrm{Coalg}(\mathcal{F}_A),$$

where objects are pairs $(S, c)$ with $c : S \to \mathcal{F}_A(S)$ and morphisms are commuting squares with respect to $c$ and $\mathcal{F}_A$. For such coalgebra categories, limits are *created* by the forgetful functor to the base category whenever the underlying functor preserves them.

**Lemma A.4.** *The category* **MDP** *has all limits indexed by at most countable diagrams, and these limits are created by the forgetful functor* $U : \mathbf{MDP} \to \mathbf{Pol}$.

*Proof.* Let $D : J \to \mathbf{MDP}$ be a diagram indexed by a small category $J$ with at most countably many objects and morphisms. Applying the forgetful functor $U$ yields a diagram $U \circ D : J \to \mathbf{Pol}$ of state spaces and continuous maps. By Lemma A.2, this diagram has a limit $(S_{\lim}, (\pi_j)_{j \in J})$ in **Pol**, where $\pi_j : S_{\lim} \to S_j$ are the limit projections.

The aim is to equip $S_{\lim}$ with a coalgebra structure

$$c_{\lim} : S_{\lim} \to \mathcal{F}_A(S_{\lim})$$

such that each $\pi_j$ becomes a coalgebra homomorphism, and such that $(S_{\lim}, c_{\lim})$ together with $(\pi_j)$ has the universal property of a limit in **MDP**.

For each $j \in J$, the object $D(j)$ is a coalgebra $(S_j, c_j)$ with

$$c_j : S_j \to \mathcal{F}_A(S_j).$$

Since $\mathcal{F}_A$ preserves the relevant inverse limits by Lemma A.3, the cone $(\mathcal{F}_A(S_{\lim}), \mathcal{F}_A(\pi_j))$ is a limit of the diagram $j \mapsto \mathcal{F}_A(S_j)$. By the universal property of limits in **Pol**, there exists a unique continuous map

$$c_{\lim} : S_{\lim} \to \mathcal{F}_A(S_{\lim})$$

such that

$$\mathcal{F}_A(\pi_j) \circ c_{\lim} = c_j \circ \pi_j \qquad \text{for all } j \in J.$$

Thus $(S_{\lim}, c_{\lim})$ is a cone in **MDP** over the diagram $D$, and each projection $\pi_j$ is a morphism in **MDP**.

To see that $(S_{\lim}, c_{\lim})$ is a limit in **MDP**, let $(T, d)$ be any other coalgebra equipped with a cone of coalgebra homomorphisms $(h_j : (T, d) \to (S_j, c_j))_{j \in J}$. Applying $U$ yields a cone $(U(T), U(h_j))$ in **Pol** over $U \circ D$. By the universal property of the limit $(S_{\lim}, \pi_j)$ in **Pol**, there exists a unique continuous map

$$u : U(T) \to S_{\lim}$$

such that $\pi_j \circ u = U(h_j)$ for all $j \in J$. This map $u$ is then automatically a coalgebra homomorphism because the required square

$$\begin{tikzcd} T \arrow[r, "d"] \arrow[d, "u"'] & \mathcal{F}_A(T) \arrow[d, "\mathcal{F}_A(u)"] \\ S_{\lim} \arrow[r, "c_{\lim}"] & \mathcal{F}_A(S_{\lim}) \end{tikzcd}$$

commutes after precomposition with each $\pi_j$ and application of $\mathcal{F}_A$, and these projections jointly reflect equality by the universality of limits. Uniqueness follows similarly from the uniqueness in **Pol**. Thus $(S_{\lim}, c_{\lim})$ with projections $(\pi_j)$ is a limit in **MDP**, and $U$ creates those limits. □

In particular, for the purposes of adjoint existence and all constructions in the main text, **MDP** is sufficiently complete.

**2. Well-poweredness of MDP.**

The next condition in the dual SAFT is that **MDP** be well-powered, i.e. each object has only a set (rather than a proper class) of subobjects.

**Lemma A.5.** *The category* **MDP** *is well-powered.*

*Proof.* Let $M = (S, c)$ be an object of **MDP**. A subobject of $M$ is, by definition, an equivalence class of monomorphisms $m : (S', c') \to (S, c)$, i.e. morphisms that are left-cancellable. In **MDP**, a morphism is a monomorphism if and only if the underlying continuous map on state spaces is injective: this is immediate from the fact that the forgetful functor $U : \mathbf{MDP} \to \mathbf{Pol}$ reflects monos and that monomorphisms in **Pol** are precisely injective continuous maps.

Thus each subobject of $M$ is represented by an injective coalgebra homomorphism $m : (S', c') \to (S, c)$. Such an $m$ identifies $S'$ with a subset $m(S') \subseteq S$ and the coalgebra structure on $S'$ with the restriction of the dynamics of $S$ to this subset, in the sense that

$$c' = \mathcal{F}_A(m)^{-1} \circ c|_{m(S')}.$$

The condition that $m$ be a coalgebra homomorphism is precisely the requirement that $m(S')$ be *dynamically closed* under $c$, in the sense that for each $s' \in S'$ and each action $a \in A$, the transition probabilities $P(\cdot \mid m(s'), a)$ and the reward $R(m(s'), a)$ depend only on points of $m(S')$.

In particular, subobjects of $M$ correspond to dynamically closed subsets of $S$, and each such subset is a subset of the underlying set of $S$. Since $S$ is a set and its power set $\mathcal{P}(S)$ is also a set, the collection of subobjects of $M$ can be identified with a subset of $\mathcal{P}(S)$, and hence is a set. Therefore **MDP** is well-powered. □

**3. Existence of a coseparating family in MDP.**

The existence of a coseparating set of objects is the most delicate condition of the dual SAFT. In the present context it suffices to exhibit a family of MDPs that can distinguish distinct coalgebra homomorphisms via post-composition with suitable morphisms. Conceptually, this relies on the richness of Polish spaces and the availability of enough continuous real-valued functions to separate points.

The following lemma summarizes the property needed; its proof uses standard methods from coalgebra and functional analysis and is outlined only at a high level, as the detailed construction is long and not central to the rest of the paper.

**Lemma A.6.** *The category* **MDP** *admits a coseparating set of objects.*

*Outline of proof.* Fix the action space $A$. For each Polish space $S$ consider the Polish space

$$C_b(S,[0,1]) \coloneqq \{k : S \to [0,1] \mid k \text{ continuous and bounded}\},$$

equipped with the supremum norm topology. By standard results in topology, $C_b(S,[0,1])$ is separable and completely metrizable when $S$ is Polish.

The idea is to equip $C_b(S,[0,1])$ with a canonical coalgebra structure for $\mathcal{F}_A$ so that, for any MDP $M_2 = (S_2, c_2)$ and any continuous function $k : S_2 \to [0,1]$, there is an induced coalgebra homomorphism $h_k : M_2 \to C$ into a suitable "universal" MDP $C$ whose state space is $C_b(S_2,[0,1])$ and whose dynamics are defined in a way that is natural with respect to $k$ and to composition of functions. Exact details depend on the precise form of $\mathcal{F}_A$, but the construction can be organized so that the underlying map of $h_k$ sends a state $s \in S_2$ to the constant function $z \mapsto k(s)$ in $C_b(S_2,[0,1])$, and the coalgebra condition amounts to compatibility of this embedding with expected future values under the dynamics of $M_2$.

Now consider any pair of distinct parallel morphisms $f, g : M_1 \rightrightarrows M_2$ in **MDP**. Since the underlying continuous maps $U(f), U(g) : S_1 \to S_2$ are distinct, there exists a point $s_1 \in S_1$ such that $f(s_1) \neq g(s_1)$. As $S_2$ is Polish (and therefore completely regular and Hausdorff), Urysohn's lemma provides a continuous function $k : S_2 \to [0,1]$ with $k(f(s_1)) = 1$ and $k(g(s_1)) = 0$. The choice of the MDP $C$ with state space $C_b(S_2,[0,1])$ and dynamics as above, together with the homomorphism $h_k : M_2 \to C$, then yields

$$h_k \circ f(s_1) \neq h_k \circ g(s_1),$$

since their images under $k$ differ. Hence $h_k \circ f \neq h_k \circ g$ as morphisms.

The collection of all such MDPs $C$ arising from Polish spaces $S_2$ in a fixed set of representatives (e.g. one for each isomorphism class of Polish spaces used as state spaces in **MDP**) forms a coseparating family. The technical work required to spell out the coalgebra structures on these $C$ and the induced homomorphisms $h_k$ is routine but lengthy and follows established patterns in coalgebra theory; for that reason, only the conceptual outline is recorded here. □

For the purpose of invoking the dual SAFT, it suffices that such a coseparating family exists. The specific choice or explicit description of a minimal family is immaterial for the rest of the paper; only its existence is used.

**4. Preservation of limits by $\mathrm{Q}_\varepsilon$.**

The final requirement of the dual SAFT is that the functor $\mathrm{Q}_\varepsilon : \mathbf{MDP} \to \mathbf{Spec}$ preserve the relevant limits (again, limits over at most countable diagrams suffice for all applications here).

Recall that for an MDP $M = (S, c)$, the functor $\mathrm{Q}_\varepsilon$ assigns the behavioral quotient

$$\mathrm{Q}_\varepsilon(M) \coloneqq (S/{\sim_\varepsilon}, d_{M,\varepsilon}),$$

where $\sim_\varepsilon$ identifies states at behavioral distance at most $\varepsilon$ according to the fixed-point pseudometric $d_M$ defined in the main text, and $d_{M,\varepsilon}$ is the induced metric (or pseudometric) on the quotient. Morphisms in **MDP** are mapped by $\mathrm{Q}_\varepsilon$ to the induced non-expansive maps between these quotients.

The key property needed is that behavioral pseudometrics behave well with respect to limits.

**Lemma A.7.** *Let $D : J \to$ **MDP** be a diagram indexed by a small countable category $J$, and let $(S_{\lim}, c_{\lim})$ be its limit in* **MDP**, *with projections $\pi_j : S_{\lim} \to S_j$. Denote by $d_{M_j}$ the behavioral pseudometric on $S_j$ associated with $D(j)$, and by $d_{M_{\lim}}$ the behavioral pseudometric on $S_{\lim}$. Then for all $s, t \in S_{\lim}$,*

$$d_{M_{\lim}}(s,t) = \sup_{j \in J} d_{M_j}\big(\pi_j(s), \pi_j(t)\big).$$

*Proof.* The behavioral pseudometric $d_M$ on an MDP $M = (S, c)$ is defined as the least fixed point of a monotone contraction operator (the Bellman operator)

$$\mathcal{K}_c : \mathsf{PM}(S) \to \mathsf{PM}(S),$$

where $\mathsf{PM}(S)$ denotes the space of bounded pseudometrics on $S$. In the present context, $\mathcal{K}_c$ is constructed as follows: for a pseudometric $d$ on $S$ and states $s, t \in S$,

$$\mathcal{K}_c(d)(s,t) \coloneqq \sup_{a \in A} \Big( |R(s,a) - R(t,a)| + \gamma\, W_1^{(d)}\big(P(s,a,\cdot), P(t,a,\cdot)\big) \Big),$$

where $P$ is the transition kernel, $R$ the reward, $\gamma$ the discount factor, and $W_1^{(d)}$ is the Wasserstein distance on the successor distributions defined using $d$ as the ground metric. The operator $\mathcal{K}_c$ is monotone and contracts the supremum norm with factor $\gamma < 1$, so it has a unique fixed point $d_M$.

For the limit coalgebra $(S_{\lim}, c_{\lim})$, the Bellman operator $\mathcal{K}_{c_{\lim}}$ is related to those of the components via the universal property of the limit and the preservation of limits by $\mathcal{F}_A$. Specifically, for any pseudometric $d$ on $S_{\lim}$, define for each $j \in J$ the induced pseudometric $d_j$ on $S_j$ by

$$d_j(x,y) \coloneqq \inf\{ d(s,t) \mid \pi_j(s) = x,\ \pi_j(t) = y \}.$$

Conversely, given a family $(d_j)_{j \in J}$ of pseudometrics on $S_j$ that is compatible with the projections (in the sense that for each morphism $\alpha : j \to k$ in $J$, the map $D(\alpha)$ is non-expansive from $(S_j, d_j)$ to $(S_k, d_k)$), one obtains a pseudometric on $S_{\lim}$ by

$$\bar{d}(s,t) \coloneqq \sup_{j \in J} d_j\big(\pi_j(s), \pi_j(t)\big).$$

The behaviour of the Bellman operator with respect to these constructions reflects the fact that the dynamics on the limit are compatible with those of the components via $\mathcal{F}_A$ and the projections $\pi_j$. Concretely, one has

$$\mathcal{K}_{c_{\lim}}(\bar{d}) = \overline{\sup_{j \in J} \mathcal{K}_{c_j}(d_j)},$$

where the bar denotes the reconstruction of a pseudometric on $S_{\lim}$ from the family $(\mathcal{K}_{c_j}(d_j))_{j \in J}$ via the supremum over $j$. Since $\mathcal{K}_{c_{\lim}}$ and each $\mathcal{K}_{c_j}$ are contractions with the same factor $\gamma$, the correspondence between families $(d_j)$ and pseudometrics on $S_{\lim}$ is compatible with taking least fixed points. Therefore the least fixed point $d_{M_{\lim}}$ satisfies

$$d_{M_{\lim}}(s,t) = \sup_{j \in J} d_{M_j}\big(\pi_j(s), \pi_j(t)\big)$$

for all $s, t \in S_{\lim}$. The details follow from standard fixed-point arguments in complete metric spaces and the uniqueness of fixed points of contractions. $\square$

With this lemma in hand, the preservation of limits by $\mathrm{Q}_\varepsilon$ can be verified.

**Lemma A.8.** *The functor* $\mathrm{Q}_\varepsilon : \mathbf{MDP} \to \mathbf{Spec}$ *preserves limits indexed by at most countable diagrams.*

*Proof.* Let $D : J \to \mathbf{MDP}$ be a diagram and $(S_{\lim}, c_{\lim})$ its limit in $\mathbf{MDP}$, with limit projections $\pi_j : S_{\lim} \to S_j$. Denote by $M_{\lim}$ this limit object and by $M_j$ the objects $D(j)$. Let

$$Q_{\lim} \coloneqq \mathrm{Q}_\varepsilon(M_{\lim}), \qquad Q_j \coloneqq \mathrm{Q}_\varepsilon(M_j).$$

The objects $Q_{\lim}$ and $Q_j$ are metric (or pseudometric) spaces obtained as quotients of the respective state spaces by the equivalence relation identifying points at behavioral distance at most $\varepsilon$.

By Lemma A.7, the behavioral pseudometric on $S_{\lim}$ satisfies

$$d_{M_{\lim}}(s,t) = \sup_{j \in J} d_{M_j}\big(\pi_j(s), \pi_j(t)\big).$$

In particular, if $d_{M_{\lim}}(s,t) \leq \varepsilon$, then $d_{M_j}(\pi_j(s), \pi_j(t)) \leq \varepsilon$ for all $j$. Thus each $\pi_j : S_{\lim} \to S_j$ is an $\varepsilon$-collapsing map: whenever two states are identified in $Q_{\lim}$, their images are identified in $Q_j$. By the universal property of the quotient maps

$$\psi_{\lim} : S_{\lim} \to Q_{\lim}, \qquad \psi_j : S_j \to Q_j,$$

there exists a unique map $\bar{\pi}_j : Q_{\lim} \to Q_j$ such that

$$\bar{\pi}_j \circ \psi_{\lim} = \psi_j \circ \pi_j.$$

The collection of maps $(\bar{\pi}_j)_{j \in J}$ is a cone from $Q_{\lim}$ to the diagram $j \mapsto Q_j$ in $\mathbf{Spec}$.

To see that this cone is limiting, let $(Z, d_Z)$ be any object of $\mathbf{Spec}$ equipped with a cone of non-expansive maps $(q_j : Z \to Q_j)_{j \in J}$ over the diagram $(Q_j)_{j \in J}$. For each $z \in Z$, consider the family $(q_j(z))_{j \in J}$ in $\prod_j Q_j$. Using the fact that the $Q_j$ arose as quotients of $S_j$ and that the maps $\pi_j$ are jointly faithful at the level of behavioral equivalence classes, it is possible to reconstruct a unique behavioral equivalence class $[s] \in Q_{\lim}$ for each $z$ that is compatible with the $(q_j(z))_{j \in J}$. Equipping $Z$ with the map $u : Z \to Q_{\lim}$ sending $z$ to this equivalence class yields a non-expansive map such that

$$\bar{\pi}_j \circ u = q_j \qquad \forall j \in J.$$

Uniqueness of $u$ follows from the fact that if two maps from $Z$ to $Q_{\lim}$ agree after composition with all $\bar{\pi}_j$, then their values must coincide because the projections faithfully reflect behavioral equivalence.

Thus $Q_{\lim}$ with the maps $\bar{\pi}_j$ is a limit of the diagram $j \mapsto Q_j$ in $\mathbf{Spec}$, and since this is precisely the image of $M_{\lim}$ under $\mathrm{Q}_\varepsilon$, it follows that $\mathrm{Q}_\varepsilon$ preserves these limits. $\square$

**Conclusion.**

Lemmas A.4, A.5, A.6, and A.8 establish that:

- $\mathbf{MDP}$ has the relevant limits (in particular all countable limits used in the paper) and is well-powered,
- $\mathbf{MDP}$ has a coseparating set of objects,
- $\mathrm{Q}_\varepsilon : \mathbf{MDP} \to \mathbf{Spec}$ preserves these limits.

These are exactly the hypotheses required by a suitable version of the dual Special Adjoint Functor Theorem. Therefore $\mathrm{Q}_\varepsilon$ admits a right adjoint $\mathrm{R}_\varepsilon : \mathbf{Spec} \to \mathbf{MDP}$. This completes the justification of Theorem 4.5. □

## Appendix B. Empirical Reproducibility Notes

For reproducibility of the experiments described in the main text, this appendix records additional technical details of the experimental setup, data generation, solver configuration, and algorithmic implementation. Unless stated otherwise, all default settings described below are used throughout.

**MDP Generation.** For the foundational verification sweep, random Markov decision processes were generated according to the following procedure.

- **State and action counts.** The numbers of states and actions, denoted $n$ and $m$ respectively, were drawn independently and uniformly from their prescribed ranges (as specified in the main text). This yields a broad variety of problem sizes while keeping computational demands manageable.
- **Reward distributions.** For each state–action pair $(s, a)$, the reward $R(s, a)$ was sampled independently from the uniform distribution $\mathcal{U}[-1, 1]$. This choice ensures bounded rewards and avoids any particular bias towards positive or negative reward structures.
- **Transition probabilities.** For each state–action pair $(s, a)$, an $n$-dimensional vector $(\tilde{p}_1, \ldots, \tilde{p}_n)$ was drawn from the Dirichlet distribution with all concentration parameters equal to 1, i.e. $\mathrm{Dirichlet}(\alpha_1, \ldots, \alpha_n)$ with $\alpha_i = 1$ for all $i$. The resulting vector is almost surely in the interior of the probability simplex
$$\Delta_n \coloneqq \left\{ (p_1, \ldots, p_n) \in [0,1]^n \;\middle|\; \sum_{i=1}^{n} p_i = 1 \right\},$$
and the $\mathrm{Dirichlet}(1, \ldots, 1)$ distribution is uniform over this simplex. This choice yields a diverse family of transition kernels without favouring any particular next-state structure.
- **Discount factor.** The discount factor $\gamma$ was sampled from a fixed finite set of values prescribed in the main text (for example, $\{0.7, 0.8, 0.9, 0.95\}$). Each value was chosen with equal probability. The requirement $\gamma \in (0, 1)$ ensures that the Bellman operators involved in value iteration and in the behavioral metric computation are contractions.

All random draws above were taken to be independent across different MDP instances and across different components (rewards and transition probabilities). This independence simplifies statistical reasoning about the ensemble of generated MDPs and ensures that the observed behaviour does not hinge on any particular dependence structure.

**Solver Configuration.** The computation of 1-Wasserstein distances between successor distributions, which is required at each iteration of the fixed-point algorithm for the behavioral metric, was formulated as a linear program (LP) and solved using standard numerical optimization software.

- **LP formulation.** For each pair of states $(s_1, s_2)$ and each action $a \in A$, the Wasserstein distance between the successor distributions $\mu = P(s_1, a, \cdot)$ and $\nu = P(s_2, a, \cdot)$ was expressed as

$$W_1^{(d)}(\mu,\nu) = \min_{\pi\in\Pi(\mu,\nu)} \sum_{i,j=1}^{n} d(i,j)\,\pi_{ij},$$

where $\pi = (\pi_{ij})$ is a coupling matrix with nonnegative entries satisfying the marginal constraints $\sum_j \pi_{ij} = \mu_i$ and $\sum_i \pi_{ij} = \nu_j$. These constraints and the nonnegativity constraints on $\pi_{ij}$ define a standard LP.

- **Numerical solver.** All LPs were solved using the `linprog` routine in `scipy.optimize` with the `highs` method enabled. This method combines primal and dual simplex algorithms with interior-point techniques and is widely used for medium- to large-scale LPs.
- **Tolerances and termination.** The default numerical tolerance of `linprog` was used, which is $10^{-9}$ for feasibility and optimality. This choice strikes a balance between numerical robustness and computational efficiency and is more stringent than what is strictly required for theoretical convergence of the behavioral metric computation.
- **Value iteration.** The optimal value function $V^*$ for each generated MDP was computed via classical value iteration. Starting from an initial function (for example, the zero function), the Bellman operator

$$(\mathcal{T}V)(s) \coloneqq \max_{a\in A}\left(R(s,a) + \gamma\sum_{s'} P(s,a,s')V(s')\right)$$

was applied repeatedly until successive iterates differed by less than a prescribed tolerance in the sup-norm. Specifically, iterations were terminated when

$$\max_{s\in S} |V_{k+1}(s) - V_k(s)| \le 10^{-9}.$$

This stopping rule ensures that the resulting value function is within $10^{-9}/(1-\gamma)$ of the exact fixed point, by the standard contraction argument.

**Adversarial Reward Search Setup.** In the adversarial experiments, the reward function was optimized in order to stress-test the behavioral metric and related quantities. The following setup was used.

- **Search domain.** The reward function $R$ was parameterized as an array in $\mathbb{R}^{n\times m}$, with entries $R(s,a)$ corresponding to the reward at state $s$ under action $a$. The search space for $R$ was the hypercube

$$[-10,10]^{n\times m},$$

meaning that each reward entry $R(s,a)$ was constrained to lie in $[-10,10]$. This range is substantially wider than the range $[-1,1]$ used for random MDP generation, thus allowing the adversarial search procedure to explore more extreme reward structures.
- **Optimization algorithm.** The adversarial search was carried out using the `differential_evolution` algorithm from `scipy.optimize`. This is a population-based global optimization method that does not rely on gradient information and is particularly suitable for nonconvex and nonsmooth objective functions, such as those involving fixed points and nested optimizations.
- **Iteration budget and parameters.** The algorithm was run for 1000 iterations with the default strategy parameters provided by the library. These defaults include the population size, mutation factor, crossover probability, and bounds handling. Using the default

configuration ensures that the setup is reproducible and can be easily replicated without manual tuning.

In all adversarial experiments, the objective function was evaluated deterministically for each candidate reward configuration, using the same fixed-point and LP machinery as in the random MDP experiments. This guarantees consistency between the two sets of experiments.

**Pseudo-code for the Behavioral Metric Computation.** The core algorithm for computing the behavioral pseudometric $d_M$ via fixed-point iteration is reproduced below for completeness. It is assumed that the state space $S$ is finite, with $|S| = n$, and that the action space $A$ is finite with $|A| = m$. The pseudocode describes the standard Picard iteration for a contraction mapping on the complete metric space of bounded pseudometrics on $S$.

**Algorithm 1:** Fixed-point iteration for the behavioral pseudometric $d_M$

**Input:** MDP $M = (S, A, P, R, \gamma)$ with finite $S$ and $A$, tolerance $\delta > 0$
**Output:** Behavioral pseudometric $d_M$ on $S$

**Initialization:** Set $d_{\text{curr}}[s,t] \leftarrow 0$ for all $s,t \in S$;
**repeat**
    $d_{\text{prev}} \leftarrow d_{\text{curr}}$;
    **for** *each pair of states* $(s_1, s_2) \in S \times S$ **do**
        $max\_dist \leftarrow 0$;
        **for** *each action* $a \in A$ **do**
            $r\_diff \leftarrow |R(s_1, a) - R(s_2, a)|$;
            $\mu \leftarrow P(s_1, a, \cdot)$;
            $\nu \leftarrow P(s_2, a, \cdot)$;
            $w\_dist \leftarrow \text{ComputeWasserstein}(d_{\text{prev}}, \mu, \nu)$ ; `// via LP solver with cost matrix induced by` $d_{\texttt{prev}}$
            $dist \leftarrow r\_diff + \gamma \cdot w\_dist$;
            $max\_dist \leftarrow \max(max\_dist, dist)$;
        **end**
        $d_{\text{curr}}[s_1, s_2] \leftarrow max\_dist$;
    **end**
    change $\leftarrow \max_{s,t \in S} |d_{\text{curr}}[s,t] - d_{\text{prev}}[s,t]|$;
**until** *change* $\leq \delta$;
**return** $d_{curr}$

The procedure `ComputeWasserstein` takes as input the current pseudometric $d_{\text{prev}}$ and the two discrete probability distributions $\mu$ and $\nu$ on $S$ (represented as vectors of length $n$) and returns the 1-Wasserstein distance between $\mu$ and $\nu$ with respect to the ground cost $d_{\text{prev}}$. This is computed by solving a linear program whose variables represent a coupling matrix $\pi$ and whose objective is the expected cost

$$\sum_{i,j=1}^{n} d_{\text{prev}}(s_i, s_j)\, \pi_{ij}$$

subject to the usual marginal constraints.

The outer loop implements the Picard iteration of the Bellman-like operator

$$\mathcal{K}_M(d)(s_1, s_2) := \sup_{a \in A} \Big( |R(s_1, a) - R(s_2, a)| + \gamma\, W_1^{(d)}\big(P(s_1, a, \cdot), P(s_2, a, \cdot)\big) \Big),$$

starting from the zero pseudometric. Because $\mathcal{K}_M$ is a contraction with factor $\gamma$, the sequence $(d_{\mathrm{curr}})$ converges in the sup-norm to the unique fixed point $d_M$, and the termination criterion change $\leq \delta$ yields an approximation whose sup-norm error is at most $\delta/(1-\gamma)$.

The details given in this section are sufficient to reconstruct and reproduce the experimental procedures described in the main text, provided that the same software environment (or an equivalent numerical linear programming solver) is available. Any deviations from these defaults in specific experiments are stated explicitly in the corresponding sections.