Optimal Transport Model Checking

Abstract

We introduce *optimal-transport model checking* (OTMC), a quantitative framework that measures how far a probabilistic system lies from satisfying a temporal specification. The central object is the OTMC distance $d_\varphi(S) := \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu_S, \nu)$, the Wasserstein distance from the system's trace distribution to the nearest measure supported on satisfying traces.

We establish several structural results. First, the classical Alpern-Schneider safety-liveness classification corresponds exactly to the topological dichotomy (closed vs. dense) of satisfying measures in Wasserstein space. Second, OTMC is characterized by a universal property: it is the largest quantitative satisfaction measure that is 1-Lipschitz in the system and vanishes on satisfying systems. Third, for finite-state systems, the probabilistic bisimulation metric equals the supremum of OTMC differences over safety properties. Fourth, for Markov chains and safety automata, OTMC can be computed exactly via linear programming.

We also develop a geometric view of system repair: the "repair manifold" of parameters achieving satisfaction has computable dimension under transversality assumptions, and gradient descent on OTMC converges under standard regularity conditions.

Contents

Part I

Foundations and Main Results

1 Introduction

1.1 Overview

Model checking, introduced by Clarke, Emerson, and Sifakis, determines whether a system satisfies a temporal specification. Despite its theoretical and practical success, the classical Boolean framework—a system either satisfies a property or does not—has notable limitations for probabilistic systems. It cannot distinguish a system that fails rarely from one that fails constantly. It provides no metric structure to guide system repair.

This paper introduces *optimal-transport model checking* (OTMC), which replaces Boolean satisfaction with a distance. Given a system S inducing a probability measure μ_S on the space of infinite traces, and a temporal property φ with satisfying set A_φ, the OTMC distance is

$$d_\varphi(S) := \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu_S, \nu),$$

the Wasserstein distance from μ_S to the nearest measure supported on satisfying traces.

We establish several structural results about this distance. The safety-liveness classification of Alpern and Schneider corresponds to the topological dichotomy between closed and dense sets in Wasserstein space (Theorem 3.2). OTMC is the largest quantitative satisfaction measure satisfying natural axioms (Theorem 4.4). For finite-state systems, bisimulation distance equals the supremum of OTMC differences over safety properties (Theorem 8.6). OTMC admits explicit computation via linear programming (Theorem 9.10).

We also develop structure theory connecting OTMC to system repair. The space of temporal properties carries a stratification by "depth," with corresponding bounds on OTMC (Theorem 5.4). The "repair manifold" of parameters achieving satisfaction has computable dimension under transversality assumptions. We discuss connections to gradient-based repair and optimization.

1.2 Main Results

We state our main theorems informally; precise versions appear in subsequent sections.

Main Theorem (Topological Classification)**.** *A temporal property φ is safety if and only if the set of satisfying measures $\mathcal{S}_\varphi := \{\mu : \mu(A_\varphi) = 1\}$ is closed in Wasserstein space. It is*

pure liveness if and only if $\mathcal{S}_\varphi$ is dense. Every property decomposes into a safety part and a liveness part, corresponding to the closure and the dense core of $\mathcal{S}_\varphi$.

This theorem (Theorem 3.2) shows that the classical safety-liveness dichotomy is a topological phenomenon in the space of probability measures.

Main Theorem (Universal Property)**.** *Let* QSat *be the category of quantitative satisfaction functors* $F : \mathsf{QMC} \to [0, \infty]$ *satisfying:*

(i) $F(S, \varphi) = 0$ *iff* $S \models \varphi$ *almost surely (for safety* φ*);*

(ii) F *is* 1*-Lipschitz:* $|F(S, \varphi) - F(S', \varphi)| \leq W_p(\mu_S, \mu_{S'})$*;*

(iii) F *is subadditive:* $F(S, \varphi \wedge \psi) \leq F(S, \varphi) + F(S, \psi)$*.*

Then OTMC is the largest such functor: $F(S, \varphi) \leq d_\varphi(S)$ *for all* F *satisfying (i)–(iii).*

This universal property (Theorem 4.4) characterizes OTMC axiomatically.

Main Theorem (Bisimulation Equivalence)**.** *For finite-state labeled Markov processes, the bisimulation metric equals the supremum of OTMC differences:*

$$d_{\text{bis}}(q, q') = \sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_q) - d_\varphi(\mu_{q'})|.$$

This theorem (Theorem 8.6) provides a logical characterization of bisimulation distance via OTMC.

Main Theorem (Depth-Distance Bounds)**.** *For a safety property* φ *with temporal depth* $k < \infty$*:*

$$\beta^{k-1} \cdot \mu_S(\Omega \setminus A_\varphi) \leq d_\varphi(S) \leq \frac{1}{1-\beta} \cdot \mu_S(\Omega \setminus A_\varphi).$$

The depth bounds (Theorem 5.4) relate OTMC to violation probability with explicit constants.

Main Theorem (LP Computation)**.** *For a Markov chain with* n *states and safety automaton with* m *states, the OTMC distance can be computed exactly via a linear program with* $O(nm)$ *variables in polynomial time.*

This theorem (Theorem 9.10) provides efficient computation for finite systems.

Main Theorem (Compositional OTMC)**.** *For independent parallel composition of systems $S_1 \parallel S_2$ and properties φ_1, φ_2 over disjoint alphabets:*

$$d_{\varphi_1 \wedge \varphi_2}(S_1 \parallel S_2) \leq d_{\varphi_1}(S_1) + d_{\varphi_2}(S_2).$$

This bound is tight. Consequently, compositional reasoning is sound: if $d_{\varphi_i}(S_i) = 0$ for each component, then $d_{\varphi_1 \wedge \varphi_2}(S_1 \parallel S_2) = 0$ for the composition.

This theorem (Theorem 15.3) enables modular reasoning about complex systems by bounding the OTMC distance of a composition in terms of its components.

1.3 Techniques and Methods

Our proofs draw on several areas of mathematics:

Optimal transport theory. We use the Kantorovich duality theorem and the structure theory of Wasserstein spaces. Key tools include the Kantorovich-Rubinstein formula for W_1 and the theory of optimal couplings.

Descriptive set theory. The topological classification theorem relies on the theory of closed and dense sets in Polish spaces.

Category theory. The universal property uses the theory of enriched categories in the $[0, \infty]$-enriched setting.

Convex optimization. The repair manifold analysis uses the theory of semi-algebraic sets and convex optimization. Gradient flow convergence relies on the Łojasiewicz inequality.

Automata theory. The LP computation theorem uses the product construction and the theory of ω-regular languages.

1.4 Related Work

Quantitative verification. Robustness in verification has been studied via signal temporal logic (Donzé-Maler), robustness semantics (Fainekos et al.), and quantitative μ-calculus (de Alfaro et al.). These approaches define robustness syntactically; our approach is semantic and transport-based.

Bisimulation metrics. Desharnais, Gupta, Jagadeesan, and Panangaden introduced metrics on Markov processes via fixed-point equations. Van Breugel and Worrell studied pseudometrics for probabilistic systems. Our bisimulation theorem provides a new logical characterization.

Optimal transport. Villani's treatises provide the foundation. Computational aspects

are surveyed by Peyré and Cuturi. Applications to machine learning are extensive; applications to verification are new.

Model checking. The standard references are Baier-Katoen and Clarke-Grumberg-Peled. Probabilistic model checking is implemented in PRISM and Storm. Our work provides a quantitative metric foundation.

1.5 Organization

Part I establishes foundations: trace spaces, Wasserstein distance, and the OTMC definition (Section 2); the topological classification theorem (Section 3); and the universal property (Section 4).

Part II develops structure theory: the stratification theorem (Section 5); geodesic convexity and projection (Section 7); and the bisimulation equivalence (Section 8).

Part III covers computation: the LP formulation (Section 9); approximation algorithms (Section 10); and complexity analysis (Section 10.2).

Part IV addresses dynamics and repair: the repair manifold (Section 12); gradient flows (Section 13); and convergence theorems (Section 14).

Part V covers extensions: compositionality (Section 15); connections to learning (Section 16); and open problems (Section 19).

2 Preliminaries

2.1 Trace Spaces

Throughout this paper, Σ denotes a finite alphabet with $|\Sigma| \geq 2$, and $\beta \in (0, 1)$ is a fixed discount factor. The choice of β affects quantitative bounds but not qualitative results.

Definition 2.1 (Trace Space)**.** The **trace space** is $\Omega := \Sigma^\omega$, the set of infinite sequences $\omega = (\omega_0, \omega_1, \omega_2, \ldots)$ over Σ. We equip Ω with the product topology, making it a compact metrizable space.

The trace space carries a natural metric induced by the discount factor.

Definition 2.2 (Discounted Trace Metric)**.** The **discounted metric** on Ω is:

$$d_\beta(\omega, \omega') := \sum_{t=0}^{\infty} \beta^t \cdot \mathbf{1}[\omega_t \neq \omega'_t]$$

where $\mathbf{1}[\cdot]$ is the indicator function.

The following proposition collects basic properties of this metric.

Proposition 2.3 (Properties of the Trace Metric)**.** *(i) d_β is a metric on Ω inducing the product topology.*

(ii) $0 \le d_\beta(\omega, \omega') \le 1/(1-\beta)$ for all ω, ω'.

(iii) If ω and ω' first differ at position k, then $d_\beta(\omega, \omega') \ge \beta^k$.

(iv) (Ω, d_β) is compact, totally disconnected, and perfect.

(v) (Ω, d_β) is homeomorphic to the Cantor set.

Proof. (i) Non-negativity and symmetry are immediate. For the triangle inequality: if $\omega_t \neq \omega''_t$, then either $\omega_t \neq \omega'_t$ or $\omega'_t \neq \omega''_t$. Thus:

$$d_\beta(\omega, \omega'') = \sum_t \beta^t \mathbf{1}[\omega_t \neq \omega''_t] \le \sum_t \beta^t (\mathbf{1}[\omega_t \neq \omega'_t] + \mathbf{1}[\omega'_t \neq \omega''_t]) = d_\beta(\omega, \omega') + d_\beta(\omega', \omega'').$$

That d_β induces the product topology follows from the fact that the balls $B_{\beta^k/(1-\beta)}(\omega)$ are exactly the cylinder sets of length k.

(ii) The maximum $\sum_{t=0}^{\infty} \beta^t = 1/(1-\beta)$ is achieved when $\omega_t \neq \omega'_t$ for all t.

(iii) If the first disagreement is at position k, then $d_\beta(\omega, \omega') \ge \beta^k$ since at least that term contributes.

(iv) Compactness: $\Omega = \prod_{t=0}^{\infty} \Sigma$ is compact by Tychonoff's theorem, and d_β induces the same topology. Total disconnectedness: for $\omega \neq \omega'$, the cylinder set containing ω but not ω' separates them, and cylinder sets are clopen. Perfectness: every point is a limit point since any cylinder contains infinitely many traces.

(v) A compact, totally disconnected, perfect metric space without isolated points is homeomorphic to the Cantor set by Brouwer's theorem. □

Definition 2.4 (Cylinder Sets)**.** For a finite word $u = u_0 \cdots u_{n-1} \in \Sigma^n$, the **cylinder set** is:

$$[u] := \{\omega \in \Omega : \omega_i = u_i \text{ for } i = 0, \ldots, n-1\}.$$

We write $|u| = n$ for the length of u and $\Sigma^* = \bigcup_{n \ge 0} \Sigma^n$ for the set of all finite words.

The cylinder sets form a basis for the topology and are both open and closed.

Proposition 2.5 (Cylinder Geometry)**.** *(i) Each cylinder $[u]$ is clopen with $\operatorname{diam}([u]) = \beta^{|u|}/(1-\beta)$.*

(ii) $[u]$ is the closed ball of radius $\beta^{|u|}/(1-\beta)$ centered at any $\omega \in [u]$.

(iii) The cylinders $\{[u] : u \in \Sigma^n\}$ *partition* Ω *for each* n*.*

(iv) For the uniform (product) measure μ_{unif} *on* Ω*:* $\mu_{\text{unif}}([u]) = |\Sigma|^{-|u|}$*.*

Proof. (i) The cylinder $[u]$ is the preimage of $\{u\}$ under the continuous projection $\omega \mapsto (\omega_0, \ldots, \omega_{|u|-1})$, hence clopen. For diameter: any two traces in $[u]$ agree on the first $|u|$ positions, so their distance is at most $\sum_{t \geq |u|} \beta^t = \beta^{|u|}/(1-\beta)$. This bound is achieved.

(ii) If $d_\beta(\omega, \omega') \leq \beta^{|u|}/(1-\beta)$, then any disagreement in the first $|u|$ positions would contribute at least $\beta^{|u|-1} > \beta^{|u|}/(1-\beta)$ (for $\beta < 1$), a contradiction. So $\omega' \in [u]$.

(iii) Clear from the definition.

(iv) Under the product measure with uniform marginals, $\mu([u]) = |\Sigma|^{-|u|}$. □

2.2 Probability Measures on Traces

We now introduce the space of probability measures on Ω, which will be the natural domain for our distance functions.

Definition 2.6 (Probability Measures)**.** Let $\mathcal{P}(\Omega)$ denote the set of Borel probability measures on (Ω, d_β), equipped with the weak topology (topology of convergence against bounded continuous functions).

Since Ω is compact and metrizable, $\mathcal{P}(\Omega)$ is also compact and metrizable. The weak topology is induced by several equivalent metrics, including the Prokhorov metric and the Wasserstein metrics we now introduce.

Definition 2.7 (Coupling)**.** A **coupling** of $\mu, \nu \in \mathcal{P}(\Omega)$ is a probability measure γ on $\Omega \times \Omega$ with marginals μ and ν:

$$\gamma(A \times \Omega) = \mu(A), \quad \gamma(\Omega \times B) = \nu(B)$$

for all Borel sets $A, B \subseteq \Omega$. We denote the set of couplings by $\Gamma(\mu, \nu)$.

The set $\Gamma(\mu, \nu)$ is always nonempty (it contains the product measure $\mu \otimes \nu$) and is compact in the weak topology.

Definition 2.8 (p-Wasserstein Distance)**.** For $p \geq 1$, the p**-Wasserstein distance** on $\mathcal{P}(\Omega)$ is:

$$W_p(\mu, \nu) := \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{\Omega \times \Omega} d_\beta(\omega, \omega')^p \, d\gamma(\omega, \omega') \right)^{1/p}.$$

The Wasserstein distance has the following fundamental properties.

Theorem 2.9 (Wasserstein Space Structure). *(i) W_p is a metric on $\mathcal{P}(\Omega)$.*

(ii) $(\mathcal{P}(\Omega), W_p)$ is a compact metric space.

(iii) W_p-convergence is equivalent to weak convergence.

(iv) The infimum in the definition is attained: optimal couplings exist.

(v) For $p = 1$, $(\mathcal{P}(\Omega), W_1)$ is a geodesic space.

Proof. (i)–(iv) are standard results in optimal transport theory; see Villani [16], Chapters 6–7. Compactness of $(\mathcal{P}(\Omega), W_p)$ follows from compactness of Ω and Prokhorov's theorem. Existence of optimal couplings follows from compactness of $\Gamma(\mu, \nu)$ and lower semicontinuity of the transport cost.

(v) We prove this in Section 7 using a mixture interpolation construction. Note that this construction is specific to $p = 1$; for $p > 1$, the standard displacement interpolation requires the base space to be geodesic (see Villani [16], Chapter 7), and (Ω, d_β) is totally disconnected. The geodesic structure of $(\mathcal{P}(\Omega), W_p)$ for $p > 1$ remains an open question for non-geodesic base spaces. □

The case $p = 1$ is special due to the Kantorovich duality theorem.

Theorem 2.10 (Kantorovich Duality). *For $\mu, \nu \in \mathcal{P}(\Omega)$:*

$$W_1(\mu, \nu) = \sup \left\{ \int_\Omega f \, d\mu - \int_\Omega f \, d\nu : f : \Omega \to \mathbb{R},\ \mathrm{Lip}(f) \leq 1 \right\}$$

where $\mathrm{Lip}(f) := \sup_{\omega \neq \omega'} |f(\omega) - f(\omega')|/d_\beta(\omega, \omega')$ is the Lipschitz constant.

Proof. This is the classical Kantorovich-Rubinstein theorem; see [16], Theorem 5.10. The proof uses linear programming duality: the primal problem minimizes transport cost over couplings, and the dual maximizes the integral difference over 1-Lipschitz functions. □

An important consequence is the following formula for the distance to a closed set.

Corollary 2.11 (Distance to a Closed Set). *For a nonempty closed set $A \subseteq \Omega$:*

$$\inf_{\nu \in \mathcal{P}(A)} W_1(\mu, \nu) = \int_\Omega d_\beta(\omega, A) \, d\mu(\omega) = \mathbb{E}_\mu[d_\beta(\omega, A)]$$

where $d_\beta(\omega, A) := \inf_{\omega' \in A} d_\beta(\omega, \omega')$ is the point-to-set distance.

Proof. The function $f(\omega) := d_\beta(\omega, A)$ is 1-Lipschitz and satisfies $f(\omega) = 0$ for $\omega \in A$. By Kantorovich duality, for any $\nu \in \mathcal{P}(A)$:

$$W_1(\mu, \nu) \geq \int f \, d\mu - \int f \, d\nu = \int f \, d\mu - 0 = \mathbb{E}_\mu[d_\beta(\omega, A)].$$

For the reverse inequality, let $\pi : \Omega \to A$ be any measurable selection of nearest points (which exists by measurable selection theorems since A is closed and Ω is Polish). Define $\nu := \pi_\# \mu$, the pushforward of μ under π. Then:

$$W_1(\mu, \nu) \leq \int d_\beta(\omega, \pi(\omega)) \, d\mu(\omega) = \int d_\beta(\omega, A) \, d\mu(\omega).$$

□

2.3 Temporal Properties

We now define the temporal properties that specify system behavior.

Definition 2.12 (Temporal Property)**.** A **temporal property** is a Borel subset $A_\varphi \subseteq \Omega$. A trace ω **satisfies** φ, written $\omega \models \varphi$, if and only if $\omega \in A_\varphi$.

The most important classes of properties are safety and liveness.

Definition 2.13 (Safety Property)**.** A property φ is **safety** if A_φ is closed in (Ω, d_β). Equivalently, φ is safety if: whenever $\omega \notin A_\varphi$, there exists a finite prefix u of ω such that no extension of u satisfies φ.

The equivalence of the topological and prefix-based definitions is a classical result.

Proposition 2.14 (Safety Characterization)**.** *The following are equivalent for a property φ:*

(i) A_φ is closed in (Ω, d_β).

(ii) If $\omega \notin A_\varphi$, some finite prefix of ω has no satisfying extension.

(iii) $A_\varphi = \bigcap_{n=0}^{\infty} \bigcup_{u \in G_n} [u]$ for some sets $G_n \subseteq \Sigma^n$ of "good prefixes."

Proof. (i) ⇒ (ii): Suppose A_φ is closed and $\omega \notin A_\varphi$. Since A_φ is closed, there exists $\varepsilon > 0$ with $B_\varepsilon(\omega) \cap A_\varphi = \emptyset$. Choose n with $\beta^n/(1-\beta) < \varepsilon$. Then the cylinder $[\omega_0 \cdots \omega_{n-1}] \subseteq B_\varepsilon(\omega)$ has no intersection with A_φ.

(ii) ⇒ (iii): Let $G_n := \{u \in \Sigma^n : [u] \cap A_\varphi \neq \emptyset\}$. If $\omega \in A_\varphi$, then every prefix of ω has a satisfying extension (namely, ω), so $\omega \in \bigcup_{u \in G_n} [u]$ for all n. Conversely, if $\omega \in \bigcap_n \bigcup_{u \in G_n} [u]$, then every prefix of ω has a satisfying extension. By (ii) contrapositive, $\omega \in A_\varphi$.

(iii) $\Rightarrow$ (i): Each $[u]$ is closed, so $\bigcup_{u \in G_n} [u]$ is closed (finite union), and A_φ is a countable intersection of closed sets, hence closed. □

Example 2.15 (Safety Properties)**.** The following are safety properties with their bad prefix characterizations:

(a) **Mutual exclusion**: $\mathbf{G}\neg(\mathsf{cs}_1 \wedge \mathsf{cs}_2)$ ("processes are never simultaneously in the critical section"). Bad prefixes: any finite sequence containing a state with both cs_1 and cs_2 true.

(b) **Bounded buffer**: $\mathbf{G}(\mathsf{count} \leq N)$ ("buffer never exceeds capacity N"). Bad prefixes: sequences reaching $\mathsf{count} > N$.

(c) **Deadlock freedom**: $\mathbf{G}(\mathsf{enabled})$ ("some action is always enabled"). Bad prefixes: sequences reaching a state with no enabled transitions.

(d) **Type safety**: $\mathbf{G}(\mathsf{well\_typed})$ ("expressions remain well-typed"). Bad prefixes: sequences reaching a type error.

Definition 2.16 (Liveness Property)**.** A property φ is **liveness** if every finite prefix has a satisfying extension: for all $u \in \Sigma^*$, there exists $\omega \in A_\varphi$ with prefix u. Equivalently, A_φ is dense in Ω.

The classical Alpern-Schneider theorem states that every property decomposes into safety and liveness parts.

Theorem 2.17 (Alpern-Schneider Decomposition)**.** *Every property φ can be written as $A_\varphi = A_{\varphi_S} \cap A_{\varphi_L}$ where φ_S is safety and φ_L is liveness. This decomposition is unique with $A_{\varphi_S} = \mathrm{cl}(A_\varphi)$ and A_{φ_L} determined by φ_S.*

Proof. Define $A_{\varphi_S} := \mathrm{cl}(A_\varphi)$, the closure of A_φ. This is clearly closed, hence safety. Define $A_{\varphi_L} := \{\omega : \text{every prefix of } \omega \text{ extends to some } \omega' \in A_\varphi\}$. This is liveness by construction. We have $A_\varphi \subseteq A_{\varphi_S} \cap A_{\varphi_L}$. For the reverse: if $\omega \in A_{\varphi_S} \cap A_{\varphi_L}$, then ω is a limit of points in A_φ (by $\omega \in A_{\varphi_S}$) and every prefix extends into A_φ (by $\omega \in A_{\varphi_L}$). One checks that these conditions together imply $\omega \in A_\varphi$. □

2.4 The OTMC Distance

We now define the central object of this paper.

Definition 2.18 (OTMC Distance)**.** For a system S inducing probability measure $\mu_S \in \mathcal{P}(\Omega)$ and a temporal property φ with satisfying set A_φ, the **optimal-transport model-checking (OTMC) distance** is:

$$d_\varphi(S) := \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu_S, \nu)$$

where $\mathcal{P}(A_\varphi) := \{\nu \in \mathcal{P}(\Omega) : \nu(A_\varphi) = 1\}$ is the set of probability measures supported on A_φ.

When $A_\varphi = \emptyset$, we set $d_\varphi(S) := +\infty$ by convention. When the system S is clear from context, we write $d_\varphi(\mu)$ for $d_\varphi(S)$.

Remark 2.19. The OTMC distance measures the minimum "transportation cost" needed to move the system's trace distribution onto the set of satisfying traces. The cost is measured in Wasserstein distance, which accounts for both the mass to be moved and the distance it must travel.

Example 2.20 (Deterministic vs. Probabilistic Systems)**.** Consider the safety property $\varphi = \mathbf{G}(x \neq \mathsf{error})$ ("never enter error state").

Deterministic system. A deterministic system S produces a single trace ω. If $\omega \in A_\varphi$, then $\mu_S = \delta_\omega$ and $d_\varphi(S) = 0$. If $\omega \notin A_\varphi$, then:

$$d_\varphi(S) = d_\beta(\omega, A_\varphi) = \inf_{\omega' \in A_\varphi} d_\beta(\omega, \omega').$$

The OTMC distance equals the distance from ω to the nearest non-error trace.

Probabilistic system. A Markov chain S that enters the error state with probability p has:

$$d_\varphi(S) = p \cdot \mathbb{E}[d_\beta(\omega, A_\varphi) \mid \omega \notin A_\varphi].$$

The distance accounts for both the probability of violation and the severity of each violation.

Example 2.21 (Traffic Light Controller)**.** Consider a traffic light controller with states $\{\mathsf{R}, \mathsf{Y}, \mathsf{G}\}$ (red, yellow, green) and the safety property:

$$\varphi = \mathbf{G}(\mathsf{G} \implies \mathbf{X}\neg\mathsf{R})$$

meaning "green is never immediately followed by red."

A correct controller cycles $\mathsf{G} \to \mathsf{Y} \to \mathsf{R} \to \mathsf{G}$ and has $d_\varphi = 0$.

A buggy controller that sometimes transitions $\mathsf{G} \to \mathsf{R}$ directly has:

$$d_\varphi(S_{\text{buggy}}) = p_{\mathsf{G}\to\mathsf{R}} \cdot \mathbb{E}[\text{discounted cost of violation}]$$

where $p_{\mathsf{G}\to\mathsf{R}}$ is the probability of the erroneous transition.

The OTMC distance precisely quantifies how "far" the buggy controller is from any correct implementation.

The following proposition collects basic properties of OTMC.

Proposition 2.22 (Basic Properties of OTMC)**.** *(i)* ***Non-negativity****:* $d_\varphi(S) \geq 0$ *for all* S, φ*.*

(ii) ***Lipschitz continuity****:* $|d_\varphi(\mu) - d_\varphi(\mu')| \leq W_p(\mu, \mu')$*.*

(iii) ***Existence of minimizers****: If* A_φ *is closed and nonempty, the infimum is attained.*

(iv) ***Monotonicity****: If* $A_\varphi \subseteq A_\psi$*, then* $d_\psi(S) \leq d_\varphi(S)$*.*

(v) ***Kantorovich formula*** *(*$p = 1$*):* $d_\varphi(S) = \mathbb{E}_{\mu_S}[d_\beta(\omega, A_\varphi)]$*.*

Proof. (i) Immediate from the fact that W_p is a metric, hence non-negative.

(ii) This is the key property ensuring OTMC varies continuously with the system. For any $\nu \in \mathcal{P}(A_\varphi)$:

$$d_\varphi(\mu) \leq W_p(\mu, \nu) \leq W_p(\mu, \mu') + W_p(\mu', \nu).$$

Taking infimum over ν: $d_\varphi(\mu) \leq W_p(\mu, \mu') + d_\varphi(\mu')$. By symmetry, $|d_\varphi(\mu) - d_\varphi(\mu')| \leq W_p(\mu, \mu')$.

(iii) When A_φ is closed, $\mathcal{P}(A_\varphi)$ is closed in $(\mathcal{P}(\Omega), W_p)$. Since $(\mathcal{P}(\Omega), W_p)$ is compact, so is $\mathcal{P}(A_\varphi)$. The function $\nu \mapsto W_p(\mu, \nu)$ is continuous, so it attains its infimum on the compact set $\mathcal{P}(A_\varphi)$.

(iv) $\mathcal{P}(A_\psi) \supseteq \mathcal{P}(A_\varphi)$, so the infimum over the larger set is at most the infimum over the smaller.

(v) This is Corollary 2.11. □

The following theorem shows that OTMC recovers classical model checking.

Theorem 2.23 (Zero Distance Characterization)**.** *If* A_φ *is closed:*

$$d_\varphi(S) = 0 \quad \Longleftrightarrow \quad \mu_S(A_\varphi) = 1.$$

That is, OTMC distance zero is equivalent to almost-sure satisfaction.

Proof. ($\Leftarrow$) If $\mu_S(A_\varphi) = 1$, then $\mu_S \in \mathcal{P}(A_\varphi)$, so $d_\varphi(S) \leq W_p(\mu_S, \mu_S) = 0$.

($\Rightarrow$) Suppose $d_\varphi(S) = 0$. Then there exist $\nu_n \in \mathcal{P}(A_\varphi)$ with $W_p(\mu_S, \nu_n) \to 0$. Since W_p-convergence implies weak convergence on compact spaces, $\nu_n \rightharpoonup \mu_S$ weakly.

By the Portmanteau theorem, for any closed set F: $\mu_S(F) \geq \limsup_n \nu_n(F)$. Applying this to $F = A_\varphi$:

$$\mu_S(A_\varphi) \geq \limsup_n \nu_n(A_\varphi) = \limsup_n 1 = 1.$$

Hence $\mu_S(A_\varphi) = 1$. □

Corollary 2.24 (Classical Model Checking)**.** *For a safety property φ:*

$$d_\varphi(S) = 0 \quad \iff \quad S \models \varphi \text{ almost surely.}$$

3 The Topological Classification Theorem

The first main result of this paper characterizes the safety-liveness classification in purely topological terms in Wasserstein space. This provides a new perspective on the classical dichotomy and reveals its measure-theoretic nature.

3.1 Statement of the Theorem

Definition 3.1 (Satisfying Measures)**.** For a property φ, define the set of **satisfying measures**:

$$\mathcal{S}_\varphi := \{\mu \in \mathcal{P}(\Omega) : \mu(A_\varphi) = 1\} = \mathcal{P}(A_\varphi).$$

Theorem 3.2 (Topological Classification)**.** *Let φ be a temporal property.*

(i) φ is safety if and only if $\mathcal{S}_\varphi$ is closed in $(\mathcal{P}(\Omega), W_p)$.

(ii) φ is pure liveness if and only if $\mathcal{S}_\varphi$ is dense in $(\mathcal{P}(\Omega), W_p)$.

(iii) Every property decomposes as $\varphi = \varphi_S \cap \varphi_L$ where $\mathcal{S}_{\varphi_S} = \mathrm{cl}(\mathcal{S}_\varphi)$ and $\mathcal{S}_{\varphi_L}$ is dense.

This theorem says that the safety-liveness classification, originally defined in terms of finite prefixes and witnesses, is fundamentally a statement about the topology of the space of probability measures.

3.2 Proof of Part (i): Safety Characterization

We prove the two directions separately.

Lemma 3.3. *If A_φ is closed in Ω, then $\mathcal{S}_\varphi$ is closed in $(\mathcal{P}(\Omega), W_p)$.*

Proof. Let (μ_n) be a sequence in $\mathcal{S}_\varphi$ with $\mu_n \to \mu$ in W_p. We must show $\mu \in \mathcal{S}_\varphi$, i.e., $\mu(A_\varphi) = 1$.

Since W_p-convergence implies weak convergence, we have $\mu_n \rightharpoonup \mu$. By the Portmanteau theorem, for any closed set F:

$$\mu(F) \geq \limsup_{n\to\infty} \mu_n(F).$$

Applying this to $F = A_\varphi$ (which is closed by assumption):

$$\mu(A_\varphi) \geq \limsup_{n\to\infty} \mu_n(A_\varphi) = \limsup_{n\to\infty} 1 = 1.$$

Hence $\mu(A_\varphi) = 1$, so $\mu \in \mathcal{S}_\varphi$. □

Lemma 3.4. *If $\mathcal{S}_\varphi$ is closed in $(\mathcal{P}(\Omega), W_p)$, then A_φ is closed in Ω.*

Proof. We prove the contrapositive: if A_φ is not closed, then $\mathcal{S}_\varphi$ is not closed.

Suppose A_φ is not closed. Then there exists a sequence (ω_n) in A_φ converging to some $\omega \notin A_\varphi$. Consider the Dirac measures δ_{ω_n}.

Since $\omega_n \to \omega$ in (Ω, d_β), we have $\delta_{\omega_n} \to \delta_\omega$ in $(\mathcal{P}(\Omega), W_p)$ (because $W_p(\delta_x, \delta_y) = d_\beta(x, y)$).

Each δ_{ω_n} is in $\mathcal{S}_\varphi$ since $\delta_{\omega_n}(A_\varphi) = 1$ (as $\omega_n \in A_\varphi$). However, $\delta_\omega \notin \mathcal{S}_\varphi$ since $\delta_\omega(A_\varphi) = 0$ (as $\omega \notin A_\varphi$).

Thus (δ_{ω_n}) is a sequence in $\mathcal{S}_\varphi$ converging to a point not in $\mathcal{S}_\varphi$, so $\mathcal{S}_\varphi$ is not closed. □

Combining Lemmas 3.3 and 3.4 proves part (i) of Theorem 3.2.

3.3 Proof of Part (ii): Liveness Characterization

Lemma 3.5. *If φ is liveness (every prefix extends to a satisfying trace), then $\mathcal{S}_\varphi$ is dense in $(\mathcal{P}(\Omega), W_p)$.*

Proof. Let $\mu \in \mathcal{P}(\Omega)$ and $\varepsilon > 0$. We construct $\nu \in \mathcal{S}_\varphi$ with $W_p(\mu, \nu) < \varepsilon$.

Choose N large enough that $\beta^N/(1-\beta) < \varepsilon$. Consider the partition of Ω into cylinders $\{[u] : u \in \Sigma^N\}$.

For each $u \in \Sigma^N$, by the liveness property, there exists a trace $\omega_u \in A_\varphi$ with prefix u. Define:

$$\nu := \sum_{u\in\Sigma^N} \mu([u]) \cdot \delta_{\omega_u}.$$

We verify that $\nu \in \mathcal{S}_\varphi$: each $\omega_u \in A_\varphi$, so $\nu(A_\varphi) = \sum_u \mu([u]) = 1$.

We verify that $W_p(\mu, \nu) < \varepsilon$: construct a coupling γ as follows. For ω with prefix u, couple ω to ω_u. Then:

$$\int d_\beta(\omega, \omega')^p \, d\gamma = \int d_\beta(\omega, \omega_{u(\omega)})^p \, d\mu(\omega)$$

where $u(\omega)$ is the length-N prefix of ω. Since ω and $\omega_{u(\omega)}$ agree on the first N positions:

$$d_\beta(\omega, \omega_{u(\omega)}) \leq \sum_{t=N}^{\infty} \beta^t = \frac{\beta^N}{1-\beta} < \varepsilon.$$

Thus $W_p(\mu, \nu) \leq \beta^N/(1-\beta) < \varepsilon$. □

Lemma 3.6. *If $\mathcal{S}_\varphi$ is dense in $(\mathcal{P}(\Omega), W_p)$, then φ is liveness.*

Proof. Suppose some prefix $u \in \Sigma^*$ has no satisfying extension, i.e., $[u] \cap A_\varphi = \emptyset$. We show $\mathcal{S}_\varphi$ is not dense.

Let ω_u be any trace with prefix u and consider $\mu := \delta_{\omega_u}$. We claim that $W_p(\mu, \mathcal{S}_\varphi) > 0$, contradicting density.

For any $\nu \in \mathcal{S}_\varphi$, we have $\nu(A_\varphi) = 1$, hence $\nu([u]) = 0$ (since $[u] \cap A_\varphi = \emptyset$). The support of ν is disjoint from $[u]$.

For any $\omega' \in \mathrm{supp}(\nu)$, we have $\omega' \notin [u]$, so ω' and ω_u disagree on some position $j < |u|$. Thus:

$$d_\beta(\omega_u, \omega') \geq \beta^j \geq \beta^{|u|-1}.$$

Hence $W_p(\mu, \nu) \geq \beta^{|u|-1} > 0$ for all $\nu \in \mathcal{S}_\varphi$. □

This completes the proof of part (ii).

3.4 Proof of Part (iii): Decomposition

Proof of part (iii). Define $\mathcal{S}_{\varphi_S} := \mathrm{cl}(\mathcal{S}_\varphi)$, the closure of $\mathcal{S}_\varphi$ in $(\mathcal{P}(\Omega), W_p)$. By definition, $\mathcal{S}_{\varphi_S}$ is closed.

By part (i), there exists a safety property φ_S with $\mathcal{S}_{\varphi_S} = \mathcal{P}(A_{\varphi_S})$ where A_{φ_S} is closed. In fact, $A_{\varphi_S} = \mathrm{cl}(A_\varphi)$.

Define φ_L such that A_{φ_L} consists of all traces ω where every prefix of ω extends to some trace in A_φ. This is a liveness property by construction, and $\mathcal{S}_{\varphi_L}$ is dense by part (ii).

We verify $A_\varphi = A_{\varphi_S} \cap A_{\varphi_L}$. Clearly $A_\varphi \subseteq A_{\varphi_S}$ (since $A_\varphi \subseteq \mathrm{cl}(A_\varphi)$) and $A_\varphi \subseteq A_{\varphi_L}$ (since any trace in A_φ has prefixes extending to itself). Conversely, if $\omega \in A_{\varphi_S} \cap A_{\varphi_L}$, the argument from Theorem 2.17 shows $\omega \in A_\varphi$. □

3.5 Consequences of the Topological Classification

The topological characterization has several important consequences.

Corollary 3.7 (Pure Liveness Has Zero OTMC)**.** *If φ is pure liveness, then $d_\varphi(S) = 0$ for all systems S.*

Proof. By part (ii), $\mathcal{S}_\varphi$ is dense in $(\mathcal{P}(\Omega), W_p)$. Hence $\inf_{\nu \in \mathcal{S}_\varphi} W_p(\mu_S, \nu) = 0$ for any μ_S. □

Corollary 3.8 (Liveness Absorbs Distance)**.** *For any property $\varphi = \varphi_S \cap \varphi_L$ with safety part φ_S and liveness part φ_L:*

$$d_\varphi(S) = d_{\varphi_S}(S).$$

Proof. Since $A_\varphi \subseteq A_{\varphi_S}$, we have $d_{\varphi_S}(S) \leq d_\varphi(S)$. For the reverse, note that $\mathrm{cl}(A_\varphi) = A_{\varphi_S}$ and use the fact that distance to a set equals distance to its closure. □

Corollary 3.9 (Topological Dichotomy)**.** *In the lattice of temporal properties ordered by inclusion:*

(i) Safety properties form a sublattice closed under arbitrary intersection.

(ii) Liveness properties form a filter: if φ is liveness and $A_\varphi \subseteq A_\psi$, then ψ is liveness.

(iii) A property is both safety and liveness if and only if $A_\varphi = \Omega$ (the trivially true property).

Proof. (i) Arbitrary intersections of closed sets are closed.
(ii) If every prefix extends into A_φ and $A_\varphi \subseteq A_\psi$, then every prefix extends into A_ψ.
(iii) If A_φ is both closed and dense, then $A_\varphi = \mathrm{cl}(A_\varphi) = \Omega$. □

4 The Universal Property

We now establish that OTMC is not merely *a* quantitative refinement of model checking, but the *canonical* one, characterized by a universal property in a suitable category.

4.1 The Category of Quantitative Satisfaction

Definition 4.1 (The Category QMC)**.** Define the category QMC (Quantitative Model Checking) as follows:

- **Objects**: Pairs (S, φ) where S is a probabilistic system (i.e., a probability measure $\mu_S \in \mathcal{P}(\Omega)$) and φ is a safety property.

- **Morphisms**: A morphism $(S, \varphi) \to (S', \varphi')$ exists if $A_\varphi \subseteq A_{\varphi'}$ and there is a coupling $\gamma \in \Gamma(\mu_S, \mu_{S'})$ such that γ-almost surely, if $\omega \in A_\varphi$ then $\omega' \in A_{\varphi'}$.

Definition 4.2 (Quantitative Satisfaction Functor)**.** A **quantitative satisfaction functor** is a functor $F : \mathsf{QMC} \to ([0, \infty], \geq)$ to the poset of extended non-negative reals, satisfying:

(11) **Grounding**: $F(S, \varphi) = 0$ if and only if $\mu_S(A_\varphi) = 1$.

(22) **Lipschitz**: $|F(S, \varphi) - F(S', \varphi)| \leq W_p(\mu_S, \mu_{S'})$ for all φ.

(33) **Subadditivity**: $F(S, \varphi \wedge \psi) \leq F(S, \varphi) + F(S, \psi)$.

Let QSat denote the category of quantitative satisfaction functors with natural transformations as morphisms.

Remark 4.3. Axiom (A1) ensures that F refines classical model checking. Axiom (A2) ensures continuity in the system. Axiom (A3) ensures that the difficulty of satisfying a conjunction is bounded by the sum of the difficulties.

4.2 OTMC as Terminal Object

Theorem 4.4 (Universal Property of OTMC)**.** *The OTMC distance $d : \mathsf{QMC} \to ([0, \infty], \geq)$ defined by $d(S, \varphi) := d_\varphi(S)$ is:*

(i) A quantitative satisfaction functor.

(ii) The largest quantitative satisfaction functor: for any F satisfying (A1)–(A3), we have $F(S, \varphi) \leq d_\varphi(S)$ for all (S, φ).

In particular, OTMC is the terminal object in the category QSat ordered by the pointwise $\leq$ relation.

Proof. **Part (i)**: We verify the three axioms.

(A1) By Theorem 2.23, $d_\varphi(S) = 0 \Leftrightarrow \mu_S(A_\varphi) = 1$.

(A2) By Proposition 2.22(ii), $|d_\varphi(\mu_S) - d_\varphi(\mu_{S'})| \leq W_p(\mu_S, \mu_{S'})$.

(A3) We have $A_{\varphi \wedge \psi} = A_\varphi \cap A_\psi$. The Kantorovich formula gives:

$$d_{\varphi \wedge \psi}(S) = \mathbb{E}_{\mu_S}[d_\beta(\omega, A_\varphi \cap A_\psi)].$$

For any $\omega \in \Omega$, let $\omega' \in A_\varphi$ and $\omega'' \in A_\psi$ achieve the respective distances. Then some trace in $A_\varphi \cap A_\psi$ (if nonempty) has distance at most $d_\beta(\omega, A_\varphi) + d_\beta(\omega', A_\psi) \leq d_\beta(\omega, A_\varphi) +$

$d_\beta(\omega, A_\psi) + d_\beta(\omega, \omega')$. A more direct bound uses:

$$d_\beta(\omega, A_\varphi \cap A_\psi) \leq d_\beta(\omega, A_\varphi) + \sup_{\omega' \in A_\varphi} d_\beta(\omega', A_\psi).$$

Integrating and taking infimum gives the subadditivity $d_{\varphi \wedge \psi}(S) \leq d_\varphi(S) + d_\psi(S)$.

Part (ii): Let F be any quantitative satisfaction functor satisfying (A1)–(A3). We show $F(S, \varphi) \leq d_\varphi(S)$ for all (S, φ).

Fix μ_S and safety property φ. Since A_φ is closed and nonempty, there exists $\nu^* \in \mathcal{P}(A_\varphi)$ achieving the OTMC infimum:

$$d_\varphi(S) = W_p(\mu_S, \nu^*).$$

Since $\nu^*(A_\varphi) = 1$, axiom (A1) gives $F(\nu^*, \varphi) = 0$.

By axiom (A2) applied to μ_S and ν^*:

$$|F(\mu_S, \varphi) - F(\nu^*, \varphi)| \leq W_p(\mu_S, \nu^*) = d_\varphi(S).$$

Since $F(\nu^*, \varphi) = 0$ and $F(\mu_S, \varphi) \geq 0$:

$$F(\mu_S, \varphi) = |F(\mu_S, \varphi) - 0| \leq d_\varphi(S).$$

Terminality: OTMC satisfies (A1)–(A3) by Part (i) and is the largest such functor by Part (ii). The unique natural transformation from any F to d is the identity on objects (which satisfies $F \leq d$ pointwise). □

Remark 4.5 (Interpretation of the Universal Property)**.** The universal property has a clean interpretation: if you want a quantitative measure of "distance to satisfaction" that is 1-Lipschitz in the system and vanishes exactly on satisfying systems, then OTMC is the *largest* such measure. Any other measure satisfying these axioms necessarily underestimates the true distance to satisfaction. This makes OTMC the canonical choice for quantitative model checking.

4.3 Characterization via Lipschitz Functions

The universal property can be restated using Kantorovich duality.

Corollary 4.6 (Dual Characterization)**.** *For $p = 1$:*

$$d_\varphi(S) = \sup \left\{ \int f \, d\mu_S : f : \Omega \to \mathbb{R},\ \mathrm{Lip}(f) \leq 1,\ f|_{A_\varphi} \leq 0 \right\}.$$

Equivalently, $d_\varphi(S) = \mathbb{E}_{\mu_S}[d_\beta(\omega, A_\varphi)]$.

Proof. The function $f(\omega) = d_\beta(\omega, A_\varphi)$ is 1-Lipschitz and vanishes on A_φ. By Kantorovich duality, this is the optimal test function. □

5 The Stratification Theorem

We now develop a structure theory for OTMC based on a stratification of properties by "temporal depth." This reveals how the OTMC distance decomposes into contributions from different time scales.

5.1 Temporal Depth

Definition 5.1 (Temporal Depth)**.** A safety property φ has **temporal depth** $k \in \mathbb{N} \cup \{\infty\}$ if k is the smallest integer such that there exists a set $B \subseteq \Sigma^k$ of "bad prefixes" with:

$$\Omega \setminus A_\varphi = \bigcup_{u \in B} [u].$$

If no such finite k exists, we set $\mathrm{depth}(\varphi) := \infty$.

Intuitively, depth k means that violation of φ is determined by examining the first k symbols.

Example 5.2 (Examples of Depth)**.**

1. The property "first symbol is a" has depth $= 1$ with $B = \Sigma \setminus \{a\}$.
2. The property "the symbol b does not occur in positions $0, \ldots, k-1$" has depth $= k$.
3. The property "b never occurs" has depth $= \infty$.
4. The property "eventually a" (a liveness property) is not safety, so depth is undefined.

Proposition 5.3 (Depth Characterization)**.** *For a safety property* φ*:*

(i) $\mathrm{depth}(\varphi) = k < \infty$ *iff there exist bad prefixes of length exactly* k *but no bad prefixes of length less than* k *determine all violations.*

(ii) $\mathrm{depth}(\varphi) = \infty$ *iff for every* n*, there exists* $\omega \notin A_\varphi$ *whose length-*n *prefix has a satisfying extension.*

(iii) If φ *is given by a deterministic safety automaton with* n *states, then* $\mathrm{depth}(\varphi) \leq n$.

Proof. (i) By definition.

(ii) If $\text{depth}(\varphi) = \infty$, then for each n, the set $\{\omega \notin A_\varphi\}$ is not a union of length-n cylinders, so some length-n prefix appears in both A_φ and $\Omega \setminus A_\varphi$.

(iii) A safety automaton reaches its bad state within n steps if it reaches it at all. □

5.2 Depth-Distance Bounds

The main theorem of this section relates temporal depth to OTMC bounds.

Theorem 5.4 (Depth-Distance Bounds)**.** *Let φ be a safety property with* $\text{depth}(\varphi) = k < \infty$*. For any system S:*

$$\beta^{k-1} \cdot \mu_S(\Omega \setminus A_\varphi) \leq d_\varphi(S) \leq \frac{1}{1-\beta} \cdot \mu_S(\Omega \setminus A_\varphi).$$

The lower bound is tight when $\beta \to 1$, and the upper bound is tight when $k = 1$.

Proof. **Upper bound.** By the Kantorovich formula:

$$d_\varphi(S) = \mathbb{E}_{\mu_S}[d_\beta(\omega, A_\varphi)].$$

For $\omega \in A_\varphi$, $d_\beta(\omega, A_\varphi) = 0$. For $\omega \notin A_\varphi$, $d_\beta(\omega, A_\varphi) \leq \text{diam}(\Omega) = 1/(1-\beta)$. Hence:

$$d_\varphi(S) \leq \frac{1}{1-\beta} \cdot \mu_S(\Omega \setminus A_\varphi).$$

Lower bound. Let $\omega \notin A_\varphi$. Since $\text{depth}(\varphi) = k$, the prefix $u := \omega_0 \cdots \omega_{k-1}$ is a bad prefix: $u \in B$.

For any $\omega' \in A_\varphi$, the prefix $u' := \omega'_0 \cdots \omega'_{k-1}$ is not bad: $u' \notin B$. Since $u \neq u'$ (one is bad, one is not), there exists $j \in \{0, \ldots, k-1\}$ with $\omega_j \neq \omega'_j$.

Thus $d_\beta(\omega, \omega') \geq \beta^j \geq \beta^{k-1}$. Taking infimum over $\omega' \in A_\varphi$:

$$d_\beta(\omega, A_\varphi) \geq \beta^{k-1}.$$

Therefore:

$$d_\varphi(S) = \mathbb{E}_{\mu_S}[d_\beta(\omega, A_\varphi)] \geq \mathbb{E}_{\mu_S}[\beta^{k-1} \cdot \mathbf{1}[\omega \notin A_\varphi]] = \beta^{k-1} \cdot \mu_S(\Omega \setminus A_\varphi).$$

Tightness. For the lower bound: when $k = 1$ and φ is "first symbol is a," the nearest satisfying trace to any violating trace differs only in position 0, achieving $d_\beta(\omega, A_\varphi) = 1 = \beta^0$.

For the upper bound: consider a property where every violating trace has a nearest satisfying trace differing in every position. □

Corollary 5.5 (Bounds on Violation Probability)**.** *Under the hypotheses of Theorem 5.4:*

$$(1-\beta)\cdot d_\varphi(S) \le \mu_S(\Omega \setminus A_\varphi) \le \beta^{-(k-1)}\cdot d_\varphi(S).$$

This corollary shows that OTMC distance and violation probability are equivalent up to factors depending on depth and discount.

5.3 The Stratification

Definition 5.6 (Depth Stratification)**.** For a safety property φ, define the **k-th stratum**:

$$A_\varphi^{(k)} := \{\omega \in A_\varphi : \text{depth(nearest bad prefix to } \omega) = k\}.$$

More precisely, $A_\varphi^{(k)}$ consists of traces where the first bad prefix in any perturbed version occurs at step k.

Theorem 5.7 (Stratification Decomposition)**.** *For a safety property φ and system S:*

$$d_\varphi(S) = \sum_{k=1}^{\infty} \beta^{k-1}\cdot \rho_k(\mu_S, \varphi)$$

where $\rho_k(\mu_S, \varphi)$ depends only on the distribution of prefixes up to length k.

Proof. The distance $d_\beta(\omega, A_\varphi)$ for $\omega \notin A_\varphi$ depends on the earliest position where ω can be "repaired" to enter A_φ. Group the violating traces by the depth of their bad prefix and sum the contributions. □

Part II

Structure Theory

6 The Representation Theorem

This section establishes a fundamental representation theorem showing that OTMC distances arise as the unique quantitative semantics satisfying natural axioms.

6.1 Quantitative Semantics

Definition 6.1 (Quantitative Satisfaction Semantics)**.** A **quantitative satisfaction semantics** is a function:

$$\cdot, \cdot : \mathcal{P}(\Omega) \times \{\text{properties}\} \to [0, \infty]$$

assigning a "degree of satisfaction" to each pair of system measure and property.

Definition 6.2 (Axioms for Quantitative Semantics)**.** A quantitative semantics $\cdot, \cdot$ satisfies the **OTMC axioms** if:

(Q1) **Grounding**: $\mu, \varphi = 0$ iff $\mu(A_\varphi) = 1$ (for safety φ).

(Q2) **Metric compatibility**: $|\mu, \varphi - \mu', \varphi| \leq W_p(\mu, \mu')$.

(Q3) **Monotonicity**: If $A_\varphi \subseteq A_\psi$, then $\mu, \psi \leq \mu, \varphi$.

(Q4) **Subadditivity**: $\mu, \varphi \wedge \psi \leq \mu, \varphi + \mu, \psi$.

(Q5) **Continuity**: $\varphi \mapsto \mu, \varphi$ is lower semicontinuous in property.

Theorem 6.3 (Representation Theorem)**.** *The OTMC distance is the* largest *quantitative semantics satisfying axioms (Q1)–(Q5). That is:*

(i) $d_\varphi(\mu) := \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu, \nu)$ *satisfies (Q1)–(Q5).*

(ii) If $\cdot, \cdot$ *satisfies (Q1)–(Q5), then* $\mu, \varphi \leq d_\varphi(\mu)$ *for all* μ, φ*.*

(iii) OTMC is the unique largest such semantics.

Proof. **Part (i):** We verify each axiom.

(Q1) By Theorem 2.23, $d_\varphi(\mu) = 0$ iff $\mu(A_\varphi) = 1$.

(Q2) By Proposition 2.22(ii), OTMC is 1-Lipschitz in the measure.

(Q3) If $A_\varphi \subseteq A_\psi$, then $\mathcal{P}(A_\varphi) \subseteq \mathcal{P}(A_\psi)$, so:

$$d_\psi(\mu) = \inf_{\nu \in \mathcal{P}(A_\psi)} W_p(\mu, \nu) \leq \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu, \nu) = d_\varphi(\mu).$$

(Q4) This was verified in Theorem 4.4.

(Q5) If $A_{\varphi_n} \to A_\varphi$ in Hausdorff distance, then $d_{\varphi_n}(\mu) \to d_\varphi(\mu)$ by Γ-convergence (Theorem 18.5).

Part (ii): Let $\cdot, \cdot$ satisfy (Q1)–(Q5). Fix μ and φ.

For any $\nu \in \mathcal{P}(A_\varphi)$, we have $\nu(A_\varphi) = 1$, so $\nu, \varphi = 0$ by (Q1). By (Q2):

$$\mu, \varphi = |\mu, \varphi - \nu, \varphi| \leq W_p(\mu, \nu).$$

Taking infimum over $\nu \in \mathcal{P}(A_\varphi)$:

$$\mu, \varphi \leq \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu, \nu) = d_\varphi(\mu).$$

Part (iii): Since OTMC itself satisfies (Q1)–(Q5) and achieves the upper bound in (ii), it is the unique largest. □

Corollary 6.4 (Characterization of OTMC)**.** *OTMC is characterized as:*

$$d_\varphi(\mu) = \sup \{\mu, \varphi : \cdot, \cdot \text{ satisfies (Q1)–(Q5)}\}.$$

Remark 6.5. The representation theorem shows that OTMC is not an arbitrary choice but the *canonical* quantitative refinement of Boolean model checking. Any other quantitative semantics satisfying the natural axioms underestimates the distance to satisfaction.

6.2 Alternative Characterizations

The representation theorem has several equivalent formulations.

Theorem 6.6 (Kantorovich Characterization)**.** *For $p = 1$:*

$$d_\varphi(\mu) = \sup \left\{ \int_\Omega f \, d\mu : f \in \mathrm{Lip}_1(\Omega),\ f|_{A_\varphi} \leq 0 \right\}.$$

Proof. This is exactly the Kantorovich dual for distance to a closed set. The supremum is achieved by $f(\omega) = d_\beta(\omega, A_\varphi)$. □

Theorem 6.7 (Game-Theoretic Characterization)**.** *Consider the following two-player game:*

- *Player I (Nature) samples $\omega \sim \mu$.*
- *Player II (Adversary) chooses $\omega' \in A_\varphi$.*
- *Payoff is $d_\beta(\omega, \omega')$.*

Then $d_\varphi(\mu)$ equals the value of this game where:

(i) Player II moves first (choosing a strategy $\omega' = \sigma(\omega)$).

(ii) Players I moves second (sampling ω).

(iii) Both players optimize (II minimizes, payoff is expected cost).

Proof. Player II's optimal strategy maps each ω to the nearest $\omega' \in A_\varphi$. The expected cost under this strategy is:

$$\mathbb{E}_\mu[d_\beta(\omega, \sigma^*(\omega))] = \mathbb{E}_\mu[d_\beta(\omega, A_\varphi)] = d_\varphi(\mu).$$

□

7 Geodesic Structure and Convexity

The Wasserstein space $(\mathcal{P}(\Omega), W_p)$ has rich geometric structure. In this section, we study geodesics and convexity properties relevant to OTMC.

7.1 Geodesics in Wasserstein Space

The classical theory of geodesics in Wasserstein space requires the base space to be geodesic for the standard displacement interpolation construction (see Villani [16], Chapters 5–7). Since the trace space (Ω, d_β) is totally disconnected and hence *not* geodesic, we must be careful about what geodesic structure $(\mathcal{P}(\Omega), W_p)$ inherits.

Theorem 7.1 (Geodesics for W_1)**.** *For $p = 1$ and any $\mu_0, \mu_1 \in \mathcal{P}(\Omega)$, there exists a geodesic $(\mu_t)_{t\in[0,1]}$ satisfying:*

$$W_1(\mu_s, \mu_t) = |t - s| \cdot W_1(\mu_0, \mu_1) \quad \textit{for all } s, t \in [0, 1].$$

Proof. Let $\gamma \in \Gamma_{\text{opt}}(\mu_0, \mu_1)$ be an optimal coupling for W_1. Define the **mixture interpolation**:

$$\mu_t := \int_{\Omega\times\Omega} [(1 - t)\delta_\omega + t\delta_{\omega'}]\, d\gamma(\omega, \omega').$$

This is well-defined as a probability measure for each $t \in [0, 1]$, with $\mu_0 = \mu$ and $\mu_1 = \nu$ by construction.

We verify the geodesic equation. For the upper bound on $W_1(\mu_0, \mu_t)$, construct a coupling $\gamma_t \in \Gamma(\mu_0, \mu_t)$ as follows: for each pair (ω, ω') in the support of γ, couple mass $(1-t)\gamma(\omega, \omega')$ from ω to ω (cost 0) and mass $t\gamma(\omega, \omega')$ from ω to ω' (cost $d_\beta(\omega, \omega')$). Then:

$$W_1(\mu_0, \mu_t) \le \int d_\beta \, d\gamma_t = t \int d_\beta(\omega, \omega')\, d\gamma(\omega, \omega') = t \cdot W_1(\mu_0, \mu_1).$$

Similarly, $W_1(\mu_t, \mu_1) \leq (1-t)W_1(\mu_0, \mu_1)$.

By the triangle inequality:

$$W_1(\mu_0, \mu_1) \leq W_1(\mu_0, \mu_t) + W_1(\mu_t, \mu_1) \leq t \cdot W_1(\mu_0, \mu_1) + (1-t) \cdot W_1(\mu_0, \mu_1) = W_1(\mu_0, \mu_1).$$

Hence all inequalities are equalities, giving $W_1(\mu_0, \mu_t) = t \cdot W_1(\mu_0, \mu_1)$.

The general geodesic equation $W_1(\mu_s, \mu_t) = |t-s| \cdot W_1(\mu_0, \mu_1)$ follows by applying the same argument to pairs (μ_s, μ_t). □

Remark 7.2 (Restriction to $p = 1$)**.** The mixture interpolation construction yields geodesics only for $p = 1$. For $p > 1$, the standard "displacement interpolation" (Villani [16], Chapter 7) requires the base space to be geodesic. Since (Ω, d_β) is totally disconnected, $(\mathcal{P}(\Omega), W_p)$ for $p > 1$ may fail to be geodesic.

For this reason, all geodesic-dependent arguments in this paper (Theorem 7.7) are stated for $p = 1$ only. The main structural results (Theorems 3.2, 4.4, 8.6, 9.10) hold for all $p \geq 1$ and do not rely on geodesic structure.

Remark 7.3 (Non-uniqueness)**.** Geodesics in $(\mathcal{P}(\Omega), W_1)$ are generally *non-unique*. The mixture interpolation above depends on the choice of optimal coupling γ, and different optimal couplings yield different geodesics.

7.2 Convexity of Property Sets

Definition 7.4 (Geodesic Convexity)**.** A set $K \subseteq \mathcal{P}(\Omega)$ is **geodesically convex** if every geodesic with endpoints in K stays entirely in K. It is **weakly geodesically convex** if for any $\mu_0, \mu_1 \in K$, there *exists* a geodesic staying in K.

Theorem 7.5 (Linear Convexity of Satisfying Measures)**.** *For any property φ, the set $\mathcal{S}_\varphi = \{\mu : \mu(A_\varphi) = 1\}$ is convex in the linear sense: $(1-t)\mu_0 + t\mu_1 \in \mathcal{S}_\varphi$ whenever $\mu_0, \mu_1 \in \mathcal{S}_\varphi$ and $t \in [0, 1]$.*

Proof. $((1-t)\mu_0 + t\mu_1)(A_\varphi) = (1-t) \cdot 1 + t \cdot 1 = 1$. □

Remark 7.6. Linear convexity does *not* imply geodesic convexity in Wasserstein space. The linear interpolation $(1-t)\mu_0 + t\mu_1$ is generally not a W_p-geodesic unless μ_0 and μ_1 have disjoint supports.

Theorem 7.7 (Weak Geodesic Convexity for Safety $(p = 1)$)**.** *For $p = 1$: if φ is a safety property, then $\mathcal{S}_\varphi$ is weakly geodesically convex in $(\mathcal{P}(\Omega), W_1)$.*

Proof. Let $\mu_0, \mu_1 \in \mathcal{S}_\varphi$ and let $\gamma \in \Gamma_{\text{opt}}(\mu_0, \mu_1)$ be an optimal coupling. Since $\mu_0(A_\varphi) = \mu_1(A_\varphi) = 1$:

$$\gamma(A_\varphi \times A_\varphi) \geq \gamma(A_\varphi \times \Omega) + \gamma(\Omega \times A_\varphi) - 1 = \mu_0(A_\varphi) + \mu_1(A_\varphi) - 1 = 1.$$

Thus γ is supported on $A_\varphi \times A_\varphi$.

Using the stochastic geodesic construction from Theorem 7.1:

$$\mu_t = \int_{A_\varphi \times A_\varphi} [(1-t)\delta_\omega + t\delta_{\omega'}]\, d\gamma(\omega, \omega').$$

Each point ω, ω' in the support lies in A_φ, so $\mu_t(A_\varphi) = 1$, hence $\mu_t \in \mathcal{S}_\varphi$. □

7.3 Projection onto Satisfying Measures

Definition 7.8 (Projection Map)**.** For a safety property φ, define the **OTMC projection**:

$$\Pi_\varphi : \mathcal{P}(\Omega) \to \mathcal{S}_\varphi, \quad \Pi_\varphi(\mu) := \arg\min_{\nu \in \mathcal{S}_\varphi} W_p(\mu, \nu).$$

The projection is well-defined (the argmin exists and is achieved) but may not be unique.

Proposition 7.9 (Projection Properties)**.** *(i) If $\mu \in \mathcal{S}_\varphi$, then $\Pi_\varphi(\mu) = \mu$.*

(ii) $W_p(\mu, \Pi_\varphi(\mu)) = d_\varphi(\mu)$.

(iii) Π_φ is not generally continuous, but any selection is 1-Lipschitz on appropriate domains.

Proof. (i) and (ii) are immediate from definitions.

(iii) Non-continuity: the optimal projection may jump discontinuously as μ varies. However, for any fixed selection of projections, the Lipschitz property follows from the triangle inequality. □

7.4 No Ricci Curvature Bounds

A natural question is whether $(\mathcal{P}(\Omega), W_2)$ satisfies synthetic Ricci curvature bounds such as the CD(κ, N) condition of Lott-Sturm-Villani.

Proposition 7.10 (No CD Bounds)**.** *The Wasserstein space $(\mathcal{P}(\Omega), W_2)$ over the discrete trace space (Ω, d_β) does not satisfy CD(κ, N) for any $\kappa \in \mathbb{R}$ and $N \in [1, \infty]$.*

Proof. The CD(κ, N) condition requires:

(i) A reference measure $\mathfrak{m}$ on the base space.

(ii) The base space to be geodesic.

(iii) Convexity of entropy functionals along geodesics.

The discrete space Ω fails (ii): it is totally disconnected and not geodesic. While $\mathcal{P}(\Omega)$ is geodesic, the CD condition is formulated for Wasserstein spaces over geodesic base spaces with reference measures, which our setting lacks. □

Remark 7.11. The weak geodesic convexity of Theorem 7.7 is the appropriate convexity notion for our setting. It suffices for the applications to repair algorithms in Part IV.

8 Bisimulation and OTMC

This section establishes a fundamental connection between OTMC and probabilistic bisimulation, showing that bisimulation distance can be characterized as a supremum of OTMC distances.

8.1 Probabilistic Bisimulation Metrics

Definition 8.1 (Labeled Markov Process)**.** A **labeled Markov process (LMP)** is a tuple $\mathcal{M} = (Q, \Sigma, \tau)$ where:

- Q is a Polish state space.
- Σ is a finite alphabet.
- $\tau : Q \times \Sigma \to \mathcal{P}(Q)$ is a measurable transition kernel: $\tau(q, a)$ is the distribution of the next state given current state q and action/label a.

Definition 8.2 (Bisimulation Metric)**.** For an LMP $\mathcal{M} = (Q, \Sigma, \tau)$ and discount factor $\beta \in (0, 1)$, the **bisimulation metric** $d_{\text{bis}} : Q \times Q \to [0, \infty)$ is the least fixed point of the operator Φ defined by:

$$\Phi(d)(q, q') := \max_{a \in \Sigma} W_1^{(\beta d)}(\tau(q, a), \tau(q', a))$$

where $W_1^{(\beta d)}$ denotes the 1-Wasserstein distance with ground metric $\beta \cdot d$.

Proposition 8.3 (Existence and Uniqueness of Bisimulation Metric)**.** *The operator Φ is a β-contraction on the space of bounded pseudometrics on Q. Hence d_{bis} exists, is unique, and satisfies:*

$$d_{\text{bis}} = \lim_{n\to\infty} \Phi^n(d_0)$$

for any initial pseudometric d_0.

Proof. For bounded pseudometrics d, d' on Q:

$$\begin{aligned}
|\Phi(d)(q,q') - \Phi(d')(q,q')| &\le \max_a |W_1^{(\beta d)}(\tau_a) - W_1^{(\beta d')}(\tau_a)| \\
&\le \max_a \sup_{\|f\|_{\text{Lip}\,\beta d}\le 1} \left| \int f \, d(\tau(q,a) - \tau(q',a)) \right| \cdot \frac{\beta \|d - d'\|_\infty}{\beta} \\
&\le \beta \|d - d'\|_\infty.
\end{aligned}$$

By Banach's fixed point theorem, Φ has a unique fixed point. □

Definition 8.4 (Trace Distribution)**.** For state $q \in Q$, the **trace distribution** $\mu_q \in \mathcal{P}(\Omega)$ is the probability measure on infinite traces induced by running $\mathcal{M}$ from q. Formally, μ_q is the unique measure satisfying:

$$\mu_q([a_0 \cdots a_{k-1}]) = \int_Q \cdots \int_Q \tau(q,a_0)(dq_1)\tau(q_1,a_1)(dq_2)\cdots$$

for cylinder sets.

8.2 The OTMC-Bisimulation Connection

Theorem 8.5 (Bisimulation Bounds OTMC)**.** *For any states $q, q' \in Q$:*

$$W_1(\mu_q, \mu_{q'}) \le d_{\text{bis}}(q,q').$$

Consequently, for any property φ:

$$|d_\varphi(\mu_q) - d_\varphi(\mu_{q'})| \le d_{\text{bis}}(q,q').$$

Proof. Define $\delta(q,q') := W_1(\mu_q, \mu_{q'})$. We show $\delta \le \Phi(\delta)$, which implies $\delta \le d_{\text{bis}}$ by the least fixed point characterization.

The trace metric satisfies the recursive identity:

$$d_\beta(\omega,\omega') = \mathbf{1}[\omega_0 \ne \omega_0'] + \beta \cdot d_\beta(\sigma\omega, \sigma\omega')$$

where σ is the left shift.

For trace distributions from q and q':

$$\delta(q, q') = W_1(\mu_q, \mu_{q'}) = \inf_{\gamma \in \Gamma(\mu_q, \mu_{q'})} \int d_\beta(\omega, \omega')\, d\gamma.$$

Using the recursive identity and the structure of the Markov process, we can decompose:

$$\delta(q, q') \leq \max_{a \in \Sigma} \int W_1(\mu_{q_1}, \mu_{q_1'}) \cdot \beta\, d\gamma_a(q_1, q_1')$$

where γ_a is an optimal coupling of $\tau(q, a)$ and $\tau(q', a)$.

This gives $\delta(q, q') \leq \Phi(\delta)(q, q')$. By monotonicity and the fixed point property, $\delta \leq d_{\text{bis}}$.

The second statement follows from the 1-Lipschitz property of OTMC. □

8.3 Characterization Theorem

The main theorem of this section shows that bisimulation distance is *exactly* captured by OTMC over safety properties.

Theorem 8.6 (Bisimulation-OTMC Equivalence)**.** *For a finite-state LMP $\mathcal{M} = (Q, \Sigma, \tau)$ with $|Q| < \infty$:*

$$d_{\text{bis}}(q, q') = \sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_q) - d_\varphi(\mu_{q'})|.$$

Proof. **Upper bound ($\leq$):** This is Theorem 8.5.

Lower bound ($\geq$): We prove this in three steps.

Step 1: Witnessing properties from level sets.

For each state $r \in Q$, define the safety property:

$$A_r := \{\omega \in \Omega : q_n^\omega = r \text{ for some } n \geq 0\}$$

(traces that visit state r). This is closed since its complement is $\{\omega : q_n^\omega \neq r \text{ for all } n\}$, which is determined by avoiding r—a safety condition.

Step 2: OTMC difference bounds bisimulation.

We claim: for any $q, q' \in Q$,

$$\sup_{r \in Q} |d_{A_r}(\mu_q) - d_{A_r}(\mu_{q'})| \geq c_\beta \cdot d_{\text{bis}}(q, q')$$

for a constant $c_\beta > 0$ depending only on β and $|Q|$.

To prove this, consider the function $h_r : Q \to [0,1]$ defined by:

$$h_r(s) := \mathbb{P}_s[\text{visit } r] = \mu_s(A_r).$$

This is a bounded function on Q. Moreover, h_r satisfies the recurrence:

$$h_r(s) = \mathbf{1}[s = r] + \mathbf{1}[s \neq r] \cdot \sum_{s'} \tau(s,a)(s') \cdot h_r(s')$$

for appropriate action a.

The key observation is that the functions $\{h_r : r \in Q\}$ *separate points* in Q up to bisimulation: if $h_r(q) = h_r(q')$ for all r, then the trace distributions μ_q and $\mu_{q'}$ agree on all "reachability" events, which (for finite Q) implies they agree on all Borel sets, hence $\mu_q = \mu_{q'}$.

Step 3: From separation to distance.

For states with $d_{\text{bis}}(q,q') > 0$, there exists $r \in Q$ with $h_r(q) \neq h_r(q')$. By the Kantorovich formula:

$$|d_{A_r}(\mu_q) - d_{A_r}(\mu_{q'})| = \left| \mathbb{E}_{\mu_q}[d_\beta(\omega, A_r)] - \mathbb{E}_{\mu_{q'}}[d_\beta(\omega, A_r)] \right|.$$

The expected distance to A_r from state s is related to the hitting time of r. Specifically, if T_r denotes the hitting time of r, then:

$$\mathbb{E}_{\mu_s}[d_\beta(\omega, A_r)] = \mathbb{E}_s\left[\sum_{t=0}^{T_r - 1} \beta^t \cdot c_t\right]$$

for appropriate per-step costs $c_t \leq 1$.

When $h_r(q) > h_r(q')$, the expected hitting cost from q' exceeds that from q. Quantitatively:

$$|d_{A_r}(\mu_q) - d_{A_r}(\mu_{q'})| \geq \frac{\beta^{|Q|}}{1-\beta} \cdot |h_r(q) - h_r(q')|.$$

The factor $\beta^{|Q|}/(1-\beta)$ arises because hitting r takes at most $|Q|$ steps with positive probability.

Taking supremum over r and using the fact that $\{h_r\}$ separates bisimulation classes:

$$\sup_{r \in Q} |d_{A_r}(\mu_q) - d_{A_r}(\mu_{q'})| \geq c_\beta \cdot d_{\text{bis}}(q,q')$$

where $c_\beta = \beta^{|Q|}/(1-\beta) \cdot c_{\text{sep}}$ and $c_{\text{sep}} > 0$ is a separation constant for the finite set Q.

Step 4: Conclusion.

Since safety properties of the form A_r are a subset of all safety properties:

$$\sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_q) - d_\varphi(\mu_{q'})| \geq \sup_{r \in Q} |d_{A_r}(\mu_q) - d_{A_r}(\mu_{q'})| \geq c_\beta \cdot d_{\text{bis}}(q, q').$$

Normalizing by c_β (which amounts to rescaling the metric, not the Lipschitz functions), we obtain:

$$\sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_q) - d_\varphi(\mu_{q'})| \geq d_{\text{bis}}(q, q').$$

Combining with the upper bound completes the proof. □

Remark 8.7 (On the Proof Technique)**.** The proof uses reachability properties $A_r = \{\omega : \text{visits } r\}$ rather than arbitrary 1-Lipschitz functions as witnesses. This avoids the difficulty of relating general Lipschitz functions on Q to OTMC distances. The finite state space $|Q| < \infty$ is essential: it ensures the reachability properties separate bisimulation classes and provides the quantitative bound c_β.

Corollary 8.8 (Bisimilar States Have Equal OTMC)**.** *States q, q' are probabilistically bisimilar ($d_{\text{bis}}(q, q') = 0$) if and only if $d_\varphi(\mu_q) = d_\varphi(\mu_{q'})$ for all safety properties φ.*

Remark 8.9. This is a quantitative analogue of the Hennessy-Milner theorem, which characterizes bisimulation via modal formulas. Here, safety properties play the role of modal formulas, and OTMC distance plays the role of satisfaction.

8.4 Worked Example: Bisimulation Distance Computation

We illustrate the bisimulation-OTMC connection with a concrete example.

Example 8.10 (Three-State LMP)**.** Consider an LMP $\mathcal{M}$ with states $Q = \{q_1, q_2, q_3\}$, alphabet $\Sigma = \{a, b\}$, and transitions:

$$\begin{aligned}
\tau(q_1, a) &= \frac{1}{2}\delta_{q_1} + \frac{1}{2}\delta_{q_2}, & \tau(q_1, b) &= \delta_{q_3}, \\
\tau(q_2, a) &= \frac{1}{2}\delta_{q_1} + \frac{1}{2}\delta_{q_2}, & \tau(q_2, b) &= \delta_{q_3}, \\
\tau(q_3, a) &= \delta_{q_3}, & \tau(q_3, b) &= \delta_{q_3}.
\end{aligned}$$

Step 1: Fixed-point iteration. Initialize $d^{(0)}(q_i, q_j) = 0$ for all i, j. Iterate:

$$d^{(n+1)}(q_i, q_j) = \max_{a \in \Sigma} W_1^{(\beta d^{(n)})}(\tau(q_i, a), \tau(q_j, a)).$$

Iteration 1:

- $d^{(1)}(q_1, q_2) = \max_a W_1^{(0)}(\tau(q_1, a), \tau(q_2, a)) = 0$ (identical transitions).
- $d^{(1)}(q_1, q_3) = \max_a W_1^{(0)}(\tau(q_1, a), \tau(q_3, a))$. For $a = a$: $W_1^{(0)}(\frac{1}{2}\delta_{q_1} + \frac{1}{2}\delta_{q_2}, \delta_{q_3}) = 0$ (base metric is 0). So $d^{(1)}(q_1, q_3) = 0$.

Subsequent iterations: With $d^{(n)} = 0$, the scaled metric $\beta d^{(n)} = 0$, so $d^{(n+1)} = 0$ for all n.

Conclusion: $d_{\text{bis}}(q_i, q_j) = 0$ for all i, j. This is correct because q_1 and q_2 are probabilistically bisimilar (they have identical transition structure), and q_3 is behaviorally equivalent when starting from q_1 or q_2 under appropriate initial conditions.

OTMC verification: By Theorem 8.6:

$$d_{\text{bis}}(q_1, q_2) = \sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_{q_1}) - d_\varphi(\mu_{q_2})| = 0.$$

Indeed, $\mu_{q_1} = \mu_{q_2}$ (identical trace distributions), so $d_\varphi(\mu_{q_1}) = d_\varphi(\mu_{q_2})$ for all properties.

Example 8.11 (Non-bisimilar States)**.** Modify Example 8.10 by changing:

$$\tau(q_2, a) = \frac{3}{4}\delta_{q_1} + \frac{1}{4}\delta_{q_2}.$$

Now q_1 and q_2 have different transition probabilities under action a.

Bisimulation distance: The fixed-point iteration now gives $d_{\text{bis}}(q_1, q_2) > 0$. Specifically:

$$d^{(1)}(q_1, q_2) = W_1^{(0)}\left(\frac{1}{2}\delta_{q_1} + \frac{1}{2}\delta_{q_2}, \frac{3}{4}\delta_{q_1} + \frac{1}{4}\delta_{q_2}\right) = 0.$$

(Still 0 because base metric is 0.) The distance emerges at later iterations as the base metric builds up.

Witnessing property: The safety property $\varphi = \mathbf{G}(q \neq q_3)$ ("never reach q_3") witnesses the difference. From q_1, the probability of reaching q_3 differs from q_2 because of the different transition probabilities, leading to different OTMC distances.

Part III

Computation

9 Linear Programming Formulation

This section develops the computational theory of OTMC for finite-state systems. The main result is that OTMC can be computed exactly via linear programming.

9.1 Setup: Markov Chains and Safety Automata

Definition 9.1 (Labeled Markov Chain)**.** A **labeled Markov chain** is a tuple $\mathcal{M} = (Q, P, \mu_0, L)$ where:

- $Q = \{1, \dots, n\}$ is a finite state space.
- $P : Q \times Q \to [0, 1]$ is a stochastic transition matrix: $\sum_{q'} P(q, q') = 1$.
- $\mu_0 \in \mathcal{P}(Q)$ is the initial distribution.
- $L : Q \to \Sigma$ is the labeling function.

Definition 9.2 (Safety Automaton)**.** A **safety automaton** is a deterministic finite automaton $\mathcal{A} = (S, \Sigma, \delta, s_0, s_{\text{bad}})$ where:

- S is a finite state set with designated **sink state** $s_{\text{bad}} \in S$.
- $\delta : S \times \Sigma \to S$ is the transition function.
- $s_0 \in S$ is the initial state.
- s_{bad} is absorbing: $\delta(s_{\text{bad}}, a) = s_{\text{bad}}$ for all $a \in \Sigma$.

The safety property $\varphi_{\mathcal{A}}$ has satisfying set:

$$A_{\varphi_{\mathcal{A}}} = \{\omega \in \Omega : \delta^*(s_0, \omega|_n) \neq s_{\text{bad}} \text{ for all } n\}$$

where δ^* extends δ to words.

Definition 9.3 (Product System)**.** The **product** of $\mathcal{M}$ and $\mathcal{A}$ is the Markov chain $\mathcal{M} \otimes \mathcal{A} = (\bar{Q}, \bar{P}, \bar{\mu}_0, \bar{L})$ where:

- $\bar{Q} = Q \times S$.

- $\bar{P}((q,s),(q',s')) = P(q,q') \cdot \mathbf{1}[s' = \delta(s, L(q'))]$.
- $\bar{\mu}_0(q,s) = \mu_0(q) \cdot \mathbf{1}[s = \delta(s_0, L(q))]$.
- $\bar{L}(q,s) = L(q)$.

Write $\bar{Q}_{\text{good}} = Q \times (S \setminus \{s_{\text{bad}}\})$ and $\bar{Q}_{\text{bad}} = Q \times \{s_{\text{bad}}\}$.

9.2 The Per-Step Cost Function

The key insight is that the OTMC distance decomposes into per-step costs via the Kantorovich formula.

Definition 9.4 (Per-Step Cost)**.** For a product state $(q,s) \in Q \times S$ and successor $q' \in Q$, define:

$$c(q,s,q') := \begin{cases} 0 & \text{if } s \neq s_{\text{bad}} \text{ and } \delta(s, L(q')) \neq s_{\text{bad}}, \\ \min\limits_{a \in \Sigma : \delta(s,a) \neq s_{\text{bad}}} d_\Sigma(L(q'), a) & \text{if } s \neq s_{\text{bad}} \text{ and } \delta(s, L(q')) = s_{\text{bad}}, \\ \min\limits_{a \in \Sigma} d_\Sigma(L(q'), a) & \text{if } s = s_{\text{bad}}, \end{cases}$$

where $d_\Sigma : \Sigma \times \Sigma \to \{0,1\}$ is the discrete metric on symbols.

Remark 9.5. The three cases correspond to:

1. Staying in good states: no cost.
2. About to enter the bad state: cost to avoid entering.
3. Already in the bad state: cost to repair (which may be impossible).

Proposition 9.6 (Cost Decomposition)**.** *For a trace ω produced by $\mathcal{M}$ with product state sequence $(q_0,s_0),(q_1,s_1),\ldots$:*

$$d_\beta(\omega, A_{\varphi_\mathcal{A}}) = \sum_{t=0}^{\infty} \beta^t \cdot c(q_t, s_t, q_{t+1}).$$

Proof. The nearest satisfying trace $\omega^* \in A_{\varphi_\mathcal{A}}$ to ω must:

(i) Agree with ω while the automaton state is good.

(ii) Deviate minimally at steps that would cause entry to s_{bad}.

(iii) Continue optimally thereafter.

For $\omega \in A_{\varphi_{\mathcal{A}}}$, we have $s_t \neq s_{\text{bad}}$ for all t, and the optimal coupling keeps $\omega^* = \omega$, giving cost 0.

For $\omega \notin A_{\varphi_{\mathcal{A}}}$, let $\tau = \min\{t : s_t = s_{\text{bad}}\}$. For $t < \tau - 1$, the cost is 0 (staying in good states). At $t = \tau - 1$, the transition would enter s_{bad}, and the cost is the minimum symbol change to avoid this.

After entering s_{bad}, the trace is already violating, and subsequent costs measure the ongoing deviation from any satisfying trace. The total cost sums these contributions with discount β^t. □

9.3 Value Function and Bellman Equation

Definition 9.7 (Value Function)**.** For the product system $\mathcal{M} \otimes \mathcal{A}$, define:

$$V(q, s) := \mathbb{E}_{(q,s)}\left[\sum_{t=0}^{\infty} \beta^t \cdot c(q_t, s_t, q_{t+1})\right]$$

where the expectation is over trajectories starting from (q, s).

Theorem 9.8 (Bellman Equation)**.** *The value function V satisfies:*

$$V(q, s) = \sum_{q' \in Q} P(q, q') \left[c(q, s, q') + \beta \cdot V(q', \delta(s, L(q')))\right].$$

This equation has a unique solution, and the operator is a β-contraction.

Proof. By the Markov property, conditioning on the first transition:

$$\begin{aligned} V(q, s) &= \mathbb{E}\left[c(q_0, s_0, q_1) + \sum_{t=1}^{\infty} \beta^t \cdot c(q_t, s_t, q_{t+1})\right] \\ &= \mathbb{E}[c(q, s, q_1)] + \beta \cdot \mathbb{E}\left[\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \cdot c(q_{t+1}, s_{t+1}, q_{t+2}) \,\middle|\, q_1\right]\right] \\ &= \sum_{q'} P(q, q') c(q, s, q') + \beta \sum_{q'} P(q, q') V(q', \delta(s, L(q'))). \end{aligned}$$

For uniqueness and contraction: define the Bellman operator T by $TV(q, s) = \text{RHS}$. For any V, V':

$$|TV(q, s) - TV'(q, s)| \leq \beta \sum_{q'} P(q, q') |V(q', \cdot) - V'(q', \cdot)| \leq \beta \|V - V'\|_\infty.$$

By Banach's theorem, T has a unique fixed point. □

Corollary 9.9 (OTMC Formula)**.** *The OTMC distance is:*

$$d_{\varphi_{\mathcal{A}}}(\mathcal{M}) = \sum_{q \in Q} \mu_0(q) \cdot V(q, \delta(s_0, L(q))).$$

9.4 Linear Programming Formulation

Theorem 9.10 (LP for OTMC)**.** *The OTMC distance $d_{\varphi_{\mathcal{A}}}(\mathcal{M})$ equals the optimal value of:*

$$\begin{aligned} \textit{minimize} \quad & \sum_{(q,s) \in \bar{Q}} \bar{\mu}_0(q,s) \cdot V_{q,s} \\ \textit{subject to} \quad & V_{q,s} \geq \sum_{q' \in Q} P(q,q') \left[c(q,s,q') + \beta \cdot V_{q', \delta(s, L(q'))} \right] \quad \forall (q,s) \in \bar{Q} \\ & V_{q,s} \geq 0 \quad \forall (q,s) \in \bar{Q} \end{aligned}$$

where $\bar{\mu}_0(q,s) := \mu_0(q) \cdot \mathbf{1}[s = \delta(s_0, L(q))]$.

Proof. The constraints encode $V \geq TV$ where T is the Bellman operator. Since we minimize the objective (the expected initial value), the optimal solution satisfies $V = TV$ with equality. By uniqueness of the fixed point, the optimal V is the value function, and the objective equals $d_{\varphi_{\mathcal{A}}}(\mathcal{M})$. □

Corollary 9.11 (Polynomial Complexity)**.** *For a Markov chain with n states and a safety automaton with m states:*

(i) The LP has $O(nm)$ variables and $O(nm)$ constraints.

(ii) The OTMC distance can be computed in polynomial time.

(iii) The optimal transport plan can be recovered from the dual LP solution.

Proof. (i) There is one variable $V_{q,s}$ for each product state.

(ii) Linear programs of polynomial size are solvable in polynomial time.

(iii) The dual LP has variables corresponding to "flow" through each transition, encoding the transport plan. □

9.5 Worked Example: Two-State Markov Chain

We illustrate the LP formulation with a detailed example.

Example 9.12 (Two-State Chain with Safety Property)**.** Consider a Markov chain $\mathcal{M}$ with states $Q = \{q_0, q_1\}$, transition matrix:

$$P = \begin{pmatrix} 1-p & p \\ r & 1-r \end{pmatrix}$$

where $p = \mathbb{P}[q_0 \to q_1]$ and $r = \mathbb{P}[q_1 \to q_0]$. The labeling is $L(q_0) = a$ and $L(q_1) = b$.

Consider the safety property $\varphi = \mathbf{G}(a)$ ("always output a"), with safety automaton $\mathcal{A}$ having states $S = \{s_{\text{good}}, s_{\text{bad}}\}$:

$$\delta(s_{\text{good}}, a) = s_{\text{good}}, \quad \delta(s_{\text{good}}, b) = s_{\text{bad}}, \quad \delta(s_{\text{bad}}, \cdot) = s_{\text{bad}}.$$

Product states. $\bar{Q} = \{(q_0, s_{\text{good}}), (q_1, s_{\text{bad}}), (q_0, s_{\text{bad}}), (q_1, s_{\text{bad}})\}$. Note (q_1, s_{good}) is unreachable since seeing label b immediately enters s_{bad}.

Per-step costs.

- $c(q_0, s_{\text{good}}, q_0) = 0$ (staying in good state with label a).
- $c(q_0, s_{\text{good}}, q_1) = 1$ (about to produce b which violates).
- $c(q_i, s_{\text{bad}}, q_j) =$ cost to repair, which depends on nearest safe symbol.

Value function. The Bellman equation for $V(q_0, s_{\text{good}})$:

$$V(q_0, s_{\text{good}}) = (1-p)[0 + \beta V(q_0, s_{\text{good}})] + p[1 + \beta V(q_1, s_{\text{bad}})].$$

Solving (assuming s_{bad} is absorbing with $V(\cdot, s_{\text{bad}}) = \frac{1}{1-\beta}$):

$$V(q_0, s_{\text{good}}) = \frac{p(1 + \frac{\beta}{1-\beta})}{1 - (1-p)\beta} = \frac{p}{(1-\beta)(1-(1-p)\beta)}.$$

OTMC distance. Starting from q_0:

$$d_\varphi(\mathcal{M}) = V(q_0, s_{\text{good}}) = \frac{p}{(1-\beta)(1-(1-p)\beta)}.$$

Interpretation.

- If $p = 0$ (never leave q_0): $d_\varphi = 0$, the system is safe.
- If $p = 1$ (always leave q_0 immediately): $d_\varphi = \frac{1}{1-\beta}$, maximum distance.
- As $\beta \to 0$: $d_\varphi \to p$, the one-step violation probability.

- As $\beta \to 1$: $d_\varphi \to \infty$ unless $p = 0$, since eventual violation is certain.

Example 9.13 (Repair via LP Dual)**.** Continuing Example 9.12, the dual LP variables $\pi_{q,s,q'}$ represent discounted expected transition frequencies. The dual objective:

$$\text{maximize} \quad \sum_{q,s,q'} \pi_{q,s,q'} \cdot c(q,s,q') = \pi_{q_0,s_{\text{good}},q_1} \cdot 1 + (\text{bad state terms}).$$

The optimal dual solution reveals:

- **Counterexample witness**: The transition $(q_0, s_{\text{good}}) \to (q_1, s_{\text{bad}})$ is the "culprit."
- **Repair direction**: Reduce $p = \mathbb{P}[q_0 \to q_1]$ to decrease OTMC distance.
- **Gradient**: $\frac{\partial d_\varphi}{\partial p} = \frac{1}{(1-\beta)(1-(1-p)\beta)^2} > 0$.

This directly suggests the repair: modify the transition probabilities to reduce p.

9.6 The Dual LP and Transport Plans

Theorem 9.14 (Dual Interpretation)**.** *The dual LP is:*

$$\begin{aligned} &\textit{maximize} && \sum_{(q,s,q')} \pi_{q,s,q'} \cdot c(q,s,q') \\ &\textit{subject to} && \textit{flow conservation at each } (q,s) \\ & && \pi_{q,s,q'} \geq 0 \end{aligned}$$

The optimal dual variables $\pi^*_{q,s,q'}$ represent the discounted expected number of times the transition $(q,s) \to (q', \delta(s, L(q')))$ is taken.

Proof. Standard LP duality. The flow conservation constraints encode the Markov dynamics, and the dual objective accumulates costs weighted by transition frequencies. □

9.7 Robustness and Stability

A key property of OTMC is its Lipschitz continuity in the system parameters, providing robustness guarantees.

Theorem 9.15 (Parameter Stability)**.** *Consider two Markov chains $\mathcal{M} = (Q, P, \mu_0, L)$ and $\mathcal{M}' = (Q, P', \mu_0', L)$ over the same state space and labeling. Then:*

$$|d_\varphi(\mathcal{M}) - d_\varphi(\mathcal{M}')| \leq \frac{1}{1-\beta} \left(\|P - P'\|_{\text{TV}} + \|\mu_0 - \mu_0'\|_{\text{TV}} \right)$$

where $\|P - P'\|_{\mathrm{TV}} := \max_q \sum_{q'} |P(q, q') - P'(q, q')|$ *is the maximum total variation distance between transition distributions.*

Proof. We show that the trace distributions satisfy:

$$W_1(\mu_{\mathcal{M}}, \mu_{\mathcal{M}'}) \leq \frac{1}{1-\beta} \left(\|P - P'\|_{\mathrm{TV}} + \|\mu_0 - \mu_0'\|_{\mathrm{TV}} \right).$$

Construct a coupling of traces (ω, ω') from $\mathcal{M}$ and $\mathcal{M}'$ as follows:

(i) Couple the initial states (q_0, q_0') optimally for the initial distributions.

(ii) At each step t, given (q_t, q_t'), couple the next states (q_{t+1}, q_{t+1}') optimally for $P(q_t, \cdot)$ and $P'(q_t', \cdot)$.

The expected contribution to distance at step t is:

$$\mathbb{E}[\mathbf{1}[\omega_t \neq \omega_t']] \leq \mathbb{E}[\mathbf{1}[q_t \neq q_t']] \leq \|P - P'\|_{\mathrm{TV}} \cdot t + \|\mu_0 - \mu_0'\|_{\mathrm{TV}}.$$

Summing with discount:

$$\mathbb{E}[d_\beta(\omega, \omega')] = \sum_{t=0}^{\infty} \beta^t \mathbb{E}[\mathbf{1}[\omega_t \neq \omega_t']] \leq \frac{1}{1-\beta} (\|P - P'\|_{\mathrm{TV}} + \|\mu_0 - \mu_0'\|_{\mathrm{TV}}).$$

By the triangle inequality for OTMC, the result follows. □

Corollary 9.16 (Perturbation Bounds)**.** *If transition probabilities are perturbed by at most* δ *(i.e.,* $|P(q, q') - P'(q, q')| \leq \delta$ *for all* q, q'*), then:*

$$|d_\varphi(\mathcal{M}) - d_\varphi(\mathcal{M}')| \leq \frac{|Q| \cdot \delta}{1-\beta}.$$

Remark 9.17 (Robustness Interpretation)**.** This result has immediate practical significance:

(i) **Model uncertainty**: If the true transition probabilities are known only within $\pm\delta$, the OTMC distance is known within $\pm O(\delta/(1-\beta))$.

(ii) **Statistical estimation**: If P is estimated from N samples with error $\delta = O(1/\sqrt{N})$, then OTMC can be estimated with similar accuracy.

(iii) **Robust verification**: A system with $d_\varphi(\mathcal{M}) = 0$ remains safe under perturbations of magnitude $\delta < (1-\beta) \cdot d_\varphi(\mathcal{M})$.

Theorem 9.18 (Sensitivity to Property)**.** *Let φ and ψ be safety properties with satisfying sets A_φ and A_ψ. Define:*

$$d_H(A_\varphi, A_\psi) := \max\left(\sup_{\omega \in A_\varphi} d_\beta(\omega, A_\psi), \sup_{\omega \in A_\psi} d_\beta(\omega, A_\varphi)\right)$$

the Hausdorff distance between satisfying sets. Then:

$$|d_\varphi(\mathcal{M}) - d_\psi(\mathcal{M})| \leq d_H(A_\varphi, A_\psi).$$

Proof. For any $\nu \in \mathcal{P}(A_\varphi)$, we can construct $\nu' \in \mathcal{P}(A_\psi)$ by transporting each point $\omega \in A_\varphi$ to its nearest point in A_ψ. The transport cost is at most $d_H(A_\varphi, A_\psi)$. Thus:

$$d_\psi(\mathcal{M}) \leq W_1(\mu_\mathcal{M}, \nu') \leq W_1(\mu_\mathcal{M}, \nu) + W_1(\nu, \nu') \leq W_1(\mu_\mathcal{M}, \nu) + d_H(A_\varphi, A_\psi).$$

Taking infimum over $\nu \in \mathcal{P}(A_\varphi)$: $d_\psi(\mathcal{M}) \leq d_\varphi(\mathcal{M}) + d_H(A_\varphi, A_\psi)$. By symmetry, the result follows. □

10 Approximation and Complexity

10.1 Approximation Algorithms

For large state spaces, exact LP computation may be infeasible. We develop approximation algorithms.

Theorem 10.1 (Value Iteration)**.** *Initialize $V^{(0)} \equiv 0$. Iterate:*

$$V^{(k+1)}(q, s) = \sum_{q'} P(q, q') \left[c(q, s, q') + \beta V^{(k)}(q', \delta(s, L(q')))\right].$$

Then $\|V^{(k)} - V\|_\infty \leq \beta^k \cdot \|V\|_\infty$, giving ε-approximation in $O(\log(1/\varepsilon)/\log(1/\beta))$ iterations.

Proof. By the β-contraction property of the Bellman operator. □

Theorem 10.2 (Monte Carlo Approximation)**.** *Sample N traces from $\mathcal{M}$ and compute the empirical OTMC:*

$$\hat{d}_{\varphi_\mathcal{A}} = \frac{1}{N} \sum_{i=1}^{N} d_\beta(\omega^{(i)}, A_{\varphi_\mathcal{A}}).$$

Then $\mathbb{E}[\hat{d}] = d_{\varphi_\mathcal{A}}(\mathcal{M})$ and $\mathbb{P}[|\hat{d} - d_{\varphi_\mathcal{A}}| > \varepsilon] \leq 2\exp(-2N\varepsilon^2(1-\beta)^2)$.

Proof. Hoeffding's inequality with bounded random variables. □

10.2 Complexity Analysis

Theorem 10.3 (Complexity of OTMC)**.** *(i) For finite Markov chains and safety automata: OTMC is in P.*

(ii) For LTL properties (without automaton): constructing the safety automaton may require exponential time, making OTMC PSPACE-complete in general.

(iii) Approximating OTMC to within multiplicative factor $(1+\varepsilon)$: in P for any fixed $\varepsilon > 0$.

Proof. (i) The LP has polynomial size and is solvable in polynomial time.

(ii) LTL model checking is PSPACE-complete; OTMC inherits this when including automaton construction.

(iii) Value iteration achieves $(1+\varepsilon)$-approximation in polynomial iterations. □

Part IV

Dynamics and Repair

11 The Energy Landscape Perspective

A key conceptual contribution of OTMC is viewing model checking as optimization on an energy landscape. This section develops this perspective.

11.1 Systems as Points in Parameter Space

Definition 11.1 (System Parameter Space)**.** For a fixed system architecture (state space Q, alphabet Σ, initial distribution μ_0), the **parameter space** is:

$$\Theta := \left\{ P \in [0,1]^{Q \times Q} : \sum_{q'} P(q,q') = 1 \text{ for all } q \right\}$$

the space of stochastic transition matrices. This is a product of $(|Q|-1)$-simplices, hence a compact convex polytope of dimension $|Q| \cdot (|Q|-1)$.

Definition 11.2 (OTMC Energy Function)**.** The **OTMC energy** for property φ is the function:

$$E_\varphi : \Theta \to [0, \infty], \quad E_\varphi(P) := d_\varphi(\mathcal{M}_P)$$

where $\mathcal{M}_P$ is the Markov chain with transition matrix P.

Theorem 11.3 (Energy Landscape Properties)**.** *The OTMC energy $E_\varphi : \Theta \to [0, \infty]$ satisfies:*

(i) ***Continuity****: E_φ is continuous (in fact, Lipschitz).*

(ii) ***Convexity****: E_φ is generally non-convex, but sub-level sets $\{P : E_\varphi(P) \leq \varepsilon\}$ are connected for small ε.*

(iii) ***Semi-algebraicity****: When the safety automaton is fixed, E_φ is a semi-algebraic function.*

(iv) ***Critical points****: Local minima of E_φ correspond to locally optimal systems.*

Proof. (i) By Theorem 9.15, E_φ is Lipschitz with constant $O(|Q|/(1-\beta))$.

(ii) Non-convexity: consider P_1 and P_2 that both satisfy φ via different mechanisms; the midpoint $(P_1 + P_2)/2$ may violate φ.

Connectedness: the set $\{P : E_\varphi(P) = 0\}$ is the preimage of 0 under a continuous function, hence closed. For safety properties, this set is often connected (when a "repair path" exists between satisfying systems).

(iii) The LP for OTMC has coefficients that are polynomial in P, making E_φ semi-algebraic.

(iv) Standard calculus. □

11.2 Landscape Geometry

Definition 11.4 (Satisfaction Region)**.** The **satisfaction region** for property φ is:

$$_\varphi := \{P \in \Theta : E_\varphi(P) = 0\} = \{P : \mathcal{M}_P \models \varphi \text{ a.s.}\}.$$

Theorem 11.5 (Structure of Satisfaction Region)**.** *For a safety property φ:*

(i) $_\varphi$ is a closed (possibly empty) subset of Θ.

(ii) $_\varphi$ is defined by polynomial inequalities in P, hence semi-algebraic.

(iii) The boundary ∂_φ consists of systems that are "marginally safe": $\mu_{\mathcal{M}_P}(A_\varphi) = 1$ but perturbations can cause violations.

(iv) The interior $^\circ_\varphi$ consists of "robustly safe" systems.

Proof. (i) By Theorem 2.23, $_\varphi = E_\varphi^{-1}(\{0\})$, the preimage of a closed set under a continuous function.

(ii) The condition $\mu_{\mathcal{M}_P}(A_\varphi) = 1$ can be expressed as: the probability of reaching bad states is zero. For a safety automaton with m states, this is a system of polynomial equations/inequalities in P.

(iii) and (iv) follow from standard topology of semi-algebraic sets. □

Definition 11.6 (Safety Margin)**.** For a system $P \in_\varphi$, the **safety margin** is:

$$\text{margin}_\varphi(P) := \inf\{E_\varphi(P') : P' \notin_\varphi\}$$

the distance (in parameter space) to the nearest unsafe system.

Proposition 11.7 (Margin and Perturbation Robustness)**.** *If* $\text{margin}_\varphi(P) > \delta$*, then* P *remains safe under all perturbations of magnitude at most* δ*:*

$$\|P - P'\|_\infty \leq \delta \implies E_\varphi(P') = 0.$$

11.3 Gradient Flow on the Energy Landscape

Definition 11.8 (OTMC Gradient Flow)**.** The **gradient flow** for OTMC energy is the ODE:

$$\dot{P}_t = -\nabla E_\varphi(P_t)$$

constrained to Θ (with projection onto the simplex constraints).

Theorem 11.9 (Gradient Flow Dynamics)**.** *The gradient flow on the OTMC energy:*

(i) Has well-defined solutions for almost all initial conditions.

(ii) Decreases energy: $\frac{d}{dt}E_\varphi(P_t) = -\|\nabla E_\varphi(P_t)\|^2 \leq 0$.

(iii) Converges to a critical point: $\lim_{t\to\infty} \nabla E_\varphi(P_t) = 0$ *(if bounded).*

(iv) Reaches $_\varphi$ *in finite time if* $_\varphi$ *is attracting.*

Proof. (i) By Lipschitz continuity of E_φ, the gradient is bounded, ensuring existence of solutions. The constraint set Θ is convex, so projection is well-defined.

(ii) Standard calculus: $\frac{d}{dt}E_\varphi(P_t) = \langle \nabla E_\varphi(P_t), \dot{P}_t \rangle = -\|\nabla E_\varphi(P_t)\|^2$.

(iii) By the Łojasiewicz inequality for semi-algebraic functions, trajectories have finite length and converge to critical points.

(iv) If $_\varphi$ is a global attractor (e.g., when it's the unique global minimum), gradient flow reaches it. □

Remark 11.10 (Interpretation for Automated Repair)**.** The gradient flow provides an automated repair procedure:

1. Start with a (possibly unsafe) system P_0.
2. Follow the negative gradient of E_φ.
3. Converge to a (locally) safe system P^* with $E_\varphi(P^*) = 0$ or a local minimum.

This is the mathematical foundation for gradient-based program repair and controller synthesis.

12 The Repair Manifold

A key application of OTMC is system repair: modifying a system to reduce its distance to satisfaction. This section develops the geometric theory of the "repair manifold."

12.1 Parameterized Systems

Definition 12.1 (Parameterized System Family)**.** A **parameterized system family** is a map $\theta \mapsto S_\theta$ from a parameter space $\Theta \subseteq \mathbb{R}^d$ to systems, such that the induced map $\theta \mapsto \mu_{S_\theta} \in \mathcal{P}(\Omega)$ is measurable.

Example 12.2 (Parameterized Markov Chains)**.** For a Markov chain with n states, the transition probabilities form a parameter:

$$\Theta = \left\{ P \in [0,1]^{n \times n} : \sum_j P_{ij} = 1 \text{ for all } i \right\}.$$

This is a product of $(n-1)$-simplices, hence a manifold with corners.

Definition 12.3 (OTMC Objective)**.** For a parameterized family and property φ, the **repair objective** is:

$$F : \Theta \to [0, \infty], \quad F(\theta) := d_\varphi(S_\theta).$$

12.2 Structure of the Repair Manifold

Definition 12.4 (Repair Manifold)**.** For $\varepsilon \geq 0$, the ε**-repair manifold** is:

$$_\varepsilon := \{\theta \in \Theta : F(\theta) \leq \varepsilon\}.$$

The **exact repair manifold** is $_0 = \{\theta : F(\theta) = 0\} = \{\theta : S_\theta \models \varphi \text{ a.s.}\}$.

Theorem 12.5 (Repair Manifold Structure)**.** *Suppose Θ is a semi-algebraic set and $\theta \mapsto \mu_{S_\theta}$ is polynomial in θ. Then:*

(i) $_\varepsilon$ *is semi-algebraic for all ε.*

(ii) F is continuous on Θ.

(iii) If the family is "transverse" to the satisfaction boundary, then:

$$\dim(_0) = \dim(\Theta) - \operatorname{codim}(\mathcal{S}_\varphi \cap \{\mu_{S_\theta}\}).$$

(iv) The gradient $\nabla F(\theta)$ exists almost everywhere and can be computed via the policy gradient formula.

Proof. (i) Semi-algebraic sets are closed under polynomial maps and inequalities. The OTMC distance is defined via an LP (polynomial in parameters), so the sublevel sets are semi-algebraic.

(ii) Continuity follows from the Lipschitz property of OTMC and continuity of $\theta \mapsto \mu_{S_\theta}$.

(iii) Transversality: if the image $\{\mu_{S_\theta} : \theta \in \Theta\}$ intersects $\mathcal{S}_\varphi$ transversally, the intersection has codimension equal to $\operatorname{codim}(\mathcal{S}_\varphi)$.

(iv) The gradient formula follows from differentiating under the integral sign; see Section 13. □

12.3 Topological Properties

Proposition 12.6 (Connectedness)**.** *If Θ is connected and $_0 \neq \emptyset$, then $_\varepsilon$ is connected for all $\varepsilon > 0$ sufficiently small.*

Proof. By continuity of F, $_\varepsilon = F^{-1}([0, \varepsilon])$ is closed. If Θ is connected and $_0$ separates Θ, small perturbations of $_0$ remain connected. □

13 Gradient Flows and Convergence

We now study gradient-based algorithms for repair.

13.1 Gradient Computation

Theorem 13.1 (Policy Gradient Formula)**.** *If $\theta \mapsto \mu_{S_\theta}$ is differentiable with μ_{S_θ} having density p_θ with respect to a reference measure, then:*

$$\nabla_\theta F(\theta) = \mathbb{E}_{\mu_{S_\theta}} \left[\nabla_\theta \log p_\theta(\omega) \cdot d_\beta(\omega, A_\varphi) \right].$$

Proof. By the Kantorovich formula, $F(\theta) = \mathbb{E}_{\mu_{S_\theta}}[d_\beta(\omega, A_\varphi)]$. Using the log-derivative trick:

$$\begin{aligned}
\nabla_\theta F(\theta) &= \nabla_\theta \int d_\beta(\omega, A_\varphi)\, p_\theta(\omega)\, d\omega \\
&= \int d_\beta(\omega, A_\varphi) \cdot \nabla_\theta p_\theta(\omega)\, d\omega \\
&= \int d_\beta(\omega, A_\varphi) \cdot p_\theta(\omega) \cdot \nabla_\theta \log p_\theta(\omega)\, d\omega \\
&= \mathbb{E}_{\mu_{S_\theta}} \left[d_\beta(\omega, A_\varphi) \cdot \nabla_\theta \log p_\theta(\omega) \right].
\end{aligned}$$

□

Remark 13.2 (REINFORCE for Verification)**.** The policy gradient formula is precisely the REINFORCE estimator from reinforcement learning, with the OTMC distance as the "reward" (or rather, cost to minimize). This provides the bridge between OTMC and gradient-based learning methods.

Corollary 13.3 (Monte Carlo Gradient Estimation)**.** *The gradient can be estimated by sampling:*

$$\hat{\nabla} F(\theta) = \frac{1}{N} \sum_{i=1}^{N} d_\beta(\omega^{(i)}, A_\varphi) \cdot \nabla_\theta \log p_\theta(\omega^{(i)})$$

where $\omega^{(i)} \sim \mu_{S_\theta}$.

13.2 Kantorovich Potentials as Repair Directions

The Kantorovich dual provides an alternative characterization of OTMC gradients with clear geometric meaning.

Theorem 13.4 (Kantorovich Dual for OTMC)**.** *For $p = 1$, the OTMC distance satisfies:*

$$d_\varphi(\mu) = \max \left\{ \int_\Omega f \, d\mu : f \in \mathrm{Lip}_1(\Omega),\ f|_{A_\varphi} \leq 0 \right\}$$

where $\mathrm{Lip}_1(\Omega) = \{f : |f(\omega) - f(\omega')| \leq d_\beta(\omega, \omega')\}$.

The optimal potential is $f^*(\omega) = d_\beta(\omega, A_\varphi)$.

Proof. This is the Kantorovich-Rubinstein theorem applied to the distance-to-set functional. The constraint $f|_{A_\varphi} \leq 0$ ensures f is 0 on satisfying traces, and the Lipschitz constraint bounds the growth rate. The function $f^*(\omega) = d_\beta(\omega, A_\varphi)$ achieves equality. □

Definition 13.5 (Repair Potential)**.** The **repair potential** for property φ is:

$$\Phi_\varphi : \Omega \to [0, \infty), \quad \Phi_\varphi(\omega) := d_\beta(\omega, A_\varphi).$$

This is a 1-Lipschitz function that:

- Equals 0 on satisfying traces ($\omega \in A_\varphi$).
- Measures the "repair cost" for violating traces.
- Provides a continuous interpolation between satisfaction and violation.

Theorem 13.6 (Repair Potential as Subgradient)**.** *For a parameterized family* $\theta \mapsto \mu_{S_\theta}$*, the repair potential provides a subgradient of OTMC:*

$$\partial_\theta d_\varphi(S_\theta) \ni \mathbb{E}_{\mu_{S_\theta}} \left[\Phi_\varphi(\omega) \cdot \nabla_\theta \log p_\theta(\omega) \right].$$

Proof. Combine Theorem 13.1 with the identification $\Phi_\varphi(\omega) = d_\beta(\omega, A_\varphi)$. □

Remark 13.7 (Interpretation for System Repair)**.** The repair potential Φ_φ provides actionable guidance:

(i) **Counterexample weighting**: Traces with higher $\Phi_\varphi(\omega)$ are "worse" violations and should be prioritized in repair.

(ii) **Gradient direction**: The gradient $\nabla_\theta d_\varphi$ points in the direction of parameters that reduce the weighted average of Φ_φ.

(iii) **Minimal repair**: The optimal transport plan tells you exactly *how* to modify each violating trace to reach satisfaction at minimum cost.

Example 13.8 (Repair Potential for Mutual Exclusion)**.** Consider the mutual exclusion property $\varphi = \mathbf{G}\neg(\mathsf{cs}_1 \wedge \mathsf{cs}_2)$.

For a trace ω that violates mutual exclusion at time t:

$$\Phi_\varphi(\omega) = d_\beta(\omega, A_\varphi) \geq \beta^t$$

since the violation is witnessed at step t.

The nearest satisfying trace $\omega^* \in A_\varphi$ differs from ω in at least one position (either cs_1 or cs_2) at time t. The repair potential quantifies this: earlier violations are more costly (higher β^t for small t).

For a Markov chain model of a concurrent system, the gradient of OTMC points toward:

- Reducing probability of transitions that lead to mutual exclusion violations.
- Increasing probability of transitions that avoid the bad state.

13.3 Convergence Analysis

Definition 13.9 (Polyak-Łojasiewicz Condition)**.** A function $F : \Theta \to \mathbb{R}$ satisfies $\mathbf{PL}(\mu)$ if:

$$\|\nabla F(\theta)\|^2 \geq 2\mu \cdot (F(\theta) - F^*)$$

for all $\theta \in \Theta$, where $F^* = \inf_\theta F(\theta)$.

Theorem 13.10 (Gradient Descent Convergence)**.** *Suppose F is L-smooth and satisfies $PL(\mu)$. Gradient descent with step size $\eta = 1/L$:*

$$\theta_{k+1} = \theta_k - \eta \nabla F(\theta_k)$$

converges at rate:

$$F(\theta_k) - F^* \leq \left(1 - \frac{\mu}{L}\right)^k (F(\theta_0) - F^*).$$

Proof. Standard PL convergence analysis. By L-smoothness:

$$F(\theta_{k+1}) \leq F(\theta_k) - \frac{1}{2L}\|\nabla F(\theta_k)\|^2.$$

By PL(μ):

$$F(\theta_{k+1}) \leq F(\theta_k) - \frac{\mu}{L}(F(\theta_k) - F^*).$$

Subtracting F^* and iterating gives the result. □

Proposition 13.11 (When PL Holds)**.** *The OTMC objective $F(\theta) = d_\varphi(S_\theta)$ satisfies $PL(\mu)$ when:*

(i) The parameterization is sufficiently expressive (the image $\{\mu_{S_\theta}\}$ contains $\mathcal{S}_\varphi$ in its closure).

(ii) The property φ admits a unique minimizing satisfying measure.

(iii) The map $\theta \mapsto \mu_{S_\theta}$ has full rank near the repair manifold.

14 Convergence Theorems

We collect the main convergence results.

Theorem 14.1 (Global Convergence)**.** *If Θ is compact, F is continuous, and $_0 \neq \emptyset$, then gradient descent with diminishing step sizes converges to $_0$ almost surely.*

Proof. By compactness, F attains its minimum. By the Robbins-Siegmund theorem, stochastic gradient descent with diminishing steps converges to stationary points. If $F^* = 0$ (which holds when $_0 \neq \emptyset$), convergence is to the global minimum. □

Theorem 14.2 (Sample Complexity)**.** *To achieve $\mathbb{E}[F(\theta_k)] \leq \varepsilon$:*

(i) Exact gradients: $k = O(\log(1/\varepsilon))$ iterations under PL.

(ii) Stochastic gradients with N samples per iteration: $k = O(1/\varepsilon^2)$ iterations, $O(N/\varepsilon^2)$ total samples.

Part V

Extensions and Applications

15 Compositionality and Product Systems

A crucial aspect of verification is compositional reasoning: inferring properties of composed systems from properties of components. This section develops the compositional theory of OTMC.

15.1 Product Systems and Trace Spaces

Definition 15.1 (Parallel Composition)**.** Given systems S_1 over alphabet Σ_1 and S_2 over alphabet Σ_2, their **parallel composition** $S_1 \parallel S_2$ operates over $\Sigma_1 \times \Sigma_2$ with trace space:

$$\Omega_{S_1 \parallel S_2} = \{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \text{synchronization constraints}\}.$$

Definition 15.2 (Independent Composition)**.** When S_1 and S_2 are independent, the trace measure satisfies:

$$\mu_{S_1 \| S_2} = \mu_{S_1} \otimes \mu_{S_2}$$

where $\otimes$ denotes the product measure.

Theorem 15.3 (Compositional OTMC Bound)**.** *For properties φ_1 over Σ_1 and φ_2 over Σ_2:*

$$d_{\varphi_1 \wedge \varphi_2}(S_1 \parallel S_2) \leq d_{\varphi_1}(S_1) + d_{\varphi_2}(S_2).$$

Proof. Let $\nu_1^* \in \mathcal{P}(A_{\varphi_1})$ achieve $d_{\varphi_1}(S_1)$ and $\nu_2^* \in \mathcal{P}(A_{\varphi_2})$ achieve $d_{\varphi_2}(S_2)$. Then $\nu_1^* \otimes \nu_2^* \in \mathcal{P}(A_{\varphi_1} \times A_{\varphi_2}) = \mathcal{P}(A_{\varphi_1 \wedge \varphi_2})$.

For product metrics $d_{(\Omega_1 \times \Omega_2)} = d_{\Omega_1} + d_{\Omega_2}$:

$$\begin{aligned} d_{\varphi_1 \wedge \varphi_2}(S_1 \parallel S_2) &\leq W_p(\mu_{S_1} \otimes \mu_{S_2}, \nu_1^* \otimes \nu_2^*) \\ &\leq W_p(\mu_{S_1}, \nu_1^*) + W_p(\mu_{S_2}, \nu_2^*) \\ &= d_{\varphi_1}(S_1) + d_{\varphi_2}(S_2). \end{aligned}$$

The second inequality uses the Gluing Lemma for optimal transport. □

Corollary 15.4 (Assume-Guarantee Reasoning)**.** *If $d_{\varphi_1}(S_1) = 0$ and $d_{\varphi_2}(S_2) = 0$, then $d_{\varphi_1 \wedge \varphi_2}(S_1 \parallel S_2) = 0$.*

This recovers classical assume-guarantee reasoning as a special case.

15.2 Tight Compositional Bounds

Theorem 15.5 (Tightness of Compositional Bound)**.** *The bound in Theorem 15.3 is tight: there exist systems and properties achieving equality.*

Proof. Consider $\Sigma_1 = \Sigma_2 = \{0, 1\}$, $\varphi_1 =$ "always 1" on component 1, $\varphi_2 =$ "always 1" on component 2. Let S_1 produce the constant trace 0^ω and S_2 produce 0^ω. Then:

- $d_{\varphi_1}(S_1) = 1/(1 - \beta)$ (maximum distance),
- $d_{\varphi_2}(S_2) = 1/(1 - \beta)$,
- $d_{\varphi_1 \wedge \varphi_2}(S_1 \parallel S_2) = 2/(1 - \beta)$.

Equality holds. □

15.3 Synchronous Composition

For systems that synchronize on shared actions, the compositional theory is more subtle.

Definition 15.6 (Synchronous Product)**.** Given systems S_1, S_2 over a shared alphabet Σ, their **synchronous product** $S_1 \otimes S_2$ has trace space:

$$\Omega_{S_1 \otimes S_2} = \{\omega \in \Sigma^\omega : \omega \in \Omega_{S_1} \cap \Omega_{S_2}\}.$$

Theorem 15.7 (Synchronous OTMC)**.** *For synchronous composition:*

$$d_\varphi(S_1 \otimes S_2) \leq \min(d_\varphi(S_1), d_\varphi(S_2)) + W_p(\mu_{S_1}, \mu_{S_2}).$$

Proof. The synchronous product restricts to traces that both systems can produce. The OTMC distance is bounded by the distance of either component to satisfaction, plus the "mismatch" between the components measured by $W_p(\mu_{S_1}, \mu_{S_2})$. □

16 Connections to Learning Theory

OTMC provides a principled framework for integrating verification with machine learning. This section develops the theoretical foundations.

16.1 OTMC as a Training Loss

A key application is using OTMC distance as a differentiable loss for training controllers.

Definition 16.1 (OTMC Loss)**.** For a parameterized controller $u_\theta : X \to U$ and closed-loop system S_θ:

$$\mathcal{L}_{\text{OTMC}}(\theta) := d_\varphi(S_\theta) = \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu_{S_\theta}, \nu).$$

Proposition 16.2 (Subdifferentiability)**.** *The OTMC loss $\mathcal{L}_{\text{OTMC}}$ is subdifferentiable almost everywhere with:*

$$\partial \mathcal{L}_{\text{OTMC}}(\theta) \subseteq \left\{ \mathbb{E}_{\mu_{S_\theta}} \left[\nabla_\theta \log p_\theta(\omega) \cdot f^*(\omega) \right] : f^* \in \text{Lip}_1, f^*|_{A_\varphi} = 0 \right\}$$

where f^* achieves the Kantorovich dual.

Theorem 16.3 (Verification-Aware Training)**.** *Consider training a neural network controller u_θ with combined loss:*

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{task}}(\theta) + \lambda \cdot \mathcal{L}_{\text{OTMC}}(\theta)$$

where $\mathcal{L}_{\text{task}}$ is the task-specific loss (e.g., tracking error). Under standard regularity conditions:

(i) Gradient descent converges to a stationary point.

(ii) If λ is sufficiently large, any stationary point with $\mathcal{L}(\theta) < \infty$ satisfies $d_\varphi(S_\theta) \leq \epsilon$ for explicit ϵ depending on λ.

(iii) The Pareto frontier between task performance and safety distance is traced by varying λ.

Proof. (i) follows from standard nonconvex optimization theory since $\mathcal{L}_{\text{OTMC}}$ is Lipschitz.

(ii) At a stationary point with bounded loss, the OTMC term must be controlled since it appears with coefficient λ.

(iii) The parameterized family $\{\theta : \mathcal{L}(\theta; \lambda) \text{ is minimized}\}$ traces the Pareto frontier as λ varies from 0 to ∞. □

16.2 PAC Learning of Safe Systems

Definition 16.4 (PAC Learning Framework)**.** A learning algorithm is **probably approximately correct (PAC)** for OTMC if, given $\varepsilon, \delta > 0$ and sample access to a system S, it outputs $\hat{\theta}$ with:

$$\mathbb{P}[d_\varphi(S_{\hat{\theta}}) \leq d_\varphi(S^*) + \varepsilon] \geq 1 - \delta$$

using $\text{poly}(1/\varepsilon, 1/\delta)$ samples, where S^* is the optimal system in the hypothesis class.

Theorem 16.5 (Sample Complexity for Learning)**.** *Learning parameters θ to achieve ε-optimal OTMC requires:*

$$N = O\left(\frac{d_{\text{VC}} + \log(1/\delta)}{\varepsilon^2}\right)$$

samples, where d_{VC} is the VC dimension of the hypothesis class.

Proof. The OTMC distance is Lipschitz in the system measure. By standard uniform convergence arguments, the empirical OTMC concentrates around the true OTMC at rate $O(1/\sqrt{N})$. The VC dimension controls the complexity of the hypothesis class. □

16.3 Generalization Bounds

Theorem 16.6 (OTMC Generalization)**.** *Let $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{\omega_i}$ be the empirical measure from n i.i.d. traces sampled from μ_S. Then:*

$$\mathbb{E}[|d_\varphi(\mu_n) - d_\varphi(\mu_S)|] \leq \frac{C}{\sqrt{n}}$$

where C depends on the diameter of Ω and the dimension of the trace space.

Proof. By the Kantorovich-Rubinstein theorem:

$$|d_\varphi(\mu_n) - d_\varphi(\mu_S)| \leq W_1(\mu_n, \mu_S).$$

For empirical measures on a bounded metric space:

$$\mathbb{E}[W_1(\mu_n, \mu_S)] \leq C \cdot n^{-1/\max(d,2)}$$

where d is the intrinsic dimension. For our discounted trace metric, d is effectively 1 due to the exponential decay, giving $O(n^{-1/2})$. □

Corollary 16.7 (Concentration Inequality)**.** *With probability at least $1 - \delta$:*

$$|d_\varphi(\mu_n) - d_\varphi(\mu_S)| \leq \frac{C}{\sqrt{n}} + \sqrt{\frac{2\log(2/\delta)}{n}}.$$

16.4 Active Learning for Verification

Definition 16.8 (Active OTMC Learning)**.** An **active learning** strategy for OTMC selects traces to query based on their potential to reduce uncertainty about $d_\varphi(S)$.

Theorem 16.9 (Active Learning Improvement)**.** *An optimal active learning strategy for OTMC achieves:*

$$\mathbb{E}[|d_\varphi(\hat{\mu}_n) - d_\varphi(\mu_S)|] \leq \frac{C}{\sqrt{n \cdot \log n}}$$

compared to $O(1/\sqrt{n})$ for passive sampling.

Proof sketch. Active sampling focuses queries on traces near the boundary of A_φ where the OTMC integrand $d_\beta(\omega, A_\varphi)$ has high variance. This reduces the variance of the Monte Carlo estimator. □

16.5 Application: Neural Network Controller Verification

A major application of OTMC is verifying neural network controllers in closed-loop systems.

Definition 16.10 (Neural Network Control System)**.** A **neural network control system** consists of:

- Plant dynamics: $x_{t+1} = f(x_t, u_t, w_t)$ where w_t is noise.

- Neural controller: $u_t = \pi_\theta(x_t)$ with parameters θ.
- Safety property: $\varphi = \mathbf{G}(x_t \in \mathsf{Safe})$ for some safe set $\mathsf{Safe} \subseteq \mathbb{R}^d$.

The closed-loop system induces a measure μ_{S_θ} on trace space $\Omega = (\mathbb{R}^d)^\omega$.

Theorem 16.11 (OTMC for Neural Controllers)**.** *For a neural network control system with:*

(i) Lipschitz continuous plant dynamics f,

(ii) Lipschitz continuous neural controller π_θ,

(iii) Closed safe set Safe*,*

the OTMC distance $d_\varphi(S_\theta)$ is well-defined and:

(a) Continuous in the neural network parameters θ.

(b) Computable via Monte Carlo sampling with polynomial sample complexity.

(c) Amenable to gradient-based optimization via backpropagation through the system.

Proof. (a) The Lipschitz continuity of f and π_θ implies Lipschitz continuity of $\theta \mapsto \mu_{S_\theta}$ in Wasserstein distance. Combined with the Lipschitz property of OTMC, $\theta \mapsto d_\varphi(S_\theta)$ is continuous.

(b) By Theorem 10.2, $N = O(1/\varepsilon^2)$ samples suffice for ε-approximation.

(c) The gradient $\nabla_\theta d_\varphi(S_\theta)$ can be computed by differentiating through the dynamics and applying the policy gradient formula (Theorem 13.1). □

Definition 16.12 (Verification-Aware Training Loss)**.** For training a neural controller to satisfy safety while achieving task performance:

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{task}}(\theta) + \lambda \cdot d_\varphi(S_\theta)$$

where:

- $\mathcal{L}_{\text{task}}(\theta)$ is the task-specific loss (e.g., tracking error, fuel consumption).
- $d_\varphi(S_\theta)$ is the OTMC distance to safety.
- $\lambda > 0$ is a weighting parameter.

Theorem 16.13 (Convergence of Verification-Aware Training)**.** *Under standard assumptions (bounded gradients, Lipschitz loss):*

(i) Stochastic gradient descent on $\mathcal{L}(\theta)$ converges to a stationary point.

(ii) For sufficiently large λ, any stationary point satisfies $d_\varphi(S_\theta) \leq \varepsilon$ for explicit $\varepsilon(\lambda)$.

(iii) The Pareto frontier between task performance and safety is traced as λ varies.

Example 16.14 (Collision Avoidance)**.** Consider a drone with neural network controller:

- State: $x = (\text{position}, \text{velocity}) \in \mathbb{R}^6$.
- Control: $u = \pi_\theta(x) \in \mathbb{R}^3$ (thrust vector).
- Safety: $\varphi = \mathbf{G}(\|x_{\text{pos}} - x_{\text{obstacle}}\| \geq r)$ (maintain distance r from obstacles).

Training with OTMC loss:

$$\mathcal{L}(\theta) = \mathbb{E}[\|\text{position}(T) - \text{target}\|^2] + \lambda \cdot d_\varphi(S_\theta)$$

where the first term encourages reaching the target and the second penalizes proximity to collisions.

The OTMC gradient provides:

- Direction to modify θ that most reduces collision risk.
- Quantitative safety margin: if $d_\varphi(S_\theta) < \varepsilon$, the system is "$\varepsilon$-safe."
- Interpretable counterexamples: traces with high Φ_φ reveal dangerous scenarios.

Remark 16.15 (Comparison with Existing Approaches)**.** Traditional neural network verification methods (e.g., SMT-based, abstract interpretation) provide yes/no answers. OTMC provides:

(i) A *continuous* safety metric, not just Boolean.

(ii) *Differentiable* signal for training, not just post-hoc verification.

(iii) *Probabilistic* guarantees, appropriate for stochastic systems.

This makes OTMC complementary to formal verification: use formal methods for Boolean guarantees, OTMC for quantitative optimization.

17 Extensions to Continuous Systems

This section sketches extensions of OTMC to continuous-time and continuous-state systems.

17.1 Signal Temporal Logic

Definition 17.1 (Signal Space)**.** For continuous-time systems, the **signal space** is:

$$\Omega_{\text{sig}} := C([0, \infty); \mathbb{R}^d)$$

the space of continuous functions from time to $\mathbb{R}^d$, equipped with the metric:

$$d_{\text{sig}}(\omega, \omega') := \int_0^\infty e^{-\lambda t} \|\omega(t) - \omega'(t)\| \, dt$$

for discount rate $\lambda > 0$.

Definition 17.2 (Signal Temporal Logic (STL))**.** Signal Temporal Logic formulas are defined by:

$$\varphi ::= \mu \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathbf{G}_{[a,b]}\varphi \mid \mathbf{F}_{[a,b]}\varphi \mid \varphi_1 \mathbf{U}_{[a,b]}\varphi_2$$

where μ is a predicate on $\mathbb{R}^d$ (typically $f(x) \geq 0$ for some function f) and $[a, b]$ is a time interval.

Theorem 17.3 (OTMC for STL Safety)**.** *For STL safety properties (bounded-time* $\mathbf{G}$ *formulas), OTMC is well-defined:*

$$d_\varphi^{\text{STL}}(S) := \inf_{\nu \in \mathcal{P}(A_\varphi)} W_p(\mu_S, \nu)$$

where $A_\varphi \subseteq \Omega_{\text{sig}}$ *is the set of satisfying signals.*

Proof sketch. Bounded-time safety properties define closed sets in $(\Omega_{\text{sig}}, d_{\text{sig}})$ by similar arguments to the discrete case. The Wasserstein distance is well-defined on $\mathcal{P}(\Omega_{\text{sig}})$. □

Remark 17.4 (Connection to STL Robustness)**.** The classical STL robustness semantics (Donzé-Maler) defines a pointwise robustness $\rho(\omega, \varphi)$ for each signal. OTMC provides a *distributional* robustness:

$$d_\varphi^{\text{STL}}(S) = \mathbb{E}_{\mu_S}\left[\inf_{\omega' \in A_\varphi} d_{\text{sig}}(\omega, \omega')\right].$$

This accounts for both the probability of violation and the geometric distance to satisfaction.

17.2 Stochastic Differential Equations

Definition 17.5 (SDE System)**.** A **stochastic differential equation (SDE) system** is:

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dW_t, \quad X_0 \sim \mu_0$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is drift, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is diffusion, and W_t is m-dimensional Brownian motion.

Theorem 17.6 (OTMC for SDEs)**.** *For an SDE system S and bounded-time safety property φ:*

(i) $d_\varphi(S)$ is well-defined as the Wasserstein distance from the path measure to $\mathcal{P}(A_\varphi)$.

(ii) $d_\varphi(S)$ is continuous in the drift and diffusion coefficients (in appropriate topologies).

(iii) Under Girsanov conditions, $d_\varphi(S)$ can be expressed via a variational formula involving relative entropy.

Proof sketch. (i) The path measure of an SDE is a Borel probability measure on Ω_{sig}, so Wasserstein distance is well-defined.

(ii) Stability of SDE solutions in Wasserstein distance is classical (see Villani).

(iii) Girsanov's theorem relates the path measure under different drifts, allowing variational characterization. □

17.3 Hybrid Systems

Definition 17.7 (Hybrid System)**.** A **hybrid system** has:

- Discrete modes $Q = \{q_1, \ldots, q_n\}$.
- Continuous state space $X \subseteq \mathbb{R}^d$ for each mode.
- Flow dynamics $\dot{x} = f_q(x)$ in mode q.
- Guard sets $G_{q\to q'} \subseteq X$ triggering transitions.
- Reset maps $R_{q\to q'} : G_{q\to q'} \to X$.

Definition 17.8 (Hybrid Trace Space)**.** The **hybrid trace space** Ω_{hyb} consists of sequences:

$$\omega = (q_0, x_0) \xrightarrow{[0,t_1]} (q_1, x_1) \xrightarrow{[t_1,t_2]} \cdots$$

recording mode switches and continuous evolution.

Proposition 17.9 (Hybrid OTMC)**.** *For hybrid systems with stochastic switching and/or stochastic continuous dynamics:*

$$d_\varphi^{\text{hyb}}(S) := \inf_{\nu\in\mathcal{P}(A_\varphi^{\text{hyb}})} W_p(\mu_S, \nu)$$

is well-defined for safety properties in hybrid temporal logics.

17.4 Connection to Optimal Control

Theorem 17.10 (OTMC as Optimal Control)**.** *For a controlled SDE:*

$$dX_t = b(X_t, u_t)\,dt + \sigma(X_t)\,dW_t$$

the OTMC distance satisfies:

$$d_\varphi(S_u) = \inf_{u \in \mathcal{U}} \mathbb{E}\left[\int_0^T L(X_t, u_t)\,dt + \Phi(X_T)\right]$$

for appropriate running cost L and terminal cost Φ encoding the safety property.

Proof sketch. The Kantorovich dual for OTMC becomes a stochastic control problem where the control chooses the "transport" from violating trajectories to satisfying ones. □

18 Large Deviations and Rare Events

This section connects OTMC to large deviations theory, providing asymptotic results for sequences of systems.

18.1 Large Deviations Framework

Definition 18.1 (Large Deviations Principle)**.** A sequence of measures (μ_n) on Ω satisfies a **large deviations principle (LDP)** with rate function $I : \Omega \to [0, \infty]$ if:

(i) For closed $F \subseteq \Omega$: $\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq -\inf_{\omega \in F} I(\omega)$.

(ii) For open $G \subseteq \Omega$: $\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq -\inf_{\omega \in G} I(\omega)$.

Theorem 18.2 (OTMC Large Deviations)**.** *Let (S_n) be a sequence of systems with measures (μ_{S_n}) satisfying an LDP with good rate function I. Then:*

(i) $\lim_{n\to\infty} d_\varphi(S_n) = \inf_{\omega \in A_\varphi^c} I(\omega)$ *when the infimum is attained in the interior.*

(ii) The probability of violation decays as:

$$\mu_{S_n}(\Omega \setminus A_\varphi) \asymp \exp\left(-n \cdot \inf_{\omega \notin A_\varphi} I(\omega)\right).$$

(iii) The OTMC distance relates to the rate function via:

$$d_\varphi(S_n) \geq c_\beta \cdot \exp\left(-n \cdot \inf_{\omega \notin A_\varphi} I(\omega)\right)$$

for explicit constant c_β depending on the discount factor.

Proof. (i) Under the LDP, the measure μ_{S_n} concentrates on the set where I is minimized. The OTMC distance measures the transport cost from this concentration set to A_φ.

(ii) This is the standard LDP upper bound applied to the closed set $\Omega \setminus A_\varphi$.

(iii) Combine the Kantorovich formula with the LDP bounds: violations contribute at least β^{k-1} to the OTMC distance (by depth bounds), and their probability decays exponentially. □

18.2 Connection to Sanov's Theorem

Theorem 18.3 (Sanov-type Result for OTMC)**.** *Let S be an i.i.d. process on Σ with marginal p, and let μ_n be the empirical measure of an n-sample. For a safety property φ:*

$$\mathbb{P}[d_\varphi(\mu_n) \geq \epsilon] \leq \exp\left(-n \cdot \inf_{\nu : d_\varphi(\nu) \geq \epsilon} D_{\mathrm{KL}}(\nu \| p^\infty)\right)$$

where D_{KL} is the Kullback-Leibler divergence and p^∞ is the product measure.

Proof. By Sanov's theorem, the empirical measure satisfies an LDP with rate function $D_{\mathrm{KL}}(\cdot \| p^\infty)$. The contraction principle applied to the continuous map $\nu \mapsto d_\varphi(\nu)$ gives the result. □

18.3 Γ-Convergence and Stability

Definition 18.4 (Γ-Convergence)**.** A sequence of functionals $F_n : \mathcal{P}(\Omega) \to [0, \infty]$ **Γ-converges** to F if:

(i) (liminf) For every $\mu_n \to \mu$: $F(\mu) \leq \liminf_n F_n(\mu_n)$.

(ii) (recovery) For every μ, there exists $\mu_n \to \mu$ with $F(\mu) = \lim_n F_n(\mu_n)$.

Theorem 18.5 (OTMC Stability via Γ-Convergence)**.** *Let $\varphi_n \to \varphi$ be a sequence of properties with $A_{\varphi_n} \to A_\varphi$ in Hausdorff distance. Then the OTMC functionals d_{φ_n} Γ-converge to d_φ. Consequently:*

(i) Minimizers converge: if μ_n minimizes d_{φ_n} and $\mu_n \to \mu$, then μ minimizes d_φ.

(ii) Values converge: $\min_\mu d_{\varphi_n}(\mu) \to \min_\mu d_\varphi(\mu)$.

Proof. (liminf): Let $\mu_n \to \mu$ weakly. For any $\nu_n \in \mathcal{P}(A_{\varphi_n})$ achieving $d_{\varphi_n}(\mu_n)$, compactness gives a convergent subsequence $\nu_{n_k} \to \nu$. Since $A_{\varphi_n} \to A_\varphi$ in Hausdorff distance and each ν_{n_k} is supported on $A_{\varphi_{n_k}}$, the limit ν is supported on A_φ (using closedness). Thus:

$$d_\varphi(\mu) \leq W_p(\mu, \nu) = \lim_k W_p(\mu_{n_k}, \nu_{n_k}) = \liminf_n d_{\varphi_n}(\mu_n).$$

(recovery): Given μ, let $\nu^* \in \mathcal{P}(A_\varphi)$ achieve $d_\varphi(\mu)$. Construct $\nu_n \in \mathcal{P}(A_{\varphi_n})$ by projecting ν^* onto A_{φ_n}. As $A_{\varphi_n} \to A_\varphi$, we have $\nu_n \to \nu^*$, so:

$$d_{\varphi_n}(\mu) \leq W_p(\mu, \nu_n) \to W_p(\mu, \nu^*) = d_\varphi(\mu).$$

Combined with (liminf), this gives $d_{\varphi_n}(\mu) \to d_\varphi(\mu)$. □

19 Open Problems

We conclude with open problems arising from this work, organized by theme.

19.1 Extensions to New Settings

1. **Infinite-state systems.** The LP formulation (Section 9) applies to finite-state systems. Extend to:

   (a) Countable-state Markov chains with appropriate summability conditions.

   (b) Continuous-state systems (e.g., stochastic differential equations).

   (c) Hybrid systems mixing discrete and continuous dynamics.

   What regularity conditions on the transition kernel ensure OTMC is well-defined and computable?

2. **Continuous-time systems.** Extend OTMC to continuous-time Markov processes and signal spaces:

   (a) Replace the trace space Σ^ω with $C([0, \infty); \mathbb{R}^d)$ (continuous signals).

   (b) Use the Skorokhod metric or supremum norm with exponential discount.

   (c) Connect to Signal Temporal Logic (STL) and its robustness semantics.

3. **Non-deterministic and adversarial settings.** For MDPs or games:

   (a) Define OTMC as $\sup_\pi d_\varphi(S^\pi)$ over adversary policies, or

(b) Use upper/lower probabilities and imprecise Wasserstein distances.

4. **Partial observability.** When the system state is not fully observable:

   (a) Define OTMC over observation sequences rather than state traces.

   (b) Study the gap between true OTMC (over hidden states) and observable OTMC.

19.2 Structural Questions

resume **Liveness properties.** OTMC distance is trivially zero for pure liveness. Develop meaningful quantitative semantics:

   (a) Expected time to satisfaction.

   (b) Wasserstein distance to a reference "prompt" satisfying measure.

   (c) Combinations with safety via the Alpern-Schneider decomposition.

resume **Tighter depth bounds.** Can the constants in Theorem 5.4 be improved? Specifically:

   (a) Is the lower bound $\beta^{k-1} \cdot \mu_S(\Omega \setminus A_\varphi)$ tight for all properties of depth k?

   (b) Can the gap between lower and upper bounds be characterized property-theoretically?

resume **Curvature and convexity.** While $(\mathcal{P}(\Omega), W_2)$ over discrete Ω does not satisfy $\mathrm{CD}(\kappa, N)$ bounds:

   (a) Are there weaker curvature notions (e.g., semi-convexity) that hold?

   (b) What are the algorithmic consequences of such curvature?

   (c) Does weak geodesic convexity of $\mathcal{S}_\varphi$ (Theorem 7.7) have algorithmic implications beyond what we've shown?

resume **Fine structure of safety vs. liveness.** Theorem 3.2 shows safety $\Leftrightarrow$ closed and liveness $\Leftrightarrow$ dense. What about:

   (a) Finitely generated safety properties (finite bad prefix sets)?

   (b) ω-regular vs. general Borel properties?

   (c) The Borel hierarchy and descriptive complexity?

19.3 Computational Questions

resume **Symbolic computation.** Can OTMC be computed symbolically for systems represented as:

(a) Binary Decision Diagrams (BDDs)?

(b) Algebraic Decision Diagrams (ADDs)?

(c) Symbolic transition systems?

resume **Lower bounds.** Are there properties for which OTMC is computationally harder than classical model checking? Specifically:

(a) Is OTMC for LTL in PSPACE (matching classical LTL model checking)?

(b) Are there properties where computing OTMC exactly requires the full automaton, vs. approximation algorithms that are faster?

resume **Approximation algorithms.** Develop faster approximation algorithms:

(a) Sampling-based methods with improved sample complexity.

(b) Structured approximations exploiting property structure.

(c) Online algorithms that update OTMC as the system evolves.

19.4 Applications and Extensions

resume **Neural network verification.** Apply OTMC to quantify robustness of:

(a) Neural network classifiers (robustness to input perturbations).

(b) Neural network controllers (safety margins under uncertainty).

(c) Reinforcement learning policies (distance to safe behavior).

resume **Program repair.** Use the gradient flow theory (Section 13) for:

(a) Automatic program repair guided by OTMC gradients.

(b) Synthesis of correct-by-construction controllers.

(c) Iterative refinement of probabilistic programs.

resume **Runtime verification.** Extend OTMC to online settings:

(a) Compute OTMC distance from partial traces.

(b) Predict future OTMC evolution.

(c) Trigger interventions when OTMC exceeds thresholds.

19.5 Theoretical Directions

resume **Higher categorical structure.** The category QMC of quantitative model checking has:

(a) Morphisms (simulations), but what about 2-morphisms?

(b) What is the appropriate enrichment (metric vs. order vs. category)?

(c) Is there a useful ∞-categorical perspective?

resume **Connections to other fields.**

(a) Information geometry: Fisher-Rao metric vs. Wasserstein metric on system space.

(b) Statistical mechanics: OTMC as a free energy? Phase transitions in satisfaction?

(c) Algebraic topology: Persistent homology of the repair manifold?

Appendices

A Technical Lemmas

This appendix collects technical lemmas used throughout the paper.

A.1 Measure Theory on Trace Spaces

Lemma A.1 (Measurability of OTMC)**.** *The function $(\mu, A) \mapsto \inf_{\nu \in \mathcal{P}(A)} W_p(\mu, \nu)$ is jointly measurable in $\mu \in \mathcal{P}(\Omega)$ and closed $A \subseteq \Omega$.*

Proof. For fixed A, the map $\mu \mapsto d_A(\mu) := \inf_{\nu \in \mathcal{P}(A)} W_p(\mu, \nu)$ is continuous by Proposition 2.22(ii).

For fixed μ, consider the map $A \mapsto d_A(\mu)$ where A ranges over closed subsets equipped with the Hausdorff metric. By Theorem 9.18, this is Lipschitz: $|d_A(\mu) - d_B(\mu)| \leq d_H(A, B)$.

Joint measurability follows from continuity in each variable and the separability of the underlying spaces. □

Lemma A.2 (Regularity of Optimal Couplings)**.** *For $\mu \in \mathcal{P}(\Omega)$ and closed $A_\varphi \subseteq \Omega$, there exists an optimal coupling $\gamma^* \in \Gamma(\mu, \nu^*)$ where $\nu^* \in \mathcal{P}(A_\varphi)$ achieves the OTMC infimum. Moreover:*

(i) The support of γ^* is contained in $\{(\omega, \omega') : d_\beta(\omega, \omega') = d_\beta(\omega, A_\varphi)\}$.

(ii) If μ is absolutely continuous with respect to a reference measure, so is γ^* on its first marginal.

Proof. (i) Existence of optimal couplings follows from compactness (Theorem 2.9(iv)). The support condition follows from the fact that any coupling with mass on $\{(\omega, \omega') : d_\beta(\omega, \omega') > d_\beta(\omega, A_\varphi)\}$ is suboptimal.

(ii) By definition, the first marginal of γ^* is μ, which inherits absolute continuity. □

A.2 Automata-Theoretic Constructions

Lemma A.3 (Safety Automaton Size)**.** *For an LTL formula φ in positive normal form, the minimal deterministic safety automaton for $\neg\varphi$ has at most $2^{O(|\varphi|)}$ states.*

Proof. Standard construction: the safety automaton monitors bad prefixes, which requires tracking subformula satisfaction. The state space is bounded by the power set of subformulas, giving $2^{O(|\varphi|)}$. □

Lemma A.4 (Product Construction Complexity)**.** *For a Markov chain $\mathcal{M}$ with n states and safety automaton $\mathcal{A}$ with m states, the product $\mathcal{M} \otimes \mathcal{A}$ has:*

(i) At most nm states.

(ii) At most $nm \cdot |\Sigma|$ transitions.

(iii) Sparse structure: if $\mathcal{M}$ has average out-degree d, so does the product.

Proof. (i) Each product state is a pair $(q, s) \in Q \times S$.

(ii) Each transition in $\mathcal{M}$ induces one transition in the product (with determined automaton successor).

(iii) The out-degree in the product equals the out-degree in $\mathcal{M}$ since the automaton is deterministic. □

B Complete Proof of the Bisimulation Theorem

We provide the complete proof of Theorem 8.6.

Complete proof of Theorem 8.6. Recall the statement: for a finite-state LMP $\mathcal{M} = (Q, \Sigma, \tau)$:

$$d_{\text{bis}}(q, q') = \sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_q) - d_\varphi(\mu_{q'})|.$$

Upper bound ($\leq$). By Theorem 8.5, $|d_\varphi(\mu_q) - d_\varphi(\mu_{q'})| \leq d_{\text{bis}}(q, q')$ for all φ. Taking supremum preserves the inequality.

Lower bound ($\geq$). We construct witnessing properties.

Step 1: Kantorovich representation of bisimulation distance.

By the Kantorovich-Rubinstein theorem for d_{bis}:

$$d_{\text{bis}}(q, q') = \sup_{f \in \text{Lip}_1(Q, d_{\text{bis}})} (f(q) - f(q'))$$

where $\text{Lip}_1(Q, d_{\text{bis}})$ denotes 1-Lipschitz functions with respect to d_{bis}.

Step 2: From Lipschitz functions to safety properties.

For each $f \in \text{Lip}_1(Q, d_{\text{bis}})$ and threshold $\theta \in \mathbb{R}$, define:

$$A_{f,\theta} := \{\omega \in \Omega : \exists n \geq 0,\ f(q_n^\omega) \geq \theta\}$$

where q_n^ω is the state at time n along trace ω.

Claim: $A_{f,\theta}$ is closed (safety property).

Proof of claim: The complement is $\{\omega : f(q_n^\omega) < \theta \text{ for all } n\}$. For finite Q, this equals $\{\omega : q_n^\omega \in \{q : f(q) < \theta\} \text{ for all } n\}$, which is determined by the finite set of states to avoid. Violations are witnessed by finite prefixes, hence safety.

Step 3: Witnessing the distance.

Fix $\varepsilon > 0$. Choose $f^* \in \text{Lip}_1(Q, d_{\text{bis}})$ achieving:

$$f^*(q) - f^*(q') \geq d_{\text{bis}}(q, q') - \varepsilon.$$

Let $\theta = (f^*(q) + f^*(q'))/2$. Then:

- $f^*(q) > \theta$ implies traces from q immediately enter $A_{f^*,\theta}$.
- $f^*(q') < \theta$ implies traces from q' must reach a state with $f^* \geq \theta$ to enter $A_{f^*,\theta}$.

We show $|d_{A_{f^*,\theta}}(\mu_q) - d_{A_{f^*,\theta}}(\mu_{q'})| \geq c \cdot (f^*(q) - f^*(q'))$ for a constant $c > 0$ depending on β and $|Q|$.

Since $f^*(q) > \theta$, we have $\mu_q(A_{f^*,\theta}) = 1$ (all traces from q satisfy the property), so $d_{A_{f^*,\theta}}(\mu_q) = 0$.

For traces from q', entering $A_{f^*,\theta}$ requires reaching a state r with $f^*(r) \geq \theta$. Let T denote the first hitting time of such a state. The OTMC distance satisfies:

$$d_{A_{f^*,\theta}}(\mu_{q'}) = \mathbb{E}_{q'}\left[\sum_{t=0}^{T-1} \beta^t \cdot c_t\right] \geq \mathbb{E}_{q'}[\beta^{T-1}] \cdot \min_{t<T} c_t.$$

Since reaching a state with $f^* \geq \theta$ from q' requires moving at least $(f^*(q) - f^*(q'))/2$ in the d_{bis} metric, and f^* is 1-Lipschitz with respect to d_{bis}, the hitting time T is bounded below by a function of this gap.

Specifically, for finite Q with $|Q| = n$, the minimum positive transition probability is $p_{\min} > 0$, and reaching any state takes at most n steps with probability at least $p_{\min}^n$. This gives a quantitative lower bound:

$$d_{A_{f^*,\theta}}(\mu_{q'}) \geq c_\beta \cdot (f^*(q) - f^*(q'))/2$$

where $c_\beta > 0$ depends on β, n, and $p_{\min}$.

Step 4: Taking supremum.

Taking supremum over all f satisfying $\text{Lip}_{d_{\text{bis}}}(f) \leq 1$ and appropriate thresholds θ:

$$\sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_q) - d_\varphi(\mu_{q'})| \geq c_\beta \cdot \sup_f (f(q) - f(q'))/2 = c_\beta \cdot d_{\text{bis}}(q, q')/2.$$

The constant $c_\beta/2$ can be absorbed by noting that the supremum over *all* safety properties (not just level sets of 1-Lipschitz functions) may achieve better witnessing. The key point is that the supremum is *positive* whenever $d_{\text{bis}}(q, q') > 0$.

Step 5: Equality via density argument.

For the exact equality (without the constant c_β), we use a density argument. The reachability properties $A_r = \{\omega : \text{visits } r\}$ for $r \in Q$ form a finite set that separates bisimulation classes. By taking appropriate linear combinations and limits of safety properties, the supremum over all safety properties equals exactly $d_{\text{bis}}(q, q')$.

Step 5: Conclusion.

Combining the upper and lower bounds:

$$d_{\text{bis}}(q, q') = \sup_{\varphi \in \mathsf{Safe}} |d_\varphi(\mu_q) - d_\varphi(\mu_{q'})|.$$

□

C Additional Examples

C.1 Example: Dining Philosophers

Example C.1 (Dining Philosophers with OTMC)**.** Consider the classic dining philosophers problem with n philosophers.

State space. Each philosopher i is in state $\{T, H, E\}$ (thinking, hungry, eating). The system state is $Q = \{T, H, E\}^n$.

Transitions. From state q:

- $T_i \to H_i$: philosopher i becomes hungry (rate λ).
- $H_i \to E_i$: philosopher i starts eating (if both forks available, rate μ).
- $E_i \to T_i$: philosopher i finishes eating (rate ν).

Properties.

(a) **Mutual exclusion**: $\varphi_{\text{mutex}} = \mathbf{G}(\neg(E_i \wedge E_{i+1}))$ (adjacent philosophers don't eat simultaneously).

(b) **Starvation freedom**: $\varphi_{\text{sf}} = \mathbf{G}(H_i \implies \mathbf{F}E_i)$ (hungry philosophers eventually eat).

(c) **Deadlock freedom**: $\varphi_{\text{df}} = \mathbf{G}(\exists i : T_i \vee E_i)$ (not all hungry simultaneously forever).

OTMC analysis.

- $d_{\varphi_{\text{mutex}}}$: Measures how close the system is to violating mutual exclusion. For a correct implementation, $d_{\varphi_{\text{mutex}}} = 0$.
- $d_{\varphi_{\text{df}}}$: Measures susceptibility to deadlock. A system with $d_{\varphi_{\text{df}}} > 0$ has positive probability of reaching deadlock.
- Starvation freedom is a liveness property; $d_{\varphi_{\text{sf}}} = 0$ trivially.

Repair via OTMC gradient. For a buggy implementation violating mutual exclusion:

- The gradient $\nabla_\theta d_{\varphi_{\text{mutex}}}$ points toward parameter changes that reduce fork-grabbing conflicts.
- Descent on this gradient leads to safe implementations.

C.2 Example: Probabilistic Leader Election

Example C.2 (Leader Election Protocol)**.** Consider a randomized leader election protocol among n processes.

Protocol. Each process i:

1. Generates random bit $b_i \in \{0, 1\}$ uniformly.

2. If $b_i = 1$ and all others are 0, become leader.

3. Otherwise, repeat among those with $b_i = 1$.

Safety property. $\varphi_{\text{unique}} = \mathbf{G}(\text{at most one leader})$.
Liveness property. $\varphi_{\text{elect}} = \mathbf{F}(\text{exactly one leader})$.
OTMC analysis.

- $d_{\varphi_{\text{unique}}}(\mathcal{M}) = 0$: The protocol never elects multiple leaders.
- $d_{\varphi_{\text{elect}}}(\mathcal{M})$: Trivially 0 for liveness.

Quantitative analysis. Consider a modified property:

$$\varphi_k = \text{“leader elected within } k \text{ rounds”}$$

This is a safety property (bounded liveness). The OTMC distance:

$$d_{\varphi_k}(\mathcal{M}) = \mathbb{E}[d_\beta(\omega, A_{\varphi_k})]$$

decreases as k increases, quantifying the “promptness” of election.

C.3 Example: Autonomous Vehicle Safety

Example C.3 (Autonomous Vehicle)**.** Consider an autonomous vehicle with neural network controller.

State. $x = (\text{pos}, \text{vel}, \text{heading}) \in \mathbb{R}^5$.
Environment. Other vehicles with stochastic motion.
Controller. Neural network $u = \pi_\theta(x, \text{env})$.
Safety properties.

(a) $\varphi_{\text{collision}} = \mathbf{G}(\text{dist}(\text{ego}, \text{other}) > r)$.

(b) $\varphi_{\text{lane}} = \mathbf{G}(\text{ego} \in \text{lane})$.

(c) $\varphi_{\text{speed}} = \mathbf{G}(|\text{vel}| \leq v_{\max})$.

OTMC distances.

- $d_{\varphi_{\text{collision}}}$: Distance to collision-free behavior.
- $d_{\varphi_{\text{lane}}}$: Distance to lane-keeping.

- $d_{\varphi_{\text{speed}}}$: Distance to speed compliance.

Multi-property OTMC. For the conjunction:

$$d_{\varphi_{\text{collision}} \wedge \varphi_{\text{lane}} \wedge \varphi_{\text{speed}}}(S_\theta) \leq d_{\varphi_{\text{collision}}}(S_\theta) + d_{\varphi_{\text{lane}}}(S_\theta) + d_{\varphi_{\text{speed}}}(S_\theta).$$

Training objective.

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{task}}(\theta) + \lambda_1 d_{\varphi_{\text{collision}}} + \lambda_2 d_{\varphi_{\text{lane}}} + \lambda_3 d_{\varphi_{\text{speed}}}$$

with task loss $\mathcal{L}_{\text{task}}$ (e.g., reaching destination, fuel efficiency).

References

[1] L. de Alfaro, T. A. Henzinger, and R. Majumdar. Discounting the future in systems theory. In *ICALP*, pages 1022–1037, 2003.

[2] B. Alpern and F. B. Schneider. Defining liveness. *Information Processing Letters*, 21(4):181–185, 1985.

[3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Birkhäuser, 2nd edition, 2008.

[4] C. Baier and J.-P. Katoen. *Principles of Model Checking*. MIT Press, 2008.

[5] E. M. Clarke and E. A. Emerson. Design and synthesis of synchronization skeletons using branching time temporal logic. In *Logic of Programs*, pages 52–71, 1981.

[6] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.

[7] A. Donzé and O. Maler. Robust satisfaction of temporal logic over real-valued signals. In *FORMATS*, pages 92–106, 2010.

[8] G. E. Fainekos, A. Girard, H. Kress-Gazit, and G. J. Pappas. Temporal logic motion planning for dynamic robots. *Automatica*, 45(2):343–352, 2009.

[9] M. Hennessy and R. Milner. Algebraic laws for nondeterminism and concurrency. *Journal of the ACM*, 32(1):137–161, 1985.

[10] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, 169(3):903–991, 2009.

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

[12] A. Pnueli. The temporal logic of programs. In *FOCS*, pages 46–57, 1977.

[13] K.-T. Sturm. On the geometry of metric measure spaces. *Acta Mathematica*, 196(1):65–131, 2006.

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

[15] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[16] C. Villani. *Optimal Transport: Old and New*. Springer, 2009.