Distributionally Robust Strategy Synthesis for Switched Stochastic Systems

Gracia, Ibon; Boskos, Dimitris; Laurenti, Luca; Mazo, Manuel

DOI
10.1145/3575870.3587127

Publication date
2023

Document Version
Final published version

Published in
Proceedings of the 26th ACM International Conference on Hybrid Systems, HSCC 2023

Distributionally Robust Strategy Synthesis for Switched Stochastic Systems

Ibon Gracia
ibon.gracia@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

Luca Laurenti
l.laurenti@tudelft.nl
Delft University of Technology
Delft, Netherlands

Dimitris Boskos
d.boskos@tudelft.nl
Delft University of Technology
Delft, Netherlands

Manuel Mazo Jr.
m.mazo@tudelft.nl
Delft University of Technology
Delft, Netherlands

ABSTRACT

We present a novel framework for formal control of uncertain discrete-time switched stochastic systems against probabilistic reach-avoid specifications. In particular, we consider stochastic systems with additive noise, whose distribution lies in an ambiguity set of distributions that are ε–close to a nominal one according to the Wasserstein distance. For this class of systems we derive control synthesis algorithms that are robust against all these distributions and maximize the probability of satisfying a reach-avoid specification, defined as the probability of reaching a goal region while being safe. The framework we present first learns an abstraction of a switched stochastic system as a *robust Markov decision process (robust MDP)* by accounting for both the stochasticity of the system and the uncertainty in the noise distribution. Then, it synthesizes a strategy on the resulting robust MDP that maximizes the probability of satisfying the property and is robust to all uncertainty in the system. This strategy is then refined into a switching strategy for the original stochastic system. By exploiting tools from optimal transport and stochastic programming, we show that synthesizing such a strategy reduces to solving a set of linear programs, thus guaranteeing efficiency. We experimentally validate the efficacy of our framework on various case studies, including both linear and non-linear switched stochastic systems. Our results represent the first formal approach for control synthesis of stochastic systems with uncertain noise distribution.

CCS CONCEPTS

• **Theory of computation** → **Abstraction**; **Logic and verification**; • **Mathematics of computing** → *Stochastic processes*; • **Computer systems organization** → *Robotic autonomy*.

KEYWORDS

Switched stochastic systems, Formal synthesis, Safe autonomy, Uncertain Markov decision processes, Wasserstein distance

ACM Reference Format:
Ibon Gracia, Luca Laurenti, Dimitris Boskos, and Manuel Mazo Jr.. 2023. Distributionally Robust Strategy Synthesis for Switched Stochastic Systems. In *Proceedings of the 26th ACM International Conference on Hybrid Systems: Computation and Control (HSCC '23), May 09–12, 2023, San Antonio, TX, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3575870.3587127

HSCC '23, May 09–12, 2023, San Antonio, TX, USA

ACM ISBN 979-8-4007-0033-0/23/05…$15.00
https://doi.org/10.1145/3575870.3587127

1 INTRODUCTION

Switched stochastic systems are a class of stochastic hybrid systems that are composed by a finite set of modes and a controller that can freely switch between them [7, 40]. Because of their modelling flexibility, switched stochastic systems are currently employed in many real-world applications, including robotics [25] and cyber-physical systems [17]. Many of these applications have two features in common: 1) they are *safety-critical*, hence formal guarantees of correctness are required, 2) the noise characteristics of the system are *uncertain*, as we often have only partial knowledge on the statistical properties of the system due to the use of statistical estimation techniques and *distributional shifts*, i.e., the noise distribution of the system may change [30]. However, existing formal control synthesis and verification methods for switched stochastic systems all assume that the noise distribution is known exactly. This leads to the fundamental research question that we aim to address in this paper: *how can we derive formal guarantees for stochastic systems whose noise distribution is uncertain?*

In this paper we present a formal control framework to synthesize robust strategies for discrete-time switched stochastic systems with uncertain additive noise. In particular, we assume that in each mode the system evolves according to possibly non-linear dynamics and is affected by an additive noise term whose distribution belongs to a Wasserstein ambiguity set, i.e., a set of distributions that are closer than a given $\epsilon > 0$, according to the Wasserstein distance, to a nominal distribution [15, 26]. For instance, such a set could be estimated using data-driven techniques [26]. We consider a finite-time probabilistic reach-avoid specification, defined as a lower bound on the probability that the system reaches a goal region while avoiding bad states. Building on a robust control synthesis framework, we synthesize a strategy that maximizes the probability that the system

satisfies the specification for the worst-case choice of adversarial distributions from the ambiguity set.

Our approach proposes to abstract the original system into a finite-state uncertain Markov decision process (MDP) [27, 37], namely a *robust MDP* [27], whose uncertainty in the transition probabilities also accounts for the distributional ambiguity in the original system. In particular, by relying on recent results from distributional robust optimization [30], we show that dynamic programming for the resulting robust MDP reduces to solving a set of linear programs, thus guaranteeing efficiency. We formally prove the correctness of our framework and test our approach on two case studies including both linear and non-linear systems and for various ambiguity sets. Note that while in this paper we focus on reach-avoid specifications, this is not limiting. In fact, probabilistic reach-avoid specification are the key building block for model-checking algorithms of various temporal logics, such as PCTL [19, 22] or LTL [9, 20]. Consequently, to the best of our knowledge, our results represent the first step to obtain formal methods for stochastic systems with uncertain or partially unknown noise characteristics.

Related Works. Various formal verification and synthesis algorithms have been developed for switched stochastic systems, with approaches including stochastic barrier functions [32] and abstractions to *finite* Markov models [9, 13, 24, 37], including *interval Markov decision processes (IMDPs)*), which are a class of Markov decision processes in which the transition probabilities belong to intervals [16, 21] and admits efficient control synthesis algorithms [9, 22]. However, all of these works assume that both the dynamics and the noise distribution of the system are well known, which is often an unrealistic assumption due to e.g., unmodelled dynamics, distributional shifts, or data-driven components. In order to close this gap recent works have started to employ machine learning algorithms, including neural networks and Gaussian processes, to devise formal control strategies in the case where the dynamics are (partially) unknown or simply too complex to be modelled [1, 20]. Nevertheless, none of these works consider the case when the distribution of the system is uncertain and lies in an ambiguity set.

Ambiguity sets are commonly used in distributionally robust optimization (DRO) problems to represent a set of probability distributions with respect to which the decision-maker wants to be robust [33]. An ambiguity set is defined as a set of probability distributions that are close to a nominal distribution, which represents our approximate knowledge of the uncertainty model. According to the way closeness is quantified, ambiguity sets are typically constructed based on moment constraints [11, 28], statistical divergences [8], and optimal transport discrepancies [4, 5, 15] like the Wasserstein distance. Wasserstein ambiguity sets, such as those considered in this paper, constitute a convenient choice to group ambiguous distributions, especially for data-driven problems. This is justified by the fact that the Wasserstein metric penalizes horizontal dislocations between distributions [31], it provides ambiguity sets that have finite-sample guarantees of containing the true distribution [14], and it enables the formulation of tractable DRO problems [26]. Dynamic aspects of distributional uncertainty with optimal transport ambiguity are studied in [6], which tracks the evolution of Wasserstein ambiguity sets for systems with an unknown state disturbance distribution, and [18], which develops a risk-aware robot control scheme to avoid dynamic obstacles that evolve according to an ambiguous distribution.

While in this work we focus on abstracting our system to a robust MDP, another class of Markov processes that is closely related to our work is distributionally robust Markov decision processes (DR-MDPs) [10, 38, 39], which are MDPs whose transition probabilities depend on some parameters that are uncertain and lie in some ambiguity set. These are substantially different from the robust MDPs considered in this paper because we do not consider any additional probabilistic structure over the ambiguous distributions to signify which uncertainty model is more likely to occur. Planning algorithms against complex specifications for various classes of robust Markov models have been already considered in the literature [22, 27, 29, 37]. However, how to combine these algorithms with tools of optimal transport to abstract and perform formal synthesis of continuous-space dynamical systems affected by noise of uncertain distribution is not considered in these works and represents a key contribution of our work.

2 BASIC NOTATION

Let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Given a set A, we denote by $|A|$ its cardinality. Given $\ell, m \in \mathbb{N}_0$ with $\ell \leq m$, we use the notation $[\ell : m]$ for the set $\{\ell, \ell+1, \dots, m\}$. For a separable metric space X, we denote by $\mathcal{B}(X)$ its Borel σ-algebra and by $\mathcal{D}(X)$ the set of probability distributions on $(X, \mathcal{B}(X))$. When X is discrete and $\gamma \in \mathcal{D}(X)$ we also denote $\gamma(x) := \gamma(\{x\})$ the probability of the event described by the singleton $\{x\}$. Let $c : X \times X \to \mathbb{R}_{\geq 0}$ be a continuous cost function defined over the product space $X \times X$. The optimal transport discrepancy between two probability distributions $p, p' \in \mathcal{D}(X)$ is defined as

$$\mathcal{T}_c(p, p') := \inf_{\pi \in \Pi(p,p')} \int_{X \times X} c(x, y) d\pi(x, y), \tag{1}$$

where $\Pi(p, p')$ is the set of all transport plans between p and p', a.k.a. couplings, i.e., probability distributions $\pi \in \mathcal{D}(X \times X)$, with marginals p and p', respectively. Since the cost c is nonnegative, $\mathcal{T}_c$ provides a discrepancy measure between distributions in $\mathcal{D}(X)$. By continuity of c, there always exists a transport plan π for which the infimum in (1) is attained [35, Theorem 1.3]. Assume that X is equipped with a metric d. Given $s \geq 1$ we denote by $\mathcal{D}_s(X)$ the set of probability distributions on X with finite s-th moment, i.e., $\mathcal{D}_s(X) = \{p \in \mathcal{D}(X) : \int_X d(x, y)^s dp(x) < \infty \text{ for some } y \in X\}$. Then the discrepancy $\mathcal{W}_s := (\mathcal{T}_{d^s})^{\frac{1}{s}}$ is also a metric in the space $\mathcal{D}_s(X)$ coined as the s-Wasserstein distance [35].

3 PROBLEM FORMULATION

We consider a partially-known discrete-time switched stochastic process described as:

$$\mathbf{x}_{k+1} = f_{\mathbf{u}_k}(\mathbf{x}_k) + \mathbf{v}_k, \tag{2}$$

where $k \in \mathbb{N}$, $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{u}_k \in U$, and $U = \{1, ..., m\}$ is a finite set of *modes* or *actions*. For every $u \in U$, $f_u : \mathbb{R}^n \to \mathbb{R}^n$ is a possibly nonlinear continuous function. The noise term $\mathbf{v}_k$ is an independent random variable with a distribution p_v^{true} that is identically distributed at each time step. While the exact distribution is unknown we do assume the following:

Assumption 1. *The distribution p_v^{true} is ε-close (in the s-Wasserstein sense) to a known distribution $\widehat{p}_v \in \mathcal{D}_s(\mathbb{R}^n)$, which we call* nominal, *i.e., $p_v^{true} \in \mathcal{P}_v := \{p \in \mathcal{D}_s(\mathbb{R}^n) : \mathcal{W}_s(p, \widehat{p}_v) \leq \varepsilon\}$, where $\mathcal{W}_s$ is determined by the metric $d(x, y) = \|x - y\|$, where $\|\cdot\|$ is a norm on $\mathbb{R}^n$ that is fixed throughout the paper, and some choice of $s \geq 1$.*

Intuitively, $\mathbf{x}_k$ is a stochastic process driven by an additive noise process $\mathbf{v}_k$, whose distribution is uncertain and is close to a nominal one and our goal is to devise control strategies that are robust to all distributions in $\mathcal{P}_v$. As a consequence, system (2) represents a large class of controlled stochastic systems with additive and uncertain noise. For instance, such a system arises in a data-driven setting, where measure concentration results [14] can be used to build a Wasserstein ambiguity set from data of $\mathbf{v}_k$ with high confidence [26], or in a distributionally robust setting, where one wants to synthesize control strategies that are robust against distributional shifts of the system.

Let $\omega_{\mathbf{x}} = \mathbf{x}_0 \xrightarrow{\mathbf{u}_0} \mathbf{x}_1 \xrightarrow{\mathbf{u}_1} \dots$ be a path (trajectory) of System (2) and denote by $\omega_{\mathbf{x}}(k) = \mathbf{x}_k$ the state of $\omega_{\mathbf{x}}$ at time k. Given a path $\omega_{\mathbf{x}}$, we denote by $\omega_{\mathbf{x}}^k$ the prefix of finite length $k+1$ of $\omega_{\mathbf{x}}$. We also denote by $\Omega_{\mathbf{x}}^{\text{fin}}$ the set of all sample paths with finite length, i.e, the set of prefixes $\omega_{\mathbf{x}}^k = \mathbf{x}_0 \xrightarrow{\mathbf{u}_0} \mathbf{x}_1 \xrightarrow{\mathbf{u}_1} \dots \xrightarrow{\mathbf{u}_{k-1}} \mathbf{x}_k$ for all $k \in \mathbb{N}$. Given a finite path, a *switching strategy* chooses the mode (action) of System (2).

Definition 1 (Switching Strategy). *A switching strategy $\sigma_x : \Omega_{\mathbf{x}}^{\text{fin}} \to U$ is a function that maps each finite path $\omega_{\mathbf{x}}^k \in \Omega_{\mathbf{x}}^{\text{fin}}$ to an action $u \in U$.*

For any $p_v \in \mathcal{P}_v$, $u \in U$, $X \in \mathcal{B}(\mathbb{R}^n)$, and $x \in \mathbb{R}^n$, let

$$T_{p_v}^u(X \mid x) = \int \mathbf{1}_X(f_u(x) + \bar{v}) p_v(\bar{v}) d\bar{v} \tag{3}$$

be the stochastic transition function induced by system (2) with noise fixed to p_v in mode $u \in U$, where $\mathbf{1}_X$ is the indicator function with $\mathbf{1}_X(x) = 1$, if $x \in X$ and $\mathbf{1}_X(x) = 0$, otherwise. From the definition of $T_{p_v}^u(X \mid x)$ it follows that, given a strategy σ_x, a noise distribution p_v, an initial condition x_0, and a time horizon $[0 : K]$, system (2) defines a stochastic process on the canonical space $\Omega = (\mathbb{R}^n)^{K+1}$ with the Borel sigma-algebra $\mathcal{B}(\Omega)$ [3]. In particular, there is a unique probability distribution $P_{p_v}^{x_0,\sigma_x}$ generated by $T_{p_v}^u$ such that for $k \in \{1, ..., K\}$

$$P_{p_v}^{x_0,\sigma_x}[\omega_{\mathbf{x}}^K(0) \in X] = \mathbf{1}_X(x_0),$$
$$P_{p_v}^{x_0,\sigma_x}[\omega_{\mathbf{x}}^K(k) \in X \mid \omega_{\mathbf{x}}^{k-1}] = T_{p_v}^{\sigma_x(\omega_{\mathbf{x}}^{k-1})}(X \mid \omega_{\mathbf{x}}^K(k-1)).$$

3.1 Problem Formulation

In this paper we consider finite-time probabilistic reach-avoid specifications for System (2) regarding the probability that a trajectory of System (2) reaches a goal region, whilst always avoiding a given set of bad states. In particular, for a time horizon $K \in \mathbb{N}_0$, a bounded safe set X, a target region $X_{\text{tgt}} \subset X$ and an initial state $x_0 \in X$, the reach-avoid probability is formally defined as

$$P_{\text{reach}}(X, X_{\text{tgt}}, K \mid x_0, \sigma_x, p_v) := P_{p_v}^{x_0,\sigma_x}[\exists k \in [0 : K] \ s.t. \ \omega_{\mathbf{x}}^k(k) \in X_{\text{tgt}} \wedge \forall\, k' < k \ \omega_{\mathbf{x}}^k(k') \in X]. \tag{4}$$

We are now ready to formally state the problem we consider in this paper.

Problem 1 (Switching Strategy Synthesis). *Consider the switched stochastic system (2), its corresponding ambiguity set $\mathcal{P}_v$, a bounded safe set X, and a target region $X_{tgt} \subset X$. Given an initial state $x_0 \in X$, a probability threshold $p_{th} \in [0, 1]$, and a horizon $K \in \mathbb{N}_0$, synthesize a switching strategy σ_x that allows us to determine if*

$$P_{\text{reach}}(X, X_{\text{tgt}}, K \mid x_0, \sigma_x, p_v) \geq p_{\text{th}}, \tag{5}$$

for all $p_v \in \mathcal{P}_v$.

Note that our focus on reach-avoid specifications in Problem 1 is not limiting: algorithms to compute more complex specifications, such as *linear temporal logic under finite traces* (LTLf), syntactically co-safe linear temporal logic (sc-LTL) or *bounded linear temporal logic* (BLTL), often reduce to reachability computations [1, 2, 9, 19].

Overview of the Approach. To solve Problem 1, in Section 5 we construct a finite-state abstraction of System (2) in terms of a robust MDP. In Section 6 we synthesize an optimal strategy for the resulting abstraction via the solution of a set of linear programs. Finally, we refine this strategy into a strategy for system (2) and derive upper and lower bounds on the probability that the system satisfies the specification under the refined strategy.

4 PRELIMINARIES

4.1 Robust Markov Decision Processes

We abstract system (2) into a *robust Markov decision process (robust MDP)* $\mathcal{M}$. Robust MDPs are a generalization of Markov decision processes in which the transition probability distributions between states are constrained to belong to an ambiguity set [27], [36].

Definition 2 (Robust MDP). *A robust Markov decision process ($\mathcal{M}$) is a tuple $\mathcal{M} = (Q, A, \Gamma)$, where*

- *Q is a finite set of states,*
- *A is a finite set of actions, and $A(q)$ denotes the set of available actions at state $q \in Q$,*
- *$\Gamma = \{\Gamma_{q,a}\}_{q \in Q, a \in A}$ are the sets of possible transition probability distributions of $\mathcal{M}$, namely, $\Gamma_{q,a} \subseteq \mathcal{D}(Q)$.*[1]

A path of a robust MDP is a sequence of states $\omega = q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \xrightarrow{a_2} \dots$ such that $a_k \in A(q_k)$ and there exists $\gamma \in \Gamma_{q_k,a_k}$ with $\gamma(q_{k+1}) > 0$ for all $k \in \mathbb{N}$. We denote the i-th state of a path ω by $\omega(i)$, a finite path of length $k+1$ by ω^k and the last state of a finite path ω^{fin} by $last(\omega^{\text{fin}})$. The set of all finite paths is denoted by $Paths^{\text{fin}}$.

Definition 3 (IMDP). *An interval Markov decision process (IMDP) $\mathcal{I}$ [9], [23], also known as bounded parameter MDP (BMDP) [16], [21], is a class of robust MDP $\mathcal{I} = (Q, A, \Gamma)$ where Γ has the following form:*

$$\Gamma_{q,a} = \{\gamma \in \mathcal{D}(Q) : \underline{P}(q, a, q') \leq \gamma(q') \leq \overline{P}(q, a, q') \text{ for all } q' \in Q\}, \tag{6}$$

for every $q \in Q$, $a \in A(q)$. The bounds $\underline{P}, \overline{P}$ are called transition probability bounds and must fulfill, for every state $q \in Q$ and action $a \in$

[1] Note that the sets of transition probability distributions of the robust MDP are independent for each state and action. This is known as *rectangular* property of the set of transition probability distributions [27], [36].

$A(q)$, that $0 \leq \underline{P}(q,a,q') \leq \overline{P}(q,a,q') \leq 1$ and $\sum_{q' \in Q} \underline{P}(q,a,q') \leq 1 \leq \sum_{q' \in Q} \overline{P}(q,a,q')$.

The actions of robust MDPs and IMDPs are chosen according to a strategy σ which is defined below.

Definition 4 (Strategy). *A strategy σ of a robust MDP model $\mathcal{M}$ is a function $\sigma : Paths^{\text{fin}} \to A$ that maps a finite path ω^k, with $k \in \mathbb{N}$, of $\mathcal{M}$ onto an action in $A(last(\omega^{\text{fin}}))$. If a strategy depends only on $last(\omega^{\text{fin}})$ and k, it is called a memoryless (Markovian) strategy.*

Given an arbitrary strategy σ, we are restricted to the set of robust Markov chains defined by the set of transition probability distributions induced by σ. In order to reduce this to a Markov chain, we define the adversary function [16], also referred to as "nature" [27], which assigns a transition probability distribution to each state-action pair.

Definition 5 (Adversary). *For a robust MDP $\mathcal{M}$, an adversary is a function $\xi : Paths^{\text{fin}} \times A \to \mathcal{D}(Q)$ that, for each finite path $\omega^{\text{fin}} \in Paths^{\text{fin}}$, state $q = last(\omega^{\text{fin}})$, and action $a \in A(q)$, assigns an admissible distribution $\gamma_{q,a} \in \Gamma_{q,a}$. The set of all adversaries is denoted by Ξ.*

For an initial condition $q_0 \in Q$, under a strategy and a valid adversary $\xi \in \Xi$, the robust MDP collapses to a Markov chain and a probability distribution $Prob_{\xi}^{q_0,\sigma}$ is induced on its paths.

5 ROBUST MDP ABSTRACTION

In order to solve Problem 1, we start by abstracting system (2) into the IMDP $\widehat{\mathcal{I}} = (Q, A, \widehat{\Gamma})$ with the noise distribution fixed to the nominal one, $\widehat{p}_v$. In this way, we embed the error caused by the state discretization into $\widehat{\mathcal{I}}$. After that, we expand the set of transition probabilities $\widehat{\Gamma}$ of $\widehat{\mathcal{I}}$ to also capture the distributional ambiguity into the abstraction, obtaining the robust MDP $\mathcal{M} = (Q, A, \Gamma)$. Note that the sets of states Q and actions A are the same in $\widehat{\mathcal{I}}$ and $\mathcal{M}$. Next we describe how we obtain Q and A, and in Section 5.2 we consider the set of transition probability distributions Γ.

5.1 States and Actions

The state-space Q of $\mathcal{M}$ is constructed as follows: consider a set of non-overlapping regions $Q_{\text{safe}} = \{q_1, q_2, \ldots, q_{|Q_{\text{safe}}|}\}$ partitioning the set X so that either $q \cap X_{\text{tgt}} = \emptyset$ or $q \cap (X \setminus X_{\text{tgt}}) = \emptyset$ for all $q \in Q_{\text{safe}}$. We denote by Q_{tgt} the subset of Q_{safe} for which $q \cap X_{\text{tgt}} = q$ and assume that it is a partition of Q_{tgt}. The states of the abstraction comprise of Q_{safe} and the unsafe region $q_u := \mathbb{R}^n \setminus X$, namely, $Q := Q_{\text{safe}} \cup \{q_u\}$. We index Q by $\mathcal{N} = \{1, \ldots, N\}$, where $N := |Q|$ and denote the actions of the abstraction as $A := U$.

5.2 Transition Probability Distributions

Accounting for the Discretization Error. To capture the state discretization error into the abstraction, we first consider an IMDP abstraction of system (2) for a fixed distribution: the nominal probability distribution $\widehat{p}_v$. Since this IMDP is constructed for the nominal distribution $\widehat{p}_v$, we call it "nominal" IMDP, and use the notation $\widehat{\mathcal{I}} = (Q, A, \widehat{\Gamma})$. Note that building an IMDP abstraction of a stochastic system with disturbances of a known distribution has been widely studied in the literature [1, 9, 12, 22] and we report the full procedure here below for completeness. We have already defined the state Q and action A spaces of $\widehat{\mathcal{I}}$ in Section 5.1. We now describe the set $\widehat{\Gamma}$ of $\widehat{\mathcal{I}}$. According to Definition 3, the set $\widehat{\Gamma}$ is defined by the transition probability bounds $\underline{P}$ and $\overline{P}$ of $\widehat{\mathcal{I}}$. To formally account for the discretization error, the bounds must satisfy for all $q \in Q_{\text{safe}}$, $q' \in Q$ and $a \in A = U$:

$$\begin{aligned} \underline{P}(q,a,q') &\leq \min_{x \in q} T^a_{\widehat{p}_v}(q' \mid x) \\ \overline{P}(q,a,q') &\geq \max_{x \in q} T^a_{\widehat{p}_v}(q' \mid x). \end{aligned} \tag{7}$$

Since we are interested in the paths of system (2) that do not exit set X, we make state q_u absorbing, i.e.,

$$\underline{P}(q_u,a,q_u) = \overline{P}(q_u,a,q_u) = 1 \tag{8}$$

for all $a \in A$. In this way we include the "avoid" part of the specification into the definition of the abstraction: the paths that reach q_u, will remain there forever and, therefore, will never reach the target set Q_{tgt}. Consequently, for each $q \in Q$ and $a \in A$ we obtain

$$\widehat{\Gamma}_{q,a} = \{\gamma \in \mathcal{D}(Q) : \underline{P}(q,a,q') \leq \gamma(q') \leq \overline{P}(q,a,q') \text{ for all } q' \in Q\}. \tag{9}$$

Accounting for the Distributional Uncertainty. Now, we expand the sets $\{\widehat{\Gamma}_{q,a}\}_{q \in Q, a \in A}$ of transition probabilities of $\widehat{\mathcal{I}}$ to also embed the distributional uncertainty into the abstraction. With this objective, we first define the following cost between states in Q:

$$c(q,q') := \inf\{\|x - y\|^s : x \in q, y \in q'\}, \tag{10}$$

for $q, q' \in Q$, and where $\|\cdot\|$ and s are the same as for $\mathcal{W}_s$ in Assumption 1. The cost $c(q,q')^{\frac{1}{s}}$ is the minimum distance, in the sense of norm $\|\cdot\|$, between any pair of points in the regions q and q', respectively . Using this cost and the exponent s in $\mathcal{W}_s$, we define the optimal transport discrepancy $\mathcal{T}_c$ between distributions over Q as in (1).[2] Given a probability distribution $\gamma \in \mathcal{D}(Q)$ and $\epsilon \geq 0$, we denote by $\mathcal{T}_c^{\epsilon}(\gamma)$ the set of all distributions to which mass can be transported from γ incurring a c-transport cost lower than ϵ. Using the previous elements, we are finally able to define Γ.

Definition 6. *The discrete uncertainty set Γ is defined for every state $q \in Q_{\text{safe}}$ and action $a \in A$ as:*

$$\Gamma_{q,a} := \bigcup_{\gamma \in \widehat{\Gamma}_{q,a}} \mathcal{T}_c^{\epsilon}(\gamma), \tag{11}$$

with $\epsilon = \varepsilon^s$. For state q_u and action $a \in A$, let $\Gamma_{q,a} := \widehat{\Gamma}_{q,a}$ (preserving the absorbing property of the unsafe state).

Each $\Gamma_{q,a}$ is the set of probability distributions over Q that are ε-close to $\widehat{\Gamma}_{q,a}$, in the sense of the optimal transport discrepancy $\mathcal{T}_c$. Once we have obtained the sets of transition probabilities Γ, along with the state Q and action A spaces, our robust MDP abstraction $\mathcal{M}$ is fully defined. The following proposition ensures that the abstraction captures all possible transition probabilities of system (2) to regions in the partition.

[2] Notice that, since c is also not a metric, the resulting optimal transport discrepancy $\mathcal{T}_c$ is not a distance.

Proposition 1 (Consistency of the Robust MDP Abstraction). *Consider the robust MDP abstraction $\mathcal{M} = (Q, A, \Gamma)$ of system (2). Let $q \in Q_{\text{safe}}$, $a \in A$, $x \in q$, $p_v \in \mathcal{P}_v$ and define $\gamma_{x,a} \in \mathcal{D}(Q)$ such that*

$$\gamma_{x,a}(q') := T^a_{p_v}(q'|x)$$

for all $q' \in Q$. Then $\gamma_{x,a} \in \Gamma_{q,a}$.

The proof of Proposition 1 is given in Appendix 9. The intuition behind Proposition 1 is that set $\Gamma_{q,a}$ contains the transition probabilities $\gamma_{x,a}$ obtained by starting from any $x \in q$, with $q \in Q_{\text{safe}}$ under $a \in A$ and for any $p_v \in \mathcal{P}_v$.

Remark 1 (Model Choice for the Abstraction). *An alternative way to include the distributional ambiguity into the abstraction is to use an IMDP abstraction $\mathcal{I} = (Q, A, \Gamma^{\text{IMDP}})$, which has the same state Q and action A spaces as $\mathcal{M}$, and in which Γ^{IMDP} is defined by transition probability bounds that fulfill:*

$$\begin{aligned} \underline{P}^{\text{IMDP}}(q,a,q') &\leq \min_{p_v \in \mathcal{P}_v} \min_{x \in q} T^a_{p_v}(q' \mid x) \\ \overline{P}^{\text{IMDP}}(q,a,q') &\geq \max_{p_v \in \mathcal{P}_v} \max_{x \in q} T^a_{p_v}(q' \mid x) \end{aligned} \tag{12}$$

for all $q \in Q_{\text{safe}}$, $q' \in Q$ $a \in A$, and for which q_u is again absorbing. Therefore, for fixed $q \in Q$ and $a \in A$, set $\Gamma^{\text{IMDP}}_{q,a}$ of $\mathcal{I}$ is defined as

$$\begin{aligned} \Gamma^{\text{IMDP}}_{q,a} = \{\gamma \in \mathcal{D}(Q) : \underline{P}^{\text{IMDP}}(q,a,q') \leq \gamma(q') \leq \overline{P}^{\text{IMDP}}(q,a,q') \\ \text{for all } q' \in Q\}. \end{aligned} \tag{13}$$

This choice of the abstraction model allows to use efficient synthesis algorithms for IMDPs [16], [21] to solve Problem 1. By the definition of the transition probability bounds of $\mathcal{I}$ in (12), Γ^{IMDP} satisfies Proposition 1, effectively capturing all possible transition probabilities of system (2). However, IMDP $\mathcal{I}$ describes the uncertainty in a very loose way, i.e., set Γ^{IMDP} of $\mathcal{I}$ can be excessively big for many ambiguity sets $\mathcal{P}_v$. Intuitively, this is caused by $\Gamma^{\text{IMDP}}_{q,a}$ being only defined through decoupled interval constraints for every successor state $q' \in Q$, This is likely to result in more conservative solutions to the reach-avoid problem as we will show in Section 7.

6 STRATEGY SYNTHESIS

Our goal is to synthesize a switching strategy $\sigma^*_{\mathbf{x}}$ for system (2) that maximizes (5). To capture the distributional uncertainty and the effect of quantization, we consider the proposed abstraction $\mathcal{M}$. We synthesize the robustly maximizing strategy σ^* for the abstraction $\mathcal{M}$, and refine it, retaining formal guarantees of correctness, when mapped back to the concrete system. In Section 6.1 and 6.2 we show how an optimal strategy σ^* for the abstraction $\mathcal{M}$ can be efficiently computed via linear programming, then in Section 6.3 we prove the correctness of our strategy synthesis approach.

6.1 Robust Dynamic Programming

Once we have obtained a robust MDP abstraction $\mathcal{M} = (Q, A, \Gamma)$, the uncertainties of $\mathcal{M}$ are characterized by an adversary ξ that at each time step, given a path of $\mathcal{M}$ and an action, selects a feasible distribution from Γ (see Definition 5). As a consequence, in order to be robust against all uncertainties, as common in the literature [16, 27], we aim to synthesize a strategy σ^* such that:

$$\sigma^* \in \arg\max_{\sigma \in \Sigma} \min_{\xi \in \Xi} P_{\text{reach}}(Q_{\text{safe}}, Q_{\text{tgt}}, K \mid q, \sigma, \xi), \tag{14}$$

for all $q \in Q$, where $P_{\text{reach}}(Q_{\text{safe}}, Q_{\text{tgt}}, K \mid q, \sigma, \xi)$ is defined as in (4) for system (2).

We denote by $\underline{p}^K$ and $\overline{p}^K$, respectively, the worst and best-case probabilities of the paths of $\mathcal{M}$ satisfying the reach-avoid specification under optimal strategy σ^*:

$$\begin{aligned} \underline{p}^K(q) &:= \min_{\xi \in \Xi} P_{\text{reach}}(Q_{\text{safe}}, Q_{\text{tgt}}, K \mid q, \sigma^*, \xi) \\ \overline{p}^K(q) &:= \max_{\xi \in \Xi} P_{\text{reach}}(Q_{\text{safe}}, Q_{\text{tgt}}, K \mid q, \sigma^*, \xi) \end{aligned} \tag{15}$$

for all $q \in Q$. The following proposition from [22] guarantees that the probabilities in (15) and the optimal strategy in (14) can be obtained through dynamic programming.

Proposition 2. *[22] Let $\underline{p}^K$ be as defined in (15) and $k \in [0 : K-1]$. Then, it holds that*

$$\underline{p}^{k+1}(q) = \begin{cases} 1 & \text{if } q \in Q_{\text{tgt}} \\ \max_{a \in A} \min_{\gamma \in \Gamma_{q,a}} \sum_{q' \in Q} \gamma(q') \underline{p}^k(q') & \text{otherwise,} \end{cases} \tag{16}$$

with initial condition $\underline{p}^0(q) = 1$ for all $q \in Q_{\text{tgt}}$ and 0 otherwise. Furthermore, for each path ω^k with $k \in [0 : K-1]$, it holds that

$$\sigma^*(\omega^k) \in \arg\max_{a \in A} \left\{ \min_{\gamma \in \Gamma_{last(\omega^k),a}} \sum_{q' \in Q} \gamma(q') \underline{p}^{K-k-1}(q') \right\}. \tag{17}$$

A consequence of Proposition 2 is that in our setting there exists an optimal policy that is Markovian and time dependent. Hence, we can restrict our search for σ^* to this class of strategies. Furthermore, once σ^* is fixed, $\overline{p}^{k+1}$, an upper bound of P_{reach}, can be readily computed via the dynamic programming recursion

$$\overline{p}^{k+1}(q) = \begin{cases} 1 & \text{if } q \in Q_{\text{tgt}} \\ \max_{\gamma \in \Gamma_{q,\sigma^*}} \sum_{q' \in Q} \gamma(q') \overline{p}^k(q') & \text{otherwise,} \end{cases} \tag{18}$$

which is analogous to that in (16) and has initial condition $\overline{p}^0(q) = 1$ for all $q \in Q_{\text{tgt}}$ and $\overline{p}^0(q) = 0$ otherwise.

6.2 Computation of Robust Dynamic Programming via Linear Programming

We now show how for each state $q \in Q$ and time $k \in [0 : K-1]$, recursion (16) reduces to solving $|A|$ linear programs. In particular, the following theorem guarantees that the inner problem in recursion (16) can be solved via linear programming. While in the theorem we explicitly consider $\underline{p}^k$, the upper bound $\overline{p}^k$ follows similarly.

Theorem 1 (Robust Dynamic Programming as a Linear Program). *Consider the robust dynamic programming recursion (16) for the robust MDP $\mathcal{M} = (Q, A, \Gamma)$. Then, for any $k \in [0 : K-1]$, $q \in Q$, and $a \in A$ the inner minimization problem in (16) is equivalent to the following linear program:*

$$\min_{\gamma_i, \widehat{\gamma}_j, \pi_{ij}} \sum_{i \in \mathcal{N}} \gamma_i \underline{p}^k(q_i), \tag{19}$$

$$\text{s.t. } \underline{P}(q,a,q_j) \le \widehat{\gamma}_j \le \overline{P}(q,a,q_j) \qquad j \in \mathcal{N} \tag{20a}$$

$$\sum_{j\in\mathcal{N}} \widehat{\gamma}_j = 1 \tag{20b}$$

$$\pi_{ij} \ge 0, \qquad i,j \in \mathcal{N} \tag{20c}$$

$$\sum_{i\in\mathcal{N}} \pi_{ij} = \widehat{\gamma}_j, \qquad j \in \mathcal{N} \tag{20d}$$

$$\sum_{j\in\mathcal{N}} \pi_{ij} = \gamma_i, \qquad i \in \mathcal{N} \tag{20e}$$

$$\sum_{i,j\in\mathcal{N}} \pi_{ij} c(q_i,q_j) \le \varepsilon^s, \tag{20f}$$

where $\underline{P}$, $\overline{P}$, and c are defined in (7), (8), *and* (10), *respectively, and s is given in Assumption 1.*

Proof. We show that, for a fixed state $q \in Q$ and action $a \in A$, the set of transition probabilities $\Gamma_{q,a}$ defined in (11) is the polytope described by the linear equations (20). To this end, let $\gamma \equiv (\gamma_1,\dots,\gamma_N) \in \Gamma_{q,a}$. From the definition of $\Gamma_{q,a}$ in (11), there exist $\widehat{\gamma} \equiv (\widehat{\gamma}_1,\dots,\widehat{\gamma}_N) \in \widehat{\Gamma}_{q,a}$ and an optimal transport plan $\pi \equiv (\pi_{ij})_{i,j=1,\dots,N}$ that transports mass from $\widehat{\gamma}$ to γ with a cost $\mathcal{T}_c(\gamma,\widehat{\gamma})$ smaller than ε^s. Consider now the set $\widehat{\Gamma}_{q,a}$ as defined in (9) with the transition probability bounds $\underline{P}$ and $\overline{P}$ given by (7) and (8). Then $\widehat{\gamma}$ satisfies the constraints (20a) and (20b). Since the optimal transport cost $\mathcal{T}_c(\gamma,\widehat{\gamma})$ is attained by the transport plan π, it follows from (1) with $X \equiv Q$ and c as in (10) that

$$\mathcal{T}_c(\gamma,\widehat{\gamma}) = \sum_{i,j\in\mathcal{N}} \pi_{ij} c(q_i,q_j).$$

Thus, since $\mathcal{T}_c(\gamma,\widehat{\gamma})$ is less than ε^s, we deduce that γ, $\widehat{\gamma}$, and π satisfy the linear constraints (20c)-(20f). Conversely, one can check along the same lines that for any $\gamma \equiv (\gamma_1,\dots,\gamma_N)$, $\widehat{\gamma} \equiv (\widehat{\gamma}_1,\dots,\widehat{\gamma}_N)$, and $\pi \equiv (\pi_{ij})_{i,j=1,\dots,N}$ satisfying (20), it also holds that $\gamma \in \Gamma_{q,a}$. The proof is complete. □

Intuitively, for each state $q \in Q$ and action $a \in A$, the constraints (20a)-(20f) capture the union of the c-transport cost ambiguity balls in $\mathcal{D}(Q)$ that have radius ε^s and centers all possible distributions of the nominal IMDP. Specifically, (20a) and (20b) represent the distributions $\widehat{\gamma}$ of the nominal IMDP, i.e., the set $\widehat{\Gamma}_{q,a}$. The constraints (20c)-(20e) describe a transport plan π, i.e., a nonnegative measure on $Q \times Q$ (cf. (20c)), which has as its marginals the distribution $\widehat{\gamma}$ of the nominal IMDP (cf. (20d)), and the target distribution γ (cf. (20e)), respectively. Finally, (20f) implies that transport cost to reach the target distribution γ is bounded by ε^s, namely, that γ belongs to the c-transport cost ambiguity ball of radius ε^s.

Remark 2. *Theorem 1 guarantees that, similarly to* IMDP*s without distributional uncertainty [16, 22], optimal policies can be computed by solving a set of linear programs. In particular, for any $k \in [0 : K-1]$ and $q \in Q$, we can solve the linear program in Theorem 1 for each $a \in A$ and take the action that maximizes the resulting value function. However, in the IMDP case the resulting linear program has substantially fewer variables and constraints compared to the problem in Theorem 1 (order of N for IMDPs, against order of N^2 for Theorem 1). Nevertheless, as we detail in Appendix A, the number of variables and constraints in our approach can often be substantially reduced when the support of the noise is bounded.*

6.3 Correctness

In this section we prove the correctness of our abstraction for system (2). We begin by refining the strategy σ^* to system (2). Let $J : \mathbb{R}^n \to Q$ map the continuous state $x \in \mathbb{R}^n$ to the corresponding discrete state $q \in Q$ of $\mathcal{M}$, i.e., for any $x \in \mathbb{R}^n$,

$$J(x) = q \iff x \in q. \tag{21}$$

Given a finite path $\omega^k_{\mathbf{x}} = \mathbf{x}_0 \xrightarrow{\mathbf{u}_0} \mathbf{x}_1 \xrightarrow{\mathbf{u}_1} \dots \xrightarrow{\mathbf{u}_{k-1}} \mathbf{x}_k$ of system (2), we define by

$$J(\omega^k_{\mathbf{x}}) = J(\mathbf{x}_0) \xrightarrow{\mathbf{u}_0} J(\mathbf{x}_1) \xrightarrow{\mathbf{u}_1} \dots \xrightarrow{\mathbf{u}_{k-1}} J(\mathbf{x}_k)$$

the corresponding finite path of the MDP abstraction. Consequently, a strategy σ^* for $\mathcal{M}$ is refined to a switched strategy σ^*_x for system (2) such that:

$$\sigma^*_x(\omega^k_{\mathbf{x}}) := \sigma^*(J(\omega^k_{\mathbf{x}})). \tag{22}$$

The following theorem, which is a direct consequence of Theorem 1 and Theorem 2 in [19], ensures that the guarantees obtained for the robust MDP abstraction also hold for system (2).

Theorem 2 (Correctness). *Let $\mathcal{M}$ be a robust MDP abstraction of system* (2), *$\sigma^* \in \Sigma$ be an optimal strategy for $\mathcal{M}$ and $\sigma^*_{\boldsymbol{x}}$ the corresponding switching strategy. Then, for any $q \in Q$, $x \in q$, and $p_v \in \mathcal{P}_v$ it holds that*

$$\overline{p}^K(q) \ge P_{\text{reach}}(X, X_{\text{tgt}}, K \mid x, \sigma^*_{\boldsymbol{x}}, p_v) \ge \underline{p}^K(q),$$

where $\underline{p}^K$ and $\overline{p}^K$ are defined in (15).

Theorem 2 guarantees that in order to solve Problem 1 we can synthesize an optimal strategy σ^* for a robust MDP abstraction of system (2) and then simply check if $\underline{p}^K(J(x))$ is greater than the given threshold.

7 CASE STUDIES

In this section we evaluate our method on two case studies of reach-avoid specifications using the abstractions in this paper, for a linear and a nonlinear system, respectively. We consider data-driven ambiguity sets which are built from i.i.d. samples $\mathbf{v}^1,\dots,\mathbf{v}^M$ of a Gaussian mixture. Using these samples, we construct an ambiguity ball centered on the empirical distribution

$$\widehat{p}_v = \frac{1}{M}\sum_{i=1}^{M} \delta_{\mathbf{v}^i}, \tag{23}$$

of the data [26], where $\delta_{\mathbf{v}^i}$ denotes the Dirac distribution that assigns unit mass to $\mathbf{v}^i$. For the metric of the ball we consider the 2-Wasserstein distance, which penalizes more distributions that have considerable mass far from the samples. This choice is appropriate for our case studies, which consider sufficiently light-tailed distributions. We present results for multiple values M of the sample size and the radius ε. To synthesize the respective strategies, we run the robust dynamic programming algorithm K_{lower} times for the case of $\underline{p}^k$ and K_{upper} times for the case of $\overline{p}^k$. Making use of $\underline{p}^{K_{\text{lower}}}$ and $\overline{p}^{K_{\text{upper}}}$ we define the "average error"

#	Approach	ε	e_{avg}	Synthesis Time (h)
#1	IMDP	0	0	2.8
#2	Robust MDP	5×10^{-3}	0.05	23.2
#3	Robust MDP	10^{-2}	0.18	23.8
#4	Robust MDP	1.5×10^{-2}	0.33	26.2
#5	IMDP	5×10^{-3}	0.83	2.8

Table 1: Summary of the results obtained for system (25). The robust dynamic programming algorithm was run for $K_{\text{lower}} = 40$ and $K_{\text{upper}} = 9$ iterations.

The label "robust MDP" denotes our proposed approach, while the label "IMDP" refers to the alternative approach pointed out in Remark 1.

$$e_{\text{avg}} := \frac{1}{N} \sum_{q \in Q} (\overline{p}^{K_{\text{upper}}}(q) - \underline{p}^{K_{\text{lower}}}(q)), \tag{24}$$

which allows us to assess the conservatism of the solution.[3]

All results were obtained with an Intel Core i5-7300HQ CPU at 2.50GHz-2.50 GHz with 8GB of RAM.

7.1 Linear System

For this case study we consider a discrete-time version of the unicycle model [34], which is obtained using a first order Euler discretization with a time step $\Delta t = 1$. We fix the velocity of the vehicle to the constant value 1 and consider its orientation angle u as the control input. Thus we get the switched system:

$$x_{k+1} = x_k + \Delta t \left(\begin{bmatrix} \cos(u_k) \\ \sin(u_k) \end{bmatrix} + \mathbf{v}_k \right). \tag{25}$$

The state of the system is the vehicle position $x_k \in \mathbb{R}^2$ and its control input u_k takes values in $U = \{0, \frac{2\pi}{8}, \ldots, 7\frac{2\pi}{8}\}$.

For system (25), we define the safe set X as the rectangle $[0,1] \times [0,1] \subset \mathbb{R}^2$ by further excluding the obstacles contained therein (cf. Figure 2). We partition the rectangle $[0,1] \times [0,1]$ into a uniform grid, which together with the complement of the rectangle yields $N = 1601$ regions. The ambiguity set is centered at an empirical distribution of $M = 10$ samples, which are drawn from a Gaussian mixture with two components, centered at $[-0.01, 0]$ and $[0.01, 0]$, respectively, and with the same covariance matrix $\text{diag}(2.5 \times 10^{-5}, 2.5 \times 10^{-5})$. As a result, the centers of both components are separated by a distance close to the size of the state discretization. The obtained results are summarized in Table 1. The derivation of the abstraction took 10 minutes for all cases, since they all rely on the same nominal IMDP abstraction $\widehat{\mathcal{I}}$.[4] It can be observed in Table 1 that both the average error e_{avg} as well as the time required to perform the strategy synthesis increase as the ambiguity set grows.

The lower bound on the reachability probability obtained for experiments #1 – #4 (representing the different ε values in Table 1) is presented in Figure 1. This figure highlights how accounting for the distributional ambiguity leads to more realistic bounds: notice that the lower bound obtained for Experiment #1, where $\varepsilon = 0$, is 1 everywhere, which means that every trajectory will satisfy the specification. This trivial lower bound is the result of assuming that the empirical distribution is the true one, which is unrealistic. The corresponding abstraction only embeds the ambiguity that arises from discretizing the state space. Indeed, considering a value of ε greater than zero leads to a more realistic lower bound, as observed for Experiments #2 – #4. It is clear from the figure that the larger the ambiguity set is, the more conservative the lower bound becomes. Figure 2 shows the vector field of System (25) when the optimal strategy synthesized for experiment #2 is applied. The synthesized strategy is almost the same for experiments #2 – #4.

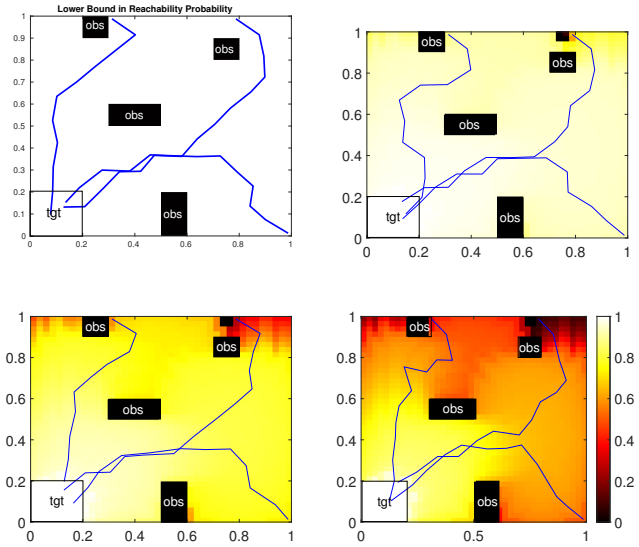


Figure 1: Results of experiments #1 – #4 in Table 1. Lower bound in the probability of reachability. The plotted trajectories correspond to Monte Carlo simulations of System (25) taking samples from a distribution $\widetilde{p}_v \in \mathcal{P}_v$.

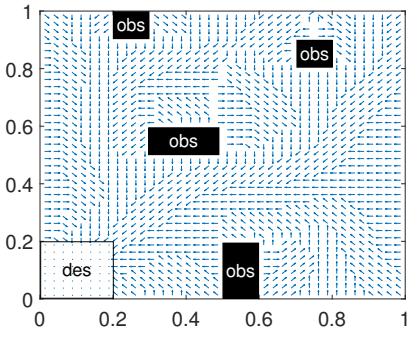
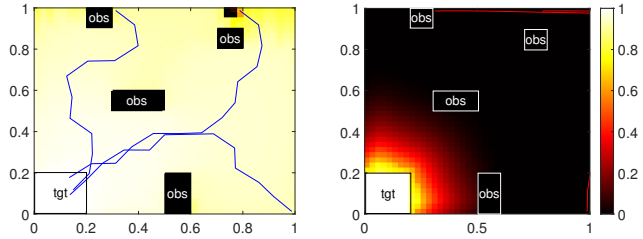


Figure 2: Vector field of System (25) in closed loop with the optimal strategy obtained for experiment #2 in Table 1.

[3] Here the conservatism of the solution has two sources: the distributional ambiguity and the accuracy of the abstraction.

[4] Notice that in the case of Experiment #1, the abstraction is simply the nominal IMDP, since no distributional ambiguity is present.

To compare our approach with the one that relies on IMDP abstractions described in Remark 1, we present the results obtained with the latter in the last row of Table 1. For this experiment, which we refer to as experiment #5, the abstraction was constructed in 13 minutes. We compare the obtained lower bounds on the reachability probability for experiments #2 and #5 in Figure 3, since they are obtained for the same ambiguity set. The results show how our proposed approach, despite requiring a larger amount of time to perform strategy synthesis, is able to provide non-trivial satisfaction guarantees, unlike the approach from Remark 1. Furthermore, our approach is able to synthesize a strategy that satisfies the reachability task, in contrast to the approach based on the IMDP abstraction. To empirically verify this argument, we computed 1000 MC simulations starting from the same states as the ones shown in Figures 1 and 3. In particular, all the trajectories of Experiments #1 − #4 satisfied the specification, unlike in Experiment #5, where no trajectory satisfied the specification.



(a) Robust MDP abstraction approach **(b) IMDP abstraction approach**

Figure 3: Results of experiments #2 and #5 in Table 1. Lower bound in the probability of reachability. The trajectories in both figures correspond to Monte Carlo simulations of System (25) taking samples from a distribution $\widetilde{p}_v \in \mathcal{P}_v$. The ones that satisfy the specification are presented in blue, while the ones that do not are presented in red.

Remark 3 (Synthesis time). *To justify the long synthesis times presented in Table 1, note that the algorithm of Section 6 was implemented in* MATLAB*, without paralelization, and the linear program presented in Section 6.2 was solved using the solver* Linprog*. We used the same solver to synthesize strategies for the IMDP abstractions to obtain a fair comparison, instead of exploiting faster, dedicated algorithms for IMDP models [16].*

7.2 Nonlinear System with 4 Modes

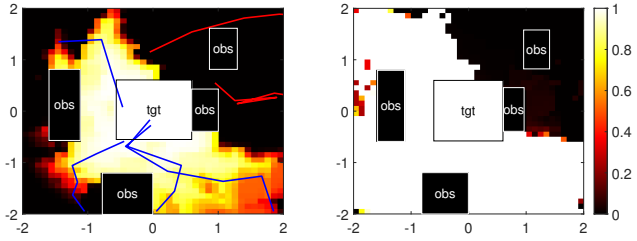
For the second case study we consider the nonlinear system from [1, 20] with dynamics:

$$x_{k+1} = x_k + \widetilde{f}_{u_k}(x_k) + \mathbf{v}_k. \tag{26}$$

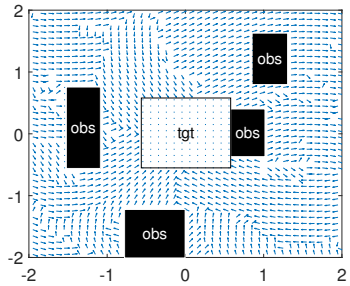
Denoting by $x^{(i)}$ the i-th component of the state, the map $\widetilde{f}_{u_k}$ is given by

$$\widetilde{f}_u(x) = \begin{cases} [0.5 + 0.2\sin(x^{(2)}), 0.4\cos(x^{(1)})]^T & \text{if } u = 1 \\ [-0.5 + 0.2\sin(x^{(2)}), 0.4\cos(x^{(1)})]^T & \text{if } u = 2 \\ [0.4\cos(x^{(2)}), 0.5 + 0.2\sin(x^{(1)})]^T & \text{if } u = 3 \\ [0.4\cos(x^{(2)}), -0.5 + 0.2\sin(x^{(1)})]^T & \text{if } u = 4. \end{cases} \tag{27}$$

The system has state $x_k \in \mathbb{R}^2$ and its control input u_k switches between the discrete values 1, 2 , 3, and 4. In analogy to the first case study, we define the safe set X as the rectangle $[-2, 2]\times[-2, 2] \subset \mathbb{R}^2$ and exclude the obstacles contained therein and discretize it into a uniform grid, which results in abstraction with $N = 1601$ states. The ambiguity set here is centered at an empirical distribution of $M = 20$ samples, that are again drawn from a Gaussian mixture with two components, centered at $[-0.05, 0]$ and $[0.05, 0]$, respectively, with covariance matrix $\mathrm{diag}(4\times10^{-4}, 4\times10^{-4})$. Again, the centers of both components are separated by a distance close to the size of the state discretization. The results were obtained for an ambiguity radius $\varepsilon = 5 \times 10^{-2}$, which yield $e_{\text{avg}} = 0.25$. Furthermore, the abstraction time was around 20 minutes, and the synthesis process took 2.43 hours: the value iteration algorithm was run $K_{\text{lower}} = 15$ times for the case of $\underline{p}^k$ and $K_{\text{upper}} = 6$ times for the case of $\overline{p}^k$. The results obtained are shown in Figure 4.



(a) Lower bound on the reachability probability **(b) Upper bound on the reachability probability**



(c) Optimal strategy

Figure 4: Results obtained for the nonlinear system (26). The trajectories in Figure 4a) correspond to Monte Carlo simulations of System (26) taking samples from a distribution $\widetilde{p}_v \in \mathcal{P}_v$. The ones that satisfy the specification are presented in blue, while the ones that do not are presented in red.

8 CONCLUSION AND FUTURE WORK

In this paper we presented a framework for the formal control of switched stochastic systems with additive, random disturbances whose probability distribution belongs to a Wasserstein ambiguity set. To this end, we derived a robust MDP abstraction of the original system and proposed an algorithm, termed robust dynamic programming, to synthesize robust strategies that maximize the probability of satisfying a reach-avoid specification. The obtained results demonstrate the effectiveness of our approach in systems with both linear and nonlinear dynamics, and show its superiority with respect to leveraging directly IMDP abstractions.

Future work includes the exploitation of dedicated robust dynamic programming algorithms to speed up the abstraction process, as well as refinement strategies to obtain smaller abstraction without resorting to additional insight regarding the system or the domain. In addition, we aim to extend the approach to richer specifications expressed as, for example, LTLf, BLTL or sc-LTL formulas.

9 PROOF OF PROPOSITION 1

To prove the proposition we will use two technical lemmas that link couplings and optimal transport discrepancies in the continuous and abstract space.

Lemma 1 (Induced Coupling on the Discrete Space). *Consider the coupling $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ with marginals p and p', a finite (measurable) partition Q of $\mathbb{R}^n$, and the induced distributions $\gamma, \gamma' \in \mathcal{D}(Q)$ with $\gamma(q) := p(q)$ and $\gamma'(q) := p(q)$ for all $q \in Q$. Then $\nu \in \mathcal{P}(Q \times Q)$, defined by*

$$\nu(q, q') := \int_{q \times q'} d\pi(x, y) \tag{28}$$

is a coupling between γ and γ.

Proof. The proof follows directly from the fact that

$$\sum_{q' \in Q} \nu(q, q') = \sum_{q' \in Q} \int_{q \times q'} d\pi(x, y) = \int_{q \times \mathbb{R}^n} d\pi(x, y) = \gamma(q),$$

and analogously for the other marginal. □

Next, given two distributions on the continuous space $\mathbb{R}^n$ we establish bounds on the optimal transport discrepancy $\mathcal{T}_c$ of their induced distributions on Q, based on their s-Wasserstein distance in the continuous space.

Lemma 2 (Induced optimal transport discrepancy). *Let $p, p' \in \mathcal{D}_s(\mathbb{R}^n)$ and consider the induced distributions $\gamma, \gamma' \in \mathcal{D}(Q)$ with $\gamma(q) := p(q)$ and $\gamma'(q) := p(q)$ for all $q \in Q$. Then for any $s \geq 1$ and $\varepsilon \geq 0$ it holds that*

$$\mathcal{W}_s(p, p') \leq \varepsilon \Rightarrow \mathcal{T}_c(\gamma, \gamma') \leq \varepsilon^s,$$

where c is given in (10).

Proof. Consider the map J in (21) and note that due to (10),

$$\|x - y\|^s \geq c(J(x), J(y)) \tag{29}$$

for all $x, y \in \mathbb{R}^n$. Let π be an optimal coupling for the s-Wasserstein distance $\mathcal{W}_s(p, p')$ and ν be the induced coupling on Q given by (28). Then we get from (1), (21), and (29) that

$$\begin{aligned}
\mathcal{W}_s(p, p')^s &= \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|^s d\pi(x, y) \\
&\geq \int_{\mathbb{R}^n \times \mathbb{R}^n} c(J(x), J(y)) d\pi(x, y) \\
&= \sum_{q, q' \in Q} c(q, q') \int_{q \times q'} d\pi(x, y) \\
&= \sum_{q, q' \in Q} c(q, q') \nu(q, q') \geq \mathcal{T}_c(\gamma, \gamma'),
\end{aligned}$$

which implies the result. The last inequality follows from (1) and Lemma 1, which asserts that ν is a coupling between γ and γ'. The proof is complete. □

The intuition behind Lemma 2 is the following: if the s-Wasserstein distance between two distributions in $\mathbb{R}^n$ is at most ε, then the optimal transport discrepancy (based on c) between their induced distributions on Q is not more than ε^s.

Proof of Proposition 1. Let q, a, x, and p_v as given in the statement and define

$$\widehat{\gamma}_{x,a}(q') := T^a_{\widehat{p}_v}(q' \mid x)$$

for all $q' \in Q$. Then it follows from (7) and (9) that

$$\widehat{\gamma}_{x,a} \in \widehat{\Gamma}_{q,a}. \tag{30}$$

Next, we get from (3) and Assumption 1 that $T^a_{p_v}(\cdot \mid x)$ and $T^a_{\widehat{p}_v}(\cdot \mid x)$ are distributions in $\mathcal{D}_s(\mathbb{R}^n)$ and that

$$\mathcal{W}_s\left(T^a_{p_v}(\cdot \mid x), T^a_{\widehat{p}_v}(\cdot \mid x)\right) \leq \varepsilon.$$

Since the induced distributions of $T^a_{p_v}(\cdot \mid x)$ and $T^a_{\widehat{p}_v}(\cdot \mid x)$ on Q are $\gamma_{x,a}$ and $\widehat{\gamma}_{x,a}$, respectively, it follows from Lemma 2 that $\mathcal{T}_c(\gamma_{x,a}, \widehat{\gamma}_{x,a}) \leq \varepsilon^s \equiv \epsilon$, namely, $\gamma_{x,a} \in \mathcal{T}_c^{\epsilon}(\widehat{\gamma}_{x,a})$. Thus, we deduce from (11) and (30) that $\gamma_{x,a} \in \Gamma_{q,a}$ and conclude the proof. □

ACKNOWLEDGMENTS

Manuel Mazo Jr.'s work was partially supported by the European Research Council through the SENTIENT project, Grant No. ERC-2017-STG #755953 (https://cordis.europa.eu/project/id/755953)

REFERENCES

[1] Steven Adams, Morteza Lahijanian, and Luca Laurenti. 2022. Formal control synthesis for stochastic neural network dynamic models. *IEEE Control Systems Letters* (2022).
[2] Calin Belta, Boyan Yordanov, and Ebru Aydin Gol. 2017. *Formal methods for discrete-time dynamical systems*. Vol. 89. Springer.
[3] Dimitir P Bertsekas and Steven Shreve. 2004. *Stochastic optimal control: the discrete-time case*.
[4] Jose Blanchet and Karthyek Murthy. 2019. Quantifying Distributional Model Risk via Optimal Transport. *Mathematics of Operations Research* 44, 2 (2019), 565–600.
[5] Jose Blanchet, Karthyek Murthy, and Fan Zhang. 2022. Optimal transport-based distributionally robust optimization: Structural properties and iterative schemes. *Mathematics of Operations Research* 47, 2 (2022), 1500–1529.
[6] Dimitris Boskos, Jorge Cortés, and Sonia Martínez. 2020. Data-driven ambiguity sets for linear systems under disturbances and noisy observations. In *2020 American Control Conference (ACC)*. 4491–4496.
[7] El-Kébir Boukas. 2007. *Stochastic switching systems: analysis and design*. Springer Science & Business Media.
[8] Giuseppe Carlo Calafiore and L El Ghaoui. 2006. On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications* 130, 1 (2006), 1–22.

[9] Nathalie Cauchi, Luca Laurenti, Morteza Lahijanian, Alessandro Abate, Marta Kwiatkowska, and Luca Cardelli. 2019. Efficiency through Uncertainty: Scalable Formal Synthesis for Stochastic Hybrid Systems. In *Proceedings of the 2019 22nd ACM International Conference on Hybrid Systems: Computation and Control*. ACM, Montreal, QC, Canada. https://doi.org/10.1145/3302504.3311805
[10] Julien Grand Clement and Christian Kroer. 2021. First-order methods for Wasserstein distributionally robust MDP. In *International Conference on Machine Learning*. PMLR, 2010–2019.
[11] Erick Delage and Yinyu Ye. 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research* 58, 3 (2010), 595–612.
[12] Maxence Dutreix and Samuel Coogan. 2020. Specification-guided verification and abstraction refinement of mixed monotone stochastic systems. *IEEE Trans. Automat. Control* 66, 7 (2020), 2975–2990.
[13] Maxence Dutreix, Jeongmin Huh, and Samuel Coogan. 2022. Abstraction-based synthesis for stochastic systems with omega-regular objectives. *Nonlinear Analysis: Hybrid Systems* 45 (2022), 101204.
[14] Nicolas Fournier and Arnaud Guillin. 2015. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields* 162, 3 (2015), 707–738.
[15] Rui Gao and Anton Kleywegt. 2022. Distributionally robust stochastic optimization with Wasserstein distance. *Mathematics of Operations Research* (2022).
[16] Robert Givan, Sonia Leach, and Thomas Dean. 2000. Bounded-parameter Markov decision processes. *Artificial Intelligence* 122, 1-2 (2000), 71–109.
[17] Sofie Haesaert, Nathalie Cauchi, and Alessandro Abate. 2017. Certified policy synthesis for general Markov decision processes: An application in building automation systems. *Performance Evaluation* 117 (2017), 75–103.
[18] Astghik Hakobyan and Insoon Yang. 2021. Wasserstein distributionally robust motion control for collision avoidance using conditional value-at-risk. *IEEE Transactions on Robotics* 38, 2 (2021), 939–957.
[19] John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. 2021. Formal verification of unknown dynamical systems via Gaussian process regression. *arXiv preprint arXiv:2201.00655* (2021).
[20] John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. 2021. Strategy synthesis for partially-known switched stochastic systems. In *Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control*. 1–11.
[21] Xenofon Koutsoukos and Derek Riley. 2006. Computational methods for reachability analysis of stochastic hybrid systems. In *International Workshop on Hybrid Systems: Computation and Control*. Springer, 377–391.
[22] Morteza Lahijanian, Sean B Andersson, and Calin Belta. 2015. Formal verification and synthesis for discrete-time stochastic systems. *IEEE Trans. Automat. Control* 60, 8 (2015), 2031–2045.
[23] Luca Laurenti, Morteza Lahijanian, Alessandro Abate, Luca Cardelli, and Marta Kwiatkowska. 2020. Formal and efficient synthesis for continuous-time linear stochastic hybrid processes. *IEEE Trans. Automat. Control* 66, 1 (2020), 17–32.
[24] Abolfazl Lavaei and Majid Zamani. 2022. From Dissipativity Theory to Compositional Synthesis of Large-Scale Stochastic Switched Systems. *IEEE Trans. Automat. Control* (2022).
[25] Ryan Luna, Morteza Lahijanian, Mark Moll, and Lydia E Kavraki. 2015. Asymptotically optimal stochastic motion planning with temporal goals. In *Algorithmic Foundations of Robotics XI*. Springer, 335–352.
[26] Peyman Mohajerin Esfahani and Daniel Kuhn. 2018. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming* 171, 1 (2018), 115–166.
[27] Arnab Nilim and Laurent El Ghaoui. 2005. Robust control of Markov decision processes with uncertain transition matrices. *Operations Research* 53, 5 (2005), 780–798.
[28] Ioana Popescu. 2007. Robust mean-covariance solutions for stochastic optimization. *Operations Research* 55, 1 (2007), 98–112.
[29] Alberto Puggelli, Wenchao Li, Alberto L Sangiovanni-Vincentelli, and Sanjit A Seshia. 2013. Polynomial-time verification of PCTL properties of MDPs with convex uncertainties. In *International Conference on Computer Aided Verification*. Springer, 527–542.
[30] Hamed Rahimian and Sanjay Mehrotra. 2019. Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659* (2019).
[31] Filippo Santambrogio. 2015. Optimal transport for applied mathematicians. *Birkäuser, NY* 55, 58-63 (2015), 94.
[32] Cesar Santoyo, Maxence Dutreix, and Samuel Coogan. 2021. A barrier function approach to finite-time stochastic system verification and control. *Automatica* 125 (2021), 109439.
[33] Alexander Shapiro. 2017. Distributionally robust stochastic programming. *SIAM Journal on Optimization* 27, 4 (2017), 2258–2275.
[34] Paulo Tabuada. 2008. An approximate simulation approach to symbolic control. *IEEE Trans. Automat. Control* 53, 6 (2008), 1406–1418.
[35] Cédric Villani. 2003. *Topics in optimal transportation*. Vol. 58. American Mathematical Society.
[36] Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. 2013. Robust Markov decision processes. *Mathematics of Operations Research* 38, 1 (2013), 153–183.
[37] Eric M Wolff, Ufuk Topcu, and Richard M Murray. 2012. Robust control of uncertain Markov decision processes with temporal logic specifications. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 3372–3379.
[38] Huan Xu and Shie Mannor. 2010. Distributionally robust Markov decision processes. *Advances in Neural Information Processing Systems* 23 (2010).
[39] Insoon Yang. 2017. A convex optimization approach to distributionally robust Markov decision processes with Wasserstein distance. *IEEE control systems letters* 1, 1 (2017), 164–169.
[40] George Yin and Chao Zhu. 2010. *Hybrid switching diffusions: properties and applications*. Vol. 63. Springer New York.

A REFORMULATION OF THE LINEAR PROGRAM IN THEOREM 1

Consider the transition probability bounds (7) and (8) of the nominal IMDP $\widehat{\mathcal{I}}$. For each $q \in Q$ and $a \in A$ let

$$\mathcal{N}_{q,a} := \{i \in \mathcal{N} : \overline{P}(q,a,q_i) > 0\} \tag{31}$$

be the set of outgoing transitions in $\widehat{\mathcal{I}}$ that may have nonzero transition probability. The following theorem provides a reformulation of the linear program (19)-(20), which has a reduced complexity when $\mathcal{N}_{q,a}$ is strictly smaller than $\mathcal{N}$.

Theorem 3. *Under the assumptions of Theorem 1, for each $q \in Q$ and $a \in A$ the linear program* (19)-(20) *is equivalent to:*

$$\min_{\pi_{ij}} \sum_{i\in\mathcal{N}, j\in\mathcal{N}_{q,a}} \pi_{ij}\underline{p}^k(q_i), \tag{32}$$

$$\text{s.t. } \underline{P}(q,a,q_j) \le \sum_{i\in\mathcal{N}} \pi_{ij} \le \overline{P}(q,a,q_j) \qquad j \in \mathcal{N}_{q,a} \tag{33a}$$

$$\sum_{i\in\mathcal{N}, j\in\mathcal{N}_{q,a}} \pi_{ij} = 1 \tag{33b}$$

$$\pi_{ij} \ge 0, \qquad i \in \mathcal{N}, j \in \mathcal{N}_{q,a} \tag{33c}$$

$$\sum_{i\in\mathcal{N}, j\in\mathcal{N}_{q,a}} \pi_{ij} c(q_i,q_j) \le \varepsilon^s, \tag{33d}$$

with $\mathcal{N}_{q,a}$ as given in (31).

Proof. Notice that one can directly eliminate the optimization variables γ_i and $\widehat{\gamma}_i$ by substituting their expressions (20d) and (20e) in (19), (20a), and (20b). Thus, since each $\widehat{\gamma}_j$, and therefore also each π_{ij} is identically fixed to zero when $j \notin \mathcal{N}_{q,a}$, we obtain the equivalent optimization problem (32)-(33) after eliminating all the redundant decision variables. □