Distributionally Robust Abstraction and Strategy Synthesis with Formal Guarantees

Ibón Gracia Merino (5358779)

Master of Science Thesis

Delft University of Technology

Delft Center for Systems and Control

Distributionally Robust Abstraction and Strategy Synthesis with Formal Guarantees

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Ibón Gracia Merino (5358779)

October 6, 2022

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Distributionally Robust Abstraction and Strategy Synthesis with Formal Guarantees

by

Ibón Gracia Merino (5358779)

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: October 6, 2022

Supervisor(s):

Prof. Dr. Dimitris Boskos. Supervisor

Prof. Dr. Luca Laurenti. Second Supervisor

Reader(s):

Prof. Dr. Manuel Mazo Jr. First Reader

Prof. Dr. Laura Ferranti. Second Reader

Abstract

There is growing interest to control cyber-physical systems under complex specifications while retaining formal performance guarantees. In this thesis we present a framework for formal control of uncertain systems under complex specifications. We consider dynamical systems with random disturbances, whose probability distribution is unknown. When it comes to the specifications, we focus on those given as syntactically co-safe linear temporal logic (scLTL) formulas. Such formulas resemble natural language, and allow us to reason over complex behaviours of the system.

We follow an abstraction-based approach: we abstract the original system to a finite-state Markov model, in which the state discretization error as well as the distributional ambiguity, are embedded as uncertainties. To do so we make use of tools from optimal transport and ambiguity sets of probability distributions. After that, we obtain a strategy for the abstraction and obtain probabilistic guarantees that the abstraction satisfies the specification. Finally, we correctly refine the strategy to one that the original system can follow, and prove that the guarantees we obtained for the abstraction also hold for the original system.

We propose two approaches to obtain the abstraction. First, we propose a data-driven approach to abstract the system into an interval Markov decision process (IMDP) when samples from the unknown distribution is available. Then we use already existing algorithms to obtain a strategy for the IMDP. Secondly, we propose an approach to abstract the original system into a robust Markov decision process (robust MDP). This second approach is applicable to more general uncertainty models besides the data-driven one, and reduces conservatism of the abstraction. Furthermore, we propose an algorithm to obtain robust strategies for robust MDPs, which also renders the guarantees that the abstraction satisfies the specification. Finally, we demonstrate the usefulness of our proposed approaches through several case studies that involve both linear and nonlinear systems.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I want to thank my mother for her constant support and love. I would not have come this far or become the person I am without her. She is the source of my strength.

I also want to thank my supervisors Dimitris and Luca, for all the suggestions and tips that they gave me during our frequent meetings. This is probably the first work in which I put huge emphasis not only in the coherence and correctness of its content, but also in presenting the ideas I discuss in the easiest way to understand. Without the constant direction of my supervisors this would not have been possible.

Finally, I want to thank all my friends and peers from Systems and Control, since we have helped each other many times through this experience.

Delft, University of Technology
October 6, 2022

Ibón Gracia Merino (5358779)

Chapter 1

Introduction

We begin this chapter by motivating our approach through a toy example of controlling a UAV that is affected by random, uncertain wind, when the specifications are complex in 1-1. After that, we state our objectives and briefly describe our approach in Section 1-2. Next, in Section 1-4 we give an outline of the document. Finally, in Section 1-5 we introduce basic notation that we will use throughout the document.

1-1 Motivation

Consider a physical dynamical system affected by a random disturbance. An example of such system can be a UAV moving in a two-dimensional space, and disturbed by random gusts of wind, which we depict in Figure 1-1. However, consider that the probability distribution of the disturbance is unknown to us. Instead, we have access to a nominal probability distribution, which is an estimation of the true, unknown, one. Furthermore, consider that our problem is that of controlling the UAV in such a way that the probability of satisfying a given complex

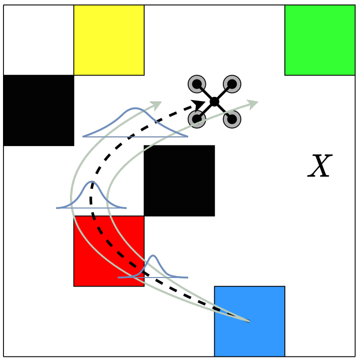


Figure 1-1: UAV example. The region $X \subset \mathbb{R}^2$ contains the different regions of interest `blue`, `green`, `yellow`, `red` and `black`, each one represented by a coloured region.

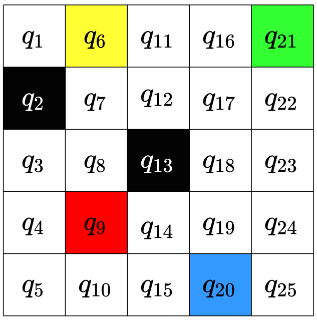
specification is satisfied. An example of such a specification can be the following one, inspired by the applications in [1] and [2]: “eventually reach red, yellow and green in no particular order, and then eventually reach blue while avoiding black and never exiting the region X throughout all the trajectory".

From the point of view of classical control theory, it is not clear how we should approach this problem: first, the specification is complex and, second, the number of possible motions of the UAV, due to the continuous nature of its dynamics, is infinite. An approach based on a single optimization problem is intractable, even if we consider only part of the specification. Furthermore, the dynamics of the system are not deterministic, but random, and there is ambiguity with respect to this randomness. Moreover, if the UAV surveillance problem is safety-critical for example, we are interested in obtaining guarantees that the system satisfies the specification. However, we cannot achieve this objective with tools from classical control theory.

Let us start simple, considering that the probability of the disturbance is known, to give a brief survey of the existing approaches. In this setting, formal methods help us solve problems like the one previously described [3], [4], [5], in a systematic way. By using tools from this field, we are able to formulate the motion specification of the UAV as a linear temporal logic (LTL) formula, which resembles natural language. This kind of formulas are used for model checking in computer science where, typically, systems with a finite number of states are analysed. However, physical systems like our UAV are often described by sets of differential or difference equations and an uncountable state space, which makes it impossible for us to directly apply tools from formal methods to them. In order to do so, we can construct a finite state model of the system that can be analyzed with these methods [6], [7]. This procedure is called abstraction. Furthermore, after analyzing the abstraction, formal methods allow us to relate the properties of the abstraction to the original system. Such properties can be, for example, guarantees of satisfying the specification, which we wanted to obtain in the first place.

Let us go back to the example of the UAV affected by random wind, and let us obtain an abstraction of said system. In this uncertain setting, we should consider abstractions that capture the stochasticity of the motion of the UAV due to the wind, such as Markov decision processes (MDP) [8], [9]. To obtain an MDP abstraction, first, we partition the continuous space X into a rectangular grid that respects the regions of interest (red, blue, etc.) as shown in Figure 1-2. Then we define the states of the MDP in such a way that each state represents a cell q_i in the previous partition. We say that current state of the MDP is q_i if the UAV's position is inside the corresponding cell. With an abuse of notation we refer by q_i to a state of the MDP abstraction and also to a region in the continuous space X. The correct interpretation should be clear from the context. The MDP depicted in Figure 1-2 has 25 states, which corresponds to a 5×5 grid. We can assume that the initial position of the UAV is inside cell q_{20}, for example, which gives an initial state to our MDP. Next, we associate an observation to each cell. These observations come from the observations associated to the continuous state space prior to its discretization. The observation corresponding to each cell is indicated by its color, and the observation associated to the white cells is the empty set $\emptyset$.

After that, we define the possible transitions between the states of the abstraction, which must resemble the physical motion of the UAV between cells. We can do this by considering only commanding the UAV to move between adjacent cells in the horizontal and vertical



q_1	q_6	q_{11}	q_{16}	q_{21}
q_2	q_7	q_{12}	q_{17}	q_{22}
q_3	q_8	q_{13}	q_{18}	q_{23}
q_4	q_9	q_{14}	q_{19}	q_{24}
q_5	q_{10}	q_{15}	q_{20}	q_{25}

Figure 1-2: Partition of the continuous space $X \in \mathbb{R}^2$ and corresponding set of states of the MDP abstraction.

directions and also commanding it to stay at rest in the current cell. The available actions of the MDP for every state are therefore a_f for move forward, a_b for move backwards, a_r for move right, a_l for move left and a_c for staying at the current cell, i.e., staying at rest. However, we must notice that for a fixed current state of the UAV and action, the successor state is not deterministic, but it is described by a probability distribution. This is due to the wind disturbances, which are probabilistic. In this setting we define the transition probabilities of our MDP abstraction in the following way: when the UAV lies at the center of a cell, we compute how much probability mass of the distribution of the successor state lies over each cell [10], as we illustrate on the left of Figure 1-3. That is the probability of transitioning to each state associated to those cells.

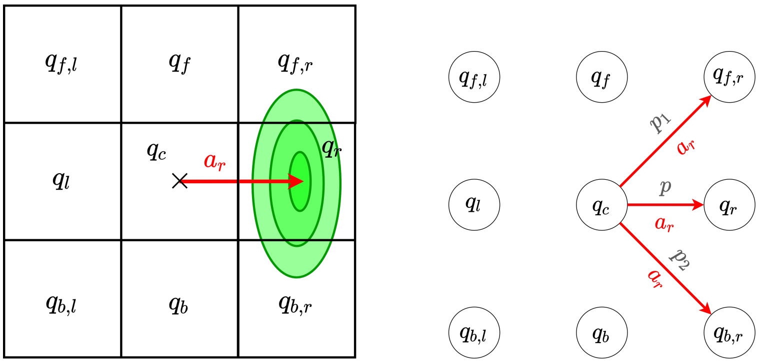


Figure 1-3: Transitions of the MDP abstraction of the UAV problem. On the left, probability distribution of the successor state of the UAV starting from the center of cell q_c and for action a_r. On the right, possible transitions of the MDP abstraction at state q_c enabled by action a_r. The probability of each transition is indicated in grey above each red arrow.

On the right of Figure 1-3, we show the probability distribution of the continuous state of the UAV after one time step while commanding it to move to the right. It is important to notice that the state transitions in the abstraction are also stochastic, and commanding the UAV to move right, for example, may also lead to the cells denoted by $q_{f,r}$ and $q_{b,r}$, in Figure 1-3. For example, consider that the UAV is at the current state q_4. Then, commanding the UAV to move to the cell on its right, q_9, can enable the transition to this cell with probability p but also the transitions to q_8 with probability p_1 and q_{10} with probability p_2.

Remember that we want to use this MDP abstraction to obtain a strategy for the UAV that enforces the specification. Since the abstraction is a stochastic model, the algorithms for MDPs allow us to compute the strategy that maximizes the probability of satisfying said specification. Furthermore, these algorithms also provide us with such probabilistic guarantees. Then if, with the computed strategy, we obtained a probability of satisfying the specification that is below a safe threshold defined by us, we would consider the performance of the UAV as not safe. In this case, we should redesign either the strategy, the system or the specification.

Regarding MDP abstractions, we must highlight that an assumption is always made in order to build the abstraction: in our example we have assumed that the current continuous state of the UAV is at the center of the cell that corresponds to state q_c. Had this state been a different one inside the cell, the transition probabilities we computed would be different. By making this assumption, we incur a discretization error in the abstraction, which means that the abstraction is an approximate model of the original system [9], [10]. This fact implies that the results we obtain by analyzing an MDP abstraction, such as guarantees of satisfying the specification, only hold approximately for the original system. In any case, it is always possible to obtain a more precise abstraction, which leads to more precise results. In our UAV example we could make the abstraction more precise if the partition of the state space was finer, and the performance guarantees would be closer to those of the original system [9]. However, a more precise abstraction comes at the expense of a higher computational burden required to construct the abstraction and also to analyze it using tools from formal methods [3].

An additional disadvantage of performing an abstraction is that, if a strategy is obtained for the abstraction such that it maximizes the probability of satisfying some specification, this strategy may not be optimal for the original system [11]. This is because the abstraction is either an approximate or a conservative model of the original system. Furthermore, generally, the performance level that the abstraction attains for a fixed strategy is not an exact guarantee of the performance level that the UAV will attain with the same strategy. However, for some classes of abstractions we are able to obtain such guarantees through correctness proofs.

A smart way to completely eliminate the discretization error is to make use of IMDP abstractions [3], [4]. These are a generalization of MDP models, and can be viewed as a family of MDPs since, in IMDPs, the transitions have associated ranges of probabilities instead of single probabilities. In Figures 1-4 and 1-5 we illustrate how we can obtain an IMDP abstraction of our UAV example. There we can appreciate how we define the bounds in the probability of transitioning between states q_c and q_r under action a_r: they are defined as the minimum and

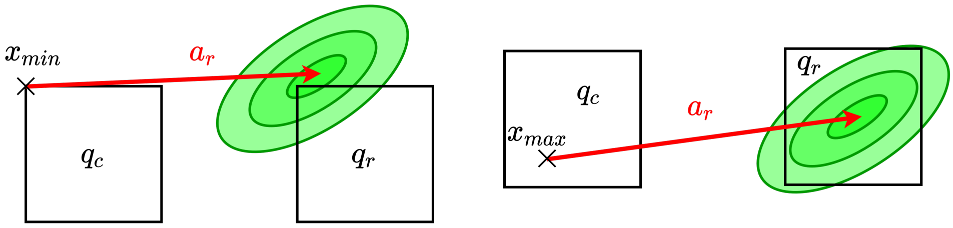


Figure 1-4: Position of the UAV after one time step starting from $x_{min} \in q_c$ on the left and from $x_{max} \in q_c$ on the right when commanded to move right. For clarity, these cells, although adjacent to each other, have been depicted as separated.

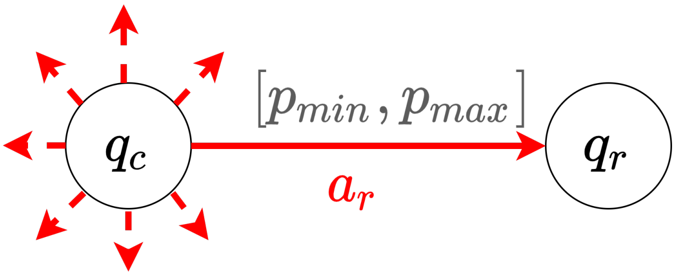


Figure 1-5: Transition of the IMDP abstraction $\mathcal{I}$ corresponding to the continuous dynamics in Figure 1-4. The bounds in the probability of transitioning from state q_c to q_r under action a_r are indicated in the figure as p_{min} and p_{max}, respectively. Note that only the transition to the cell on the right of q_c is represented here for clarity, while the additional transitions to other cells are represented only by dashed arrows.

maximum fraction of mass of the successor state that falls inside cell q_r when starting from a point in cell q_c. Furthermore, in Figure 1-4 we denote by x_{min} and x_{max} the points inside q_c for which the probability of the UAV reaching cell q_r in one time step by starting from cell q_c and under action a_r is minimized and maximized, respectively. Moreover, in Figure 1-5 we denote the corresponding probabilities of such events by p_{min} and p_{max}. In this way, the range of transition probabilities of the abstraction includes the transition probabilities that correspond to all possible current positions of the UAV. This is done for every cell and, therefore, the abstraction considers all behaviours of the original system. Once we have constructed the IMDP abstraction, we are able to, again, use tools from formal methods to obtain a strategy that enforces the specification. Additionally, since we did not incur any discretization error, we are able to prove that the guarantees obtained for the abstraction also hold for the original system [4], [7], [12]. Note that, in IMDP abstractions, since the transition probabilities are ranges, the probabilistic guarantees of the UAV satisfying the specification are also a range of probabilities.

Previously in this section, we have described how we deal with systems whose stochastic behaviour is known. Now, let us address the problem that the stochastic behaviour of the system is uncertain. In our UAV example, this can be because we do not know the exact probability distribution of the wind disturbance. However, we can always obtain information about this behaviour if we measure the wind speed a finite number of times. Note that these samples only give us partial knowledge about the true probability distribution of the disturbances. Another setting is that of robust control: we have a nominal, approximate probability distribution of the random disturbance, but we want to tolerate small deviations from this one to obtain guarantees of increased robustness. In order to obtain a strategy for these ambiguous systems, we can start by obtaining an abstraction that accounts for this ambiguity. If we make use of IMDP abstractions, then we can use the already existing approaches to obtain a strategy that enforced the specification while being robust with respect to the distributional uncertainty. However, in this thesis we will show that usually IMDP abstractions are not enough, since they provide trivial guarantees in many practical cases. Therefore we will also make use of more complicated Markovian abstractions, and develop algorithms that allow us to obtain strategies for these ones. To perform abstractions that account for the ambiguity in the distribution of the disturbance, we will make use of tools from optimal transport and ambiguity sets.

1-2 Objectives and Approach

Our objective is to synthesize a strategy for a discrete-time, continuous-state system with additive disturbance, whose probability distribution is uncertain, under complex specifications. We consider specifications given as scLTL formulas, since it is easier to check if the system satisfies such specifications and these are expressive enough. To synthesize a strategy for the original system, we compute an abstraction of that system. We make use of either IMDPs or Markov models with a more complex uncertainty set of transition probabilities, which we call robust MDP. After that, we synthesize a strategy for the abstraction that maximizes the probability of satisfying the specification, when the transitions between states are such that reduce the most this probability. Finally, we refine the synthesized strategy to one that the original system can use. To account for distributional uncertainty about the probability distribution of the disturbance, we leverage tools from optimal transport and ambiguity sets.

We consider two possible scenarios. In the first one, we consider a data-driven setting, in which we only have samples from the disturbance, and we want to account for this uncertainty about its true distribution. However, in this thesis we always consider given ambiguity sets. We leave the problem of how to compute such an ambiguity set that contains the true distribution to further research. A possible approach to estimate the size of the ambiguity set from data is to make use of the measure concentration results from [13]. Using this approach we would obtain sets that are guaranteed to contain the true probability distribution with some confidence. Other practical approaches to determine the size of the ambiguity set are cross-validation tests, bootstrapping or goodness-of-fit tests [14]. On the other hand, we consider a robust control problem, in which the probability distribution of the disturbance is known, but we want to tolerate small deviations from this one. The result of both approaches, under our assumptions, is a formal abstraction, which includes the distributional ambiguity, and which leads to formal guarantees. We highlight that this second setting is more general, and the nominal probability distribution does not need to be built from data, but can be any distribution.

1-3 Contribution

The contributions of this thesis are the following:

- We propose a framework to perform formal strategy synthesis for dynamical systems with additive, random and ambiguous disturbances under complex specifications.
- We present a novel approach to compute data-driven abstractions of such systems to IMDPs when we have a finite amount of samples of the disturbance. In this approach, the size of the ambiguity set is given: the problem of finding the size of the ambiguity set such that it contains the true, unknown probability of the disturbance is out of the scope of this thesis.

- We propose a new class of Markovian abstractions which account for small variations in the probability distribution of the disturbance, according to the Wasserstein distance. We denote said abstractions as robust MDPs,

- We propose a novel synthesis algorithm for robust MDPs that allows us to obtain robust strategies and satisfaction guarantees under specifications given as scLTL formulas.

- We prove that the satisfaction guarantees obtained for the abstractions hold for the original system,

- We demonstrate the efficacy of our approaches through several case studies with both linear and nonlinear systems.

1-4 Outline

This thesis document is structured as follows: first, we give an overview, in Chapter 2 of the problem of synthesizing a strategy that enforces a complex specification for stochastic systems, when the stochastic behaviour of the original system is known. For that, we define tools from formal methods such as LTL in Section 2-1 and automata in Section 2-2, which allow us to formulate complex specifications and to synthesize a strategy that enforces those. Additionally, in Sections 2-3 and 2-4, we define Markov models such as MDPs and IMDPs, respectively, and discuss their use as abstractions. Furthermore, in Section 2-4-1, we describe how to synthesize a strategy that maximizes the probability of satisfying a given complex specification for IMDPs. At the end of each section of this chapter, we give a small review of the previous works related to those Sections. Next, we introduce tools from optimal transport, such as the Wasserstein distance between probability measures, and the concept of ambiguity sets in Chapter 3. We need these tools for our purpose of performing abstractions of stochastic, uncertain systems.

We propose our first approach to synthesize strategies of uncertain stochastic systems in Chapter 4: we describe how to obtain data-driven distributionally robust (DR) IMDP abstractions of stochastic, uncertain systems. We leverage data from said systems to build IMDP abstractions that are robust with respect to the uncertainty about the stochastic behaviour of the system. Additionally, in Chapter 5, we present our second approach. We describe how to obtain robust MDP abstractions of stochastic, uncertain systems. We aim to leverage tools from optimal transport to perform abstractions of said systems, when the probability of the disturbance belongs to a given ambiguity set. Furthermore, we modify the interval value iteration algorithm used used for strategy synthesis in IMDPs to be able to synthesize a strategy for a robust MDP. The latter approach is general: it is not limited to the data-driven setting. However, if used in said setting, it leads to less conservative guarantees when there is high uncertainty about the stochastic behaviour of the original system. After that, we show in Chapter 6 the results of the approaches described in Chapters 4 and 5 in two case studies: a linear system and a nonlinear one. Finally, we summarize our results in 7, where we also point out to possible future work.

1-5 Basic Notation

Through this document, we make use of the following notation and mathematical symbols:

We denote by $\mathbb{N}_{\geq 0} = \mathbb{N} \cup \{0\}$ and by $[n]$ the set of integers $\{1, 2, ..., n\}$, for all $n \in \mathbb{N}$. We also denote the complement of set $\mathcal{Y}$ in $\mathcal{X}$ by $\mathcal{X} \setminus \mathcal{Y}$. Additionally, we denote by $\mathbf{1}_{\mathcal{X}}$ the indicator function of set $\mathcal{X}$ For a Polish [1] (separable complete metric) space Ξ equipped with a metric d, $\mathcal{B}(\Xi)$ and $\mathcal{P}(\Xi)$ denote the Borel σ-algebra of Ξ and the set of Borel probability measures on Ξ, respectively. Furthermore, for distance d and $p \geq 1$, we denote by $\mathcal{P}_p(\Xi)$ the set of Borel probability measures on Ξ, with finite p-th moment :

$$\mathcal{P}_p(\Xi) = \{P \in \mathcal{P}(\Xi) : \int_{\Xi} d^p(\xi, \zeta_0) P(d\xi) < \infty \text{ for some } \zeta_0 \in \Xi\},$$

where the set $\mathcal{P}_p(\Xi)$ does not depend on ζ_0 due to the triangle inequality [15]. Notice that we omit the dependence of $\mathcal{P}_p(\Xi)$ on d, since this distance is always clear from the context. We denote the probability (under probability distribution P) of event A happening by $P(A)$, and the expectation operator under the same probability distribution is denoted by $\mathbb{E}^P[\cdot]$. We denote as M^T the transpose of a matrix $M \in \mathbb{R}^{n \times m}$ and by $\langle x, y \rangle = x^T y$ the inner product between two vectors x, $y \in \mathbb{R}^n$. For a vector $x \in \mathbb{R}^n$, $\|x\|_p$ denotes its p-norm and $\|x\|_{p,*} = \sup_{\|\xi\|_p \leq 1} x^T \xi$ the corresponding dual norm. If the norm is the 1-norm, we omit the subscript p. For a metric space $(\mathcal{X}, d)$, for $x \in \mathcal{X}$ and $\mathcal{Y} \subset \mathcal{X}$, we denote by $d(x, \mathcal{Y}) = \inf\{d(x, y) : y \in \mathcal{Y}\}$ the minimum distance between x and set $\mathcal{Y}$. We also denote by $B_r(c; \mathcal{X}, d) \subset \mathcal{X}$, $r > 0$, $c \in \mathcal{X}$ a ball of radius r with center on c. When $(\mathcal{X}, d)$ is clear from the context, we omit it as an argument and simply write $B_r(c)$. Finally, for a probability space $\mathcal{P}_p(\Xi)$ and distance between probabilities on this space $\mathcal{W}$, we denote by $\mathbb{B}_r(P; \mathcal{P}_p(\Xi), \mathcal{W}) \subset \mathcal{P}_p(\Xi)$,$r > 0$, $P \in \mathcal{P}_p(\Xi)$, a ball based on $\mathcal{W}$ with center on P and radius r. If $\mathcal{P}_p(\Xi)$ and $\mathcal{W}$ are clear from the context, we omit them as arguments, and use the simpler notation $\mathbb{B}_r(P)$.

[1] We consider Polish spaces to make use of measures over continuous as well as discrete spaces.

Chapter 2

Background on Formal Strategy Synthesis for Markov Processes

In this chapter we give an introduction to formal methods. Formal methods involve analysis and design of complex systems under complex specifications. The idea is that performing appropriate mathematical analysis can contribute to the reliability and robustness of a system, and allow the derivation of correct-by-design control systems [16]. Beyond classical specifications in control theory such as stability and controllability of a system or invariance of a set, etc. [3], in formal methods we use rich specifications. We typically formulate such specifications using tools like Linear Temporal Logic (LTL), since LTL formulas are able to represent a broad class of complex properties. Examples of such properties are nothing bad ever happening (safety), something good eventually happening (liveness) and even more elaborate behaviours in time [6]. Furthermore it is possible to formally check if a system satisfies such specifications using tools like automata or graph theory. Using tools like these is becoming increasingly necessary with the new developments in cyber-physical systems. In such systems, where physical elements and software are closely coupled, tools from formal methods are needed for analysis and controller synthesis purposes. For example, a crucial aspect in safety-critical applications is the procurement of formal guarantees [4], which can be achieved with these tools, and not just with those coming from classical control theory.

Formal methods come from computer science, where they are applied to finite state systems such as transition systems or MDPs. Therefore, in order to use these tools on more complex systems such as cyber-physical ones, we need to find adequate, finite state representations of such systems. Such representations are called abstractions. Roughly, an abstract model can be seen as a finite transition graph, whose states represent aggregate states of the original system. Furthermore,the transitions of the abstraction correspond to state trajectories of the original system [6]. Once we have constructed the abstraction, we can use it to check if the original system satisfies the specifications and to synthesize a strategy that achieves that task. Verification and synthesis tasks can be performed by using model-checking tools such as automata.

A particular type of abstractions which are useful for studying stochastic systems are finite state Markov models. For such models, both verification and strategy synthesis are performed in a probabilistic setting: through verification we obtain probabilistic guarantees that the system meets its specifications, and through synthesis we construct a strategy that maximizes the previous guarantees. Among such abstractions we find MDPs and their variants, like IMDPs.

This section is organized as follows. First, we introduce LTL in Section 2-1, as a way to formulate complex specifications. After that, we describe automata theory in Section 2-2 as the means to check if the behaviour a system satisfies a specification given as LTL. Next, we describe in detail Markov models such as MDPs and a generalisation of these, called IMDPs, in Sections 2-3 and 2-4 respectively. Furthermore, we tackle the problem of synthesizing a strategy that maximizes the probability that an IMDP satisfies a class of LTL formulas, called syntactically co-safe LTL (scLTL) formulas, in Section 2-4-1.

2-1 LTL

In this section we focus on temporal logic, which we will use in our approach to formulate complex specifications for our system. Linear temporal logic (LTL) is a logic formalism used in formal methods that allows to reason about the temporal behaviour of a system. This is achieved through formulas that include observations, temporal operators and logic operators, and such formulas allow to formulate complex specifications. An example of the behaviour that an LTL formula can describe is the following motion plan for a UAV, taken from [1]: “always avoid obstacles and visit regions `a`, `b`, `c` and `d` infinitely often", where `a`, `b`, `c` and `d` are regions in the physical space. In this section we define the syntax and semantics of these formulas, and focus on a subclass of LTL: scLTL. The theory that we introduce in this section is based on [6]. We next, we give a review of temporal logic and we highlight some useful applications.

Consider the Boolean operators “true", “and" and “negation", denoted as $\top$, $\wedge$ and $\neg$, respectively. Consider also the temporal operators “next" and “until", which are denoted as $\bigcirc$ and $\mathbb{U}$, respectively.

Definition 2-1.1. ***(LTL Syntax)*** *A linear temporal logic (LTL) formula ϕ defined over a set of observations O is recursively defined as:*

$$\phi = \top \mid o \mid \phi_1 \wedge \phi_2 \mid \neg\phi \mid \bigcirc \phi \mid \phi_1 \mathbb{U} \phi_2,$$

where o is an observation and ϕ , ϕ_1 and ϕ_2 are LTL formulas.

Using the previous rules, the logical operator “or" ($\vee$) and the additional temporal operators “eventually" ($\Diamond$) and “always" ($\Box$) are defined as

$$\begin{aligned} \phi_1 \vee \phi_2 &= \neg(\phi_1 \neg\phi_2) \\ \Diamond\phi &= \top\mathbb{U}\phi \\ \Box\phi &= \neg\Diamond\neg\phi. \end{aligned}$$

LTL formulas are interpreted over infinite words which are made of observations from some set O of observations. The LTL semantics are defined as follows:

Definition 2-1.2. ***(LTL Semantics)*** *The satisfaction of a formula ϕ over a set of observations O at position $k \in \mathbb{N}_+$ by word $w_O = w_O(1)w_O(2)w_O(3)\cdots \in O^\omega$, denoted by $w_O(k) \models \phi$, is defined recursively as follows [6]:*

- $w_O(k) \models \top$,
- $w_O(k) \models o$ *for some* $o \in O$ *if* $w_O(k) = o$,
- $w_O(k) \models \neg\phi$ *if* $w_O(k) \not\models \phi$,
- $w_O(k) \models \phi_1 \wedge \phi_2$ *if if* $w_O(k) \models \phi_2$ *and if* $w_O(k) \models \phi_2$,
- $w_O(k) \models \bigcirc\phi$ *if* $w_O(k+1) \models \phi$,
- $w_O(k) \models \phi_1\mathbb{U}\phi_2$ *if there exists* $j \geq k$ *such that* $w_O(j) \models \phi_2$ *and, for all* $k \leq i < j$*, we have* $w_O(i) \models \phi_1$.

A word w_O satisfies an LTL formula ϕ, denoted as $w_O \models \phi$, if $w_O(1) \models \phi$. The language of infinite words that satisfy formula ϕ is denoted by $\mathcal{L}_\phi$.

We next give an informal interpretation of the rules of LTL semantics for typical LTL formulas:

- $\bigcirc\phi$ is satisfied at at current step k if ϕ is satisfied at the “next" step $k+1$,
- $\phi_1\mathbb{U}\phi_2$ is satisfied at current step k if ϕ_1 is satisfied “until" ϕ_2 becomes satisfied,
- $\Box\phi$ is satisfied at current step k if ϕ is satisfied at each future step (this is, ϕ is “always" satisfied),
- $\Box\neg\phi$ is satisfied at current step k if $\neg\phi$ is satisfied at each future step k (this is, ϕ is “never" satisfied),
- $\Diamond\phi$ is satisfied at current step k if ϕ becomes satisfied at some future step (this is, ϕ is “eventually" satisfied),
- $\Diamond\Box\phi$ is satisfied at current step k if ϕ becomes satisfied at some future step and it remains satisfied for all the following steps (this is, ϕ is satisfied “eventually forever"),
- $\Box\Diamond\phi$ is satisfied at current step k if ϕ “always" becomes satisfied at some future step (this is, ϕ is satisfied “infinitely often").

Specifically, there exists a class of LTL formulas denoted syntactically co-safe LTL (scLTL) formulas in which only the operators $\bigcirc$, $\mathbb{U}$, $\Diamond$, $\wedge$, $\vee$ and $\neg$ are used. Furthermore, the negation $\neg$ operator only precedes observations (except for constructing the $\vee$ operator). Therefore, the “always" operator is not present in scLTL formulas. The syntax of scLTL formulas is defined as follows:

Definition 2-1.3. ***(scLTL Syntax)*** *A syntactically co-safe LTL (scLTL) formula ϕ defined over a set of observations O is recursively defined as:*

$$\phi = \top \mid o \mid \neg o \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2 \mid \bigcirc \phi \mid \phi_1\mathbb{U}\phi_2,$$

where o is an observation and ϕ , ϕ_1 and ϕ_2 are scLTL formulas.

This class of LTL is useful for the following reason: even though scLTL formulas are defined over infinite words, the satisfaction of one of such formulas by some word, if the latter satisfies the formula, always happens in finite time. This makes it easier to check if a word satisfies a formula, as we explain in Section 2-2. Furthermore, if an infinite word $w_O \in O^\omega$ satisfies an scLTL formula ϕ, then that word must contain a finite prefix $w_O(1)w_O(2)\ldots w_O(n) \in O^n$ called good prefix. Furthermore, any infinite word with this prefix satisfies ϕ. The language of all finite good prefixes of an scLTL formula ϕ is denoted $\mathcal{L}_{pref,\phi}$.

Back to our UAV example, we can define the specification stated in Section 1 as the scLTL formula (2-1) over the set of observations $O = \{\texttt{black}, \texttt{red}, \texttt{green}, \texttt{yellow}, \texttt{blue}\}$:

$$\phi = (\neg\texttt{black}\,\mathbb{U}\,\texttt{blue}) \wedge (\neg\texttt{blue}\,\mathbb{U}\,\texttt{red}) \wedge (\neg\texttt{blue}\,\mathbb{U}\,\texttt{green}) \wedge (\neg\texttt{blue}\,\mathbb{U}\,\texttt{yellow}). \tag{2-1}$$

Note that this one belongs to the class of scLTL formulas since it is constructed according to the rules described in Definition 2-1.3.

We now give a review of formal languages and highlight their applications. We can specify complex tasks through temporal logic formulas. Such tasks can be avoiding an undesired set of states, sequentially visiting different sets of states, transitioning between sets with the highest probability in probabilistic settings, etc, or combinations of these. Therefore, many authors have suggested the use of, for example, LTL and computation tree logic (CTL), to define the specifications for several problems in control, such as motion planning [8]. LTL formulas have been used to formulate specifications in multi-agent motion planning problems [1], and in verification and synthesis for hybrid systems [11], [4], for example. On the other hand, CTL can be used to check if all executions of the system satisfy some specification, or if there exists at least one that does. Since CTL is more complex than LTL, less intuitive, and the latter already captures a wide spectrum of properties, in this thesis we only use LTL formulas. Furthermore, despite CTL being more efficient than LTL for model checking (polynomial vs exponential time in the size of the formula), CTL formulas are longer than LTL ones when representing the same specification. For this reason, the performance of both is similar in practice [6].

Besides LTL, another classes of temporal logic exist, such as bounded LTL (BLTL), metric TL (MTL), propositional TL over the reals (RTL) and the already mentioned CTL. We can use both BLTL and MTL to specify time intervals on the temporal operators of the formulas. Such formulas allow us to formulate specifications such as “something eventually happens in a number of steps between 2 and 4" [6]. RTL, the continuous time version of LTL, is used in [17] in a motion planning problem. LTL also has probabilistic versions, such as probabilistic LTL (PLTL), in which the probability of satisfying the formula is stated. Furthermore, the semantics of a PLTL formula are defined with respect to a Markov decision process (MDP) or a variant of these models. Analogously, probabilistic CTL (PCTL) also exists. In [3], PCTL is used along with interval Markov decision processes for strategy synthesis and verification of stochastic systems.

Temporal logic formulas are commonly used in computer science, and allow us to reason about the temporal behaviour of discrete state systems. Examples of this kind of systems are transition systems and MDPs, and in later sections we will define the latter. In this thesis we focus on scLTL since we consider these formulas to be expressive enough, being able to describe, although finite, complex behaviours [4], [11]. Furthermore, synthesizing strategies that enforce such formulas is easier. Moreover, reasoning only over finite behaviours makes sense in highly uncertain systems [11]. Nevertheless, straightforward extensions to, for example, BLTL, exist [4], if specifying bounded-time properties was desired.

2-2 Automata

An automaton is a mathematical object that resembles a finite transition graph and which has a set of "special" states denoted accepting states. Its state evolves when the automaton receives observations as inputs. When these observations correspond to an input word, the evolution of the state of the automaton determines if the word is accepted or not. Due to this ability of automata, we can use them to check whether the paths of systems with an observation map, such as MDPs and their variants, satisfy temporal logic formulas. In this section we formally define automata. We focus on finite state automata but we also mention other classes like Büchi and Rabin automata and highlight their differences when it comes to representing an LTL formula. We also cite some off-the-shelf software tools for constructing automata from LTL formulas and give some applications.

Definition 2-2.1. ***(Finite State Automaton** [6]) A finite state automaton (FSA) is a tuple $\mathcal{A} = (Z,\ z_0,\ O,\ \Delta,\ Z_{ac})$, where*

- *Z is a finite set of states,*
- *z_0 is the initial state,*
- *O is the input alphabet,*
- *$\Delta : Z \times O \to Z$ is a transition function, and*
- *$Z_{ac} \subseteq Z$ is the set of accepting states.*

Now we give the intuition behind the operation of an automaton as in [18]. The automaton starts at the initial state z_0. Then, its state is forced to evolve by an input word, which is a sequence of observations from the input alphabet. The input word is read observation by observation from left to right. Therefore, the state of the automaton will change from z to $z' = \Delta(z, o)$ if it receives observation o. This process is then repeated by taking as an input the following observation of the input word. In the case that the current state q does not have outgoing transitions for the upcoming observation o, this is, $\Delta(z, o) = \emptyset$, the state of the automaton is stuck, and the word is said to be rejected. Once the automaton has read the complete word, the latter is said to be accepted if and only if the final state reached by the automaton belongs to the set of accepting states Z_{ac}. If this is not the case, the word also is rejected.

In the following we formally describe the functioning of an automaton as in [4]. A run of automaton $\mathcal{A}$ over a finite word $w_O = w_O(1)w_O(2)\dots w_O(n) \in O^n$ is a sequence $w_Z = w_Z(0)w_Z(1)\dots w_Z(n) \in Z^{n+1}$ where $w_Z(0) = z_0$ and $w_Z(k) = \Delta(w_Z(k-1),\ w_O(k))$ for all $k = 1,\ 2, \dots n$. The word w_O is said to be accepted by the automaton $\mathcal{A}$ if and only if the corresponding run ends in an accepting state, this is, $w_Z(n) \in Z_{ac}$. The set of all words accepted by $\mathcal{A}$ is called the language of $\mathcal{A}$, $\mathcal{L}_{\mathcal{A}}$. The FSA we have presented is deterministic. Nevertheless, non-deterministic finite state automata (NFA) also exist. In NFA, the transition function Δ can be non deterministic and the initial state can be a set of initial states instead. However, every NFA has an equivalent FSA representation, and therefore in this document we only focus on FSA.

To illustrate the previous theoretical concepts, consider again our UAV example. We can construct a deterministic FSA that captures the language of formula (2-1), which we show in Figure 2-1. As we can observe, the FSA has 8 states: $Z = \{z_1, z_2, \ldots, z_8\}$. The set of observations is $O = \{\texttt{red}, \texttt{green}, \texttt{yellow}, \texttt{blue}, \texttt{black}\}$. Furthermore, the accepting state is $Z_{ac} = z_8$. Finally, the transitions are represented by the arrows in the figure, with the observations that trigger said transitions near each arrow.

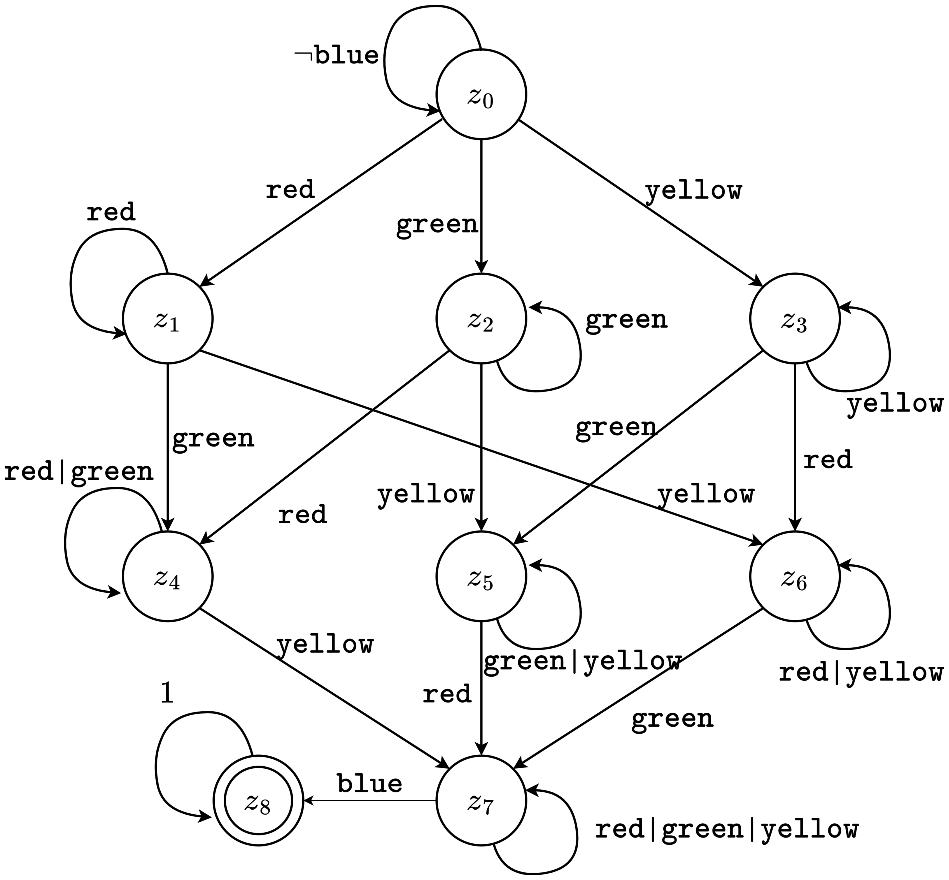


Figure 2-1: Deterministic FSA that captures the language of the scLTL specification in (2-1), for the UAV example.

In this section we have denoted by O the set of inputs of the automata. This is the same notation we used for the sets of observations we defined in Section 2-1 for LTL. Furthermore, it is also the same notation that we will use for the observations associated to MDPs and IMDPs in Sections 2-3 and 2-4, respectively. This is because the automata takes the traces from a system such as an MDP as an input, to check if they satisfy a temporal logic formula defined over that set of observations. Given an scLTL formula ϕ, a FSA can always be constructed such that it accepts all and only prefixes that satisfy ϕ, i.e., $\mathcal{L}_{\mathcal{A}} = \mathcal{L}_{pref,\phi}$ [6]. Nevertheless, this is not the case for general LTL formulas. On the contrary, for an arbitrary LTL formula, both a non-deterministic Büchi automaton and a deterministic Rabin automaton can always be constructed such that they accept all and only words from $\mathcal{L}_\phi$. There are slight differences between Büchi and Rabin automata and FSA, which make these more complex and lead to higher computational burden. Nevertheless, their semantics are defined over infinite words $w_O \in O^\omega$, which makes them able to consider infinite behaviours. We can translate an LTL formula ϕ to the corresponding type of automaton, FSA, Büchi or Rabin, by using off-the-shelf software tools like "LTL2BA", "LTL2DSTAR" and "SCHECK2". We do not give a formal definition of Büchi and Rabin automata, since in this thesis we will only make use of (deterministic) FSA. This is because we focus on specifications that can be given as scLTL formulas, which can be fully described by FSA, as we have already pointed out. For more information about Büchi and Rabin automata see [6].

To mention some applications of automata, in [11], FSA are used to check if the traces of an IMDP satisfy specifications given as scLTL formulas in a problem of strategy synthesis for unknown systems. Furthermore, FSA are also built from scLTL and BLTL formulas in [4] for strategy synthesis of IMDP abstractions of hybrid systems. Finally, Büchi automata are used in [1] together with transition systems in a decentralized motion planning problem.

2-3 MDPs

We have already highlighted the importance of finite-state system models in formal methods. In this thesis we need a class of models that is useful in stochastic settings. Imagine, for example, the problem of sending a command to a system through a malfunctioning communication channel that loses packets with given probability [18]. A kind of simple but powerful stochastic finite-state models are MDPs, which we describe in this section. In these ones, the actions or inputs to the MDP enable transitions between its states with given probability distributions [6]. In our example of the malfunctioning channel, the system could be modelled by an MDP in which the action "send command" has two possible effects. On the one hand it can make the system transition from the state "wait for command" to the state "command received" with probability p, which represents the possibility of the system receiving the command due to a correct functioning of the channel. On the other hand, the same action can make the system remain in the same state "wait for command" with probability $1-p$, which models a packet loss.

MDPs are able to model stochastic decision-making problems in a straightforward way. In this document we focus on labelled MDPs. These ones have an observation map associated to them, which assigns an observation to each state. In this section we formally define MDPs and the concepts of paths and strategies of an MDP. Furthermore, we give the intuition behind these formal definitions. We will also highlight the usefulness and implications of MDPs in abstractions of stochastic systems.

Definition 2-3.1. ***(Markov Decision Process** [4], [18], [19]) A Markov decision process (MDP) is defined as a tuple $\mathcal{M} = (Q, A, P, O, L)$ where:*

- *Q is a finite set of states,*
- *A is a finite set of actions,*

- $P : Q \times A \times Q \to [0, 1]$ *is the transition probability function such that for all states* $q \in Q$ *and actions* $a \in A$*,*

$$\sum_{q' \in Q} P(q, a, q') \in \{0, 1\},$$

- O *is a finite set of atomic propositions or observations,*
- $L : Q \to 2^O$ *is a labelling function or observation map that assigns to each state in* Q *a subset of atomic propositions in* O*.*

An action $a \in A$ is said to be enabled at state $q \in Q$ if and only if $\sum_{q' \in Q} P(q, a, q') = 1$. The set of actions enabled at q is denoted by $A(q) = \{a \in A \, : \, \sum_{q' \in Q} P(q, a, q') = 1\} \subseteq A$. This set contains the actions that can be chosen at each state, which are those that have outgoing transitions with nonzero probability. The transition function has the property of being a probability distribution over Q for given $q \in Q$ and $a \in A(q)$, i.e., $\sum_{q' \in Q} P(q, a, q') = 1$. Without any loss of generality, in this thesis we consider an arbitrary deterministic initial state $q_0 \in Q$. Furthermore, we only consider IMDPs for which the state-dependent action set is always non-empty: $A(q) \neq \emptyset$ for all $q \in Q$. The intuition behind the functioning of a MDP is the following. The MDP starts at the initial state q_0. An action is then chosen from the state-dependent action set $A(q)$, which assigns the transition probability $P(q, a, \cdot)$ that is a distribution over the state space. The following state q' is chosen from Q with probability $P(q, a, q')$. This process is repeated forever.

Now we define the notion of paths and traces of the MDP. A path of an MDP $\mathcal{M}$ is a sequence of states [4] $w_Q = q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \xrightarrow{a_2} \ldots$ such that $a_t \in A(q_t)$, and $P(q_t, a_t, q_{t+1}) > 0$ for all $t \in \mathbb{N}_{\geq 0}$. We denote a path of finite length as $w_Q^{\mathtt{fin}}$. Furthermore, a path of finite length $k+1$, for $k \in \mathbb{N}_{\geq 0}$ is denoted w_Q^k, and the sets of all infinite and finite paths are called Paths_Q and $\mathrm{Paths}_Q^{\mathtt{fin}}$ respectively. We also denote the last state of a finite path w_Q^k by $last(w_Q^k)$ and the $t+1$-th state of a path w_Q by $w_Q(t)$. A path $w_Q = q_0 q_1 q_2 \ldots$ produces a trace $w_O = o_0 o_1 o_2 \ldots \in O^\omega$ in $\mathcal{M}$ such that $o_t = L(q_t)$. The set of all infinite traces produced by paths of $\mathcal{M}$ is called the language of $\mathcal{M}$, $\mathcal{L}_{\mathcal{M}}$. This language has infinitely many words.

Next we define the notion of strategy of an MDP.

Definition 2-3.2. ***(Strategy of an MDP*** *[4]) A strategy* σ *of an MDP* $\mathcal{M}$ *is defined as a function* $\sigma : Paths_Q^{\mathtt{fin}} \to A$ *that maps a finite path* $w_Q^{\mathtt{fin}}$ *of* $\mathcal{M}$ *to an action* $a \in A(last(w_Q^{\mathtt{fin}}))$*. If the strategy only depends on the last state* $last(w_Q)$ *of the path, it is called a memoryless or Markovian strategy. Moreover, if the strategy is the same every time step it is called a stationary strategy. The set of all possible strategies is denoted by* Σ*.*

For a fixed strategy σ, MDP $\mathcal{M}$ becomes a Markov chain (MC) [19], with transition probabilities induced by σ. This means that given a strategy σ, a probability measure is induced over the set of all paths of the resulting Markov chain. Informally, a strategy, also known as policy or controller, is a function that depends on the finite path that the system has followed up to the current time step, this is, its history. At every time step, the strategy selects an action based on the history up to that time step, which makes the state evolve in time. When equipped with a strategy that makes the decision of selecting actions at each time step, an MDP becomes a Markov chain, which is an MDP for which only one action is available at each state. Therefore a Markov chain is a system that evolves in a stochastic fashion and in which no decision-making is contemplated. When the strategy is fixed, the paths of the resulting Markov chain have a uniquely defined probability associated to them.

Once we have introduced MDPs, we motivate their use in abstractions, and explain some features of MDP abstractions. Continuous-state stochastic systems can be abstracted to MDPs, by partitioning the original state space, which allows us to use tools from formal methods on these models. This is due to the finite-state nature of MDP models we described in Section 2-3. However, by performing the abstraction, we incur an abstraction error due to the discretization of the continuous state-space. If we compare the difference between the probability of the paths of the original system and those of the MDP abstraction satisfying a given specification, we can observe that this error grows with time [9]. Furthermore, this error is bigger the coarser is the discretization of the state space. This means that we can make the error as small as we want by making use of a finer discretization. However, since the error bound obtained in [9] is conservative, it is often found that in order to achieve an acceptable error, the number of states of the MDP abstraction needs to be huge. This is known as the state-explosion issue of MDP abstractions [4], and is translated into an increased difficulty for tools from formal methods to work with the abstraction. Therefore, we encounter a trade-off between precision of the abstraction and its computational complexity.

To effectively deal with the state-explosion issue, in [10], a tighter bound was found for the discretization error when abstracting continuous state systems to Markov chains. That same research proposed an adaptive refinement algorithm that takes into account the dynamics of the system and the geometry of the state-space partition. Furthermore, since no actions to the system were considered, the applications of this research were limited to verification. Furthermore, also to deal with the state-explosion issue, the use of IMDP abstractions was proposed in [3]. By using these, we do not incur a discretization error error even when we also discretize the state-space. Instead, we consider every possible behaviour of the real system to construct a conservative abstraction, which effectively mitigates the state explosion issue [3]. We introduce IMDPs in Section 2-4.

Now, let us give examples of applications of MDPs as abstractions. In [9], a stochastic hybrid system is abstracted to a Markov chain, which is used for verification in a probabilistic set invariance problem. The performance of the approach is assessed on a multi-room heating problem. In [8], an MDP abstraction is constructed in order to perform motion planning of a robot under specifications given as temporal logic. In this formulation, the transitions represent several-steps-ahead motions of the robot instead of one-step-ahead predictions, and each action corresponds to a feedback control primitive. The transition probabilities are computed using Monte Carlo sampling from simulations in which different sources of uncertainty, like sensor and actuator noise, are included.

2-4 IMDPs

In this section we describe a generalisation of the MDP formalism called interval Markov decision process (IMDP), which is particularly useful as an abstraction of a stochastic systems. An IMDP [4], also referred to as Bounded parameter Markov decision process (BMDP) [3], is a generalisation of an MDP in which the transition probabilities are not fixed, but instead lay inside of a bounded region. In this document, in the same way that we did with MDPs, we consider labelled IMDPs. IMDPs can be seen as a family of MDPs that share the same state space, action space, observation set and observation map, but which have different transition probabilities. This difference in the transition probabilities can represent uncertainty about the stochastic behaviour of the system.

As an example, consider again the problem of sending a command over a malfunctioning communication channel described in Section 2-3. Consider also that now we are uncertain about the probability of the channel malfunctioning. In this scenario we could consider a family of MDP models for this system which differ in the probability of transitioning from state “wait for command" to “command received", to account for the uncertainty about this probability. Furthermore, the transition probabilities of each MDP could lay on some bounded interval that represents the uncertainty of the system. This family of MDPs can be represented as an IMDP. Furthermore, we can then use the theory of IMDPs to perform analysis and synthesis tasks on this system. Consider that the performance of the system is defined through some criterion such as the probability of satisfying some specification. Then, we could, for example, determine if the performance of the channel is still satisfactory in the worst case possible, this is, if an MDP from the family of MDPs is picked such that the probability of a malfunction is the highest possible. This is a verification task, and verification of IMDP models is based on this worst-case analysis. Furthermore, we could also synthesize a strategy for the IMDP that maximizes the probability of satisfying the specification in a way that it is robust with respect to the uncertainty about its transition probabilities. This is done by computing the strategy that maximizes the satisfaction probability in the worst-case scenario. In our example, this happens when the probability of a malfunction in the channel is the highest possible.

In this section we formally define IMDPs and the concepts of paths, strategy and adversary of an IMDP. Furthermore, we give the intuition behind such formal definitions and the way an IMDP model works. After that, we give a review of the works related to IMDPs and IMDP abstractions, highlighting some of their applications. After that, we describe the procedure of synthesizing a strategy for an IMDP under specifications given as scLTL formulas in Section 2-4-1.

Let us begin by formally defining an IMDP.

Definition 2-4.1. ***(Interval Markov Decision Process** [4]) An interval Markov decision process (IMDP) is a tuple $\mathcal{I} = (Q, A, \underline{P}, \overline{P}, O, L)$ where:*

- *Q is a finite set of states,*
- *A is a finite set of actions.*
- *$\underline{P} : Q \times A \times Q \to [0, 1]$ is a function where $\underline{P}(q, a, q')$ represents a lower bound in the probability of transitioning from state q to state q' under action $a \in A$, such that for all states $q \in Q$ and actions $a \in A$, $\sum_{q' \in Q} \underline{P}(q, a, q') \leq 1$,*
- *$\overline{P} : Q \times A \times Q \to [0, 1]$ is a function where $\overline{P}(q, a, q')$ represents an upper bound in the probability of transitioning from state q to state q' under action $a \in A$, such that for all states $q \in Q$ and actions $a \in A$, $\sum_{q' \in Q} \overline{P}(q, a, q')$ is either 0 or ≥ 1,*
- *O is a finite set of atomic propositions or observations,*

- *$L : Q \to 2^O$ is a labeling function or observation map that assigns to each state in Q a subset of observations in O.*

For all $q, q' \in Q$ and $a \in A$, it holds that $\underline{P}(q, a, q') \leq \overline{P}(q, a, q')$. Furthermore, an action $a \in A$ is said to be enabled at state $q \in Q$ if and only if $\sum_{q' \in Q} \overline{P}(q, a, q') \neq 0$. We refer to the set of actions enabled at q as state-dependent action set and we denote it by $A(q) = \{a \in A : \sum_{q' \in Q} \overline{P}(q, a, q') \neq 0\} \subseteq A$. This set is equivalent to the one we defined for MDPs, and it contains the actions that can be chosen at each state. Such actions are those that have outgoing transitions leaving that state in the sense of the upper bound in the transition probabilities $\overline{P}$. That means that, for such actions, an outgoing transition leaving that state is possible for some MDP inside the IMDP. In this document we consider an arbitrary deterministic initial state $q_0 \in Q$. Furthermore, we only consider IMDPs for which the state-dependent action set is always non-empty: $A(q) \neq \emptyset$ for all $q \in Q$.

Furthermore, the transition probability bounds must satisfy that

$$\sum_{q' \in Q} \underline{P}(q, a, q') \leq 1 \leq \sum_{q' \in Q} \overline{P}(q, a, q')$$

for all $q \in Q$ and $a \in A(q)$. Let $\mathcal{D}(Q)$ be the set of discrete probability distributions over Q. Then, using the nomenclature that is typical in IMDPs [4], [20],[11], given state $q \in Q$ and action $a \in A(q)$, we say that $\gamma_{q,a} \in \mathcal{D}(Q)$ is a feasible probability distribution reachable from q by a if it fulfills:

$$\underline{P}(q, a, q') \leq \gamma_{q,a}(q') \leq \overline{P}(q, a, q')$$

for each state q'. We denote by $\Gamma_{q,a}$ the set of all feasible transition probabilities from $q \in Q$ by $a \in A(q)$:

$$\Gamma_{q,a} = \{\gamma_{q,a} \in \mathcal{D}(Q) : \underline{P}(q, a, q') \leq \gamma_{q,a}(q') \leq \overline{P}(q, a, q'), \text{ for all } q' \in Q\}, \tag{2-2}$$

for all $q \in Q$ by $a \in A(q)$. The nomenclature of “feasible" transition probabilities and set comes from the fact that set $\Gamma_{q,a}$ will appear as the feasible set of the optimization problems that we need to solve to synthesize strategies for IMDPs. This is a formal way to define the set of MDPs contained in the IMDP $\mathcal{I}$: such MDPs are those which, while sharing the same state and action spaces, observation set and observation map, have transition probabilities in the range defined by the lower and upper bounds $\underline{P}$ and $\overline{P}$.

The intuition behind the operation of an IMDP is quite similar to that of a MDP. The IMDP starts at the initial state q_0. At a current state q, an action is chosen from the state-dependent action set $A(q)$, which assigns the transition probability bounds $\underline{P}(q, a, \cdot)$ and $\overline{P}(q, a, \cdot)$ over the state space. Then, a feasible transition probability distribution γ_q^a over the state space Q is selected in a non-deterministic fashion from the set of feasible transition probability distributions Γ_q^a. After that, the following state q' is chosen with probability $\gamma_q^a(q')$. This process is repeated forever.

Now we define the notions of paths and strategy of an IMDP, which are analogous to those of MDPs. A path of an IMDP $\mathcal{I}$ is a sequence of states [11] $w_Q = q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \xrightarrow{a_2} \dots$ such that $a_t \in A(q_t)$, and $\overline{P}(q_t, a_t, q_{t+1}) > 0$ for all $t \in \mathbb{N}_{\geq 0}$. We denote a path of finite length as $w_Q^{\texttt{fin}}$. Furthermore, a path of finite length $k+1$, for $k \in \mathbb{N}_{\geq 0}$ is denoted w_Q^k, and the sets of all infinite and finite paths Paths_Q and $\text{Paths}_Q^{\texttt{fin}}$ respectively. We also denote the last state of a finite path w_Q^k by $last(w_Q^k)$ and the $t+1$-th state of a path w_Q by $w_Q(t)$. Now, a transition of $\mathcal{I}$ is feasible if and only if the upper bound $\overline{P}(q_t, a_t, q_{t+1})$ is bigger than zero, this is, there is a MDP $\mathcal{M}$ inside $\mathcal{I}$ for which this transition has nonzero probability of happening. Additionally, a strategy σ of an IMDP $\mathcal{I}$ is the same as that explained in Definition 2-3.2 for MDPs.

As we have already pointed out, in addition to the stochastic behaviour that is also present in MDPs, IMDPs are in some sense non-deterministic models. This is because the transition probabilities are not fixed as in MDPs, but uncertain, and we have no further information about how likely it is that any feasible probability distribution is chosen in the end [21]. Therefore, it is necessary that we define an additional ingredient that selects feasible transition probabilities and which is not present in MDPs. This new element is known as the adversary.

Definition 2-4.2. ***((Adversary of an IMDP[4])*** *Given an IMDP $\mathcal{I}$, an adversary is a function $\pi : paths^{\texttt{fin}} \times A \to D(Q)$ that, for each finite path $w_Q^{\texttt{fin}} \in Paths^{\texttt{fin}}$ and action $a \in A(last(w_Q^{\texttt{fin}}))$, assigns a feasible probability distribution $\gamma_q^a \in \Gamma_q^a$, where $q = last(w_Q^{\texttt{fin}})$. The set of all adversaries is denoted by* Π.

Adversaries choose, at each time step, a feasible probability distribution, which defines the probability of each transition. It can be seen as a means to pick one of the MDPs that form the IMDP $\mathcal{I}$ at each time step. Given a strategy σ and an adversary π for the IMDP $\mathcal{I}$, the IMDP reduces to a Markov chain, and a probability measure over its set of paths is induced. As we already explained in Section 2-3, a Markov chain is an MDP for which only one action is available at each state. This makes it a stochastic process in which no decision-making is contemplated.

IMDPs were introduced in [21], where their usefulness in aggregate schemes was already pointed out. With these models, we can capture the variation in the transition probabilities for different base states which are aggregated together in the same aggregated state can be captured. In that work, a modified value iteration algorithm for strategy synthesis was proposed, using the Expected Total Discounted Reward Criterion, as it is typical in MDPs [19]. We should highlight the efficiency of this algorithm. These results were extended in [22] to the undiscounted case, which is needed to formulate reachability problems and, ultimately, to perform synthesis under complex specifications.

In [3], abstractions to IMDPs were used as an alternative to MDP abstractions, since the former do not incur any discretization error. In this work, verification and synthesis algorithms were proposed for switching linear systems with additive noise under specifications given as Probabilistic Computation Tree Logic formulas. Furthermore, a local refinement scheme was introduced to reduce conservatism. In the previous research, however, little attention was devoted to the development of tractable algorithms of computing the abstractions. This, on the contrary, was shown to be the most computationally expensive part. Concerning this issue, in [4] an efficient procedure of computing IMDP abstractions from stochastic hybrid systems was proposed, based on convex optimization. The approach assumed that the continuous dynamics were, for each action, linear in the state and in the noise, being Gaussian the latter. Furthermore, in [5], an efficient approach to compute IMDP abstractions for Neural network dynamic models (NNDMs) with additive Gaussian noise is proposed. These abstractions are later used to synthesize switching strategies under specifications given as scLTL. This approach heavily relies on the results of [4] and also in those related to verification of neural networks [23], [24]. IMDP abstractions have also found applications in event-triggered control. in [12], the sampling behaviour of periodic event-triggered Control systems is analyzed by making use of interval Markov chains, since these abstractions allow to compute bounds on the sampling performance indicators of the system.

In the previous years, interest for verification of unknown systems has also arisen. In [20], verification of systems whose dynamics are unknown but real data from such systems is available is treated. The proposed framework makes use of data-driven IMDP abstractions of Gaussian Process regression models and derives formal guarantees for the unknown system. This approach is applied to systems with switching dynamics. This research is further developed in [11], where the problem of strategy synthesis for switched systems is tackled. Opposite to the framework of [11], in our data-driven approach we assume full knowledge of the dynamics of the system, except for the disturbance. Therefore, we expect to exploit this partial knowledge about the system dynamics, instead of modelling it entirely or partially as a Gaussian Process: by using our approach that relies on tools from optimal transport and DRO, we aim to obtain a less conservative abstraction, and tighter bounds in the performance guarantees.

Also in the data-driven setting, we must highlight the research carried out in [25], since the setting of that work is the closest to ours. That research tackles the problem of synthesizing strategies for systems with additive i.i.d. disturbances with unknown probability distribution by making use of IMDP abstractions. The approach uses tools from the scenario approach that allow to leverage information from samples of the disturbance. Using this approach, the transition probability bounds of the IMDP are computed with some confidence. The approach proposed in that research are applied to several scenarios: a UAV motion planning problem, a building temperature control and a spacecraft rendezvous problem. Opposite to that approach, we make use of tools from optimal transport and ambiguity sets to account for the distributional ambiguity. This has a remarkable advantage with respect to the scenario approach: by using the Wasserstein distance we take into account the distances between samples and sets to compute the bounds in the transition probabilities. On the other hand, the scenario approach computes the bounds just by taking into account the number of samples that lay inside the sets. For example, if we consider that the samples fall outside two cells q and q', the scenario approach assigns the same upper bound to both cells, irrespective to their distance to the samples. This is counter-intuitive, and accounting for these distances might lead to tighter bounds, or might allow us to construct a practical abstraction when the number of samples is small. Note that in [25], the number of samples required to construct the abstraction is always very high. However, using tools from optimal transport and DRO has the following drawback: we need to determine the size of the ambiguity set that we use to account for the distributional ambiguity which, in this thesis, we assume given. A second key difference between the approach proposed in [25] and our approach is that, in order to compute the transition probability bounds, the approach of [25] relies on a reachability analysis. However, this one can be computationally expensive, and it is not a trivial problem when the system dynamics are nonlinear. On the contrary, we do not rely on reachability computations: instead, we define the set of actions of the abstractions as a finite partition of the set of control inputs of the continuous system. Furthermore, our abstractions account for both the distributional robustness and the partitioning of the state-space, which allows us to easily deal with nonlinear systems.

On the other hand, beyond IMDPs, several works have studied the problem of synthesizing strategies for Markov models with sets of transition probabilities more complex than intervals, characteristic of IMDPs. In [26] the IMDP model was generalized. This work considers transition probabilities that belong to convex sets. Moreover, convex optimization methods are proposed to perform verification with respect to probabilistic computational tree logic (PCTL) specifications. The approach is illustrated in a ZeroConf Dynamic Configuration Protocol problem for IPv4 Link-Local Addresses. Furthermore, in this example a convex MDP (CMDP) is obtained as an abstraction of the real system, using a likelihood-based uncertainty set estimated from data, with some confidence level. This work is further generalized in [27], in which the case that the transition probabilities may belong to several, even nonconvex, sets. The framework of that work is that of robust stochastic dynamic programming using discounted rewards. Another work that uses Markov models whose transition probabilities belong to ambiguity sets is that in [28]. In this one, algorithms for strategy synthesis for these models under LTL specifications are proposed. However, to the best of the knowledge of the author of this document, none of these robust MDP formulations have been used as formal abstractions of continuous-state stochastic systems.

Related to MDP abstractions with uncertain transition probabilities, is the research in [29]. In that work, a Wasserstein distance-based distributionally robust approach for MDP abstractions with parameter ambiguity is presented. First, an approximate MDP abstraction is constructed, and then its transition probability is treated as an ambiguous parameter governed by an uncertain probability distribution. The results of the approach show improved robustness with respect to ambiguity in the transition probabilities. However, since the abstraction is just an approximation of the original system, no formal guarantees regarding the original system can be obtained from this approach. Opposite to that research, we are formal: we propose actual approaches to build correct Markovian abstractions from continuous-state systems, and we treat the transition probabilities of our abstraction as actual probabilities. These belong to a given ambiguity set, defined using the Wasserstein distance. This means that the guarantees we obtain are formal, and also hold for the original system. Furthermore, we consider complex specifications given as scLTL formulas.

2-4-1 Strategy Synthesis for IMDPs

The process of obtaining a strategy that enforces some specification is called synthesis, or strategy synthesis. Back to our UAV example, consider that an IMDP abstraction of this one is available. Moreover, consider that our objective is that of synthesizing a strategy that maximizes the probability of satisfying some specification given as an scLTL formula, while being robust to the uncertainties of the IMDP model. The specification can be the one we considered in Section 1-1: “eventually reach `red`, `yellow` and `green` in no particular order, and then eventually reach `blue` while avoiding `black` throughout all the trajectory", which can be represented by scLTL formula (2-1). Typically, we are interested in the motions of the UAV that remain in the bounded set $X \subset \mathbb{R}^n$. To embed into the IMDP model the additional specification “while remaining in set X throughout all the trajectory", we make the state q_u that represents region X, absorbing. In this way, we exclude the paths of the IMDP that exit X, since these will remain there forever, not satisfying the specification.

In this section we describe the process of synthesizing a strategy for an IMDP $\mathcal{I}$ that maximizes the probability of satisfying a complex specification given as an scLTL formula ϕ, while being robust to all uncertainties in the model. This process can be seen as a two-player stochastic game: at each time step, player one, the strategy, chooses an action to maximize the probability of satisfaction of ϕ, while player two, the adversary, selects the transition probabilities to minimize said probability. The approach that we describe in this section, consists in reformulating the synthesis problem as a maximal reachability probability problem over a different IMDP: the one obtained by taking the product between the IMDP model and a FSA that captures the language of ϕ. Then we solve the maximal reachability probability problem via value iteration [30]. We must highlight that strategy synthesis, when compared to the abstraction process, is way more efficient.

We start denoting by $P(w_Q^k \models \phi | w_Q^k(0) = q, X, \sigma_{\mathcal{I}}, \pi_{\mathcal{I}})$ the probability that the paths of IMDP $\mathcal{I}$ satisfy the scLTL formula ϕ within $k \geq 0$ steps, while never exiting set X, when the initial state is q, the memory-dependent strategy $\sigma_{\mathcal{I}}$ is followed and the transition probabilities are chosen by adversary $\pi_{\mathcal{I}}$. After that, as described in [4], we construct a deterministic FSA $\mathcal{A} = (Z,\ z_0,\ O,\ \Delta,\ Z_{ac})$ such that it captures the language of ϕ as explained in Section 2-2. Then, an additional IMDP $\mathcal{I}_\phi$ is constructed by taking the product of the initial IMDP $\mathcal{I}$ with FSA $\mathcal{A}$:

Definition 2-4.3. *(**Product IMDP** [4]) Given an IMDP $\mathcal{I}$ as in Definition 2-4.1 and a FSA $\mathcal{A}$ as in Definition 2-2.1, the product IMDP is another IMDP $\mathcal{I}_\phi = \mathcal{I} \times \mathcal{A}$ defined as the tuple $\mathcal{I}_\phi = (Q_\phi, A_\phi, \underline{P}_\phi, \overline{P}_\phi, Q_{\phi,ac})$ where:*

- *The set of states Q_ϕ is defined as $Q_\phi = Q \times Z$,*
- *The set of actions A_ϕ is defined as $A_\phi = A$,*
- *The lower and upper bounds, $\underline{P}_\phi$ and $\overline{P}_\phi$ respectively, of the transition probabilities are defined as:*

$$\begin{aligned} \underline{P}_\phi((q,z),a,(q',z')) &= \begin{cases} \underline{P}(q,a,q') & \textit{if } z' = \Delta(z, L(q')), \\ 0 & \textit{otherwise} \end{cases} \\ \overline{P}_\phi((q,z),a,(q',z')) &= \begin{cases} \overline{P}(q,a,q') & \textit{if } z' = \Delta(z, L(q')), \\ 0 & \textit{otherwise,} \end{cases} \end{aligned} \tag{2-3}$$

  for all $(q,z),(q',z') \in Q_\phi$, $a \in A_\phi$.

- *The set of accepting states $Q_{\phi,ac}$ is defined as $Q_{\phi,ac} = Q \times Z_{ac}$,*

The set of accepting states, $Q_{\phi,ac}$ of $\mathcal{I}_\phi$ correspond to the set of accepting states of $\mathcal{A}$. Intuitively, the transition probability bounds of $\mathcal{I}_\phi$ are defined by only taking into account transitions between states of $\mathcal{I}_\phi$ that generate transitions in $\mathcal{A}$. In this way, we only consider the paths of $\mathcal{I}_\phi$ that are able to produce an accepting run in $\mathcal{A}$. In this product IMDP, the adversary can pick transition probabilities from the following set:

$$\Gamma_{q_\phi}^a = \{\gamma_{q_\phi}^a \in \mathcal{D}(Q_\phi) : \underline{P}_\phi(q_\phi, a, {q_\phi}') \leq \gamma_{q_\phi}^a({q_\phi}') \leq \overline{P}_\phi(q_\phi, a, {q_\phi}'),\ \text{for all}\ {q_\phi}' \in Q_\phi\},$$

for all $q_\phi \in Q_\phi$, where $\mathcal{D}(Q_\phi)$ is the set of probability measures over Q_ϕ.

We must highlight that, from the definition of the FSA $\mathcal{A}$, a path of IMDP $\mathcal{I}$ satisfies ϕ if and only if its trace produces an accepting run in $\mathcal{A}$. Therefore, it becomes clear that we are interested in the paths of $\mathcal{I}_\phi$ that reach $Q_{\phi,ac}$. For this reason, solving the strategy synthesis problem for $\mathcal{I}$ is the same as finding the strategy σ_ϕ that maximizes the probability of the paths of $\mathcal{I}_\phi$ reaching $Q_{\phi,\texttt{ac}}$, in the two-player game point of view we already described. This is known as maximal reachability probability problem [11], and can be solved via value iteration [30], as we describe in Section 2-4-1.

Maximal Reachability Probability Problem for IMDPs

The maximal reachability probability problem, defined for MDPs and its variants [30], allows us to synthesize a strategy that maximizes the probability of reaching a target set of states. As we described in Section 2-4-1, this problem also arises in strategy synthesis problems under specifications described as scLTL formulas. This is because such problems can be formulated as maximal reachability probability problems. In this section we explain how to solve such problems and describe the characteristics of the solution.

We formulate the maximal reachability problem for an IMDP $\mathcal{I} = (Q, A, \underline{P}, \overline{P}, O, L)$ and target set $Q_{\texttt{tgt}} \subset Q$. To translate this setting to that of strategy synthesis for an scLTL specification ϕ, it suffices to consider $\mathcal{I}_\phi$ and $Q_{\phi,\texttt{ac}}$, as defined in Section 2-4-1, instead of $\mathcal{I}$ and $Q_{\texttt{tgt}}$. We start by considering the worst and best-case probabilities of the paths of $\mathcal{I}$, under strategy $\sigma \in \Sigma$, reaching $Q_{\texttt{tgt}}$ within k steps by starting on $q \in Q$:

$$\begin{aligned}&\min_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma, \pi)\\ &\max_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma, \pi)\end{aligned} \tag{2-4}$$

Then, we define the optimal strategy $\sigma^* \in \Sigma$ as the one that maximizes the first expression in (2-4), this is, the strategy that, maximizes the lower bound in the probability of reachability:

$$\sigma^*(q) := \arg\max_{\sigma\in\Sigma} \min_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma, \pi), \tag{2-5}$$

for all $q \in Q,\ k \in \mathbb{N} \cup \{\infty\}$. This is a pessimistic way of defining the optimal strategy, which corresponds to the two-player game we described in Section 2-4-1. An alternative optimistic criterion, in which σ^* is defined as the strategy that maximizes the upper bound in (2-4) also exists [30]. However, in our setting of formal strategy synthesis we are interested in obtaining the strategy that maximizes the worst-case probability that the abstraction satisfies the specification. This is because this lower bound is the performance guarantee that we are looking for [4] and, therefore we want it to be as high as possible. Let us denote by $\underline{p}^k$ and $\overline{p}^k$, respectively, the worst and best-case probabilities of the paths of $\mathcal{I}$ reaching $Q_{\texttt{tgt}}$ within k steps under strategy $\sigma^* \in \Sigma$:

$$\begin{aligned}\underline{p}^k(q) &:= \max_{\sigma\in\Sigma} \min_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma, \pi)\\ \overline{p}^k(q) &:= \max_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma^*, \pi)\end{aligned} \tag{2-6}$$

for all $q \in Q,\ k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$. Now we state Theorem 2-4.1, which allows us to obtain the probabilities in (2-6) and the strategy in (2-5).

Theorem 2-4.1. ***(Interval Value Iteration)*** *Consider the IMDP $\mathcal{I}$. Then, probabilities $\underline{p}^k$ in* (2-6) *are obtained for all $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ recursively, starting from $\underline{p}^0(q) = 1$ for all $q \in Q_{\texttt{tgt}}$ and $\underline{p}^0(q) = 0$ otherwise:*

$$\underline{p}^{k+1}(q) = \begin{cases} 1 & \textit{if}\, q \in Q_{\texttt{tgt}} \\ \max_{a \in A(q)} \min_{\gamma_{q,a} \in \Gamma_{q,a}} \sum_{q' \in Q} \gamma_{q,a}(q') \underline{p}^k(q') & \textit{otherwise} \end{cases}. \tag{2-7}$$

Furthermore, with a small abuse of notation, strategy σ^* in (2-5) *is the Markovian, but time-dependent strategy $\sigma^* = \{\sigma^*(\cdot; k)\}_{k=1}^{\infty}$ that fulfills:*

$$\sigma^*(q; k+1) = \arg \max_{a \in A(q)} \min_{\gamma_{q,a} \in \Gamma_{q,a}} \sum_{q' \in Q} \gamma_{q,a}(q') \underline{p}^k(q'), \tag{2-8}$$

for all $q \in Q$, $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$. Using σ^*, probabilities $\overline{p}^k$ in (2-6) *are also obtained for all $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ recursively, starting from $\overline{p}^0(q) = 1$ for all $q \in Q_{\texttt{tgt}}$ and $\overline{p}^0(q) = 0$ otherwise:*

$$\overline{p}^{k+1}(q) = \begin{cases} 1 & \textit{if}\, q \in Q_{\texttt{tgt}} \\ \max_{\gamma_{q,\sigma^*(q;k)} \in \Gamma_{q,\sigma^*(q;k)}} \sum_{q' \in Q} \gamma_{q,\sigma^*(q;k)}(q') \overline{p}^k(q') & \textit{otherwise} \end{cases}. \tag{2-9}$$

Proof. The proof follows a similar reasoning that the one presented in [30] for stationary strategies. It is easy to prove that the probabilities $\underline{p}^k(q)$ and $\overline{p}^k(q)$ correspond to the probabilities defined in (2-6), for all $q \in Q$, $k \in \mathbb{N}_{\geq 0}$. Furthermore, since they are monotonically increasing and upper bounded by 1 for all $q \in Q$, sequences $\underline{p}^k(q)$ and $\overline{p}^k(q)$ converge to the following fixed points of recursions (2-7) and (2-9), respectively:

$$\begin{aligned} \underline{p}(q) &= \lim_{k \to \infty} \underline{p}^k(q) = \max_{\sigma \in \Sigma} \min_{\pi \in \Pi} P(\exists k \in \mathbb{N}_{\geq 0} \text{ s.t. } w_Q(k) \in Q_{\texttt{tgt}} | w_Q(0) = q, \sigma, \pi) \\ \overline{p}(q) &= \lim_{k \to \infty} \overline{p}^k(q) = \max_{\pi \in \Pi} P(\exists k \in \mathbb{N}_{\geq 0} \text{ s.t. } w_Q(k) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma^*, \pi) \end{aligned} \tag{2-10}$$

for all $q \in Q$. The previous expressions are the probabilities of the paths of infinite length of $\mathcal{I}$ ever reaching $Q_{\texttt{tgt}}$ under σ^*. Furthermore

$$\lim_{k \to \infty} \sigma^*(q; k) = \sigma^*_{\texttt{stat}}(q),$$

for all $q \in Q$. Notice that strategy $\sigma^*_{\texttt{stat}}$ is stationary, because strategy $\sigma^*(\cdot; k)$ only depends on $\underline{p}^{k-1}$, as shown in (2-8), and the latter converges for $k \to \infty$. □

Notice that, due to Bellman's optimality principle [19], strategy σ^* is memoryless, since it does not depend on the whole path w_Q^k of $\mathcal{I}$, but only on $last(w_Q^k)$. Furthermore, it is the same strategy for all initial conditions $q \in Q$. However, this strategy is also time-dependent, since it depends on current value function $\underline{p}^k$, which changes over time.

Now, let us discuss the computational features of the value iteration algorithm (2-7), (2-9). First of all, we highlight that, since iteration (2-7) is independent of $\overline{p}^k$ for all $k \geq 0$, we can compute this recursion, and strategy σ^* first. Once we have achieved convergence we can compute the second recursion (2-9) and obtain $\overline{p}^k$, which is useful to analyze the error of the solution [4]. Secondly, notice that computing, for example, $\underline{p}^{k+1}$, involves solving an optimization problem over $\Gamma_{q,a}$ for each state $q \in Q$ and action $a \in A(q)$. In these problems the decision variables are the transition probabilities. Since these appear linearly in (2-7) and set $\Gamma_{q,a}$ is described by linear constraints, as we showed in (2-2), said problems are linear programs (LPs), which can be solved using off-the-shelf LP solvers. However, there is a alternative way to solve such problems by taking into account their structure [21]: we begin by ordering the states $q \in Q$ according to the value of $\underline{p}^k$ at each state q in an increasing fashion, this is, the first states in the ordering are those with the lowest value. After that the transition probabilities that minimize the output of (2-7) are found as the ones that assign as much probability as possible to the first states in the ordering. In the case of the maximization in (2-9) problem, the ordering should be the opposite. This procedure is way more efficient than solving the linear program with the standard LP solvers.

Results of Strategy Synthesis for IMDPs Under scLTL Specifications

As a starting point, consider that we have solved the maximal reachability problem described in Section 2-4-1 for the product IMDP $\mathcal{I}_\phi$ and target set $Q_{\phi,ac}$, defined in Section 2-4-1. As a result, we have obtained the upper and lower bounds in probability $\underline{p}^k_\phi$ and $\overline{p}^k_\phi$ of the paths of $\mathcal{I}_\phi$ reaching $Q_{\phi,ac}$. Furthermore, we have obtained the strategy σ^*_ϕ of $\mathcal{I}_\phi$ that maximizes the previous lower bound at each time step. In this section we describe how we can translate those results to the strategy and guarantees of the IMDP $\mathcal{I}$ that we seek. In fact, the results we describe hold for any Markovian strategy σ_ϕ of $\mathcal{I}_\phi$, and when $\underline{p}^k_\phi$ and $\overline{p}^k_\phi$ have been obtained by solving (2-7) and (2-9) for fixed strategy σ_ϕ.

First, we prove that any Markovian, and possibly time-dependent strategy σ_ϕ in $\mathcal{I}_\phi$ maps to a history dependent strategy $\sigma_\mathcal{I}$ in the initial IMDP abstraction $\mathcal{I}$.

Lemma 2-4.1. ***(Translating a Markovian Strategy of the Product IMDP to the IMDP**[11]**)*** *Consider the product IMDP $\mathcal{I}_\phi$ obtained by taking the product between IMDP $\mathcal{I}$ and FSA $\mathcal{A}$, where the latter captures the language of the scLTL formula ϕ. Consider also the Markovian strategy σ_ϕ of $\mathcal{I}_\phi$. Then, strategy σ_ϕ can be translated to a memory-dependent strategy $\sigma_\mathcal{I}$ of $\mathcal{I}$.*

Proof. We first express the state z of FSA $\mathcal{A}$ at time step t as a function of of the finite paths w^t_Q of $\mathcal{I}$:

$$\begin{aligned} z(t) &= \Delta(z(t-1), L(q(t))) \\ &= \Delta(\Delta(z(t-2), L(q(t-1)), L(q(t)) \\ &= f(z_0, L(q(1)), L(q(2)), \dots, L(q(t)) \\ &= \widetilde{f}(z_0, w^t_Q), \end{aligned} \tag{2-11}$$

where $f(\cdot)$ and $\widetilde{f}(\cdot)$ simply denote dependence on the argument. Using Expression (2-11), we find that the corresponding strategy of $\mathcal{I}$ is $\sigma_\mathcal{I} = \{\sigma_\mathcal{I}(w^t_Q; k)\}^\infty_{k=1}$, where:

$$\sigma_\mathcal{I}(w^t_Q; k) = \sigma_\phi(q_\phi(t); k) = \sigma_\phi((q(t), z(t)); k) = \sigma_\phi((q(t), \widetilde{f}(z_0, w^t_Q)); k), \tag{2-12}$$

where k is the number of time steps until the end of the horizon. which is a memory-dependent strategy. □

Secondly, we relate the bounds in the probability of $\mathcal{I}_\phi$ reaching $Q_{\phi,ac}$ to those in the probability of $\mathcal{I}$ satisfying the specification ϕ.

Lemma 2-4.2. *(**Guarantees Of the IMDP Abstraction** [4],[11]) Consider the strategy $\sigma^*_{\mathcal{I}}$ of $\mathcal{I}$ obtained from σ^*_ϕ as described in Lemma 2-4.1. Then it holds that the bounds in the probability of the paths of $\mathcal{I}$ satisfying ϕ within k steps while never exiting X by following strategy $\sigma^*_{\mathcal{I}}$ and starting from state $q \in Q$ are:*

$$\underline{p}^k_{\mathcal{I}}(q) := \inf_{\pi\in\Pi} P(w^k_Q \models \phi | w^k_Q(0) = q, X, \sigma^*_{\mathcal{I}}, \pi) = \underline{p}^k_\phi(q, z_0) \tag{2-13}$$

$$\overline{p}^k_{\mathcal{I}}(q) := \sup_{\pi\in\Pi} P(w^k_Q \models \phi | w^k_Q(0) = q, X, \sigma^*_{\mathcal{I}}, \pi) = \overline{p}^k_\phi(q, z_0) \tag{2-14}$$

for all $q \in Q$.

Notice that, in practice, we are only interested in the probabilities obtained with initial state z_0 of $\mathcal{A}$, since the runs of the automaton always start at its initial state z_0. The probability bounds in (2-13) are the guarantees of the IMDP abstraction satisfying the specification we were looking for. Furthermore, we have synthesized the strategy $\sigma_{\mathcal{I}^*}$ that maximizes the lower bound. The complexity of the interval value iteration algorithm when used on $\mathcal{I}_\phi$ is polynomial in the number of states of $\mathcal{I}_\phi$ and exponential in the size of ϕ in the worst case. Additionally, for infinite horizon $k \to \infty$, although $\sigma^*_\phi(\cdot; k)$ converges to a stationary strategy for $\mathcal{I}_\phi$ [30], $\sigma^*_{\mathcal{I}}(\cdot; k)$ is still a memory-dependent strategy for $\mathcal{I}$ [4].

Chapter 3

Background on Optimal Transport and Ambiguity Sets

Continuous-state systems can be abstracted to IMDPs so that we can use tools from formal methods on such systems [3], [4],[11]. However, the approaches proposed in those researches assume that we have perfect knowledge of the random terms in the dynamics of our system. An instance in which this is not the case is our UAV system, if we only have an approximate description of the disturbance obtained from samples of this one. In this uncertain setting we choose to include the distributional ambiguity into the abstraction, which we denote by “robust" abstraction. To build such robust abstractions, we use tools from optimal transport and ambiguity sets, which we introduce in this chapter. Particularly, we focus on defining ambiguity sets using the Wasserstein distance, and in solving distributionally robust (DR) uncertainty quantification problems.

This Chapter is organised as follows: in Section 3-1 we give a small survey of optimal transport. In Section 3-1 we describe Kantorovich' formulation of the optimal transport problem, and we define the Wasserstein distance. In Section 3-2 we introduce ambiguity sets and we highlight their usefulness to represent uncertainty about the probability distribution of a random variable. In Section 3-2, we discuss tractable reformulations of distributionally robust uncertainty quantification problems.

3-1 Optimal Transport

In this section we briefly introduce optimal transport, which is necessary to formulate the Wasserstein distance. This distance is, in turn, necessary to define a class of ambiguity sets that account for distributional ambiguity. First, we briefly describe optimal transport. After that, in Section 3-1 we introduce Kantorovich' formulation of the optimal transport problem, and relate this formulation to the definition of the Wasserstein distance. Finally, we describe the dual reformulation of this distance.

Optimal transportation theory is the discipline that deals with the transportation of resources between two different configurations. This theory includes problems such as finding the least expensive way to transport a group of objects from one configuration to another. A feature of this problem is that the cost associated to moving each object from the first configuration to each spot in the second configuration is generally different. Optimal transport has applications in logistics, transportation networks, economics, etc. Furthermore, it also has applications in statistics: we might seek the cheapest way of moving the probability mass from one measure to another, when moving each unit of mass between two points in space incurs a cost. Furthermore, optimal transport allows us to define distances in probability spaces such as the Wasserstein distance, that quantify how "far away" two probability measures are from each other. Such distances have applications in image processing, machine learning and distributionally robust optimization (DRO), where the last one is of particular interest for this thesis.

Traditionally, solutions to the optimal transport problem were proposed as assignment problems: each element of mass in the first measure is assigned to one element in the support of the second measure. This was achieved, initially, by using permutations, and after that through Monge maps [31], which allowed for more generality. However, both approaches are difficult to solve. This is because the the first approach we described has a combinatorial nature, and in the second one the feasible set of the problem is non-convex. Additionally, both approaches are not able to provide feasible solutions in the case that the weight vectors of each measure are not compatible. This implies that, when comparing discrete probability measures, the mass from one location of one measure is not allowed to be split to several different locations of the second measure [31]. For example, consider the problem of transporting probability mass from a Dirac measure to a finite number of different locations. Using the previous approaches, we would not be able to solve this problem, since we would not be able to split the mass from the original location to the desired ones.

A huge development in of optimal transport was made by the work of Kantorovich, who proposed to tackle the mass transport problem in a “probabilistic" way, which made mass splitting possible [32]. His idea was to use couplings to represent the possible assignment of mass from one location to several others. This formulation allows us to split mass to more than one destination, which the previous formulations did not. Furthermore, this formulation results in a linear program, as we describe in Section 3-1.

Kantorovich’ Formulation

Here we describe Kantorovich’ formulation of the optimal transport problem between arbitrary probability measures P and P'. In this formulation, the optimal transport cost is defined as the minimum transport cost between these measures that can be achieved by using a feasible coupling. This one is a joint measure whose marginals are equal to the probability measures P and P'. We also use this formulation to define the Wasserstein distance, which we will, in turn, use later to define ambiguity sets of probability distributions. Additionally, we state the dual definition of the Wasserstein distance.

First, we define the concept of a push-forward measure, which is needed to formulate the optimal transport problem in Kantorovich’ formulation.

Definition 3-1.1. *(**Push-forward Measure**[15], [31]) Consider the measurable spaces Ξ and Ξ'. Consider a probability measure $P \in \mathcal{P}(\Xi)$ and a measurable map $T : \Xi \to \Xi'$. The push-forward measure $T_{\#}P \in \mathcal{P}(\Xi')$ of P through T is defined as:*

$$T_{\#}P(A) = P(T^{-1}(A)) = P(\{\xi \in \Xi : T(\xi) \in A\}),$$

for any measurable set $A \in \mathcal{B}(\Xi')$.

The intuition behind the concept of a push-forward measure is that it is the probability measure obtained by transporting the mass of P from Ξ to Ξ' by T.

Now, we finally state the optimal transport problem in Kantorovich' formulation.

Definition 3-1.2. *(**Kantorovich' Formulation of the Optimal Transport Problem** [31]) Denote by $T^i : \Xi \times \Xi \to \Xi$, $i = 1, 2$, the projections $T^1(\xi_1, \xi_2) = \xi_1$ and $T^2(\xi_1, \xi_2) = \xi_2$. Consider also the coupling $\pi \in \mathcal{P}(\Xi \times \Xi)$, which is a joint probability measure on the product space $\Xi \times \Xi$, and the cost $c(\xi, \zeta)$ defined over the same product space. The optimal transport cost between probability measures $P, P' \in \mathcal{P}$ is:*

$$\mathcal{L}_c(P, P') = \min_{\pi \in \mathcal{U}(P,P')} \int_{\Xi \times \Xi} c(\xi, \zeta) d\pi(\xi, \zeta), \tag{3-1}$$

where

$$\mathcal{U}(P, P') = \{\pi \in \mathcal{P}(\Xi \times \Xi) : T^1_{\#}\pi = P,\ T^2_{\#}\pi = P'\}.$$

Taking into account the definition of push-forward measure in Definition 3-1.1, the constraints that define the set $\mathcal{U}(P, P')$ are equivalent to imposing the marginal constraints

$$\begin{aligned} \pi(A \times \Xi) &= P(A) \\ \pi(\Xi \times B) &= P'(B), \end{aligned}$$

for any measurable sets $A, B \subset \Xi$. Problem (3-1) is a linear program over the space of probability measures $\mathcal{P}(\Xi \times \Xi)$. Note that if $|\Xi| < \infty$, the previous linear program is finite-dimensional, whereas it is infinite-dimensional otherwise. Furthermore, for continuous cost c and compact space Ξ, problem (3-1) always has solutions [31].

Once we have formulated a generic optimal transport problem using Kantorovich's formulation, we define the Wasserstein distance between two probability measures as the solution of an optimal transport problem.

Definition 3-1.3. *(p-**Wasserstein Distance**[15]) Let the cost in (3-1) be $c(\xi, \zeta) = d^p(\xi, \zeta)$, where d is a distance in Ξ and $p \in [1, \infty)$. The p-Wasserstein distance (based on d), $\mathcal{W}_p(P, P')$, between probability measures $P, P' \in \mathcal{P}^d_p(\Xi)$ is defined as*

$$\mathcal{W}_p(P, P') = \mathcal{L}_{d^p}(P, P')^{\frac{1}{p}}. \tag{3-2}$$

Notice that in Definition 3-1.3 we omit the dependence of $\mathcal{W}_p$ on d, since this distance is always clear from the context. Intuitively, the Wasserstein distance is defined as the minimum cost of transporting, via a coupling, probability mass from measure P to measure P'. This distance is bigger the more different are both probability measures. The solution of (3-2) is also a distance in the measure space $\mathcal{P}_p$. Furthermore, there exists an alternative way of defining the Wasserstein distance using a dual approach. Consider the space $\mathcal{L}^1(P)$ of P-integrable functions. The dual reformulation of the p-Wasserstein distance is:

$$\mathcal{W}_p^p(P, P') = \sup_{u \in \mathcal{L}^1(P), v \in \mathcal{L}^1(P')} \left\{ \int_\Xi u(\xi) P(d\xi) + \int_\Xi v(\xi) P'(d\xi) : u(\xi) + v(\zeta) \leq d^p(\xi, \zeta), \forall \xi, \zeta \in \Xi \right\}.$$

The functions u, v are called dual functions or potentials. Additionally, sets $\mathcal{L}^1(P)$ and $\mathcal{L}^1(P')$ can be replaced with $\mathcal{C}_b(\Xi)$, where $\mathcal{C}_b(\Xi)$ is the set of continuous and bounded real-valued functions in Ξ. Now, denote by $\mathcal{L}$ the space of Lipschitz functions with $|f(\xi) - f(\xi')| \leq ||\xi - \xi'||$ for all ξ, $\xi' \in \Xi$. When $d(\xi, \zeta) = ||\xi - \zeta||$ is the 1-norm on Ξ and $p = 1$, the $\mathcal{W}_1(P, P')$ is called the 1-Wasserstein distance, and it has the following dual reformulation, [14]:

$$\mathcal{W}_1(P, P') = \sup_{u \in \mathcal{L}} \left\{ \int_\Xi u(\xi) dP(\xi) - \int_\Xi u(\xi) dP'(\xi) \right\}, \tag{3-3}$$

being u the dual function or potential.

3-2 Ambiguity Sets

In this section we use the theory of optimal transport and specifically the Wasserstein distance to define ambiguity sets. We also compare different ambiguity sets, and we describe the advantages of using Wasserstein distance based ones.

We start by motivating the use of ambiguity sets. In many problems involving random quantities, their exact probability distributions is not known exactly. Instead, we might have an approximate idea of this distribution. However, we want to account for this uncertainty in our problems, for example, to make decisions that are robust respect to the unknown probability. In this scenario, ambiguity sets are useful. An ambiguity set is a set of probability distributions, which we expect to contain the true, unknown probability of the random variable. Since we have an approximate idea of the unknown distribution, we define an ambiguity set as the one that contains all probability distributions that are "close" to the approximate one, according to some criterion. We call the approximate probability distribution that was our starting point by "nominal" probability distribution. The size of the ambiguity set is crucial: it should contain the true, unknown distribution, since we want to account for it, but it should not be excessively big. The last part is because, the bigger the ambiguity set, the more conservative we are. The optimal size of the ambiguity set is the smallest that contains the unknown probability distribution.

Ambiguity sets are widely used in DRO problems, which are optimization problems in which random parameters of unknown distribution are present. In DRO problems, ambiguity sets are used to capture said ambiguity, and a decision that is robust with respect to the ambiguity is seek. Consider our UAV example in which random gusts of wind disturb the dynamics of the system. If we are unsure about the probability distribution of these disturbances, we can opt to consider an ambiguity set of distributions, in which we expect the unknown distribution to lie. Then we can seek a strategy that enforces some specification while being robust to all probabilities in the ambiguity set. In this way, we ensure that the UAV will satisfy the specification even under the distributional ambiguity of the wind.

We have made clear how ambiguity sets allow us to capture the stochastic behaviour of random, ambiguous parameters. Furthermore, we have highlighted that the choice of the ambiguity set is key to make a good decision. Now let us discuss different classes of ambiguity sets that we can define. When it comes to defining an ambiguity set, several criteria are possible like those related to the moments of the probability distributions in the set or to the distance between such distributions. There is a class of ambiguity sets denoted as moment-ambiguity sets, which contain all probability distributions that satisfy certain constraints on their moments. Such constraints are usually related to a nominal distribution, like a most likely one, in the following way: some moments of all distributions in the ambiguity set should either be the same as those of the nominal one, or at least close to those according to some tolerance. However, moment ambiguity sets, despite leading to tractable convex reformulations in many DRO problems, do not make use of all the available information about the unknown probability distribution. This is because either no data or just sampled data is available in practice, and the conditions on some moments do not describe everything that is known about the random variable [15]. Furthermore consider a DRO problem in the data-driven setting, this is, when the nominal distribution is an empirical one, built from N samples of the uncertain variable. In this setting and, when the number of samples approaches infinity [33], the moment ambiguity set does not shrink to the unknown probability. This class of ambiguity sets have been used for example in [34] for policy synthesis in a distributionally robust safety problem for continuous state systems.

On the other hand, there is a class of ambiguity sets based on statistical metrics. These sets are defined as the sets of all distributions such that the metric between any distribution in the set and the nominal one is less than some tolerance. Examples of such metrics are the ϕ-divergences or the Wasserstein distance. The KL-divergence, a class of ϕ-divergence, has become popular lately. However, it has been shown that it is unable to give raise to ambiguity sets constructed from sampled data that contain the true probability distribution if that distribution is continuous [14]. Furthermore, ϕ-divergences in general have the disadvantage of not taking into account how far away the mass is being transported from one probability distribution to another [15], which is not intuitive. Conversely, the Wasserstein distance is able to give raise to ambiguity sets that contain both continuous and discrete probability distributions, while including a distance-based transport cost. Furthermore, consider a Wasserstein distance based ambiguity set centered on an empirical distribution built from data. In this scenario, modern results involving measure concentration show that the true probability distribution is guaranteed to belong, with some confidence, to the ambiguity set. Finally, tractable reformulations of DRO problems with Wasserstein ambiguity sets have been developed lately [14], [15], in the data-driven setting. Because of their desirable features, in this thesis we make use of Wasserstein distance-based ambiguity sets. Wasserstein ambiguity sets have found applications in distributionally robust formulations of different problems like motion planning for robotics [35], [36] and distributionally robust MDPs with parameter ambiguity [29], for example. In this thesis we use Wasserstein distance-based ambiguity sets of the form $\mathcal{P}_{\text{amb}} = \mathbb{B}_{\varepsilon}(\widehat{P})$. The previous one is a ball in the space of probability distributions with finite p-th moment $\mathcal{P}_p(\Xi)$, of radius $\varepsilon > 0$ and centered on the nominal distribution $\widehat{P}$. We illustrate this set in Figure 3-1.

Consider a Wasserstein ambiguity set in the data-driven setting, and a fixed confidence level. In [14] it is proven that the ambiguity set shrinks to the true probability distribution as the number of samples, N, approaches infinity if ε is defined as an appropriate function inversely proportional to N. Additionally, there exist tractable reformulations of data-driven DRO problems for some classes of uncertainty sets and cost functions [14], [15]. We explain such properties of Wasserstein distance-based data-driven DRO problems in the following Sections.

Measure Concentration Results and Wasserstein Ball

In this section we state the properties of data-driven, Wasserstein distance-based ambiguity sets. To do so, we start by describing modern measure concentration results found in [13]. Consider the random variable $\xi \in \Xi \subset \mathbb{R}^n$ with unknown distribution P^{true}. A key assumption is that P^{true} is a light-tailed distribution [14], i.e., there exists an exponent $a > 1$ such that:

$$a' = \int_{\Xi} \exp \|\xi\|^a P^{\text{true}}(d\xi) < \infty. \tag{3-4}$$

Consider the empirical distribution built from N samples $\{\widehat{\xi^i}\}_{i=1}^N$ of ξ:

$$\widehat{P}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\widehat{\xi^i}},$$

where $\delta_{\widehat{\xi^i}}$ is the Dirac delta that represents a unit of mass concentrated in $\widehat{\xi^i}$. Theorem 2 in [13] states that, under assumption (3-4), the following inequality holds in the case of the 1-Wasserstein distance [14]:

$$Pr(\mathcal{W}_1(P^{\text{true}}, \widehat{P}^N) \geq \varepsilon) \leq \begin{cases} c_1 \exp(-c_2 N \varepsilon^{\max\{n,2\}}) & \text{if } \varepsilon \leq 1 \\ c_1 \exp(-c_2 N \varepsilon^a) & \text{if } \varepsilon > 1 \end{cases}, \tag{3-5}$$

for all $N \geq 1$, $n \neq 2$ and $\varepsilon > 0$, where c_1, c_2 are positive constants that only depend on a, a' and n [14]. Furthermore, for $n = 2$ similar results hold, but, since the expressions are more complicated, we do not state them here. The previous results mean that the distance (in the sense of the 1-Wasserstein metric) between the true distribution P^{true} and the empirical one $\widehat{P}^N$ being bigger than a certain ε has a probability that is upper bounded. Additionally, this upper bound decreases exponentially with the number of samples N. Now, consider the following Wasserstein distance-based ambiguity set:

$$\mathbb{B}_\varepsilon(\widehat{P}^N) = \{P \in \mathcal{P}_p(\Xi) : \mathcal{W}_1(P, \widehat{P}^N) \geq \varepsilon\}.$$

This ambiguity set is a ball of radius ε centered on the empirical distribution $\widehat{P}^N$, built from i.i.d. samples $\{\widehat{\xi^i}\}_{i=1}^N$ of ξ. From the structure of this Wasserstein ball, it follows that the measure concentration results (3-5) provide a lower bound on the probability of the true distribution P^{true} being inside of $\mathbb{B}_\varepsilon(\widehat{P}^N)$. This means that we can build data-driven ambiguity sets in a meaningful way, that is, in such a way that they are guaranteed to contain the true probability distribution with high probability. We illustrate this fact in Figure 3-1. We can use (3-5) to select the smallest radius of the ambiguity set that is guaranteed to contain the true, unknown distribution with some confidence. Consider the data-driven

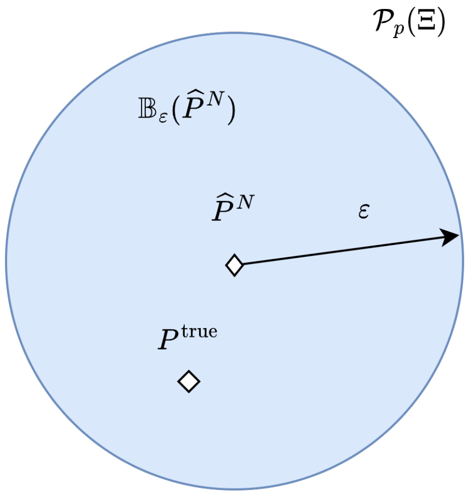


Figure 3-1: Wasserstein ambiguity set $\mathbb{B}_\varepsilon(\widehat{P}^N)$. The ambiguity set is a ball of radius ε centered on $\widehat{P}^N$. Note that in this illustration, an empirical distribution has been taken as the nominal one. A good choice of the radius ε allows the ball to include the true distribution P^{true}, without being unnecessarily large.

ambiguity set $\mathcal{P}_{\text{amb}} = \mathbb{B}_{\varepsilon_N}(\widehat{P}^N)$ and a fixed confidence level $1-\beta$, with $\beta \in (0,1)$. Then, this radius is

$$\varepsilon_N = \begin{cases} (\frac{\log(c_1\beta^{-1})}{c_2 N})^{\frac{1}{\max\{m,2\}}} & \text{if } N \geq \frac{\log(c_1\beta^{-1})}{c_2} \\ (\frac{\log(c_1\beta^{-1})}{c_2 N})^{\frac{1}{a}} & \text{if } N < \frac{\log(c_1\beta^{-1})}{c_2} \end{cases}. \tag{3-6}$$

Notice that the radius of the ambiguity ball ε_N is a decreasing function of N. This means that the higher the number of samples, the smaller the ambiguity set needs to be to guarantee that it contains the true distribution with confidence $1-\beta$. Furthermore, in [14] it was proved that if β_N is defined as an adequately decreasing function of the number of samples, then DRO problems using such set are asymptotically consistent: the solution of the DRO problem using such set converges to the solution of the ambiguity-free problem as N approaches infinity.

However, despite the theoretical results here described, in practice the procedure of selecting a confidence level $1-\beta$ and then obtaining ε such that said confidence level is guaranteed is not followed. This is due to the following practical reasons [14]:

- The constants c_1, c_2 in (3-5) are very difficult to obtain.
- Even if we computed said constants, they would lead to an excessively big ambiguity set, which means high conservatism.

Due to the previous reasons, in this thesis we always consider, even in the data-driven setting, ambiguity sets of fixed radius ε, which we choose arbitrarily. Some practical approaches to determine this parameter are cross-validation tests, bootstrapping or goodness-of-fit tests [14]. Furthermore, in the data-driven setting, we leave the problem of obtaining a value of ε that leads to formal guarantees, such as those in (3-5), to future researches.

Uncertainty Quantification

In this section we describe an algorithm to solve uncertainty quantification problems for variables with ambiguous distributions that belong to data-driven ambiguity sets. These kind of problems arise, for example, when we want to perform an abstraction of a stochastic system whose dynamics are ambiguous. Uncertainty quantification is concerned about the problem of evaluating how much probability mass from a probability measure is inside or outside of a particular set. This problem may arise for example when we want to quantify what are the odds that the value obtained by throwing a dice are between 3 and 5. Another example is the problem of determining the odds that the position of a robot, whose dynamics are modelled by some stochastic process, ends up in a particular region in space. If we have no knowledge about the stochastic behaviour of the dice and robot in the previous examples but, instead, we want to infer it from data, we can use a data-driven DR formulation. This means using an ambiguity set of probability distributions, and computing the best or worst-case probabilities of the random variable belonging to the set. In this section we formally state the DR-version of the uncertainty quantification problem, and describe how to solve it. We must highlight that, when using Wasserstein ambiguity sets, tractable solutions exist when the center of the ambiguity set is an empirical distribution built from data [14], [15].

Consider the arbitrary set $S \subset \mathbb{R}^n$ and the random variable $\xi \in \mathbb{R}^n$. Assume that we want to compute the probability of that random variable belonging to said set. If the probability distribution of ξ is unknown but, instead, we have access to N samples of ξ, $\{\widehat{\xi}_i\}_{i=1}^N$, we can make use a data-driven DR-formulation. Two key approaches to solve this kind of problems are given in [14] and [15], respectively. First, in [14], an exact reformulation of DR-uncertainty quantification problems was proposed in the form of a convex programs, which becomes linear if the 1-Wasserstein distance is used. Secondly, in [15] it was found that, for data-driven DR-uncertainty quantification problems, a finitely-supported worst-case distribution always exists. Additionally, such structure makes it possible to solve the DR-uncertainty quantification problem in a very simple, and intuitive way: it is only required to find distances between the samples $\{\widehat{\xi}_i\}_{i=1}^N$ and the boundary of set S. The solution is then obtained by adding up these distances until a threshold is exceeded. In this thesis we will only make use of this second approach, since it is more advantageous when used to obtain our desired data-driven DR-IMDP abstractions. We will state these advantages in Section 4-2, after we describe the problems we need to solve to obtain said abstractions. In the following, we will limit ourselves to describe this approach.

In [15], an approach is presented to to solve problems of the form

$$\inf_{P \in \mathbb{B}_\varepsilon(\widehat{P}^N)} P[\xi \in S]. \tag{3-7}$$

Consider the case of the nominal distribution being empirical and constructed from the sample set $\{\widehat{\xi}_i\}_{i=1}^N \subset \mathbb{R}^n$, this is, $\widehat{P} = \widehat{P}^N = \sum_{i=1}^N \delta_{\widehat{\xi}^i}$. Without loss of generality, assume that the samples are ordered according to their distance $d(\widehat{\xi}^k, \mathbb{R}^n \setminus S)$ to $\mathbb{R}^n \setminus S$, namely:

$$d(\widehat{\xi}^i, \mathbb{R}^n \setminus S) \leq d(\widehat{\xi}^j, \mathbb{R}^n \setminus S) \text{ for all } 1 \leq i \leq j \leq N. \tag{3-8}$$

Consider the set $[I]$ with $I \leq N$ as the set of indices of the samples that lay outside set S. Furthermore, we say $I = 0$ if and only if $d(\widehat{\xi}^i, \mathbb{R}^n \setminus S) > 0$ for all $i \leq N$. Additionally, for all i

such that $I < i \leq N$, this is, for samples inside S, we define ξ_*^i as the point in the boundary of set S that is the closest to $\widehat{\xi}^i$:

$$\xi_*^i \in \arg\min_{\xi \in \partial S} d(\xi, \widehat{\xi}^i). \tag{3-9}$$

Theorem 3-2.1. ***(Geometric Solution of DR-Uncertainty Quantification Problems** [15]) Let $j^* = max\{j \in [N] \cup \{0\} : \sum_{i=1}^{j} d(\widehat{\xi}^i, \mathbb{R}^n \setminus S) \leq N\varepsilon\}$. Then problem (3-7) has the following solution:*

- *If j^*=N, then $\inf_{P \in \mathbb{B}_\varepsilon(\widehat{P}^N)} P[\xi \in S] = P^*[\xi \in S] = 0$, which is attained by a worst-case distribution*

$$P^* = \frac{1}{N}\sum_{i=1}^{I} \delta_{\widehat{\xi}^i} + \sum_{i=I+1}^{N} \delta_{\xi_*^i}.$$

- *If $j^* < N$, then $\inf_{P \in \mathbb{B}_\varepsilon(\widehat{P}^N)} P[\xi \in S] = P^*[\xi \in S] = 1 - \frac{j^*+p^*}{N}$, which is attained by a worst-case distribution*

$$P^* = \frac{1}{N}\sum_{i=1}^{I} \delta_{\widehat{\xi}^i} + \sum_{i=I+1}^{j^*} \delta_{\xi_*^i} + \frac{p^*}{N}\delta_{\xi_*^{j^*+1}} \frac{1-p^*}{N}\delta_{\widehat{\xi}^{j^*+1}} + \frac{1}{N}\sum_{i=j^*+2}^{N} \delta_{\widehat{\xi}^i},$$

where $p^* = (N\varepsilon - \sum_{i=I+1}^{j^*} d(\widehat{\xi}^i, \mathbb{R}^n \setminus S))/d(\widehat{\xi}^{j^*+1}, \mathbb{R}^n \setminus S)$.

The previous worst-case distribution P^* corresponds to transporting probability mass from the samples to the boundary of the set S in a greedy fashion: all the mass from the first sample $\widehat{\xi}^1$ is transported first since its distance to the boundary is the shortest. This procedure is repeated for the next samples in the ordering until the mass of sample $j^* + 1$ cannot be fully transported since the transport budget ε is exceeded. This constraint violation means that that distribution would not belong to the Wasserstein set $\mathbb{B}_\varepsilon(\widehat{P}^N)$. Therefore, only a fraction p^* of the mass of this last sample is transported, to obtain a distribution that remains in the ambiguity set. The rest of the samples are not transported and, therefore, are left untouched. In Figure 3-2 we depict this geometric approach.

Notice that despite the fact that the results stated in Theorem 3-2.1 are given for DR-uncertainty quantification problems of the form (3-7), Theorem 3-2.1 also allows us to solve problems of the form

$$\sup_{P \in \mathbb{B}_\varepsilon(\widehat{P}^N)} P(\xi \in q'), \tag{3-10}$$

for an arbitrary set $q' \subset \mathbb{R}^n$. We are also interested in solving such problems because similar ones will arise in Chapter 4, when we describe how to obtain data-driven abstractions. Expression (3-11) shows how we can obtain a solution to problem (3-10) by using the results of Theorem 3-2.1:

$$\begin{aligned} &\sup_{P \in \mathbb{B}_\varepsilon(\widehat{P}^N)} P(\xi \in q') \\ =&\sup_{P \in \mathbb{B}_\varepsilon(\widehat{P}^N)} 1 - P(\xi \in \mathbb{R}^n \setminus q') \\ =&1 - \inf_{P \in \mathbb{B}_\varepsilon(\widehat{P}^N)} P(\xi \in S), \end{aligned} \tag{3-11}$$

by saying $S = \mathbb{R}^n \setminus q'$.

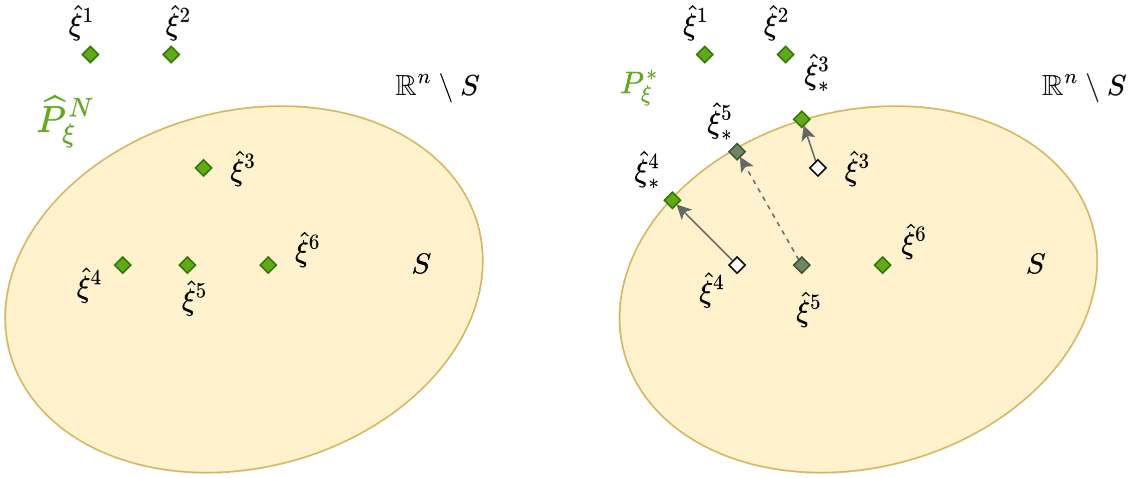


Figure 3-2: Example of the geometric approach. On the left, the empirical distribution $\widehat{P}^N$ consisting of 6 samples: two samples lay outside set S while the remaining 4 are inside set S. On the right, the greedy, worst-case distribution P^*. This worst-case distribution is obtained by transporting all the mass from samples $\widehat{\xi}^3$ and $\widehat{\xi}^4$ to the closest points in the boundary, $\widehat{\xi}^3_*$ and $\widehat{\xi}^4_*$. Furthermore, a fraction of mass $\frac{p^*}{N}$ of $\widehat{\xi}^5$ is transported to $\widehat{\xi}^5_*$, while the remaining $\frac{1-p^*}{N}$ fraction is left in $\widehat{\xi}^5$.

Chapter 4

Strategy Synthesis via Abstractions to DR-IMDPs

In this chapter we introduce our first approach of synthesizing a strategy for an uncertain system. We follow an abstraction-based approach: we abstract the system into an IMDP in a way that the abstraction accounts for the uncertain behaviour of the system. After that, we synthesize a strategy for the IMDP abstraction using the tools for IMDPs already described in Section 2-4-1. The setting of this chapter is a data-driven one: we have access to data from the uncertain system, which we leverage to build the abstraction. We call the process of obtaining a robust IMDP abstraction of such systems by using samples of their disturbances “data-driven DR-abstractions to IMDPs". Furthermore, we assume that the true distribution of said disturbance lies at a distance (in the Wasserstein sense) of at most ε from the distribution constructed from the samples: the empirical distribution.

This chapter is organized as follows: first, in Section 4-1 we formally define the class of systems that we abstract to IMDPs, and we formally state our problem. Next, in Section 4-2 we describe in detail our proposed approach. After that, we describe the algorithms that allow us to synthesize a strategy for the DR-IMDP abstraction in Section 4-3. Furthermore, in Section 4-4 we correctly translate the strategy and satisfaction guarantees obtained for the abstraction to the original system.

4-1 Problem Statement

Let us begin by describing the class of systems that we consider. Consider the following discrete-time dynamical system:

$$x_{t+1} = f(x_t, u_t) + \xi_t, \tag{4-1}$$

where $x_t \in \mathbb{R}^n$ is the state, $u_t \in U_{\mathcal{C}}$ is the control input and $\xi_t \in \mathbb{R}^n$ is a random disturbance. This disturbance process is i.i.d. and has probability distribution $P_\xi^{\mathtt{true}}$, which is unknown to us. Furthermore, notice that, for simplicity and without loss of generality, we consider that the dimension of the disturbance is the same as that of the state space, $n \in \mathbb{N}$. Additionally, we associate to system (4-1) a set of observations $O_{\mathcal{C}}$, and an observation function $L_{\mathcal{C}}$ that assigns an observation $o \in O_{\mathcal{C}}$ to every state $x \in \mathbb{R}^n$ of the system. We now define the transition kernel $\mathcal{T}_{\mathcal{C}}$, parametric in the probability $P_\xi \in \mathcal{P}_p(\mathbb{R}^n)$, which for any $x \in \mathbb{R}^n$, $u \in U_{\mathcal{C}}$ assigns a probability measure on the Borel space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. This is:

$$\mathcal{T}_{\mathcal{C}}(D|x, u; P_\xi) := \int_D (f(x,u) + \xi) dP_\xi, \tag{4-2}$$

for any measurable set D. We need such a parametric kernel because probability distribution $P_\xi^{\mathtt{true}}$ of ξ is unknown. We also define the notion of paths and traces of system (4-1). A path of system (4-1) is a sequence of states [4] $w_{\mathbb{R}^n} = x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} x_2 \xrightarrow{u_2} \dots$ that, for fixed P_ξ satisfy the dynamics (4-1) and where $u_t \in U_{\mathcal{C}}$ for all $t \in \mathbb{N}_{\geq 0} \cup \{\infty\}$. We denote the prefix of length $k+1$, for $k \in \mathbb{N}_{\geq 0}$, of a path $w_{\mathbb{R}^n}$ by $w_{\mathbb{R}^n}^k$, and the sets of all finite and infinite paths are called $\mathrm{Paths}_{\mathbb{R}^n}^{\mathtt{fin}}$ and $\mathrm{Paths}_{\mathbb{R}^n}$, respectively. We also denote the last state of a finite path $w_{\mathbb{R}^n}^{\mathtt{fin}}$ by $last(w_{\mathbb{R}^n}^{\mathtt{fin}})$ and the $t+1$-th state of a path $w_{\mathbb{R}^n}$ by $w_{\mathbb{R}^n}(t)$. A path $w_{\mathbb{R}^n} = x_0 x_1 x_2 ...$ produces a trace $w_{o_{\mathcal{C}}} = o_0 o_1 o_2 ... \in O^\omega$ such that $o_t = L_{\mathcal{C}}(x_t)$ for all $t \in \mathbb{N}_{\geq 0} \cup \{\infty\}$.

Now let us introduce the concept of a strategy of system (4-1).

Definition 4-1.1. ***(Strategy of a Continuous-State Dynamical System)*** *A strategy $\sigma_{\mathcal{C}} : Paths_{\mathbb{R}^n}^{\mathtt{fin}} \to U_{\mathcal{C}}$ of system* (4-1) *is a function that maps a finite path $w_{\mathbb{R}^n}^k$ into a control input $u \in U_{\mathcal{C}}$. The set of all strategies $\sigma_{\mathcal{C}}$ is denoted by $\Sigma_{\mathcal{C}}$.*

Then, given $P_\xi \in \mathcal{P}_p(\mathbb{R}^n)$, a time horizon $[0, k]$ and a strategy $\sigma_{\mathcal{C}}$, system (4-1) defines a stochastic process in the space $\Omega = (\mathbb{R}^n)^{k+1}$, with Borel sigma algebra $\mathcal{B}(\Omega)$ [11]. Furthermore, for initial state $x_0 \in \mathbb{R}^n$, a probability P is induced over the paths of said process, which is uniquely defined by $\mathcal{T}_{\mathcal{C}}$: for any $t \in \{0, 1, \dots, k-1\}$

$$P(w_{\mathbb{R}^n}^k(0) \in D) = \mathbf{1}_D(x_0)$$
$$P(w_{\mathbb{R}^n}^k(t+1) \in D | x_t = x, u_t = \sigma_{\mathcal{C}}(w_{\mathbb{R}^n}^t), P_\xi) = \mathcal{T}_{\mathcal{C}}(D | x_t, \sigma_{\mathcal{C}}(w_{\mathbb{R}^n}^t); P_\xi)$$

for any measurable set $D \in \mathbb{R}^n$. Furthermore, for $k = \infty$, P is also uniquely defined by $\mathcal{T}_{\mathcal{C}}$ by the *Ionescu-Tulcea* extension theorem [37].

In the setting of this chapter we assume that we have access to a finite number of samples $\{\widehat{\xi}^i\}_{i=1}^N$ from disturbance ξ, that we might have obtained in two different ways: either from direct measurements of the disturbance or, if we have recorded the state and control input of a trajectory of the system, by subtracting $f(x_t, u_t)$ from the state x_{t+1} for all t available. Once we have defined the class of systems we consider, let us state our problem. Denote by $P(w_{\mathbb{R}^n}^k \models \phi | w_{\mathbb{R}^n}^k(0) = x, X, \sigma_{\mathcal{C}}, P_\xi)$ the probability of the paths of system (4-1) for fixed P_ξ satisfying an scLTL formula ϕ within $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ steps while staying inside $X \subset \mathbb{R}^n$ by starting from state $x \in X$ and following strategy $\sigma_{\mathcal{C}}$.

Problem 4-1.1. ***(Data-driven Distributionally Robust Strategy Synthesis)*** *Consider the dynamical system* (4-1)*. Assume that a finite set of samples $\{\widehat{\xi}^i\}_{i=1}^N$, $N \geq 1$, from its disturbance are available, and denote the corresponding empirical distribution $\widehat{P}_\xi^N$. Moreover, consider the p-Wasserstein distance-based ambiguity set $\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$, of radius $\varepsilon > 0$ and centered* *on $\widehat{P}_\xi^N$. Consider also a compact set $X \subset \mathbb{R}^n$ and an scLTL formula ϕ defined over the regions of interest of X. Then, find a near-optimal strategy $\sigma_\mathcal{C}^*$ that allows to determine if for given initial state $x \in X$, probability threshold $p_{\mathtt{th}}$ and horizon $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$*

$$P(w_{\mathbb{R}^n}^k \models \phi | w_{\mathbb{R}^n}^k(0) = x, X, \sigma_\mathcal{C}^*, P_\xi) \geq p_{\mathtt{th}} \tag{4-3}$$

holds for all probabilities $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$ of the disturbance ξ. Then, if ε has been chosen in such a way that $P_\xi^{\mathtt{true}} \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$ holds with high confidence [1], we can easily check if (4-3) holds for the case of system (4-1) under $P_\xi^{\mathtt{true}}$, since

$$\begin{aligned} &P(w_{\mathbb{R}^n}^k \models \phi | w_{\mathbb{R}^n}^k(0) = x, X, \sigma_\mathcal{C}^*, P_\xi^{\mathtt{true}}) \\ \geq &\min_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P(w_{\mathbb{R}^n}^k \models \phi | w_{\mathbb{R}^n}^k(0) = x, X, \sigma_\mathcal{C}^*, P_\xi) \cdot (1-\beta), \end{aligned}$$

where $1-\beta$ is the confidence level. However, notice that we leave this last part outside the statement of problem 4-1.1 because in this thesis we assume given ambiguity sets. The problem of finding the size of the ambiguity set that contains $P_\xi^{\mathtt{true}}$ is out of the scope of this thesis.

We seek a near-optimal strategy because there is no way of finding the exact strategy that maximizes the probability in (4-3) [11]. Instead, as we already pointed out in the introduction of this chapter, we follow an abstraction-based approach to solve Problem 4-1.1. First, we abstract system (4-1) to an finite-state IMDP, which accounts for both the discretization of the state-space and the distributional uncertainty regarding the disturbance. Then we synthesize the strategy that maximizes the worst-case probability of the paths of said abstraction satisfying the specification. We do that using interval value iteration algorithm described in Section 2-4-1. Finally we translate said strategy into a strategy that system (4-1) is able to use, while preserving the guarantees obtained for the abstraction.

4-2 Obtaining DR-IMDP Abstractions

As we already pointed out, to solve Problem 4-1.1 we follow an abstraction-based approach. This one relies on the theory of IMDPs and ambiguity sets explained in Sections 2-4 and 3-2, respectively. In this section we describe how we obtain these DR-IMDP abstractions.

First, we abstract system (4-1) to the IMDP $\mathcal{I} = (Q, A, \underline{P}, \overline{P}, O, L)$. Let us begin with the state space Q. We focus on a compact set $X \subseteq \mathbb{R}^n$ of the state space of system (4-1). We discretize this one into a finite number of non-overlapping regions $\widetilde{Q} := \{q_1, q_2, \ldots, q_{|\widetilde{Q}|}\}$ such that:

$$\cup_{q \in \widetilde{Q}} q = X, \quad \text{and} \quad q \cap q' = \emptyset \quad \forall q, q' \in \widetilde{Q} \,\text{and}\, q \neq q'.$$

Moreover, in this chapter we restrict the set the set X to be a hyper-rectangle and we only make use of uniform partitions. Using this kind of partitions is key in our approach to obtain the transition probability bounds of the abstraction, as we explain later in Section 4-2-1.

[1] see Section 3-2.

Now, we assign each region q of the discretization to a different state in the IMDP $\mathcal{I}$. With an abuse of notation we refer by q to both, a state $q \in \widetilde{Q}$ of the IMDP, and to a region $q \subset X$. The correct interpretation should be clear from the context. Furthermore, for the sake of simplicity, we restrict to the case that the discretization respects the regions of interest: for every region $q \in \widetilde{Q}$, all $x \in q$ must share the same observations. IMDP abstractions in which the regions of interest are not respected are also possible [4]. Additionally, we denote by q_u the set $\mathbb{R}^n \setminus X$, which corresponds to the rest of the state space of system (4-1). Taking this extra state into account, we define the state space of the IMDP $\mathcal{I}$ as $Q := \widetilde{Q} \cup \{q_u\}$. In this way we have discretized the entirety of the continuous state space $\mathbb{R}^n$. Next, when it comes to the observation set O of the $\mathcal{I}$, we let this set be the same as in the original system: $O := O_{\mathcal{C}}$.

Regarding the set of actions A of $\mathcal{I}$, we define it as a finite subset of $U_{\mathcal{C}}$. Furthermore we make all actions available at each state q, i.e., $A(q) = A$ for all $q \in Q$. This is not necessarily the only way to define the set of actions of the IMDP, and state-dependent action sets can exist. However, to simplify the problem, in this document we only consider this case. Next, we define the observation map of the IMDP $\mathcal{I}$ as $L(q) := L_{\mathcal{C}}(x)$ for any $x \in q$ and for all $q \in Q$. This means that the observation map of the IMDP assigns to each cell the same observation that was assigned to that region in the continuous system model. In this way, we leverage the fact that the discretization respects the regions of interest.

Finally, we state the expressions that the transition probability bounds $\underline{P}, \overline{P}$ of $\mathcal{I}$ must satisfy. Such quantities must bound the worst and best-case probabilities of transitioning between states taking into account the ambiguity introduced by the state discretization and the distributional uncertainty. Therefore, taking into account the dynamics (4-1) and the ambiguity set $\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$, said bounds must fulfill

$$\begin{aligned}
\underline{P}(q,a,q') &\leq \min_{x_t \in q} \inf_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x_{t+1} \in q' | x_t, a_t = a) \\
&= \min_{x_t \in q} \inf_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} \mathcal{T}_{\mathcal{C}}(q' | x_t, u_t = a; P_\xi) \\
\overline{P}(q,a,q') &\geq \max_{x_t \in q} \sup_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x_{t+1} \in q' | x_t, u_t = a) \\
&= \max_{x_t \in q} \sup_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} \mathcal{T}_{\mathcal{C}}(q' | x_t, a; P_\xi),
\end{aligned} \tag{4-4}$$

for all $q, q' \in Q \setminus \{q_u\}$, $a \in A$, and

$$\begin{aligned}
\underline{P}(q,a,q_u) &\leq 1 - \max_{x_t \in X} \sup_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x_{t+1} \in X | x_t, a_t = a) \\
&= 1 - \max_{x_t \in q} \sup_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} \mathcal{T}_{\mathcal{C}}(X | x_t, u_t = a; P_\xi) \\
\overline{P}(q,a,q_u) &\geq 1 - \min_{x_t \in q} \inf_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x_{t+1} \in X | x_t, u_t = a) \\
&= 1 - \min_{x_t \in q} \inf_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} \mathcal{T}_{\mathcal{C}}(X | x_t, a; P_\xi),
\end{aligned} \tag{4-5}$$

for all $q \in Q \setminus \{q_u\}$, $a \in A$. Furthermore, according to Problem 4-1.1, we are only interested on the paths that never exit X. To account for that extra specification, we make the state q_u

absorbing by defining $\overline{P}(q_u, a, q_u) = \underline{P}(q_u, a, q_u) = 1$ and $\overline{P}(q_u, a, q') = \underline{P}(q_u, a, q') = 0$ for all $q' \in Q \setminus \{q_u\}$, $a \in A$.

Obtaining the bounds in expressions (4-4) and (4-5) is the main problem we need to solve to compute our DR-IMDP abstractions. Now, before describing how we obtain such bounds, let us begin by looking at the inner optimization problems over the ambiguity set in (4-4) and (4-5). These problems are, DR-uncertainty quantification problems similar to those in (3-10) and (3-7) from Section 3-2. Let us remember that in that section we pointed out to two approaches to solve data-driven DR-uncertainty quantification problems: the one introduced in [14] based on convex (or linear) programming, and the one presented in [15] that has a geometrical interpretation. As we already stated in Section 3-2, in our approach, we only make use of the second method due to the advantages it presents:

- This approach can be used with a generic distance d,
- This approach has a geometric interpretation, and it is easy to find relaxations to the problem, to speed up computations, as we describe in Section 4-2.3,
- Furthermore, this approach allows us to combine the inner and outer optimization problems in and (4-4) and (4-5) into a single one when we use said relaxations. This allows us to obtain bounds analytically, as we discuss in Section 4-2-2,
- Computing distances from sample to set can be done in a very efficient way if the set is a hyperrectangle: it does not require to solve a number of optimization problems in which the number of variables scales with the number of samples, N, which can be high. Instead, we can obtain the distances by using a series of `if` statements and computing distances between points and planes in closed form.

Let us remember from Section 3-2 that, DR-uncertainty quantification problems like those in and (4-4) and and (4-5) only have tractable reformulations when the nominal distribution is finitely supported. Fortunately, this is the case in the data-driven setting that we consider in this chapter. Once that we have discussed the tools we have to solve the inner problems in and (4-4) and and (4-5), let us notice that the solution of these DR-uncertainty quantification problems is parametric in the variable x_t. Then, we need to also optimize the result of the inner problems over x_t. Unfortunately, in general this is not a convex problem, since it depends on the distribution of the samples, which is random. This means that to get good bounds for $\underline{P}(q, a, q')$ and $\overline{P}(q, a, q')$, we need to find tractable ways to deal with this non-convexity.

Taking into account the previous considerations, now we briefly describe how we obtain the bounds in expressions (4-4) and (4-5). To obtain those in (4-4) we rely on precomputing a lookup table. Consider a generic state $q^* \in Q \setminus \{q_u\}$ of the partition. The lookup table contains the solution of DR-uncertainty quantification problems for the case that the samples are centered at every point in a grid around this region q^*. After that, we obtain the bounds by performing a grid search in the lookup table, making use of the reachable set of current state $q \in Q \setminus \{q_u\}$ by action $a \in A$. Since the disturbance is additive and the regions $q \subset \mathbb{R}^n$ are of the same size (except q_u), we prove that we only need one lookup table to determine the bounds between any states $q, q' \in Q \setminus \{q_u\}$ and action $a \in A$. However, notice that an additional table over X is needed to obtain the bounds of transitioning to state q_u in expressions (4-5). This is the first way in which we obtain the bounds in expressions (4-4) and (4-5), which we describe in detail in Section 4-2-2.

On the other hand, the approach in [15] to solve the DR-uncertainty quantification problems allows us to formulate efficient relaxations of problems (4-4) and (4-5). Such relaxations lead to analytical bounds in the transition probabilities of $\mathcal{I}$ that are tight when the distance between the samples of x_{t+1} and the set q' is large. Therefore, we make use of these relaxations when the previous distance goes over a threshold D_{th}. This is the second way to obtain the transition probability bounds, which we describe in detail in Section 4-2-2.

4-2-1 Lookup Table

In this section we describe the main way in which we obtain the transition probability bounds in (4-4) and (4-5) by searching on a lookup table. Actually, we need one lookup table to compute the bounds in (4-4) and a second one to compute the bounds in (4-5).

Let us first clarify the notation we use in this section. We only make use of distance $d = \|\cdot\|_p$ and Wasserstein distance $\mathcal{W}_p$, with $p \geq 1$. Furthermore, we make use of balls in the metric space $(\mathbb{R}^n, \|\cdot\|_p)$, and balls in the probability space $\mathcal{P}_p(\mathbb{R}^n)$ based on the distance $\mathcal{W}_p$. According to Section 1-5, we use the notation B and $\mathbb{B}$, respectively, to refer to such balls. Furthermore, we omit the spaces $\mathbb{R}^n$, $\mathcal{P}_p(\mathbb{R}^n)$ and the distances $\|\cdot\|_p$, $\mathcal{W}_p$ as arguments of such balls, since these are clear from the context.

Now, let us start by describing how we obtain the lookup table. After that we describe how we obtain the bounds in (4-4) by searching in the table.

Definition 4-2.1. ***(Lookup table)*** *Consider a uniform grid over a region of the state space which contains an arbitrary rectangle $q^* \subseteq X$ from the partition of the state space. Denote the set of centers of the sub-cells of this grid by $\{y_k^{q^*}\}_{k=1}^{n_{\texttt{table}}}$, where $n_{\texttt{table}}$ is the total number of centers. Let $B_{\delta^*}(y_k^{q^*}) \subset \mathbb{R}^n$ be the smallest ball, centered on any $y_k^{q^*}$, that contains the corresponding sub-cell. We define the lookup table for q^* as a table with $n_{\texttt{table}}$ entries, where each entry contains both the position of the respective center $y_k^{q^*}$ and the solution of the following DR-uncertainty quantification problems:*

$$\begin{aligned} \underline{p}(y_k^{q^*}) &:= \inf_{P_\xi \in \mathbb{B}_{\varepsilon+\delta^*}(\widehat{P}_\xi^N)} P_\xi(y_k^{q^*} + \xi \in q^*) \\ \overline{p}(y_k^{q^*}) &:= \sup_{P_\xi \in \mathbb{B}_{\varepsilon+\delta^*}(\widehat{P}_\xi^N)} P_\xi(y_k^{q^*} + \xi \in q^*), \end{aligned} \tag{4-6}$$

for all $k \in [n_{\texttt{table}}]$.

Uncertainty quantification problems (4-6) correspond to the worst and best-case probabilities of ξ belonging to set q^* when the samples of ξ are translated by the position of $y_k^{q^*}$. For clarity, we depict the grid, sub-cells, centers of the sub-cells and ball $B_{\delta^*}(y_k^{q^*})$ in Figure 4-1. According to Definition 4-2.1, each entry of the lookup table contains the following information: the position of the corresponding center $y_k^{q^*}$ of the grid, and the solution to the DR-uncertainty quantification problems (4-6) corresponding to that center. Without loss of generality, we define q^* as the a rectangle centered at the origin. We should also highlight that we have

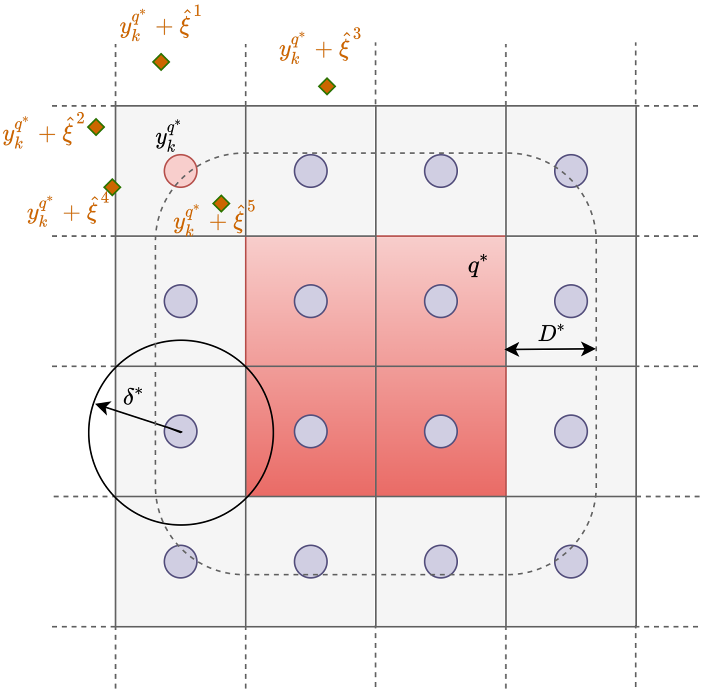


Figure 4-1: Illustration of the grid surrounding q^* and how we obtain the lookup table. The figure also shows the points that lie at a distance D^* from q^* and the disposition of the samples of $y_k^{q^*}+\xi$.

increased the transport budget of the ambiguity balls in (4-6) to $\varepsilon+\delta^*$. This is to account for the case that the samples are centered in any point inside the sub-cells that is not the corresponding center, as we describe in Proposition 4-2.1:

Proposition 4-2.1. *Consider the lookup table as described in Definition 4-2.1. Take the entry of the table with position $y_k^{q^*}$, which is the center of the corresponding sub-cell subcell$(y_k^{q^*})$. Consider also the corresponding quantities $\underline{p}(y_k^{q^*})$ and $\overline{p}(y_k^{q^*})$ in that entry of the table. Then, the latter are, respectively, lower and upper bounds on the solution of the DR-uncertainty quantification problems for set q^* that correspond to the case that the samples are centered at any point $x\in$ subcell$(y_k^{q^*})$:*

$$\begin{aligned}\underline{p}(y_k^{q^*}) &\leq \min_{x\in subcell(y_k^{q^*})}\ \inf_{P_\xi\in\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x+\xi\in q^*)\\ \overline{p}(y_k^{q^*}) &\geq \max_{x\in subcell(y_k^{q^*})}\ \sup_{P_\xi\in\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x+\xi\in q^*)\end{aligned} \tag{4-7}$$

Proof. Consider the sub-cell of the grid centered on $y_k^{q^*}$. Then we get that

$$\min_{x\in\mathrm{subcell}(y_k^{qj})}\ \inf_{P_\xi\in\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x+\xi\in q^*) \geq \min_{x\in B_{\delta*}(y_k^{q^*})}\ \inf_{P_\xi\in\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(x+\xi\in q^*). \tag{4-8}$$

We now define the random variable $z:=x+\xi$ and its empirical distribution $\widehat{P}_{x+\xi}^N$, constructed from samples of $\{x+\widehat{\xi^i}\}_{i=1}^N$. These ones allow us to reformulate the expression on the right hand of (4-8) as

$$\min_{x \in B_{\delta^*}(y_k^{q^*})} \inf_{P_z \in \mathbb{B}_\varepsilon(\widehat{P}^N_{x+\xi})} P_z(z \in q^*) \tag{4-9}$$

Now, consider the set of allowed P_z in problem (4-9):

$$\bigcup_{x \in B_{\delta^*}(y_k^{q^*})} \mathbb{B}_\varepsilon(\widehat{P}^N_{x+\xi}). \tag{4-10}$$

By following the same procedure as in Lemma A.2 of [38], it is possible to obtain an over-approximation of the set (4-10). First, we define another empirical probability distribution, $\widehat{P}^N_{y_k^{q^*}+\xi}$, constructed from samples of $\{y_k^{q^*} + \widehat{\xi}^i\}_{i=1}^N$. Then, we seek to bound in the transport of mass from $\widehat{P}^N_{y_k^{q^*}+\xi}$ to $\widehat{P}^N_{x+\xi}$ for all $x \in B_{\delta_{q^*}}(y_k^{q^*})$ by using the Wasserstein distance and the coupling $\pi^* := \frac{1}{N}\sum_{i=1}^N \delta_{(y_k^{q^*}+\widehat{\xi}^i, x+\widehat{\xi}^i)} \in \mathcal{P}_p(\mathbb{R}^n \times \mathbb{R}^n)$. It is trivial to check that the marginals of this coupling are the measures $\widehat{P}^N_{y_k^{q^*}+\xi}$ and $\widehat{P}^N_{x+\xi}$, so it is a feasible coupling. Therefore, the p-Wasserstein distance between these two measures fulfills:

$$\mathcal{W}_p^p(\widehat{P}^N_{y_k^{q^*}+\xi}, \widehat{P}^N_{x+\xi}) \leq \int_{\mathbb{R}^n \times \mathbb{R}^n} ||\xi - \zeta||_p^p \pi^*(d\xi, d\zeta) = \frac{1}{N}\sum_{i=1}^N \|x - y_k^{q^*}\|_p^p = (\delta^*)^p, \tag{4-11}$$

for all $x \in B_{\delta^*}(y_k^{q^*})$. Now, consider an arbitrary probability $P_z \in \mathbb{B}_\varepsilon(\widehat{P}^N_{x+\xi})$. Using the result in (4-11) and the triangle inequality for $\mathcal{W}_p$, we get that the Wasserstein distance between probabilities P_z and $\widehat{P}^N_{y_k^{q^*}+\xi}$ is

$$\mathcal{W}_p(P_z, \widehat{P}^N_{y_k^{q^*}+\xi}) \leq \underbrace{\mathcal{W}_p(P_z, \widehat{P}^N_{x+\xi})}_{\leq \varepsilon} + \underbrace{\mathcal{W}_p(\widehat{P}^N_{y_k^{q^*}+\xi}, \widehat{P}^N_{x+\xi})}_{\leq \delta^*} \leq \varepsilon + \delta^*.$$

This means that the probability P_z that we considered belongs to a Wasserstein ball of radius $\varepsilon + \delta^*$ and centered on distribution $\widehat{P}^N_{y_k^{q^*}+\xi}$, as we illustrate in Figure 4-2. Using these results, we over-approximate the set (4-10) of problem (4-9) as follows:

$$\{P_z \in \mathbb{B}_\varepsilon(\widehat{P}^N_{x+\xi}), x \in B_{\delta^*}(y_k^{q^*})\} \subseteq B_{\varepsilon+\delta^*}(\widehat{P}^N_{y_k^{q^*}+\xi}).$$

Using this set instead of (4-10) we get a lower bound for equation (4-9), and therefore the first expression in (4-7) follows. We could also prove that the second expression in (4-9) is also true by following a similar procedure. However, since this is trivial we only provide a proof for the first inequality. □

Once we have computed the lookup table, we use it to obtain the transition probability bounds $\underline{P}(q, a, q')$, $\overline{P}(q, a, q')$ for any combination of current state $q \in Q$, action $a \in A$ and successor state $q' \in Q$. For that we make use of the forward reachable set of state q by action a. First, we over approximate such set, and then translate it to be able to perform a search in the sub-cells of the lookup table whose union contains the resulting set. First, let us define the forward reachable set $\mathcal{R}(q, a)$ of f from q by a as the set

$$\mathcal{R}(q, a) := \{f(x, a) \in \mathbb{R}^n : x \in q\},$$

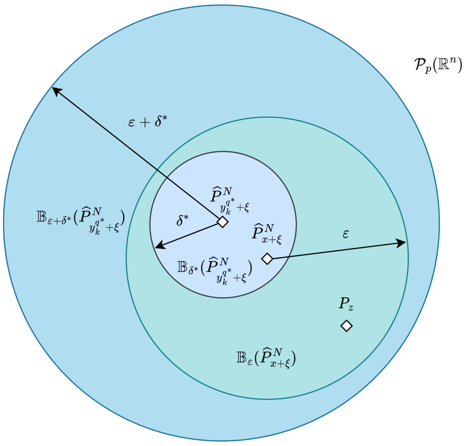

$\mathcal{P}_p(\mathbb{R}^n)$
$\varepsilon + \delta^*$
$\widehat{P}^N_{y_k^{q^*}+\xi}$
$\mathbb{B}_{\varepsilon+\delta^*}(\widehat{P}^N_{y_k^{q^*}+\xi})$
δ^*
$\widehat{P}^N_{x+\xi}$
ε
$\mathbb{B}_{\delta^*}(\widehat{P}^N_{y_k^{q^*}+\xi})$
P_z
$\mathbb{B}_{\varepsilon}(\widehat{P}^N_{x+\xi})$

Figure 4-2: Graphical representation of the p-Wasserstein distances we considered in the proof of Equation (4-7)

for all $q \in Q$, $a \in A$. This is the set of successor states that can be reached (deterministically) starting in $x \in q$ by action $a \in A$. Let us now define affine under and overapproximations $\underline{f}_q(x, a)$ and $\overline{f}_q(x, a)$ of f in q such that for fixed a $\underline{f}_q(x, a) \leq f(x, a) \leq \overline{f}_q(x, a)$ for all $x \in q$, $q \in Q$. Using these functions we obtain the following overapproximation of $\mathcal{R}(q, a)$:

$$\mathcal{R}_{app}(q, a) := \{y \in \mathbb{R}^n : \underline{f}_q(x, a) \leq y \leq \overline{f}_q(x, a), x \in q\} \supseteq \mathcal{R}(q, a). \tag{4-12}$$

Both sets $\mathcal{R}(q, a)$ and $\mathcal{R}_{app}(q, a)$ are illustrated in Figure 4-3. As shown in [5], we easily obtain the set $\mathcal{R}_{app}(q, a)$ as described in Proposition 4-2.2:

Proposition 4-2.2. *Consider the vertices $\{v_1, v_2, \dots, v_{(n^2)}\}$ of the rectangle q of the discretization of the state-space. Then it holds that $\mathcal{R}_{app}(q, a)$ corresponds to the convex hull of the rectangles $\{rect(\underline{f}(v_k, a), \overline{f}(v_k, a))\}_{k=1}^{(2^n)}$:*

$$\mathcal{R}_{app}(q, a) = conv(\{rect(\underline{f}(v_k, a), \overline{f}(v_k, a))\}_{k=1}^{(2^n)}),$$

Set $\mathcal{R}_{app}(q, a)$ is a polytope, and way easier to work with than $\mathcal{R}(q, a)$. Now, consider the rectangle $q' \in Q \setminus \{q_u\}$ and its center $c_{q'}$. Then, we translate the set $\mathcal{R}_{app}(q, a)$ by quantity $c_{q'}$. We denote by $\mathcal{R}_{app}(q, a) - c_{q'}$ the translation of set $\mathcal{R}_{app}(q, a)$ by the quantity $c_{q'}$. This translated set is useful to obtain the bounds of transitioning between states $q, q' \in Q \setminus \{q_u\}$ under action $a \in A$ by searching in the lookup table. Furthermore, denote by $Y_{q^*}(q, a, q') \subseteq \{y_k^{q^*}\}_{k=1}^{n_{\texttt{table}}}$ the subset of all centers $y_k^{q^*}$ of the table such that the intersection between the sub-cells of the grid with centers in $y_k^{q^*}$ and the translated set $\mathcal{R}_{app}(q, a) - c_{q'}$ is nonempty:

$$Y_{q^*}(q, a, q') := \{y_k^{q^*} \in \{y_k^{q^*}\}_{k=1}^{n_{\texttt{table}}} : \text{subcell}(y_k^{q^*}) \cap (\mathcal{R}_{app}(q, a) - c_{q'}) \neq \emptyset\}, \tag{4-13}$$

for all $q, q' \in Q \setminus \{q_u\}$, $a \in A$. Then, the following theorem allows us to find bounds on the probability of transitioning from states $q \in Q$ to $q' \in Q$ under action $a \in A$ by performing a grid search on the lookup table computed for the region q^*:

Theorem 4-2.1. ***(Tight Bounds in the Transition Probabilities by Searching on the Lookup Table).*** *Consider the lookup table computed as described in 4-2.1 and the set*

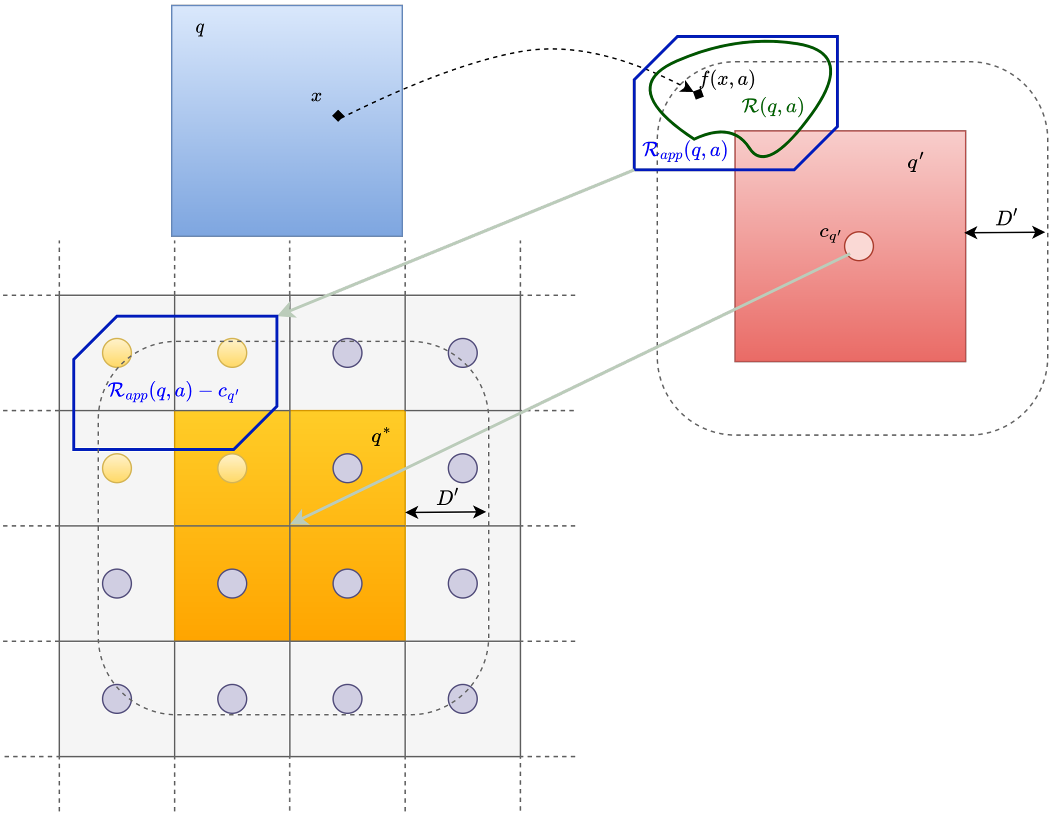


Figure 4-3: Illustration of how we compute the transition probability bounds between q and q' under a by using $\mathcal{R}_{app}(q,a)$ and searching in the lookup table. The points in the lookup table that we should consider, this is, set $Y_{q^*}(q,a,q')$, are represented in light yellow.

$Y_{q^*}(q,a,q')$ as defined in (4-13) for all $q,q' \in Q \setminus \{q_u\}$, $a \in A$. Then the bounds in (4-4) are obtained by performing a simple grid search:

$$\underline{P}(q,a,q') := \min_{y_k^{q^*} \in Y_{q^*}(q,a,q')} \underline{p}(y_k^{q^*})$$

$$\overline{P}(q,a,q') := \max_{y_k^{q^*} \in Y_{q^*}(q,a,q')} \overline{p}(y_k^{q^*}).$$

Proof. For the proof of Theorem 4-2.1, we need to take into account two things: the forward reachable set of q by a, and the additivity of the disturbance ξ. First, let us take into account the reachable set. Consider now the problem of computing the lower bound $\underline{P}(q,a,q')$. The upper bound follows a similar reasoning. Using $\mathcal{R}_{app}(q,a)$ we reformulate the first problem in (4-4) as:

$$\min_{x\in q} \inf_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(f(x,a)+\xi \in q'|x,a) \geq \min_{y\in\mathcal{R}_{app}(q,a)} \inf_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(y+\xi \in q'). \tag{4-14}$$

Furthermore, taking into account that the disturbance is additive, we perform a (negative) translation of both sets $\mathcal{R}_{app}(q,a)$ and set q' by the center of cell q', $c_{q'}$, so that we get $q' - c_{q'} = q^*$. This means that we obtain the solution of the problem in the right hand side of

(4-14) by solving a DR-uncertainty quantification problem in set q^* and where the samples are centered at a point in the translated reachable set $\mathcal{R}_{app}(q,a) - c_{q'}$:

$$\min_{y\in\mathcal{R}_{app}(q,a)} \inf_{P_\xi\in\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(y+\xi\in q') = \min_{y\in\mathcal{R}_{app}(q,a)-c_{q'}} \inf_{P_\xi\in\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(y+\xi\in q^*). \tag{4-15}$$

This equivalence is shown in Figure 4-3. This formulation allows us to make use of the lookup table we presented in Definition 4-2.1, since it contains the solution of DR-uncertainty quantification problems in a grid that covers q^*. Now, we want to formulate the outer minimization in the right hand side of expression (4-15) as a finite search on the table. To do so, we start by further over approximating $\mathcal{R}_{app}(q,a) - c_{q'}$ by the union of all the sub-cells of the grid that have a nonempty intersection with that set:

$$\bigcup_{y_k^{q^*}\in Y_{q^*}(q,a,q')} \text{subcell}(y_k^{q^*}) \supset \mathcal{R}_{app}(q,a) - c_{q'}. \tag{4-16}$$

By using the set in (4-16) in the outer minimization of expression (4-15), we get a further lower bound on its solution. Now, by construction of the table and by taking into account Proposition 4-2.1, instead of searching in region $\mathcal{R}_{app}(q,a) - c_{q'}$, it suffices to search in the finite subset of the centers of the grid $Y_{q^*}(q,a,q')$ defined in (4-13). With this, we have proved Theorem 4-2.1. □

Now, let us remember that we have only described the procedure to obtain the bounds in (4-4). In order to also compute the bounds in (4-5), we follow the same approach, with the only difference that now we make use of a second lookup table. This second table is analogous to the one in Definition 4-2.1:

Definition 4-2.2. ***(Lookup table Over X)*** *Consider a uniform grid over a region of the state space which contains the rectangle X that defines the workspace. Denote the set of centers of the sub-cells of this grid by $\{y_k^X\}_{k=1}^{n_{\texttt{table,X}}}$, where $n_{\texttt{table,X}}$ is the total number of centers. Let $B_{\delta^X}(y_k^X) \subset \mathbb{R}^n$ be the smallest ball centered on any y_k^X that contains the corresponding sub-cell. We denote he radius of this ball by $\delta^X > 0$. We define the lookup table for X as a table with $n_{\texttt{table,X}}$ entries, where each entry contains both the position of the respective center y_k^X and the solution of the following DR-uncertainty quantification problems:*

$$\begin{aligned} \underline{p}(y_k^X) &:= \inf_{P_\xi\in\mathbb{B}_{\varepsilon+\delta^*}(\widehat{P}_\xi^N)} P_\xi(y_k^X+\xi\in X) \\ \overline{p}(y_k^X) &:= \sup_{P_\xi\in\mathbb{B}_{\varepsilon+\delta^*}(\widehat{P}_\xi^N)} P_\xi(y_k^X+\xi\in X), \end{aligned} \tag{4-17}$$

for all $k\in[n_{\texttt{table,X}}]$.

Uncertainty quantification problems (4-6) correspond to the worst and best-case probabilities of ξ belonging to set X when the samples of ξ are translated by the position of y_k^X. Then, for state $q,\in Q\setminus\{q_u\}$ and action $a\in A$ we make use of the reachable set of q by a, $\mathcal{R}_{app}(q,a)$, to obtain the bounds in (4-5) by searching in the lookup table of Definition 4-2.2. Furthermore, we denote by $Y_X(q,a) \subseteq \{y_k^X\}_{k=1}^{n_{\texttt{table,X}}}$ the subset of all centers y_k^X of the table over X such that the intersection between the sub-cells of the grid with centers in y_k^X and set $\mathcal{R}_{app}(q,a)$ is nonempty:

$$Y_X(q,a) := \{y_k^X \in \{y_k^X\}_{k=1}^{n_{\texttt{table,X}}} : \text{subcell}(y_k^x) \cap \mathcal{R}_{app}(q,a) \neq \emptyset\}, \tag{4-18}$$

for all $q, \in Q \setminus \{q_u\}$, $a \in A$. Finally, we state a theorem, analogous to Theorem 4-2.1, that allows us to compute the transition probability bounds in (4-5):

Theorem 4-2.2. ***(Tight Bounds in the Transition Probabilities to q_u by Searching on the Lookup Table Over X).*** *Consider the lookup table computed as described in 4-2.2 and the set $Y_X(q,a)$ as defined in* (4-13) *for all $q \in Q \setminus \{q_u\}$, $a \in A$. Then we obtain the bounds in* (4-5) *by performing a simple grid search:*

$$\underline{P}(q,a,q_u) := 1 - \max_{y_k^X \in Y_X(q,a)} \underline{p}(y_k^X)$$
$$\overline{P}(q,a,q_u) := 1 - \min_{y_k^X \in Y_X(q,a)} \overline{p}(y_k^X).$$

The proof of Theorem 4-2.2 is analogous to that of Theorem 4-2.1 and, therefore, we omit it for simplicity.

Using the lookup tables defined in this section results in a remarkable increase in the efficiency of the abstraction. This is because we only need to compute the lookup tables for q^* and X once, and then use these same tables find every transition probability bounds. Note that the computational burden needed to compute the lookup table is way higher than the one required to search for values on it. Furthermore, once we have obtained the lookup table, the burden that computing the rest of the abstraction requires is independent of the number of samples, since we have already included this information on the lookup table. Therefore, the increase in efficiency of using the lookup table is more noticeable the higher the number of samples is. However, this increases the time required to compute the table. However, let us remember that we also make use of an alternative way to obtain the transition probability bounds, that we denote “overapproximation method" which we describe in detail in Section 4-2-2. Since the latter is less computationally demanding than searching in the lookup table, in practice, we proceed as follows: first, we compute the distance between samples and the successor set q', D_{min}. Then, if we find out that this distance is bigger than some threshold, D_{th}, we use the overapproximation method to obtain the bounds. If that is not the case, then we use the lookup table.

To finish this subsection, we must state the minimum extension of the grids of the lookup tables. These should fully cover the area in which any successor point might land (deterministically) when we do not make use of the overapproximation method, this is, when $D_{min} \leq D_{th}$. However, since we have not defined the previous parameters yet, we discuss the extension of the grids in Section 4-2-2, and not here.

4-2-2 Overapproximation Method

In section 4-2-1 we described a way to obtain the transition probability bounds in (4-4) and (4-5). However, the solution relied on precomputing a lookup table, and then searching on it, which can be relatively inefficient if used to find the transition probabilities between every pair of states $q, q' \in Q$ and under every action $a \in A$. In this section we describe a second way to obtain the bounds in (4-4) and (4-5) in an alternative way. We rely on relaxing the DR-uncertainty quantification problems on the right hand side of (4-4) and (4-5) to obtain analytical bounds on those problems. Therefore, we refer to this alternative way of obtaining the transition probability bounds as “overapproximation method". Obtaining the bounds in this way is way more efficient than searching in the lookup table. However, this procedure only yields tight bounds in the transition probabilities when the distances from the samples of the successor state x_{t+1} to the successor region q' is large.

This section is structured as follows. We focus on the inner problem of the second expression in (4-4). First, we prove that, if all the mass from the random variable x_{t+1} was concentrated in the sample of this same variable that is the closest to q', then the problem would have an analytical solution. Furthermore, we prove that this analytical solution is an upper bound on the inner problem in the second expression of (4-4). Additionally, we leverage this analytical solution to obtain a closed-form bound on the outer problem in this same expression. We prove that obtaining the desired upper bound $\overline{P}(q, a, q')$ is as simple as finding a distance. Furthermore, we also leverage this distance to obtain the lower bound $\underline{P}(q, a, q')$ in the first expression of (4-4), and the bounds in expression (4-5).

Let us start by considering the following DR-uncertainty quantification problem for a fixed current state x_t and action a_t:

$$\sup_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(f(x_t, a_t) + \xi \in q'), \tag{4-19}$$

where q' is an arbitrary rectangle of the discretization of the state space. Problem (4-19) is the inner one in (4-4), which we need to solve when we want to compute the upper bound of a transition probability. Now, we formulate problem (4-19) in the form of (3-10). To do so, consider the empirical probability distribution $\widehat{P}_{x_{t+1}}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\widehat{x}_{t+1}^i}$ constructed from the set of samples $\{\widehat{x}_{t+1}^i\}_{i=1}^N = \{f(x_t, a_t) + \widehat{\xi}^i\}_{i=1}^N$. Using this probability distribution we obtain that (4-19) is the same as:

$$\sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\widehat{P}_{x_{t+1}}^N)} P_{x_{t+1}}(x_{t+1} \in q'), \tag{4-20}$$

Problem (4-20) is a DR-uncertainty quantification problem in the form of (3-10), which we can solve using identity (3-11) and the results of Theorem 3-2.1. However, obtaining the solution requires computing a high number of distances between the samples $\{\widehat{x}_{t+1}^i\}_{i=1}^N$ and the set q'. To reduce this computational burden, we propose an approximated way to solve (4-20), which turns out quite effective when said samples are far away from q'. We illustrate this procedure in Figure 4-4 and we describe it in the following. We start by considering a bounded set $T \subset \mathbb{R}^n$ which contains the set of samples: $\{\widehat{\xi}^i\}_{i=1}^N \subset T$. This new set should be easy to work with, so we define it as the smallest ball $T = B_{r_s}(c_s)$ centered on c_s and with radius r_s that contains all the samples. Furthermore, for fixed state x_t and action a_t, ball $B_{r_s}(c_s)$ allows us to over-approximate the samples of $\{\widehat{x}_{t+1}^i\}_{i=1}^N$ by performing a translation:

$$\{\widehat{x}_{t+1}^i\}_{i=1}^N \subset \{x = f(x_t, a_t) + \zeta : \zeta \in B_{r_s}(c_s)\} = B_{r_s}(c_s + f(x_t, a_t)).$$

This new ball corresponds to a translation of $B_{r_s}(c_s)$ by quantity $f(x_t, a_t)$. Once we have obtained the overapproximating set $B_{r_s}(c_s + f(x_t, a_t))$, consider the problem of transporting as much mass as possible from said set to q'. This is a simplification of problem (4-20). To formulate such problem, let us define the points ξ_T and $\xi_{q'}$ as the points in $B_{r_s}(c_s + f(x_t, a_t))$ and q' that are the closest to q' and $B_{r_s}(c_s + f(x_t, a_t))$, respectively. This is:

$$\xi_s, \xi_{q'} := \arg \min_{\xi \in B_{r_s}(c_s + f(x_t, a_t)), \zeta \in q'} d(\xi, \zeta).$$

Then it is possible to obtain analytical bounds on the problem of transporting as much mass as possible from $B_{r_s}(c_s + f(x_t, a_t))$ to q', as we state in Proposition 4-2.3:

Proposition 4-2.3. ***(Analytical Solution of DR-Uncertainty Quantification Problems when the Nominal Distribution is a Singleton)*** *Denote by δ_{ξ_s} the probability that concentrates a unit of probability mass at ξ_s. Furthermore, let q' be an arbitrary rectangle of the discretization of the state space. Additionally, denote by $D(x_t) := d(\xi_s, \xi_{q'})$ the distance between ξ_s and $\xi_{q'}$, which is the minimum distance between points of sets $B_{r_s}(c_s + f(x_t, a_t))$ and q', for action $a \in A$. Finally, consider the maximum probability mass that we can transport from probability δ_{ξ_s} to set q' without exceeding a p-Wasserstein distance of ε. Then the following holds:*

$$\sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})} P_{x_{t+1}}(x_{t+1} \in q') = \begin{cases} (\frac{\varepsilon}{D(x_t)})^p \iff \frac{\varepsilon}{D(x_t)} \leq 1 \\ 1 \iff \frac{\varepsilon}{D(x_t)} > 1 \end{cases}, \tag{4-21}$$

Proof. The analytical solution to problem (4-21) follows from taking into account (3-11) and using the geometric interpretation described in Theorem 3-2.1. □

Now, we use the results described in Proposition 4-2.3 as an upper bound on (4-20):

Theorem 4-2.3. *The solution of* (4-21) *is an upper bound of problem* (4-20)*, that is:*

$$\sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\widehat{P}^N_{x_{t+1}})} P_{x_{t+1}}(x_{t+1} \in q') \leq \sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})} P_{x_{t+1}}(x_{t+1} \in q'). \tag{4-22}$$

Proof. Let us start by considering the approximated problem (4-21). From this definition we obtain that:

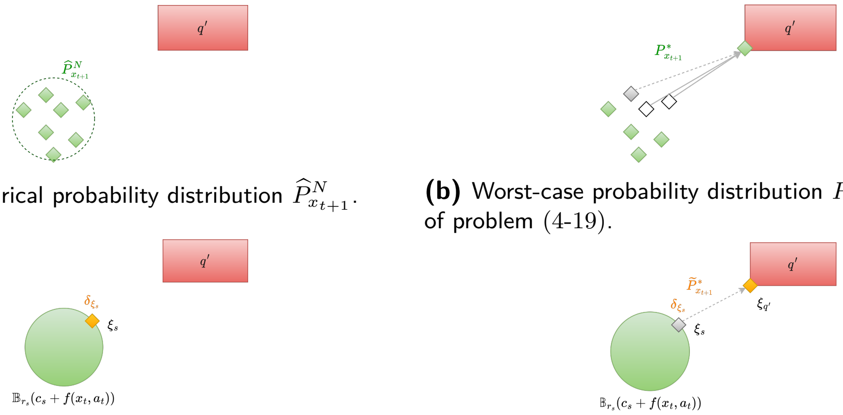
$$\sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})} P_{x_{t+1}}(x_{t+1} \in q') \geq \widetilde{P}_{x_{t+1}}(x_{t+1} \in q') \text{ for all } \widetilde{P}_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s}).$$

To prove (4-22), we just need to find a probability $\widetilde{P}_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})$ whose fraction of mass inside set q' is the same as the solution of problem (4-20), this is:

$$\widetilde{P}_{x_{t+1}}(x_{t+1} \in q') = \sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\widehat{P}^N_{x_{t+1}})} P_{x_{t+1}}(x_{t+1} \in q'). \tag{4-23}$$

Then, (4-22) would follow, since:

$$\sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\widehat{P}^N_{x_{t+1}})} P_{x_{t+1}}(x_{t+1} \in q') = \widetilde{P}_{x_{t+1}}(x_{t+1} \in q') \leq \sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})} P_{x_{t+1}}(x_{t+1} \in q').$$


q'
$\widehat{P}^N_{x_{t+1}}$
q'
$P^*_{x_{t+1}}$
q'
δ_{ξ_s}
ξ_s
$\mathbb{B}_{r_s}(c_s + f(x_t, a_t))$
q'
$\widetilde{P}^*_{x_{t+1}}$
δ_{ξ_s}
$\xi_{q'}$
ξ_s
$\mathbb{B}_{r_s}(c_s + f(x_t, a_t))$

(a) Empirical probability distribution $\widehat{P}^N_{x_{t+1}}$.

(b) Worst-case probability distribution $P^*_{x_{t+1}}$ of problem (4-19).

(c) Set $B_{r_s}(c_s + f(x_t, a_t))$ and probability distribution δ_{ξ_s}.

(d) Worst-case probability distribution $\widetilde{P}^*_{x_{t+1}}$ of problem (4-21).

Figure 4-4: Approximated way of solving (4-20) when q' is a polytope. In 4-4 and 4-4a we depict problem (4-20). In 4-4a the point on the corner of $B_{r_s}(c_s + f(x_t, a_t))$ contains mass transported from two samples of the empirical distribution and a fraction of a third sample. In the same way, in 4-4b and 4-4c we depict problem (4-21). In 4-4c a fraction of all the mass at $\xi_{\widetilde{T}(x_t)}$ is transported to $\xi_{q'}$.

Now, from Theorem 3-2.1, we know that worst case distributions that attain the supremum in (4-20) and (4-21) exist. We denote these worst-case distributions by $P^*_{x_{t+1}}$ and $\widetilde{P}^*_{x_{t+1}}$, respectively:

$$\sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\widehat{P}^N_{x_{t+1}})} P_{x_{t+1}}(x_{t+1} \in q') = P^*_{x_{t+1}}(x_{t+1} \in q')$$

$$\sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})} P_{x_{t+1}}(x_{t+1} \in q') = \widetilde{P}^*_{x_{t+1}}(x_{t+1} \in q').$$

Furthermore, from Theorem 3-2.1, we know that the worst case distributions $P^*_{x_{t+1}}$ $\widetilde{P}^*_{x_{t+1}}$ are supported on at most $N+1$ points. Additionally, we know that these points correspond to either points in the support of the empirical distributions $\widehat{P}^N_{x_{t+1}}$ and δ_{ξ_s}, respectively, or points that are the closest in q' to said support points. We denote these worst-case distributions by $P^*_{x_{t+1}}$ and $\widetilde{P}^*_{x_{t+1}}$, respectively. Therefore, for this proof it suffices to consider p-Wasserstein distances between discrete distributions. Let us express the empirical $\widehat{P}^N_{x_{t+1}}$ and worst-case $P^*_{x_{t+1}}$ distributions as the following discrete distributions:

$$\widehat{P}^N_{x_{t+1}} = \sum_{i=1}^{N} a_i \delta_i = \frac{1}{N} \sum_{i=1}^{N} \delta_i, \quad P^*_{x_{t+1}} = \sum_{j=1}^{N+1} b_j \delta_j,$$

where a_i, b_j are the weights of such measures and δ_i represents unit mass concentrated at point i of the support of the corresponding measure. This allows us to express the p-Wasserstein distance between measures $\widehat{P}^N_{x_{t+1}}$ and $P^*_{x_{t+1}}$ in Kantorovich' formulation as:

$$\mathcal{W}_p^p(\widehat{P}^N_{x_{t+1}}, P^*_{x_{t+1}}) = \min_{\pi \in \mathcal{U}(\widehat{P}^N_{x_{t+1}}, P^*_{x_{t+1}})} \sum_{i=1}^{N} \sum_{j=1}^{N+1} d^p_{i,j} \pi_{i,j}, \tag{4-24}$$

where

$$\mathcal{U}(\widehat{P}^N_{x_{t+1}}, P^*_{x_{t+1}}) = \tag{4-25}$$
$$\{\pi \in \mathbb{R}^{N\times(N+1)} : \sum_{i=1}^{N} \pi_{i,j} = b_j \text{ for all } j \in [N+1], \sum_{j=1}^{N+1} \pi_{i,j} = a_i \text{ for all } i \in [N]\}. \tag{4-26}$$

In the expressions (4-24) and (4-25), $\pi_{i,j}$ is the coupling that determines how much probability mass is transported from sample $\widehat{x}^i_{t+1}$ of the empirical distribution $\widehat{P}^N_{x_{t+1}}$ to point j of $P^*_{x_{t+1}}$. Furthermore, we denote by $d_{i,j}$ the distance between these two points.

Now, remember that we needed to find a probability $\widetilde{P}_{x_{t+1}}$ such that it contains the same fraction of mass inside q' than $P^*_{x_{t+1}}$. For this reason, we choose $\widetilde{P}_{x_{t+1}}$ as the probability supported in two points, ξ_s and $\xi_{q'}$, as shown in Figure 4-4d:

$$\widetilde{P}_{x_{t+1}} := P^*_{x_{t+1}}(x_{t+1} \in q')\delta_{\xi_{q'}} + P^*_{x_{t+1}}(x_{t+1} \notin q')\delta_{\xi_s}.$$

This probability is the result of concentrating all mass of the worst-case probability $P^*_{x_{t+1}}$ that lies inside q' in $\xi_{q'}$, and all the mass that lies outside of q' in ξ_s. Measure $\widetilde{P}_{x_{t+1}}$ can also be considered as a probability supported on $N+1$ points in the same way that $P^*_{x_{t+1}}$ is, but where these points overlap in either ξ_s or $\xi_{q'}$. Figure 4-4d helps to visualize this fact. Note that, by definition of probability $\widetilde{P}_{x_{t+1}}$, its probability mass inside of q' is the same as the solution of problem (4-20). Now, we prove that this probability lies inside the p-Wasserstein ball $\mathbb{B}_\varepsilon(\delta_{\xi_s})$. The intuition behind this proof is related to noticing that the amount of mass this distribution concentrates in $\xi_{q'}$, $P^*_{x_{t+1}}(x_{t+1} \in q')$, is lower than the one $\widetilde{P}^*_{x_{t+1}}$ concentrates in that same point. This is due to the fact that the distance $d(\xi_{q'}, \xi_s)$ is lower than the one between any sample x^i_{t+1} for all $i \leq N$ and set q'. This allows us to transport more mass to q' with the same budget ε. We denote by π^* the coupling that attains the minimum in (4-24). The p-Wasserstein distance between the measures δ_{ξ_s} and $\widetilde{P}_{x_{t+1}}$ is:

$$\mathcal{W}^p_p(\delta_{\xi_s}, \widetilde{P}_{x_{t+1}}) = \min_{\pi \in \mathcal{U}(\delta_{\xi_s}, \widetilde{P}_{x_{t+1}})} \sum_{i=1}^{N} \sum_{j=1}^{N+1} \widetilde{d}^p_{i,j}\pi_{i,j}, \tag{4-27}$$

where $\widetilde{d}_{i,j}$ denotes the distance between points i and j in the support of both measures. Now, we notice that the coupling π^* that attained the minimum transport cost between measures $\widehat{P}^N_{x_{t+1}}$ and $\widetilde{P}^*_{x_{t+1}}$ is a feasible coupling in problem (4-27), that is, $\pi^* \in \mathcal{U}(\delta_{\xi_s}, \widetilde{P}_{x_{t+1}})$. Taking this consideration into account in (4-27) we get that:

$$\min_{\pi \in \mathcal{U}(\delta_{\xi_s}, \widetilde{P}_{x_{t+1}})} \sum_{i=1}^{N} \sum_{j=1}^{N+1} \widetilde{d}^p_{i,j}\pi_{i,j} \leq \sum_{i=1}^{N} \sum_{j=1}^{N+1} \widetilde{d}^p_{i,j}\pi^*_{i,j}. \tag{4-28}$$

Furthermore, we notice that all transport distances $\widetilde{d}_{i,j}$ between points in the support of δ_{ξ_s} and $\widetilde{P}_{x_{t+1}}$, when compared to distances $d_{i,j}$ between points in the support of $\widehat{P}^N_{x_{t+1}}$ and $P^*_{x_{t+1}}$ are always smaller: $\widetilde{d}_{i,j} \leq d_{i,j}$ for all $i \in [N]$, $j \in [N+1]$. This follows from the way we have defined ξ_s and $\widetilde{P}_{x_{t+1}}$. Using this result and the right-hand side of expression (4-28), we have that:

$$\sum_{i=1}^{N} \sum_{j=1}^{N+1} \widetilde{d}^p_{i,j}\pi^*_{i,j} \leq \sum_{i=1}^{N} \sum_{j=1}^{N+1} d^p_{i,j}\pi^*_{i,j}.$$

Notice that the right-hand side of the last expression is the p-Wasserstein distance (4-24). From this result it follows that:

$$\mathcal{W}_p(\delta_{\xi_s}, \widetilde{P}_{x_{t+1}}) \leq \mathcal{W}_p(\widehat{P}^N_{x_{t+1}}, P^*_{x_{t+1}}) \leq \varepsilon,$$

which means that probability $\widetilde{P}_{x_{t+1}}$ fulfills that $\widetilde{P}_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})$. Therefore, Theorem 4-2.3 follows. □

Since the approximated problem (4-21) has an analytical solution, using his method effectively reduces the computational burden of the DR-uncertainty quantification problem: we do not need to compute N distances to solve (4-20). As a drawback, the approximation is conservative: it leads to a higher upper bound in the transition probabilities, especially when the distances between samples $\{\widehat{x}^i_{t+1}\}^N_{i=1}$ and set q' are large. We should also highlight that any choice of the overapproximating set T is valid. Furthermore, we could also over-approximate the set q' by a simpler set, and then work with this set instead. The objective of doing so is to reduce the complexity of the problem we need to solve to compute the distance $D(x_t)$. However, this reduction in computational burden entails, as a trade-off, an increase in the conservatism of the results obtained. This means that choosing simpler over-approximating sets leads to a higher value of the solution of problem (4-20). For example, the simplest sets we can think of are balls that over approximate $\{\widehat{\xi}^i\}^N_{i=1}$ and q'. Computing the distance $D(x_t)$ in this case is trivial, but the obtained value of $D(x_t)$ is relatively conservative with respect to other choices. A different choice could be to choose polytopes to over-approximate our sets. With this choice, computing $D(x_t)$ requires that we solve a convex optimization problem, which is quadratic if the distance is the euclidean one ($p = 2$). These shapes lead to a less conservative result, since the distance $D(x_t)$ is smaller than if we use spheres.

When computing the transition probability bounds in (4-4) and (4-5), the set q' is always a rectangle. However, we make the final decision to only use balls to overapproximate the set of samples since, in this case, we are able to compute the minimum distance to set q' without having to solve any optimization problem: we only need to compute the distance from the center of the ball $B_{r_s}(c_s + f(x_t, a_t))$ that over approximates the set $\{\widehat{x}^i_{t+1}\}^N_{i=1}$ to q' by using a set of logical “ifs". This is because the distance from the center $c_s + f(x_t, a_t)$ to the faces that define the rectangle q' has a closed form expression. Then we obtain the distance $D(x_t)$ by subtracting the radius r_s of the overapproximating ball. The complexity of this approximated way of obtaining the bounds in (4-4) is the following: for fixed $x_t \in q$, $a_t \in A$, we compute once the over-approximating ball $B_{r_s}(c_s + f(x_t, a_t))$ and then $D(x_t)$ for each $q, q' \in Q$. When compared to also having to compute N distances and additional computations in the exact case, the approximated method highly reduces the computational burden.

Analytical Transition Probability Bounds Using the Approximate Approach

Previously in this section we have obtained an analytical upper bound on problem (4-19). Consider that we have computed said upper bound as described in in Proposition 4-2.3. Now, to obtain the upper bound in the transition probabilities $\overline{P}(q, a, q')$ we need to solve the outer optimization problem over x_t, as we notice in the second inequality of (4-4). In this section we leverage the results we obtained previously to obtain the transition probability bounds in (4-4) and (4-5) analytically.

The organisation of this section is the following. First, we show that, when making use of the overapproximation method described previously in this section, solving the maximization over x_t is equivalent to finding the smallest distance $\min D(x_t)$ when $x_t \in q$. We also show that, fortunately, this outer optimization problem is, in fact, a convex program when the dynamics f of the system are linear. Furthermore, we prove that, in the general case that the dynamics are nonlinear, we are still able to obtain the upper bound $\overline{P}(q, a, q')$ in the second problem of (4-4) in an efficient way by obtaining an under estimator of the distance previously mentioned. Additionally, we prove that obtaining this distance also allows us to find the lower bound $\underline{P}(q, a, q')$ in (4-4) in some cases. Moreover, we prove that we are able to obtain also the bounds in (4-5) in a similar way. Furthermore, we define a threshold condition that must be fulfilled to use this method instead of the one that relies on the lookup tables. Finally, we also define the minimum extension that the grids for which we have computed the lookup tables in Section 4-2-1 must cover.

We begin by stating the following theorem, which defines the upper bound $\overline{P}(q, a, q')$ as a function of the smallest distance $\min D(x_t)$ when $x_t \in q$:

Theorem 4-2.4. ***(Analytical upper Bound in the Transition Probabilities using the overapproximation Method)*** *Consider that we make use of the overapproximation method from Proposition 4-2.3 to obtain an upper bound on the solution of the inner problems in the second expression of* (4-4)*. Finally, for fixed $q, q' \in Q \setminus \{q_u\}$, $a \in A$, denote by $\underline{D}$ the minimum distance between ball $B_{r_s}(c_s + f(x_t, a_t))$ and set q', for all possible values of $x_t \in q$. Then $\overline{P}(q, a, q')$ has the following analytical expression:*

$$\overline{P}(q, a, q') = \begin{cases} (\frac{\varepsilon}{\underline{D}})^p & \textit{if } \frac{\varepsilon}{\underline{D}} \leq 1 \\ 1 & \textit{if } \frac{\varepsilon}{\underline{D}} > 1 \end{cases}. \tag{4-29}$$

Proof. The proof follows by using the results in Theorem 4-2.3 and Proposition 4-2.3 in the second expression of (4-4):

$$\max_{x_t \in q} \sup_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(f(x_t, a_t) + \xi \in q') \leq \max_{x_t \in q} \sup_{P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\delta_{\xi_s})} P_{x_{t+1}}(x_{t+1} \in q') \tag{4-30}$$

$$= \begin{cases} \max_{x_t \in q}(\frac{\varepsilon}{D(x_t)})^p = (\frac{\varepsilon}{\underline{D}})^p & \text{if } \frac{\varepsilon}{\underline{D}} \leq 1 \\ 1 & \text{if } \frac{\varepsilon}{\underline{D}} > 1 \end{cases}. \tag{4-31}$$

□

Therefore, if we are able to find the x_t that minimizes the distance $D(x_t)$ from set $B_{r_s}(c_s + f(x_t, a_t))$ that contains the samples to the polytope q', we could easily compute the upper bound $\overline{P}(q, a, q')$. When the set q' is convex, we can obtain the minimum distance $\underline{D}$ by solving the following optimization problem:

$$\underline{D} := \min_{x_t \in q} D(x_t) = \min_{x_t \in q} \min_{\xi \in B_{r_s}(c_s + f(x_t, a_t)), \zeta \in q'} d(\xi, \zeta) = \min_{x_t \in q, \xi \in B_{r_s}(c_s + f(x_t, a_t)), \zeta \in q'} d(\xi, \zeta), \tag{4-32}$$

which is convex if the dynamics are linear in x_t, since d is a convex function and the constraints, including $\xi \in B_{r_s}(c_s + f(x_t, a_t))$, are convex.

In practice, all sets $q \in Q \setminus \{q_u\}$ are rectangles for simplicity. Furthermore, if we used a polytopic approximation of the samples instead of a ball, and we use the euclidean distance, problem (4-32) would become a quadratic program. However, we chose to use balls since, it allows us to further reformulate the problem in such a way that no optimization problem needs to be solved, as we already pointed out at the end of Section 4-2-2. This method works also when the dynamics are nonlinear and, therefore, we describe it in detail only in the context of such systems. However, remember that we can also make use of that method when the dynamics are linear in the state. Opposite to the linear case, when the dynamics are nonlinear in f for any fixed action a, the problem becomes nonconvex due to the non-convexity of constraint $\xi \in B_{r_s}(c_s + f(x_t, a_t))$ in (4-32). However, obtaining $\underline{D}$ exactly as defined in that expression is not needed: it suffices to find an under estimator D_{min} of that distance. Then we could obtain the upper transition probability bound by using D_{min} instead of $\underline{D}$ in (4-29). In the following theorem we define such under estimator D_{min}:

Theorem 4-2.5. ***(Under estimator of*** $\underline{D}$***)*** *Consider the same states $q, q' \in Q \setminus \{q_u\}$ and action $a \in A$ as in Theorem 4-2.4. Denote by L_a the Lipschitz constant of the function $f(\cdot, a)$ for fixed action a. Consider also the smallest ball $B_{\delta_q}(c_q) \supset q$ with center at some point c_q and radius δ_q that contains the set q. Then the distance*

$$D_{min} := \min_{\xi \in B_{L_a \delta_q + r_s}(f(c_q, a) + c_s), \zeta \in q'} d(\xi, \zeta) \tag{4-33}$$

is an under estimator of $\underline{D}$.

Proof. In this proof we drop the time index in state x and action a for simplicity. Using the Lipschitz constant L_a we obtain an overapproximation of the set reached from q under a: $\mathcal{R}(q, a) = \{f(x, a) \in \mathbb{R}^n : x \in q\} \subset B_{L_a \delta_q}(f(c_q, a))$. Using this reachable set, we obtain an overapproximation of the set to which ξ belongs in (4-32):

$$\bigcup_{x \in q} B_{r_s}(c_s + f(x, a)) \tag{4-34}$$

$$\subset \bigcup_{x \in B_{\delta_q}(c_q)} B_{r_s}(c_s + f(x, a)) \tag{4-35}$$

$$\subset \bigcup_{y \in B_{L_a \delta_q}(f(c_q, a))} B_{r_s}(c_s + y) \tag{4-36}$$

$$= B_{L_a \delta_q + r_s}(f(c_q, a) + c_s). \tag{4-37}$$

From expressions (4-34) we get:

$$\underline{D} \geq \min_{\xi \in B_{L_a \delta_q + r_s}(f(c_q, a) + c_s), \zeta \in q'} d(\xi, \zeta).$$

Therefore Theorem 4-2.5 follows. □

In Figure 4-5 we depict the distance D_{min}. Computing D_{min} as defined in Theorem 4-2.5 requires solving a convex program. Furthermore, since the set of allowed values of ξ is a ball, as we described before, we can compute this distance without the need to solve any optimization problem: we can use a set of “ifs", and obtain the distances from the center $f(c_q, a) + c_s$ of the ball to the faces that define the rectangle q' in closed form. After that we

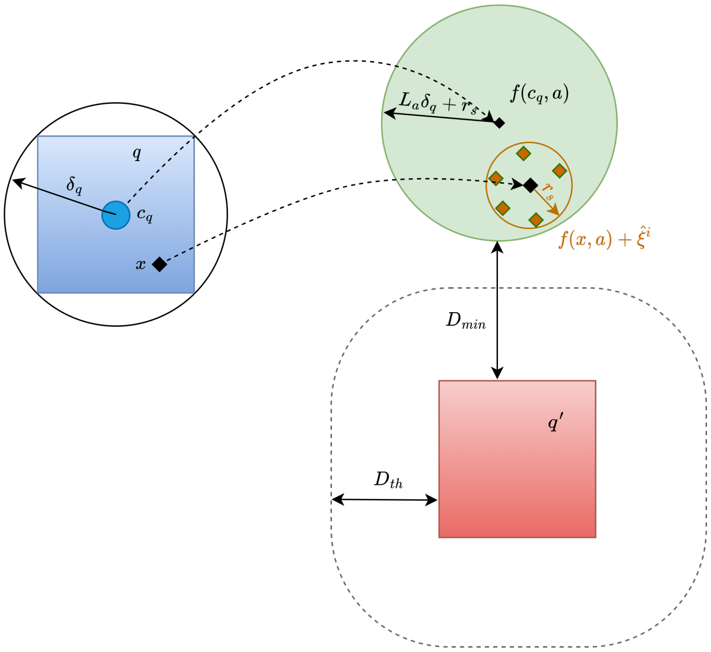


Figure 4-5: Graphical representation of the procedure used to obtain an lower bound of $\underline{D}$: D_{min}. Note that in this figure, $c_s = 0$ for clarity.

just need to pick the minimum distance from those and subtract the radius $r_s + L_a\delta_q$ of the ball. Throughout the entirety of Section 4-2-2 we have only focused on obtaining an upper bound in the transition probabilities in (4-4). However, now we prove that the results of this section allow us to also obtain a tight lower bound, without the need to resort to the more expensive method that relies on a lookup table.

Theorem 4-2.6. ***(Analytical Lower bound in the Transition Probabilities using the Overapproximation Method)*** *Consider the states $q, q' \in Q \setminus \{q_u\}$ and action $a \in A$. Consider also the distance D_{min} as defined in Theorem 4-2.5. Then the following holds: if $D_{min} > 0$, then $\underline{P}(q, a, q') = 0$.*

Proof. Consider the right-hand-side of the upper inequality in (4-4). Then we get that:

$$\underline{P}(q,a,q') \leq \min_{x\in q} \inf_{P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)} P_\xi(f(x,a)+\xi \in q') \leq \min_{x\in q} \widehat{P}_\xi^N(f(x,a)+\xi \in q').$$

Furthermore, the condition $\underline{D}_{min} > 0$ means that all samples $\{f(x,a)+\widehat{\xi}^i\}_{i=1}^N$ lay outside q' for all $x \in q$:

$$d(f(x,a)+\widehat{\xi}^i, q') \geq D_{min} \text{ for all } x \in q,\ i \in [N].$$

Therefore, Theorem 4-2.6 follows. □

Finally, we describe how we obtain the bounds in (4-5), which are those in the probability of transitioning to q_u. We follow a similar approach than the one we used to compute the bounds in (4-4). We first notice that the results of Proposition 4-21 and Theorems 4-2.3 and 4-2.6 still hold when q' is an arbitrary set, for example, X, or $\mathbb{R}^n \setminus X$. Then we state the following theorem, which is analogous to Theorem 4-2.4:

Theorem 4-2.7. ***(Analytical Bounds in the Transition Probabilities to q_u using the Overapproximation Method)*** *Consider a fixed state $q \in Q \setminus \{q_u\}$ and action $a \in A$. Furthermore, consider that we have obtained the under estimators $D_{min,X}$ and D_{min,q_u} of the distances between samples $\{f(x_t, a_t)+\widehat{\xi}^i\}_{i=1}^N$ and set X and $\mathbb{R}^n \setminus X$, respectively, for all $x_t \in q$. Then, if $D_{min,q_u} > 0$, the following holds*

$$\begin{aligned} \overline{P}(q,a,q_u) &= \begin{cases} (\frac{\varepsilon}{D_{min,q_u}})^p & \mathit{if} \frac{\varepsilon}{D_{min,q_u}} \leq 1 \\ 1 & \mathit{if} \frac{\varepsilon}{D_{min,q_u}} > 1 \end{cases} \\ \underline{P}(q,a,q_u) &= 0 \end{aligned} \tag{4-38}$$

On the other hand, if $D_{min,X} > 0$, the following holds

$$\begin{aligned} \overline{P}(q,a,q_u) &= 1 \\ \underline{P}(q,a,q_u) &= \begin{cases} 1-(\frac{\varepsilon}{D_{min,X}})^p & \mathit{if} \frac{\varepsilon}{D_{min,X}} \leq 1 \\ 0 & \mathit{if} \frac{\varepsilon}{D_{min,X}} > 1 \end{cases} \end{aligned} \tag{4-39}$$

The proof of Theorem 4-2.7 is similar to that of Theorems 4-2.4 and 4-2.6. Therefore, we omit it for simplicity. Notice that we have not specified the way in which we have found the under estimators $D_{\text{min},X}$ and D_{min,q_u}, which allows for generality: the results of Theorem 4-2.7 are useful when the dynamics are both linear and nonlinear.

By using the overapproximation method, we reduce the computational burden required to compute the bounds, since no search in the lookup table needs to be performed. Furthermore, all the auxiliary operations required to compute the set $\mathcal{R}_{app}(q,a)$ and its translated equivalent are no longer required. We just need to use the results of Theorem 4-2.5 and Theorem 4-2.4 to obtain the desired upper bound $\overline{P}(q,a,q')$. We need to solve problem (4-33) only once for each $q, q' \in Q$ and $a \in A$. Furthermore, using the results from Theorems 4-2.4 and 4-2.6, we obtain the desired lower and upper bounds $\underline{P}(q,a,q')$ and $\overline{P}(q,a,q')$, if D_{min} is found to be bigger than zero. Furthermore, the complexity of this method is independent of the number of samples N, since we only need to compute the overapproximation T of the sample set once. This is due to the additive nature of the disturbance. Practical results show that this methods is way faster than the one that relies on the lookup table. However, since it only leads to tight bounds when the computed distance D_{min} is relatively large, we only make use of it when this distance exceeds some threshold: $D_{min} \geq D_{th}$. This threshold depends of the problem at hand: mainly in how fine the state discretization is and in how spread the samples are.

To finish this section, we use the parameters that we defined here to compute a lower bound in the extension of the grids of the lookup tables that we used in Section 4-2-1. Let us take into account the biggest Lipschitz constant L of the dynamics $f(\cdot, a)$ for all actions: $L = \max\{L_a : a \in A\}$. Let us also consider the radius δ_q of the smallest ball $B_{\delta_q}(c_q)$ containing the region q of the state space discretization and the radius r_s of the ball that contains the samples. Then, we find a lower bound for the distance D^*: $D^* \geq D_{th} + 2L\delta_q + r_s$, up to which the grids should extend from q^* and X, respectively. This distance is also depicted in Figure 4-1 for the case of the grid that covers q^*.

4-3 Strategy Synthesis for DR-IMDP Abstractions

In Section 4-2 we described the process of obtaining a DR-IMDP abstraction $\mathcal{I}$ of system (4-1) that accounts for the distributional ambiguity in its disturbance. Once we have obtained said abstraction, we synthesize a strategy for this one by making use of the approach described in Section 2-4-1 for IMDPs, which yields the following results: strategy $\sigma_{\mathcal{I}}^*$, and the bounds $\underline{p}_{\mathcal{I}}^k(q)$ and $\overline{p}_{\mathcal{I}}^k(q)$ in the probability of the paths of $\mathcal{I}$ satisfying ϕ within $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ steps while never exiting X by following strategy $\sigma_{\mathcal{I}}^*$ and starting from state $q \in Q$. Furthermore, strategy $\sigma_{\mathcal{I}}^*$ is a memory-dependent strategy that maximizes the lower bound $\underline{p}_{\mathcal{I}}^k(q)$ for each $q \in Q$, $k \in \mathbb{N} \cup \{\infty\}$.

4-4 Correctness

Once we have synthesized a strategy for the DR-IMDP abstraction by following the procedure we described in Section 2-4-1, we need to address the following questions. The first one is: how can the original system use this strategy of the DR-IMDP? Furthermore, the second one is: how are we sure that, when the original system uses this strategy, the probabilistic guarantees found for the abstraction hold also for the original system? In this section we address the previous two problems: first we translate strategy $\sigma_{\mathcal{I}}^*$ of $\mathcal{I}$ to the (also memory-dependent) strategy $\sigma_{\mathcal{C}}^*$ of system (4-1). Secondly, we prove that the probabilistic guarantees $\underline{p}_{\mathcal{I}}^k$ and $\overline{p}_{\mathcal{I}}^k$ obtained for the DR-IMDP abstraction hold for the original system, when this one follows strategy $\sigma_{\mathcal{C}}^*$. This is referred to as correctness of the abstraction [11], [4],[7].

First, we refine strategy $\sigma_{\mathcal{I}}^*$ of $\mathcal{I}$ into a strategy $\sigma_{\mathcal{C}}^*$ of system (4-1) that maps a finite path of the latter into an action $a \in A$. To do so we first define a function $J : \mathbb{R}^n \to Q$ that maps the continuous state $x \in \mathbb{R}^n$ to the corresponding discrete state $q \in Q$ of $\mathcal{I}$. Formally, for any $x \in \mathbb{R}^n$, $J(x) = q$ if and only if $x \in q$. We can also use this function to map finite paths $w_{\mathbb{R}^n}^t = x(0)x(1)\ldots x(t)$ of the system (4-1) to finite paths $w_Q^t = q(0)q(1)\ldots q(t)$ of $\mathcal{I}$. With a small abuse of notation, we say

$$w_Q^t = J(w_{\mathbb{R}^n}^t) = J(x(0))J(x(1))\ldots J(x(t)) = q(0)q(1)\ldots q(t)$$

when $J(x(i)) = q(i)$ for all $i \in \{0, 1, \ldots, t\}$. The previous result allows us to refine strategy $\sigma_{\mathcal{I}}^*$ into strategy $\sigma_{\mathcal{C}}^* = \{\sigma_{\mathcal{C}}^*(w_{\mathbb{R}^n}^t; k)\}_{k=1}^{\infty}$ of the original system as follows:

$$\sigma_{\mathcal{C}}^*(w_{\mathbb{R}^n}^t; k) := \sigma_{\mathcal{I}}^*(J(w_{\mathbb{R}^n}^t); k), \tag{4-40}$$

where $k \in \mathbb{N} \cup \{\infty\}$ is the number of steps until the horizon of the bellman recursion (2-7).

Secondly, we want to prove that the guarantees obtained for the abstraction hold for the original system. To do that, we begin by stating Lemma 4-4.1:

Lemma 4-4.1. *Consider the ambiguity set $\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$ of distributions of the disturbance ξ, and the arbitrary distribution $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$. Assume that system* (4-1) *has been abstracted into the DR-IMDP $\mathcal{I}$ as described in Section 4-2. Consider, for fixed state $q \in Q$ and action $a \in A$ the transition probability bounds $\underline{P}(q, a, \cdot)$ and $\overline{P}(q, a, \cdot)$ of $\mathcal{I}$. Moreover, consider the continuous state $x \in q \subset \mathbb{R}^n$. Now, denote by $P_{x_{t+1}}$ the probability of the (continuous) successor state*

x_{t+1} *from* x *under* a *and for* P_ξ*. Using the kernel* $\mathcal{T}_\mathcal{C}$ *in* (4-2)*, the previous distribution is obtained as:*

$$P_{x_{t+1}}(D) := \mathcal{T}_\mathcal{C}(D|x, a; P_\xi)$$

for any measurable set D*. Then the state of system* (4-1) *after one time step,* x_{t+1}*, lies on cell* $q' \in Q$ *with probability* $P_{x_{t+1}}(q') \in [\underline{P}(q, a, q'), \overline{P}(q, a, q')]$*, for all* $q' \in Q$*,* $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$*.*

Proof. The proof follows directly from the fact that the transition probability bounds of $\mathcal{I}$ must fulfill expressions (4-4) and (4-5). □

Denote by $P(w^k_{\mathbb{R}^n} \models \phi | w^k_{\mathbb{R}^n}(0) = x, X, \sigma^*_\mathcal{C}, P_\xi)$ the probability of the paths of system (4-1), for P_ξ, satisfying ϕ within k steps, while remaining in the safe region $X \subset \mathbb{R}^n$, by following strategy $\sigma^*_\mathcal{C}$ and starting from $x \in \mathbb{R}^n$. Now we state the theorem that ensures that the guarantees obtained for the abstraction hold for the original system:

Theorem 4-4.1. ***(Correctness of the Probabilistic Guarantees Of the DR-IMDP Abstraction)*** *Consider system* (4-1)*, ambiguity set* $\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$*, and the DR-IMDP abstraction* $\mathcal{I}$ *of such system obtained as in Section 4-2. Furthermore, consider an scLTL formula* ϕ*, and the strategy* $\sigma^*_\mathcal{I}$ *and bounds* $\underline{p}^k_\mathcal{I}$ *and* $\overline{p}^k_\mathcal{I}$ *of* $\mathcal{I}$ *obtained for such specification as explained in Section 4-3. Consider also strategy* $\sigma^*_\mathcal{C}$ *of system* (4-1)*, as defined in expression* (4-40)*. Then for any* $x \in \mathbb{R}^n$ *where* $x \in q$*, and for any* $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$ *it holds that*

$$P(w^k_{\mathbb{R}^n} \models \phi | w^k_{\mathbb{R}^n}(0) = x, X, \sigma^*_\mathcal{C}, P_\xi) \in [\underline{p}^k_\mathcal{I}(q), \overline{p}^k_\mathcal{I}(q)]$$

for all $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$*.*

We provide a detailed proof of Theorem 4-4.1 in Appendix C. Said proof, despite being tailored to the more general setting of robust MDP abstractions, also holds for DR-IMDPs: we only need to consider $\Gamma_{q,a}$, for all $q \in Q$, $a \in A$ as the feasible sets of transition probabilities.

Now, if ε has been chosen in such a way that $P_\xi^{\texttt{true}} \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$ holds with high confidence [2], we can easily compute a lower bound in the satisfaction probability of system (4-1) under $P_\xi^{\texttt{true}}$: for any $x \in \mathbb{R}^n$ where $x \in q$ it holds that

$$P(w^k_{\mathbb{R}^n} \models \phi | w^k_{\mathbb{R}^n}(0) = x, X, \sigma^*_\mathcal{C}, P_\xi^{\texttt{true}}) \geq (1 - \beta) \cdot \underline{p}^k_\mathcal{I}(q),$$

where $1 - \beta$ is the confidence level.

[2] see Section 3-2.

Chapter 5

Strategy Synthesis via Abstractions to Robust MDPs

In this chapter we introduce our second approach of synthesizing a strategy for an uncertain system. Again, we follow an abstraction-based approach: we abstract the system to a class of MDP with uncertain transition probabilities in a way that the abstraction accounts for the uncertain behaviour of the system. After that, we synthesize a strategy for the abstraction. However, the setting we consider here is more general that the data-driven one: we have access to a nominal distribution for the disturbance, which does not need to be finitely-supported, and we want to be robust with respect to small deviations from this one. This includes distributions built from samples, Gaussian distributions, etc. We call such abstractions of systems under this distributional ambiguity “robust MDPs". Once we have obtained the robust MDP abstraction, we synthesize a strategy for this one that is robust with respect to all ambiguities of the abstraction. With that purpose, we propose a modified value iteration algorithm for reachability, which is analogous to the one described in 2-4-1 for IMDPs.

This chapter is structured as follows: first, in Section 5-1 we formally define the class of systems that we abstract to robust MDPs, and we formally state our problem. Next, in Section 5-2, we formally define said robust MDP abstractions, and we describe how to obtain them. After that, in Section 5-3 we describe the algorithms that allow us to synthesize a strategy that enforces a complex specification for these abstractions. Then, in Section 5-4 we correctly translate the strategy and satisfaction guarantees obtained for the abstraction to the original system.

5-1 Problem Statement

In this section we formally state our problem, however, let us begin by describing the class of systems that we consider. Consider the following discrete-time dynamical system:

$$x_{t+1} = f(x_t, u_t) + \xi_t, \tag{5-1}$$

where $x_t \in \mathbb{R}^n$ is the state, $u_t \in U_{\mathcal{C}}$ is the control input and $\xi_t \in \mathbb{R}^n$ is a random disturbance. This disturbance process is i.i.d. and has probability distribution $P_\xi^{\texttt{true}}$, which is unknown to us. Furthermore, notice that, for simplicity and without loss of generality, we consider that the dimension of the disturbance is the same as that of the state space, $n \in \mathbb{N}$. Additionally, we associate to system (5-1) a set of observations $O_{\mathcal{C}}$, and an observation function $L_{\mathcal{C}}$ that assigns an observation $o \in O_{\mathcal{C}}$ to every state $x \in \mathbb{R}^n$. Furthermore, we also associate to system (5-1) the transition kernel (4-2). The concept of paths, traces and strategy of (5-1) are the same as those defined in 4-1 and, therefore, we do not sate them again to avoid repetition.

Once we have defined the class of systems we consider, let us state the problem that we need to solve. Denote by $P(w_{\mathbb{R}^n}^k \models \phi | w_{\mathbb{R}^n}^k(0) = x, X, \sigma_{\mathcal{C}}, P_\xi)$ the probability of the paths of system (5-1), for probability P_ξ of ξ, satisfying the scLTL formula ϕ within $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ steps while staying inside $X \subset \mathbb{R}^n$ by starting from state $x \in X$ and following strategy $\sigma_{\mathcal{C}}$.

Problem 5-1.1. ***(Distributionally Robust Strategy Synthesis)*** *Consider the dynamical system* (5-1)*. Consider also the p-Wasserstein distance-based ambiguity set $\mathbb{B}_\varepsilon(\widehat{P}_\xi)$, of radius $\varepsilon > 0$ and centered on a nominal distribution $\widehat{P}_\xi$. Consider also a compact set $X \subset \mathbb{R}^n$ and an scLTL formula ϕ defined over the regions of interest of X. Then, find a near-optimal strategy $\sigma_{\mathcal{C}}^*$ that allows to determine if for given initial state $x \in X$, probability threshold $p_{\texttt{th}}$ and horizon $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$*

$$P(w_{\mathbb{R}^n}^k \models \phi | w_{\mathbb{R}^n}^k(0) = x, X, \sigma_{\mathcal{C}}^*, P_\xi) \geq p_{\texttt{th}} \tag{5-2}$$

holds for all probabilities $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$.

Then, if the ambiguity set is such that $P_\xi^{\texttt{true}} \in \mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$, we have formal guarantees that (5-2) will hold for the original system (5-1) under the true distribution $P_\xi^{\texttt{true}}$ of ξ. However, again, notice that we leave this last part outside the statement of problem 5-1.1 because in this thesis we assume given ambiguity sets. The problem of finding the size of the ambiguity set that contains $P_\xi^{\texttt{true}}$ is out of the scope of this thesis.

We seek a near-optimal strategy because there is no way of finding the exact strategy that maximizes the probability in (4-3) [11]. Instead, as we already pointed out in the introduction of this chapter, we follow an abstraction-based approach to solve Problem 5-1.1. First, we abstract system (5-1), for fixed distribution of the disturbance $\widehat{P}_\xi$, into an IMDP $\widehat{\mathcal{I}}$, which accounts only for the discretization of the state-space. We denote this IMDP as “nominal". Then we expand the set of feasible transition probabilities of $\widehat{\mathcal{I}}$ to account for the distributional uncertainty about the disturbance. We denote the resulting abstraction “robust MDP". Then we synthesize the strategy that maximizes the worst-case probability of the paths of said robust MDP satisfying the specification. To do so, we use a modified value iteration algorithm, which we name “robust value iteration". Finally we translate said strategy into a strategy that system (5-1) is able to use, while preserving the guarantees obtained for the abstraction.

5-2 Obtaining Robust MDP Abstractions

As we already pointed out, we solve Problem 5-1.1 by following an abstraction-based approach. This one relies on the theory of IMDPs and the Wasserstein distance explained in Sections 2-4 and 3-1 respectively. In this section we describe how we obtain these robust MDP abstractions. We rely on precomputing an IMDP abstraction of the original system when the probability of the disturbance is assumed to be the fixed to the nominal probability $\widehat{P}_\xi$. This nominal abstraction accounts for the ambiguity introduced by the state discretization. After that, we obtain an additional abstraction $\mathcal{M}^R$ that also accounts for the distributional ambiguity by expanding the set of feasible transition probabilities of $\widehat{\mathcal{I}}$. We denote the resulting abstraction "robust MDP". We start by formally defining the nominal IMDP abstraction in Section 5-2-1. After that, we describe how to use it to obtain the desired robust MDP abstraction in Section 5-2-2.

5-2-1 Nominal IMDP Abstraction

The process of obtaining a robust MDP abstraction relies in precomputing an IMDP abstraction of the original system (5-1) when the probability of the disturbance is assumed to be the fixed nominal probability $\widehat{P}_\xi$. We denote this IMDP abstraction by "nominal" IMDP, $\widehat{\mathcal{I}} = (Q, A, \underline{P}, \overline{P}, O, L)$, which we formally define in this section.

We obtain the state space Q, action A and observation O sets and observation function L of $\widehat{\mathcal{I}}$ in the following way. Let us begin with the state space. We focus on a compact set $X \subseteq \mathbb{R}^n$ of the state space of system (5-1). We discretize this one into a finite number of non-overlapping regions $\widetilde{Q} := \{q_1, q_2, \ldots, q_{|\widetilde{Q}|}\}$ such that:

$$\cup_{q \in \widetilde{Q}} q = X, \quad \text{and} \quad q \cap q' = \emptyset \quad \forall q, q' \in \widetilde{Q} \, \text{and} \, q \neq q'.$$

In the setting of this chapter, unlike in Chapter 4, we allow the partition to be non-uniform, since there is no advantage from using such partitions. Now, we assign each region q of the discretization to a different state in the nominal IMDP $\widehat{\mathcal{I}}$. With an abuse of notation we refer by q to both, a state $q \in \widetilde{Q}$ of $\widehat{\mathcal{I}}$, and to a region $q \subset X$. The correct interpretation should be clear from the context. Furthermore, for the sake of simplicity, we restrict to the case that the discretization respects the regions of interest: for every region $q \in \widetilde{Q}$, all $x \in q$ must share the same observations. However, IMDP abstractions in which the regions of interest are not respected are also possible [4]. Additionally, we denote by q_u the set $\mathbb{R}^n \setminus X$, which corresponds to the rest of the state space of system (5-1). Taking this extra state into account, we define the state space of $\widehat{\mathcal{I}}$ as $Q := \widetilde{Q} \cup \{q_u\}$. In this way we have discretized the entirety of the continuous state space $\mathbb{R}^n$. Next, when it comes to the observations of $\widehat{\mathcal{I}}$, we let this observation set be the same as that of the original system: $O := O_\mathcal{C}$.

Regarding the set of actions A of $\widehat{\mathcal{I}}$, we define it as a finite subset of $U_\mathcal{C}$. Furthermore we make all actions available at each state q, i.e., $A(q) = A$ for all $q \in Q$. This is not necessarily the only way to define the set of actions of the nominal IMDP, and state-dependent action sets can exist. However, to simplify the problem, in this document we only consider this case. Next, we define the observation map of $\widehat{\mathcal{I}}$ as $L(q) := L_\mathcal{C}(x)$ for any $x \in q$ and for all $q \in Q$. This means that the L assigns to each cell the same observation that was assigned to that region in the continuous system model. In this way, we leverage the fact that the discretization respects the regions of interest.

Now we define the transition probability bounds $\underline{P}, \overline{P}$ of $\widehat{\mathcal{I}}$. Consider states $q, q' \in Q \setminus \{q_u\}$ and action $a \in A$. Furthermore, consider the transition kernel $\mathcal{T}_\mathcal{C}$ in (4-2), for fixed probability

$\widehat{P}_\xi$. Then we define $\underline{P}, \overline{P}$ as the bounds in the probability of system (5-1) transitioning from region q to region q' under action a:

$$\begin{aligned} \underline{P}(q,a,q') &\leq \min_{x\in q} \widehat{P}_\xi(x_{t+1} \in q' | x_t = x, a_t = a) = \min_{x\in q} \mathcal{T}_\mathcal{C}(q'|x,a;\widehat{P}_\xi), \\ \overline{P}(q,a,q') &\geq \max_{x\in q} \widehat{P}_\xi(x_{t+1} \in q' | x_t = x, a_t = a) = \max_{x\in q} \mathcal{T}_\mathcal{C}(q'|x,a;\widehat{P}_\xi). \end{aligned} \tag{5-3}$$

Additionally, since $X = \mathbb{R}^n \setminus q_u$, we define the bounds in probability of transitioning to q_u from any $q \in Q \setminus \{q_u\}$ and for any $a \in A$ as:

$$\begin{aligned} \underline{P}(q,a,q_u) &\leq 1 - \max_{x\in q} \widehat{P}_\xi(x_{t+1} \in X | x_t = x, a_t = a) = 1 - \max_{x\in q} \mathcal{T}_\mathcal{C}(X|x,a;\widehat{P}_\xi), \\ \overline{P}(q,a,q_u) &\leq 1 - \min_{x\in q} \widehat{P}_\xi(x_{t+1} \in X | x_t = x, a_t = a) = 1 - \min_{x\in q} \mathcal{T}_\mathcal{C}(X|x,a;\widehat{P}_\xi). \end{aligned} \tag{5-4}$$

Furthermore, according to Problem 5-1.1, we are only interested in the paths of $\mathcal{C}$ that never exit X. To account for that extra specification, we make the state q_u of $\widehat{\mathcal{I}}$ absorbing by defining $\overline{P}(q_u,a,q_u) = \underline{P}(q_u,a,q_u) = 1$ for all $a \in A$ and $\overline{P}(q_u,a,q) = \underline{P}(q_u,a,q) = 0$ for all $q \in Q \setminus \{q_u\}$.

Now, let us remember from the theory of IMDPs in Section 2-4 the concept of feasible transition probabilities of an IMDP. Consider the set of probability distributions over the state space Q of $\widehat{\mathcal{I}}$, $\mathcal{D}(Q)$. We denote by $\widehat{\Gamma}_{q,a}$ the set of transition probabilities of $\widehat{\mathcal{I}}$ from $q \in Q$ by $a \in A$ that fulfill:

$$\widehat{\Gamma}_{q,a} = \{\widehat{\gamma}_{q,a} \in \mathcal{D}(Q) : \underline{P}(q,a,q') \leq \widehat{\gamma}_{q,a}(q') \leq \overline{P}(q,a,q'), \text{ for all } q' \in Q\}, \tag{5-5}$$

for all $q \in Q$ by $a \in A$. The nomenclature of "feasible" transition probabilities and set comes from the fact that set $\widehat{\Gamma}_{q,a}$ will appear as the feasible set of the optimization problems that we need to solve in this chapter to synthesize a strategy.

We must highlight that in this chapter we only give the formal definition of the transition probability bounds of $\widehat{\mathcal{I}}$, but we do not describe the procedure we follow to obtain them. This is because obtaining an IMDP abstraction for an arbitrary nominal distribution is not the goal of this thesis. Therefore, in this chapter we assume that the nominal distribution is such that it allows us to easily compute a nominal IMDP abstraction using existing approaches. For example, efficient algorithms to obtain IMDP abstractions have already been proposed when the disturbance is Gaussian, for systems with switched linear dynamics [4] and for systems modelled as Neural Network Dynamic Models [5]. Furthermore, when the distribution of the nominal disturbance is finitely-supported, which is the case in the data-driven scenario, we could obtain an IMDP abstraction as follows: we could follow the approach described in Chapter 4, being the nominal distribution the empirical one $\widehat{P}^N_\xi$, but setting the transport budget ε to zero. Therefore, our approach is also general with respect to the way we obtain the IMDP abstraction.

5-2-2 Robust MDP Abstraction

In this section we formally define our desired robust MDP abstractions, which account for both, the state discretization and the distributional ambiguity. We do that by expanding the set $\widehat{\Gamma}_{q,a}$ of feasible transition probabilities of the nominal IMDP $\widehat{\mathcal{I}}$, to also account for distributional uncertainty. We start by defining a Wasserstein distance between probabilities supported on the discrete state space Q of the $\widehat{\mathcal{I}}$. After that, we use that distance to define the set of feasible distributions of $\mathcal{M}^R$. Finally, we discuss the advantages and disadvantages of robust MDP abstractions, when compared to DR-IMDP ones, in the data-driven setting.

Let us begin by defining a p-Wasserstein distance between distributions over Q. This p-Wasserstein distance is based on distance (see Definition 3-1.3) d_{ij}, which we define as the minimum distance between states $q_i, q_j \in Q$:

$$d_{ij} := \min\{\|x - y\|_p : x \in q_i, y \in q_j\}, \tag{5-6}$$

for $p \geq 1$. To simplify the notation we denote it by $\mathcal{W}_p$, which is the same notation we use for the p-Wasserstein distance between distributions over $\mathbb{R}^n$. Whether we refer to one or the other is clear from the context, since one takes as arguments distributions over $\mathbb{R}^n$ whereas the other is defined for distributions over Q. Furthermore, notice that when distributions P and P' in Definition 3-1.3 are finitely supported, $\mathcal{W}_p(P, P')$ is defined as the solution of a finite linear program [31]. This is because the transport of mass between a finite number of points in Q is carried out through finitely supported couplings defined over $Q \times Q$. This result is useful to define the algorithms that allow us to synthesize strategies for robust MDPs.

Now, we define the feasible set of probability distributions of our robust MDP abstraction $\mathcal{M}^R$. We begin by considering an IMDP abstraction $\widehat{\mathcal{I}}$ of the original system under the nominal probability $\widehat{P}_\xi$. Once we have obtained such abstraction, consider its set $\widehat{\Gamma}_q^a \subset \mathcal{D}(Q)$ of feasible transition probabilities for every $q \in Q$ by $a \in A$ as defined in (5-5). We recall that the nomenclature of “feasible" is related to the fact that such set appears as the feasible set of the strategy synthesis algorithms for IMDPs. Then, to include the distributional ambiguity into the abstraction we expand set $\widehat{\Gamma}_q^a$. We denote the set of feasible transition probabilities of $\mathcal{M}^R$ for fixed $q \in Q$, $a \in A$ by $\widehat{\Gamma}_q^a \oplus \varepsilon$. Furthermore, we define this set as the subset of $\mathcal{D}(Q)$ that contains all distributions which lay at a distance of at most ε from set $\widehat{\Gamma}_q^a$:

$$\widehat{\Gamma}_q^a \oplus \varepsilon = \bigcup_{\widehat{\gamma}_{q,a} \in \widehat{\Gamma}_q^a} \mathbb{B}_\varepsilon(\widehat{\gamma}_{q,a}). \tag{5-7}$$

The feasible set $\widehat{\Gamma}_q^a \oplus \varepsilon$, together with the sets Q, A and O and observation function L of IMDP $\widehat{\mathcal{I}}$ uniquely define the robust MDP $\mathcal{M}^R$. Notice that $\mathcal{M}^R$ only differs from $\widehat{\mathcal{I}}$ in its set of feasible transition probabilities. The concepts of paths, strategy and adversary of a robust MDP are identical to the case of MDPs (see Section 2-3). However, notice that in a robust MDP the adversary chooses probabilities from set $\widehat{\Gamma}_q^a \oplus \varepsilon$, which is defined by more constraints than just intervals.

To finish this section, we discuss the usefulness of robust MDP abstractions in the data-driven setting, despite the setting that we consider in this chapter being general. In this setting, a finite number of samples $\{\widehat{\xi}^i\}_{i=1}^N$ from the disturbance are available. We leverage this information and build a robust MDP abstraction in the following way: we begin by constructing an empirical distribution from the samples, which we take as the nominal one. After that, we obtain the nominal IMDP abstraction using this empirical distribution and following the approach described in Chapter 4, but setting the radius of the ambiguity set ε to zero. In this way the nominal IMDP only accounts for the state discretization. Then, we build a robust MDP abstraction as we described in this section.

The advantage of using, in a data-driven setting, robust MDPs instead of the DR-IMDPs discussed in Chapter 4 is that the set of feasible transition probabilities of the former is tighter than the one of the latter. In the case of DR-IMDPs, this set can be excessively big for some practical values of the radius ε of the ambiguity set. This is translated into a big range of probabilities of satisfying the specification, which does not provide useful information. In other words, by making use of the DR-IMDP abstractions described in Chapter 4, we are providing the adversary with too much freedom. The cause of this undesired behaviour is the way in which we have defined the transition probability bounds of the DR-IMDP in expressions (4-4). As an example, consider the upper bound $\overline{P}(q,a,\cdot)$ for fixed $q \in Q$ and $a \in A$: we compute it allowing to transport as much mass as possible, with budget ε from the samples of x_{t+1} to region q', and we do this for every $q' \in Q$. However, this is counter-intuitive since, for example, if transporting as much probability mass to a state $q' \in Q$ already consumes all the budget, it should not be possible to further transport mass to other states. Nevertheless, there is no constraint that enforces this behaviour in the approach we described in Chapter 4. Opposite to DR-IMDP abstractions, robust MDPs avoid this issue, since the set of feasible transition probabilities is defined by more constraints. These constraints couple the mass transported across the state space of the abstraction, and enforce the budget limitation. On the other hand and, as a trade-off to this reduction in conservatism, synthesizing a strategy for a robust MDP abstraction is way less efficient than for the case of DR-IMDPs, as we show in Section 5-3.

5-3 Strategy Synthesis for Robust MDP Abstractions

In Section 5-2 we described how to obtain a robust MDP abstraction of the system in (5-1) that accounts for both the state discretization and the distributional uncertainty. In this section we focus on the problem of synthesizing a strategy for this abstraction that enforces the satisfaction of an scLTL formula ϕ. Remember that, in Section 2-4-1, we explained how synthesizing such a strategy for an IMDP boils down to to solving a maximal reachability probability problem for a more complex IMDP. For this reason, in this section we just describe how to synthesize a strategy for robust MDPs in the setting of reachability. The extension to the setting of enforcing more complex specifications given as scLTL formulas follows the same reasoning described in Section 2-4-1.

Maximal Reachability Probability Problem for Robust MDPs

Consider the robust MDP $\mathcal{M}^R$ and target set $Q_{\texttt{tgt}} \subset Q$. We start by considering the worst and best-case probabilities of the paths of $\mathcal{M}^R$, under strategy $\sigma \in \Sigma$, reaching $Q_{\texttt{tgt}}$ within k steps by starting on $q \in Q$:

$$\begin{aligned} &\min_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma, \pi) \\ &\max_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma, \pi) \end{aligned} \tag{5-8}$$

Then, we define the optimal strategy $\sigma^* \in \Sigma$ the one that maximizes the first expression in (5-8), this is, the strategy that, maximizes the lower bound in the probability of reachability:

$$\sigma^* := \arg\max_{\sigma\in\Sigma} \min_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\texttt{tgt}} | w_Q^k(0) = q, \sigma, \pi), \tag{5-9}$$

for all $q \in Q$, $k \in \mathbb{N} \cup \{\infty\}$. This is a pessimistic way of defining the optimal strategy, which corresponds to the two-player game we described in Section 2-4-1 for IMDPs. However, in our setting of formal strategy synthesis we are interested in obtaining the strategy that maximizes the worst-case probability that the abstraction satisfies the specification. This is because this lower bound is the performance guarantee that we are looking for [4] and, therefore we want it to be as high as possible. Let us denote by $\underline{p}^k$ and $\overline{p}^k$, respectively, the worst and best-case probabilities of the paths of $\mathcal{M}^R$ reaching $Q_{\mathtt{tgt}}$ within k steps under strategy $\sigma^* \in \Sigma$:

$$\begin{aligned}\underline{p}^k(q) &:= \max_{\sigma\in\Sigma}\min_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\mathtt{tgt}} | w_Q^k(0) = q, \sigma, \pi)\\ \overline{p}^k(q) &:= \max_{\pi\in\Pi} P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\ w_Q^k(t) \in Q_{\mathtt{tgt}} | w_Q^k(0) = q, \sigma^*, \pi)\end{aligned} \tag{5-10}$$

for all $q \in Q$, $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$.

Now we state Theorem 5-3.1, which allows us to obtain the probabilities in (5-10) and the strategy in (5-9).

Theorem 5-3.1. ***(Robust Value Iteration)*** *Consider the robust MDP $\mathcal{M}^R$, whose set of feasible transition probabilities is $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ for each $q \in Q$, $a \in A$. Then, probabilities $\underline{p}^k$ in* (5-10) *are obtained for all $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ recursively, starting from $\underline{p}^0(q) = 1$ for all $q \in Q_{\mathtt{tgt}}$ and $\underline{p}^0(q) = 0$ otherwise:*

$$\underline{p}^{k+1}(q) = \begin{cases} 1 & \mathit{if}\, q \in Q_{\mathtt{tgt}} \\ \max_{a\in A(q)} \min_{\gamma_{q,a}\in\widehat{\Gamma}_{q,a}\oplus\varepsilon} \sum_{q'\in Q} \gamma_{q,a}(q')\underline{p}^k(q') & \mathit{otherwise}, \end{cases} \tag{5-11}$$

Furthermore, with a small abuse of notation, strategy σ^* in (5-9) *is the Markovian, but time-dependent strategy $\sigma^* = \{\sigma^*(\cdot; k)\}_{k=1}^{\infty}$ that fulfills:*

$$\sigma^*(q; k+1) = \arg\max_{a\in A(q)} \min_{\gamma_{q,a}\in\widehat{\Gamma}_{q,a}\oplus\varepsilon} \sum_{q'\in Q} \gamma_{q,a}(q')\underline{p}^k(q'), \tag{5-12}$$

for all $q \in Q$, $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$. Using σ^*, probabilities $\overline{p}^k$ in (5-10) *can be obtained for all $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$ recursively, starting from $\overline{p}^0(q) = 1$ for all $q \in Q_{\mathtt{tgt}}$ and $\overline{p}^0(q) = 0$ otherwise:*

$$\overline{p}^{k+1}(q) = \begin{cases} 1 & \mathit{if}\, q \in Q_{\mathtt{tgt}} \\ \max_{\gamma_{q,\sigma^*(q;k)}\in\widehat{\Gamma}_{q,\sigma^*(q;k)}\oplus\varepsilon} \sum_{q'\in Q} \gamma_{q,\sigma^*(q;k)}(q')\overline{p}^k(q') & \mathit{otherwise} \end{cases}. \tag{5-13}$$

We denote the previous iterative process by robust value iteration.

Proof. The proof follows a similar reasoning that the one presented in [30] for IMDPs under stationary strategies. Notice that, despite the fact that here the set of feasible transition probabilities is $\widehat{\Gamma}_{q,a} \oplus \varepsilon$, this fact does not affect the proof. It is easy to prove that the probabilities $\underline{p}^k(q)$ and $\overline{p}^k(q)$ correspond to the probabilities defined in (5-10), for all $q \in Q$, $k \in \mathbb{N}_{\geq 0}$. Furthermore, since they are monotonically increasing and upper bounded by 1 for all $q \in Q$, sequences $\underline{p}^k(q)$ and $\overline{p}^k(q)$ converge to the following fixed points of recursions (5-11) and (5-13), respectively:

$$\begin{aligned}\underline{p}(q) &= \lim_{k\to\infty} \underline{p}^k(q) = \max_{\sigma\in\Sigma}\min_{\pi\in\Pi} P(\exists k \in \mathbb{N}_{\geq 0} \text{ s.t. } w_Q(k) \in Q_{\mathtt{tgt}} | w_Q(0) = q, \sigma, \pi)\\ \overline{p}(q) &= \lim_{k\to\infty} \overline{p}^k(q) = \max_{\pi\in\Pi} P(\exists k \in \mathbb{N}_{\geq 0} \text{ s.t. } w_Q(k) \in Q_{\mathtt{tgt}} | w_Q^k(0) = q, \sigma^*, \pi)\end{aligned} \tag{5-14}$$

for all $q \in Q$. The previous expressions are the probabilities of the paths of infinite length of $\mathcal{M}^R$ ever reaching $Q_{\mathtt{tgt}}$ under σ^*. Furthermore

$$\lim_{k \to \infty} \sigma^*(q; k) = \sigma^*_{\mathtt{stat}}(q),$$

for all $q \in Q$. This is because strategy σ^* only depends on $\underline{p}^k$, as shown in (5-12), and the latter converges for $k \to \infty$. □

Notice that, due to Bellman's optimality principle [19], strategy σ^* is memoryless, since it does not depend on the whole path w_Q^k of $\mathcal{M}^R$, but only on $last(w_Q^k)$. Furthermore, it is the same strategy for all initial conditions $q \in Q$. However, this strategy is also time-dependent, since it depends on current value function $\underline{p}^k$, which changes over time.

Results of Strategy Synthesis for robust MDPs Under scLTL Specifications

As a starting point, consider the FSA $\mathcal{A}$ that captures the language of ϕ, and the product robust MDP $\mathcal{M}^R_\phi := \mathcal{M}^R \times \mathcal{A}$. The latter is defined in a similar way as the product IMDP was defined in Definition 2-4.3. The only difference is that now, we need to define the set of feasible transition probabilities of $\mathcal{M}^R_\phi$, which we denote by $\widehat{\Gamma}^\phi_{(q,z),a} \oplus \varepsilon$. This set corresponds to set $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ whenever a transition in $\mathcal{M}^R$ generates a transition in $\mathcal{A}$[1]. Consider also target set $Q_{\phi,ac}$. This set is analogous to the one defined in Definition 2-4.3 for the product IMDP. Now, assume that we have solved the maximal reachability problem described in Section 5-3 for $\mathcal{M}^R_\phi$. As a result, we have obtained the upper and lower bounds in probability $\underline{p}^k_\phi$ and $\overline{p}^k_\phi$ of the paths of $\mathcal{M}^R_\phi$ reaching $Q_{\phi,ac}$. Furthermore, we have obtained the strategy σ^*_ϕ of $\mathcal{M}^R_\phi$ that maximizes the previous lower bound at each time step. In this section we describe how we can translate those results to the strategy and guarantees of the IMDP $\mathcal{I}$ that we seek. In fact, the results we describe hold for any Markovian strategy σ_ϕ of $\mathcal{I}_\phi$, and when $\underline{p}^k_\phi$ and $\overline{p}^k_\phi$ have been obtained by solving (5-11) and (5-13) for fixed strategy σ_ϕ.

First, we prove that any Markovian, and possibly time-dependent strategy σ_ϕ in $\mathcal{M}^R_\phi$ maps to a history dependent strategy $\sigma_{\mathcal{M}^R}$ in the initial robust MDP abstraction $\mathcal{M}^R$.

Lemma 5-3.1. ***(Translating a Markovian Strategy of the Product Robust MDP to the Robust MDP)*** *Consider the product robust MDP $\mathcal{M}^R_\phi$ obtained by taking the product between robust MDP $\mathcal{M}^R$ and FSA $\mathcal{A}$, where the latter captures the language of the scLTL formula ϕ. Consider also the Markovian strategy σ_ϕ of $\mathcal{M}^R_\phi$. Then, strategy σ_ϕ can be translated to a memory-dependent strategy $\sigma_{\mathcal{M}^R}$ of $\mathcal{M}^R$.*

Proof. The proof is identical to the proof of Lemma 2-4.1 for IMDPs. □

Secondly, we relate the bounds in the probability of $\mathcal{M}^R_\phi$ reaching $Q_{\phi,ac}$ to those in the probability of $\mathcal{M}^R$ satisfying the specification ϕ.

[1] This is analogous to how we define the transition probabilities of a product IMDP in (2-3). However, for simplicity, we do not state the definition of $\widehat{\Gamma}^\phi_{(q,z),a} \oplus \varepsilon$ here.

Lemma 5-3.2. ***(Guarantees Of the robust MDP Abstraction)*** *Consider the strategy $\sigma^*_{\mathcal{M}^R}$ of $\mathcal{M}^R$ obtained from σ^*_ϕ as described in Lemma 5-3.1. Then it holds that the bounds in the probability of the paths of $\mathcal{M}^R$ satisfying ϕ within k steps while never exiting X by following strategy $\sigma^*_{\mathcal{M}^R}$ and starting from state $q \in Q$ are:*

$$\underline{p}^k_{\mathcal{M}^R}(q) := \inf_{\pi \in \Pi} P(w^k_Q \models \phi | w^k_Q(0) = q, X, \sigma^*_{\mathcal{M}^R}, \pi) = \underline{p}^k_\phi(q, z_0) \tag{5-15}$$

$$\overline{p}^k_{\mathcal{M}^R}(q) := \sup_{\pi \in \Pi} P(w^k_Q \models \phi | w^k_Q(0) = q, X, \sigma^*_{\mathcal{M}^R}, \pi) = \overline{p}^k_\phi(q, z_0) \tag{5-16}$$

for all $q \in Q$.

Proof. The proof follows the same reasoning as in the case of IMDPs: a path of $\mathcal{M}^R$ satisfies ϕ if and only if it generates an accepting run in $\mathcal{A}$. Since this is the same as the corresponding path of $\mathcal{M}^R_\phi$ reaching $Q_{\phi,ac}$, Lemma 5-3.2 follows. □

Notice that, in practice, we are only interested in the probabilities obtained with initial state z_0 of $\mathcal{A}$, since the runs of the automaton always start at its initial state z_0. The probability bounds in (5-15) are the guarantees of the robust MDP abstraction satisfying the specification we were looking for. Furthermore, we have synthesized the strategy $\sigma^*_{\mathcal{M}^R}$ that maximizes the lower bound. The complexity of the interval value iteration algorithm used on $\mathcal{M}^R_\phi$ is polynomial in the number of states of $\mathcal{M}^R_\phi$ and exponential in the size of ϕ in the worst case. Additionally, for infinite horizon $k \to \infty$, although $\sigma^*_\phi(\cdot; k)$ becomes a stationary strategy for $\mathcal{M}^R_\phi$, $\sigma^*_{\mathcal{M}^R}(\cdot; k)$ is still a memory-dependent strategy for $\mathcal{M}^R$, as a result of Lemma 5-3.1.

Complexity of Robust Value Iteration

Now, let us discuss the computational features of the value iteration algorithm introduced in Theorem 5-3.1 to solve a maximal reachability probability problem. First of all, we highlight that, since iteration (5-11) is independent of the lower bound in probability, we can compute this recursion and obtain the strategy in (5-12) first. Once we have achieved convergence we can compute the second recursion (5-13) and obtain the upper bound in probability, which is useful to analyze the error of the solution [4]. Secondly, formulate the inner problem present in recursion (5-11) as a linear program. The same theorem applies to (5-13) by just changing the min operator to a max.

Theorem 5-3.2. ***(Robust Value Iteration as a Linear Program)*** *Consider the robust value iteration recursion* (5-11) *for the robust MDP $\mathcal{M}^R$. Consider also, for fixed state $q \in Q$, action $a \in A$, and probability $\underline{p}^k$, the inner optimization problem over set $\widehat{\Gamma}_{q,a} \oplus \varepsilon$. Then the folowing holds:*

$$\min_{\gamma_{q,a} \in \widehat{\Gamma}_{q,a} \oplus \varepsilon} \sum_{q_i \in Q} \gamma_{q,a}(q_i) \underline{p}^k(q_i) = \min_{\gamma_{q,a} \in \mathbb{R}^{|Q|}, \widehat{\gamma}_{q,a} \in \mathbb{R}^{|Q|}, \pi \in \mathbb{R}^{|Q|^2}} \sum_{q_i \in Q} \gamma_{q,a}(q_i) \underline{p}^k(q_i), \tag{5-17}$$

where $\gamma_{q,a}, \widehat{\gamma}_{q,a}$ and π must satisfy the following constraints:

$$\underline{P}(q,a,q_j) \leq \widehat{\gamma}_{q,a}(q_j) \leq \overline{P}(q,a,q_j) \qquad q_j \in Q \tag{5-18a}$$

$$\sum_{q_j \in Q} \widehat{\gamma}_{q,a}(q_j) = 1 \tag{5-18b}$$

$$\pi_{ij} \geq 0, \qquad q_i \in Q,\ q_j \in Q \tag{5-18c}$$

$$\sum_{q_i \in Q} \pi_{ij} = \widehat{\gamma}_{q,a}(q_j), \qquad q_j \in Q \tag{5-18d}$$

$$\sum_{q_j \in Q} \pi_{ij} = \gamma_{q,a}(q_i), \qquad q_i \in Q \tag{5-18e}$$

$$\sum_{q_i \in Q} \sum_{q_j \in Q} \pi_{ij} d_{ij}^p \leq \varepsilon^p. \tag{5-18f}$$

Proof. First, we prove that, for fixed state $q \in Q$ and action $a \in A$, the set of feasible transition probabilities $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ defined in (5-7) is a polytope. We do this by defining this set in the following alternative way: $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ is defined as the set of all transition probabilities $\gamma_{q,a}$ such that there exists a $\widehat{\gamma}_{q,a} \in \widehat{\Gamma}_{q,a}$ and a coupling π that transports mass between $\widehat{\gamma}_{q,a}$ and $\gamma_{q,a}$ with a cost smaller than ε. Consider now set $\widehat{\Gamma}_{q,a}$ in (5-5) and the p-Wasserstein distance in Definition 3-1.3. Set $\widehat{\Gamma}_{q,a}$ is already defined by linear constraints. Furthermore, $\mathcal{W}_p(\gamma_{q,a}, \widehat{\gamma}_{q,a})$ is defined as the minimum cost of an LP, which for the case of finitely supported distributions is finitely dimensional [31]. This means that $\gamma_{q,a}$ satisfying constraints (5-18) for some $\widehat{\gamma}_{q,a}$, π is equivalent to $\gamma_{q,a} \in \widehat{\Gamma}_{q,a} \oplus \varepsilon$. Secondly, consider, for fixed state $q \in Q$, action $a \in A$ and probability $\underline{p}^k$, the inner minimization over the set $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ in (5-11). Since the transition probabilities $\gamma_{q,a}$ appear linearly in the cost of this expression, this completes the proof. □

Now we describe the constraints in (5-18). The first and second constraints are those that come from the feasible set of distributions of the IMDP $\widehat{\mathcal{I}}$. These constraints are the same as picking a feasible distribution from the IMDP $\widehat{\mathcal{I}}$, this is, $\widehat{\gamma}_{q,a} \in \widehat{\Gamma}_{q,a}$. The remaining constraints are those that come from Definition 3-1.3 of the p-Wasserstein distance $\mathcal{W}_p$, between probabilities over Q: the third constraint imposes non-negativity of the coupling π. The fourth and fifth constraints impose that the mass transported to every state $q_i \in Q$ must come from the mass of $\widehat{\gamma}_{q,a}$. This transport is carried out through the component π_{ij} of the coupling, which transports mass from $\widehat{\gamma}_{q,a}(q_j)$ to $\gamma_{q,a}(q_i)$, for all $q_i, q_j \in Q$. Finally, the last constraint imposes that the cost of transporting mass from $\widehat{\gamma}_{q,a}$ to $\gamma_{q,a}$ should be less than the budget ε.

Despite the problem (5-17) - (5-18) being an LP, we cannot solve it by using the efficient algorithm that we described in Section 2-4-1 for IMDPs. Instead, standard LP solvers like the interior point algorithm must be employed. In practice, this means that synthesizing a strategy for robust MDPs is computationally heavier than doing so for IMDPs. Furthermore, the number of variables and constraints in Problem (5-17) - (5-18) is way higher, as we show in Table 5-1:

As we can observe in Table 5-1, the complexity of LP (5-17) - (5-18) grows with the number of states $|Q|$ of the abstraction: both the number of variables and the number of inequality constraints are quadratic in $|Q|$. Moreover, most of the computational complexity comes from

	IVI	Robust VI	Relaxed Robust VI
# Variables	$\lvert Q\rvert$	$\lvert Q\rvert\cdot(\lvert Q\rvert+2)$	$\lvert Q\rvert+(\lvert Q_c^{q,a}\rvert+2)\cdot\lvert Q_s^{q,a}\rvert$
# Inequality constraints	$2\lvert Q\rvert$	$\lvert Q\rvert\cdot(\lvert Q\rvert+2)+1$	$\lvert Q_s^{q,a}\rvert\cdot(\lvert Q_c^{q,a}\rvert+3)+1$
# Equality constraints	1	$2\cdot\lvert Q\rvert+1$	$\lvert Q_c^{q,a}\rvert+\lvert Q_s^{q,a}\rvert+2$

Table 5-1: Complexity of the inner problems present in different value iteration recursions. The comparison is made when no redundant constraints have been eliminated in (5-18) and (5-20).

the fact that the coupling is defined over the product space $Q \times Q$, which is a space of high cardinality. In the following we will describe how we can reduce said complexity. First, we state the following definition:

Definition 5-3.1. ***(Support of the Nominal Transition Probabilities)*** *We define the support $Q_s^{q,a}$ of the nominal transition probabilities $\widehat{\gamma}_{q,a} \in \widehat{\Gamma}_q^a$, for fixed current state $q \in Q$ and action $a \in A$, as*

$$Q_s^{q,a} = \bigcup_{\overline{P}(q,a,q')>0} q'.$$

intuitively, this is the set that contains all states to which the IMDP abstraction $\widehat{\mathcal{I}}$ is able to transition in one step. Notice that $Q_s^{q,a} \subseteq Q$ holds always, and $Q_s^{q,a} \subset Q$ only if the support of the disturbance ξ is bounded. This is the case, for example, in the data-driven setting, in which the nominal distribution is empirical, this is, built from a finite number of samples of ξ. If $Q_s^{q,a} \subset Q$, then we are able to reduce the complexity of LP (5-17) - (5-18) by considering couplings over the space $Q_s^{q,a} \times Q$, whose cardinality is smaller than $|Q \times Q|$. To further reduce the complexity of LP (5-17) - (5-18), we consider couplings defined over a simplification of state space Q. First, consider a smaller, state and action-dependent, subset $Q_c^{q,a}$, with cardinality $|Q_c^{q,a}|$ of the state space such that $Q_s^{q,a} \subset Q_c^{q,a} \subset Q$. The set $Q_c^{q,a}$ should contain states closer to those in $Q_s^{q,a}$ in the sense of distance d_{ij}, up to a threshold distance. This makes sense because we expect mass to only be transported a few cells away from $Q_s^{q,a}$, since the cost of transporting mass is higher the longer the transport distance. In addition to set $Q_c^{q,a}$, from the point of view of mass transport, we consider all remaining states $q_i \in Q \setminus Q_c^{q,a}$ as a single state q_r. Taking into account sets $Q_s^{q,a}$, $Q_c^{q,a}$ and state q_r, we consider couplings π defined over $(Q_c^{q,a} \cup q_r) \times Q_s^{q,a}$, which are lower-dimensional than the ones used in (5-18). This means that we consider mass transport from set $Q_s^{q,a}$ to set $Q_c^{q,a}$, and also to states $Q \setminus Q_c^{q,a}$ as if it was a single state q_r. The mass transported to q_r then will be shared between states in $Q \setminus Q_c^{q,a}$, while fulfilling the budget constraint. Let us also define the following distance:

$$d_{is} := \min\{d_{ij} : q_j \in Q_s^{q,a}\}, \tag{5-19}$$

for all $q_i \in Q \setminus Q_c^{q,a}$. Notice that d_{is} is an under estimator of d_{ij} for all $q_i \in Q \setminus Q_c^{q,a}$, $q_j \in Q_s^{q,a}$. Taking into account the previous elements we can define the following set:

Definition 5-3.2. *Consider, for fixed state $q \in Q$ and action $a \in A$, the support of the nominal transition probabilities $Q_s^{q,a}$ as defined in Definition 5-3.1. Furthermore, consider a set of states $Q_c^{q,a}$ such that $Q_s^{q,a} \subset Q_c^{q,a} \subset Q$. Consider also the additional state q_r and a discrete coupling π over the product space $(Q_c^{q,a} \cup q_r) \times Q_s^{q,a}$. Additionally, consider the* *distance d_{is} as defined in* (5-19). *We define set Γ_q^a as the set of all $\gamma_{q,a} \in \mathbb{R}^n$ for which there exist $\widehat{\gamma}_{q,a} \in \mathbb{R}^{|Q_s^{q,a}|}$ and $\pi \in \mathbb{R}^{(|Q_c^{q,a}|+1)\times|Q_s^{q,a}|}$ that satisfy the following constraints:*

$$\underline{P}(q,a,q_j) \le \widehat{\gamma}_{q,a}(q_j) \le \overline{P}(q,a,q_j) \qquad q_j \in Q_s^{q,a} \tag{5-20a}$$

$$\sum_{q_j \in Q_s^{q,a}} \widehat{\gamma}_{q,a}(q_j) = 1 \tag{5-20b}$$

$$\pi_{ij} \ge 0, \qquad q_i \in Q_c^{q,a} \cup q_r^{q,a},\ q_j \in Q_s^{q,a} \tag{5-20c}$$

$$\sum_{q_i \in Q_c^{q,a}} \pi_{ij} + \pi_{rj} = \widehat{\gamma}_{q,a}(q_j), \qquad q_j \in Q_s^{q,a} \tag{5-20d}$$

$$\sum_{q_j \in Q_s^{q,a}} \pi_{ij} = \gamma_{q,a}(q_i), \qquad q_i \in Q_c^{q,a} \tag{5-20e}$$

$$\sum_{q_j \in Q_s^{q,a}} \pi_{r,j} = \sum_{q_i \in Q \setminus Q_c^{q,a}} \gamma_{q,a}(q_i) \tag{5-20f}$$

$$\sum_{q_i \in Q_c^{q,a}} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij}^p + \sum_{q_i \in Q \setminus Q_c^{q,a}} \gamma_{q,a}(q_i) d_{is}^p \le \varepsilon^p. \tag{5-20g}$$

Now we describe the meaning of the constraints that describe $\Gamma_{q,a}$ in Definition 5-3.2. The first and second constraints are the same as those in (5-18). However, now the first one is defined over a reduced space $Q_s^{q,a}$, and the sum in the second one is performed over this same set. The third constraint imposes non-negativity of the coupling, now defined over the reduced space $(Q_c^{q,a} \cup q_r) \times Q_s^{q,a}$. The fourth constraint imposes that the amount of mass transported by the coupling to all the state space comes from $\widehat{\gamma}_{q,a}(q_j)$. In the same way, the fifth constraint imposes that the amount of mass that state $q_i \in Q_c^{q,a}$ receives comes from the coupling. Additionally, the sixth constraint imposes that the amount of mass that is transported to the rest of the state space $Q \setminus Q_c^{q,a}$, also comes from the coupling. Finally, the seventh constraint imposes that the cost of transporting mass must be less than the budget ε^p. Notice that the mass that is transported to state $q_i \in Q \setminus Q_c^{q,a}$ is weighed by distance d_{is} defined in (5-19). In this way, we are not considering the individual cost of transporting mass from $q_j \in Q_s^{q,a}$ to $q_i \in Q \setminus Q_c^{q,a}$, but a reduced cost, based on d_{is}. We do this because we are not considering individual couplings between such states. Notice that, from the structure of the fourth and fifth constraints in (5-20), we could eliminate the variables $\widehat{\gamma}_{q,a}$ and $\gamma_{q,a}(q_i)$, $q_i \in Q_c^{q,a}$ to further reduce the complexity of the problem. However, for clarity we leave the set of constraints (5-20) as it is. The different sets considered in Definition 5-3.2 are illustrated in Figure 5-1.

Using set $\Gamma_{q,a}$, we define an alternative iterative algorithm to the robust value iteration one introduced in Theorem 5-3.1 to solve reachability problems:

Definition 5-3.3. ***(Relaxed Robust Value Iteration)*** *Consider robust MDP $\mathcal{M}^R$ and the value iteration algorithm introduced in Theorem 5-3.1. We denote by relaxed robust value iteration the algorithm we obtain by using set $\Gamma_{q,a}$ instead of $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ in expressions* (5-11),

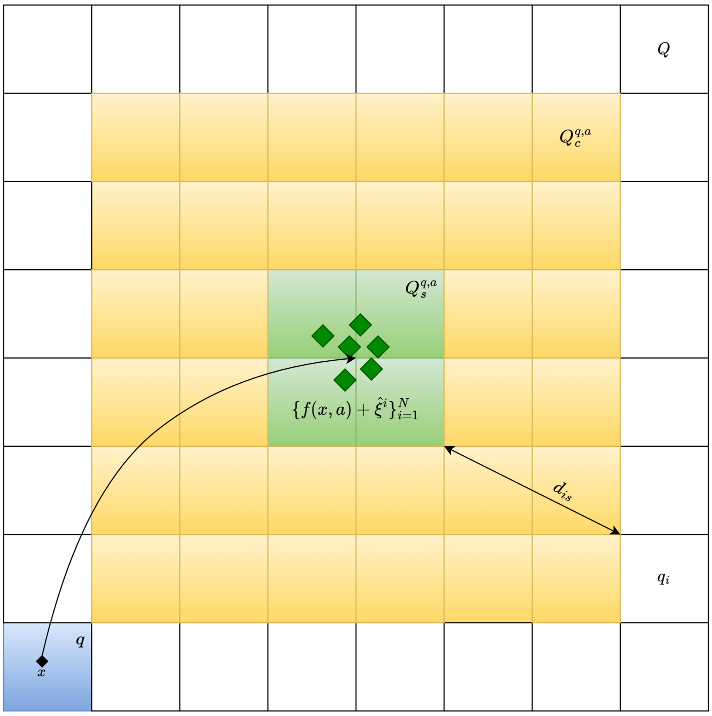


Figure 5-1: Graphical representation of the regions $Q_s^{q,a}$ and $Q_c^{q,a}$ for fixed state q and action a. In this illustration, the nominal IMDP $\widehat{\mathcal{I}}$ has been constructed considering an empirical distribution of the disturbance as the nominal one. All distributions in the nominal set of feasible distributions $\widehat{\Gamma}_q^a$ are supported on the four states that compose set $Q_s^{q,a}$, since those are the states that contain the green samples. Furthermore, the cardinality of set $Q_c^{q,a}$ is $|Q_c^{q,a}| = 36$. Additionally, distance d_{is} is depicted in the figure for an arbitrary state $q_i \in Q \setminus Q_c^{q,a}$.

(5-13) *and* (5-12) *for every* $q \in Q$ *and* $a \in A$. *This algorithm yields the sequences* $\{\underline{p}_{\texttt{rel}}\}_{k=0}^{\infty}$ *and* $\{\overline{p}_{\texttt{rel}}\}_{k=0}^{\infty}$ *and strategy* $\sigma_{\texttt{rel}}^*$.

Furthermore, relaxed robust value iteration yields bounds in the satisfaction probabilities that contain the ones obtained from robust value iteration. We formally state this result in Theorem 5-3.3.

Theorem 5-3.3. *Consider robust MDP* $\mathcal{M}^R$ *and sequences* $\{\underline{p}\}_{k=0}^{\infty}$ *and* $\{\overline{p}\}_{k=0}^{\infty}$ *obtained by performing robust value iteration as described in Theorem 5-3.1. Furthermore, assume sequences* $\{\underline{p}_{\texttt{rel}}\}_{k=0}^{\infty}$ *and* $\{\overline{p}_{\texttt{rel}}\}_{k=0}^{\infty}$ *have been computed by performing relaxed robust value iteration as described in Definition 5-3.3. Then, the following holds:*

$$\begin{aligned} \underline{p}_{\texttt{rel}}^{k}(q) \leq \underline{p}^{k}(q) \\ \overline{p}_{\texttt{rel}}^{k}(q) \geq \overline{p}^{k}(q) \end{aligned} \tag{5-21}$$

for all $q \in Q$, $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$.

Proof. We provide the proof of Theorem 5-3.3 in Appendix A. □

The complexity of the LPs that we need to solve to perform relaxed robust value iteration is way lower than that of robust value iteration. We compare the complexity of such LPs in

Table 5-1. This reduction is further expressed as a percentage in Figure 5-2, where we compare the two LPs in the following case: first, the coupling π in (5-18) is defined over a reduced space $Q \times Q_s^{q,a}$. Secondly, all redundant variables in (5-18) and (5-20), have been eliminated. Note that, as the dimension of $Q_c^{q,a}$ grows, the reduction is less relevant. However, for a practical value of $|Q_c^{q,a}| \approx 200$ cells of a state space of $|Q| = 1601$, this reduction is still remarkable: both the number of variables and the number of inequality constraints are reduced by a factor of around 87%. This indicates how beneficial it can be to perform relaxed robust value iteration. However note that, in Figure 5-2, the number of inequalities practically unaltered. This is because, as Table 5-1 shows, the number of equality constraints is only increased by one.

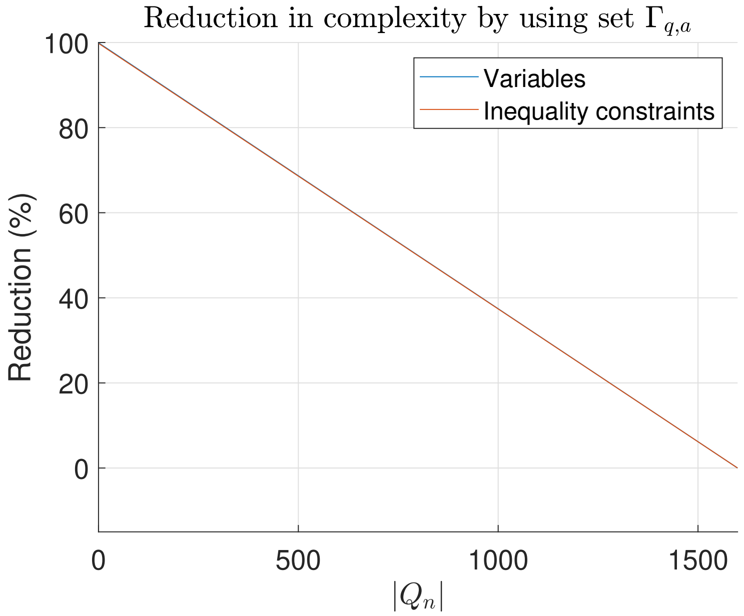


Figure 5-2: Reduction in the complexity of the LPs involved in relaxed robust value iteration when compared to those present in robust value iteration. These results were obtained for a robust MDP with $|Q| = 1601$, $|Q_s^{q,a}| = 9$.

However, despite benefits of performing relaxed robust value iteration instead of value iteration, the former entails a disadvantage: as Theorem 5-3.3 points out, the relaxed robust value iteration algorithm yields more conservative results than robust value iteration. Therefore, there exists a trade-off between computational complexity and conservatism when using this relaxed algorithm to perform strategy synthesis.

5-4 Correctness

Once we have synthesized a strategy for the robust MDP abstraction by following the procedure we described in Section 5-3, we need to address the following questions. The first one is: how can the original system use this strategy of the robust MDP? Furthermore, the second one is: how are we sure that, when the original system uses this strategy, the probabilistic guarantees found for the abstraction hold also for the original system? In this section we address the previous two problems: first we translate strategy $\sigma^*_{\mathcal{M}^R}$ of $\mathcal{M}^R$ to the (also memory-dependent) strategy $\sigma^*_{\mathcal{C}}$ of system (5-1). Secondly, we prove that the probabilistic guarantees $\underline{p}^k_{\mathcal{M}^R}$ and $\overline{p}^k_{\mathcal{M}^R}$ obtained for the robust MDP abstraction hold for the original system, when this one follows strategy $\sigma^*_{\mathcal{C}}$. This is referred to as correctness of the abstraction [11], [4],[7].

The procedure we follow is exactly the same we described for DR-IMDP abstractions in 4-4. First, we obtain strategy $\sigma^*_{\mathcal{C}}$ of system (5-1) using an expression analogous to (4-40). Secondly, we want to prove that the guarantees obtained for the abstraction hold for the original system. To do that, we begin by stating Lemma 5-4.1:

Lemma 5-4.1. *Consider an arbitrary distribution $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$ of the disturbance ξ. Assume that system* (5-1) *under probability $\widehat{P}_\xi$ of the disturbance has been abstracted into the nominal IMDP $\widehat{\mathcal{I}}$, as described in Section 5-2-1. Consider also, for fixed state $q \in Q$ and action $a \in A$ the set of feasible probability distributions $\widehat{\Gamma}_{q,a}$ of $\widehat{\mathcal{I}}$. Moreover, consider the continuous state $x \in q \subset \mathbb{R}^n$. Now, denote by $\gamma_{x,a}$ the transition probability from from x under a and for distribution P_ξ of ξ:*

$$\gamma_{x,a}(q') := \mathcal{T}_{\mathcal{C}}(q'|x, a; P_\xi)$$

for all $q' \in Q$. Finally, consider the set $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ as defined in (5-7)*. Then, $\gamma_{x,a} \in \widehat{\Gamma}^a_q \oplus \varepsilon$.*

Proof. A detailed proof of Lemma 5-4.1 is given in Appendix B. □

The intuition behind Theorem 5-4.1 is that set $\widehat{\Gamma}_{q,a} \oplus \varepsilon$ contains the transition probability $\gamma_{x,a}$ obtained by starting from any $x \in q$, under $a \in A$ and for any $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$. Notice that Lemma 5-4.1 is analogous to Lemma 4-4.1 for DR-IMDPs.

Denote by $P(w^k_{\mathbb{R}^n} \models \phi | w^k_{\mathbb{R}^n}(0) = x, X, \sigma^*_{\mathcal{C}}, P_\xi)$ the probability of the paths of system (5-1), for probability P_ξ of ξ, satisfying ϕ within k steps, while remaining in the safe region $X \subset \mathbb{R}^n$, by following strategy $\sigma^*_{\mathcal{C}}$ and starting from $x \in \mathbb{R}^n$. Now we state the theorem that ensures that the guarantees obtained for the abstraction hold for the original system:

Theorem 5-4.1. ***(Correctness of the Probabilistic Guarantees Of the Robust MDP Abstraction)*** *Consider system* (5-1) *and the robust MDP abstraction $\mathcal{M}^R$ of the former obtained as in Section 5-2. Furthermore, consider an scLTL formula ϕ, and the strategy $\sigma^*_{\mathcal{M}^R}$ and bounds $\underline{p}^k_{\mathcal{M}}$ and $\overline{p}^k_{\mathcal{M}^R}$ of $\mathcal{M}^R$ obtained for such specification as explained at the beginning of Section 5-4. Consider also strategy $\sigma^*_{\mathcal{C}}$ of system* (5-1)*, obtained from $\sigma^*_{\mathcal{M}^R}$ in a similar way to* (4-40)*. Then for any $x \in \mathbb{R}^n$ where $x \in q$, and for any $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$ it holds that*

$$P(w^k_{\mathbb{R}^n} \models \phi | w^k_{\mathbb{R}^n}(0) = x, X, \sigma^*_{\mathcal{C}}, P_\xi) \in [\underline{p}^k_{\mathcal{M}^R}(q), \overline{p}^k_{\mathcal{M}^R}(q)]$$

for all $k \in \mathbb{N}_{\geq 0} \cup \{\infty\}$.

We present a detailed proof of Theorem 5-4.1 in Appendix C.

Chapter 6

Experimental Results

In this Chapter we show the results, through simulations, of our proposed approaches to synthesize strategies for uncertain systems under complex specifications. Although the approach described in Chapter 5 is useful in a more general setting, here we always consider a data-driven one for simplicity: the disturbance ξ has unknown probability distribution $P_\xi^{\texttt{true}}$, and a finite amount of samples $\{\widehat{\xi}^i\}_{i=1}^N$ from ξ is available. Furthermore, the ambiguity sets we consider are Wasserstein balls $\mathbb{B}_\varepsilon(\widehat{P}_\xi^N)$ centered on the empirical distribution $\widehat{P}_\xi^N$ built from the samples, and with given radius $\varepsilon > 0$. Furthermore, we always use the 2-Wasserstein distance for the two following reasons: first, it penalizes more the transport of mass from the samples, allowing us to use bigger values of ε. Secondly, this means that the distance over the state space $\mathbb{R}^n$ will be the euclidean norm, $\|\cdot\|_2$. This, in turn, allows us to implement very fast algorithms to find distances between samples and rectangles, leading to shorter abstraction times.

We prove the effectiveness of our approaches in two different systems, a linear and a nonlinear one. Furthermore, the disturbances we consider in this section are either Gaussian or Gaussian Mixtures with two components. We provide results for several sizes N of the sample set and radius ε. The complex specification we consider is always the same: $\phi := \neg\texttt{obs}\mathbb{U}\texttt{des}$, where `obs` and `des` are the observations that correspond, respectively, to an obstacle and to the desired region. The previous specification is commonly denoted as "reach-avoid" problem. We synthesize strategies by running the value iteration algorithms until we achieve convergence. Consider the bounds in probability $\underline{p}^k$, $\overline{p}^k$ obtained in the value iteration algorithms. We consider that the solution of said algorithms has converged when the following stopping criteria is fulfilled:

$$\begin{aligned} &\sup_{q\in Q}(\underline{p}^{k+1}(q) - \underline{p}^k(q)) \leq \texttt{tol}_{\texttt{VI}} \\ &\sup_{q\in Q}(\overline{p}^{k+1}(q) - \overline{p}^k(q)) \leq \texttt{tol}_{\texttt{VI}}, \end{aligned}$$

where $\texttt{tol}_{\texttt{VI}}$ is the tolerance in value iteration. We always set $\texttt{tol}_{\texttt{VI}} = 0.02$. Furthermore, consider the converged bounds in the probability of satisfying the specification $\underline{p}$, $\overline{p}$. We define the “Average error" $e_{\texttt{avg}}$ as:

$$e_{\texttt{avg}} = \frac{1}{|Q|} \sum_{q \in Q} (\overline{p}(q) - \underline{p}(q)). \tag{6-1}$$

This error allows us to assess the conservatism of the solution.

First, in Section 6-1 we describe the dynamics of the systems that we consider here. After that, in Section 6-2 we present the results of the approach we described in Chapter 4. Next, in Section 6-3 we show the results of the approach we explained in Chapter 5.

6-1 System Dynamics

Linear System

In this section we describe the linear system that we use in our simulations. This one is a simplified model of the kinematic unicycle system. We consider a discrete-time version of the continuous-time unicycle model [39], which we obtain by using the Euler approach with a time discretization of $\Delta t = 1$. To obtain a simpler system, we consider that the velocity v is fixed, and we also consider no dynamics in the orientation angle θ, being this one the control input:

$$x_{t+1} = x_t + \Delta t (v \begin{bmatrix} \cos(\theta_t) \\ \sin(\theta_t) \end{bmatrix} + \xi_t). \tag{6-2}$$

The states of the system are the position in the 2-dimensional space, $x_t \in \mathbb{R}^2$. The control input is the angle θ, as we already highlighted, which can take values in the continuous range $U_{\mathcal{C}} = [0, 2\pi)$. Finally, the random term $\xi \in \mathbb{R}^2$ is an external i.i.d. disturbance, for example, wind, that disturbs the position of the unicycle with probability P_ξ^{true} for all t. In this section we always consider a velocity of $v = 1$.

Nonlinear System with 4 Modes

The nonlinear system we chose is the same one employed in [11] to synthesize strategies for partially-known switched stochastic systems using IMDP abstractions. The dynamics of said system are the following:

$$x_{t+1} = x_t + \widetilde{f}(x_t, u_t) + \xi_t, \tag{6-3}$$

where

$$\widetilde{f}(x, u) = \begin{cases} [0.5 + 0.2\sin(x^{(2)}), 0.4\cos(x^{(1)})]^T & \text{if } u = 1 \\ [-0.5 + 0.2\sin(x^{(2)}), 0.4\cos(x^{(1)})]^T & \text{if } u = 2 \\ [0.4\cos(x^{(2)}), 0.5 + 0.2\sin(x^{(1)})]^T & \text{if } u = 3 \\ [0.4\cos(x^{(2)}), -0.5 + 0.2\sin(x^{(1)})]^T & \text{if } u = 4 \end{cases} \tag{6-4}$$

In (6-4), $x^{(i)}$ denotes the i-th component of the state. The states of the system are the position in the 2-dimensional space, $x_t \in \mathbb{R}^2$. The control input is the switching control input u, which switches between the modes in (6-4) by takes values in $U_{\mathcal{C}} = \{1, 2, 3, 4\}$. Finally, the random term $\xi \in \mathbb{R}^2$ is an external i.i.d. disturbance, that disturbs the position of the system with probability P_ξ^{true} for all t.

# Exp.	ε	$e_{\mathtt{avg}}$	Abstraction Time	Synthesis Time
1	10^{-3}	0.021	$8.6 + 5.7$ min	1 min (24+12 it)
2	2×10^{-3}	0.094	$8.76 + 4.8$ min	1.2 min (31+11 it)
3	3×10^{-3}	0.338	$9.5 + 5.5$ min	1.2 min (33+11 it)
4	5×10^{-3}	0.83	$8.5 + 4.5$ min	0.6 min (10+10 it)

Table 6-1: Summary of the experiments performed for the linear system (6-2). Side note: the abstraction time is indicated as $t_1 + t_2$, being t_1 and t_2 the time required to compute the lookup tables and to perform the rest of the abstraction, respectively. Furthermore, we state both the time and the number of iterations required to perform strategy synthesis. We express the latter as ($n_{\mathtt{lower}} + n_{\mathtt{upper}}$ it), being $n_{\mathtt{lower}}$ and $n_{\mathtt{upper}}$, respectively, the number of iterations required to achieve convergence of the lower and upper bounds in probability.

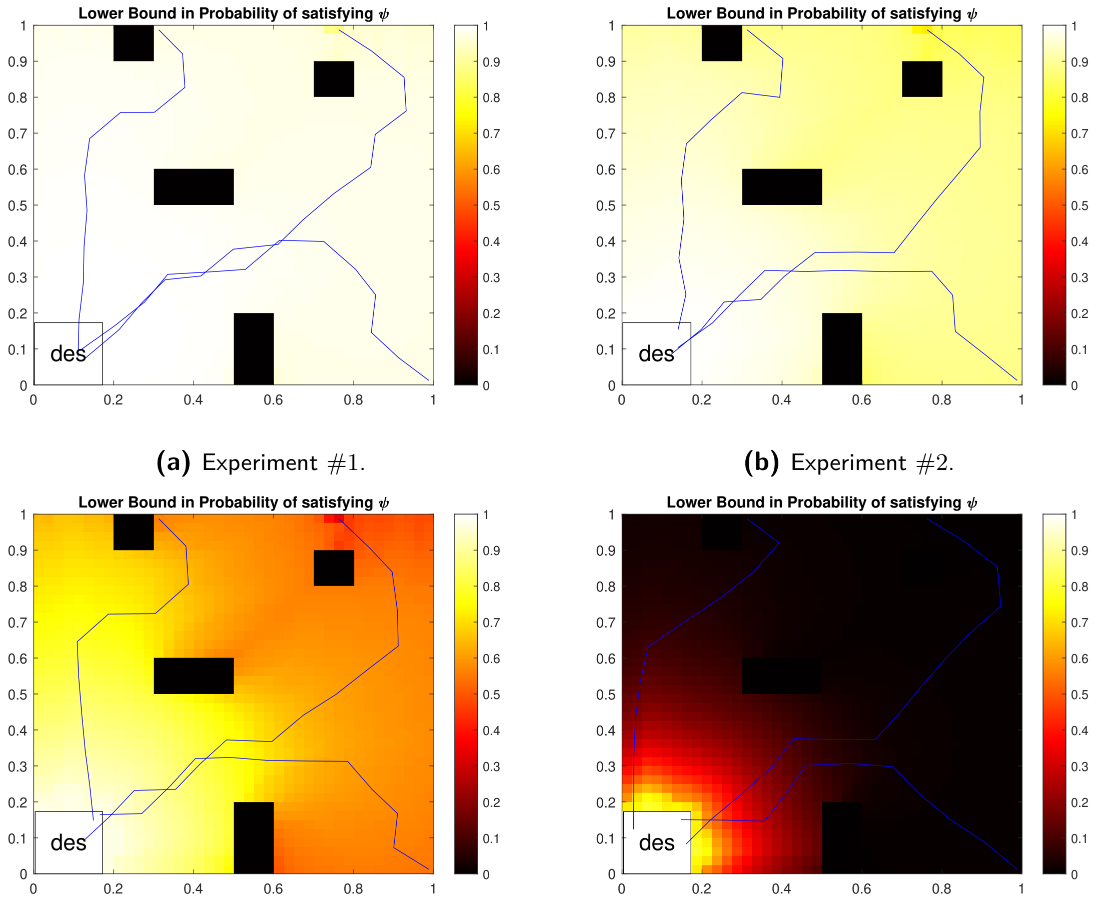
6-2 Results of the Approach Based on DR-IMDP Abstractions

In this section we describe the results of the approach we described in Chapter 4. We start by showing the results in the linear system of Section 6-1, and then we focus on the results obtained for the nonlinear system in that same section.

Results of the Approach Based on DR-IMDP Abstractions on the Linear System

For this system, the workspace we consider is always the set $X = [0, 1] \times [0, 1] \subset \mathbb{R}^2$. Furthermore, we always use the same discretization of the state space, which is a uniform grid that leads to a state space of the abstraction Q of cardinality $|Q| = 1601$. The set of actions of the DR-IMDP abstraction is obtained as a uniform partition of $U_{\mathcal{C}}$: $A = \{0, 2\pi\frac{1}{n_a}, \ldots, 2\pi\frac{n_a-1}{n_a}\}$. Furthermore, we choose $n_a = 8$. Additionally, $P_{\xi}^{\mathtt{true}}$ is a Gaussian Mixture with two components, centered at $[-0.01, 0]$ and $[0.01, 0]$, respectively, where both components have a standard deviation of 0.005. This means that the centers of both components are separated by a distance close to the size of the state discretization. Moreover, we always use a set of $N = 500$ samples to obtain the abstraction. The lookup tables we computed for q^* and X have, respectively, 279596 and 83716 entries, which we arranged as follows: we made the grid finer in regions closer to q^* and X, and coarser in the more distant regions. We also highlight that, to speed up the process of searching in the lookup tables, we clustered the data of said tables: the final size of the tables is of 3700 and 5852 entries, respectively. This highly reduces the abstraction time. Furthermore, we set the value of the distance threshold that switches between using the lookup table and the overapproximation as described in Chapter 4 to $D_{\mathtt{th}} = 0.04$. Then we used the approach in4 to perform strategy synthesis. We give a summary of the results in Table 6-1. Furthermore, the lower bounds we obtained in the probability of satisfying ϕ are shown in Figure 6-1. Additionally, since the same upper bound in the probability of satisfying ϕ was found for Experiments #1-#4, we show this single plot in Figure 6-2.

The results observed both in Table 6-1 and in Figures 6-1 and 6-2 show, as we expected, that a bigger size of the ambiguity set leads to looser bounds of satisfying the specification. In this case, we observe that the bigger the ambiguity, the smaller is the value of the lower bound in the probability of satisfying ϕ. Regarding the time required to build the abstraction, its similar for all the experiments. This is because we have not changed the number of samples



(a) Experiment #1.

(b) Experiment #2.

(c) Experiment #3.

(d) Experiment #4.

Figure 6-1: Results of Experiments in Table 6-1. Lower bound in the probability of satisfying ϕ. The three blue trajectories correspond to Monte Carlo simulations of the system taking samples from the true probability distribution of the disturbance.

N, $D_{\mathtt{th}}$, or the number of entries of the lookup tables, which greatly influence the abstraction time. Furthermore, we observe that the number of iterations needed to achieve convergence is almost the same for every choice of ε in experiments #1-#3. However, this is not the case in Experiment #4, for which convergence is quickly achieved. This is because the bounds in the probability of satisfying ϕ are almost trivial for $\varepsilon = 5 \times 10^{-3}$: we observe a lower bound closer to zero in most of the state space in Figure 6-1d, and an upper bound of one where there are no obstacles in Figure 6-2.

We also plot the vector field corresponding to the system in closed-loop with the synthesized strategies in Experiments #1-#4 in Figure 6-3. The strategies we observe are almost the same for all the experiments, even when the bounds we obtained are loose, as in Experiment #4. Moreover, we found these bounds very conservative, by performing Monte Carlo simulations in which we took new samples from the true distribution every time step: for 1000 MC simulations starting on the lower-right corner of the workspace, we found that all of them satisfied ϕ, for strategies #1 to #4.

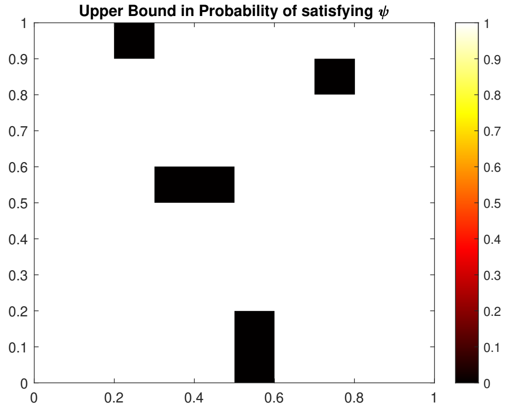
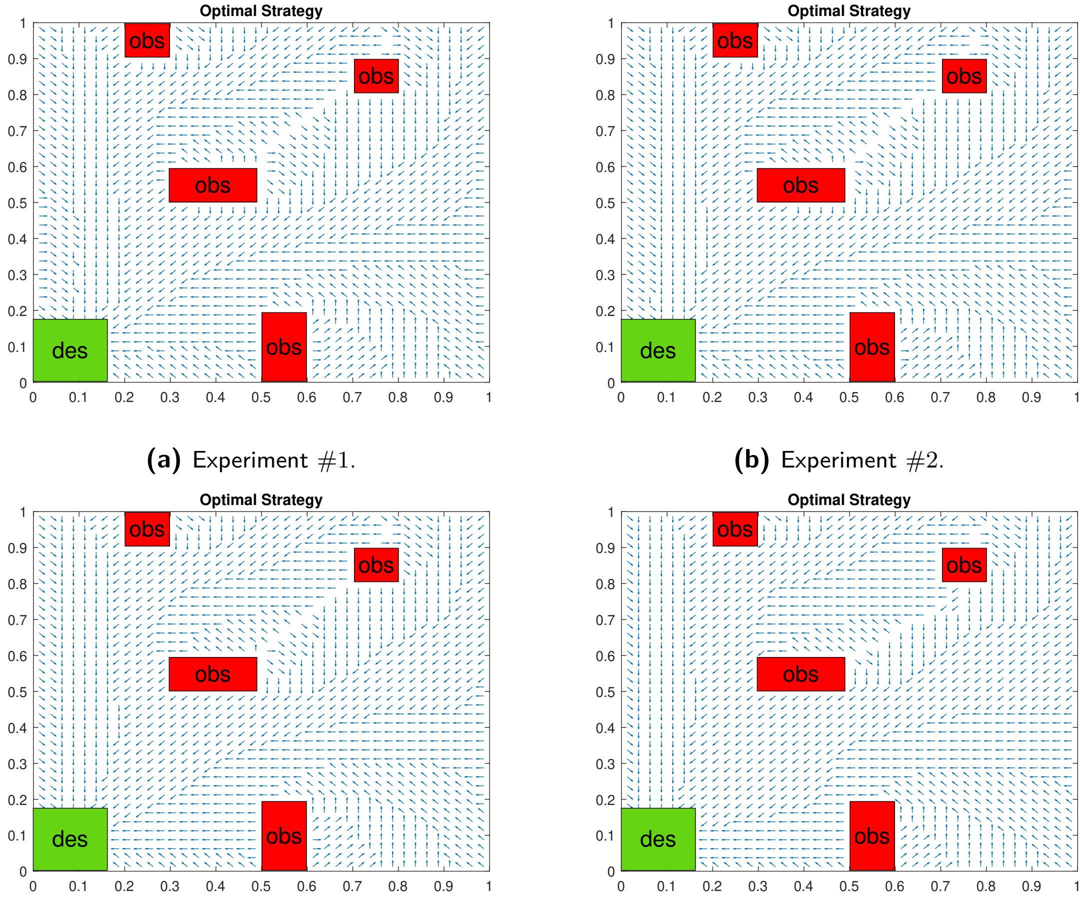


Figure 6-2: Results of the experiments in Table 6-1. Upper bound in the probability of satisfying ϕ.



(a) Experiment $\#1$.

(b) Experiment $\#2$.

(c) Experiment $\#3$.

(d) Experiment $\#4$.

Figure 6-3: Results of Experiments in Table 6-1. Synthesized strategy.

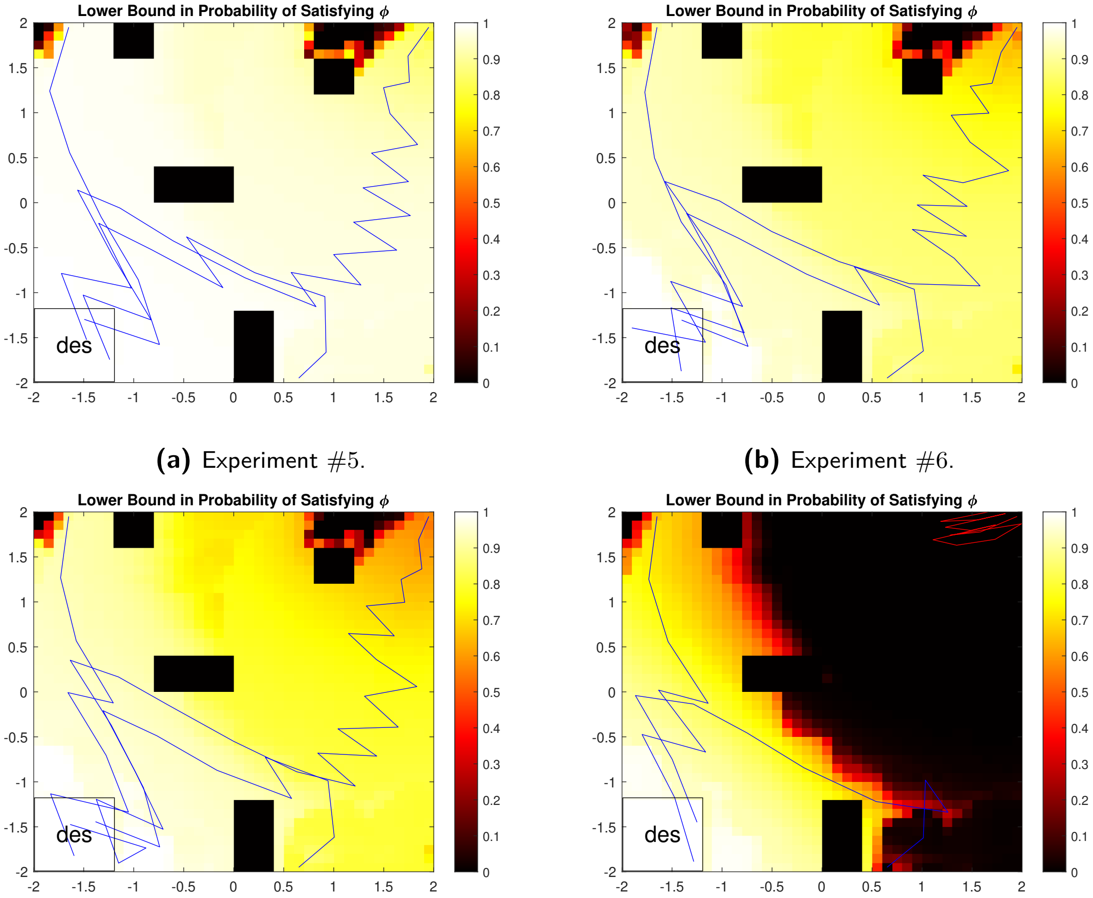
# Experiment	ε	$e_{\texttt{avg}}$	Abstraction Time	Synthesis Time
5	2×10^{-3}	0.04	$1.2 + 12$ min	0.96 min (63+19 it)
6	4×10^{-3}	0.13	$1.2 + 13.2$ min	0.9 min (61 + 17 it)
7	5×10^{-3}	0.16	$1.3 + 11.7$ min	1 min (63+17 it)
8	7×10^{-3}	0.6	$1.2 + 11.7$ min	0.6 min (33+16 it)

Table 6-2: Summary of the experiments performed for the nonlinear system (6-3). Side note: the abstraction time is indicated as $t_1 + t_2$, being t_1 and t_2 the time required to compute the lookup tables and the to perform the rest of the abstraction, respectively. Furthermore, we state both the time and the number of iterations required to perform strategy synthesis. We express the latter as ($n_{\texttt{lower}} + n_{\texttt{upper}}$ it), being $n_{\texttt{lower}}$ and $n_{\texttt{upper}}$, respectively, the number of iterations required to achieve convergence of the lower and upper bounds in probability.

Results of the Results of the Approach Based on DR-IMDP Abstractions on the Nonlinear System

For this system, the workspace we consider is always the set $X = [-2, 2] \times [-2, 2] \subset \mathbb{R}^2$. Furthermore, we always use the same discretization of the state space, which is a uniform grid, yielding a state space of the abstraction Q of cardinality $|Q| = 1601$. The set of actions of the DR-IMDP is the same as the set of modes of system (6-3): $A = U_C$. Additionally, $P_\xi^{\texttt{true}}$ is a Gaussian Mixture with two components, centered at $[-0.05, 0]$ and $[0.05, 0]$, respectively, where both components have a standard deviation of 0.02. Again, this means that the centers of both components are separated by a distance close to the size of the state discretization. Moreover, we always use a set of $N = 20$ samples to obtain the abstraction. The lookup tables we computed for q^* and X have, respectively, 844204 and 286336 entries, which we arranged as follows: we made the grid finer in regions closer to q^* and X, and coarser in the more distant regions. As we did in the case of the linear system, to speed up the process of searching in the lookup tables, we clustered the data of said tables: the final size of the tables is of 62052 and 24136 entries, respectively. Furthermore, we set $D_{\texttt{th}} = 0.15$. Then we used the approach in4 to perform strategy synthesis. We give a summary of the results in Table 6-2. Furthermore, the lower bounds we obtained in the probability of satisfying ϕ are shown in Figure 6-1. Additionally, since we obtained the same upper bound in the probability of satisfying ϕ as in the case of the linear system, we do not show this result here: we limit ourselves to point out to Figure 6-2.

The results observed both in Table 6-2 and in Figures 6-4 show that the radius of the ambiguity set ε has the same effect as in the case of system (6-2): the bigger the ambiguity set, the looser are the bounds in probability of satisfying the specification. Regarding the time required to build the abstraction, its similar in Experiments #5-#8. This is because the number of samples N, $D_{\texttt{th}}$ and the number of entries of the lookup tables is the same in these experiments, which highly influence the abstraction time. However, we note that, when compared to the results obtained in Table 6-1, the lookup tables require less time to be computed, whereas the abstraction takes more time. The first result is due to the fact that in this section we set $N = 20$, as opposed to the $N = 500$ samples used for system (6-2). Furthermore, in this section, the clustered tables have a bigger number of entries than those used for system (6-2). For this reason, the time required to search in the tables is bigger in this case, even when the number of actions of system (6-4) is half the number of those of system (6-2). The previous results highlight the effect of the number of samples and the size



(a) Experiment #5.

(b) Experiment #6.

(c) Experiment #7.

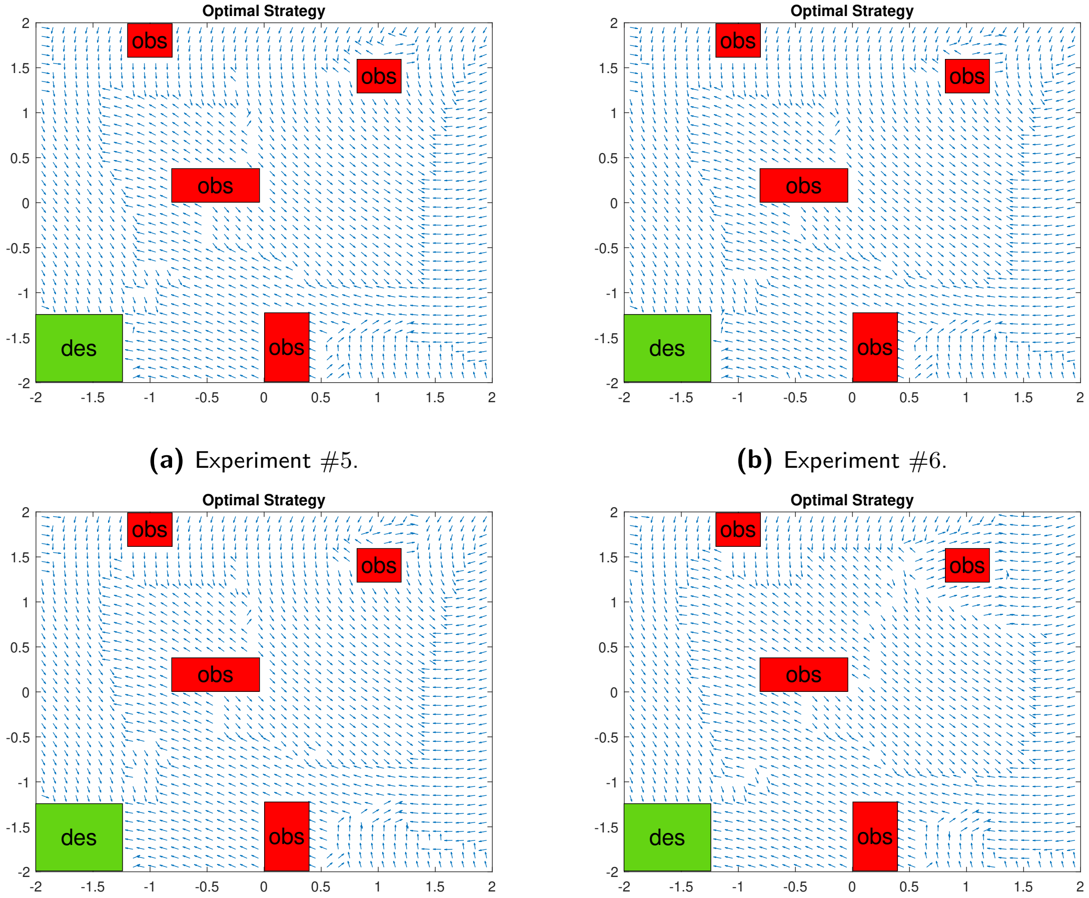
(d) Experiment #8.

Figure 6-4: Results of Experiments in Table 6-1. Lower bound in the probability of satisfying ϕ. The three blue trajectories correspond to Monte Carlo simulations of the system taking samples from the true probability distribution of the disturbance.

of the lookup tables in the complexity of obtaining the abstraction.

The effect of ε in the number of iterations required to achieve convergence of interval value iteration is the same observed in Table 6-2: it is almost the same for Experiments #5-#7, but different for Experiment #8, for which convergence is achieved faster. This last part is because the bounds in the satisfaction probability of Experiment #8 are almost trivial in a big portion of the state space: we observe a lower bound closer to zero in almost half the state space in Figure 6-4d, and an upper bound of one where there are no obstacles in Figure 6-2.

We also plot the vector field corresponding to the system in closed-loop with the synthesized strategies in Figure 6-5. The strategies we observe are almost the same for Experiments #5-#7. However, the one obtained in Experiment #8 is quite different, and does not lead to a good closed-loop performance: the paths that start in the dark regions in Figure 6-4d under this strategy do not satisfy the specification in most of the cases. Notice that in Figure 6-4d, the trajectory starting in the upper-right corner does not satisfy ϕ. Furthermore, from 1000 Monte Carlo simulations performed by starting in the upper-right corner, none satisfied the



(a) Experiment #5.

(b) Experiment #6.

(c) Experiment #7.

(d) Experiment #8.

Figure 6-5: Results of Experiments in Table 6-2. Synthesized strategy.

specification. This was not the case, however, for Experiments #5-#7, for which all Monte Carlo trajectories starting in that region satisfied ϕ.

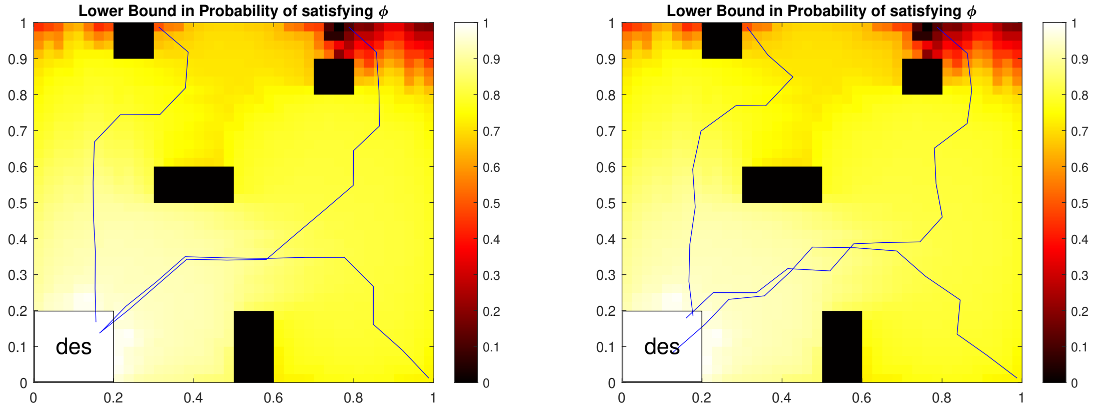
6-3 Results of the Approach Based on Robust MDP Abstractions

In this section we describe the results of the approach we described in Chapter 5. We limit to show the effectiveness of this approach in the linear system (6-2). Furthermore, all results we show in this section are obtained for a radius of the ambiguity ball $\varepsilon = 0.01$. The reason behind the reduced number of results we present here, when compared to Section 6-2 is that the approach of Chapter 5 is very expensive to use. Furthermore, for simplicity, we also use this approach in a data-driven setting, this is, when the center of the ambiguity set is an empirical distribution built from data.

For system (6-2), we consider the same workspace, state discretization and actions that we considered in Section 6-2 for this same system. However, here we obtain results for two classes of unknown distributions $P_\xi^{\mathtt{true}}$: the first one a Gaussian with zero mean and standard deviation of 0.005. The second one is a Gaussian Mixture with two components, centered at $[-0.01, 0]$ and $[0.01, 0]$, respectively, where both components have a standard deviation of 0.005. Again, this means that the centers of both components are separated by a distance close to the size of the partition of the state space. Moreover, we always use a set of $N = 10$ samples to obtain the empirical distribution. Then, we obtain the nominal IMDP by using the approach in Chapter 4 and setting $\varepsilon = 0$ and $D_{\texttt{th}} = 0$. The lookup tables we computed for q^* and X have, respectively, 2057500 and 9966780 entries, which we arranged as follows: we made the grid finer in regions closer to q^* and X, and coarser in the more distant regions. We also clustered the data of said tables: the final size of the tables is of 12176 and 59632 entries, respectively. Then we used the approach in Chapter 5 to perform strategy synthesis. We give a summary of the results in Table 6-3. Furthermore, the lower bounds we obtained in the probability of satisfying ϕ are shown in Figure 6-6. Additionally, since the same upper bound in the probability of satisfying ϕ was found for Experiments #9-#10, we show this single plot in Figure 6-7.

# Experiment	$P_\xi^{\texttt{true}}$	$e_{\texttt{avg}}$	Abstraction Time	Synthesis Time
9	Gaussian	0.18	$7 + 2.4$ min	24.4 h (40 + 9 it)
10	Gaussian Mixture	0.18	$7 + 3$ min	23.8 h (40 + 8 it)

Table 6-3: Summary of the experiments performed for the linear system (6-2). Side note: the abstraction time is indicated as $t_1 + t_2$, being t_1 and t_2 the time required to compute the lookup tables and the to perform the rest of the nominal IMDP abstraction, respectively. Furthermore, we state both the time and the number of iterations required to perform strategy synthesis. We express the latter as ($n_{\texttt{lower}} + n_{\texttt{upper}}$ it), being $n_{\texttt{lower}}$ and $n_{\texttt{upper}}$, respectively, the number of iterations required to achieve convergence of the lower and upper bounds in probability.



(a) Experiment #9.

(b) Experiment #10.

Figure 6-6: Results of Experiments in Table 6-3. Lower bound in the probability of satisfying ϕ. The three blue trajectories correspond to Monte Carlo simulations of the system taking samples from the true probability distribution of the disturbance.

The results observed both in Table 6-3 and in Figures 6-6 and 6-7 show, as we expected, that the approach based on robust MDP abstractions is able to provide good results even when the ambiguity set is relatively big. Notice that the results in this section are obtained for

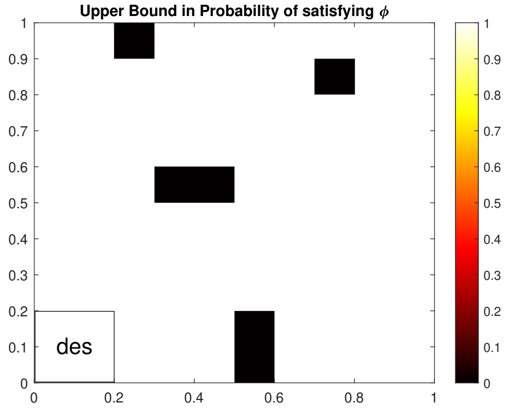
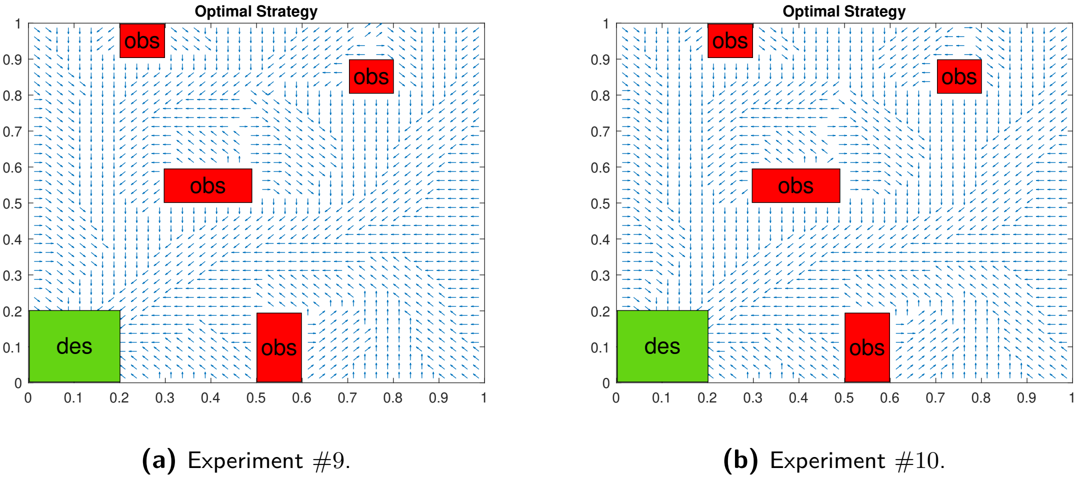


Figure 6-7: Results of the experiments in Table 6-3. Upper bound in the probability of satisfying ϕ.

$\varepsilon = 0.01$, whereas Experiments #3 and #4 in Section 6-2 already yielded more conservative results for $\varepsilon = 3 \times 10^{-3}$ and $\varepsilon = 5 \times 10^{-3}$, respectively. Regarding the time required to build the abstraction, its similar for both experiments. This is because we have not changed the number of samples N, $D_{\mathtt{th}}$, or the number of entries of the lookup tables, which greatly influence the abstraction time. Furthermore, we observe that it takes way more time to synthesize a strategy for these abstractions than for DR-IMDPs used in Section 6-2, as we expected.

We also plot the vector field corresponding to the system in closed-loop with the synthesized strategies in Experiments #9-#10 in Figure 6-8. The strategies we observe are almost identical for both experiments.



(a) Experiment #9.

(b) Experiment #10.

Figure 6-8: Results of Experiments in Table 6-3. Synthesized strategy.

Chapter 7

Conclusion

In this thesis we presented two approaches to synthesize strategies for dynamical systems with random, uncertain disturbances, under specifications given as scLTL formulas. In order to capture the uncertainty in the distribution of the disturbance we used ambiguity sets based on the Wasserstein distance. The two proposed approaches rely on performing an abstraction of the original system: the first one relies on abstracting the system into an IMDP, which we denote DR-IMDP. After that, a strategy that enforces an scLTL formula is synthesized using already existing algorithms for IMDPs. On the other hand, the second approach employs a robust MDP abstraction of the original system. In order to synthesize a strategy that enforces an scLTL specification, we proposed a modified value iteration algorithm called robust value iteration. The proposed approaches effectively account for the ambiguity in the distribution of the disturbance, leading to robust strategies that enforce complex specifications. Furthermore, we proved that the probabilistic guarantees that the abstractions satisfy the specifications also hold for the original, uncertain system.

While the first approach relies on the assumption that the center of the ambiguity set is an empirical distribution, the second one works in a more general setting: the the center of the ambiguity set does not need to be an empirical distribution. Furthermore, as we described in Chapter 5, robust MDPs are less conservative abstractions than DR-IMDPs, and yield less conservative satisfaction guarantees. Moreover, as the results from Chapter 6 show, it is easier to perform a robust MDP abstraction than a DR-IMDP one. This is because, to obtain a robust MDP abstraction, the only computationally intensive part is to compute the nominal IMDP. Furthermore, less problems need to be solved to obtain a nominal IMDP than to obtain a DR-IMDP abstraction. On the other hand, synthesizing a strategy for a DR-IMDP is way more efficient than doing so for a robust MDP. This is because the interval value iteration algorithm for IMDPs and DR-IMDPs is way more efficient than robust value iteration for robust MDPs. The overall result is that using the approach that relies in robust MDP abstractions is way less efficient. Therefore, we find a trade-off between conservatism and efficiency: using robust MDP abstractions for strategy synthesis we obtain less conservative results than using DR-IMDPs, but the approach based on the latter is faster to use.

In Chapter 6 we tested our approaches in the data-driven setting, this is, when the ambiguity ball is centered in an empirical distribution obtained from data. This was for the sake of simplicity, since we already highlighted in Chapter 5 that our second approach can also be employed when the uncertainty model is more general. We showed how our approaches are able to synthesize robust strategies that enforce scLTL specifications for both linear and nonlinear systems. Moreover, we presented results for different sizes of the sample set and shapes of the unknown data-generating distribution. Furthermore, we showed how, the bigger the ambiguity, the more conservative the satisfaction guarantees become. One of the key results of Chapter 6 is that we demonstrated experimentally the previously mentioned trade-off between conservatism of the solution and efficiency. We showed that using the approach in Chapter 5 we obtain tighter guarantees than using the approach in Chapter 4 even when the ambiguity ball is bigger. Furthermore, we demonstrated that synthesizing a strategy for a robust MDP is way less efficient than doing so for a DR-IMDP.

7-1 Future Work

We now give a brief summary of the possible future research directions regarding the topics involved in this thesis.

- Obtain results for different ambiguity models than the data-driven one considered in Chapter 6. For example, we could consider a Wasserstein-based ambiguity set centered on a gaussian distribution. Then, we could use the approaches in [4] and [5] to obtain the nominal IMDP abstraction, and then easily define the robust MDP that accounts for all the probabilities in the ambiguity set.

- Find more efficient ways of performing robust value iteration. Making use of entropic regularization of optimal transport [31] might be useful.

- Obtain robust MDP abstractions using different ambiguity sets than Wasserstein balls.

- Since the unknown probability distribution of the disturbance is always the same, find a way to enforce this constraint. This would be translated into a coupling of the transition probabilities that are chosen by the adversary at each state, reducing conservatism of the solution.

- Obtain results for higher order systems.

- Propose an algorithm to deal with systems in which the disturbance only affects part of the dynamics. Exploit this structure to formulate more efficient algorithms that allow to obtain abstractions of higher order systems.

- Combine robust value iteration with adaptive refinement algorithms that allow to obtain acceptable results while employing a smaller state space, as in [3], [5]. Doing this would reduce the complexity of the abstraction and increase the efficiency of the robust value iteration algorithm.

Appendix A

Proof of Theorem 5-3.3

First of all, if the nominal distribution $\widehat{P}_\xi$ of ξ has a bounded support, we only consider (discrete) nominal distributions $\widehat{\gamma}_{q,a}$ defined over $Q_s^{q,a}$. Therefore, the couplings π we consider are defined over the product space $Q \times Q_s^{q,a}$. Taking into account the previous considerations, (5-18) becomes:

$$\begin{aligned}
\underline{P}(q,a,q_j) \leq \widehat{\gamma}_{q,a}(q_j) \leq \overline{P}(q,a,q_j) \qquad & q_j \in Q_s^{q,a} && \text{(A-1a)}\\
\sum_{q_j \in Q_s^{q,a}} \widehat{\gamma}_{q,a}(q_j) = 1 \qquad & && \text{(A-1b)}\\
\pi_{ij} \geq 0, \qquad & q_i \in Q,\ q_j \in Q_s^{q,a} && \text{(A-1c)}\\
\sum_{q_i \in Q} \pi_{ij} = \widehat{\gamma}_{q,a}(q_j), \qquad & q_j \in Q_s^{q,a} && \text{(A-1d)}\\
\sum_{q_j \in Q_s^{q,a}} \pi_{ij} = \gamma_{q,a}(q_i), \qquad & q_i \in Q && \text{(A-1e)}\\
\sum_{q_i \in Q} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij}^p \leq \varepsilon^p. \qquad & && \text{(A-1f)}
\end{aligned}$$

We now leave the first and second constraints in (A-1) are left as they are. However, we write the fourth constraint as:

$$\begin{aligned}
& \sum_{q_i \in Q} \pi_{ij} = \widehat{\gamma}_{q,a}(q_j), \qquad q_j \in Q_s^{q,a}\\
\iff & \sum_{q_i \in Q_c^{q,a}} \pi_{ij} + \sum_{q_i \in Q \setminus Qc^{q,a}} \pi_{ij} = \widehat{\gamma}_{q,a}(q_j), \qquad q_j \in Q_s^{q,a}.
\end{aligned} \tag{A-2}$$

Furthermore, we separate the fifth constraint in (A-1) as the following two sets of constraints:

$$\begin{aligned}
& \sum_{q_j \in Q_s^{q,a}} \pi_{ij} = \gamma_{q,a}(q_i), \qquad q_i \in Q\\
\iff & \begin{cases} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} & = \gamma_{q,a}(q_i), \qquad q_i \in Q_c^{q,a} \\ \sum_{q_j \in Q_s^{q,a}} \pi_{ij} & = \gamma_{q,a}(q_i), \qquad q_i \in Q \setminus Q_c^{q,a} \end{cases}.
\end{aligned} \tag{A-3}$$

Additionally, we relax the second constraint in (A-3) as follows:

$$\sum_{q_j \in Q_s^{q,a}} \pi_{ij} = \gamma_{q,a}(q_i), \qquad q_i \in Q \setminus Q_c^{q,a} \Longrightarrow \sum_{q_j \in Q_s^{q,a}} \sum_{q_i \in Q \setminus Q_c^{q,a}} \pi_{ij} = \sum_{q_i \in Q \setminus Q_c^{q,a}} \gamma_{q,a}(q_i). \tag{A-4}$$

Furthermore, when it comes to the last constraint in (A-1), we relax it by using the distance d_{is} defined in (5-19), which is an under estimator of d_{ij} for all $q_i \in Q \setminus Q_c^{q,a}$, $q_j \in Q_s^{q,a}$:

$$\begin{aligned} & \sum_{q_i \in Q} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij} \leq \varepsilon \\ \Longleftrightarrow & \sum_{q_i \in Q_c^{q,a}} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij} + \sum_{q_i \in Q \setminus Q_c^{q,a}} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij} \leq \varepsilon \\ \Longrightarrow & \sum_{q_i \in Q_c^{q,a}} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij} + \sum_{q_i \in Q \setminus Q_c^{q,a}} d_{is} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} \leq \varepsilon. \end{aligned} \tag{A-5}$$

Furthermore, using the last expression in (A-3) we get

$$\begin{aligned} & \sum_{q_i \in Q_c^{q,a}} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij} + \sum_{q_i \in Q \setminus Q_c^{q,a}} d_{is} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} \leq \varepsilon \\ = & \sum_{q_i \in Q_c^{q,a}} \sum_{q_j \in Q_s^{q,a}} \pi_{ij} d_{ij} + \sum_{q_i \in Q \setminus Q_c^{q,a}} d_{is} \gamma_{q,a}(q_i) \leq \varepsilon. \end{aligned} \tag{A-6}$$

Taking a look at expressions (A-2) and (A-4), we notice that the variables π_{ij} for all $q_i \in Q \setminus Q_c^{q,a}$, $q_j \in Q_s^{q,a}$ never appear separated, but always as a sum. Therefore, we get rid of said individual variables by defining

$$\pi_{rj} := \sum_{q_i \in Q \setminus Q_c^{q,a}} \pi_{ij}, \qquad q_j \in Q_s^{q,a}. \tag{A-7}$$

Note that π_{rj} is non-negative by the third expression in (A-1). Using this new variable, expressions (A-2), and (A-4) become:

$$\begin{aligned} \sum_{q_i \in Q_c^{q,a}} \pi_{ij} + \pi_{rj} = \widehat{\gamma}_{q,a}(q_j), \qquad q_j \in Q_s^{q,a} \\ \sum_{q_j \in Q_s^{q,a}} \pi_{rj} = \sum_{q_i \in Q \setminus Q_c^{q,a}} \gamma_{q,a}(q_i) \end{aligned} \tag{A-8}$$

We choose to use the notation of π_{rj} for the following reason: this variable concentrates all probability mass assigned from state $q_j \in Q_s^{q,a}$ to all states outside of $Q_c^{q,a}$. This is the same as regarding the whole set $Q \setminus Q_c^{q,a}$ as a single state q_r, to which mass can be transported from $Q_s^{q,a}$. Therefore, we can consider the variable π_{rj} as a component of a coupling defined over a reduced product space $(Q_c^{q,a} \cup q_r^{q,a}) \times Q_s^{q,a}$, where $q_r^{q,a}$ represents set $Q \setminus Q_c^{q,a}$.

Now, consider the first and second constraints in (A-1). Consider also the non-negativeness constraint on π and the constraint obtained in (A-3) for all $q_i \in Q_c^{q,a}$. Finally, consider the relaxed constraints in (A-6) and (A-8). It becomes clear that said set of constraints is the same as $\Gamma_{q,a}$, as defined in (5-20). Furthermore, since we have obtained the constraints in

(5-20) by relaxing those in (5-18), it follows that $\Gamma_{q,a} \supset \widehat{\Gamma}_{q,a} \oplus \varepsilon$. Therefore, for any $p : Q \to \mathbb{R}$ we obtain

$$\min_{\gamma_{q,a} \in \Gamma_{q,a}} \sum_{q_i \in Q} \gamma_{q,a}(q_i) p(q_i) \leq \min_{\gamma_{q,a} \in \widehat{\Gamma}_{q,a} \oplus \varepsilon} \sum_{q_i \in Q} \gamma_{q,a}(q_i) p(q_i),$$

for all $q \in Q$, $a \in A$. The previous result allows us to prove that sequence $\{\underline{p}_{\mathtt{rel}}(q)\}_{k=0}^{\infty}$ is lower bounded by sequence $\{\underline{p}(q)\}_{k=0}^{\infty}$, for all $q \in Q$. Therefore, Theorem 5-3.3 follows.

Appendix B

Proof of Lemma 5-4.1

Let us first establish a relation between the p-Wasserstein distance between probabilities over $\mathbb{R}^n$ and the one between probabilities over Q:

Lemma B-0.1. *Consider two distributions, P, and P', defined over $\mathbb{R}^n$. Let us define the discrete distributions γ and γ' over Q, such that each entry $\gamma(q')$ and $\gamma'(q')$ contains the fraction of mass of P and P' inside state q', respectively: $\gamma(q') = P(q')$ and $\gamma'(q') = P(q')$ for all $q' \in Q$. Finally, consider the p-Wasserstein distance $\mathcal{W}_p(P, P')$ based on the p-norm, and the p-Wasserstein distance $\mathcal{W}_p(\gamma, \gamma')$ based on distance* (5-6)*. Then, for $\varepsilon > 0$, the following holds:*

$$\mathcal{W}_p(P', P) \leq \varepsilon \Rightarrow \mathcal{W}_p(\gamma', \gamma) \leq \varepsilon.$$

Proof. From the definition of Wasserstein distance in (3-2) between the distributions P' and P we get:

$$\mathcal{W}_p^p(P', P) = \inf_{\pi \in \mathcal{U}(P',P)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|_p^p d\pi(x, y) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|_p^p d\pi^*(x, y),$$

where π^* is the coupling that attains the infimum. Let us define the function $J(x)$ that assigns a point $x \in q$ to the region it belongs, q: $J(x) = q \iff x \in q$, for all $q \in Q$. Using this definition we get that

$$\begin{aligned}\mathcal{W}_p^p(P', P) &= \inf_{\pi \in \mathcal{U}(P',P)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|_p^p d\pi(x, y) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \|x - y\|_p^p d\pi^*(x, y) \\ &\geq \int_{\mathbb{R}^n \times \mathbb{R}^n} d^p_{J(x),J(y)} d\pi^*(x, y),\end{aligned}$$

since $d^p_{J(x),J(y)} \leq \|x - y\|_p^p$ for all $x, y \in \mathbb{R}^n$. Furthermore, we prove that

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} d^p_{J(x),J(y)} d\pi^*(x, y) \geq \inf_{\pi \in \mathcal{U}(P',P)} \int_{\mathbb{R}^n \times \mathbb{R}^n} d^p_{J(x),J(y)} d\pi(x, y). \tag{B-1}$$

Now, consider a bigger feasible set for the problem on the right of expression (B-1):

$$\begin{aligned}&\mathcal{U}(P',P) := \{\pi \in \mathcal{P}_p(\mathbb{R}^n \times \mathbb{R}^n) : \pi(\mathbb{R}^n, A) = P'(A),\ \pi(B, \mathbb{R}^n) = P(B) \text{ for all } A, B \in \mathcal{B}(\mathbb{R}^n)\}\\ \subset &\mathcal{U}'(P',P) := \{\pi \in \mathcal{P}_p(\mathbb{R}^n \times \mathbb{R}^n) : \pi(\mathbb{R}^n, q_i) = P'(q_i),\ \pi(q_j, \mathbb{R}^n) = P(q_j) \text{ for all } q_i, q_j \in Q\}.\end{aligned}$$

Using this set we get

$$\inf_{\pi \in \mathcal{U}(P',P)} \int_{\mathbb{R}^n \times \mathbb{R}^n} d^p_{J(x),J(y)} d\pi(x,y) \geq \inf_{\pi \in \mathcal{U}'(P',P)} \int_{\mathbb{R}^n \times \mathbb{R}^n} d^p_{J(x),J(y)} d\pi(x,y). \tag{B-2}$$

Given the special shape of the cost function in the problem on the right of (B-2), it suffices to look for discrete couplings of the form $\pi = \pi_{ij}\delta_{ij} \in \mathcal{P}_p(\mathbb{R}^n \times \mathbb{R}^n)$, where δ_{ij} is the Dirac measure concentrated at any points x, y such that $x \in q_i$, $y \in q_j$:

$$\inf_{\pi \in \mathcal{U}'(P',P)} \int_{\mathbb{R}^n \times \mathbb{R}^n} d^p_{J(x),J(y)} d\pi(x,y), \tag{B-3}$$

which is equivalent to the following finite linear program:

$$\begin{aligned}&\inf_{\pi \in \mathbb{R}^{|Q|\times|Q|}_{\geq 0}} && \sum_{q_i \in Q, q_j \in Q} d^p_{ij}\pi_{ij} && && \text{(B-4a)}\\ &s.t. && \sum_{q_j \in Q} \pi_{ij} = P'(q_i) && \forall q_i \in Q && \text{(B-4b)}\\ & && \sum_{q_i \in Q} \pi_{ij} = P(q_j) && \forall q_j \in Q. && \text{(B-4c)}\end{aligned}$$

Note that, in problem (B-4), with an abuse of notation, we refer as π to the discrete coupling: $\pi_{i,j}$ represents the amount of probability mass transported from state $q_i \in Q$ to state $q_j \in Q$. This is the notation that we employ from now on when considering couplings between discrete measures. Finally, notice that LP (B-4) is precisely the definition of the Wasserstein distance $\mathcal{W}_p$ between distributions over Q based on distance d_{ij}. Therefore we conclude that $\mathcal{W}_p(P', P) \geq \mathcal{W}_p(\gamma', \gamma)$, which completes the proof. □

The intuition behind Lemma B-0.1 is the following: if two measures over $\mathbb{R}^n$ lay at a Wasserstein distance distance of at most ε, then their discrete equivalents over Q do not lay at a bigger distance (based on d_{ij}).

Now we are finally able to prove Lemma 5-4.1. Since the disturbance in (5-1) is additive, $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$ implies $P_{x_{t+1}} \in \mathbb{B}_\varepsilon(\widehat{P}_{x_{t+1}})$ for fixed $x \in \mathbb{R}^n$, $a \in A$. This last expression is the same as $\mathcal{W}_p(P_{x_{t+1}}, \widehat{P}_{x_{t+1}}) \leq \varepsilon$. Now, consider the transition probabilities from $x \in q$, $q \in Q$, under $a \in A$ and for probability $\widehat{P}_\xi$ of the disturbance:

$$\widehat{\gamma}_{x,a}(q') := \mathcal{T}_{\mathcal{C}}(q'|x, a; \widehat{P}_\xi)$$

for all $q' \in Q$. By construction of the IMDP $\widehat{\mathcal{I}}$, $\widehat{\gamma}_{x,a} \in \widehat{\Gamma}^a_q$ for all $x \in q$. Then, from Lemma B-0.1, we get that $\mathcal{W}_p(\gamma_{x,a}, \widehat{\gamma}_{x,a}) \leq \varepsilon$ for all $x \in q$, since $\gamma_{x,a}, \widehat{\gamma}_{x,a}$ are the equivalent probabilities over Q of $P_{x_{t+1}}, \widehat{P}_{x_{t+1}}$. Therefore Theorem 5-4.1 follows.

Appendix C

Proof of Theorem 5-4.1

This proof has been inspired by the one presented in [12] and has been adapted to the setting of reachability. This consideration means that we need to take into account different theorems that guarantee convergence of value iteration. We begin the proof by considering the original system and its robust MDP abstraction. Then we define a (deterministic) FSA that captures the language of the specification, given as a scLTL formula. After that we define the products between the original system (and its abstraction) with the FSA. Next we define strategies for all the previous systems. Then we highlight that proving correctness for the case of specifications given as scLTL formulas is equivalent to proving correctness in the simpler case of reachability. Therefore, we limit to this scenario. For this one, we prove that the sequence of value functions of the original system obtained via value iteration is bounded by those obtained by performing robust value iteration on the abstraction. This completes the proof.

Consider the stochastic system (5-1) defined in Section 5-1. We can express this class of systems as a parametric, continuous-state MDP $\mathcal{C}$. To do so, we define as $A_{\mathcal{C}} = U_{\mathcal{C}}$ the set of actions of said MDP, where $U_{\mathcal{C}}$ can be uncountable. Furthermore, we consider kernel $\mathcal{T}_{\mathcal{C}}$ in (4-2) as the transition kernel of said MDP. Using the previous elements, we formally define system (5-1) as the following parametric, continuous-state MDP:

Definition C-0.1. ***(Parametric, Continuous-State MDP)*** *A parametric, continuous-state MDP is a tuple $\mathcal{C} = (\mathbb{R}^n, A_{\mathcal{C}}, \mathcal{T}, O_{\mathcal{C}}, L_{\mathcal{C}})$, where:*

- *$\mathbb{R}^n$, with $n \in \mathbb{N}$, is the (uncountable) set of states,*
- *$A_{\mathcal{C}}$ is a (possibly uncountable) set of actions,*
- *$\mathcal{T} : \mathcal{B}(\mathbb{R}^n) \times \mathbb{R}^n \times A_{\mathcal{C}} \times \mathcal{P}_p(\mathbb{R}^n) \to [0, 1]$ is a stochastic transition kernel such that it assigns to each $x \in \mathbb{R}^n$, $a \in A_{\mathcal{C}}$ and $P_\xi \in \mathcal{P}_p(\mathbb{R}^n)$ a probability measure on the Borel space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$,*
- *$O_{\mathcal{C}}$ is a finite set of atomic propositions or observations,*
- *$L_{\mathcal{C}} : \mathbb{R}^n \to 2^{O_{\mathcal{C}}}$ is a labelling function or observation map that assigns to each state in $\mathbb{R}^n$ a subset of atomic propositions in $O_{\mathcal{C}}$.*

The concept of paths, traces and strategy of $\mathcal{C}$ are the same as those defined in 4-1 and, therefore, we do not sate them again to avoid repetition.

Next, consider the FSA $\mathcal{A} = (Z,\, z_0,\, O_{\mathcal{C}},\, \delta,\, Z_{ac})$ that represents the scLTL formula ϕ. We can now define a stochastic hybrid system as the product $\mathcal{C}_\phi := \mathcal{C} \times \mathcal{A}$, which can be described as another (parametric) MDP over the hybrid state $\mathbb{R}^n \times Z$ [9]. Note that MDP $\mathcal{C}_\phi$ has a set of accepting states, $\mathbb{R}^n \times Z_{ac}$, which correspond to states in the accepting set Z_{ac} of $\mathcal{A}$ [4]. The transition kernel of $\mathcal{C}_\phi$, for fixed P_ξ is equivalent to kernel (4-2) whenever this one leads to a transition that corresponds to an actual transition in $\mathcal{A}$[1]. Next, we obtain a robust MDP abstraction $\mathcal{M}^R$ of the continuous system as we explained in Section 5-2, and we build the product robust MDP $\mathcal{M}^R_\phi := \mathcal{M}^R \times \mathcal{A}$ as we described in Section 5-3. Notice that the robust MDP $\mathcal{M}^R_\phi$ also has a set of accepting states, $Q \times Z_{ac}$, which correspond to states in the accepting set Z_{ac} of $\mathcal{A}$. Using the robust value iteration algorithm described in that section, we obtain the (time-dependent) Markovian strategy σ^*_ϕ of $\mathcal{M}^R_\phi$. Moreover, this proof is valid for any Markovian strategy of $\mathcal{M}^R_\phi$. Furthermore, let us refine said strategy to the strategy $\sigma^*_{\mathcal{C}_\phi}$ of $\mathcal{C}_\phi$ by use making use of the function J defined in Section 4-4:

$$\sigma^*_{\mathcal{C}_\phi}((x,z);k) := \sigma^*_\phi((J(x),z);k)$$

for all $(x,z) \in \mathbb{R}^n \times Z$, $k \in \mathbb{N}_{\geq 0}$. Additionally, let us translate strategy $\sigma^*_{\mathcal{C}_\phi}$ into a memory-dependent strategy $\sigma^*_{\mathcal{C}}$ over the finite paths $w^k_{\mathbb{R}^n}$ of $\mathcal{C}$, by using Lemma 5-3.1. Notice that we stated Lemma 5-3.1 for robust MDPs, but it still holds for both IMDP models, since robust MDPs are a generalization of the latter.

When it comes to satisfying ϕ, it is evident that the paths of $\mathcal{C}$ that, under strategy $\sigma^*_{\mathcal{C}}$, satisfy ϕ are those that correspond to paths in $\mathcal{C}_\phi$ that reach the accepting set $\mathbb{R}^n \times Z_{ac}$ under strategy $\sigma^*_{\mathcal{C}_\phi}$. In the same way, the paths of $\mathcal{M}^R$ that satisfy ϕ are those that correspond to paths in $\mathcal{M}^R_\phi$ that reach $Q \times Z_{ac}$ under strategy σ^*_ϕ. Therefore, proving correctness boils down to proving that the probability of $\mathcal{C}_\phi$ reaching its accepting set is bounded by the probability interval of $\mathcal{M}^R$ reaching this set, both under their respective strategies. Since the exact proof would be very difficult for a reader to understand due to the complex notation, we follow an alternative approach: inspired by the proof of Theorem 2 in [7], we provide a proof in the scenario of reachability for $\mathcal{C}$ and its robust MDP abstraction $\mathcal{M}^R$, forgetting about their products with $\mathcal{A}$. Furthermore, assume that the optimal strategy $\sigma^*_{\mathcal{M}^R}$ has been obtained as described in Section 5-3 as the solution of the maximal reachability probability problem for target set $Q_{\mathtt{tgt}}$. Refine said strategy to the strategy $\sigma^*_{\mathcal{C}}$ of $\mathcal{C}$ as follows:

$$\sigma^*_{\mathcal{C}}(x;k) := \sigma^*_{\mathcal{M}^R}(J(x);k),$$

for all $x \in \mathbb{R}^n$, $k \in \mathbb{N} \cup \infty$. Denote by $X_{\mathtt{tgt}} \subset \mathbb{R}^n$ the region of the state space that corresponds to $Q_{\mathtt{tgt}}$:

$$X_{\mathtt{tgt}} := \bigcup_{q \in Q_{\mathtt{tgt}}} q.$$

Moreover, let us denote by $P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\; w^k_{\mathbb{R}^n}(t) \in X_{\mathtt{tgt}} | X, w^k_{\mathbb{R}^n}(0) = x, \sigma^*_{\mathcal{C}}, P_\xi)$ the probability of $\mathcal{C}$, for distribution P_ξ of ξ, reaching the target set $X_{\mathtt{tgt}}$ within k steps while

[1] This is analogous to how we define the transition probabilities of a product IMDP in (2-3). However, for simplicity, we do not state the definition of this kernel here.

staying in set $X \subseteq \mathbb{R}^n$ under strategy $\sigma^*_{\mathcal{C}}$ and by starting in state $x \in \mathbb{R}^n$. Furthermore, consider the value function $p^k(x)$, for all $x \in \mathbb{R}^n$, $k \geq 0$, recursively defined for an arbitrary $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$ by

$$p^{k+1}(x) = \begin{cases} 1 & \text{if } x \in X_{\texttt{tgt}} \\ 0 & \text{if } x \in \mathbb{R}^n \setminus X \\ \int_{\mathbb{R}^n} p^k(x') \mathcal{T}_{\mathcal{C}}(dx'|x, \sigma^*_{\mathcal{C}}(x;k); P_\xi) & \text{otherwise} \end{cases} \tag{C-1}$$

and with $p^0(x) = 1$ if $x \in Q_{\texttt{tgt}}$ and $p^0(x) = 0$ otherwise. It is evident that, for all $k \geq 0$

$$p^k(x) = P(\exists t \in \mathbb{N}_{\geq 0} \text{ s.t. } t \leq k,\, w^k_{\mathbb{R}^n}(t) \in X_{\texttt{tgt}} | X, w^k_{\mathbb{R}^n}(0) = X, \sigma^*_{\mathcal{C}}, P_\xi)$$

for all $x \in \mathbb{R}^n$ [7]. Additionally, from Lemma 1 in [7] it holds that $\lim_{k\to\infty} p^k(x) = P(\exists k \in \mathbb{N}_{\geq 0} \cup \{\infty\} \text{ s.t. } w_{\mathbb{R}^n}(k) \in X_{\texttt{tgt}} | w^k_{\mathbb{R}^n}(0) = x, \sigma^*_{\mathcal{C}}, P_\xi)$ for all $x \in \mathbb{R}^n$. The latter is the probability of the paths of $\mathcal{C}$, for distribution P_ξ of ξ, ever reaching $X_{\texttt{tgt}}$ while always staying in X by starting from $x \in \mathbb{R}^n$ and by following strategy $\sigma^*_{\mathcal{C}}$. For ease of notation, from now on we refer to a strategy $\sigma^*_{\mathcal{M}^R}$ of $\mathcal{M}^R$ simply as σ, and to strategy $\sigma^*_{\mathcal{C}}$ of $\mathcal{C}$ simply as $\sigma_{\mathcal{C}}$. Let us state the following lemma:

Lemma C-0.1. ***(Bounds in Value Function of the Original System)*** *Consider the sequence of value functions $\{p^k\}_{k\geq 0}$ of the original system $\mathcal{C}$ as defined recursively in* (C-1) *for an arbitrary $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$. Furthermore, consider the sequences $\{\underline{p}^k\}_{k\geq 0}$ and $\{\overline{p}^k\}_{k\geq 0}$ obtained by performing robust interval value iteration on $\mathcal{M}^R$ as described in Section 5-3. Then, $\underline{p}^k(q) \leq p^k(x) \leq \overline{p}^k(q)$ for $x \in q$ and for all $q \in Q$, $k \geq 0$.*

Proof. We will prove Lemma C-0.1 for the case of the lower-bounding sequence $\{\underline{p}^k\}_{k\geq 0}$. The proof of the case of the upper bound follows a similar reasoning. Let us start by considering the robust interval value iteration introduced in Theorem 5-3.1. Now let us obtain the proof for the case that $x \in X$. Consider an arbitrary $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$. We begin by defining the following choice of the adversary at iteration k, for all $q, q' \in Q$:

$$\gamma^*_{q,\sigma(q;k)}(q';k) := \int_{q'} \mathcal{T}_{\mathcal{C}}(dx'|x^{*,k}, \sigma(q;k); P_\xi)$$

where we define $x^{*,k} := \arg\min_{x\in q} p^{k+1}(x)$, for all $k \geq 0$. Notice that, from Lemma 5-4.1, this is a feasible transition probability of $\mathcal{M}^R$, this is, $\gamma^*_{q,\sigma(q;k)}(\cdot;k) \in \widehat{\Gamma}_{q,\sigma(q)} \oplus \varepsilon$ for all $q \in Q$, $k \geq 0$. [2] We prove Lemma C-0.1 by induction: first, we assume that

$$\underline{p}^k(q) \leq \min_{x\in q} p^k(x) \tag{C-2}$$

for all $q \in Q$. Then, we prove that, under assumption (C-2) we obtain

$$\underline{p}^{k+1}(q) \leq \min_{x\in q} p^{k+1}(x) \tag{C-3}$$

for all $q \in Q$. We start from the base case at $k = 0$, where assumption (C-2) holds trivially:

$$\underline{p}^0(q) = \min_{x\in q} p^0(x) = \begin{cases} 1 & \text{if } x \in\ X_{\texttt{tgt}} \\ 0 & \text{otherwise} \end{cases} \tag{C-4}$$

[2] In the case of DR-IMDPs, Lemma 4-4.1 proves that $\gamma^*_{q,\sigma(q;k)}(\cdot;k) \in \Gamma_{q,\sigma(q)}$ for all $q \in Q$, $k \geq 0$, where $\Gamma_{q,\sigma(q)}$ is the set of feasible transition probabilities from q by $\sigma(q)$ of the IMDP.

for all $x \in q$, $q \in Q$. Using the initial condition in (C-4), we notice that Lemma C-0.1 trivially holds for $x \in X_{\texttt{tgt}}$ and $x \in \mathbb{R}^n \setminus X$: $\underline{p}^k(q) = p^k(x) = 1$ for all $x \in q$, $q \in Q_{\texttt{tgt}}$ and $\underline{p}^k(q_u) = p^k(x) = 0$ for all $x \in \mathbb{R}^n \setminus X$, for all $k \geq 0$. Now, let us perform the induction step for $q \in Q \setminus Q_{\texttt{tgt}}$:

$$\begin{aligned}
\underline{p}^{k+1}(q) &= \max_{a\in A} \min_{\gamma_{q,a}\in\widehat{\Gamma}_{q,a}\oplus\varepsilon} \sum_{q'\in Q} \gamma_{q,a}(q')\underline{p}^k(q') \\
&= \min_{\gamma_{q,\sigma(q;k)}\in\widehat{\Gamma}_{q,\sigma(q;k)}\oplus\varepsilon} \sum_{q'\in Q} \gamma_{q,\sigma(q;k)}(q')\underline{p}^k(q') \\
&\leq \sum_{q'\in Q} \gamma^*_{q,\sigma(q;k)}(q';k)\underline{p}^k(q') \\
&= \sum_{q'\in Q} \int_{q'} \mathcal{T}_{\mathcal{C}}(dx'|x^{*,k}, \sigma(q;k), P_\xi)\underline{p}^k(q') \\
&= \sum_{q'\in Q} \int_{q'} \underline{p}^k(q')\mathcal{T}_{\mathcal{C}}(dx'|x^{*,k}, \sigma_{\mathcal{C}}(x^{*,k};k), P_\xi) \\
&\leq \sum_{q'\in Q} \int_{q'} (\min_{x'\in q'} p^k(x'))\mathcal{T}_{\mathcal{C}}(dx'|x^{*,k}, \sigma_{\mathcal{C}}(x^{*,k};k), P_\xi) \\
&\leq \sum_{q'\in Q} \int_{q'} p^k(x')\mathcal{T}_{\mathcal{C}}(dx'|x^{*,k}, \sigma_{\mathcal{C}}(x^{*,k};k), P_\xi).
\end{aligned} \tag{C-5}$$

Now, let us express the last case in Equation (C-1) as:

$$p^{k+1}(x) = \int_{\mathbb{R}^n} p^k(x')\mathcal{T}_{\mathcal{C}}(dx'|x, \sigma_{\mathcal{C}}(x;k), P_\xi) = \sum_{q'\in Q} \int_{q'} p^k(x')\mathcal{T}_{\mathcal{C}}(dx'|x, \sigma_{\mathcal{C}}(x;k), P_\xi). \tag{C-6}$$

Finally, comparing the right-most expression in (C-6) with the last one in (C-5) and taking into account the definition of $x^{*,k}$, we get that

$$\sum_{q'\in Q} \int_{q'} p^k(x')\mathcal{T}_{\mathcal{C}}(dx'|x^{*,k}, \sigma_{\mathcal{C}}(x^{*,k};k), P_\xi) \leq p^{k+1}(x),$$

which completes the proof. □

Furthermore, since we have proved Lemma C-0.1 for an arbitrary distribution $P_\xi \in \mathbb{B}_\varepsilon(\widehat{P}_\xi)$, Theorem 5-4.1 follows. Notice also that Lemma C-0.1 implies that value function $\underline{p}^k(q)$ is always a lower bound of $p^k(x)$, for $x \in q$, for all $q \in Q$, $k \geq 0$. We must also highlight that, while in this appendix we have provided a proof for Theorem 5-4.1, the proof of Theorem 4-4.1 follows the same reasoning: it suffices to consider the set $\Gamma_{q,a}$ of transition probabilities of the DR-IMDP, as defined in expression (2-2) instead of $\widehat{\Gamma}_{q,a} \oplus \varepsilon$, for all $q \in Q$, $a \in A$.

Bibliography

[1] C. K. Verginis, Z. Xu, and D. V. Dimarogonas, "Decentralized motion planning with collision avoidance for a team of uavs under high level goals," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 781–787, IEEE, 2017.

[2] X. Ji and Y. Niu, "Robust strategy planning for uav with ltl specifications," in *2016 35th Chinese Control Conference (CCC)*, pp. 2890–2895, IEEE, 2016.

[3] M. Lahijanian, S. B. Andersson, and C. Belta, "Formal verification and synthesis for discrete-time stochastic systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 8, pp. 2031–2045, 2015.

[4] N. Cauchi, L. Laurenti, M. Lahijanian, A. Abate, M. Kwiatkowska, and L. Cardelli, "Efficiency through uncertainty: Scalable formal synthesis for stochastic hybrid systems," in *Proceedings of the 22nd ACM International Conference on Hybrid Systems: Computation and Control*, pp. 240–251, 2019.

[5] S. Adams, M. Lahijanian, and L. Laurenti, "Formal control synthesis for stochastic neural network dynamic models," *IEEE Control Systems Letters*, 2022.

[6] C. Belta, B. Yordanov, and E. A. Gol, *Formal methods for discrete-time dynamical systems*, vol. 15. Springer, 2017.

[7] J. Jackson, L. Laurenti, E. Frew, and M. Lahijanian, "Formal verification of unknown dynamical systems via gaussian process regression," *arXiv preprint arXiv:2201.00655*, 2021.

[8] M. Lahijanian, S. B. Andersson, and C. Belta, "Temporal logic motion planning and control with probabilistic satisfaction guarantees," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 396–409, 2011.

[9] A. Abate, J.-P. Katoen, J. Lygeros, and M. Prandini, "Approximate model checking of stochastic hybrid systems," *European Journal of Control*, vol. 16, no. 6, pp. 624–641, 2010.

[10] M. Lahijanian, S. B. Andersson, and C. Belta, “Approximate markovian abstractions for linear stochastic systems,” in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 5966–5971, IEEE, 2012.

[11] J. Jackson, L. Laurenti, E. Frew, and M. Lahijanian, “Strategy synthesis for partially-known switched stochastic systems,” in *Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control*, pp. 1–11, 2021.

[12] G. Delimpaltadakis, L. Laurenti, and M. Mazo Jr, “Abstracting the sampling behaviour of stochastic linear periodic event-triggered control systems,” *arXiv preprint arXiv:2103.13839*, 2021.

[13] N. Fournier and A. Guillin, “On the rate of convergence in wasserstein distance of the empirical measure,” *Probability Theory and Related Fields*, vol. 162, no. 3, pp. 707–738, 2015.

[14] P. M. Esfahani and D. Kuhn, “Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations,” *Mathematical Programming*, vol. 171, no. 1, pp. 115–166, 2018.

[15] R. Gao and A. J. Kleywegt, “Distributionally robust stochastic optimization with wasserstein distance,” *arXiv preprint arXiv:1604.02199*, 2016.

[16] J. Dahl, G. R. de Campos, C. Olsson, and J. Fredriksson, “Collision avoidance: A literature review on threat-assessment techniques,” *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 1, pp. 101–113, 2018.

[17] G. E. Fainekos, A. Girard, H. Kress-Gazit, and G. J. Pappas, “Temporal logic motion planning for dynamic robots,” *Automatica*, vol. 45, no. 2, pp. 343–352, 2009.

[18] C. Baier and J.-P. Katoen, *Principles of model checking.* MIT press, 2008.

[19] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons, 2014.

[20] J. Jackson, L. Laurenti, E. Frew, and M. Lahijanian, “Safety verification of unknown dynamical systems via gaussian process regression,” in *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 860–866, IEEE, 2020.

[21] R. Givan, S. Leach, and T. Dean, “Bounded-parameter markov decision processes,” *Artificial Intelligence*, vol. 122, no. 1-2, pp. 71–109, 2000.

[22] D. Wu and X. Koutsoukos, “Probabilistic verification of uncertain systems using bounded-parameter markov decision processes,” in *International Conference on Modeling Decisions for Artificial Intelligence*, pp. 283–294, Springer, 2006.

[23] H. Zhang, T.-W. Weng, P.-Y. Chen, C.-J. Hsieh, and L. Daniel, “Efficient neural network robustness certification with general activation functions,” *Advances in neural information processing systems*, vol. 31, 2018.

[24] K. Xu, Z. Shi, H. Zhang, Y. Wang, K.-W. Chang, M. Huang, B. Kailkhura, X. Lin, and C.-J. Hsieh, "Automatic perturbation analysis for scalable certified robustness and beyond," *Advances in Neural Information Processing Systems*, vol. 33, pp. 1129–1141, 2020.

[25] T. S. Badings, A. Abate, N. Jansen, D. Parker, H. A. Poonawala, and M. Stoelinga, "Sampling-based robust control of autonomous systems with non-gaussian noise," *arXiv preprint arXiv:2110.12662*, 2021.

[26] A. Puggelli, W. Li, A. L. Sangiovanni-Vincentelli, and S. A. Seshia, "Polynomial-time verification of pctl properties of mdps with convex uncertainties," in *International Conference on Computer Aided Verification*, pp. 527–542, Springer, 2013.

[27] A. Nilim and L. El Ghaoui, "Robust control of markov decision processes with uncertain transition matrices," *Operations Research*, vol. 53, no. 5, pp. 780–798, 2005.

[28] E. M. Wolff, U. Topcu, and R. M. Murray, "Robust control of uncertain markov decision processes with temporal logic specifications," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 3372–3379, IEEE, 2012.

[29] I. Yang, "A convex optimization approach to distributionally robust markov decision processes with wasserstein distance," *IEEE control systems letters*, vol. 1, no. 1, pp. 164–169, 2017.

[30] D. Wu and X. Koutsoukos, "Reachability analysis of uncertain systems using bounded-parameter markov decision processes," *Artificial Intelligence*, vol. 172, no. 8-9, pp. 945–954, 2008.

[31] G. Peyré, M. Cuturi, *et al.*, "Computational optimal transport: With applications to data science," *Foundations and Trends® in Machine Learning*, vol. 11, no. 5-6, pp. 355–607, 2019.

[32] L. Kantorovich, "On the transfer of masses (in russian)," in *Doklady Akademii Nauk*, vol. 37, pp. 227–229, 1942.

[33] S. Garatti and M. Campi, "Risk and complexity in scenario optimization," *Mathematical Programming*, pp. 1–37, 2019.

[34] I. Yang, "A dynamic game approach to distributionally robust safety specifications for stochastic systems," *Automatica*, vol. 94, pp. 94–101, 2018.

[35] A. Hakobyan and I. Yang, "Wasserstein distributionally robust motion planning and control with safety constraints using conditional value-at-risk," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 490–496, IEEE, 2020.

[36] A. Hakobyan and I. Yang, "Distributionally robust risk map for learning-based motion planning and control: A semidefinite programming approach," *arXiv preprint arXiv:2105.00657*, 2021.

[37] A. Abate, F. Redig, and I. Tkachev, "On the effect of perturbation of conditional probabilities in total variation," *Statistics & Probability Letters*, vol. 88, pp. 1–8, 2014.

[38] D. Boskos, J. Cortés, and S. Martínez, "Data-driven ambiguity sets with probabilistic guarantees for dynamic processes," *IEEE Transactions on Automatic Control*, vol. 66, no. 7, pp. 2991–3006, 2020.

[39] O. Hussien and P. Tabuada, "Lazy controller synthesis using three-valued abstractions for safety and reachability specifications," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 3567–3572, IEEE, 2018.

Glossary

List of Acronyms