

# Scenario-Based Verification of Uncertain MDPs

★

Murat Cubuktepe<sup>1</sup>, Nils Jansen<sup>2</sup>, Sebastian Junges<sup>3</sup>,  
Joost-Pieter Katoen<sup>3</sup>, Ufuk Topcu<sup>1</sup>

<sup>1</sup> The University of Texas at Austin, Austin, USA

<sup>2</sup> Radboud University Nijmegen, Nijmegen, The Netherlands

<sup>3</sup> RWTH Aachen University, Aachen, Germany

**Abstract.** We consider Markov decision processes (MDPs) in which the transition probabilities and rewards belong to an uncertainty set parametrized by a collection of random variables. The probability distributions for these random parameters are unknown. The problem is to compute the probability to satisfy a temporal logic specification within any MDP that corresponds to a sample from these unknown distributions. In general, this problem is undecidable, and we resort to techniques from so-called scenario optimization. Based on a finite number of samples of the uncertain parameters, each of which induces an MDP, the proposed method estimates the probability of satisfying the specification by solving a finite-dimensional convex optimization problem. The number of samples required to obtain a high confidence on this estimate is independent from the number of states and the number of random parameters. Experiments on a large set of benchmarks show that a few thousand samples suffice to obtain high-quality confidence bounds with a high probability.

**Keywords:** MDP, Uncertainty, Verification, Scenario optimisation

## 1 Introduction

*MDPs.* Markov decision processes (MDPs) model sequential decision-making problems in stochastic dynamic environments [51]. They are widely used in areas like planning [52], reinforcement learning [53], formal verification [48], and robotics [24]. Mature model checking tools like PRISM [21] and Storm [35] employ efficient algorithms to verify the correctness of MDPs against temporal logic specifications [2] provided all transition probabilities and cost functions are exactly known. In many applications, however, this assumption may be unrealistic, as certain system parameters are typically not exactly known and under control by external sources.

---

\* Supported by the grants ARL # ACC-APG-RTP W911NF, NASA # 80NSSC19K0209, NSF # 1646522, and NSF # 1652113.

*Uncertain MDPs.* A common approach to deal with unknown system parameters is to let transition probabilities and cost functions of an MDP belong to uncertainty sets, resulting in so-called *uncertain* MDPs [25,14,28], which generalize *interval* MDPs [20,27,7]. However, solution approaches, e.g., in [25,14,28], usually rely on the potentially limiting assumption that the uncertainty sets at different states of the MDP are independent from each other.

Consider a simple motion planning scenario where an unmanned aerial vehicle (UAV) is tasked to transport a certain payload to a target location. The problem is to compute a policy for the UAV to successfully deliver the payload while taking into account the weather conditions. External factors like wind strength or direction may affect the movement of the UAV. The assumption that such weather conditions are independent between the different possible states of UAV is unrealistic, and does not adequately model the scenario at hand.

For settings in which the uncertainties at different states depend on each other, an option is to account for all possible—albeit infinitely many—values in the uncertainty sets. The policy synthesis problem can be formulated as a so-called semi-infinite convex optimization problem, which includes finitely many variables but infinitely many constraints [28]. This problem, however, is NP-hard [28,18]. Furthermore, it fails to exploit additional information that may be available as random variables over the uncertainty sets [36], and may be very conservative. For instance, weather-data in the form of probability distributions may provide additional information on potential changes during the mission.

In this paper, we study a setting in which the fact the uncertain parameters are random variables and the dependencies between them are accounted for explicitly. Furthermore, each random parameter follows an unknown probability distribution from which we can sample the parameter values.

*Problem statement.* Compute the probability with which there exists a policy such that a reachability or an expected-cost specification is satisfied for any randomly drawn parameter value.

We call this probability the *satisfaction probability*. The intuition is that the question of whether all (or some) parameter values satisfy a specification—as is often done in parameter synthesis [46]—is replaced by the question of *how much we expect the (sampled) model to satisfy a specification*. For example, a satisfaction probability of 80% tells that, if we randomly sample the parameters, with a probability of 80% there exists a policy for the resulting MDP that satisfies the specification. Computing the satisfaction probability is in general undecidable, even for known probability distributions over the parameter values [37].

*Scenario-based verification.* Therefore, we resort to *sampling-based* algorithms that yield a confidence (probability) on the bounds of the satisfaction probability. Referring back to the UAV example, we want to compute a confidence probability in the probability that there exists a policy for the UAV to successfully finish the mission. As a first step, we take the aforementioned semi-infinite optimization problem that accounts for all possible parameter values as a basis. Each concrete

parameter value is referred to as a *scenario* in the convex optimization literature [15]. For specific problems where a distribution over individual scenarios is present, a technique called *scenario-based optimization* provides guarantees on the satisfaction probability via efficient sampling techniques [15,16]. The basic idea is to consider a finite set of samples from the distribution over the scenarios and restrict the semi-infinite problem to these samples. The resulting convex optimization problem with finitely many constraints can be solved efficiently [50].

For our setting, we first sample a finite number of parameter instantiations each of which induces a concrete MDP. We can solve the synthesis problem for this MDP efficiently using, e.g., a probabilistic model checker. Based on the results, we compute a satisfaction probability and an estimate of its potential error. For example, a 90% estimate in a satisfaction probability of 80%, means that the error is at most 10%. We show that the error in the estimate diminishes to zero exponentially rapidly with increasing number of samples. Moreover, we show that the number of required samples does neither depend on the size of the state space nor the number of random parameters. We validate the theoretical results using several MDPs that have different sizes of state and parameter spaces and demonstrate experimentally that the required number of samples is indeed not sensitive to the dimension of the state and parameter space. In addition, we show the effectiveness of our method with a new dedicated case study based on the aforementioned UAV example which incorporates 2 500 random parameters.

*Related work.* The so-called *parameter synthesis* problem is concerned with computing parameter values such that there exists a policy in the induced non-parametric MDP that satisfies the specifications. Most of the work in parameter synthesis focus on finding one parameter value that satisfies the specification. The approaches involve computing a rational function of the reachability probabilities [11,17,41], utilizing convex optimization [34,40], and sampling-based methods [26,29]. The problem of whether there exists a value in the parameter space that satisfies a reachability specification is ETR-complete<sup>4</sup> [47], and finding a satisfying parameter value is exponential in the number of parameters.

The work in [45] considers the analysis of Markov models in the presence of uncertain rewards, utilizing statistical methods to reason about the probability of a parametric MDP satisfying an expected cost specification. This approach is restricted to reward parameters and does not explicitly compute confidence bounds. [43] computes bounds on the long-run probability of satisfying a specification with probabilistic uncertainty for Markov chains. Other related techniques include multi-objective model checking to maximize the average performance with probabilistic uncertainty sets [36], sampling-based methods which minimize the *regret* with uncertainty sets [33], and Bayesian reasoning to compute parameter values that satisfy a metric temporal logic specification on a continuous-time Markov chain [38]. [37] considers a variant of the problem in this paper where

---

<sup>4</sup> The ETR satisfiability problem is to decide if there exists a satisfying assignment to the real variables in a Boolean combination of a set of polynomial inequalities. It is known that  $\text{NP} \subseteq \text{ETR} \subseteq \text{PSPACE}$ .

the probability distribution of the uncertainty sets is assumed to be known. The paper formulates the policy synthesis problem as an (undecidable [30]) partially observable Markov decision process (POMDP) synthesis problem and use off-the-shelf point-based POMDP methods [10,6]. The work in [27,25] consider the verification of MDPs with convex uncertainties. However, the uncertainty sets for different states in an MDP are restricted to be independent, which does not hold in our problem setting where we have parameter dependencies.

Uncertainties in MDPs have received quite some attention in the artificial intelligence and planning literatures. Interval MDPs [27,7] use probability intervals in the transition probabilities. Dynamic programming, robust value iteration and robust policy iteration have been developed for MDPs with uncertain transition probabilities whose parameters are statistically independent, also referred to as rectangular, to find a policy ensuring the highest expected total reward at a given confidence level [14,25]. The work in [28] relaxes this independence assumption a bit and determines a policy that satisfies a given performance with a pre-defined confidence provided an observation history of the MDP is given by using conic programming. State-of-the art exact methods can handle models of up to a few hundred of states [42]. Multi-model MDPs [44] treat distributions over probability and cost parameters and aim at finding a single policy maximizing a weighted value function. For deterministic policies this problem is NP-hard, and it is PSPACE-hard for history-dependent policies.

## 2 Preliminaries

A *probability distribution* over a finite set  $X$  is a function  $\mu: X \rightarrow [0, 1] \subseteq \mathbb{R}$  with  $\sum_{x \in X} \mu(x) = 1$ . The set of all distributions on  $X$  is denoted by  $Distr(X)$ . Let  $V = \{x_1, \dots, x_n\}$  be a finite set of *parameters* over  $\mathbb{R}^n$ . The set of polynomials over  $V$  is denoted by  $\mathbb{Q}[V]$ . We denote the cardinality of a set  $\mathcal{U}$  by  $|\mathcal{U}|$ .

### 2.1 Parametric Models

**Definition 1 (pMDP).** A parametric Markov decision process (*pMDP*)  $\mathcal{M}$  is a tuple  $\mathcal{M} = (S, Act, s_I, V, \mathcal{P})$  with a finite set  $S$  of states, a finite set  $Act$  of actions, an initial state  $s_I \in S$ , a finite set  $V$  of real-valued variables (parameters) and a transition function  $\mathcal{P}: S \times Act \times S \rightarrow \mathbb{Q}[V]$ .

For  $s \in S$ ,  $ActS(s) = \{\alpha \in Act \mid \exists s' \in S, \mathcal{P}(s, \alpha, s') \neq 0\}$  is the set of *enabled* actions at  $s$ . Without loss of generality, we require  $ActS(s) \neq \emptyset$  for  $s \in S$ . If  $|ActS(s)| = 1$  for all  $s \in S$ ,  $\mathcal{M}$  is a *parametric discrete-time Markov chain (pMC)*. We denote the transition function for pMCs by  $\mathcal{P}(s, s')$ .

A pMDP  $\mathcal{M}$  is a *Markov decision process (MDP)* if the transition function yields *well-defined* probability distributions, i.e.,  $\mathcal{P}: S \times Act \times S \rightarrow [0, 1]$  and  $\sum_{s' \in S} \mathcal{P}(s, \alpha, s') = 1$  for all  $s \in S$  and  $\alpha \in ActS(s)$ . We denote the *parameter space* of  $\mathcal{M}$  by  $\mathcal{V}_{\mathcal{M}}$ . Applying an *instantiation*  $u \in \mathcal{V}_{\mathcal{M}}$  to a pMDP  $\mathcal{M}$  yields the *instantiated* MDP  $\mathcal{M}[u]$  by replacing each  $f \in \mathbb{Q}[V]$  in  $\mathcal{M}$  by  $f[u]$ . An

instantiation  $u$  is *well-defined* for  $\mathcal{M}$  if the resulting model  $\mathcal{M}[u]$  is an MDP. We assume that all parameter instantiations in  $\mathcal{V}_{\mathcal{M}}$  yield well-defined MDPs. We call  $u$  *graph-preserving* if for all  $s, s' \in S$  and  $\alpha \in Act$  it holds that  $\mathcal{P}(s, \alpha, s') \neq 0 \Rightarrow \mathcal{P}(s, \alpha, s')[u] \in (0, 1]$ . If  $\mathcal{P}(s, \alpha, s') \in \{p, 1 - p \mid p \in V\} \cup \mathbb{Q}$ , then the parameter space  $\mathcal{V}_{\mathcal{M}}$  is given by the rectangle  $[0, 1]^{|V|}$ . We also consider a state-action cost function  $c: S \times Act \rightarrow \mathbb{Q}[V]$ . We denote the set of cost parameters as  $\mathcal{W}$ .

To define measures on MDPs, nondeterministic choices are resolved by a so-called *policy*  $\sigma: S \rightarrow Act$  with  $\sigma(s) \in ActS(s)$ . The set of all policies over  $\mathcal{M}$  is  $Str^{\mathcal{M}}$ . For the specifications that we consider in this paper, memoryless deterministic policies are sufficient [48]. Applying a policy to an MDP yields an *induced Markov chain* where all nondeterminism is resolved.

For an MC  $\mathcal{D}$ , the *reachability specification*  $\varphi_r = \mathbf{P}_{\leq \lambda}(\diamond T)$  asserts that a set  $T \subseteq S$  of *target states* is reached with probability at most  $\lambda \in [0, 1]$ <sup>5</sup>. If  $\varphi_r$  holds for  $\mathcal{D}$ , we write  $\mathcal{D} \models \varphi_r$ . Model checking for the more general PCTL [4] or  $\omega$ -regular specifications is often reducible to checking reachability specifications [48]. For an MDP  $\mathcal{M}$ ,  $\varphi_r$  holds if for all  $\sigma \in Str^{\mathcal{M}}$  such that the induced MC  $\mathcal{D}$  by the policy  $\sigma$  reaches the set  $T$  with a probability of at most  $\lambda$ . For an *expected cost specification*  $\varphi_c = \mathbf{EC}_{\leq \kappa}(\diamond G)$ , it holds that  $\mathcal{D} \models \varphi_c$  if and only if the expected cost of reaching a set  $G \subseteq S$  is at most  $\kappa \in \mathbb{R}$ . The expected cost of reaching  $G$  is well-defined if and only if  $\mathbf{P}(\diamond T) = 1$  for all policies in an MDP.

## 2.2 Uncertain MDPs

We now introduce the setting that we study in this paper. Specifically, we use parameters to define the uncertainty in the transition probabilities and cost functions of an MDP. Each random parameter follows an unknown probability distribution from which we can sample the parameter values.

**Definition 2 (uMDP).** *An uncertain Markov decision process  $\mathcal{M}_{\mathbb{P}}$  (uMDP) is a tuple  $\mathcal{M}_{\mathbb{P}} = (\mathcal{M}, \mathbb{P})$  where  $\mathcal{M}$  is a pMDP, and  $\mathbb{P}$  is a probability distribution over the parameter space  $\mathcal{V}_{\mathcal{M}}$ . If  $\mathcal{M}$  is a pMC, then we call  $\mathcal{M}_{\mathbb{P}}$  a uMC.*

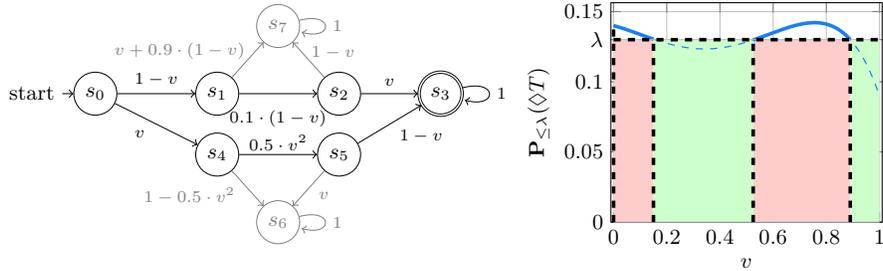
Intuitively, a uMDP is a pMDP with an associated distribution over possible (graph-preserving) parameter instantiations. That is, a realization of  $\mathbb{P}$  yields a concrete MDP  $\mathcal{M}[u]$  with the respective instantiation  $u \in \mathcal{V}_{\mathcal{M}}$  (and  $\mathbb{P}(u) > 0$ ).

*Remark 1.* In a uMDP, we distinguish *controllable* and *uncontrollable* parameters. The uncontrollable parameters follow the probability distribution  $\mathbb{P}$ . In contrast, we can actively *instantiate* the controllable parameters. In the following, we specifically allow cost parameters to be controllable.

**Definition 3 (Satisfaction Probability).** *Let  $\mathcal{M}_{\mathbb{P}} = (\mathcal{M}, \mathbb{P})$  be a uMDP and  $\varphi$  a specification. The (weighted) satisfaction probability of  $\varphi$  is*

$$F(\mathcal{M}_{\mathbb{P}}, \varphi) = \int_{\mathcal{V}_{\mathcal{M}}} I_{\varphi}(u) d\mathbb{P}(u)$$

<sup>5</sup> The theory also applies to lower bounded properties.



**Fig. 1.** Left: A uMC with parameter  $v$ . Right: The probability of satisfying the reachability specification  $\varphi_r = \mathbf{P}_{\leq \lambda}(\diamond T)$  versus the value of the parameter  $v$ . Intervals that satisfy  $\varphi_r$  are green, intervals that violate  $\varphi_r$  are red.

with  $u \in \mathcal{V}_{\mathcal{M}}$  and  $I_{\varphi}: \mathcal{V}_{\mathcal{M}} \rightarrow \{0, 1\}$  is the indicator for  $\varphi$ , i.e.  $I_{\varphi}(u) = 1$  iff  $\mathcal{M}[u] \models \varphi$ .

Note that  $I_{\varphi}$  is measurable, as  $\mathcal{V}_{\mathcal{M}}$  is the finite union of semi-algebraic sets [49]. Moreover, we have that  $F(\mathcal{M}_{\mathbb{P}}, \varphi) \in [0, 1]$  and  $F(\mathcal{M}_{\mathbb{P}}, \varphi) + F(\mathcal{M}_{\mathbb{P}}, \neg \varphi) = 1$ .

*Example 1.* Consider the uMC in the left figure of Fig. 1 with the uncontrollable parameter set  $V = \{v\}$ , initial state  $s_0$ , target set  $T = \{s_3\}$  and an uniform distribution for the parameter  $v$  over the interval  $[0, 1]$ . We plot the probability of satisfying the specification  $\varphi_r = \mathbf{P}_{\leq \lambda}(\diamond T)$  as a function of  $v$  in the right figure of Fig. 1. We also show the satisfying region and its complementary as green and red regions. The satisfying region is given by the union of the intervals  $[0.13, 0.525]$  and  $[0.89, 1.0]$ , and the satisfaction probability  $F(\mathcal{M}_{\mathbb{P}}, \varphi_r)$  is  $0.395 + 0.11 = 0.505$ .

### 3 Problem Statement

In this section, we state the problem that we study in this paper. We seek to compute the satisfaction probability of the parameter space for a reachability or an expected cost specification  $\varphi$  on a uMDP. Intuitively, we seek the probability that a randomly sampled instantiation from the parameter space induces an MDP which satisfies  $\varphi$ . Formally: Given a uMDP  $\mathcal{M}_{\mathbb{P}} = (\mathcal{M}, \mathbb{P})$ , and a specification  $\varphi$ , compute the satisfaction probability  $F(\mathcal{M}_{\mathbb{P}}, \varphi)$ . However, as mentioned, the problem is in general undecidable [37]. Therefore, we consider an approximation of computing the satisfaction probability:

*Problem 1.* Given a uMDP  $\mathcal{M}_{\mathbb{P}} = (\mathcal{M}, \mathbb{P})$ , a reachability specification  $\varphi_r = \mathbf{P}_{\leq \lambda}(\diamond T)$ , and a *tolerance probability*  $\nu$ , compute a confidence probability  $\alpha_{\nu}$  such that  $F(\mathcal{M}_{\mathbb{P}}, \varphi_r) \geq 1 - \nu$  holds with a probability of at least  $1 - \alpha_{\nu}$ .

We illustrate the problem statement with the following example.

*Example 2.* For the UAV motion planning example, consider the question “What is the probability on a given day such that there exists a policy for the UAV to successfully finish the mission.” A possible result is, e.g., 0.78 (confidence probability: 0.99) and 0.81 (confidence probability: 0.95). Then, with a confidence probability of 0.99, the actual satisfaction probability is indeed greater than 0.78, and with a (slightly lower) confidence probability of 0.95 it is greater than 0.81. Such a result shows that it is quite likely that the UAV will finish the mission successfully with a probability that is at least 81%.

Similar to Problem 1, we also consider expected cost specifications.

*Problem 2.* Given a uMDP  $\mathcal{M}_{\mathbb{P}} = (\mathcal{M}, \mathbb{P})$ , and an expected cost specification  $\varphi_c = \text{EC}_{\leq \kappa}(\diamond G)$ , a tolerance probability  $\nu$ , and a confidence probability  $\alpha_\nu$  determine if there exists an instantiation to the cost parameters such that  $F(\mathcal{M}_{\mathbb{P}}, \varphi_c) \geq 1 - \nu$  holds with a probability of at least  $1 - \alpha_\nu$ .

*Remark 2.* The main difference between Problem 1 and Problem 2 is that we consider *controllable* cost parameters. We seek to compute an instantiation to these parameters such that the satisfaction probability is greater than  $1 - \nu$  with high confidence.

## 4 Scenario-Based Verification

In this section, we present our approach to solving Problem 1 and 2, that is, to approximate the satisfaction probability with respect to a specification. We first consider the robust policy synthesis problem that accounts for *all possible values* in the uncertainty set, potentially leading to a very pessimistic result. This problem can be formulated as a semi-infinite convex optimization problem, which is NP-hard [28]. Here, we exploit the structure of this problem, which includes finitely many variables but infinitely many constraints. Our approach is based on *scenario optimization* [15,16]: We sample a finite number of parameter values and restrict the semi-infinite problem to these samples. The resulting *finite-dimensional* convex optimization problem can be solved efficiently [50]. Based on the solution of the optimization problem, we compute high confidence in the estimate of the satisfaction probability. The estimate also generalizes to the samples from the probability distribution that are not in the sample set.

*Remark 3.* For ease of presentation, we focus on uncertain Markov chains (uMCs). Our results and methods generalize to uncertain MDPs (uMDPs).

We first develop the main results for the simple setting where *all sampled* instantiated MCs from the parameter space  $\mathcal{V}_{\mathcal{D}}$  satisfy the reachability specification  $\varphi_r$ . This assumption does not imply that *all* instantiated MCs satisfy  $\varphi_r$ : The sample set does not contain an MC that violates  $\varphi_r$  even though there exists such an MC in the parameter space. In Section 4.2, we drop this assumption and allow sampled points in  $\mathcal{V}_{\mathcal{D}}$  to violate  $\varphi_r$ . This completes our treatment of Problem 1. In Section 4.3, we show how our results generalize to expected cost specifications  $\varphi_c$ , to solve Problem 2.

## 4.1 Restriction to Satisfying Samples

In this section, we assume that all instantiated MCs satisfy  $\varphi_r$ . We then generalize our method to any values of  $\nu$ . We want to check if a uMC  $\mathcal{D}$  satisfies a reachability specification  $\varphi_r = \mathbf{P}_{\leq \lambda}(\diamond T)$  for all instantiations in the sample set  $\mathcal{U}$ . For each instantiation, we can formulate a linear program (LP) that is feasible if and only if  $\varphi_r$  is satisfied [51]. For a subset  $\mathcal{U} \subseteq \mathcal{V}_{\mathcal{D}}$  of the parameter space  $\mathcal{V}_{\mathcal{D}}$  of the uMC  $\mathcal{D}$ , we can then write the conjunction of these LPs. We assume that  $|\mathcal{U}|$  is finite and sampled from the probability distribution  $\mathbb{P}$  over the parameter space  $\mathcal{V}_{\mathcal{D}}$ .

For each instantiation  $u \in \mathcal{U}$ , we introduce a set of linear constraints that are parametrized by  $u$ <sup>6</sup>. We use the following variables. For  $s \in S$  and  $u \in \mathcal{U}$ , the variable  $p_s^u \in [0, 1]$  represents the probability of reaching the target set  $T \subseteq S$  from state  $s$ . The variable  $\tau$  represents an upper bound on the probability of satisfying  $\varphi_r$  for all instantiations in  $\mathcal{U}$ . Note that  $\tau$  is a variable in our formulation, whereas  $\lambda$  is the threshold of the reachability specification, and thus constant. The set  $\neg\exists\diamond T$  represents the set of states which cannot reach the target set  $T$ . The probability of reaching  $T$  from these states is zero, and the set  $\neg\exists\diamond T$  does not change for different graph-preserving instantiations [17]. The set  $\neg\exists\diamond T$  can be found in polynomial time in the size of a uMC by using standard graph-based search algorithms [48]. We solve the following LP  $\mathcal{L}_r(\mathcal{U})$ , which is parametrized by each instantiation  $u$  in  $\mathcal{U}$ ,

$$\text{minimize } \tau \tag{1}$$

$$\text{subject to } \forall u \in \mathcal{U},$$

$$p_{s_I}^u \leq \tau, \tag{2}$$

$$p_{s_I}^u \leq \lambda, \tag{3}$$

$$p_s^u = 1, \quad \forall s \in T, \tag{4}$$

$$p_s^u = 0, \quad \forall s \in \neg\exists\diamond T, \tag{5}$$

$$p_s^u = \sum_{s' \in S} \mathcal{P}(s, s')[u] \cdot p_{s'}^u, \quad \forall s \in S \setminus (T \cup \neg\exists\diamond T). \tag{6}$$

The objective (1) minimizes the maximal probability that can be achieved by all MCs induced by  $\mathcal{U}$ . The constraint (2) represents an upper bound on the reachability probability for all instantiations. We minimize the upper bound to compute the maximal probability of satisfying  $\varphi_r$  for all instantiated MCs. The constraint (3) ensures that the probability of reaching  $T$  from the initial state  $s_I$  is below the threshold  $\lambda$ . The constraint (4) sets the probability to reach a state in  $T$  from  $T$  to 1. The constraint (5) sets the reachability probabilities from the states in  $\neg\exists\diamond T$  to zero. The constraint (6) computes the probability of satisfying the specification for each non-target state  $s \in S$  in the standard way.

There are infinitely many constraints in the semi-infinite LP  $\mathcal{L}_r(\mathcal{V}_{\mathcal{D}})$  as the cardinality of  $(\mathcal{V}_{\mathcal{D}})$  is infinite and  $\mathcal{L}_r(\mathcal{V}_{\mathcal{D}})$  has infinitely many constraints in the

<sup>6</sup> we assume that each sample has a unique index

form of (2)–(6). Our approach is based on *scenario optimization* [13,15,16], where we instantiate the parameters  $u \in \mathcal{V}_{\mathcal{D}}$  by sampling the probability distribution  $\mathbb{P}$ . Then, for a given violation probability  $\nu \in (0, 1)$ , we compute a solution that violates the constraints in the LP  $\mathcal{L}_r(\mathcal{V}_{\mathcal{D}})$  with a probability that is not larger than  $\nu$ . We first give some properties of the LP  $\mathcal{L}_r(\mathcal{U})$ .

**Theorem 1.** *Let uMC  $\mathcal{D}$  and the sample sets  $\mathcal{U} \subseteq \mathcal{V}_{\mathcal{D}}$  with  $K = |\mathcal{U}| \geq 2$ . Assume for all  $u \in \mathcal{U}$ ,  $\mathcal{D}[u] \models \varphi_r$ . For a given tolerance probability  $\nu \in [0, 1)$ , let the associated confidence probability*

$$\alpha_\nu = \sum_{i=0}^1 \binom{K}{i} (1-\nu)^{K-i} \nu^i. \quad (7)$$

Then, with a probability of at least  $1 - \alpha_\nu$ , we have

$$F(\mathcal{D}_{\mathbb{P}}, \varphi_r) \geq 1 - \nu. \quad (8)$$

*Proof.* The key idea of the proof is to relate the finite LP  $\mathcal{L}_r(\mathcal{U})$  induced by a sampled set  $\mathcal{U}$  to the semi-infinite LP  $\mathcal{L}_r(\mathcal{V}_{\mathcal{D}})$ . Then, we use the results given in [16, Theorem 1] to obtain the lower bound  $1 - \alpha_\nu$ . Let the convex set  $C_{\mathcal{U}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau)$  be generated by the set  $\mathcal{U}$  according to the probability distribution  $\mathbb{P}$  over  $\mathcal{V}_{\mathcal{D}}$  as

$$C_{\mathcal{U}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau) = \{(\lambda, \tau) \mid \forall u \in \mathcal{U} \text{ satisfying (2) – (6)}\}. \quad (\star)$$

The convex set  $C_{\mathcal{U}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau)$  constitutes the set of feasible instantiations to the LP  $\mathcal{L}_r(\mathcal{U})$  and is exactly in the form of Equation 5 in [16]. Using  $C_{\mathcal{U}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau)$ , we reformulate  $\mathcal{L}_r(\mathcal{U})$  as the convex program

$$\begin{aligned} & \text{minimize } \tau \\ & \text{subject to } (\lambda, \tau) \in C_{\mathcal{U}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau), \end{aligned} \quad (9)$$

where the last constraint denotes that for a given  $(\lambda, \tau)$ , the feasible set of  $C_{\mathcal{U}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau)$  is not empty, i.e., there exists a feasible solution pair  $(\lambda, \tau)$  to the scenario problem  $\mathcal{L}_r(\mathcal{U})$ . This convex program asserts that all MCs in  $\mathcal{U}$  should induce a reachability probability that is less than  $\tau$ , satisfying the specification  $\varphi_r$ . Moreover, the convex program constitutes a scenario approximation to the so-called *chance-constrained problem* [1]. Such an optimization problem states that the probability of satisfying a (chance) constraint is above a certain threshold:

$$\begin{aligned} & \text{minimize } \tau \\ & \text{subject to } (\lambda, \tau) \in \mathbb{R} \times \mathbb{R}, \\ & \mathbb{P} \left( (\lambda, \tau) \in C_{\mathcal{V}_{\mathcal{D}}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau) \right) \geq 1 - \nu. \end{aligned} \quad (10)$$

The chance constraint in (10) ensures that the probability that an instantiation—obtained via distribution  $\mathbb{P}$ —satisfies the specification  $\varphi_r$  is at least  $1 - \nu$ . Theorem 1 in [16] shows that any feasible solution to the problem in (9) is feasible to the problem in (10) with a confidence probability of  $1 - \alpha_\nu$ , which shows that the violation probability of the solution is at most  $\nu$ . In our case, the probability of violation is exactly the probability that the instantiated MCs do not satisfy the specification  $\varphi_r$ . Thus, the claim follows.  $\square$

*Remark 4 (Independence to model size).* The confidence probability in Theorem 1 is in fact independent from the number of states, transitions, or random parameters of the uMC. From a practical perspective, the number of samples that are needed for a certain confidence does not depend on the model size.

Finally, Theorem 1 asserts that with a probability of at least  $1 - \alpha_\nu$ , the next sampled point from  $\mathcal{V}_{\mathcal{D}}$  will satisfy the specification with a probability of at least  $1 - \nu$ . Note that  $\alpha_\nu$  is the tail probability of a binomial distribution. It converges exponentially rapidly to 0 in  $|\mathcal{U}|$  [16].

## 4.2 Satisfaction Probability by Treating Violating Samples

Theorem 1 assumes that all sampled points, that is, the induced MCs, satisfy the specification  $\varphi_r$ . This is a severe assumption in general. To lift this assumption, we consider the *discarding approach* from [19]. Specifically, after sampling a set of instantiations  $\mathcal{U}$  from  $\mathcal{V}_{\mathcal{D}}$  according to the probability distribution  $\mathbb{P}$ , we remove the constraints for the MCs that violate the specification  $\varphi_r$  from the LP. We construct the set  $\mathcal{R} = \mathcal{U} \setminus \mathcal{Q}$ , where  $\mathcal{Q}$  denotes the set of samples that induce MCs violating the specification  $\varphi_r$ . Therefore, the set  $\mathcal{R}$  denotes the set of sampled MCs that satisfy the specification  $\varphi_r$ . We then solve the LP  $\mathcal{L}_r(\mathcal{R})$

$$\begin{aligned} & \text{minimize } \tau \\ & \text{subject to } \forall u \in \mathcal{R}, \\ & \quad (2) - (6), \end{aligned} \tag{11}$$

where for  $u \in \mathcal{R}$  and  $s \in S$ ,  $p_s^u$  gives the probability of satisfying the reachability specification of the instantiated MC  $\mathcal{D}[u]$  at state  $s$ . The other constraints in the optimization problem in LP  $\mathcal{L}_r(\mathcal{R})$  are identical to the LP  $\mathcal{L}_r(\mathcal{U})$ . We give the main result of this section.

**Theorem 2.** *Let uMC  $\mathcal{D}$  and the sample sets  $\mathcal{U}, \mathcal{Q} \subseteq \mathcal{V}_{\mathcal{D}}$ , with  $K = |\mathcal{U}| \geq 2$  and  $L = |\mathcal{Q}|$ . For a given tolerance probability  $\nu \in [0, 1)$ , the associated confidence probability is*

$$\alpha_\nu = \binom{L+1}{L} \sum_{i=0}^{L+1} \binom{K}{i} (1-\nu)^{K-i} \nu^i. \tag{12}$$

Then, with a probability of at least  $1 - \alpha_\nu$ , we have

$$F(\mathcal{D}_{\mathbb{P}}, \varphi_r) \geq 1 - \nu. \tag{13}$$

*Proof.* Similar to the proof of Theorem 1, the main idea is to relate the LP  $\mathcal{L}_r(\mathcal{R})$  to the chance-constrained convex problem in (10). Then, we invoke the results from [19, Theorem 1] to get the desired result. Let the convex set  $C_{\mathcal{R}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau)$ , which is generated by the samples in  $\mathcal{R}$ , be defined by

$$C_{\mathcal{R}}^{\mathcal{D}_{\mathbb{P}}}(\lambda, \tau) = \{(\lambda, \tau) \mid \forall u \in \mathcal{R} \text{ such that } (*) \text{ is satisfied}\}.$$

The set  $C_{\mathcal{R}}^{\mathcal{D}^{\varphi}}(\lambda, \tau)$  is in the form of the Definition 2.1 in [16]. We reformulate the LP  $\mathcal{L}_r(\mathcal{R})$  as the convex program

$$\begin{aligned} & \text{minimize } \tau \\ & \text{subject to } (\lambda, \tau) \in C_{\mathcal{R}}^{\mathcal{D}^{\varphi}}(\lambda, \tau). \end{aligned} \quad (14)$$

where the last constraint denotes that the instantiated MCs from the parameter values of the set  $\mathcal{R}$  should induce a reachability probability less than  $\tau$ , and thus, satisfy the specification  $\varphi_r$ . The problem in (14) is a scenario approximation to the problem in (10). Theorem 2.1 in [19] asserts that with a probability of  $\alpha_\nu$ , the violation probability of the solution is at most  $\nu$ , which is the probability of violating the specification for the next sample. Similar to Theorem 1, the violation probability  $\nu$  is the probability that an instantiated MC does not satisfy the specification  $\varphi_r$ . Thus, the claim follows.  $\square$

### 4.3 Expected Cost Specifications

So far, we have focused on parameters that were uncontrollable, and assumed to be random. Now, we consider the case where the cost function  $c$  is parametric and the cost parameters are *controllable*. Therefore, the parameters in the cost function are now variables that we can optimize over to satisfy an expected cost specification  $\varphi_c = \text{EC}_{\leq \kappa}(\diamond G)$  for the instantiated MCs. Similar to the previous sections, we assume that we sample a set of instantiations  $\mathcal{U}_c$  from the probability distribution  $\mathbb{P}$  over the parameter space  $\mathcal{V}_{\mathcal{D}}$ . In this case, we modify the LP  $\mathcal{L}_r(\mathcal{U})$  to obtain the following LP, which we denote by  $\mathcal{L}_c(\mathcal{U}_c)$ ,

$$\begin{aligned} & \text{minimize } \tau \\ & \text{subject to } \forall u \in \mathcal{U}_c, \\ & c_{s_I}^u \leq \tau, \\ & c_{s_I}^u \leq \kappa, \\ & c_s^u = 0 \quad \forall s \in G, \\ & c_s^u = c(s) + \sum_{s' \in S} \mathcal{P}(s, s')[u] c_{s'}^u \quad \forall s \in S \setminus G, \end{aligned} \quad (15)$$

where for  $s \in S$ ,  $c(s) \in \mathbb{R}_{\geq 0}^{|\mathcal{W}|}$  is the cost function at state  $s$ ,  $|\mathcal{W}|$  is the number of the cost parameters, and for  $u \in \mathcal{U}_c$ ,  $c_s^k$  gives the expected cost of reaching the target  $G$  of the instantiated MC  $\mathcal{D}[k]$  at state  $s$ . Note that the cost parameters  $\mathcal{W}$  are in the LP  $\mathcal{L}_c(\mathcal{U}_c)$  as variables for the parametric cost function. In the scenario problem (15), we optimize over  $c(s)$  and  $c_s^k$  to minimize the maximal induced cost of the instantiated MCs. If  $c$  is an *affine* function, then the optimization problem  $\mathcal{L}_c(\mathcal{U}_c)$  is convex. In this case, the probabilistic properties of the scenario problem are given by the following theorem.

**Theorem 3.** *Let uMC  $\mathcal{D}$  and the sample set  $\mathcal{U}_c \subseteq \mathcal{V}_{\mathcal{D}}$  with  $W = |\mathcal{W}|$ , and  $K = |\mathcal{U}_c| \geq W + 1$ . Assume for all  $u \in \mathcal{U}_c$ ,  $\mathcal{D}[k] \models \varphi_c$ . For a given tolerance probability  $\nu \in [0, 1)$ , let the associated confidence probability*

$$\alpha_\nu = \sum_{i=0}^{W+1} \binom{K}{i} (1-\nu)^{K-i} \nu^i. \quad (16)$$

Then, with a probability of at least  $1 - \alpha_\nu$ , we have  $F(\mathcal{M}_\mathbb{P}, \varphi_c) \geq 1 - \nu$ .

*Proof.* Following the proof of Theorem 1, we define the convex set

$$C_{\mathcal{U}_c}^{\mathcal{D}_\mathbb{P}}(\kappa, \tau, c) = \left\{ (\kappa, \tau, c) \mid \forall u \in \mathcal{U}_c \text{ such that} \right. \\ \left. \begin{aligned} c_{s_j}^u &\leq \tau, \\ c_{s_j}^u &\leq \kappa, \\ c_s^u &= 0 \quad \forall s \in G, \\ c_s^u &= c(s) + \sum_{s' \in S} \mathcal{P}(s, s')[u] c_{s'}^u \quad \forall s \in S \setminus G \end{aligned} \right\}$$

The main difference compared to the proof of Theorem 1 is that we have cost parameters in  $c$  as the decision variables and we consider an expected cost specification instead of a reachability specification. Similarly to the proof of Theorem 1, we reformulate the LP  $\mathcal{L}_c(\mathcal{U}_c)$  as the following convex problem

$$\begin{aligned} &\text{minimize } \tau \\ &\text{subject to } (\kappa, \tau, c) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{|\mathcal{W}|}, \\ &\quad (\kappa, \tau, c) \in C_{\mathcal{U}_c}^{\mathcal{D}_\mathbb{P}}(\kappa, \tau, c). \end{aligned} \tag{17}$$

This convex problem is a scenario approximation to the chance constrained problem given by

$$\begin{aligned} &\text{minimize } \tau \\ &\text{subject to } (\kappa, \tau, c) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{|\mathcal{W}|}, \\ &\quad \mathbb{P} \left( (\kappa, \tau, c) \in C_{\mathcal{V}_\mathcal{D}}^{\mathcal{D}_\mathbb{P}}(\kappa, \tau, c) \right) \geq 1 - \nu. \end{aligned} \tag{18}$$

Therefore, similar to the Theorem 1, we obtain the desired claim.  $\square$

We now consider the case that we compute an instantiation of the cost variables, and some of the instantiated MCs satisfy the expected cost specification. We construct the set  $\mathcal{R}_c = \mathcal{U}_c \setminus \mathcal{Q}_c$ , where  $\mathcal{Q}_c$  denotes the set of samples that induce MCs which violate the specification  $\varphi_c$ . For this case, we obtain:

**Theorem 4.** *Let uMC  $\mathcal{D}$  and the sample sets  $\mathcal{U}_c, \mathcal{Q}_c \subseteq \mathcal{V}_\mathcal{D}$ , with  $W = |\mathcal{W}|$ ,  $K = |\mathcal{U}_c| \geq 2$  and  $L = |\mathcal{Q}|$ . For a given tolerance probability  $\nu \in [0, 1)$ , let the associated confidence probability*

$$\alpha_\nu = \binom{l+W+1}{l} \sum_{i=0}^{l+W+1} \binom{K}{i} (1-\nu)^{K-i} \nu^i. \tag{19}$$

Then, with a probability of at least  $1 - \alpha_\nu$ , we have  $F(\mathcal{M}_\mathbb{P}, \varphi_c) \geq 1 - \nu$ .

*Proof.* The proof is similar to the proofs of Theorem 2 and 3, and omitted.  $\square$

#### 4.4 Building Scenario-Based Algorithms

The question remains how we leverage the theoretical results to compute an estimate on the satisfaction probability to solve Problems 1 and 2. For instance, let

$\nu$  be a violation probability and  $\mathcal{U}$  the sample set. Then, we can use Theorem 2 or 4 to compute the confidence probability  $\alpha_\nu$  by using the discarding approach from [19]. Similarly, for a the sample set  $\mathcal{U}$  and a threshold on the confidence probability  $\alpha_\nu$  we do a *bisection* on  $\nu$ . Specifically, we repeatedly apply Theorem 2 or 4 for different values of  $\nu \in (0, 1)$ , to see if the corresponding confidence probability  $\alpha_\nu$  is below the threshold. We then approximate the lower and upper bounds on  $\nu$ .

The correctness of the approach is based on scenario-based optimization. However, it also applies to an obtained solution by any procedure [39]. For instance, for any obtained value for the controlled parameters, we can construct a scenario program by sampling from random parameters. We can then apply Theorem 2 or 4 to compute the confidence probability  $\alpha_\nu$  or the violation probability  $\nu$ .

*Generalization to uMDPs.* Recall that we want to compute the satisfaction probability for a uMDP. The probability that for any sampled MDP we are able to synthesize a policy that satisfies the specification  $\varphi_r$ . To generalize our results to uMDPs, we can modify the constraint (6) in the LP  $\mathcal{L}_r(\mathcal{U})$  as

$$p_s^u \leq \sum_{s' \in S} \mathcal{P}(s, \alpha, s')[u] \cdot p_{s'}^u, \quad \forall s \in S \setminus (T \cup \neg \exists \diamond T), \quad \forall \alpha \in ActS(s), \quad (20)$$

asserting that, for each non-target state  $s \in S$  and action  $\alpha \in ActS(s)$ , the probability induced by the *minimizing policy* is an upper bound to the probability variables  $p_s^u$ . The reachability specification  $\varphi_r$  is satisfied if and only if the reachability probability at the initial state induced by the minimizing policy is less than  $\lambda$ . We can assert if  $\varphi_r$  is satisfied by combining the constraints (20) with the constraints (2)–(5). Then, our theoretical results apply to the uMDPs.

## 5 Numerical Examples

We implemented the approach from Section 4 using the model checker Storm [35] to construct and analyze samples of MDPs. To solve the scenario optimization problems with cost parameters, we used the SCS solver [31]. All computations ran on a computer with 8 2.2 GHz cores, and 32 GB of RAM.

We report on a set of well-known benchmarks used in parameter synthesis [46] that are, for instance, available on the website of the tools PARAM [17] or part of the PRISM benchmark suite [23]. Moreover, we created a dedicated case study that is based on the aforementioned UAV example.

### 5.1 Parameter Synthesis Benchmarks

*Setup.* In our first set of benchmarks, we adopt parametric MDPs and MCs from [32]. Essentially, the technique from that paper allows to approximate the percentage of instantiations that satisfy (or do not satisfy) a specification. We assume a uniform distribution over the parameter space and set  $\nu$  equal to the

**Table 1.** Information for the benchmark instances taken from [32].

benchmark	instance	$\varphi$	#pars	Model Information		Satisfaction Probability	
				#states	#trans	sat ( $1 - \nu$ )	unsat ( $\nu$ )
brp	(256,5)	$\mathbb{P}$	2	19 720	26 627	0.055	0.898
	(16,5)	$\mathbb{E}$	4	1 304	1 731	0.275	0.676
	(32,5)	$\mathbb{E}$	4	2 600	3 459	0.232	0.718
crowds	(10,5)	$\mathbb{P}$	2	104 512	246 082	0.537	0.413
	(20,7)	$\mathbb{P}$	2	45 421 597	164 432 797	0.416	0.534
nand	(10,5)	$\mathbb{P}$	2	35 112	52 647	0.218	0.733
	(25,5)	$\mathbb{P}$	2	865 592	1 347 047	0.206	0.744
consensus	(2,2)	$\mathbb{P}$	2	272	492	0.280	0.669
	(4,2)	$\mathbb{P}$	4	22 656	75 232	0.063	0.888

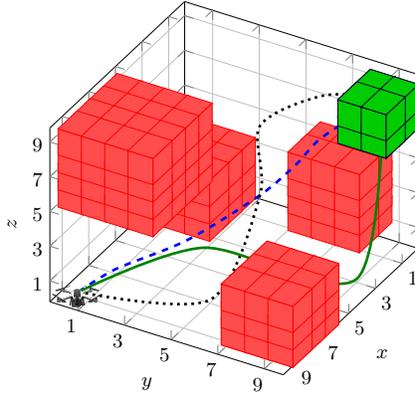
**Table 2.** Confidence probabilities  $\alpha_\nu$  for different numbers of samples.

Samples		100		1,000		10,000		Time (s)
benchmark	instance	$\alpha_\nu$ , sat	$\alpha_\nu$ , unsat	$\alpha_\nu$ , sat	$\alpha_\nu$ , unsat	$\alpha_\nu$ , sat	$\alpha_\nu$ , unsat	
brp	(256,5)	$9.99 \cdot 10^{-2}$	$7.02 \cdot 10^{-1}$	$1.60 \cdot 10^{-2}$	$7.77 \cdot 10^{-2}$	$1.12 \cdot 10^{-6}$	$3.55 \cdot 10^{-6}$	1761.45
	(16,5)	$2.72 \cdot 10^{-1}$	$1.97 \cdot 10^{-1}$	$1.14 \cdot 10^{-1}$	$3.36 \cdot 10^{-2}$	$5.52 \cdot 10^{-6}$	$1.80 \cdot 10^{-8}$	39.76
	(32,5)	$4.01 \cdot 10^{-1}$	$2.95 \cdot 10^{-1}$	$1.39 \cdot 10^{-1}$	$7.76 \cdot 10^{-2}$	$1.24 \cdot 10^{-6}$	$2.63 \cdot 10^{-6}$	78.17
crowds	(10,5)	$2.57 \cdot 10^{-1}$	$3.72 \cdot 10^{-1}$	$1.65 \cdot 10^{-1}$	$1.16 \cdot 10^{-1}$	$9.33 \cdot 10^{-7}$	$8.22 \cdot 10^{-4}$	0.19
	(20,7)	$4.18 \cdot 10^{-1}$	$1.38 \cdot 10^{-1}$	$2.41 \cdot 10^{-1}$	$9.48 \cdot 10^{-2}$	$5.81 \cdot 10^{-5}$	$2.83 \cdot 10^{-5}$	0.45
nand	(10,5)	$3.48 \cdot 10^{-1}$	$2.95 \cdot 10^{-1}$	$3.64 \cdot 10^{-2}$	$3.41 \cdot 10^{-1}$	$2.64 \cdot 10^{-9}$	$1.48 \cdot 10^{-4}$	144.26
	(25,5)	$4.42 \cdot 10^{-1}$	$3.71 \cdot 10^{-1}$	$4.12 \cdot 10^{-2}$	$3.78 \cdot 10^{-1}$	$3.49 \cdot 10^{-6}$	$2.91 \cdot 10^{-4}$	5327.82
consensus	(2,2)	$3.38 \cdot 10^{-1}$	$3.56 \cdot 10^{-1}$	$1.32 \cdot 10^{-1}$	$1.32 \cdot 10^{-1}$	$5.67 \cdot 10^{-7}$	$8.37 \cdot 10^{-4}$	0.72
	(4,2)	$1.79 \cdot 10^{-1}$	$1.41 \cdot 10^{-1}$	$6.51 \cdot 10^{-2}$	$4.75 \cdot 10^{-3}$	$4.26 \cdot 10^{-5}$	$9.29 \cdot 10^{-8}$	300.21

percentage of instantiations that do not satisfy the specification (and vice versa for  $1 - \nu$ ). We solve Problem 1 and show that the satisfaction probability is with confidence  $\alpha_\nu$  as least as high as the approximate satisfaction percentages from [32]. We adapt the *Consensus* protocol [3] and the *Bounded Retransmission Protocol* (brp) [5] to uMDPs; the *Crowds Protocol* (crowds) [12] and the *NAND Multiplexing* benchmark (nand) [8] become uMCs. In Table 1 we list the type of specification checked ( $\varphi$ ) and the number of parameters, states, and transitions. We also list the satisfaction probability (as obtained in [32]) for satisfying (sat) and falsifying (unsat) the specification  $\varphi$ .

*Results.* Table 2 shows the confidence probability  $\alpha_\nu$  for each benchmark to satisfy and falsify the specification after 100, 1 000 and 10 000 samples from the parameter space. In particular, for each number of samples, we report the average  $\alpha_\nu$  after running 10 full iterations of the same benchmark. Furthermore, we list the time to solve 1 000 samples for each instance (Time (s)).

The results in Table 2 show that for some benchmarks we get a high confidence probability already after 1 000 samples. For other benchmarks, the confidence probability is still considerably low, for instance considering nand and falsifying the specification. After 10 000 samples, we get a very high confidence in the satisfaction probability for all benchmarks. These results demonstrate that



**Fig. 2.** An example of a 3D UAV benchmark with obstacles and a target area.

we can efficiently compute a high confidence in the satisfaction probability. In particular, for the same number of samples, the obtained confidence probabilities are consistent for varying number of states and parameters of the underlying models. Therefore, no dependence on the size of models is shown (see Remark 4).

## 5.2 UAV Motion Planning

In our second benchmark, we consider the previously mentioned UAV motion planning example to model a realistic problem with a high number of random parameters. We model the problem as a uMDP, where the parameters represent how the weather conditions affect the movement of the UAV, and how the weather may change. In particular, different wind conditions induce specific satisfaction probabilities. We assume that the planning area is a certain valley where we have historic weather data which provide distributions over parameter values. The mission of the UAV is to transport a payload to a specific location and return safely to its initial position. The problem is to compute the satisfaction probability, that is, the probability that for any sampled MDP for this scenario we are able to synthesize a UAV policy that satisfies the specification.

We model the problem as follows: States encode the position of the UAV, the current weather situation, and the general wind direction in the valley. Parameters describe how the weather affects the position of the UAV for different zones in the valley, and how the weather/wind may change during the day. Fig. 2 shows an example environment with zones to avoid (red) and a target zone (green). We define four different weather conditions that each induce certain probability distributions over the eight different wind directions. The parameters of the model determine the probabilities of transitioning between different weather and wind conditions at each time step. The specification is to reach the target zone safely with a probability of at least 0.9. The number of states in our example is 266 880, and the number of parameters is 2 500.

For the distributions over parameter values, that is, over weather conditions, we consider the following cases. First, we assume a uniform distribution over the different weather conditions in each zone. Second, the probability for a weather condition inducing a wind direction that pushes the UAV into the positive  $y$ -direction is five times more likely than others. Similarly, in the third case, it is five times more likely to push the UAV into the negative  $x$ -direction. We depict some example trajectories of the UAV for three different conditions in Fig. 2. The trajectory given by the blue dashed line represents the expected trajectory for the first case, taking a direct route to reach the target area. Similarly, the trajectories given by the black dotted and solid green lines represent the expected trajectories for the second and third cases. For the second case, we observe that the UAV tries to avoid to get closer to the obstacles in  $x$  direction as the wind may push the UAV to the obstacles. For the third case, the UAV avoids the obstacle at the bottom and then reaches the target area.

We sample 1000 parameters for each case and approximate the maximal satisfaction probability with a confidence probability of at least  $1 - \alpha_\nu$ , with  $\alpha_\nu = 10^{-6}$ . The highest satisfaction probability is given by the first weather condition with 0.86, and the other conditions have a satisfaction probability of 0.78 and 0.75, showing that it may be harder to navigate around the obstacles with non-uniform probability distributions. The average time to compute the satisfaction probabilities is 1341 seconds.

Finally, we introduce costs to a 2-dimensional example, where hitting an obstacle causes (1) a cost of 100 and (2) the UAV to return to the initial position. Specifically, we introduce cost parameters for transitions that steer the UAV towards  $x$  or  $y$ -directions. We minimize the maximal possible expected cost (under all parameter values) to reach the target location. The specification asserts that the resulting expected cost should be less than 20.

We uniformly sample 1000 parameter values for weather conditions and note that the UAV policies favor on average transitioning to  $y$ -direction more compared to the  $x$ -direction to minimize the cost while ensuring that the probability of hitting an obstacle is minimized. The average expected cost of the induced MDPs is 7.41 and the satisfaction probability is 0.71. The solving time for this example is 2274 seconds.

## 6 Conclusion

We presented a new sampling-based approach to uncertain Markov models. Theoretically, we showed how to effectively and efficiently approximate the probability that any randomly drawn sample satisfies a temporal logic specification. Furthermore, we showed the computational tractability of our approaches by means of well-known benchmarks and a new, dedicated case study.

In the future, we plan to exploit our approaches for more involved models such as parametric extensions to continuous-time Markov chains [9] or Markov automata [22]. Another line of future work will be a closer integration with a parameter synthesis framework.

## References

1. Abraham Charnes and William W Cooper. Chance-Constrained Programming. *Management science*, 6(1):73–79, 1959.
2. Amir Pnueli. The Temporal Logic of Programs. In *FOCS*, pages 46–57, 1977.
3. James Aspnes and Maurice Herlihy. Fast Randomized Consensus Using Shared Memory. *Journal of Algorithms*, 15(1):441–460, 1990.
4. Hans Hansson and Bengt Jonsson. A Logic for Reasoning about Time and Reliability. *Formal Aspects of Computing*, 6(5):512–535, 1994.
5. L. Helmk, M. Sellink, and F. Vaandrager. Proof-Checking a Data Link Protocol. In *TYPES*, volume 806 of *LNCS*, pages 127–165. Springer, 1994.
6. Anthony Cassandra, Michael L Littman, and Nevin L Zhang. Incremental Pruning: A Simple, Fast, Exact Method for Partially Observable Markov Decision Processes. In *UAI*, pages 54–61, 1997.
7. Robert Givan, Sonia Leach, and Thomas Dean. Bounded-Parameter Markov Decision Processes. *Artificial Intelligence*, 122(1-2):71–109, 2000.
8. Jie Han and Pieter Jonker. A System Architecture Solution for Unreliable Nanoelectronic Devices. *IEEE Transactions on Nanotechnology*, 1:201–208, 2002.
9. Christel Baier, Boudewijn R. Haverkort, Holger Hermanns, and Joost-Pieter Katoen. Model-Checking Algorithms for Continuous-Time Markov Chains. *IEEE Trans. Software Eng.*, 29(6):524–541, 2003.
10. Joelle Pineau, Geoff Gordon, and Sebastian Thrun. Point-Based Value Iteration: an Anytime Algorithm for POMDPs. In *IJCAI*, pages 1025–1030, 2003.
11. Conrado Daws. Symbolic and Parametric Model Checking of Discrete-Time Markov chains. In *ICTAC*, volume 3407 of *LNCS*, pages 280–294. Springer, 2004.
12. Vitaly Shmatikov. Probabilistic Analysis of an Anonymity System. *Journal of Computer Security*, 12(3-4):355–377, 2004.
13. G.C. Calafiore and M.C. Campi. Uncertain Convex Programs: Randomized Solutions and Confidence Levels. *Mathematical Programming*, 102(1):25–46, 2005.
14. Arnab Nilim and Laurent El Ghaoui. Robust Control of Markov Decision Processes with Uncertain Transition Matrices. *Operations Research*, 53(5):780–798, 2005.
15. Giuseppe C. Calafiore and Marco C. Campi. The Scenario Approach to Robust Control Design. *IEEE Trans. Automat. Contr.*, 51(5):742–753, 2006.
16. Marco C. Campi and Simone Garatti. The Exact Feasibility of Randomized Solutions of Uncertain Convex Programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.
17. Ernst Moritz Hahn, Holger Hermanns, and Lijun Zhang. Probabilistic Reachability for Parametric Markov Models. *STTT*, 13(1):3–19, 2010.
18. Dimitris Bertsimas, David B Brown, and Constantine Caramanis. Theory and Applications of Robust Optimization. *SIAM review*, 53(3):464–501, 2011.
19. Marco C Campi and Simone Garatti. A Sampling-and-Discarding Approach to Chance-Constrained Optimization: Feasibility and Optimality. *Journal of Optimization Theory and Applications*, 148(2):257–280, 2011.
20. Benoît Delahaye, Kim G. Larsen, Axel Legay, Mikkel L. Pedersen, and Andrzej Wasowski. Decision problems for interval Markov chains. In *LATA*, volume 6638 of *LNCS*, pages 274–285. Springer, 2011.
21. Marta Kwiatkowska, Gethin Norman, and David Parker. PRISM 4.0: Verification of Probabilistic Real-Time Systems. In *CAV*, volume 6806 of *LNCS*, pages 585–591. Springer, 2011.

22. Hassan Hatefi and Holger Hermanns. Model Checking Algorithms for Markov Automata. *ECEASST*, 53, 2012.
23. Marta Kwiatkowska, Gethin Norman, and David Parker. The PRISM Benchmark Suite. In *QEST*, pages 203–204. IEEE CS, 2012.
24. Rowan McAllister, Thierry Peynot, Robert Fitch, and Salah Sukkarieh. Motion Planning and Stochastic Control with Experimental Validation on a Planetary Rover. In *IROS*, pages 4716–4723. IEEE, 2012.
25. Eric M. Wolff, Ufuk Topcu, and Richard M. Murray. Robust Control of Uncertain Markov Decision Processes with Temporal Logic Specifications. In *CDC*, pages 3372–3379, 2012.
26. Taolue Chen, Ernst Moritz Hahn, Tingting Han, Marta Kwiatkowska, Hongyang Qu, and Lijun Zhang. Model Repair for Markov Decision Processes. In *TASE*, pages 85–92. IEEE CS, 2013.
27. Alberto Puggelli, Wenchao Li, Alberto L Sangiovanni-Vincentelli, and Sanjit A Seshia. Polynomial-Time Verification of PCTL Properties of MDPs with Convex Uncertainties. In *CAV*, pages 527–542. Springer, 2013.
28. Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. Robust Markov Decision Processes. *Mathematics of Operations Research*, 38(1):153–183, 2013.
29. Indika Meedeniya, Irene Moser, Aldeida Aleti, and Lars Grunske. Evaluating Probabilistic Models with Uncertain Model Parameters. *Software & Systems Modeling*, 13(4):1395–1415, 2014.
30. Krishnendu Chatterjee, Martin Chmelík, and Mathieu Tracol. What is Decidable about Partially Observable Markov Decision Processes with  $\omega$ -Regular Objectives. *Journal of Computer and System Sciences*, 82(5):878–911, 2016.
31. B. O’Donoghue, E. Chu, N. Parikh, and S. Boyd. Conic Optimization via Operator Splitting and Homogeneous Self-Dual Embedding. *Journal of Optimization Theory and Applications*, 169(3):1042–1068, June 2016.
32. Tim Quatmann, Christian Dehnert, Nils Jansen, Sebastian Junges, and Joost-Pieter Katoen. Parameter Synthesis for Markov Models: Faster Than Ever. In *ATVA*, volume 9938 of *LNCS*, pages 50–67, 2016.
33. Asrar Ahmed, Pradeep Varakantham, Meghna Lowalekar, Yossiri Adulyasak, and Patrick Jaillet. Sampling Based Approaches for Minimizing Regret in Uncertain Markov Decision Processes (MDPs). *J. Artif. Intell. Res.*, 59:229–264, 2017.
34. Murat Cubuktepe, Nils Jansen, Sebastian Junges, Joost-Pieter Katoen, Ivan Papanusha, Hasan A. Poonawala, and Ufuk Topcu. Sequential Convex Programming for the Efficient Verification of Parametric MDPs. In *TACAS (2)*, volume 10206 of *LNCS*, pages 133–150, 2017.
35. Christian Dehnert, Sebastian Junges, Joost-Pieter Katoen, and Matthias Volk. A Storm is Coming: A Modern Probabilistic Model Checker. In *CAV (2)*, volume 10427 of *LNCS*, pages 592–600. Springer, 2017.
36. Dimitri Scheftelowitsch, Peter Buchholz, Vahid Hashemi, and Holger Hermanns. Multi-Objective Approaches to Markov Decision Processes with Uncertain Transition Parameters. In *VALUETOOLS*, pages 44–51, 2017.
37. Sebastian Arming, Ezio Bartocci, Krishnendu Chatterjee, Joost-Pieter Katoen, and Ana Sokolova. Parameter-Independent Strategies for pMDPs via POMDPs. In *QEST*, pages 53–70. Springer, 2018.
38. Luca Bortolussi and Simone Silveti. Bayesian Statistical Parameter Synthesis for Linear Temporal Properties of Stochastic Models. In *TACAS*, pages 396–413, 2018.
39. Marco Claudio Campi, Simone Garatti, and Federico Alessandro Ramponi. A General Scenario Theory for Nonconvex Optimization and Decision Making. *IEEE Trans. Automat. Contr.*, 63(12):4067–4078, 2018.

40. Murat Cubuktepe, Nils Jansen, Sebastian Junges, Joost-Pieter Katoen, and Ufuk Topcu. Synthesis in pmdps: A tale of 1001 parameters. In *ATVA*, volume 11138 of *LNCS*, pages 160–176. Springer, 2018.
41. Paul Gainer, Ernst Moritz Hahn, and Sven Schewe. Incremental Verification of Parametric and Reconfigurable Markov Chains. *CoRR*, abs/1804.01872, 2018.
42. Chin Pang Ho and Marek Petrik. Fast Bellman Updates for Robust MDPs. In *ICML*, 2018.
43. Yamilet R. Serrano Llerena, Marcel Böhme, Marc Brünink, Guoxin Su, and David S. Rosenblum. Verifying the Long-run Behavior of Probabilistic System Models in the Presence of Uncertainty. In *ESEC/SIGSOFT FSE*, pages 587–597. ACM, 2018.
44. Lauren N Steimle, David L Kaufman, and Brian T Denton. Multi-Model Markov Decision Processes. *Optimization Online*, 2018.
45. Giovanni Bacci, Mikkel Hansen, and Kim Guldstrand Larsen. Model Checking Constrained Markov Reward Models with Uncertainties. In *QEST*, pages 37–51, 2019.
46. Sebastian Junges, Erika Ábrahám, Christian Hensel, Nils Jansen, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. Parameter Synthesis for Markov Models. *CoRR*, abs/1903.07993, 2019.
47. Tobias Winkler, Sebastian Junges, Guillermo A. Pérez, and Joost-Pieter Katoen. On the complexity of reachability in parametric markov decision processes. In *CONCUR*, volume 140 of *LIPICs*, pages 14:1–14:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.
48. Christel Baier and Joost-Pieter Katoen. *Principles of Model Checking*. MIT Press, 2008.
49. S. Basu, R. Pollack, and M.F. Roy. *Algorithms in Real Algebraic Geometry*. Springer, 2010.
50. Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.
51. Martin L Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.
52. Stuart J Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. 2016.
53. Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT press, 2018.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

