Formal Controller Synthesis for Stochastic Dynamical Models with Epistemic Uncertainty

Thom S. Badings
Radboud University
Nijmegen, the Netherlands

Alessandro Abate
University of Oxford
Oxford, United Kingdom

Nils Jansen
Radboud University
Nijmegen, the Netherlands

David Parker
University of Birmingham
Birmingham, United Kingdom

Hasan A. Poonawala
University of Kentucky
Kentucky, USA

Marielle Stoelinga
Radboud University
Nijmegen, the Netherlands

Abstract

Capturing both aleatoric and epistemic uncertainty in models of robotic systems is crucial to designing safe controllers. Most existing approaches for synthesizing certifiably safe controllers exclusively consider aleatoric but not epistemic uncertainty, thus requiring that model parameters and disturbances are known precisely. Our contribution to overcoming this restriction is a novel abstraction-based controller synthesis method for continuous-state models with stochastic noise, uncertain parameters, and external disturbances. By sampling techniques and robust analysis, we capture both aleatoric and epistemic uncertainty, with a user-specified confidence level, in the transition probability intervals of a so-called interval Markov decision process (iMDP). We then synthesize an optimal policy on this abstract iMDP, which translates (with the specified confidence level) to a feedback controller for the continuous model, with the same performance guarantees. Our experimental benchmarks confirm that accounting for epistemic uncertainty leads to controllers that are more robust against variations in parameter values.

1 Introduction

Reach-avoid problems. Reach-avoid problems are omnipresent in robotic motion planning [6]. Consider, for example, an unmanned aerial vehicle (UAV) that must navigate to a desirable region within a given time horizon, while avoiding certain unsafe regions. The problem is to synthesize a controller, which guarantees that the UAV will reach its goal while avoiding the unsafe regions. A powerful approach to synthesizing such certifiably safe controllers leverages probabilistic verification to provide formal guarantees over reach-avoid specifications. Most robotic systems are, however, characterized by continuous state and action spaces, while formal verification is generally limited to discrete models. Thus, finite abstractions are used to make formal verification feasible for continuous-state models [1]. Being formal, verification guarantees on the finite abstraction carry over to the continuous model. In this paper, we adopt such an abstraction-based approach to controller synthesis.

Probabilities are not enough. To account for uncertainty, robotic systems are often modelled using stochastic dynamical models. Recently, the notion of uncertainty has often been distinguished in *aleatoric* (statistical) and *epistemic* (systematic) uncertainty [7, 12]. Aleatoric uncertainty captures natural randomness (i.e., stochasticity) in the outcome of transitions, while epistemic uncertainty is in particular modelled by parameters that are not precisely known [11]. A general premise is that purely probabilistic approaches fail to capture epistemic uncertainty [10]. In this work, we aim to reason *under both aleatoric and epistemic uncertainty*, in order to synthesize provably-correct controllers for safety-critical applications. Existing abstraction methods fail to achieve this novel, general goal.

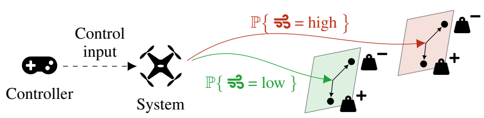

Figure 1: Aleatoric (stochastic) uncertainty in the wind (🌬) causes probability distributions over the outcomes of controls; epistemic uncertainty in the UAV's mass (⚖) causes transitions to be nondeterministic.

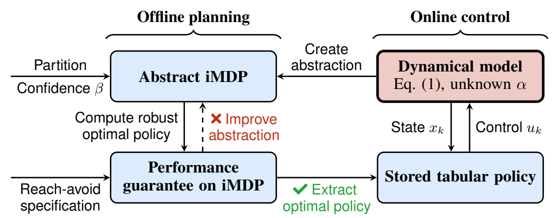

Figure 2: Our overall approach to solve the problem consists of an offline planning phase in which we create an iMDP abstraction, and an online control phase to derive a provably-correct feedback controller.

Models with epistemic uncertainty. We consider reach-avoid problems for stochastic dynamical models with continuous state and action spaces, under epistemic uncertainty due to uncertain parameters and external disturbances. These parameters and disturbances lie within a *convex uncertainty set* (in the simplest case, intervals), such as a drone whose mass is only known to lie between 0.75–$1.25\,\mathrm{kg}$. As shown in Fig. 1, the dynamics depend on uncertain factors, e.g., the wind and the drone's mass. For the wind, we may derive a probabilistic model from, e.g., weather data, to reason over the likelihood of state dynamics. For the mass (epistemic uncertainty), however, no information allows us to reason probabilistically, yielding *nondeterministic* state dynamics. Concretely, we consider a dynamical model whose state $x_k \in \mathbb{R}^n$ at time $k \in \mathbb{N}$ evolves as

$$x_{k+1} = A(\alpha)x_k + B(\alpha)u_k + q_k + \eta_k, \tag{1}$$

where $u_k \in \mathcal{U}$ is the (constrained) control input, and η_k is a stochastic process noise, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ is assumed to be unknown. The matrices $A(\alpha) = \sum_{i=1}^{r} \alpha_i A_i$ and $B(\alpha) = \sum_{i=1}^{r} \alpha_i B_i$ are a convex combination of a finite set of matrices $A_1, \dots, A_r$ and $B_1, \dots, B_r$, where $\alpha \in \Gamma$ belongs to the unit simplex Γ. Moreover, the disturbance $q_k \in \mathcal{Q}$ is only known to lie in a given convex uncertainty set $\mathcal{Q}$. Thus, the model in Eq. (1) captures epistemic uncertainty in q_k, as well as in $A(\alpha)$ and $B(\alpha)$ through model parameter α.

Problem statement. Our goal is to synthesize a feedback controller for Eq. (1) that is *robust against nondeterminism* due to parameter uncertainty and disturbances, and that *reasons over probabilities* derived from stochastic noise. We wish to synthesize a controller with a *probably approximately correct (PAC)*-style guarantee to satisfy a reach-avoid task. Thus, we solve the following problem:

Problem. Given a reach-avoid specification for the dynamical model in Eq. (1), compute a feedback controller and a *lower bound* $\lambda \in [0, 1]$ on the probability that, *under any admissible value of the parameters* (i.e., under any $\alpha \in \Gamma$, $q_k \in \mathcal{Q}\ \forall k \in \mathbb{N}$), the specification is probabilistically satisfied with this lower bound and *with at least a user-specified confidence probability* $\beta \in (0, 1)$.

We solve this problem via a discrete-state abstraction of the continuous model in the form of a so-called interval Markov decision process (iMDP), which is an extension of an MDP with intervals of transition probabilities [9]. We generate this abstraction by partitioning the continuous state space and defining actions that induce potential transitions between elements of this partition.

Robustness to capture nondeterminism. The main contribution that allows us to capture nondeterminism, is that we reason over *sets* of potential transitions (as shown by the boxes in Fig. 1), rather than *precise* transitions, e.g., as in [2]. Intuitively, for a given action, the aleatoric uncertainty creates a probability distribution over *sets of possible outcomes*. To ensure robustness against epistemic uncertainty, we consider *all possible outcomes* within these sets. We show that, for our class of models, computing these sets of all possible outcomes is computationally tractable. Building upon this reasoning, we provide the following guarantees related to the aforementioned problem.

1) PAC guarantees on abstractions. We show that both probabilities and nondeterminism can be captured in the probability intervals of an iMDP. We use sampling methods from scenario optimization [8] and concentration inequalities [4] to compute PAC bounds on these intervals. With a predefined confidence probability, the iMDP correctly captures both aleatoric and epistemic uncertainty.

2) Correct-by-construction control. For the iMDP, we compute a *robust optimal policy* that maximizes the worst-case probability of satisfying the reach-avoid specification. The iMDP policy is automatically translated to a *provably-correct feedback controller* for the original, continuous model 'on the fly'. This means that, by construction, the PAC guarantees on the iMDP carry over to the satisfaction of the specification for the continuous model, thus solving the problem stated above.

Contributions. We develop the first abstraction-based, formal controller synthesis method that simultaneously captures epistemic and aleatoric uncertainty for continuous-state/action models. We provide results on the PAC-correctness of obtained iMDP abstractions, and guarantees on the synthesized controllers for a reach-avoid specification. In the remainder of this paper, we highlight the key elements of our approach and show with a benchmark that accounting for epistemic uncertainty yields controllers that are more robust against deviations in parameter values and disturbances.

2 Abstraction-Based Controller Synthesis

Our overall approach is shown in Fig. 2 and consists of an *offline planning* phase in which we create the iMDP and compute a robust optimal policy, and an *online control* phase in which we automatically derive a provably-correct controller for the continuous model. We briefly discuss both phases.

1) Create abstraction. Given a model as in Eq. (1), we create an abstract iMDP by partitioning the state space into a set of convex polytopic regions, each of which corresponds with a state of the iMDP. We then define the iMDP actions via backward reachability computations under a so-called *nominal model* that neglects both the aleatoric and epistemic uncertainty in Eq. (1), and is thus deterministic. We compensate for the error caused by this simplification in the iMDP's transition probability intervals. Intuitively, the upper/lower bounds of the intervals correspond to the *best/worst possible outcome of the epistemic uncertainty*, respectively. Using principled sampling methods from scenario optimization [8] and concentration inequalities [4], we compute PAC bounds on these intervals, yielding an iMDP that is PAC-correct. That is, for a user-specified confidence $\beta \in (0, 1)$, it holds, for every iMDP state-action pair (s, a) and successor state s', that $\mathbb{P}^N \Big\{ \underline{p} \leq P(s,a)(s') \leq \bar{p} \Big\} \geq 1 - \beta$.

2) Compute robust optimal policy. Using tools from probabilistic model checking [3], we compute a robust optimal policy for the resulting iMDP. This robust optimal policy maximizes (over the available action) the worst-case (over the probability intervals) probability of satisfying the reach-avoid specification. Recall that the overall problem is to find a controller with a reach-avoid probability of at least λ. If this condition holds for the obtained policy, we output the policy and proceed to step 3; otherwise, we attempt to improve the abstraction in one of the following ways. First, we can refine the partition at the cost of a larger iMDP. Second, using more samples N yields an improved iMDP through tighter intervals (see, e.g., [2] for such trade-offs). Finally, the uncertainty in $\alpha \in \Gamma$ may be too large, meaning we need to reduce set Γ using learning techniques (see Sect. 4 work).

3) Online control. We use the extracted policy on the abstract iMDP to derive a provably-correct feedback controller for the dynamical system 'on the fly'. At each time step k, we determine to which polytopic region the current state x_k belongs, and we obtain the corresponding optimal action from the iMDP policy. We then compute the actual control input u_k for the dynamical model associated with this iMDP action as the solution of a convex optimization program. Due to the correctness of our abstraction procedure, the formal guarantees on the iMDP carry over to the satisfaction of the reach-avoid specification on the dynamical system. As a result, our approach leads to a *provably-correct feedback controller* for the dynamical system, thus solving the aforementioned problem.

3 Synthesizing Robust Controllers

We show that our approach yields controllers that are robust against deviations in parameter values and disturbances. Consider a UAV whose longitudinal position p_k and velocity v_k are modelled as

$$x_{k+1} = \begin{bmatrix} p_k \\ v_k \end{bmatrix} = \begin{bmatrix} 1 & \tau \\ 0 & \frac{m - 0.1\tau}{m} \end{bmatrix} x_k + \begin{bmatrix} \frac{\tau^2}{2m} \\ \frac{\tau}{m} \end{bmatrix} u_k + \eta_k,$$

with τ the discretization time, and $\mathcal{U} = [-5, 5]$. The mass $m \in \mathbb{R}$ of the UAV is uncertain, and is only known to lie in $m \in [0.75, 1.25]$. We fix the nominal value of the mass as $\hat{m} = 1$. The specification is to reach a position of $p_k \geq 8$ before time $K = 12$, while avoiding speeds of $|v_k| \geq 10$. We use a partition of 24×20 regions and 20K samples to estimate probability intervals. We compare against a baseline that builds an iMDP for the nominal model only, thus neglecting parameter uncertainty.

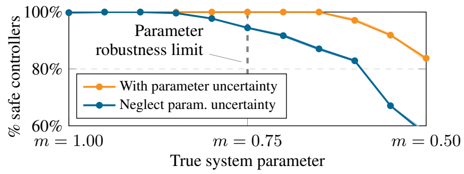

Figure 3: % of initial states with safe performance guarantees (i.e., the simulated reachability probability is above the guaranteed reach-avoid probability on the iMDP). Our approach that accounts for epistemic uncertainty is *100% safe up to the parameter robustness limit*; neglecting uncertainty is not.

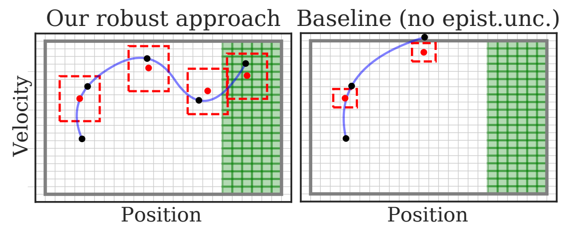

Figure 4: With our approach, the system safely reaches the goal (in green), while the baseline neglecting epistemic uncertainty leaves the safe set (gray box), as it underestimates the epistemic uncertainty (shown as the red boxes).

Neglecting epistemic uncertainty is unsafe. The run time of this benchmark is $\pm\, 3\,\mathrm{s}$. For initial state x_0, we define a controller c at a parameter value $\alpha \in \Gamma$ to be *unsafe* if the reach-avoid probability $V(x_0, \alpha, c)$ (which we estimate using Monte Carlo simulations) is below the guaranteed reach-avoid probability on the iMDP abstraction. In Fig. 3, we show the deviation of the actual mass m from its nominal value, versus the average percentage of states with a safe controller (over 10 repetitions). The *parameter robustness limit* represents the extreme values of the parameter against which our approach is guaranteed to be robust ($m = 0.75$ and 1.25 in this case).

Our approach yields *100% safe controllers* for deviations well over the robustness limit of $m \in [0.75, 1.25]$. By contrast, the baseline yields *6% unsafe controllers* at the robustness limit. We show simulated trajectories under an actual mass $m = 0.75$ in Fig. 4. These trajectories confirm that our approach safely reaches the goal region, while the baseline does not, as it neglects epistemic uncertainty. We observe similar results for models with multiple parameters.

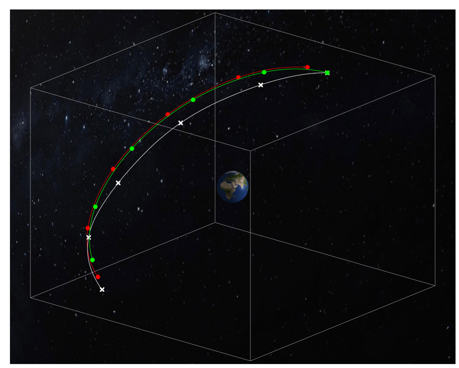

Figure 5: Simulated trajectory for a spacecraft reach-avoid problem with only aleatoric uncertainty. The chaser spacecraft (white) must navigate to the target (green), while not colliding with the one in red.

4 Conclusions and Current Research

We presented a novel abstraction-based controller synthesis method for dynamical models with aleatoric uncertainty due to stochastic noise, and epistemic uncertainty due to uncertain parameters and external disturbances. The method captures those different types of uncertainties in order to ensure certifiably safe controllers. Our experiment in this short paper shows that we derive controllers that are robust against against deviations in the uncertain model parameters.

In the future, we wish to apply our method to larger reach-avoid problems, such as the spacecraft problem shown in Fig. 5 (which is a result obtained under only aleatoric uncertainty). Moreover, we wish to integrate the abstractions in a *safe learning framework* [5], such that we use our method to synthesize controllers that are provably safe, while reducing the epistemic uncertainty by interacting with the system. We also aim to extend our method to *nonlinear systems*, such as non-holonomic robots [13], by capturing linearization errors as external disturbances. The main challenge is to obtain a set-bounded representation of the linearization error, which depends on the model dynamics.

References

[1] A. Abate, M. Prandini, J. Lygeros, and S. Sastry. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11):2724 – 2734, 2008.

[2] 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. In *AAAI (to appear)*.

AAAI Press / The MIT Press, 2022.

[3] C. Baier and J. P. Katoen. *Principles of Model Checking*. MIT Press Books, 2008.

[4] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities - A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[5] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig. Safe learning in robotics: From learning-based control to safe reinforcement learning. *CoRR*, abs/2108.06266, 2021.

[6] J. F. Fisac, M. Chen, C. J. Tomlin, and S. S. Sastry. Reach-avoid problems with time-varying dynamics, targets and constraints. In *HSCC*, pages 11–20. ACM, 2015.

[7] C. R. Fox and G. Ülkümen. Distinguishing two dimensions of uncertainty. *Fox, Craig R. and Gülden Ülkümen (2011),"Distinguishing Two Dimensions of Uncertainty," in Essays in Judgment and Decision Making, Brun, W., Kirkebøen, G. and Montgomery, H., eds. Oslo: Universitetsforlaget*, 2011.

[8] S. Garatti and M. C. Campi. L-inf Layers and the Probability of False Prediction. *IFAC Proceedings Volumes*, 42(10):1187–1192, 2009. ISSN 14746670. doi: 10.3182/20090706-3-fr-2004.00197.

[9] R. Givan, S. M. Leach, and T. L. Dean. Bounded-parameter markov decision processes. *Artif. Intell.*, 122(1-2):71–109, 2000.

[10] E. Hüllermeier and W. Waegeman. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach. Learn.*, 110(3):457–506, 2021.

[11] R. C. Smith. *Uncertainty quantification: theory, implementation, and applications*, volume 12. Siam, 2013.

[12] T. J. Sullivan. *Introduction to uncertainty quantification*, volume 63. Springer, 2015.

[13] S. Thrun, W. Burgard, and D. Fox. *Probabilistic robotics*. Intelligent robotics and autonomous agents. MIT Press, 2005.