Proceedings of Machine Learning Research vol 283:1–13, 2025 7th Annual Conference on Learning for Dynamics and Control

Temporal Logic Control for Nonlinear Stochastic Systems Under Unknown Disturbances

Ibon Gracia IBON.GRACIA@COLORADO.EDU
University of Colorado Boulder

Luca Laurenti L.LAURENTI@TUDELFT.NL
Delft University of Technology

Manuel Mazo Jr. M.MAZO@TUDELFT.NL
Delft University of Technology

Alessandro Abate ALESSANDRO.ABATE@CS.OX.AC.UK
University of Oxford

Morteza Lahijanian MORTEZA.LAHIJANIAN@COLORADO.EDU
University of Colorado Boulder

Editors: N. Ozay, L. Balzano, D. Panagou, A. Abate

Abstract

In this paper, we present a novel framework to synthesize robust strategies for discrete-time nonlinear systems with random disturbances that are unknown, against temporal logic specifications. The proposed framework is data-driven and abstraction-based: leveraging observations of the system, our approach learns a high-confidence abstraction of the system in the form of an uncertain Markov decision process (UMDP). The uncertainty in the resulting UMDP is used to formally account for both the error in abstracting the system and for the uncertainty coming from the data. Critically, we show that for any given state-action pair in the resulting UMDP, the uncertainty in the transition probabilities can be represented as a convex polytope obtained by a two-layer state discretization and concentration inequalities. This allows us to obtain tighter uncertainty estimates compared to existing approaches, and guarantees efficiency, as we tailor a synthesis algorithm exploiting the structure of this UMDP. We empirically validate our approach on several case studies, showing substantially improved performance compared to the state-of-the-art.

Keywords: Data-Driven Control, Strategy Synthesis, Uncertain MDPs, Safe Autonomy

1. Introduction

The synthesis of safe strategies for stochastic systems is critical in ensuring *reliable* and *safe* operations in domains such as robotics and cyber-physical systems (Belta et al., 2007; Lavaei et al., 2022). When the system is subject to *unknown* random disturbances, it becomes difficult to account for uncertainties while guaranteeing performance against high-level specifications. Existing methods often assume known distributions for the disturbances or rely on abstractions with overly conservative uncertainty estimates, limiting their scalability and applicability in practice. This paper addresses these gaps by presenting a novel framework to obtain strategies for nonlinear stochastic systems with unknown disturbances, ensuring both formal guarantees and computational efficiency.

Our framework employs a data-driven, abstraction-based approach to strategy synthesis for stochastic systems with unknown noise under *linear temporal logic over finite traces* (LTL_f) (De Giacomo and Vardi, 2013) specifications. Starting with data from the system's trajectories, we construct a high-confidence abstraction in the form of an uncertain Markov decision process (UMDP)

(Iyengar, 2005), a flexible model that captures complex uncertainties. Unlike existing methods relying on interval-based abstractions or conservative assumptions, our framework represents transition probability uncertainties as convex polytopes. These sets are derived through a novel two-layer discretization scheme and learning the support of the unknown disturbance. This leads to tighter uncertainty sets and less conservative results compared to existing methods. Exploiting this UMDP structure, we introduce a synthesis algorithm for LTL$_f$ specifications that simplifies the computation, reducing the complexity of standard UMDP linear programming approaches. By incorporating uncertainties from both abstraction errors and data limitations, our framework yields a strategy that is robust. Our empirical evaluations over various types of systems reveals the efficacy of this approach over existing methods, namely in data efficiency, tightness of results, and scalability.

The main contributions of this paper are fourfold: (i) a novel strategy synthesis framework for nonlinear stochastic systems under non-additive, unknown disturbances with LTL$_f$ specifications, (ii) a distribution-agnostic, data-driven construction of UMDP abstraction whose structure reduces conservatism of existing abstraction-based techniques, (iii) an efficient synthesis algorithm for this UMDP abstraction, which does not introduce additional conservatism , and (iv) a set of case studies that show superiority of the framework over the state-of-the-art, with up to 3 orders of magnitude improvement in sample complexity and an order of magnitude reduction in computation time.

Related Work Abstractions of stochastic systems to finite Markov decision processes (MDPs) are powerful tools for controller synthesis on highly-complex systems under complex logic specifications (Lavaei et al., 2022). In particular, Interval MDPs (IMDPs) (Lahijanian et al., 2015; Givan et al., 2000) abstract systems by presenting uncertain transition probabilities within intervals, capturing the full range of system behaviors. For example, (Cauchi et al., 2019) efficiently abstracts linear systems with additive Gaussian noise, while (Skovbekk et al., 2023) extends this to nonlinear dynamics. Uncertain MDPs (UMDPs) (Iyengar, 2005; El Ghaoui and Nilim, 2005) generalize IMDPs by allowing transition probabilities to belong to more complex sets and have been used for strategy synthesis against specifications such as linear temporal logic (LTL) (Wolff et al., 2012). However, these abstractions typically require system models, which are often unavailable in practice.

To address model uncertainty, existing methods leverage Gaussian processes (Jackson et al., 2021), neural networks (Adams et al., 2022), and ambiguity sets (Gracia et al., 2022), which are then abstracted as IMDPs or UMDPs. Statistical tools like the scenario approach are also used to abstract stochastic (Badings et al., 2023b,a), non-deterministic (Kazemi et al., 2024), and deterministic systems (Coppola et al., 2022, 2023). Also, techniques like super-martingales and barrier functions enable safety verification and control for general dynamics (Lechner et al., 2022; Badings et al., 2024; Mazouz et al., 2024), by assuming, however, that the disturbance distribution is known.

When disturbance distributions are uncertain, some works combine IMDP abstractions with statistical tools (Badings et al., 2023b,a; Schön et al., 2023), while others employ barrier certificates with the scenario approach for safety verification (Mathiesen et al., 2023). Others (Gracia et al., 2025a) construct Wasserstein ambiguity sets from data samples to abstract systems as UMDPs, which account for the uncertainty regarding the unknown distributions. However, these methods typically assume simple dynamics or additive noise. For general dynamics with unknown disturbance distributions, only a few works exist. (Salamati et al., 2021) uses barrier certificates for safety verification, and (Gracia et al., 2024) extends the ambiguity set approach to nonlinear dynamics, formally characterizing the guarantees of these approaches. Both assume that certain distribution-related properties, such as variance or support, are known—an assumption often unrealistic in practice, and also suffer from high sample complexity, especially under high-confidence requirements. Our work overcomes these limitations by removing assumptions about disturbance distributions and offering a data-efficient and scalable approach suitable for systems with general dynamics.

2. Problem Formulation

In this work the focus is on discrete-time stochastic systems given by

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, u_k, \boldsymbol{w}_k), \tag{1}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ denotes the state at time $k \in \mathbb{N}_0$, $u_k \in U \subset \mathbb{R}^m$ is the control input chosen from a finite set U, and $\boldsymbol{w}_k \in W \subseteq \mathbb{R}^d$ is the disturbance. The latter is a sequence of independent and identically distributed (i.i.d.) random variables on the probability space $(W, \mathcal{B}(W), P_W)$, with $\mathcal{B}$ being the Borel σ-algebra on W, being the support W and probability distribution P_W of $\boldsymbol{w}_k$ *unknown*. The (possibly nonlinear) vector field $f : \mathbb{R}^n \times U \times W \to \mathbb{R}^n$ is assumed to be Lipschitz continuous on its third argument, uniformly for all values of its first argument on some set.

Assumption 1 *There exists a set $X \subset \mathbb{R}^n$, such that, for every $u \in U$, there exists a constant $L_u > 0$ such that, for all $x \in X$, $w, w' \in W$, it holds that $\|f(x,u,w) - f(x,u,w')\| \leqslant L_u \|w - w'\|$.*

In lieu of unknown W and P_W, we assume a dataset on the disturbance is available.

Assumption 2 *A set $\{\hat{\boldsymbol{w}}^{(i)}\}_{i=1}^N$ of N i.i.d. samples from P_W is available.*

Assumption 2 is commonly made in related work (Badings et al., 2023b; Gracia et al., 2024) and can be practically satisfied through, e.g., observations of the state and control. This is evident when f is affine in $\boldsymbol{w}_k$; otherwise, it suffices for f to be injective over only a subset of $\mathbb{R}^n$ as discussed in (Gracia et al., 2024). This condition is met by many practical systems, like those in our case studies.

Given $x_0, \ldots, x_K \in \mathbb{R}^n$, $u_0, \ldots, u_{K-1} \in U$, and $K \geqslant 0$, we denote a finite *trajectory* of System (1) by $\omega_x = x_0 \xrightarrow{u_0} \ldots \xrightarrow{u_{K-1}} x_K$. We let $|\omega_x|$ denote the length of ω_x, define Ω_x as the set of all trajectories with $|\omega_x| < \infty$ and denote by $\omega_x(k)$ the state of ω_x at time $k \in \{0, \ldots, |\omega_x|\}$. A *strategy* of System (1) is a function $\sigma_x : \Omega_x \to U$ that assigns a control u to each finite trajectory ω_x. Given $x \in \mathbb{R}^n$, $u \in U$, the *transition kernel* $T : \mathcal{B}(\mathbb{R}^n) \times \mathbb{R}^n \times U \to [0,1]$ of System (1) assigns the probability $T(B \mid x, u) = \int_W \mathbb{1}_B(f(x,u,w)) P_W(dw)$, where the indicator function $\mathbb{1}_B(x) = 1$ if $x \in B$, and 0 otherwise, to each Borel set $B \in \mathcal{B}(\mathbb{R}^n)$. For a strategy σ_x and an initial condition $x_0 \in \mathbb{R}^n$, the transition kernel defines a unique probability measure $P_{x_0}^{\sigma_x}$ over the trajectories of System (1) (Bertsekas and Shreve, 1996). In this way, $P_{x_0}^{\sigma_x}[\boldsymbol{\omega}_x(k) \in B]$ denotes the probability that $\boldsymbol{x}_k$ belongs to the set $B \in \mathcal{B}(\mathbb{R}^n)$ when following strategy σ_x from initial state x_0. In this work, we are interested in the temporal behavior of System (1) w.r.t. a bounded (*safe*) set $X \in \mathcal{B}(\mathbb{R}^n)$ and a set of regions of interest R_{int}, with $r \subseteq X$ and $r \in \mathcal{B}(\mathbb{R}^n)$ for all $r \in R_{\text{int}}$. We denote by $r_{\text{unsafe}} = \mathbb{R}^n \backslash X$ the unsafe region. We consider a set $AP := \{\mathfrak{p}_1, \ldots, \mathfrak{p}_{|AP|-1}, \mathfrak{p}_{\text{unsafe}}\}$ of *atomic propositions*, and associate a subset of atomic propositions to each region $r \in R_{\text{int}} \cup \{r_{\text{unsafe}}\}$. We define the *labeling function* $L : \mathbb{R}^n \to 2^{AP}$ as the function that maps each state $x \in \mathbb{R}^n$ to the atomic propositions that are true in the region where x lies, e.g., if we associate $\{\mathfrak{p}_1\}$ to region r_1, we conclude that $\mathfrak{p}_1$ is *true* at x, denoted $\mathfrak{p}_1 \equiv \top$, if $x \in r_1$. In consequence, each trajectory $\omega_x = x_0 \xrightarrow{u_0} \ldots \xrightarrow{u_{K-1}} x_K$ results in the (observation) *trace* $\rho = \rho_0 \ldots \rho_K$, where $\rho_k := L(x_k)$.

In order to formally characterize behaviors of System (1), we use *linear temporal logic over finite traces* (LTL$_f$) (De Giacomo and Vardi, 2013), which generalizes Boolean logic to temporal behaviors. An LTL$_f$ property φ is a logical formula defined over atomic propositions AP using Boolean connectives "negation" ($\neg$) and "conjunction" ($\wedge$), and the temporal operators "until" ($\mathcal{U}$) and "next" ($\bigcirc$). The syntax of formula φ is recursively defined as $\varphi := \top \mid \mathfrak{p} \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \bigcirc\varphi \mid \varphi_1 \mathcal{U} \varphi_2$, where $\mathfrak{p} \in AP$ and φ_1, φ_2 are also LTL$_f$ formulas. The temporal operators "eventually" ($\Diamond$) and "globally" ($\Box$) are derived from the above syntax as $\Diamond\varphi := \top\mathcal{U}\varphi$ and $\Box\varphi := \neg\Diamond(\neg\varphi)$. LTL$_f$ formulae are semantically interpreted over finite traces (De Giacomo and Vardi, 2013). We say a trajectory ω_x satisfies a formula φ, i.e., $\omega_x \models \varphi$, if some prefix of its trace ρ satisfies φ.

Our goal is to synthesize a strategy for System (1) to ensure satisfaction of a given LTL$_f$ formula φ. However, note that (i) under a given strategy, the satisfaction of φ is probabilistic, and (ii) in our setting, the distribution of the disturbance is unknown. Hence, we aim to leverage data samples to generate a strategy that guarantees System (1) satisfies φ with high probability. Furthermore, note that the synthesized strategy must account for the learning gap due to the lack of knowledge of P_W.

Problem 1 *Consider stochastic System (1), a set $\{\hat{\boldsymbol{w}}^{(i)}\}_{i=1}^N$ of N i.i.d. samples from P_W, a bounded set $X \in \mathcal{B}(\mathbb{R}^n)$ on which Assumption 1 holds, and an LTL$_f$ formula φ' defined over the regions of interest R_{int}. Given a confidence level $1-\alpha \in (0,1)$, synthesize a strategy σ_x and a high probability bound function $\underline{p} : X \to [0,1]$ such that, with confidence at least $1-\alpha$, for every initial state $x_0 \in X$, σ_x guarantees that the probability that the paths $\omega_x \in \Omega_x$ satisfy $\varphi := \varphi' \wedge \Box\neg\mathfrak{p}_{unsafe}$ while remaining in X is lower bounded by $\underline{p}(x_0)$, i.e., $P_{x_0}^{\sigma_x}[\omega_x \models \varphi] \geqslant \underline{p}(x_0)$.*

We remark that P_W is unknown, and no assumptions are imposed on it. Instead, since only samples are available, the probabilistic satisfaction guarantees must hold with a *confidence*, related to the probability that the N samples are representative of P_W, and interpreted in the frequentist sense: if the process of sampling N times from P_W and synthesizing σ_x is repeated infinitely many times, the condition $P_{x_0}^{\sigma_x}[\omega_x \models \varphi] \geqslant \underline{p}(x_0)$ for all $x_0 \in X$ holds in at least $1-\alpha$ of the cases.

Overview of the approach Given the uncountable nature of the state-space of System (1) and P_W being unknown, solving Problem 1 exactly is infeasible. Therefore, we adopt an abstraction-based approach, which provides a strategy along with a conservative, high-probability bound for every initial state. The abstraction is an uncertain MDP (UMDP) constructed from a finite partition of X. We learn the transition relations between regions using f and disturbance samples, capturing the system's behavior with confidence $1-\alpha$. Our UMDP construction is specifically designed to tightly capture the learning uncertainty. Then, we devise a strategy synthesis algorithm based on robust dynamic programming to (robustly) maximize the probability of satisfying φ on this UMDP. Next, we refine this strategy to System (1) such that it guarantees that the closed-loop system satisfies φ with a probability higher than the one obtained for the abstraction with confidence $1-\alpha$.

3. Preliminaries on Uncertain Markov Decision Processes

An uncertain MDP (UMDP), also known as a robust MDP, is a stochastic system that generalizes the MDP class by allowing its transition probability distributions to be uncertain, taking values from a set (Iyengar, 2005; El Ghaoui and Nilim, 2005; Wiesemann et al., 2013).

Definition 1 (Uncertain MDP) *A labeled uncertain Markov decision process (UMDP) $\mathcal{M}$ is a tuple $\mathcal{M} = (S, A, \Gamma, s_0, AP, L)$, where S and A are finite sets of states and actions, respectively, $s_0 \in S$ is the initial state, $\Gamma = \{\Gamma_{s,a} \subseteq \mathcal{P}(S) : s \in S, a \in A\}$, where $\mathcal{P}(S)$ is the set of probability* *distributions on S, and $\Gamma_{s,a} \neq \emptyset$ is the set of transition probability distributions for state $s \in S$ and action $a \in A$, AP is a finite set of atomic propositions, and $L : S \to 2^{AP}$ is the labeling function.*

A *strategy* of a UMDP $\mathcal{M}$ is a function $\sigma : \Omega \to A$ that maps each finite path to the next action. We denote by Σ the set of all strategies of $\mathcal{M}$. Given path $\omega \in \Omega$ and $\sigma \in \Sigma$, the process evolves from $s_k = \mathrm{last}(\omega)$ under $a_k = \sigma(\omega)$ to the next state according to a distribution in Γ_{s_k,a_k}. This distribution is chosen by the adversary (Givan et al., 2000). Formally, an *adversary* is a function $\xi : S \times A \times (\mathbb{N} \cup \{0\}) \to \mathcal{P}(S)$ that maps each state s_k, action a_k, and time step $k \in \mathbb{N} \cup \{0\}$ to a transition probability distribution $\gamma \in \Gamma_{s_k,a_k}$, such that $s_{k+1} \sim \gamma$. We denote the set of all adversaries by Ξ. Given an initial condition $s_0 \in S$, a strategy $\sigma \in \Sigma$ and an adversary $\xi \in \Xi$, the UMDP collapses to a Markov chain with a unique probability distribution $Pr_{s_0}^{\sigma,\xi}$ over its paths.

4. Data-driven UMDP Abstraction

In this section, we construct a UMDP $\mathcal{M}$, whose path probabilities are guaranteed to encompass the probabilities of System (1)'s trajectories with confidence $1 - \alpha$. We define the set of states S of $\mathcal{M}$ as follows. Let $R := \{r_1, \ldots, r_{|R|}\}$ be a finite partition of the continuous state-space $\mathbb{R}^n$ into non-overlapping, non-empty regions, which respects the regions of interest R_{int} and the safe set X, and such that $r \in \mathcal{B}(\mathbb{R}^n)$ for all $r \in R$. We let region $r_{|R|} := r_{\text{unsafe}}$ represent the unsafe set. We assign each region $r \in R$ to a state $s \in S$ in the abstraction $\mathcal{M}$ through the bijective map $J : R \to S$, which ensures that $J^{-1}(s) = r \in R$ is unique. For simplicity, we abuse the notation and also say $J(x) = s$ if $x \in r$ with $J(r) = s$. We define the action set A of $\mathcal{M}$ as $A := U$. With a slight abuse of language, we also denote by L the labeling function of $\mathcal{M}$, which maps each state $s \in S$ to the atomic propositions that hold at $x \in r = J^{-1}(s)$. Next, we define the set of transition probability distributions of the abstraction. To that end, we begin by stating the following proposition, whose proof follows from (Badings et al., 2023a, Eq.12)[1], which gives uniform bounds in the probabilities that System (1) transitions from each point x in some region $r \in R$ to some region $\tilde{r} \in \mathcal{B}(\mathbb{R}^n)$.

Proposition 2 *Given a region $r \in R$, an action $a \in A$ and a realization $w \in W$ of $\boldsymbol{w}$, denote by $Reach(r, a, w) := \{f(x, a, w) : x \in r\}$ the reachable set of r under a and w. Then, the probability of transitioning from each state $x \in r$ to region $\tilde{r} \in \mathcal{B}(\mathbb{R}^n)$ under action $a \in A$ is bounded by*

$$P_W(\{w : Reach(r, a, w) \subseteq \tilde{r}\}) \leqslant T(\tilde{r} \mid x, a) \leqslant P_W(\{w : Reach(r, a, w) \cap \tilde{r} \neq \emptyset\}) \tag{2}$$

Below, we use the samples of $\boldsymbol{w}$ to derive data-driven bounds that contain the ones in (2) and leverage these bounds to define the set of transition probability distributions for $\mathcal{M}$.

4.1. Data-Driven Transition Probability Bounds

We now construct the sets $\Gamma_{s,a}$ of transition probability distributions of the abstraction by leveraging the samples from $\boldsymbol{w}$. Specifically, in our abstraction, each set $\Gamma_{s,a}$ is defined by: (i) interval bounds on the probability of transitioning to each state $s' \in S$, (ii) interval bounds on the probability of transitioning to a cluster of states in 2^S, and (iii) a bound on the probability of transitioning to states within the reachable set of the learned support of P_W. Notably, (ii) and (iii) distinguish our construction from prior work, which relies solely on (i). As a result, our UMDP incorporates additional

1. The measurability of the events in (2) is formally proved in the supplementary material (Gracia et al., 2025b).

constraints, leading to tighter uncertainty sets. This yields less conservative probabilistic guarantees, as shown in the case studies. To derive the bounds (i)-(iii), we use Proposition 2, samples from $\boldsymbol{w}$, and two well-known concentration inequalities. Proposition 3 enables us to compute bounds on transition probabilities between regions, which we later use to obtain bounds in (i)-(ii).

Proposition 3 *Consider the set* $\{\hat{\boldsymbol{w}}^{(i)}\}_{i=1}^{N}$ *of i.i.d. samples from* $\boldsymbol{w}$*. Pick* $r \in R$*,* $a \in A$*,* $\tilde{r} \in \mathcal{B}(\mathbb{R}^n)$ *and* $\beta \in (0,1)$*, and let* $\epsilon = \sqrt{\log(2/\beta)/(2N)} > 0$*. Then, with confidence at least* $1-\beta$ *we have that, for all* $x \in r$*,*

$$T(\tilde{r} \mid x, a) \geqslant \underline{P}(r,a)(\tilde{r}) := \frac{1}{N}\left|\{i \in \{1,\ldots,N\} : \mathit{Reach}(r,a,\hat{\boldsymbol{w}}^{(i)}) \subseteq \tilde{r}\}\right| - \epsilon \tag{3a}$$

$$T(\tilde{r} \mid x, a) \leqslant \overline{P}(r,a)(\tilde{r}) := \frac{1}{N}\left|\{i \in \{1,\ldots,N\} : \mathit{Reach}(r,a,\hat{\boldsymbol{w}}^{(i)}) \cap \tilde{r} \neq \varnothing\}\right| + \epsilon. \tag{3b}$$

Proof Consider the lower bound in (2). Denote $E := \{w \in W : \mathrm{Reach}(r,a,w) \subseteq \tilde{r}\}$ and note that $T(\tilde{r} \mid x,a) \geqslant P_W(E) = \mathbb{E}_{P_W}[\mathbb{1}_E(\boldsymbol{w})]$ for all $x \in r$. Therefore applying Hoeffding's inequality to the random variable $\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_E(\hat{\boldsymbol{w}}^{(i)})$ yields $P_W^N[P_W(E) \geqslant \frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_E(\hat{\boldsymbol{w}}^{(i)}) - \epsilon] \geqslant 1-\beta/2$, with $\epsilon = \sqrt{\log(2/\beta)/(2N)}$. Thus, the first expression in (3) holds for all $x \in r$ with confidence $1-\beta/2$. Employing a similar argument, we obtain a similar result for the second expression in (3). Combining both results via the union bound, we obtain the result. ■

Remark 4 *The complexity of computing the bounds in* (3) *is proportional to* N*, which is typically high to obtain tight bounds. To reduce this complexity, we cluster the* N *samples from* $\boldsymbol{w}$ *into* $N_c \ll N$ *clusters, each with center* c_j *and diameter* ϕ_j*. Substituting the sets* $\mathit{Reach}(r,a,\hat{\boldsymbol{w}}^{(j)})$ *in* (3) *by* $\{f(x,a,w) \in \mathbb{R}^n : x \in r, \|w - c_j\| \leqslant \phi_j/2)\}$*, it is evident that Proposition 3 still holds, with relaxed bounds. Note that this clustering induces a partition on* W*, allowing to overapproximate the sets* $\{f(x,a,w) \in \mathbb{R}^n : x \in r, \|w - c_j\| \leqslant \phi_j/2)\}$ *as shown by Skovbekk et al. (2023).*

Next, we estimate the support of P_W in (iii). Including this information into $\mathcal{M}$ tightens the sets $\Gamma_{s,a}$ of transition probability distributions, thus yielding a less conservative abstraction.

Proposition 5 (Confidence Region (Tempo et al., 2013)) *Let* $\hat{c} = \max\{\|\hat{\boldsymbol{w}}^{(1)}\|,\ldots,\|\hat{\boldsymbol{w}}^{(N)}\|\}$*. Then, for any* $\epsilon_c, \beta_c > 0$ *and* $N \geqslant \log(1/\beta_c)/\log(1/(1-\epsilon_c))$*, it holds, with a confidence greater than* $1-\beta_c$ *with respect to the random choice of* $\{\hat{\boldsymbol{w}}^{(i)}\}_{i=1}^{N}$*, that* $P_W(\{w \in W : \|w\| \leqslant \hat{c}\}) \geqslant 1-\epsilon_c$*.*

We denote the learned confidence region for w by $\widehat{W} := \{w \in W : \|w\| \leqslant \hat{c}\}$, which contains at least $1-\epsilon_c$ probability mass from P_W with a confidence greater than $1-\beta_c$. We also define $\widehat{\mathrm{Post}}(s,a) := \{J(f(x,a,w)) \in S : x \in J^{-1}(s), w \in \widehat{W}\}$ for each $s \in S$, $a \in A$ to be the set of abstract states that can be reached from region $J^{-1}(s)$ and under some noise realization $w \in \widehat{W}$.

We now have all the components needed to define our abstraction. Intuitively, we leverage a two-layer discretization: a fine one represented by S and a coarse one formed by clustering the elements of S (see Figure 1). Let $Q \subseteq 2^S$ be this clustering, which is non-overlapping and $\bigcup_{q \in Q} q = S$. Q is crucial for obtaining non-zero lower-bound transition probabilities in (3a), as $\mathrm{Reach}(r,a,\hat{\boldsymbol{w}}^{(i)})$ often cannot be contained within a single small region but can be captured by a cluster of them (see Figure 1). We also leverage the learned support $\widehat{W}$ of P_W to impose the constraint that the successor state of a given (s,a) pair lies on some region with high probability. As described in (Gracia et al., 2025b), this constraint is key for our approach to be effective. Note that varying R and Q yields a different abstraction, which makes them hyperparameters. We now formally define our abstraction.

Definition 6 (UMDP Abstraction) *Let $Q \subseteq 2^S$ be a non-overlapping clustering of S and, for all $s \in S \backslash \{s_{|S|}\}$, $a \in A$, let $Q(s,a) \subseteq Q$ be the subset that covers $\widehat{Post}(s,a)$, i.e., $\widehat{Post}(s,a)$ is contained in $C(s,a) := \bigcup_{q \in Q(s,a)} q$. We define the UMDP abstraction of System (1) as $\mathcal{M} = (S, A, s_0, \Gamma, AP, L)$, with, $\forall s \in S \backslash \{s_{|S|}\}$ and $\forall a \in A$,*

$$\Gamma_{s,a} := \Big\{\gamma \in \mathcal{P}(S) : \underline{P}(r_s,a)(r_{s'}) \leqslant \gamma(s') \leqslant \overline{P}(r_s,a)(r_{s'}) \qquad \forall s' \in C(s,a), \\ \underline{P}(r_s,a)(r_{q'}) \leqslant \sum_{s' \in q'} \gamma(s') \leqslant \overline{P}(r_s,a)(r_{q'}) \;\; \forall q' \in Q(s,a), \sum_{s' \in C(s,a)} \gamma(s') \geqslant 1 - \epsilon_c \Big\}, \quad (4)$$

where $\underline{P}, \overline{P}$ are defined in (3), $r_{s'} = J^{-1}(s')$, and $r_{q'} = \bigcup_{s' \in q'} J^{-1}(s')$, and $\Gamma_{s_{|S|},a} = \{\delta_{s_{|S|}}\}$ for all $a \in A$, where $\delta_{s_{|S|}}$ denotes the Dirac measure located at $s_{|S|}$.

Note that, by making the unsafe state absorbing, we embed the safety part $\Box\neg\mathfrak{p}_{\text{unsafe}}$ of φ into $\mathcal{M}$, since a path of $\mathcal{M}$ satisfies φ' only if it remains in X. In Theorem 7, we establish that the UMDP $\mathcal{M}$ is a sound abstraction of System (1), i.e., that $\mathcal{M}$ captures all 1-step behaviors of System (1).

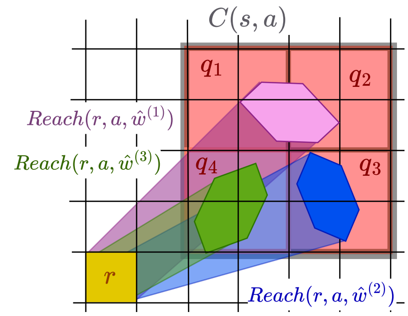

Figure 1: Illustration of the sets in Def. 1. $C(s,a) = \{q_1, q_2, q_3, q_4\}$, and each q_i contains 4 states. The probability that the successor state of $s = J(r)$ under action a will be in $C(s,a)$ is higher than $1-\epsilon$. Note that the reachable sets corresponding to $\hat{\boldsymbol{w}}^{(2)}$ and $\hat{\boldsymbol{w}}^{(3)}$ are contained in q_3 and q_4, respectively, but no single regions in the fine partition contains them completely.

Theorem 7 (Soundness of UMDP Abstraction) *For all $s \in S \backslash \{s_{|S|}\}$, $a \in A$, $x \in J^{-1}(s)$, $s' \in S$ define $\gamma_x \in \mathcal{P}(S)$ as $\gamma_x(s') := T(J^{-1}(s') \mid x, a)$ for all $s' \in S$. Then, $\gamma_x \in \Gamma_{s,a}$ for all $s \in S \backslash \{s_{|S|}\}$, $a \in A$, with confidence of at least $1-\alpha$, where $\alpha = \beta_c + \big(\sum_{s \in S \backslash \{s_{|S|}\}, a \in A} |C(s,a)| + |Q(s,a)|\big)\beta$.*

The proof of Theorem 7 is provided in the supplementary material (Gracia et al., 2025b). Note that $|C(s,a)|$ and $|Q(s,a)|$ depend on the samples from P_W trough $\widehat{W}$ (see Proposition 5). However, the expression for α in Theorem 7 follows from a union bound argument, which holds only if the number of transitions $\sum_{s \in S \backslash \{s_{|S|}\}, a \in A} |C(s,a)| + |Q(s,a)|$ is deterministic. We solve this issue by first estimating this number. Then, given $\alpha, \beta_c, \epsilon_c$ and N, we solve for β and ϵ, and obtain the corresponding abstraction $\mathcal{M}$. Finally, we check that the actual number of transitions does not exceed the estimated one, guaranteeing that $\mathcal{M}$ is sound by Theorem 7.

Corollary 8 *Given $\alpha \in (0,1)$ and $\epsilon, \epsilon_c \in (0,1)$, the sample complexity of obtaining a UMDP abstraction with confidence at least $1-\alpha$ is $N = \max\big\{\log(\frac{n_{learn}/\alpha}{2\epsilon^2}), \log(\frac{n_{learn}/\alpha}{\log(1/(1-\epsilon_c))})\big\}$, with $n_{learn} = 1 + \sum_{s \in S \backslash \{s_{|S|}\}, a \in A} |C(s,a)| + |Q(s,a)|$.*

5. Strategy Synthesis

We now synthesize a strategy for System (1) and lower bound the probability that the closed-loop system satisfies formula φ. We first show that standard synthesis procedures for general UMDPs from the literature (Wolff et al., 2012; Gracia et al., 2025a) apply to our setting and introduce a novel (tailored) algorithm that leverages the specific structure of our abstraction for faster computations.

Algorithm 1 2-layer O-maximization

Require: $\mathcal{M}, s \in S \backslash \{s_{|S|}\}, a \in A, \underline{p}^k$
Ensure: γ
Sort $\text{Post}(s, a)$ according to $\{\underline{p}^k(s')\}_{s' \in \text{Post}(s,a)}$ in increasing order
$\gamma(s') \leftarrow \underline{P}(s, a)(s')$ for all $s' \in \text{Post}(s, a)$
$\gamma(s') \leftarrow 0$ for all $s' \notin \text{Post}(s, a)$
$\gamma(q') \leftarrow \sum_{s' \in q'} \gamma(s')$ for all $q' \in Q(s, a)$
$M \leftarrow 1 - \sum_{s' \in S} \gamma(s')$
for $q' \in Q(s, a)$ **do**
  $m \leftarrow \underline{P}(s, a)(q') - \gamma(q')$
  if $m > 0$ **then**
    for $s' \in q'$ **do**
      $g \leftarrow \min\{m, \overline{P}(s, a)(s') - \gamma(s')\}$
      $\gamma(s') \leftarrow \gamma(s') + g, \gamma(q') \leftarrow \gamma(q') + g$
      $m \leftarrow m - g, M \leftarrow M - g$
for $s' \in \text{Post}(s, a)$ **do**
  $q' \leftarrow$ get cluster q' such that $s' \in q'$
  if $q' \neq \varnothing$ **then**
    $m \leftarrow \min\{M, \overline{P}(s, a)(q') - \gamma(q'), \overline{P}(s, a)(s') - \gamma(s')\}$
    $\gamma(q') \leftarrow \gamma(q') + m$
  else
    $m \leftarrow \min\{M, \overline{P}(s, a)(s') - \gamma(s')\}$
  $\gamma(s') \leftarrow \gamma(s') + m, M \leftarrow M - m$

We start by translating φ into its equivalent deterministic finite automaton $\mathcal{A}^\varphi$ (De Giacomo and Vardi, 2013) and constructing the product $\mathcal{M}^\varphi = \mathcal{M} \otimes \mathcal{A}^\varphi$ (Wolff et al., 2012). We then synthesize a strategy σ^φ on $\mathcal{M}^\varphi$ via unbounded-horizon *robust dynamic programming* (RDP) with a reachability objective, as detailed in (Gracia et al., 2025a, Theorems 6.2, 6.6). σ^φ maximizes the probability of satisfying φ under adversarial choices of transition probabilities from Γ^φ. Finally, we refine σ^φ into a strategy σ_x for System (1). Theorem 9, whose proof is given in Gracia et al. (2025b), ensures that satisfaction probability bounds are preserved under this procedure, thus solving Problem 1.

Theorem 9 (Strategy Synthesis through Product UMDP) *Let σ^φ and $\underline{p}^\varphi$ be respectively the optimal strategy and the lower bound in the probability of satisfying φ obtained via RDP on $\mathcal{M}^\varphi$. Denote by S_0^φ the set of initial states of $\mathcal{M}^\varphi$ and by Lift $: S \to S_0^\varphi$ the function that maps states of $\mathcal{M}$ to S_0^φ. Furthermore, let σ_x be the strategy obtained by refining σ^φ to System (1). Then, with confidence $1 - \alpha$, $Pr_x^{\sigma^\varphi}[\omega_x \models \varphi] \geqslant \underline{p}^\varphi(\mathit{Lift}(s))$ for all $x \in X$, where $s = J(x)$.*

Tailored Synthesis Algorithm We introduce a synthesis algorithm tailored for UMDPs as such in Definition 6, which exploits their structure for greater efficiency. The algorithm draws inspirations from IMDP value iteration (Givan et al., 2000) to speed up the computation of the optimal adversaries in RDP, which is typically formulated a linear program in standard UMDPs. We note that this approach is applicable to our product UMDP $\mathcal{M}^\varphi$ because it retains the same structure of $\mathcal{M}$, as detailed in the supplementary material Gracia et al. (2025b). Therefore and, for simplicity, we describe the algorithm in the context of a reachability problem on $\mathcal{M}$ rather than on $\mathcal{M}^\varphi$.

The algorithm first streamlines the uncertainty set of $\mathcal{M}$. Observe that the sets $\Gamma_{s,a}$ of $\mathcal{M}$ can be simplified by discarding the last constraint in (4) and adjusting the transition probability upper bounds accordingly. Specifically, we increase $\overline{P}(r_s, a)(r_{s_{|S|}})$ and $\overline{P}(r_s, a)(r_{q'})$ for all $q' \in Q(s, a)$ with $s_{|S|} \in q'$ by ϵ_c, and set $\overline{P}(r_s, a)(r_{s'}) = 0$ for all $s' \notin C(s, a) \cup \{s_{|S|}\}$. As a result, the adversary can choose a distribution with a higher probability of transitioning to the unsafe region. It can be shown that the solution $\underline{p}$ obtained via RDP in this simplified UMDP is not greater than the one obtained for $\mathcal{M}$. We provide a formal proof in (Gracia et al., 2025b), and hereby assume that $\mathcal{M}$ is simplified as described. Then, on the modified $\mathcal{M}$, Alg. 1 computes the optimal adversary for each state-action pair (s, a) by extending the O-maximizing algorithm devised for IMDP value iteration (Givan et al., 2000) to $\mathcal{M}$: via a 2-layer O-maximizing logic, the algorithm efficiently allocates probability mass to states of $\mathcal{M}$ with the lowest value function while respecting the constraints in (4). It begins by ensuring that the lower bounds $\underline{P}(r_s, a)(r_{s'})$ are satisfied for all states s' (Lines 2-3), then proceeds to allocate mass to each cluster $q' \in Q(s, a)$ to meet the required lower bounds (Lines 4-12). The algorithm ensures that the total probability mass remains feasible by maintaining the constraints $\gamma(s') \leqslant \overline{P}(r_s, a)(r_{s'})$, $\sum_{s' \in q'} \gamma(s') \leqslant \overline{P}(r_s, a)(r_{q'})$ throughout the allocation (Lines 13-20). This allocation process guarantees that as much mass as possible is assigned to states with the smallest value function while ensuring $\gamma \in \Gamma_{s,a}$. Alg. 1 terminates once all mass is allocated.Note that RDP algorithm (Gracia et al., 2025a) calls Alg. 1 for every (s, a) and in every iteration until termination. The following theorem proves its correctness and runtime complexity.

Theorem 10 (Correctness of Algorithm 1) *Let $k \in \mathbb{N} \cup \{0\}$, $s \in S \backslash \{s_{|S|}\}$, $a \in A$ and $\underline{p}^k \in \mathbb{R}^{|S|}$ be the value function obtained after k iterations of RDP. Define Post$(s, a) := \{s' \in S : \overline{P}(r_s, a)(r_{s'}) > 0\}$. Then, the output γ of Algorithm 1 satisfies $\gamma \in \arg\min_{\gamma \in \Gamma_{s,a}} \sum_{s' \in S} \gamma(s') \underline{p}^k(s')$, and it has a computational complexity of $\mathcal{O}(|\text{Post}(s, a)| \log(|\text{Post}(s, a)|))$.*

The proof of Theorem 10 is provided in (Gracia et al., 2025b). Note that the computational complexity of solving the linear program in the theorem statement using a standard Simplex algorithm is of $\mathcal{O}(|\text{Post}(s, a)|^3)$, highlighting the computational advantage of using Alg. 1.

Remark 11 *Note that classical IMDP abstractions use only the first constraint of Γ in (4), giving the adversary freedom to choose worse γ than if the abstraction were like our UMDPs. Consequently, IMDPs yield lower probabilistic guarantees, further motivating our abstraction choice.*

6. Case Studies

We now demonstrate empirically the effectiveness of our approach through 7 case studies. These include a nonlinear pendulum with non-additive disturbances, kinematic unicycle models with 2- and 3-D state-spaces and under nonlinear coulomb friction, a 2-D linear system with multiplicative noise, and a 4-D thermal regulation benchmark with multiplicative uncertainty. The considered specifications are reach-avoid (φ_1), the LTL$_f$ specification from (Gracia et al., 2024) (φ_2) and a 15-step safety specification (φ_3). For details of these setups, see (Gracia et al., 2025b).

We compare our approach against (Gracia et al., 2024), in Case Studies $1, 3$, 5-6. 2 is a more challenging version of 1, where W is bigger and the pendulum cannot swing up in one attempt due to control saturation. Similarly, 4 extends 3 with unbounded noise of larger variance and a smaller goal set. Note that (Gracia et al., 2024) relies on ambiguity set learning and cannot handle an unbounded $\boldsymbol{w}$. We also show results obtained using a naïve IMDP abstraction with and without learned support, to show tightness of our approach. Table 1 summarizes our results, highlighting the clear advantages of our approach over (Gracia et al., 2024). Our method significantly reduces sample complexity, often by orders of magnitude, allowing for smaller abstractions while achieving tighter results, thus reducing computation times in most cases. For abstractions of the same size, our synthesis times are typically smaller, sometimes by an order of magnitude. Additionally, our approach produces tighter probabilistic guarantees (less error) than using a naïve IMDP, showcasing the benefits of incorporating additional information into the definition of Γ, albeit with higher computational effort. Furthermore, UMDPs yield a smaller error, less than a half in some case studies, than when the abstraction is just an IMDP with an additional constraint related to learning the support of P_W, showing the benefit of a 2-layer partition. We also compared the performance of Alg. 1 against the

Table 1: Benchmark results. "Approach" indicates the abstraction used: UMDP (Definition 1), "IMDP (Learn Support)" (a UMDP with only the first and third constraints in (4)), and "Naïve IMDP" (traditional IMDP). e_{avg} denotes the average difference in satisfaction probabilities between the lower and upper bounds across all states. Time for abstraction and synthesis is given in minutes, and N_{cluster} denotes the number of noise samples after clustering. The confidence in all case studies is $1 - \alpha = 1 - 10^{-9}$ and $\beta_c = \alpha/2$.

System (Spec.)	Approach	$\|Q\|$	$\|A\|$	N	N_{cluster}	e_{avg}	Abstr. Time	Synth. Time
Pendulum (φ_1)	UMDP	10^4	5	5×10^3	47	0.552	1.796	3.643
Gracia et al. (2024)	UMDP	10^4	5	10^4	44	0.007	1.857	6.679
	UMDP	10^4	5	10^5	47	0.003	1.797	2.515
	Gracia et al. (2024)	4×10^4	5	10^6	49	0	5.273	61.167
Pendulum (φ_1)	UMDP	1.225×10^5	5	5×10^4	172	0.108	130.680	203.817
(Torque-Limited)	UMDP	4×10^4	5	5×10^4	173	0.440	31.328	15.203
	UMDP	4×10^4	5	10^5	179	0.186	30.952	40.241
	UMDP	4×10^4	5	10^6	179	0.048	30.626	29.936
	UMDP	4×10^4	5	10^7	191	0.033	40.520	25.308
	IMDP (Learn Support)	4×10^4	5	10^7	191	0.164	16.714	29.119
$3D$ Unicycle (φ_1)	UMDP	5.932×10^4	10	10^4	235	0.579	34.689	33.156
Gracia et al. (2024)	UMDP	5.932×10^4	10	10^5	325	0.267	45.610	36.18
	UMDP	5.932×10^4	10	10^6	358	0.156	52.733	33.751
	Gracia et al. (2024)	6.4×10^4	10	5×10^8	8869	0.447	457.431	43.342
$3D$ Unicycle (φ_1)	UMDP	7.401×10^4	10	10^5	248	0.5	68.708	302.855
(difficult)	UMDP	5.932×10^4	10	10^6	285	0.241	61.583	200.311
	UMDP	7.401×10^4	10	10^6	285	0.165	75.430	207.516
	IMDP (Learn Support)	7.401×10^4	10	10^6	285	0.166	25.109	21.606
	Naïve IMDP	5.932×10^4	10	10^7	340	0.870	45.865	3.68
	UMDP	5.932×10^4	10	10^7	353	0.170	70.593	178.599
Multiplicative	UMDP	3.6×10^3	1	4.705×10^3	247	0.369	0.156	0.051
noise (φ_1)	IMDP (Learn Support)	3.6×10^3	1	4.371×10^3	247	0.377	0.156	0.011
Skovbekk et al. (2023)	UMDP	3.6×10^3	1	4.725×10^4	306	0.296	0.194	0.064
	UMDP	3.6×10^3	1	4.722×10^5	312	0.258	0.199	0.070
	UMDP	9.22×10^3	1	4.722×10^5	309	0.243	0.594	0.306
	IMDP	9.22×10^3	1	4.722×10^5	309	0.256	0.594	0.063
	Gracia et al. (2024)	10^4	1	4.66×10^5	1066	0.323	13.149	3.863
$2D$ Unicycle (φ_2)	IMDP (Learn Support)	3.6×10^3	8	5×10^3	36	0.135	0.257	4.527
Gracia et al. (2024)	UMDP	3.6×10^3	8	5×10^3	36	0.087	0.257	9.012
	UMDP	3.6×10^3	8	10^4	35	0.038	0.254	7.013
	UMDP	3.6×10^3	8	10^5	41	0.003	0.294	5.622
	UMDP	3.6×10^3	8	10^7	45	0.000	0.321	4.341
	Gracia et al. (2024)	3.6×10^3	8	10^7	46	0.030	0.106	3.043
4-Room Heating (φ_3)	UMDP	2.074×10^4	16	5×10^4	620	0.042	143.869	5.739
Abate et al. (2010)	IMDP (Learn Support)	2.074×10^4	16	5×10^4	620	0.102	129.381	2.99
	UMDP	2.074×10^4	16	5×10^5	753	0.020	258.441	6.127

linear programming solver *Linprog* on an abstraction with $|S| = 1600$ and $|A| = 8$, achieving the same guarantees but reducing the total synthesis time from 2880s to 60s, a reduction of $48\times$.

7. Conclusion

We propose an approach to synthesize strategies for nonlinear stochastic systems with unknown disturbances via abstractions to UMDPs. We also identify pitfalls in the use of naïve abstractions for nonlinear systems and present a synthesis algorithm tailored to our UMDP class. Our extensive case studies show the efficacy and advantages of our framework w.r.t. existing works. In future research we plan to increase the tightness of our results by including additional information into the UMDP and investigate overlapping clusters.

Acknowledgments

I. Gracia and M. Lahijanian are supported in part by National Science Foundation (NSF) under grant number 2039062 and Air Force Research Lab (AFRL) under agreement number FA9453-22-2-0050. L. Laurenti is partially supported by the NWO(grant OCENW.M.22.056).

References

Alessandro Abate, Joost-Pieter Katoen, John Lygeros, and Maria Prandini. Approximate model checking of stochastic hybrid systems. *European Journal of Control*, 16(6):624–641, 2010.

Steven Adams, Morteza Lahijanian, and Luca Laurenti. Formal control synthesis for stochastic neural network dynamic models. *IEEE Control Systems Letters*, 6:2858–2863, 2022.

Thom Badings, Licio Romao, Alessandro Abate, and Nils Jansen. Probabilities are not enough: Formal controller synthesis for stochastic dynamical models with epistemic uncertainty. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 14701–14710, 2023a.

Thom Badings, Licio Romao, Alessandro Abate, David Parker, Hasan A Poonawala, Marielle Stoelinga, and Nils Jansen. Robust control for dynamical systems with non-gaussian noise via formal abstractions. *Journal of Artificial Intelligence Research*, 76:341–391, 2023b.

Thom Badings, Wietze Koops, Sebastian Junges, and Nils Jansen. Learning-based verification of stochastic dynamical systems with neural network policies. *arXiv preprint arXiv:2406.00826*, 2024.

Calin Belta, Antonio Bicchi, Magnus Egerstedt, Emilio Frazzoli, Eric Klavins, and George J Pappas. Symbolic planning and control of robot motion [grand challenges of robotics]. *IEEE Robotics & Automation Magazine*, 14(1):61–70, 2007.

Dimitri Bertsekas and Steven E Shreve. *Stochastic optimal control: the discrete-time case*, volume 5. Athena Scientific, 1996.

Nathalie Cauchi, Luca Laurenti, Morteza Lahijanian, Alessandro Abate, Marta Kwiatkowska, and Luca 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*, pages 240–251, 2019.

Rudi Coppola, Andrea Peruffo, and Manuel Mazo Jr. Data-driven abstractions for verification of deterministic systems. *arXiv preprint arXiv:2211.01793*, 2022.

Rudi Coppola, Andrea Peruffo, and Manuel Mazo. Data-driven abstractions for verification of linear systems. *IEEE Control Systems Letters*, 7:2737–2742, 2023.

Giuseppe De Giacomo and Moshe Y Vardi. Linear temporal logic and linear dynamic logic on finite traces. In *IJCAI'13 Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, pages 854–860. Association for Computing Machinery, 2013.

Laurent El Ghaoui and Arnab Nilim. Robust solutions to markov decision problems with uncertain transition matrices. *Operations Research*, 53(5):780–798, 2005.

Robert Givan, Sonia Leach, and Thomas Dean. Bounded-parameter markov decision processes. *Artificial Intelligence*, 122(1-2):71–109, 2000.

Ibon Gracia, Dimitris Boskos, Morteza Lahijanian, Luca Laurenti, and Manuel Mazo Jr. Distributionally robust strategy synthesis for switched stochastic systems. *arXiv preprint arXiv:2212.14260*, 2022.

Ibon Gracia, Dimitris Boskos, Luca Laurenti, and Morteza Lahijanian. Data-driven strategy synthesis for stochastic systems with unknown nonlinear disturbances. *arXiv preprint arXiv:2406.09704*, 2024.

Ibon Gracia, Dimitris Boskos, Morteza Lahijanian, Luca Laurenti, and Manuel Mazo Jr. Efficient strategy synthesis for switched stochastic systems with distributional uncertainty. *Nonlinear Analysis: Hybrid Systems*, 55:101554, 2025a.

Ibon Gracia, Luca Laurenti, Manuel Mazo Jr., Alessandro Abate, and Morteza Lahijanian. Supplementary material for "temporal logic control for nonlinear stochastic systems under unknown disturbances". https://mortezalahijanian.com/papers/L4DC2025.pdf, 2025b. Accessed: 2024-12-01.

Garud N Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 30(2): 257–280, 2005.

John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. Strategy synthesis for partially-known switched stochastic systems. In *Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control*, pages 1–11, 2021.

Milad Kazemi, Rupak Majumdar, Mahmoud Salamati, Sadegh Soudjani, and Ben Wooding. Data-driven abstraction-based control synthesis. *Nonlinear Analysis: Hybrid Systems*, 52:101467, 2024.

Morteza Lahijanian, Sean B Andersson, and Calin Belta. Formal verification and synthesis for discrete-time stochastic systems. *IEEE Transactions on Automatic Control*, 60(8):2031–2045, 2015.

Abolfazl Lavaei, Sadegh Soudjani, Alessandro Abate, and Majid Zamani. Automated verification and synthesis of stochastic hybrid systems: A survey. *Automatica*, 146:110617, 2022.

Mathias Lechner, D. Žikelić, Krishnendu Chatterjee, and Thomas A Henzinger. Stability verification in stochastic control systems via neural network supermartingales. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 7326–7336, 2022.

Frederik Baymler Mathiesen, Licio Romao, Simeon C Calvert, Alessandro Abate, and Luca Laurenti. Inner approximations of stochastic programs for data-driven stochastic barrier function design. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 3073–3080. IEEE, 2023.

Rayan Mazouz, Frederik Baymler Mathiesen, Luca Laurenti, and Morteza Lahijanian. Piecewise stochastic barrier functions. *arXiv preprint arXiv:2404.16986*, 2024.

Ali Salamati, Abolfazl Lavaei, Sadegh Soudjani, and Majid Zamani. Data-driven verification and synthesis of stochastic systems through barrier certificates. *arXiv preprint arXiv:2111.10330*, 2021.

Oliver Schön, Birgit van Huijgevoort, Sofie Haesaert, and Sadegh Soudjani. Bayesian approach to temporal logic control of uncertain systems. *arXiv preprint arXiv:2304.07428*, 2023.

John Skovbekk, Luca Laurenti, Eric Frew, and Morteza Lahijanian. Formal abstraction of general stochastic systems via noise partitioning. *IEEE Control Systems Letters*, 2023.

Roberto Tempo, Giuseppe Calafiore, Fabrizio Dabbene, et al. *Randomized algorithms for analysis and control of uncertain systems: with applications*, volume 7. Springer, 2013.

Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. Robust markov decision processes. *Mathematics of Operations Research*, 38(1):153–183, 2013.

Eric M Wolff, Ufuk Topcu, and Richard M Murray. Robust control of uncertain markov decision processes with temporal logic specifications. In *2012 IEEE 51st IEEE Conference on decision and control (CDC)*, pages 3372–3379. IEEE, 2012.