To appear in the 2021 Proc. of the ACM Int. Conf. on Hybrid Systems: Computation and Control (HSCC 2021)

Strategy Synthesis for Partially-known Switched Stochastic Systems

John Jackson
john.m.jackson@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

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

Eric Frew
eric.frew@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

Morteza Lahijanian
morteza.lahijanian@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

ABSTRACT

We present a data-driven framework for strategy synthesis for partially-known switched stochastic systems. The properties of the system are specified using *linear temporal logic (LTL) over finite traces* (LTL$_f$), which is as expressive as LTL and enables interpretations over finite behaviors. The framework first learns the unknown dynamics via Gaussian process regression. Then, it builds a formal abstraction of the switched system in terms of an uncertain Markov model, namely an *Interval Markov Decision Process (IMDP)*, by accounting for both the stochastic behavior of the system and the uncertainty in the learning step. Then, we synthesize a strategy on the resulting IMDP that maximizes the satisfaction probability of the LTL$_f$ specification and is robust against all the uncertainties in the abstraction. This strategy is then refined into a switching strategy for the original stochastic system. We show that this strategy is near-optimal and provide a bound on its distance (error) to the optimal strategy. We experimentally validate our framework on various case studies, including both linear and non-linear switched stochastic systems.

CCS CONCEPTS

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

KEYWORDS

Switched stochastic systems, Gaussian process regression, Formal synthesis, Safe autonomy, Uncertain Markov decision processes

ACM Reference Format:
John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. 2021. Strategy Synthesis for Partially-known Switched Stochastic Systems. In *24th ACM International Conference on Hybrid Systems: Computationand Control (HSCC '21), May 19–21, 2021, Nashville, TN, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3447928.3456649

1 INTRODUCTION

Switched stochastic systems are a class of *stochastic hybrid systems* (SHSs) that provide a powerful framework for modeling complex real-world systems. They consist of a finite set of stochastic processes that capture the uncertainty in the evolution of the underlying system with the ability to switch between these processes, representing control options. These models are employed in numerous application domains such as robotics [25], biological systems [19], and cyber-physical systems [15]. Many of the applications are in *safety-critical* domains and require formal analysis of the underlying system. Existing formal approaches to analysis and synthesis of SHSs are model based, and the resulting guarantees apply only to the model of the system. In reality, the true model of the system is often partially-known due to, e.g., the use of black-box controllers, or the lack of a closed-form analytical representation. This poses a major challenge for formal reasoning, which also relates to the classical question of *how to extend formal guarantees from models to systems?* This work investigates a data-driven approach to address this challenge.

Formal verification and synthesis for SHSs has been well studied in recent years, e.g., [8, 11, 21–23, 30]. The proposed approaches can be generally divided into two categories. One is a set of approaches based on numerical analysis of stochastic differential (difference) equations with asymptotic guarantees in terms of weak convergence [21]. The other set of approaches is based on a finite *abstraction* of the SHS to a Markov process, and their formal guarantees are on probabilistic satisfaction of temporal logic specifications, namely *linear temporal logic* (LTL) and *probabilistic computation tree logic* (PCTL) [4]. Despite the recent advances, both categories of approaches assume that the SHS model is fully known and perfectly represents the underlying system. This assumption, however, is often violated, especially in modern systems where AI-modules are increasingly employed as black-box components.

A few recent studies focus on dealing with unknown dynamical systems, e.g., [2, 12, 16, 20]. The proposed approaches are based on data-driven methods and assume some knowledge on the system. Work [16, 20] impose a strong assumption that the underlying system is linear. Then, they employ techniques such as Bayesian

HSCC '21, May 19–21,2021, Nashville, TN, USA

ACM ISBN 978-1-4503-8339-4/21/05…$15.00
https://doi.org/10.1145/3447928.3456649

inference and chance-constrained optimization to provide probabilistic guarantees for the unknown system from a finite set of data. Work [2] relaxes the linearity assumption and proposes approximation of the unknown dynamics through a piecewise-polynomial function. Then, the safety of the system is assessed through barrier certificates. While this method can deal with a more general class of systems, it is unclear how the guarantees can be extended to the underlying system.

An effective method to deal with unknown dynamics in safety-critical applications is *Gaussian process* (GP) regression [29]. The advantage of GP regression is in its ability to quantify the bound on the uncertainty in the learning process as derived in [9, 13, 24, 31]. This has led to an increased use of GPs in safe learning frameworks, e.g., [3, 5, 18, 28, 33]. In most work, the main objective is to learn safe policies via reinforcement learning with the exception of work [18], which considers a safety verification problem of unknown systems with noisy measurements. The proposed framework uses GPs to construct a Markov abstraction for an invariant set (safety) analysis from a noisy dataset. In all these work, the assumption is that the underlying system is deterministic and the specification is simple, whereas the focus here is on unknown stochastic systems with complex specifications.

This work presents a formal synthesis framework for stochastic systems with partially-known models in the form of switched stochastic processes. The framework is able to provide formal guarantees on the behavior of the underlying system from a set of data. The specification language is *LTL over finite traces* (LTL$_f$) [10], which has the same expressively as LTL, but the interpretations of its formulas are over finite behaviors making it an appropriate language for highly uncertain (unknown) systems such as those considered here. The approach is based on finite abstraction and employs GP regression for its construction. Given a set of data, the framework first learns the unknown dynamics using GP regression. Then, an abstraction is constructed in the form of an uncertain Markov process, namely *interval Markov decision process* (IMDP) using the known and learned dynamics as well as the errors bounds of the learning process. Given an LTL$_f$ property, a strategy is computed on the abstraction that maximizes the probability of satisfaction of the property and is robust against all the errors introduced in the learning and abstraction steps. This not only results in a switching (control) strategy for the underlying system, but it also provides a lower bound probability for the satisfaction of the LTL$_f$ property for every initial state.

The main contribution of this work is a theoretical and computational framework for control synthesis for partially-known stochastic systems from a given set of data. This work shows a method of harvesting the power of machine learning techniques, in particular GP regression, in a formal synthesis framework. Unlike classical model-based approaches, this framework enables the extension of the formal guarantees to the underlying system. This is achieved by formally incorporating both the uncertainty related to the stochastic behavior of the system and the uncertainty related to the partial knowledge of the system in the abstraction, and then accounting for these uncertainties in generating a robust switching strategy. As a result, this framework allows for synthesis for complex systems from simplified (low-fidelity) models, i.e., linearized models; hence, enabling the use of rich and matured techniques for simple (linear) models in control design for complex systems. Furthermore, this paper presents derivations for probabilistic bounds for the transition probabilities of the IMDP abstraction as well as proofs of correctness for the methodology. Finally, the synthesis framework is demonstrated through a series of case studies on unknown stochastic systems with both linear and nonlinear dynamics with various LTL$_f$ specifications.

2 PROBLEM FORMULATION

Consider a partially-known switched stochastic process as described below:

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

where $k \in \mathbb{N}$, $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{u}_k \in U$, and $U = \{1, ..., m\}$ is a finite set of *modes* or *actions*. For every $u \in U$, $f_u : \mathbb{R}^n \to \mathbb{R}^n$ is a (known a-priori) continuous function and $g_u : \mathbb{R}^n \to \mathbb{R}^n$ is a possibly nonlinear continuous function representing the unknown dynamics of Process (1). The noise term $\mathbf{v}_k$ is a random variable with an independent and stationary θ-sub-Gaussian distribution p_v. This class of distributions are those whose tails decay at least as fast as a Gaussian with variance θ^2, including all distributions with bounded support the Gaussian distribution itself [26].

Intuitively, $\mathbf{x}_k$ is a stochastic process driven by the noise process $\mathbf{v}_k$, where some or all the dynamics are unknown in each mode, and $\mathbf{u}_k$ indicates the current mode (and hence switching between the modes). Process (1) is a rich model that allows for the inclusion of modeling errors in addition to noise. For instance, consider a nonlinear noisy control system with a finite set of controls U. If only a linear approximate model of the system is available, then Process (1) can be used to represent it, where f_u becomes the approximate linear model of the system and g_u is all the higher-order dynamics that are not modelled under each controller u.

We assume to have a collection of state-action-state measurements $\mathrm{D} = \{(x_i, u_i, x_i^+)_{i=1}^m\}$ generated by Process (1), where $x_i^+ \in \mathbb{R}^n$ is a sample of one-step evolution of Process (1) when it is initialized at $x_i \in \mathbb{R}^n$ in mode $u_i \in U$. Our goal is to use D to learn g_u for each $u \in U$. In order to achieve this correctly, we need an assumption on the regularity of g_u. The following assumption suffices to guarantee that g_u can be learned arbitrarily well via GP regression.

Assumption 1. *For a compact set $X \subset \mathbb{R}^n$, let $\kappa : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{>0}$ be a given kernel and $\mathcal{H}_\kappa(X)$ the reproducing kernel Hilbert space (RKHS) of functions over X corresponding to κ with norm $\|\cdot\|_\kappa$ [31]. Then, for each $u \in U$ and $i \in \{1, ..., n\}$, $g_u^{(i)}(\cdot) \in \mathcal{H}_\kappa(X)$ and for a constant $B_i > 0$, it holds that $\|g_u^{(i)}(\cdot)\|_\kappa \leq B_i$, where $g_u^{(i)}$ is the i-th component of g_u.*

Assumption 1 is a standard assumption [18, 31], which is intimately related to the continuity of g_u, as discussed in Section 4. For instance, assuming that κ is the widely used squared exponential kernel, we obtain that $\mathcal{H}_\kappa(X)$ is a space of functions that is dense with respect to the set of continuous functions on a compact set $X \subset \mathbb{R}^n$, i.e., members of $\mathcal{H}_\kappa(X)$ can approximate any continuous function on X arbitrarily well [32].

Let $\omega_{\mathbf{x}} = x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} \ldots$ be a sample path (trajectory) of Process (1) and denote by $\omega_{\mathbf{x}}(k) = x_k$, the state of $\omega_{\mathbf{x}}$ at time k. Further, we denote by $\Omega_{\mathbf{x}}^{\text{fin}}$ the set of all sample paths with finite length, i.e, the set of trajectories $\omega_{\mathbf{x}}^k = x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} \ldots \xrightarrow{u_{k-1}} x_k$

for all $k \in \mathbb{N}$. With a slight abuse of notation, given path $\omega_{\mathbf{x}}$, we denote by $\omega_{\mathbf{x}}^k$ the prefix of length $k+1$ of $\omega_{\mathbf{x}}$.

Given a finite path, a *switching strategy* chooses the mode (action) of Process (1).

Definition 1 (Switching Strategy). *A switching strategy $\pi_{\mathbf{x}} : \Omega_{\mathbf{x}}^{fin} \to U$ is a function that maps a finite path $\omega_{\mathbf{x}}^k \in \Omega_{\mathbf{x}}^{fin}$ to a mode (action) $u \in U$. The set of all switching strategies is denoted by $\Pi_{\mathbf{x}}$.*

For $u \in U$, a Borel measurable set $X \subseteq \mathbb{R}^n$, and $x \in \mathbb{R}^n$, call

$$T^u(X \mid x) = \int \mathbf{1}_X(f_u(x) + g_u(x) + v)p_v(\bar{v})d\bar{v},$$

the stochastic transition function induced by Process (1) in mode $u \in U$, where

$$\mathbf{1}_X(x) = \begin{cases} 1 & \text{if } x \in X \\ 0 & \text{otherwise} \end{cases}$$

is the indicator function. From the definition of $T^u(X \mid x)$ it follows that, given a strategy $\pi_{\mathbf{x}}$, for a time horizon $[0, N]$, Process (1) defines a stochastic process on the canonical space $\Omega = (\mathbb{R}^n)^{N+1}$ with the Borel σ−algebra $\mathcal{B}(\Omega)$ induced by the product topology and with the unique probability measure P generated by $T^{\pi_{\mathbf{x}}}$ and a (fixed) initial condition $x_0 \in \mathbb{R}^n$ such that for $k \in \{1, ..., N\}$

$$P[\omega_{\mathbf{x}}^N(0) \in X] = \mathbf{1}_X(x_0),$$
$$P[\omega_{\mathbf{x}}^N(k) \in X \mid \omega_{\mathbf{x}}^N(k-1) = x, \pi_{\mathbf{x}}] = T^{\pi_{\mathbf{x}}(\omega_{\mathbf{x}}^{k-1})}(X \mid x).$$

Furthermore, for $N = \infty$, P is still uniquely defined by T^u by the *Ionescu-Tulcea extension theorem* [1].

In this paper, we are interested in the properties of Process (1) in a compact set $X \subset \mathbb{R}^n$. Specifically, we analyze the behavior of Process (1) with respect to a finite set of closed regions of interest $R = \{r_1, \ldots, r_{|R|}\}$, where $r_i \subseteq X$. To this end, we associate to each region r_i the atomic proposition $\mathfrak{p}_i$ such that $\mathfrak{p}_i = \top$ (i.e., $\mathfrak{p}_i$ is *true*) if $x \in r_i$; otherwise $\mathfrak{p}_i = \bot$ (i.e., $\mathfrak{p}_i$ is *false*). Let $AP = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_{|R|}\}$ denote the set of all atomic propositions and $L : X \to 2^{AP}$ be the labeling function that assigns to state x the set of atomic propositions that are true at x, i.e.,

$$\mathfrak{p}_i \in L(x) \quad \Leftrightarrow \quad x \in r_i.$$

Then, we define the *trace* or *observation* of path $\omega_{\mathbf{x}}^k = x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} \ldots \xrightarrow{u_{k-1}} x_k$ to be

$$\rho = \rho_0\rho_1 \ldots \rho_k,$$

where $\rho_i = L(x_i) \in 2^{AP}$ for all $i \leq k$. With an abuse of notation we use $L(\omega_{\mathbf{x}}^k)$ to denote the trace of $\omega_{\mathbf{x}}^k$.

2.1 Linear temporal logic on finite traces (LTL$_f$)

In this work, we are interested in the temporal properties of Process (1) with respect to the regions of interest in R. To express such properties, *linear temporal logic* (LTL) [4] is a popular choice of language given its rich expressivity and intuitive formalism. Here, we employ *LTL interpreted over finite traces* (LTL$_f$) [10], which has the same syntax as LTL but its semantics is defined over finite traces.

Definition 2 (LTL$_f$ Syntax). *An LTL$_f$ formula φ is built from a set of atomic propositions AP and is closed under the Boolean connectives as well as the "next" operator $\mathcal{X}$ and the "until" operator $\mathcal{U}$:*

$$\varphi ::= \top \mid \mathfrak{p} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathcal{X}\varphi \mid \varphi\mathcal{U}\varphi$$

where $\mathfrak{p} \in AP$, $\top$ is "true" or a tautology, and $\neg$ and $\wedge$ are the "negation" and "and" operators in Boolean logic, respectively.

The common temporal operators "eventually" ($\mathcal{F}$) and "globally" ($\mathcal{G}$) are defined as:

$$\mathcal{F}\,\varphi = \top\,\mathcal{U}\,\varphi \quad \text{and} \quad \mathcal{G}\,\varphi = \neg\mathcal{F}\,\neg\varphi.$$

The semantics of LTL$_f$ is defined as follows.

Definition 3 (LTL$_f$ Semantics). *The semantics of an LTL$_f$ formula φ are defined over finite traces in AP^*. The set of all finite traces is $(2^{AP})^*$. Let $|\rho|$ denote the length of trace ρ and ρ_i be the i-th symbol of ρ. Further, $\rho, i \models \varphi$ is read as: "the i-th step of trace ρ is a model of φ." Then,*

- $\rho, i \models \top$,
- $\rho, i \models \mathfrak{p}$ *iff* $\mathfrak{p} \in \rho_i$,
- $\rho, i \models \neg\varphi$ *iff* $\rho, i \not\models \varphi$,
- $\rho, i \models \varphi_1 \wedge \varphi_2$ *iff* $\rho, i \models \varphi_1$ *and* $\rho, i \models \varphi_2$,
- $\rho, i \models \mathcal{X}\varphi$ *iff* $|\rho| > i+1$ *and* $\rho, i+1 \models \varphi$,
- $\rho, i \models \varphi_1\mathcal{U}\varphi_2$ *iff* $\exists j$ *s.t.* $i \leq j < |\rho|$ *and* $\rho, j \models \varphi_2$ *and* $\forall k$, $i \leq k < j$, $\rho, k \models \varphi_1$.

Finite trace ρ satisfies φ, denoted by $\rho \models \varphi$, if $\rho, 0 \models \varphi$.

An LTL$_f$ formula φ defines a language $\mathcal{L}(\varphi)$ over the alphabet 2^{AP}. $\mathcal{L}(\varphi)$ is a regular language, more specifically,

$$\mathcal{L}(\varphi) = \{\rho \in (2^{AP})^* \mid \rho \models \varphi\}.$$

Given compact set $X \subset \mathbb{R}^n$, its set of regions of interest R and the corresponding set of atomic propositions AP, and an LTL$_f$ formula φ defined over AP, as in [34], we say that path $\omega_{\mathbf{x}}$ of Process (1) satisfies φ if there exists a prefix of $\omega_{\mathbf{x}}$ that is in the language of φ and lies entirely in X, i.e.,

$$\omega_{\mathbf{x}} \models \varphi \quad \Leftrightarrow \quad \exists k \in \mathbb{N} \;\; s.t. \;\; L(\omega_{\mathbf{x}}^k) \in \mathcal{L}(\varphi) \text{ and}$$
$$\omega_{\mathbf{x}}^k(k') \in X \;\; \forall k' \leq k, \quad (2)$$

where $L(\omega_{\mathbf{x}}^k) \in \left(2^{AP}\right)^*$ is the trace (observation) of $\omega_{\mathbf{x}}^k$.

2.2 Problem Formulation

The ideal goal of this work is, given an LTL$_f$ formula φ, to synthesize a switching strategy $\pi_{\mathbf{x}}^*$ such that under $\pi_{\mathbf{x}}^*$ the probability of the paths of Process (1) that satisfy φ is maximized. Nevertheless, we should stress that, in general, the partial knowledge of Process (1) and the limited amount of data available (not controllable a-priori) make it infeasible to find a switching strategy that maximizes such a probability. Hence, in Problem 1 we seek a near-optimal strategy such that, under this switching strategy, Process (1) is guaranteed to satisfy φ with a (high) probability greater than a given threshold with a quantifiable distance from the optimal probability.

Problem 1 (Switching Strategy Synthesis). *Given a partially-known switched stochastic system as defined in Process* (1), *a dataset* D, *a compact set X, an LTL$_f$ property φ defined over the regions* *of interest in X, and a probability threshold $\bar{p}$, find a near-optimal switching strategy $\pi_{\boldsymbol{x}}^{\varepsilon}$ that determines whether for every $x_0 \in X$*

$$P[\omega_{\boldsymbol{x}} \models \varphi \mid \pi_{\boldsymbol{x}}^{\varepsilon}, \omega_{\boldsymbol{x}}(0) = x_0] \geq \bar{p},$$

and quantify the corresponding error $\varepsilon_{x_0} \geq 0$ with respect to the optimal switching strategy, i.e.,

$$|P[\omega_{\boldsymbol{x}} \models \varphi \mid \pi_{\boldsymbol{x}}^{\varepsilon}, \omega_{\boldsymbol{x}}(0) = x_0] - p^*(x_0)| \leq \varepsilon_{x_0},$$

where $p^*(x_0) = \max_{\pi_{\boldsymbol{x}} \in \Pi_{\boldsymbol{x}}} P[\omega_{\boldsymbol{x}} \models \varphi \mid \pi_{\boldsymbol{x}}, \omega_{\boldsymbol{x}}(0) = x_0]$.

Overview of the Approach. In order to solve Problem 1 we rely on GP regression and Assumption 1 to find a function $\hat{g}_u$ such that with high probability, $|\hat{g}_u(x) - g_u(x)| \leq \epsilon_u$ for all $x \in X$ and a given $\epsilon_u > 0$. We then use $\hat{g}_u$ to build an abstraction of Process (1) in terms of a finite Markov model, where the stochastic nature of Process (1), the error in employing $\hat{g}_u$ instead of g_u, and the error corresponding to the discretization of space are all formally modelled as uncertainty. We then synthesize an optimal strategy for the resulting Markov model that maximizes the probability that the paths of the Markov model satisfy φ and is robust against the uncertainties. Finally, we derive upper and lower bounds on the probability that Process (1) satisfies φ under this strategy.

3 PRELIMINARIES

3.1 Gaussian Process Regression

Gaussian Process (GP) regression is a non-parametric Bayesian machine learning method [29]. For an unknown function $\mathrm{g} : \mathbb{R}^n \to \mathbb{R}$, the basic assumption of GP regression is that g is a sample from a GP with covariance $\kappa : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{>0}$. Consider a dataset of samples $\mathrm{D} = \{(\mathrm{x}_i, \mathrm{y}_i), i \in \{1, \ldots, m\}\}$, where y_i is a sample of an observation of $\mathrm{g}(\mathrm{x}_i)$ with independent zero-mean noise v, which is assumed to be normally distributed with variance σ^2. Let X and Y be ordered vectors with all points in D such that $\mathrm{X}_i = \mathrm{x}_i$ and $\mathrm{Y}_i = \mathrm{y}_i$. Further, call $K(\mathrm{X}, \mathrm{X})$ the matrix with $K_{i,j}(\mathrm{X}_i, \mathrm{X}_j) = \kappa(\mathrm{x}_i, \mathrm{x}_j)$, $K(\mathrm{x}, \mathrm{X})$ the vector such that $K_i(\mathrm{x}, \mathrm{X}) = \kappa(\mathrm{x}, \mathrm{X}_i)$, and $K(\mathrm{X}, \mathrm{x})$ defined accordingly. Then, the predictive distribution of g at a test point x is given by the conditional distribution of g given D, which is Gaussian and with mean μ_{D} and variance σ_{D}^2 given by

$$\mu_{\mathrm{D}}(\mathrm{x}) = K(\mathrm{x}, \mathrm{X})\big(K(\mathrm{X}, \mathrm{X}) + \sigma^2 I_m\big)^{-1} Y \tag{3}$$

$$\sigma_{\mathrm{D}}^2(\mathrm{x}) = \kappa(\mathrm{x}, \mathrm{x}) - K(\mathrm{x}, \mathrm{X})\big(K(\mathrm{X}, \mathrm{X}) + \sigma^2 I_m\big)^{-1} K(\mathrm{X}, \mathrm{x}), \tag{4}$$

where I_m is the identity matrix of size $m \times m$.

3.2 Interval Markov Decision Processes

We use a generalization of Markov decision processes to abstract the system. An *interval Markov decision process* (IMDP), also called bounded-parameter Markov decision process, uses interval-valued transition probabilities [14, 17].

Definition 4 (IMDP). *An interval Markov decision process (IMDP) is a tuple $\mathcal{I} = (Q, A, \check{P}, \hat{P}, AP, L)$, where*

- *Q is a finite set of states,*
- *A is a finite set of actions, and $A(q)$ denotes the set of available actions at state $q \in Q$.*
- *$\check{P} : Q \times A \times Q \to [0, 1]$ is a function, where $\check{P}(q, a, q')$ defines the lower bound of the transition probability from state $q \in Q$ to state $q' \in Q$ under action $a \in A(q)$,*
- *$\hat{P} : Q \times A \times Q \to [0, 1]$ is a function, where $\hat{P}(q, a, q')$ defines the upper bound of the transition probability from state $q \in Q$ to state $q' \in Q$ under action $a \in A(q)$,*
- *AP is a finite set of atomic propositions,*
- *$L : Q \to 2^{AP}$ is a labeling function that assigns to each state $q \in Q$ a subset of AP.*

For all $q, q' \in Q$ and $a \in A(q)$, it holds that $\check{P}(q, a, q') \leq \hat{P}(q, a, q')$ and $\sum_{q' \in Q} \check{P}(q, a, q') \leq 1 \leq \sum_{q' \in Q} \hat{P}(q, a, q')$.

A path of an IMDP is a sequence of states $\omega_{\mathcal{I}} = q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \xrightarrow{a_2} \ldots$ such that $a_k \in A(q_k)$ and $\hat{P}(q_k, a_k, q_{k+1}) > 0$ for all $k \in \mathbb{N}$. We denote the last state of a finite path $\omega_{\mathcal{I}}^{\mathrm{fin}}$ by $last(\omega_{\mathcal{I}}^{\mathrm{fin}})$ and the set of all finite and infinite paths by $Paths^{\mathrm{fin}}$ and $Paths$, respectively. The actions are chosen according to a strategy π which is defined below.

Definition 5 (Strategy). *A strategy π of an IMDP model $\mathcal{I}$ is a function $\pi : Paths^{\mathrm{fin}} \to A$ that maps a finite path $\omega_{\mathcal{I}}^{\mathrm{fin}}$ of $\mathcal{I}$ onto an action in $A(last(Paths^{\mathrm{fin}}))$. If a strategy depends only on $last(\omega_{\mathcal{I}}^{\mathrm{fin}})$, it is called a memoryless (stationary) strategy. The set of all strategies is denoted by Π.*

Given an arbitrary strategy π, we are restricted to the set of interval Markov chains defined by the transition probability intervals induced by π. In order to reduce this to a Markov chain, we define the adversary function, which assigns a transition probability distribution at each state.

Definition 6 (Adversary). *For an IMDP $\mathcal{I}$, an adversary is a function $\xi : Paths^{\mathrm{fin}} \times A \to \mathcal{D}(Q)$ that, for each finite path $\omega_{\mathcal{I}}^{\mathrm{fin}} \in Paths^{\mathrm{fin}}$, state $q = last(\omega_{\mathcal{I}}^{\mathrm{fin}})$, and action $a \in A(last(\omega_{\mathcal{I}}^{\mathrm{fin}}))$, assigns a feasible distribution γ_q^a which satisfies*

$$\check{P}(q, a, q') \leq \gamma_q^a(q') \leq \hat{P}(q, a, q').$$

The set of all adversaries is denoted by Ξ.

Under a strategy and a valid adversary, the IMDP collapses to a Markov chain and induces a probability measure on its paths. We use this measure as an optimization objective for synthesizing a desirable strategy.

4 IMDP ABSTRACTION

In order to solve Problem 1, we start by abstracting Process (1) into an IMDP $\mathcal{I} = (Q, A, \check{P}, \hat{P}, AP, L)$. Below we describe how we obtain Q, A, AP, and L. Then, in Section 4.2 we consider upper and lower bounds for the transition probabilities.

4.1 States and Actions

The set of states Q of $\mathcal{I}$ is obtained by discretizating the compact set X. This discretization needs to respect the set of regions of interest $R = \{r_1, ..., r_{|R|}\}$. To achieve this, we first construct a set of non-overlapping regions of interests R' from R such that

$$\cup_{r' \in R'} r' = R \quad \text{and} \quad r_i' \cap r_j' = \emptyset \quad \forall r_i', r_j' \in R' \text{ and } r_i' \neq r_j'.$$

Then, we partition each r_i' into a set of cells (regions) that are non-overlapping. Next, we partition the remainder of the compact set

$(X \setminus R)$ to a set of non-overlapping cells. Let $Q_s = \{q_1, ..., q_{|Q_s|}\}$ denote the resulting set of all cells (include R'). Then, by construction, it holds that

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

Each region is associated to a state of IMDP $\mathcal{I}$. With an abuse of notation, q denotes both the region, i.e., $q \subset X$, as well as its corresponding IMDP state, i.e, $q \in Q$. From the context, the correct interpretation of q should be clear. Furthermore, let q_u denote the remainder of the state space, i.e., $\mathbb{R}^n \setminus X$. Then, the set of states of $\mathcal{I}$ is defined as

$$Q = Q_s \cup \{q_u\}.$$

The set of actions A of $\mathcal{I}$ is given by the set of modes U, i.e., $A = U$, and all actions are available at each state of $\mathcal{I}$, i.e., $A(q) = A$ for all $q \in Q$.

The set of atomic propositions AP is the same as those defined over X. With an abuse of notation, we define the IMDP labeling function $L : Q \to 2^{AP}$ with $L(q) = L(x)$ for any choice of $x \in q$. Note that, because the discretization respects the regions of interests, the labels of the points in a discrete cell are necessarily the same, i.e., $L(x) = L(x')$ for all $x, x' \in q$.

4.2 Transition Probability Bounds

In order to compute the transition probability bounds $\check{P}$ and $\hat{P}$ for all $q, q' \in Q$ and $u \in A = U$, we need to derive the following bounds:

$$\check{P}(q, u, q') \leq \min_{x \in q} T^u(q' \mid x), \tag{5}$$

$$\hat{P}(q, u, q') \geq \max_{x \in q} T^u(q' \mid x). \tag{6}$$

However, without any knowledge about g_u in Process (1), the computation of such quantities is infeasible. In what follows we show how we can employ the data in D and GP regression to compute a function $\hat{g}_u$ such that for any $x \in X$, $g_u(x)$ and $\hat{g}_u(x)$ are provably close.

4.2.1 Regression Approach. In our setting, data in D are samples (x, u, x^+) of Process (1) such that

$$x^+ = f_u(x) + g_u(x) + v,$$

where both x and u are known and v is a sample from the noise process $\mathbf{v}$. From this we can obtain a measurement of g_u by simply noticing that f_u is known, i.e., we obtain a dataset composed by:

$$y^+ = x^+ - f_u(x) = g_u(x) + v, \tag{7}$$

where y^+, x^+, x, u are all known. Note that, in our setting, we make no assumptions on the fact that g_u is a sample from a given GP. Furthermore, the noise $\mathbf{v}_k$ is not necessarily Gaussian for any $k \in \mathbb{N}$. As a result, the assumptions for GP regression discussed in Section 3.1 are not satisfied and we cannot directly use its prediction to make probabilistic statements over g_u. Nevertheless, thanks to Assumption 1 we can rely on the properties of the RKHS space generated by κ to bound the regression error even in our more agnostic setting.

In particular, for each $g_u : \mathbb{R}^n \to \mathbb{R}^n$, we use n independent GPs to learn $g_u^{(i)}$, the i-th component of g_u. Then, for a given mode u, we consider $\hat{g}_u^{(i)} = \mu_{\mathrm{D}}$, where μ_{D} is the posterior mean of the GP as described in (3). We use the following Lemma from [9] to characterize the error in employing $\hat{g}_u$ instead of g_u.

Lemma 1 ([9], Theorem 2). *Let X be a compact set, $\delta \in (0, 1)$, γ_κ^m the maximum information gain parameter associated with κ and dataset* D *of m training points, and $B_i > 0$ such that $\|g_u^{(i)}\|_\kappa \leq B_i$. Assume that $\boldsymbol{v}$ is θ-sub-Gaussian and μ_{D} and σ_{D} are found by setting $\sigma = 1 + 2/m$. Define $\beta = (\theta/\sqrt{\sigma})\big(B_i + \theta\sqrt{2(\gamma_\kappa^m + 1 + \log 1/\delta)}\big)$. Then, it holds that*

$$P\big[\forall x \in X,\ |\mu_D(x) - g_u^{(i)}(x)| < \beta\sigma_D(x)\big] \geq 1 - \delta. \tag{8}$$

One challenge in employing Lemma 1 is in determining the values (or bounds) for the information gain constant γ_κ^m and the RKHS constant B_i. A procedure for obtaining γ_κ^m is given in [31]. The RKHS constant B_i is instead intimately related to the continuity of g_u, as shown in Theorem 3.11 of [27], where a bound of B_i in terms of the maximum value that g_u obtains in X and the kernel κ is given.

4.2.2 Transitions within Q_s. For all states $q, q' \in Q_s$, the transition probability bounds in (5) are given by Theorem 1 below. In order to state this result, we first need to introduce the notions of expansion and reduction of a closed set.

Definition 7 (Expansion and Reduction of a Set). *Given a compact set $q \subset \mathbb{R}^n$ and a set of n scalars $c = \{c_1, \ldots, c_n\}$, where $c_i \geq 0$, the expansion of q by c is defined as*

$$\overline{q}(c) = \{x \in \mathbb{R}^n \mid \exists x_q \in q \ \ s.t.\ \ |x_q^{(i)} - x^{(i)}| \leq c_i \ \ \forall i = \{1, \ldots, n\}\},$$

and the reduction of q by c is

$$\underline{q}(c) = \{x_q \in q \mid \forall x_{\partial q} \in \partial q,\ \ |x_q^{(i)} - x_{\partial q}^{(i)}| > c_i \ \ \forall i = \{1, \ldots, n\}\},$$

where ∂q denotes the boundary of q.

In addition, we define the image of region q under the learned dynamics by

$$Im(q) = \{f_u(x) + \hat{g}_u(x) \mid x \in q\}$$

and the intersection indicator function as

$$\mathbf{1}_V^W = \begin{cases} 1 & \text{if } V \cap W \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

for arbitrary sets V and W. We can now bound the transition probabilities between the IMDP states in Q_s.

Theorem 1. *Let $\|h\|_\infty^q \equiv \sup_{x \in q} |h(x)|$. Given an action (mode) $u \in A$, regions $q, q' \in Q_s$, dataset* D*, regression $\hat{g}_u$, and positive real vectors $\epsilon \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^n$, it holds that*

$$\begin{aligned} &\max_{x \in q} T^u(q' \mid x) \\ &\qquad \leq \mathbf{1}_{\overline{q}'(\epsilon+\eta)}^{Im(q)} \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i] \prod_{i=1}^{n} P[|v^{(i)}| \leq \eta_i] \\ &\qquad\quad + \prod_{i=1}^{n} (1 - P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i]), \end{aligned} \tag{9}$$

$$\min_{x \in q} T^u(q' \mid x) \geq$$
$$\mathbf{1}^{Im(q)}_{X \setminus \underline{q}'(\epsilon+\eta)} \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i] \prod_{i=1}^{n} P[|v^{(i)}| \leq \eta_i]. \tag{10}$$

Proof. Let $\|g_u - \hat{g}_u\| \leq \epsilon$ denote the event $\|g_u^{(i)} - \hat{g}_u^{(i)}\| \leq \epsilon_i$ for $i = 1, \dots, n$ (and similar for the complementary event). Define

$$P[\omega_{\mathbf{x}}(1) \in q' \mid x, u] := P[\omega_{\mathbf{x}}(1) \in q' | \omega_{\mathbf{x}}(0) = x \in q, u].$$

Then using the law of total probability

$$P[\omega_{\mathbf{x}}(1) \in q' \mid x, u] =$$
$$P[\omega_{\mathbf{x}}(1) \in q' \mid x, u, \|g_u - \hat{g}_u\| \leq \epsilon] \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i] +$$
$$P[\omega_{\mathbf{x}}(1) \in q' \mid x, u, \|g_u - \hat{g}_u\| > \epsilon] \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q > \epsilon_i]$$

The transition kernel can be upper bounded by

$$\max_{x \in q} T^u(q' \mid x) = \max_{x \in q} P[\omega_{\mathbf{x}}(1) \in q' \mid x, u] \leq$$
$$\max_{x \in q} P[\omega_{\mathbf{x}}(1) \in q' \mid x, u, \|g_u - \hat{g}_u\| \leq \epsilon] \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i]$$
$$+ 1 \cdot \prod_{i=1}^{n} (1 - P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i]).$$

To account for the uncertainty in the regression and process noise, we again use the law of total probability and expand q' by ϵ and η and check for an intersection between $Im(q)$ and $\overline{q}'(\epsilon + \eta)$:

$$\leq \mathbf{1}^{Im(q)}_{\overline{q}'(\epsilon+\eta)} \prod_{i=1}^{n} P[|v^{(i)}| \leq \eta_i] \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i] +$$
$$\prod_{i=1}^{n} (1 - P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i]).$$

Similarly, the transition kernel can be lower bounded by determining if any points lie outside of the intersection of $Im(q)$ and $\underline{q}'(\epsilon + \eta)$:

$$\min_{x \in q} T^u(q' \mid x) = \min_{x \in q} P[\omega_{\mathbf{x}}(1) \in q' \mid x, u]$$
$$\geq \mathbf{1}^{Im(q)}_{X \setminus \underline{q}(\epsilon+\eta)} \prod_{i=1}^{n} P([v^{(i)}| \leq \eta_i] \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i]$$
$$+ 0 \cdot \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q > \epsilon_i]$$
$$= \mathbf{1}^{Im(q)}_{X \setminus \underline{q}'(\epsilon+\eta)} \prod_{i=1}^{n} P[|v^{(i)}| \leq \eta_i] \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i].$$

□

Theorem 1 computes formal bounds for the transition probabilities by using the law of total probability with respect to the events $|v^{(i)}| \leq \eta_i$ (noise is bounded by η_i), $\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i$ (the supremum of the regression error is bounded by ϵ_i for all $x \in q$), and their complementary events. In particular, a bound on the probability of the latter event can be obtained by Lemma 1, while the probability of former depends on the known distribution of the noise p_v.

In order to get non-trivial transition bounds, constants η and ϵ should be selected to minimize or maximize the bounds in (9) and (10) respectively. In particular, we pick η as the smallest constants such that the noise is bounded by η with high probability, e.g., 0.99. Then, for this η, our procedure to choose a value for ϵ is as follows. We first check if $Im(q) \subset q'$. If it is the case, we pick ϵ as the greatest constants such that $Im(q) \subset \underline{q}'(\epsilon + \eta)$. Otherwise, we simply select ϵ as the smallest constants such that $\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon$ with high probability, e.g., that satisfies the bound in Lemma 1 with probability 0.99.

Note that for $q \subset X$ the above procedure requires one to compute $Im(q)$. This is equivalent to computing the maximum and minimum of (3) for all $x \in q$. Arbitrarily tight bounds for these quantities can be computed by utilizing the convexity of most used kernels, such as the the squared-exponential function, as outlined in [6, 7]. With a similar approach, a bound for $\max_{x \in q} \sigma_D(x)$ can also be computed, as this is required for the computation of Lemma 1.

4.2.3 Transitions to q_u. The probability interval for transitioning to the state $q_u \in Q$, i.e., the region outside of X, is given by

$$\check{P}(q, u, q_u) = 1 - \max_{x \in (q} T^u(X \mid x),$$
$$\hat{P}(q, u, q_u) = 1 - \min_{x \in (q} T^u(X \mid x).$$

These bounds can be calculated as a corollary of Theorem 1.

Corollary 1. *Let $q \in Q_s$, then for any $\epsilon, \eta > 0$ it holds that*

$$\check{P}(q, u, q_u)$$
$$\geq 1 - \mathbf{1}^{Im(q)}_{\overline{X}(\epsilon+\eta)} \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i] \prod_{i=1}^{n} P[|v^{(i)}| \leq \eta_i]$$
$$- \prod_{i=1}^{n} (1 - P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i]),$$
$$\hat{P}(q, u, q_u)$$
$$\leq 1 - \mathbf{1}^{Im(q)}_{X \setminus \underline{X}(\epsilon+\eta)} \prod_{i=1}^{n} P[\|g_u^{(i)} - \hat{g}_u^{(i)}\|_\infty^q \leq \epsilon_i] \prod_{i=1}^{n} P[|v^{(i)}| \leq \eta_i].$$

To complete the construction of abstraction $\mathcal{I}$, we make q_u absorbing, i.e., $\check{P}(q_u, u, q_u) = \hat{P}(q_u, u, q_u) = 1$ for all $u \in A$, to ensure that $\mathcal{I}$ does not account for the transitions to X from q_u since such paths do not satisfy φ as defined in (2).

5 STRATEGY SYNTHESIS

Given an LTL$_f$ formula φ, ideally we would like to synthesize an optimal switching strategy $\pi_{\mathbf{x}}^*$ for Process (1), under which the probability of satisfaction of φ by the paths of Process (1) is maximized. However, since g_u is unknown, this is generally infeasible. Instead, we employ the IMDP abstraction $\mathcal{I}$ as constructed above, which is a conservative model of Process (1) since the transition probabilities of $\mathcal{I}$ include uncertainties (errors) of the learning process as well as those related to the discretization. On this model, we find a strategy that is robust against all these uncertainties and maximizes the probability of satisfying φ. Then, we can refine this strategy to a switching strategy for Process (1). Note that the resulting strategy is not necessarily optimal for Process (1), however, in what follows we show how the error between the resulting strategy and optimal strategy $\pi^*_{\mathbf{x}}$ can be quantified.

5.1 Near-optimal Robust Strategy

The uncertainties in $\mathcal{I}$ are characterized by adversary ξ, which chooses a feasible transition probability from one IMDP state to another under a given action. Recall that given a strategy π and an adversary ξ, $\mathcal{I}$ becomes a Markov chain with a well-defined probability measure over its paths. Then, our (robust and near-optimal strategy) objective translates to finding a strategy that maximizes the probability of satisfying φ with the assumption that the adversary (uncertainty) attempts to minimize this probability, i.e.,

$$\pi^{\varepsilon} = \arg\max_{\pi\in\Pi}\ \min_{\xi\in\Xi} P[\omega_{\mathcal{I}} \models \varphi \mid \pi, \xi, \omega_{\mathcal{I}}(0) = q], \tag{11}$$

Under π^{ε}, the lower bound and upper bound on the probability of satisfaction are then given by

$$\check{p}(q) = \min_{\xi\in\Xi} P[\omega_{\mathcal{I}} \models \varphi \mid \pi^{\varepsilon}, \xi, \omega_{\mathcal{I}}(0) = q], \tag{12}$$

$$\hat{p}(q) = \max_{\xi\in\Xi} P[\omega_{\mathcal{I}} \models \varphi \mid \pi^{\varepsilon}, \xi, \omega_{\mathcal{I}}(0) = q], \tag{13}$$

respectively.

To correctly refine a strategy computed on $\mathcal{I}$ to a switching strategy for process $\mathbf{x}$, let $z : \mathbb{R}^n \to Q$ be a function that maps each state x of Process (1) to its corresponding discrete region $q \in Q$, i.e., $z(x) = q$ iff $x \in q$. We also use z to denote mapping from finite paths of process $\mathbf{x}$ to finite paths of $\mathcal{I}$, i.e., for a finite path $\omega^k_{\mathbf{x}} = x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} \dots \xrightarrow{u_{k-1}} x_k$, the corresponding path on $\mathcal{I}$ is given by

$$z(\omega^k_{\mathbf{x}}) = z(x_0) \xrightarrow{u_0} z(x_1) \xrightarrow{u_1} \dots \xrightarrow{u_{k-1}} z(x_k).$$

Then, strategy π^{ε} on $\mathcal{I}$ is correctly refined to a switching strategy $\pi^{\varepsilon}_{\mathbf{x}}$ for process $\mathbf{x}$ by

$$\pi^{\varepsilon}_{\mathbf{x}}(\omega^k_{\mathbf{x}}) = \pi^{\varepsilon}(z(\omega^k_{\mathbf{x}})). \tag{14}$$

Note that, the maximum probability of satisfaction of φ by Process (1) is necessarily lower bounded by $\check{p}$ in (10), i.e., $p^*(x) \geq \check{p}(z(x))$, where

$$p^*(x) = \max_{\pi_{\mathbf{x}}} P[\omega_{\mathbf{x}} \models \varphi \mid \pi_{\mathbf{x}}, \omega_{\mathbf{x}}(0) = x].$$

However, $p^*(x)$ is not necessarily upper bounded by $\hat{p}$ in (13). Probability $p^*(x)$ can instead be upper bounded by

$$\hat{p}^*(x) = \max_{\pi\in\Pi}\ \max_{\xi\in\Xi} P[\omega_{\mathcal{I}} \models \varphi \mid \pi, \xi, \omega_{\mathcal{I}}(0) = z(x)], \tag{15}$$

where the adversary (uncertainty) cooperatively chooses feasible transition probabilities to maximize the probability of satisfaction of φ. Therefore,

$$p^*(x) \in [\check{p}(z(x)), \hat{p}^*(z(x))].$$

Below, we show how the strategy in (11), its corresponding probability bounds in (12) and (13), and probability in (15) can be computed.

5.2 Synthesis

Given an LTL$_f$ formula φ, a deterministic finite automaton can be constructed that precisely accepts the language of φ per [10].

Definition 8 (DFA). *A deterministic finite automaton (DFA) constructed from an LTL$_f$ formula φ defined over atomic propositions AP is a tuple $\mathcal{A}_{\varphi} = (S, 2^{AP}, \delta, s_0, S_F)$, where S is a finite set of states, 2^{AP} is a finite set of input symbols, each of which is a set of atomic propositions in AP, $\delta : S \times 2^{AP} \to S$ is the transition function, $s_0 \in S$ is the initial state, and $S_F \subseteq S$ is the set of accepting (final) states.*

A finite *run* on a DFA is a sequence of states $s = s_0 s_1 \dots s_{n+1}$ induced by a sequence of symbols $\rho = \rho_0 \rho_1 \dots \rho_n$, where $\rho_i \in 2^{AP}$ and $s_{i+1} = \delta(s_i, \rho_i)$. Finite run s on trace ρ is accepting if $s_n \in S_F$. If s is accepting, then trace ρ is accepted by $\mathcal{A}_{\varphi}$. The set of all traces that are accepted by $\mathcal{A}_{\varphi}$ is call the language of $\mathcal{A}_{\varphi}$, denoted by $\mathcal{L}(\mathcal{A}_{\varphi})$. This language is equal to the language of φ, i.e., $\mathcal{L}(\varphi) = \mathcal{L}(\mathcal{A}_{\varphi})$.

Next, we construct a product of DFA $\mathcal{A}_{\varphi}$ with IMDP $\mathcal{I}$ to capture the paths of $\mathcal{I}$ that satisfy φ.

Definition 9 (Product IMDP). *Given an IMDP $\mathcal{I} = (Q, A, \check{P}^{\mathcal{P}}, \hat{P}^{\mathcal{P}}, AP, L)$ and DFA $\mathcal{A}_{\varphi} = (S, 2^{AP}, \delta, s_0, S_F)$, the product IMDP (PIMDP) $\mathcal{P} = \mathcal{I} \times \mathcal{A}_{\varphi}$ is itself an IMDP defined as the tuple $\mathcal{P} = (Q^{\mathcal{P}}, A, \check{P}^{\mathcal{P}}, \hat{P}^{\mathcal{P}}, Q^{\mathcal{P}}_0, Q^{\mathcal{P}}_F)$, where $Q^{\mathcal{P}} = Q \times S$, $Q^{\mathcal{P}}_F = Q \times S_F$,*

$$Q^{\mathcal{P}}_0 = \{(q, s_{init}) \mid q \in Q,\ s_{init} = \delta(s_0, L(q))\},$$

and

$$\check{P}^{\mathcal{P}}((q,s),u,(q',s')) = \begin{cases} \check{P}(q,u,q') & \textit{if } s' = \delta(s, L(q)) \\ 0 & \textit{otherwise} \end{cases}$$

$$\hat{P}^{\mathcal{P}}((q,s),u,(q',s')) = \begin{cases} \hat{P}(q,u,q') & \textit{if } s' = \delta(s, L(q)) \\ 0 & \textit{otherwise.} \end{cases}$$

In its essence, the PIMDP incorporates the historical dependencies on the runs of the DFA and couples them with the paths of the IMDP. The projection of a path of $\mathcal{P}$ that reaches a state in $Q^{\mathcal{P}}_F$ onto $\mathcal{A}_{\varphi}$ is an accepting run, and hence, the projection of this path onto abstraction $\mathcal{I}$ is a path that satisfies φ. Therefore, the synthesis problem in (11) is reduced to computing a robust (pessimistic) strategy on product $\mathcal{P}$ that maximizes the probability of reaching $Q^{\mathcal{P}}_F$. Similarly, the probability in (15) is given by an optimistic strategy on $\mathcal{P}$ that maximizes the the probability of reaching $Q^{\mathcal{P}}_F$. These problems are variations of a known problem called *maximal reachability probability problem* and can be solved using a method similar to value iteration called interval-value iteration [22, 35], whose computational complexity is polynomial. The resulting strategies are stationary on $\mathcal{P}$, which map to history dependent strategies on $\mathcal{I}$.

5.3 Correctness

The following theorem shows that $\pi^{\varepsilon}_{\mathbf{x}}$ in (14) is a ε-near-optimal switching strategy for Process (1) and quantifies its distance (error) ε to the optimal switching strategy $\pi^*_{\mathbf{x}}$.

Theorem 2. *Given a partially-known switched stochastic system as defined in Process* (1)*, a dataset D, a compact set $X \subset \mathbb{R}^n$, an LTL$_f$ formula φ defined over the regions of interest in X, let $\mathcal{I}$ be an IMDP abstraction as described in Section 4, π^{ε} be a strategy on $\mathcal{I}$ given by*

(11), *and* $\pi^{\epsilon}_{\mathbf{x}}$ *be the switching strategy for Process* (1) *obtained from* π^{ϵ} *according to mapping* z *in* (14)*. Further, let* $\check{p}$, $\hat{p}$, *and* $\hat{p}^*$ *be the probability bounds in* (12), (13), *and* (15)*, respectively. Then, it holds that*

$$P[\omega_{\mathbf{x}} \models \varphi \mid \pi^{\epsilon}_{\mathbf{x}}, \omega_{\mathbf{x}}(0) = x] \in [\check{p}(z(x)), \hat{p}(z(x))],$$

and

$$|P[\omega_{\mathbf{x}} \models \varphi \mid \pi^{\epsilon}_{\mathbf{x}}, \omega_{\mathbf{x}}(0) = x] - p^*(x)| \leq \hat{p}^*(z(x)) - \check{p}(z(x)),$$

where $p^*(x) = \max_{\pi_{\mathbf{x}} \in \Pi_{\mathbf{x}}} P[\omega_{\mathbf{x}} \models \varphi \mid \pi_{\mathbf{x}}, \omega_{\mathbf{x}}(0) = x]$.

Theorem 2 is a straightforward consequence of Theorem 1 and guarantees that the probability that Process (1) satisfies φ is contained between $\check{p}$ and $\hat{p}$. In order to quantify the distance of $\pi^{\epsilon}_{\mathbf{x}}$ from the optimal strategy $\pi^*_{\mathbf{x}}$, we compute the optimal upper bound probability $\hat{p}^*$ correspondent to the strategy that optimistically maximizes the probability of reaching $Q_F^{\mathcal{P}}$. In fact, recall that $\pi^{\epsilon}_{\mathbf{x}}$ corresponds to the strategy that maximizes the lower bound of reaching $Q_F^{\mathcal{P}}$. It follows that for any $x \in X$,

$$\varepsilon_x = |\hat{p}^*(z(x)) - \check{p}(z(x))|.$$

Given a probability bound $\bar{p}$ on the satisfaction of formula φ, we use $\check{p}$ and $\hat{p}$ to classify each initial state $x_0 \in X$ as one of the following:

$$x \in \begin{cases} Q^{\text{yes}} & \text{if } \check{p}(z(x)) \geq \bar{p} \\ Q^{\text{no}} & \text{if } \hat{p}(z(x)) < \bar{p} \\ Q^{?} & \text{otherwise.} \end{cases}$$

Given initial state x_0, we can guarantee that φ is definitely satisfied by the underlying system with at least $\bar{p}$ if $x_0 \in Q^{\text{yes}}$. If $x_0 \in Q^{\text{no}}$, then we can guarantee that the underlying system never meets the probability threshold $\bar{p}$. For $x_0 \in Q^{?}$, no guarantees relative to the threshold $\bar{p}$ can be given.

6 CASE STUDIES

We illustrate the proposed framework in three case studies using linear and nonlinear switched systems. In all the demonstrations, the compact set is $X = [-2, 2] \times [-2, 2]$. We use a uniform grid of size 0.125 over X to create Q_s for our abstraction.

6.1 Linear Switched System with Three Modes

We first demonstrate the framework on a three-mode linear switched system similar to the synthesis example presented in [22]. We assume the dynamics in all three modes are unknown, i.e.,

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

where each mode is a linear system with $g_u(\mathbf{x}_k) = A_u \mathbf{x}_k$ for all $u \in \{1, 2, 3\}$,

$$A_1 = \begin{bmatrix} 0.4 & 0.1 \\ 0 & 0.5 \end{bmatrix}, \ A_2 = \begin{bmatrix} 0.4 & 0.5 \\ 0 & 0.5 \end{bmatrix}, \ A_3 = \begin{bmatrix} 0.4 & 0 \\ 0.5 & 0.5 \end{bmatrix},$$

and $\mathbf{v}$ is drawn from a Gaussian distribution $\mathcal{N}(0, \sigma^2 I)$ truncated between $[-\sigma, \sigma]$ with $\sigma = 0.01$.

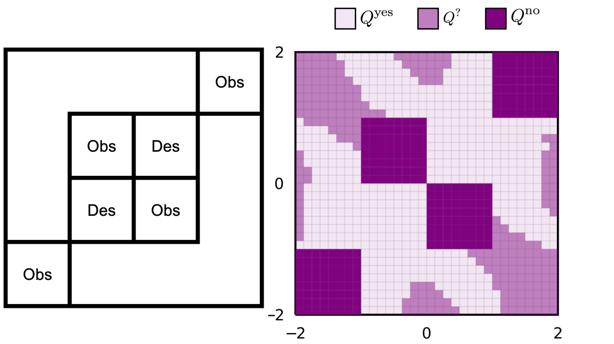


Figure 1: Region partition and classification of initial states using the strategy synthesized for the linear system and φ_1.

Two-hundred i.i.d. data points per mode were sampled and propagated through the dynamics to create the dataset for regression. Figure 1 shows the partition of X with labelled regions *Des* and *Obs* indicating "Desired" and "Obstacle" regions, respectively. With an abuse of notation, these are used to define the atomic propositions $\{Des, Obs\}$ and the LTL$_f$ specification

$$\varphi_1 = \mathcal{G}(\neg Obs) \wedge \mathcal{F}(Des),$$

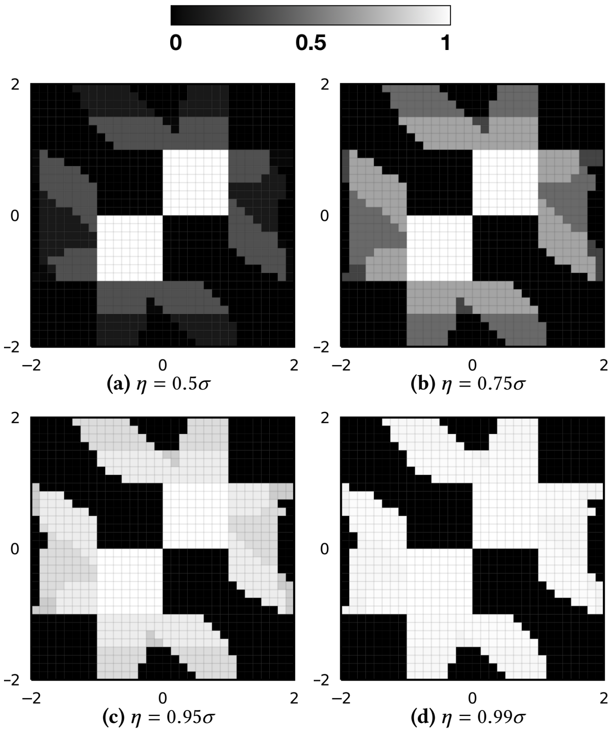
which reads, "Globally avoid Obstacles and eventually reach a Desired region".

Using our framework, we learned the unknown dynamics and synthesized a robust and near-optimal switching strategy π^{ϵ}. Figure 1 shows the classification of each initial region with threshold probability $\bar{p} = 0.95$ under this strategy. Initial states with the *Des* label belong to Q^{yes} as they satisfy φ_1 while states with the *Obs* label violate it and belong to Q^{no}. There are additional states belonging to Q^{yes} such that actions dictated by π^{ϵ} drive the system to an accepting state with a high probability. These results closely resemble the results presented in [22], which assumed full knowledge of the dynamics, whereas here the dynamics are fully unknown and are estimated from a limited set of data.

6.2 Parameter Choices

We provide a brief look at the effect of choosing different values of η, the bounds on the noise components in Theorem 1, on the synthesis results. For any choice of η, the optimal value of ϵ, the bounds on the suprema of the regression error components, is then chosen to minimize (maximize) the upper-bound (lower-bound) of the transition probability in Theorem 1 as discussed previously. In the ideal case, the noise parameter primarily effects the lower-bound of the transition probability, as the optimal choice of ϵ leaves the indicator function in (10) with a value of zero.

For the three-mode linear switched system above, the effect of changing η uniformly in a naïve manner is shown in Figure 2. The choices of η are presented as fractions of the bounds on the truncated Gaussian distribution. As expected, the lower-bound on the probability of satisfaction decreases as the value of η decreases. In a large part, this is due to the reduction in the lower-bound of the probability of staying within X. There is clearly an optimal trade-off to be made between ϵ and η, which will be considered in future work.



(a) $\eta = 0.5\sigma$

(b) $\eta = 0.75\sigma$

(c) $\eta = 0.95\sigma$

(d) $\eta = 0.99\sigma$

Figure 2: Effect on changing the parameter η on the lower bound of the probability of satisfaction from each state.

6.3 Nonlinear Switched System with Four Modes

Next, we synthesize a switching strategy for a nonlinear system with four modes, a known linear dynamics component, and an unknown nonlinear dynamics component. The form of the system is

$$\mathbf{x}_{k+1} = \mathbf{x}_k + g_{\mathbf{u}_k}(\mathbf{x}_k) + \mathbf{v}_k$$

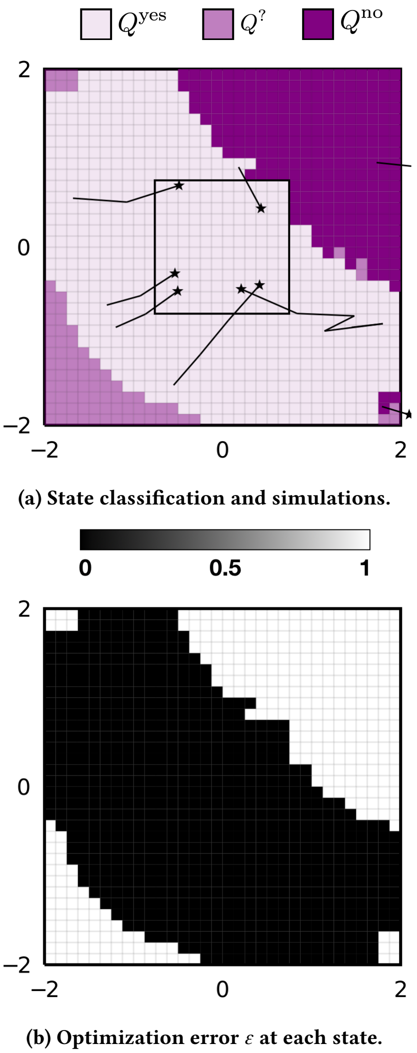
The unknown dynamics are

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

where $x^{(i)}$ indicates the i-th component of the state. Three-hundred i.i.d. data points were generated from each mode with the same noise distribution.

The region of interest is *Des*, which indicated by the black square in Figure 3a, and the specification is

$$\varphi_2 = \mathcal{F}\, Des.$$



(a) State classification and simulations.

(b) Optimization error ε at each state.

Figure 3: Synthesis results and simulations for the nonlinear system for φ_2. The black square in (a) indicates the region with label *Des*.

Figure 3a shows the synthesis and simulated result. The black lines indicate simulated paths from a set of randomly selected initial states, and the star indicates the terminal state. Individual simulations were terminated if an accepting state in the PIMDP was reached, or if the specification was violated. The size of Q^{yes} in Figure 3a shows that the strategy π^{ε} can drive many states into *Des* with a high probability. However, there are many states in Q^{no}. In these states, the *upper-bound* of the probability of satisfying φ_2 induced by employing π^{ε} does not meet $\bar{p}$. In other words, there is a significant chance of violating φ_2.

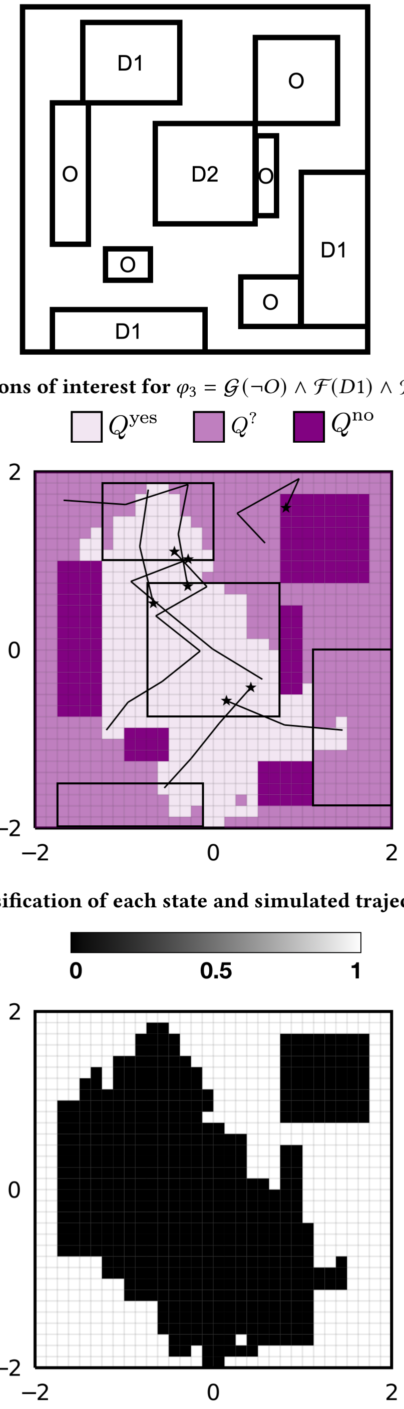
Figure 3b shows upper bounds of error ε at each state. Recall that ε bounds the satisfaction probability distance under optimal and near-optimal strategies in Problem 1. The darker regions correspond to areas with ε approaching zero, meaning π^{ε} is indeed near-optimal. The lighter regions with ε approaching one indicate that π^{ε} does not necessarily choose the optimal action. This could be mitigated by collecting more data and performing a finer abstraction, or it is possible the system does not have sufficient control authority.

Finally, we perform controller synthesis for the partially-known nonlinear switched system given the specification

$$\varphi_3 = \mathcal{G}(\neg O) \wedge \mathcal{F}(D1) \wedge \mathcal{F}(D2)$$

with two reachability objectives. Figure 4a shows the partitioning of the space with labels $D1$, $D2$ and O indicating "Desired Location 1", "Desired Location 2" and "Obstacle" regions respectively. We use the same abstraction we generated in the previous case.

Figure 4b shows several simulation results from different initial states, two of which are in $Q^?$ and the others in Q^{yes}. The Q^{no} set is comprised of only the O regions, because starting in O automatically violates the specification. Much of the Q^{yes} set is made up of a majority of $D2$ and some states starting in $D1$. There is a large amount of free space that can be driven into $D1$ with high probability. All paths but one terminate with satisfying φ_3, but a single path is driven into an obstacle. Figure 4c shows that the optimal action has been found for many of the states, but there is a significant number of states with a trivial bound of $\varepsilon = 1$. This metric can help identify areas for further data collection, or state discretization refinement.



(a) Regions of interest for $\varphi_3 = \mathcal{G}(\neg O) \wedge \mathcal{F}(D1) \wedge \mathcal{F}(D2)$.

(b) Classification of each state and simulated trajectories.

(c) Optimization error ε at each state.

Figure 4: Synthesis results and simulations for the nonlinear system with φ_3.

7 CONCLUSION

We developed a data-driven framework for synthesizing a near-optimal control strategy for partially unknown switched stochastic systems with LTL$_f$ specifications. The framework is based on abstraction to an uncertain Markov model that incorporates both the uncertainty given by the stochastic dynamics of the system and the uncertainty in learning the unknown dynamics of the system via GP regression. Our work makes a step towards formally safe and correct data-driven systems. However, many challenges are ahead in order to make our framework to scale to larger datasets and higher dimensional systems. In the future, we plan to consider sparse Gaussian processes as well as optimal techniques for parameter tuning and refinement.

8 ACKNOWLEDGEMENT

This work was supported in part by the NSF Center for Unmanned Aircraft Systems, NSF under awards IIP - 1650468, and NSF 2039062.

REFERENCES

[1] Alessandro Abate, Frank Redig, and Ilya Tkachev. 2014. On the effect of perturbation of conditional probabilities in total variation. *Statistics & Probability Letters* 88 (2014), 1–8.

[2] Mohamadreza Ahmadi, Arie Israel, and Ufuk Topcu. 2017. Safety assessemt based on physically-viable data-driven models. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 6409–6414.

[3] Anayo K Akametalu, Shahab Kaynama, Jaime F Fisac, Melanie N Zeilinger, Jeremy H Gillula, and Claire J Tomlin. 2014. Reachability-based safe learning with Gaussian processes. In *IEEE 53rd Annual Conference on Decision and Control (CDC), 2014: 15-17 Dec. 2014, Los Angeles, California, USA*. IEEE, 1424–1431.

[4] Christel Baier, Joost-Pieter Katoen, et al. 2008. *Principles of model checking*. MIT press Cambridge.

[5] Felix Berkenkamp, Matteo Turchetta, Angela Schoellig, and Andreas Krause. 2017. Safe model-based reinforcement learning with stability guarantees. In *Advances in neural information processing systems*. 908–918.

[6] Arno Blaas, Andrea Patane, Luca Laurenti, Luca Cardelli, Marta Kwiatkowska, and Stephen Roberts. 2020. Adversarial robustness guarantees for classification with gaussian processes. In *International Conference on Artificial Intelligence and Statistics*. 3372–3382.

[7] Luca Cardelli, Marta Kwiatkowska, Luca Laurenti, and Andrea Patane. 2019. Robustness guarantees for bayesian inference with gaussian processes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 7759–7768.

[8] Nathalie Cauchi, Luca Laurenti, Morteza Lahijanian, Alessandro Abate, Marta Kwiatkowska, and Luca Cardelli. 2019. Efficiency through Uncertainty: Scalable Formal Synthesis for Stochastic Hybrid Systems. In *Proceedings of the 2019 22nd ACM International Conference on Hybrid Systems: Computation and Control*. ACM, Montreal, QC, Canada. https://doi.org/10.1145/3302504.3311805

[9] Sayak Ray Chowdhury and Aditya Gopalan. 2017. On kernelized multi-armed bandits. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 844–853.

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

[11] Laurent Doyen, Goran Frehse, George J Pappas, and André Platzer. 2018. Verification of hybrid systems. In *Handbook of Model Checking*. Springer, 1047–1110.

[12] Souradeep Dutta, Susmit Jha, Sriram Sankaranarayanan, and Ashish Tiwari. 2018. Learning and verification of feedback control systems using feedforward neural networks. *IFAC-PapersOnLine* 51, 16 (2018), 151–156.

[13] Pascal Germain, Francis Bach, Alexandre Lacoste, and Simon Lacoste-Julien. 2016. PAC-Bayesian theory meets Bayesian inference. In *Advances in Neural Information Processing Systems*. 1884–1892.

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

[15] Sofie Haesaert, Nathalie Cauchi, and Alessandro Abate. 2017. Certified policy synthesis for general Markov decision processes: An application in building automation systems. *Performance Evaluation* 117 (2017), 75–103.

[16] Sofie Haesaert, Paul MJ Van den Hof, and Alessandro Abate. 2017. Data-driven and model-based verification via Bayesian identification and reachability analysis. *Automatica* 79 (2017), 115–126.

[17] Ernst Moritz Hahn, Vahid Hashemi, Holger Hermanns, Morteza Lahijanian, and Andrea Turrini. 2019. Interval Markov decision processes with multiple objectives: From Robust strategies to Pareto curves. *ACM Transactions on Modeling and Computer Simulation* 29, 4 (2019), 1–31. https://doi.org/10.1145/3309683

[18] John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. 2020. Safety verification of unknown dynamical systems via gaussian process regression. *IEEE CDC* (2020).

[19] A Agung Julius, Ádám Halász, M Selman Sakar, Harvey Rubin, Vijay Kumar, and George J Pappas. 2008. Stochastic modeling and control of biological systems: the lactose regulation system of Escherichia coli. *IEEE Trans. Automat. Control* 53, Special Issue (2008), 51–65.

[20] Joris Kenanian, Ayca Balkan, Raphael M Jungers, and Paulo Tabuada. 2019. Data driven stability analysis of black-box switched linear systems. *Automatica* 109 (2019), 108533.

[21] Harold Kushner and Paul G Dupuis. 2013. *Numerical methods for stochastic control problems in continuous time*. Vol. 24. Springer Science & Business Media.

[22] Morteza Lahijanian, Sean B Andersson, and Calin Belta. 2015. Formal verification and synthesis for discrete-time stochastic systems. *IEEE Trans. Automat. Control* 60, 8 (2015), 2031–2045.

[23] Luca Laurenti, Morteza Lahijanian, Alessandro Abate, Luca Cardelli, and Marta Kwiatkowska. 2020. Formal and efficient synthesis for continuous-time linear stochastic hybrid processes. *IEEE Trans. Automat. Control* 66, 1 (2020), 17–32.

[24] Armin Lederer, Jonas Umlauft, and Sandra Hirche. 2019. Uniform Error Bounds for Gaussian Process Regression with Application to Safe Control. In *Advances in Neural Information Processing Systems*. 657–667.

[25] Ryan Luna, Morteza Lahijanian, Mark Moll, and Lydia E. Kavraki. 2014. Asymptotically Optimal Stochastic Motion Planning with Temporal Goals. In *Int'l Workshop on the Algorithmic Foundations of Robotics (WAFR)*. Istanbul, Turkey, 335–352.

[26] Pascal Massart. 2007. *Concentration inequalities and model selection*. Vol. 6. Springer.

[27] Vern I Paulsen and Mrinal Raghupathi. 2016. *An introduction to the theory of reproducing kernel Hilbert spaces*. Vol. 152. Cambridge University Press.

[28] Kyriakos Polymenakos, Luca Laurenti, Andrea Patane, Jan-Peter Calliess, Luca Cardelli, Marta Kwiatkowska, Alessandro Abate, and Stephen Roberts. 2019. Safety Guarantees for Planning Based on Iterative Gaussian Processes. *arXiv preprint arXiv:1912.00071* (2019).

[29] Carl Edward Rasmussen. 2003. Gaussian processes in machine learning. In *Summer School on Machine Learning*. Springer, 63–71.

[30] Sadegh Esmaeil Zadeh Soudjani, Caspar Gevaerts, and Alessandro Abate. 2015. FAUST2: Formal Abstractions of Uncountable-STate STochastic Processes. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 272–286.

[31] Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias W Seeger. 2012. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory* 58, 5 (2012), 3250–3265.

[32] Ingo Steinwart. 2001. On the influence of the kernel on the consistency of support vector machines. *Journal of machine learning research* 2, Nov (2001), 67–93.

[33] Yanan Sui, Alkis Gotovos, Joel W Burdick, and Andreas Krause. 2015. Safe exploration for optimization with Gaussian processes. *Proceedings of Machine Learning Research* 37 (2015), 997–1005.

[34] Andrew M. Wells, Morteza Lahijanian, Lydia E. Kavraki, and Moshe Y. Vardi. 2020. LTLf Synthesis on Probabilistic Systems. In *11th International Symposium on Games, Automata, Logics, and Formal Verification (Electronic Proceedings in Theoretical Computer Science, Vol. 326)*, Jean-Francois Raskin and Davide Bresolin (Eds.). Open Publishing Association, 166–181.

[35] Di Wu and Xenofon Koutsoukos. 2008. Reachability analysis of uncertain systems using bounded-parameter Markov decision processes. *Artificial Intelligence* 172, 8-9 (2008), 945–954.