Formal Synthesis of Safety Controllers for Unknown Stochastic Control Systems using Gaussian Process Learning

Rameez Wajid RAMEEZ.WAJID@COLORADO.EDU
Computer Science Department, University of Colorado Boulder, USA.

Asad Ullah Awan ASAD.AWAN@TUM.DE
Department of Electrical and Computer Engineering, Technical University of Munich, Germany.

Majid Zamani MAJID.ZAMANI@COLORADO.EDU
Computer Science Department, University of Colorado Boulder, USA.
Computer Science Department, LMU Munich, Germany.

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

Abstract

Formal synthesis of controllers for stochastic control systems with unknown models is a challenging problem. In this paper, we focus on safety controller synthesis for nonlinear stochastic control systems. The approach consists of a learning step followed by a controller synthesis scheme using control barrier functions. In the learning phase, we employ Gaussian processes (GP) to learn models of unknown stochastic control systems in the presence of both process and measurement noises. In the controller synthesis phase, we compute control barrier functions together with their corresponding controllers based on the learned GP and quantify lower bounds on the probabilities of safety satisfaction for the original unknown systems equipped with the synthesized controllers. Finally, the effectiveness of the proposed approach is illustrated on a room temperature control and a vehicle lane-keeping example.

Keywords: Synthesis, safety controllers, stochastic systems with unknown models, Gaussian processes, control barrier functions

1. Introduction

Designing safety controllers for safety-critical applications is an important problem. Here, safety is considered in the sense of preventing the system from reaching a given *unsafe* set. The general approaches for synthesizing a safety controller require accurate mathematical models of the system dynamics. However, closed-form models derived from first principles for many real-world systems are complex or even not available, and hence one cannot use model-based techniques for such systems. Hence, the design of safety controllers is much more challenging for systems with unknown or partially known models.

Approaches based on Barrier functions (Prajna et al. (2007)) have been promising for synthesizing safety controllers. These discretization-free approaches usually formulate the search for barrier functions as sum-of-squares (SOS) optimization problems which are computed using existing semidefinite programming (SDP) solvers (Ames et al. (2019); Borrmann et al. (2015)). Barrier functions can also be leveraged to synthesize controllers for complex logic specifications (Yang et al. (2020); Li and Belta (2019)). Unfortunately, approaches based on barrier functions require a precise mathematical model of the system which may not be available.

Recently, Gaussian processes (GPs) have emerged as a learning-based technique for modelling unknown dynamical systems which allow for a quantification of uncertainty for the learned model (Williams and Rasmussen (2006)). The uncertainty quantification provides out-of-sample performance guarantees, making the GPs attractive tools in control applications like adaptive control (Chowdhary et al. (2014)), feedback-linearization (Umlauft et al. (2017)), and policy through reinforcement learning for robotic applications (Akametalu et al. (2014)). An accurate GP regression model can often be constructed using only a relatively small number of training samples.This property makes GPs more desirable with respect to other data-driven approaches such as those based on scenario convex problems (SCP) (Calafiore and Campi (2006)), which require large numbers of samples for providing out-of-sample performance guarantees (Salamati et al. (2021); Berger et al. (2021)).

Related Work: Recently, there have been some results to combine GPs with control barrier functions for safety controller synthesis of unknown nonlinear control-affine systems. The work in Jagtap et al. (2020a) uses GPs to model unknown continuous-time control-affine dynamics and then uses the learned GPs to compute control barrier certificates together with their corresponding controllers satisfying safety specifications. The work in Castañeda et al. (2021) uses GP regression to learn model uncertainties for control-affine systems with known nominal dynamics. The GPs are then used to adjust control barrier certificates, earlier derived from the nominal dynamics. Another approach has been proposed to combine GP learning with abstraction-based techniques for partially known stochastic systems (Jackson et al. (2020, 2021)). However, none of these existing approaches can synthesize safety controllers for fully unknown stochastic systems without discretizing state sets, while considering both process and measurement noises simultaneously.

In this paper, we provide a scheme to synthesize safety controllers for nonlinear stochastic control systems with unknown dynamics. First, we use the Gaussian process regression to learn a model of the unknown stochastic system with a probabilistic guarantee on the model accuracy. Specifically, we use the improved Gaussian process upper confidence bound (IGP-UCB) (Chowdhury and Gopalan (2017)) for establishing the probabilistic closeness between the learned GP and the original unknown model. Then, we construct a control barrier certificate together with the corresponding controller using the Counterexample Guided Inductive Synthesis framework (CEGIS) (Ravanbakhsh and Sankaranarayanan (2015)). The synthesized controller is shown to satisfy the specified safety specification on the original system with an a priori chosen confidence bound. Our approach is one of the first attempts that accounts for both process and observation noises in the controller synthesis process for unknown stochastic systems. We use a room temperature control example and a vehicle lane-keeping scenario to illustrate the effectiveness of the proposed results.

2. Preliminaries and Problem Definition

2.1. Preliminaries

We consider the probability space $(\Omega, \mathcal{F}_\Omega, \mathbb{P})$, where Ω is the sample space, $\mathcal{F}_\Omega$ is a sigma-algebra consisting of subsets of Ω as events, and $\mathbb{P}$ is the probability measure that assigns probability to those events. Random variables X are assumed to be measurable functions of form $X : (\Omega, \mathcal{F}_\Omega) \to (S_X, \mathcal{F}_X)$. Any random variable induces a probability measure on $(S_X, \mathcal{F}_X)$ as $Prob\{A\} = \mathbb{P}_\Omega\{X^{-1}(A)\}$ for any $A \in \mathcal{F}_X$.

2.2. Notations

We denote the set of positive integers by $\mathbb{N} := \{1, 2, 3, ...\}$ and the set of non-negative integers by $\mathbb{N}_0 := \{0, 1, 2, ...\}$. The set of real, positive real, and non-negative real numbers are denoted by $\mathbb{R}$, $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq 0}$, respectively. We use $\mathbb{R}^n$ to denote an n-dimensional Euclidean space and the space of real matrices with n rows and m columns is denoted by $\mathbb{R}^{n\times m}$. The (n-dimensional) multivariate normal distribution is denoted by $\mathcal{N}(\mu, C)$ with mean vector $\mu \in \mathbb{R}^n$ and covariance matrix $C \in \mathbb{R}^{n\times n}$. We use the notation $\bigcap_{i=1}^{n} M_i$, for the conjunction of events $M_1, \ldots M_n$.

A Hilbert space of square integrable functions which includes functions of the form $h(x) = \sum_i \alpha_i k(x, x_i)$, where $\alpha_i \in \mathbb{R}$, $x, x_i \in X \subset \mathbb{R}^n$, is called a reproducing kernel Hilbert space (RKHS) if $k : X \times X \mapsto \mathbb{R}_{\geq 0}$ is a symmetric positive definite function called kernel. The corresponding induced RKHS norm with respect to a kernel k is denoted by $\|h\|_k$. A more rigorous discussion on RKHS norms can be found in Paulsen and Raghupathi (2016).

A random sequence $\varepsilon_r := \{\varepsilon_r(t) : \Omega \mapsto W, t \in \mathbb{N}_0\}$ is conditionally R-sub-Gaussian for a fixed constant $R \in \mathbb{R}_{\geq 0}$ if it satisfies

$$\forall t \in \mathbb{N}_0, \forall b \in \mathbb{R}, \quad \mathbb{E}[\mathsf{e}^{b\varepsilon_r(t)} \mid \mathcal{F}_{t-1}] \leq \mathsf{e}^{\left(\frac{b^2R^2}{2}\right)}, \tag{1}$$

where $\mathcal{F}_{t-1}$ is the sigma-algebra generated by the random variables $\{\varepsilon_r(0), \varepsilon_r(1), \ldots, \varepsilon_r(t-1)\}$. We use $\varepsilon_r \sim subG(R)$ to denote such a random sequence.

2.3. Discrete-time stochastic control systems

We consider discrete-time stochastic control systems as the underlying models for unknown systems.

Definition 1 *A discrete-time stochastic control system (dt-SCS) is characterized by a tuple $S = (X, U, \varepsilon, f)$, where*

- *$X \subseteq \mathbb{R}^n$ is a Borel space as the state space of the system. We denote by $(X, \mathcal{B}(X))$ the measurable state space where $\mathcal{B}(X)$ is the Borel sigma-algebra on the state space.*
- *$U \subseteq \mathbb{R}^m$ is a Borel space as input space of the system.*
- *$\varepsilon = [\varepsilon_1, \ldots, \varepsilon_n]$ is a vector of n independent σ-sub-Gaussian random sequences, i.e. $\varepsilon_i \sim subG(\sigma)$, $\forall i \in \{1, \ldots, n\}$, $\sigma \in \mathbb{R}_{\geq 0}$.*
- *Map $f : X \times U \mapsto X$ is a measurable function characterizing the state evolution of the system.*

For a given initial state $x(0) = x_0 \in X_0 \subset X$ and input sequence $\{u(t) : \Omega \to U, t \in \mathbb{N}_0\}$, the state evolution is characterized by the following difference equation:

$$x(t+1) = f(x(t), u(t)) + \varepsilon(t), \quad t \in \mathbb{N}_0. \tag{2}$$

We assume that the safety of a dt-SCS S is enforced by a stationary policy $u : X \to U$ mapping at any time t the current state $x(t)$ to an input $u(t)$. For the main problem formulation, we consider the following assumptions.

Assumption 1 *For a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$, the map $f : X \times U \mapsto X$ is unknown.*

We also assume that the map f in $\mathcal{S}$ has low complexity, as measured under the reproducing kernel Hilbert space (RKHS) norm (Paulsen and Raghupathi (2016)) as follows:

Assumption 2 *For a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$, each component of the map f has a bounded RKHS norm with respect to the kernel k, i.e. $\exists B_j \in \mathbb{R}_{\geq 0}$ s.t. $\|f_j\|_k \leq B_j$ for all $j \in \{1, \ldots, n\}$.*

The RKHS has a property of being dense in the space of continuous functions for positive definite kernels over a compact domain X. This means that the kernel can arbitrarily approximate any continuous function over the compact domain X (Seeger et al. (2008)). Assumption 2 allows us to use Gaussian process regression to model f. Next, we have some assumptions on the availability of the training data-set.

Assumption 3 *For a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$, we have access to measurements for $x(t) \in X$, $u(t) \in U$ and to the noisy observations $y(t) = x(t+1) + w(t) = f(x(t), u(t)) + \varepsilon(t) + w(t)$, $\forall t \in \mathbb{N}_0$, where $w = [w_1, \ldots, w_n]$, is a vector of n independent θ-sub-Gaussian random sequences representing the measurement noise (independent of ε), i.e. $w_i \sim subG(\theta), i \in \{1, \ldots, n\}, \theta \in \mathbb{R}_{\geq 0}$.*

In practice, measurements $f(x(t), u(t))$ can be acquired by simulating or running the system $\mathcal{S}$ from multiple initial conditions. Then, the observations $y(t)$ can be rewritten as

$$y(t) = f(x(t), u(t)) + \nu(t), \tag{3}$$

$t \in \mathbb{N}_0$, where $\nu = [\nu_1, \ldots, \nu_n]$ and $\nu_i \sim subG(R)$ with $R^2 = \theta^2 + \sigma^2$, $i \in \{1, \ldots, n\}$.

The controller synthesis problem investigated in this paper can now be stated as follows:

Problem 1 *For a system dt-SCS $\mathcal{S}$ satisfying Assumptions 1-3, an initial set $X_0 \subset X$, and an unsafe set $X_u \subset X$, synthesize a controller that provides a lower bound on the probability that the solution process of $\mathcal{S}$ starting in X_0 does not reach X_u within a bounded time horizon.*

3. Gaussian Process Modelling

A Gaussian process (GP) is a non-parametric probabilistic framework belonging to the kernel methods family in machine learning (Bishop (2006)). It utilizes the concept of a prior probability distribution over discrete random variables and generalizes them to an infinite space of continuous functions. Its most important application is the GP regression, which is used to model unknown nonlinear functions. A GP with a domain X_{in} is completely specified by its mean function $m : X_{in} \mapsto \mathbb{R}$ and covariance function $k : X_{in} \times X_{in} \mapsto \mathbb{R}$ written as $\mathcal{GP}(m, k)$. We denote by $f \sim \mathcal{GP}(m, k)$ the approximation of function f by a GP $\mathcal{G}(m, k)$. The *a-priori* distribution (i.e. before training the GP) corresponding to $f \sim \mathcal{GP}(m, k)$ at any point $x \in X_{in}$ is Gaussian with the mean and covariance given by $m(x)$ and $k(x, x)$, respectively. The covariance function (also known as kernel) $k(x, x')$ is a similarity measure between any two inputs x, $x' \in X_{in}$. The kernel choice is largely problem-dependent with the linear, squared-exponential, and Matérn kernels being most commonly utilized ones. It is common to take the mean function m to be a zero-valued function, which we also assume here without loss of generality.

The GP approximation for an n-dimensional function $f : X \times U \to X$, where $X \subset \mathbb{R}^n, U \subset \mathbb{R}^m$, can be obtained by modeling each component f_j by n independent GPs i.e.,

$$f_j \sim \mathcal{GP}_j(0, k_j), \tag{4}$$

where the kernel is denoted by $k_j : (X \times U) \times (X \times U) \mapsto \mathbb{R}$, $j \in \{1, 2, \ldots, n\}$, and 0 represents zero-valued function.

Suppose we collect N measurements $\{y^{(1)}, \ldots, y^{(N)}\}$ and $\{(x, u)^{(1)}, \ldots, (x, u)^{(N)}\}$, where $y^{(i)} = f((x, u)^{(i)}) + \nu^{(i)}$, $i \in \{1, 2, \ldots, N\}$, (as in Assumption 3); then the *posterior* distribution corresponding to $f_j(x, u)$, for $j \in \{1, 2, \ldots, n\}$, at an arbitrary state $x \in X$ and input $u \in U$ is computed as a normal distribution $\mathcal{N}(\mu_j(x, u), \rho_j^2(x, u))$ with the mean and covariance given by

$$\begin{aligned} \mu_j(x, u) &= \overline{k}_j^T (K_j + (1 + 2/N)\mathbf{I}_N)^{-1} y_j, \\ \rho_j^2(x, u) &= k_j((x, u), (x, u)) - \overline{k}_j^T (K_j + (1+2/N)\mathbf{I}_N)^{-1} \overline{k}_j, \end{aligned} \tag{5}$$

respectively, where $\mathbf{I}_N$ is the identity matrix, $\overline{k}_j = [k_j((x, u)^{(1)}, (x, u)), \ldots, k_j((x, u)^{(N)}, (x, u))]^T \in \mathbb{R}^N$, $y_j = [y_j^{(1)}, \ldots, y_j^{(N)}]^T \in \mathbb{R}^N$, and

$$K_j = \begin{bmatrix} k_j((x, u)^{(1)}, (x, u)^{(1)}) & \ldots & k_j((x, u)^{(1)}, (x, u)^{(N)}) \\ \vdots & \ddots & \vdots \\ k_j((x, u)^{(N)}, (x, u)^{(1)}) & \ldots & k_j((x, u)^{(N)}, (x, u)^{(N)}) \end{bmatrix} \in \mathbb{R}^{N \times N}.$$

Now, the function f can be approximated by augmenting the mean and covariance functions in (5) from each GP as follows:

$$\begin{aligned} \mu(x, u) &:= [\mu_1(x, u), \mu_2(x, u), \ldots, \mu_n(x, u)]^T \\ \rho^2(x, u) &:= [\rho_1^2(x, u), \rho_2^2(x, u), \ldots, \rho_n^2(x, u)]^T. \end{aligned} \tag{6}$$

The following lemma shows that we can quantify the upper bound on the difference between the true value $f(x, u)$ and the inferred mean $\mu(x, u)$ with a probability lower bound.

Lemma 2 *Consider a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$ satisfying Assumptions 1-3, and a learned GP for f using N training points, having the posterior mean and covariance functions as in (5). Then, the following inclusion holds true with a confidence of at least $(1 - \delta)^n$:*

$$f(x, u) \in \{\mu(x, u) + d \mid d \in \mathcal{D}\}, \forall x \in X, \forall u \in U \tag{7}$$

where, $\mathcal{D} := \{[d_1, \ldots, d_n]^T | \, d_j \in [-\beta_j \bar{\rho}_j, \beta_j \bar{\rho}_j], j \in \{1, \ldots, n\}\}$, $\beta_j = B_j + R\sqrt{2\left(\alpha_j + 1 + \log(\frac{1}{\delta})\right)}$, and $\bar{\rho}_j{}^2(x, u) = \max_{x \in X, u \in U} \rho_j^2(x, u)$.

The proof is similar to that of (Umlauft et al., 2018, Lemma 2). It follows from (Chowdhury and Gopalan, 2017, Theorem 2) by extending the scalar result that $\mu_j(x, u) - \beta_j \rho_j(x, u) \leq f_j(x, u) \leq \mu_j(x, u) + \beta_j \rho_j(x, u), \forall x \in X, \forall u \in U$ holds with a confidence of at least $1 - \delta$ to an n-dimensional state-set.

Remark 3 *Computing the information-theoretic term α_j in the expression of β_j above, which quantifies the mutual information gain between the original function and the finite data samples, is an NP-hard problem in general. For commonly used kernels, e.g. the squared-exponential or the linear kernel, α_j grows sub-linearly with the number of data samples N, as detailed in Srinivas et al. (2009). In our case-study, we circumvent this problem by directly approximating d in (7) using a Monte-Carlo approach (details in Section 5).*

4. Control Barrier Functions

Here, we introduce a notion of control barrier functions which is used to find a control policy that yields a lower bound on the probability that a discrete time stochastic system avoids an unsafe set over a bounded time horizon, as formalized in the next lemma borrowed from Jagtap et al. (2020b).

Lemma 4 *Consider a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$ as in Definition 1 and sets $X_0, X_u \subseteq X$ as the initial and unsafe sets, respectively. Suppose there exists a function $B : X \mapsto \mathbb{R}_{\geq 0}$, and constants $c \in \mathbb{R}$, $\lambda \in \mathbb{R}_{\geq 0}$, $\gamma \in \mathbb{R}_{>0}$, with $\gamma > \lambda$, such that*

$$B(x) \leq \lambda, \quad \forall x \in X_0, \tag{8}$$

$$B(x) > \gamma, \quad \forall x \in X_u, \tag{9}$$

and $\forall x \in X$, $\exists u \in U$, such that

$$\mathbb{E}[B(x(t+1)) \mid x(t) = x, u(t) = u] - B(x(t)) \leq c. \tag{10}$$

Then, under a control policy u associated with B (cf. existential quantifier in condition (10)), the lower bound on the probability that the solution process of $\mathcal{S}$ starting from any initial state $x_o \in X_0$ does not reach X_u in a bounded time horizon $[0, T]$ is given by

$$\mathbb{P}\{x(t) \notin X_u, \forall t \in [0, T] \subset \mathbb{N}_0 \mid x(0) = x_0\} \geq 1 - \frac{\lambda + \max(0, c)T}{\gamma}. \tag{11}$$

Remark 5 *Condition (10) in Lemma 4 implicitly gives rise to a (stationary) control policy $u : X \mapsto U$ according to the existential quantifier on the input for any state $x \in X$.*

For a dt-SCS with unknown map f, we derive a lower bound on the probability that the solution process does not enter an unsafe set in a bounded time horizon via a learned GP using data as described in Section 3.

Theorem 6 *Consider a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$ satisfying Assumptions 1-3, a learned Gaussian process model with the posterior mean μ and covariance $\rho^2(\cdot)$ as given in (6), and the result in Lemma 2. Let $X_0, X_u \subset X$ represent the initial and unsafe sets for $\mathcal{S}$, respectively. Suppose there exists a function $B : X \mapsto \mathbb{R}_{\geq 0}$, constants $c \in \mathbb{R}$, $\lambda \in \mathbb{R}_{\geq 0}$, $\gamma \in \mathbb{R}_{>0}$, with $\gamma > \lambda$, such that*

$$B(x) \leq \lambda, \quad \forall x \in X_0, \tag{12}$$

$$B(x) > \gamma, \quad \forall x \in X_u, \tag{13}$$

and $\forall x \in X$, $\exists u \in U$ such that $\forall d \in \mathcal{D}$,

$$\mathbb{E}[B(\mu(x(t), u(t)) + d + \varepsilon(t)) \mid x(t) = x, u(t) = u] - B(x(t)) \leq c, \tag{14}$$

where $\mathcal{D}$ is the set defined in (7). Then, under a control policy u associated with B (cf. existential quantifier in (14)), the following inequality, which provides the lower bound on the probability that the solution process of $\mathcal{S}$ starting from any initial state $x_o \in X_0$ does not reach X_u within a bounded time horizon $[0, T]$, holds with a confidence of at least $(1 - \delta)^n$

$$\mathbb{P}\{x(t) \notin X_u, \forall t \in [0, T] \subset \mathbb{N}_0 \mid x(0) = x_0\} \geq (1 - \frac{\lambda + \max(0, c)T}{\gamma}). \tag{15}$$

Proof From the result in Lemma 2, we have that the inclusion $f(x,u) \in \{\mu(x,u)+d \mid d \in \mathcal{D}\}, \forall x \in X, \forall u \in U$, holds with a confidence of at least $(1-\delta)^n$. Together with (14), this implies that the condition:
$\forall x \in X, \exists u \in U$ such that,

$$\mathbb{E}[B(f(x(t),u(t))+\varepsilon(t)) \mid x(t)=x, u(t)=u] - B(x(t)) \leq c, \tag{16}$$

holds true with a confidence of at least $(1-\delta)^n$.

It then follows from Lemma 4 that the inequality

$$\mathbb{P}\{x(t) \notin X_u, \forall t \in [0,T] \subset \mathbb{N}_0 \mid x(0)=x_0\} \geq (1 - \frac{\lambda + \max(0,c)T}{\gamma}). \tag{17}$$

holds with a confidence of at least $(1-\delta)^n$. ■

4.1. Calculation of barrier certificate

The computation of a control barrier function (if existing) for a dt-SCS is a difficult task, in general. However, if the input set of a dt-SCS is assumed to be finite, i.e. $U = \{u_1, u_2, \ldots, u_k\}$, where $u_i \in \mathbb{R}^m, i \in \{1,2,\ldots,k\}$, then the search for a parametric control barrier function and its associated control policy becomes tractable. We employ the Counterexample Guided Inductive Synthesis framework (CEGIS) which has recently become popular for the synthesis of barrier functions (Solar-Lezama et al. (2006); Jagtap et al. (2020a)). The following lemma, adapted from (Jagtap et al., 2020a, Lemma 4.4), provides feasibility conditions that, if satisfied, guarantee the existence of a control barrier function for the unknown system using its learned GP.

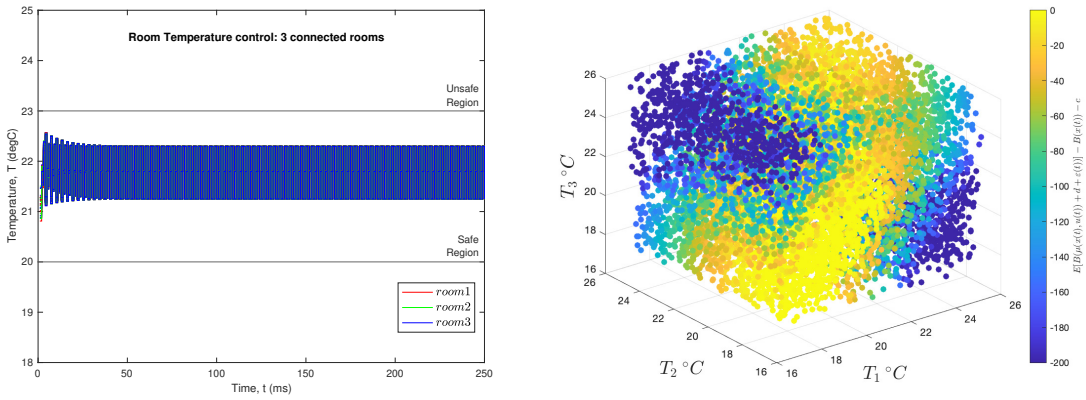
Lemma 7 *Consider a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$ satisfying Assumptions 1-3, where $U = \{u_1, \ldots, u_k\}$ with $u_i \in \mathbb{R}^m$, $\forall i \in \{1,2,\ldots,k\}$. Let $X_o, X_u \subset X$. Suppose there exists a function $B : X \mapsto \mathbb{R}_{\geq 0}$, and constants $c \in \mathbb{R}$, $\lambda \in \mathbb{R}_{>0}$, $\gamma \in \mathbb{R}_{\geq 0}$, where $\gamma > \lambda$, such that the following expression holds*

$$\bigwedge_{x \in X_0} B(x) \leq \lambda \bigwedge_{x \in X_u} B(x) > \gamma \bigwedge_{x \in X} \left(\bigvee_{u \in U} \left(\bigwedge_{d \in \mathcal{D}} \mathbb{E}[B(\mu(x,u)+d+\varepsilon) \mid x,u] - B(x) \leq c \right) \right). \tag{18}$$

Then, B satisfies the conditions in Theorem 6 and any $u : X \mapsto U$ defined as "for any $x \in X$ pick $u \in U$ satisfying $\mathbb{E}[B(\mu(x,u)+d+\varepsilon)] - B(x) \leq c$ for an arbitrary $d \in \mathcal{D}$" is the corresponding control policy.

To employ the CEGIS framework for the computation of $B(x)$ as in Lemma 7, one can consider a function of the parametric form $B(a,x) = \sum_{i=1}^{p} a_i b_i(x)$ with user-defined basis functions $b_i(x)$ and unknown coefficients $a_i \in \mathbb{R}, i \in \{1,2,\ldots,p\}$. Now, one can re-write the feasibility expression from Lemma 7 based on coefficients a_i.

The coefficients a_i can be efficiently found using Satisfiability Modulo Theories (SMT) solvers (De Moura and Bjørner (2008); Gao et al. (2013)). Detailed discussions on the CEGIS approach can be found in Jagtap et al. (2020b).



(a) Temperature evolution under control policy in (21)

(b) Third barrier condition as in (10)

Figure 1: Room temperature control

5. Case Study

5.1. Room temperature control

The safety controller synthesis approach was tested on a model, taken from Meyer et al. (2017), for temperature regulation of a circular building of three connected rooms. We define the dt-SCS for this model as $\mathcal{S} = (X, U, f, \varepsilon)$, where $X = [0, 45]^3$, $U = \{0, 0.6\}^3$, and for each $x = [x_1, x_2, x_3] \in X$, and $u = [u_1, u_2, u_3] \in U$,

$$f_i(x, u) := x_i + \alpha(x_{i+1} + x_{i-1} - 2x_i) + \beta(T_e - x_i) + \eta(T_h - x_i)u_i, \tag{19}$$

$i \in \{1, 2, 3\}$, where f_i represents the i^{th} component of f, x_i represents the temperature (in degrees Celsius) of the i^{th} room, x_{i+1} and x_{i-1} represent the temperatures of the neighbouring rooms (with $x_0 = x_3$ and $x_4 = x_1$), u_i represents the heater input of the i^{th} room. The finite set U corresponds to the heater off and on configurations respectively. The constant $T_h = 50°C$ is the heater temperature, $T_e = -1°C$ is the ambient temperature, and constants $\alpha = 0.045$, $\beta = 0.0045$, and $\eta = 0.09$ are heat exchange coefficients. The process noise is a vector $\varepsilon = [\varepsilon_1, \varepsilon_2, \varepsilon_3]$, where $\varepsilon_i \sim subG(0.01), i \in \{1, 2, 3\}$.

For the safety specification, we consider an initial state set $X_0 = [21, 22]^3$, and an unsafe region $X_u = [0, 20]^3 \cup [23, 45]^3$. For the formal synthesis of a safety controller, first, we model the unknown maps f_i, $i \in \{1, 2, 3\}$, by training three independent GPs. For each GP $f_i \sim \mathcal{GP}_i(0, k_i)$, the kernel k_i is the squared-exponential function (Srinivas et al. (2009)), defined as $k_i((x, u), (x', u')) = \sigma_{f_i}^2 \exp\left(\frac{-\|(x,u)-(x',u')\|^2}{2\sigma_{l_i}^2}\right)$, where σ_{f_i} and σ_{l_i} are the hyper-parameters of the kernel. We assume a measurement noise sequence $w_i \sim subG(1.01)$. We collect $N = 200$ samples of x, u, and $y_i = f_i(x, u) + \varepsilon_i + w_i = f_i(x, u) + \nu_i$ (as in Assumption 3), where $\nu_i \sim subG(R)$, $R = \sqrt{0.01^2 + 1.0071^2}$, $i \in \{1, 2, 3\}$, by simulating the system from multiple initial conditions and inputs chosen randomly from a uniform distribution. Using MATLAB's `fitrgp` module, we obtain the hyper-parameters of the kernel functions, resulting in $\sigma_{f_1}^2 = 560.97$, $\sigma_{f_2}^2 = 560.98$, $\sigma_{f_3}^2 = 560.95$, $\sigma_{l_1} = 1963.70$, $\sigma_{l_2} = 1963.71$, and $\sigma_{l_3} = 1963.66$. As mentioned in Remark 3, computing the information-theoretic term α is a hard problem in general. Thus, we employ Monte-Carlo approach (Asmussen and Glynn (2007)) to obtain a probability bound on the accuracy of the

learned GP provided in Lemma 2. For a fixed error bound $\beta_i\bar{\rho}_i = 0.01$ (i.e. $\mathcal{D} = [-0.01, 0.01]^3$) on the distance between the actual map $f_i(x,u)$ and the learned map $\mu_i(x,u)$, we obtain a probability interval for (7) as $[0.9987, 0.9999]$ with a confidence of $1-10^{-10}$ using 10^6 realizations. The lower bound $(1-\delta)^3$ in Lemma 2 can thus be chosen as 0.9987.

In the next step, a polynomial-type control barrier function is obtained using the CEGIS approach as described in Section 4.1. The barrier function is computed as

$$\begin{aligned} B(x) = {} & 9.506219x_1^2 + 11.369872x_2^2 + 11.847413x_3^2 - 8.073953x_1x_2 - 9.339144x_1x_3 \\ & +13.856766x_2x_3 - 35.793193x_1 - 16.10336x_2 - 10.319252x_3 + 662.86428, \end{aligned} \tag{20}$$

resulting in $\lambda = 0.5$, $\gamma = 11$, and $c = -0.55$. The corresponding control policy is chosen as

$$u(x) = \arg\min_{u\in U} ||u||, \text{ subject to: } \mathbb{E}[B(\mu(x,u)+d+\varepsilon)] - B(x) \leq c, \tag{21}$$

for an arbitrarily chosen $d \in [-0.01, 0.01]^3$. For a bounded time horizon $[0, T]$, where T = 5s, this results in a probability lower bound of 0.9545 in (15) that holds with a confidence of at least 0.9987. Figure 1(a) shows a few realizations of the evolution of temperature of the three rooms under this policy and Figure 1(b) shows the satisfaction of the last condition of the barrier certificate for the learned model.

5.2. Vehicle model

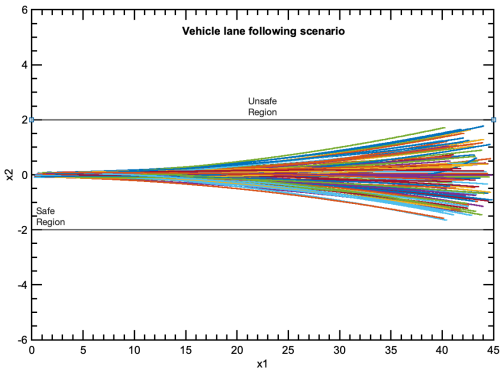
Here, we consider a vehicle lane keeping example in which the constant velocity kinematic single-track model of a vehicle (BMW 320i) is used. The discrete-time version of the dynamics (Jagtap et al. (2020b)) is formulated as a dt-SCS $\mathcal{S} = (X, U, f, \varepsilon)$ where $X = [0, 50] \times [-6, 6] \times [-0.05, 0.05] \times [-0.1, 0.1]$, $U = \{-0.5, 0, 0.5\}$, and for each $x = [x_1, x_2, x_3, x_4] \in X$, and $u \in U$,

$$\begin{aligned} f_1(x,u) &:= x_1 + \tau_s v \cos(x_4), \\ f_2(x,u) &:= x_2 + \tau_s v \sin(x_4), \\ f_3(x,u) &:= x_3 + \tau_s u, \\ f_4(x,u) &:= x_4 + \tau_s \frac{v}{l_{wb}}. \tan(x_3). \end{aligned} \tag{22}$$

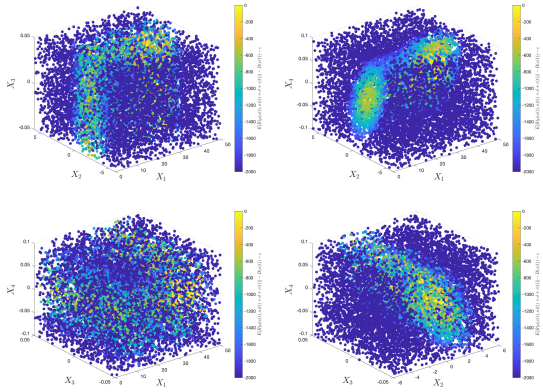
States x_1 and x_2 are the (Cartesian) position coordinates, x_2 is the steering angle, and x_4 is the heading angle of the vehicle. The constants in the above equations are sampling time $\tau_s = 0.01$ s, forward velocity $v = 10$ m/s, and length of wheel base $l_{wb} = 2.578$ m. The process noise is $\varepsilon = [\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4]$, where each $\varepsilon_i \sim subG(0.01), i \in \{1, 2, 3, 4\}$.

For the safety specification, we consider an initial state set $X_0 = [0, 5] \times [-0.1, 0.1] \times [-0.005, 0.005] \times [-0.01, 0.01]$, and an unsafe set $X_u = X_{u_1} \cup X_{u_2}$ where $X_{u_1} = [0, 50] \times [-6, -2] \times [-0.05, 0.05] \times [-0.1, 0.1]$, and $X_{u_2} = [0, 50] \times [2, 6] \times [-0.05, 0.05] \times [-0.1, 0.1]$.

We model the unknown maps $f_i, i \in \{1, 2, 3, 4\}$, by training four independent GPs. For each GP $f_i \sim \mathcal{GP}_i(0, k_i)$, the kernel k_i is the squared-exponential function (Srinivas et al. (2009)).We collect $N = 500$ samples of x, u, and $y_i = f_i(x,u) + \varepsilon_i + w_i = f_i(x,u) + \nu_i$ (as in Assumption 3), where $\nu_i \sim subG(R)$, $R = \sqrt{0.01^2 + 1.004^2}$, $i \in \{1, 2, 3, 4\}$, by simulating the system from multiple initial conditions and inputs chosen randomly from a uniform distribution. As in the previous case study, we obtain the hyper-parameters of the kernel functions using MATLAB's `fitrgp` module, resulting in $\sigma_{f_1}^2 = 55.2606$, $\sigma_{f_2}^2 = 165.8472$, $\sigma_{f_3}^2 = 1499.86$, $\sigma_{f_4}^2 = 0.0423$, $\sigma_{l_1} = 107.4961$, $\sigma_{l_2} = 359.8126$, $\sigma_{l_3} = 3653.275$, and $\sigma_{l_4} = 0.1461$. Using the Monte-Carlo sampling approach,



(*a*) Evolution of the position (states x_1 and x_2) of the vehicle using the controller in (24)



(*b*) Third barrier condition as in (10)

Figure 2: Vehicle lane following control

for a fixed error bound $\beta_i\bar{\rho}_i = 0.1$ (i.e. $\mathcal{D} = [-0.1, 0.1]^4$) on the distance between the actual map $f_i(x,u)$ and the learned map $\mu_i(x,u)$, $i \in \{1,\ldots,4\}$, we obtain a probability interval for (7) as $[0.9201, 0.9371]$ with a confidence of $1-10^{-10}$ using 10^6 realizations. The lower bound $(1-\delta)^3$ can be thus chosen as 0.9201. In the next step a polynomial control barrier function given below was obtained using the CEGIS approach as described in Section 4.1:

$$\begin{aligned} B(x) = {} & 90.45852x_1^2 + 19166.082475x_2^2 - 53595.4041x_3^2 + 185791.32465x_4^2 - 1879.96065x_1x_2 \\ & - 1896.165937x_1x_3 - 271.77839x_1x_4 - 7526.572264x_2x_3 + 6769.374605x_2x_4 \\ & + 26855.30431x_3x_4 - 1744.956593x_1 + 22161.708246x_2 - 69823.820245x_3 \\ & + 2474.137174x_4 + 5847.015382, \end{aligned} \tag{23}$$

resulting in $\lambda = 0.5$, $\gamma = 37610$, and $c = 469.3125$. The corresponding control policy is chosen as

$$u(x) = \arg\min_{u\in U} ||u||, \text{ subject to: } \mathbb{E}[B(\mu(x,u)+d+\varepsilon)] - B(x) \leq c, \tag{24}$$

for an arbitrarily chosen $d \in [-0.1, 0.1]^4$. For a bounded time horizon $[0, T]$, where T = 4s, this results in a probability lower bound 0.9501 in (15) that holds with a confidence of at least 0.9201. Figure 2(a) shows a few realizations of the position of the vehicle under this control policy and Figure 2(b) shows the last condition of barrier certificate for the learned model.

6. Conclusions

In this work, we proposed a discretization-free approach for formal synthesis of safety controllers for fully unknown stochastic systems using Gaussian process learning and control barrier certificates. In the future, we aim to extend this approach to include more complex properties. Also, we plan on exploring the idea of computation of the barrier certificates using constrained Gaussian processes which allows us to combine the learning and the barrier computation steps simultaneously while preserving the formal guarantees.

Acknowledgments

This work was supported in part by the NSF under grant CNS-2039062 and the H2020 ERC Starting Grant AutoCPS (grant agreement No 804639).

References

Anayo K Akametalu, Jaime F Fisac, Jeremy H Gillula, Shahab Kaynama, Melanie N Zeilinger, and Claire J Tomlin. Reachability-based safe learning with Gaussian processes. In *53rd IEEE Conference on Decision and Control*, pages 1424–1431. IEEE, 2014.

Aaron D Ames, Samuel Coogan, Magnus Egerstedt, Gennaro Notomista, Koushil Sreenath, and Paulo Tabuada. Control barrier functions: Theory and applications. In *2019 18th European Control Conference (ECC)*, pages 3420–3431. IEEE, 2019.

Søren Asmussen and Peter W Glynn. *Stochastic simulation: algorithms and analysis*, volume 57. Springer Science & Business Media, 2007.

Guillaume O. Berger, Raphaël M. Jungers, and Zheming Wang. Chance-constrained quasi-convex optimization with application to data-driven switched systems control. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, pages 571–583, 07 – 08 June 2021.

Christopher M Bishop. *Pattern recognition and machine learning*. Springer, 2006.

Urs Borrmann, Li Wang, Aaron D Ames, and Magnus Egerstedt. Control barrier certificates for safe swarm behavior. *IFAC-PapersOnLine*, 48(27):68–73, 2015.

Giuseppe Carlo Calafiore and Marco C Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

Fernando Castañeda, Jason J Choi, Bike Zhang, Claire J Tomlin, and Koushil Sreenath. Pointwise feasibility of Gaussian process-based safety-critical control under model uncertainty. *arXiv preprint:2106.07108*, 2021.

Girish Chowdhary, Hassan A Kingravi, Jonathan P How, and Patricio A Vela. Bayesian nonparametric adaptive control using Gaussian processes. *IEEE Transactions on Neural Networks and Learning Systems*, 26(3):537–550, 2014.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853. PMLR, 2017.

Leonardo De Moura and Nikolaj Bjørner. Z3: An efficient SMT solver. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 337–340. Springer, 2008.

Sicun Gao, Soonho Kong, and Edmund M Clarke. dReal: An SMT solver for nonlinear theories over the reals. In *International Conference on Automated Deduction*, pages 208–214. Springer, 2013.

John Jackson, Luca Laurenti, Eric Frew, and Morteza Lahijanian. Safety verification of unknown dynamical systems via Gaussian process regression. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 860–866. IEEE, 2020.

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.

Pushpak Jagtap, George J. Pappas, and Majid Zamani. Control barrier functions for unknown nonlinear systems using Gaussian processes. In *59th IEEE Conference on Decision and Control (CDC)*, pages 3699–3704, December 2020a. doi: 10.1109/CDC42340.2020.9303847. ISSN: 2576-2370.

Pushpak Jagtap, Sadegh Soudjani, and Majid Zamani. Formal synthesis of stochastic systems via control barrier certificates. *IEEE Transactions on Automatic Control*, 2020b. ISSN 1558-2523. doi: 10.1109/TAC.2020.3013916.

Xiao Li and Calin Belta. Temporal logic guided safe reinforcement learning using control barrier functions. *arXiv preprint:1903.09885*, 2019.

Pierre-Jean Meyer, Antoine Girard, and Emmanuel Witrant. Compositional abstraction and safety synthesis using overlapping symbolic models. *IEEE Transactions on Automatic Control*, 63(6): 1835–1841, 2017.

Vern I Paulsen and Mrinal Raghupathi. *An introduction to the theory of reproducing kernel Hilbert spaces*, volume 152. Cambridge university press, 2016.

Stephen Prajna, Ali Jadbabaie, and George J. Pappas. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control*, 52(8): 1415–1428, August 2007. ISSN 1558-2523. doi: 10.1109/TAC.2007.902736.

Hadi Ravanbakhsh and Sriram Sankaranarayanan. Counterexample guided synthesis of switched controllers for reach-while-stay properties. *arXiv preprint:1505.01180*, 2015.

Ali Salamati, Abolfazl Lavaei, Sadegh Soudjani, and Majid Zamani. Data-driven safety verification of stochastic systems via barrier certificates. *IFAC-PapersOnLine*, 54(5):7–12, 2021.

Matthias W Seeger, Sham M Kakade, and Dean P Foster. Information consistency of nonparametric Gaussian process methods. *IEEE Transactions on Information Theory*, 54(5):2376–2382, 2008.

Armando Solar-Lezama, Liviu Tancau, Rastislav Bodik, Sanjit Seshia, and Vijay Saraswat. Combinatorial sketching for finite programs. In *Proceedings of the 12th International Conference on Architectural Support for Programming Languages and Operating Systems*, pages 404–415, 2006.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint:0912.3995*, 2009.

Jonas Umlauft, Thomas Beckers, Melanie Kimmel, and Sandra Hirche. Feedback linearization using Gaussian processes. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 5249–5255. IEEE, 2017.

Jonas Umlauft, Lukas Pöhler, and Sandra Hirche. An uncertainty-based control lyapunov approach for control-affine systems modeled by Gaussian process. *IEEE Control Systems Letters*, 2(3): 483–488, 2018.

Christopher K Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

Guang Yang, Calin Belta, and Roberto Tron. Continuous-time signal temporal logic planning with control barrier functions. In *2020 American Control Conference (ACC)*, pages 4612–4618. IEEE, 2020.