2018 IEEE International Conference on Robotics and Automation (ICRA)
May 21-25, 2018, Brisbane, Australia

Safe Learning of Quadrotor Dynamics Using Barrier Certificates *

Li Wang, Evangelos A. Theodorou, and Magnus Egerstedt†

Abstract— **To effectively control complex dynamical systems, accurate nonlinear models are typically needed. However, these models are not always known. In this paper, we present a data-driven approach based on Gaussian processes that learns models of quadrotors operating in partially unknown environments. What makes this challenging is that if the learning process is not carefully controlled, the system will go unstable, i.e., the quadcopter will crash. To this end, barrier certificates are employed for safe learning. The barrier certificates establish a non-conservative forward invariant safe region, in which high probability safety guarantees are provided based on the statistics of the Gaussian Process. A learning controller is designed to efficiently explore those uncertain states and expand the barrier certified safe region based on an adaptive sampling scheme. Simulation results are provided to demonstrate the effectiveness of the proposed approach.**

I. Introduction

Safety is crucial to many physical control dynamical systems, such as autonomous vehicles, industrial robots, chemical reactors, and air-traffic control systems [2], [4]. The existence of model inaccuracies and unknown disturbances poses a great challenge to the design of safe controllers for these systems.

Tools such as robust control and adaptive control methods have been developed in classic control theory to ensure the safety and stability of the system, see [8] and the references therein. Meanwhile, machine learning based control approaches are becoming increasingly popular as a way to deal with inaccurate models [7], [13], due to their abilities to infer unknown models from data and actively improve the performance of the controller with the learned model. In contrast to classic control methods, learning based control approaches require only limited expert knowledge and fewer assumptions about the system [17]. However, there always exists an inherent trade-off between safety and performance in these methods [3]. Data-driven learning approaches rarely provides safety guarantees, which limits their applicability to real-world safety critical control dynamical systems [17].

A number of safe learning approaches have been proposed in the literature. Among these methods, the use of learning Control Lyapunov Functions (CLF) is shown to be a promising approach. [16] introduced a verifier to explicitly validate the learned CLF. However, when the model of the system is inaccurate, the verifier must check an infinite number of inequalities throughout the state space, which is computationally difficult [10]. [5] seeks to learn CLF and maximize the safe operation region for the system with GP model. High probability safety guarantees are provided based on Lyapunov stability and GP statistics. In addition, a reachability-based safe learning approach was presented in [1] to reduce the conservativeness of reachability analysis by learning the disturbance from data. Due to the complexity of the HJI reachability analysis, it is often computationally expensive to calculate the robust forward reachable set for a sufficiently large time horizon.

In contrast to the aforementioned methods, this paper interprets the safe operation region as general invariant sets established with barrier certificates, which permits a much richer set of safe control options, rather than Lyapunov sub-level sets. The barrier certificates formally define a forward invariant safe region, where all system trajectories starting in this region remains in this region for all time [14], [21], [2]. Barrier certificates were successfully applied to many safety critical dynamical systems [12], [18], [19]. In this paper, we construct a safe operation region with barrier certificates, and gradually expand the certified safe region as the uncertainty of the system reduces. The unknown dynamics of the system is represented with a GP model, which provides both the mean and variance of the prediction. Using the statistics of GP model, a high probability safety guarantee of the system with inaccurate model is provided. The search for maximum volume barrier certificates involves the validation of an infinite number of inequality constraints, which is computationally expensive. Inspired by the discrete sampling technique used in [4], we design an adaptive sampling algorithm to significantly reduce the computation intensity, i.e., the more certain regions in the state space are sampled less without loss of safety guarantees.

The main contributions of this paper are threefold. First, a safe learning strategy is developed based on barrier certificates, which admits a rich set of learning control options. Second, an adaptive sampling algorithm is proposed to significantly reduce the computation intensity of the learning process. Third, a recursive learning strategy based on GP is presented to learn the complex 3D nonlinear quadrotor dynamics online.

II. Preliminaries of Barrier Certificates and Gaussian Process

Preliminary results regarding the two fundamental tools, i.e., barrier certificates and Gaussian Process, used to formulate the safe learning strategy are presented in this section.

*The work by the first and third authors was sponsored by Grant No. 1544332 from the U.S. National Science Foundation.

†Li Wang and Magnus Egerstedt are with the School of Electrical and Computer Engineering, Evangelos A. Theodorou is with the School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. Email: {liwang, magnus, evangelos.theodorou}@gatech.edu

978-1-5386-3081-5/18/$31.00 ©2018 IEEE

A. Barrier Certificates and Set Invariance

Consider a control affine dynamical system

$$\dot{x} = f(x) + g(x)u, \tag{1}$$

where $x \in \mathscr{X} \subseteq \mathbb{R}^n$ and $u \in \mathscr{U} \subseteq \mathbb{R}^m$ are the state and control of the system, $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are Lipschitz continuous. Let the safe set of the system be encoded as the superlevel set of a smooth function $h : \mathbb{R}^n \to \mathbb{R}$,

$$\mathscr{C} = \{x \in \mathbb{R}^n \mid h(x) \geq 0\}. \tag{2}$$

The function $h(x)$ is termed a Control Barrier Function (CBF), if there exists an extended class-κ function ($\kappa(0) = 0$ and strictly increasing) such that

$$\sup_{u \in \mathscr{U}} \left\{ \frac{\partial h}{\partial x} f(x) + \frac{\partial h}{\partial x} g(x)u + \kappa(h(x)) \right\} \geq 0,$$

for all $x \in \mathscr{E}$ with $\mathscr{C} \subseteq \mathscr{E}$.

Given a CBF, the barrier certified safe control space $S(x)$ is defined as

$$S(x) = \left\{ u \in U \mid \frac{\partial h}{\partial x} f(x) + \frac{\partial h}{\partial x} g(x)u + \kappa(h(x)) \geq 0 \right\}, \; x \in \mathscr{E}.$$

With barrier certificates, the invariance property of $\mathscr{C}$ is established with the following theorem,

Theorem [21]: *Given a set $\mathscr{C} \subset \mathbb{R}^n$ defined by (2) and a CBF h defined on $\mathscr{E}$, with $\mathscr{C} \subseteq \mathscr{E} \subset \mathbb{R}^n$, any Lipschitz continuous controller u: $\mathscr{E} \to \mathbb{R}$ such that $u \in S(x)$ for the system (1) renders the set $\mathscr{C}$ forward invariant.*

This type of barrier certificates expands the certified safe control space significantly by allowing $h(x)$ to decrease within $\mathscr{C}$ as opposed to strictly increasing [21], [2]. Compared with Lyapunov sublevel set based safe region, barrier certificates provide a more permissive notion of safety. This fact can be illustrated with the following example.

Example 1: Consider an autonomous dynamical system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 + 0.8x_2^2 \\ -x_1 - x_2 + x_1^2 x_2 \end{bmatrix}, \tag{3}$$

the safe region of this system is estimated with both the Lyapunov sublevel set and barrier certificates.

Since (3) is a polynominal system, the safe sets can be computed directly with Sum-of-Squares programs using YALMIP [11] and SMRSOFT [6] solvers. The safe region estimated with the optimal polynomial Lyapunov function is

$$\mathscr{A}_1 = \{x \mid V^*(x) \leq 1\},$$

where $V^*(x) = 1.343x_1^2 + 0.5155x_1x_2 + 1.152x_2^2$.

The safe region estimated with barrier certificates is

$$\mathscr{A}_2 = \{x \mid h^*(x) \geq 0\},$$

where $h^*(x) = 1 - 0.4254x_1 - 0.3248x_2 - 0.7549x_2^2 - 0.8616x_1^2 - 0.2846x_1x_2$.

From Fig. 1, it can be observed that the barrier certified safe region $\mathscr{A}_2$ is much larger than the Lyapunov based safe region $\mathscr{A}_1$. Consequently, safe learning controller based on barrier certificates are allowed to explore more states of the system. In this paper, we will leverage the non-conservative safety guarantee of barrier certificates to allow a much richer set of safe learning control options.

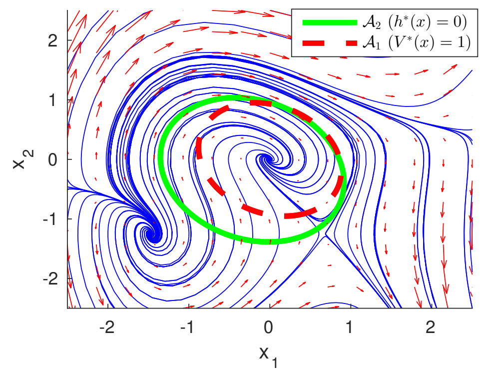


Fig. 1: Estimates of safe regions for system (3). The regions enclosed by the dashed red ellipse and solid green ellipse are estimated safe regions with optimal polynomial Lyapunov function $V^*(x)$ and barrier certificates $h^*(x)$, respectively.

B. Gaussian Processes

A GP is a nonparametric regression method that can capture complex unknown functions [15]. With a GP, every point in the state space is associated with a normally distributed random variable, which allows us to derive high probability statements about the system.

We now consider a system with partially unknown dynamics $d(x)$ in this paper, i.e.,

$$\dot{x} = f(x) + g(x)u + d(x), \tag{4}$$

where $x \in \mathscr{X} \subseteq \mathbb{R}^n$ and $u \in \mathscr{U} \subseteq \mathbb{R}^m$ are the state and control of the system. It is assumed that $d(x)$ is Lipschitz continuous. Although the proposed method applies to general dynamical systems, here we restrict our attention to the class of systems that can be addressed with existing computation tools.

Since the unmodeled dynamics $d(x)$ is n dimensional, each dimension is approximated with a GP model $\mathscr{GP}(0, k(x, x'))$ with a prior mean of zero and a covariance function of $k(x, x')$, where $k(x, x')$ is the kernel function to measure the similarity between any two states $x, x' \in \mathscr{X}$. The measurement of $d(x)$ is obtained indirectly by subtracting the inaccurate model prediction $[f(x) + g(x)u]$ from the noisy measurement of the system dynamics $[\dot{x} + \mathscr{N}(0, \sigma_n^2)]$.

Given a collection of w measurements $y_w = [\hat{d}(x_1), \hat{d}(x_2), \ldots, \hat{d}(x_w)]^T$, the mean $m(x_*)$ and variance $\sigma^2(x_*)$ of $d(x_*)$ at the query state x_* are

$$m(x_*) = k_*^T (K + \sigma_n^2 I)^{-1} y_w, \tag{5}$$

$$\sigma^2(x_*) = k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_*, \tag{6}$$

where $\lfloor K \rfloor_{(i,j)} = k(x_i, x_j)$ is the kernel matrix, and $k_* = [k(x_1, x_*), k(x_2, x_*), \ldots, k(x_w, x_*)]^T$ [15].

With the learned system dynamics based on GP, a high probability confidence interval of the unmodeled dynamics $d(x)$ can be established as

$$\mathscr{D}(x) = \{d \mid m(x) - k_\delta \sigma(x) \leq d \leq m(x) + k_\delta \sigma(x)\}, \tag{7}$$

where k_δ is a design parameter to get $(1 - \delta)$ confidence, $\delta \in (0, 1)$.

III. Safe Learning with Barrier Certificates

In order to ensure that the learning based controller never enters the unsafe region, we will learn barrier certificates for the system and use the learned certificates to regulate the controller. Starting with a conservative certificate with certified safe region $\mathscr{C}_0(x)$, this certified safe region is gradually expanded with new data until it stops growing. This incremental learning process is visualized in Fig. 2.

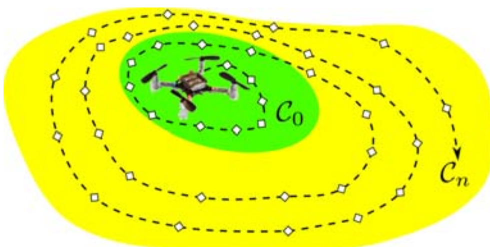


Fig. 2: Incremental learning of the barrier certificates. The green region $\mathscr{C}_0$ and the yellow regions $\mathscr{C}_n$ are the initial and final barrier certified safe regions, respectively. The barrier certified safe region gradually grows as more and more data points are sampled in the state space.

More concretely, the goal of the learning process is to maximize the volume of the barrier certified safe region $\mathscr{C}$ by adjusting $h(x)$, i.e.,

$$\begin{aligned} &\max_{h(x)} \quad \mathrm{vol}(\mathscr{C}) \\ &\text{s.t.} \quad \max_{u\in\mathscr{U}} \min_{d\in\mathscr{D}(x)} \left\{ \frac{\partial h}{\partial x}(f(x)+g(x)u+d)+\gamma h(x) \right\} \geq 0, \\ &\qquad\qquad \forall x \in \mathscr{C}. \end{aligned}$$

Since u and d are independent from each other and the high confidence interval for disturbance is $\mathscr{D}(x)$ in (7), the barrier certificates constraint can be considered as

$$\begin{aligned} &\max_{h(x)} \quad \mathrm{vol}(\mathscr{C}) \\ &\text{s.t.} \quad \max_{u\in\mathscr{U}} \left\{ \frac{\partial h}{\partial x} g(x)u \right\} + \frac{\partial h}{\partial x} m(x) - k_\delta \left| \frac{\partial h}{\partial x} \right| \sigma(x) \\ &\qquad + \frac{\partial h}{\partial x} f(x) + \gamma h(x) \geq 0, \forall x \in \mathscr{C}. \end{aligned} \tag{8}$$

When more data points are collected about the system dynamics, the uncertainty $\sigma(x)$ will gradually decrease. As a result, more states will satisfy the barrier certificates constraint. The goal of exploration is to actively collect data to reduce $\sigma(x)$ and maximize the volume of $\mathscr{C}$.

It should be pointed out that the barrier certified region maximization problem (8) is a non-convex, infinite dimensional optimization problem, which is intractable to solve in practice. We will make two simplifications to make it solvable, namely by employing adaptive sampling of the state space and parameterization of the shape of $\mathscr{C}$.

A. Adaptive Sampling of the State Space

Due to the Lipschitz continuity of the system dynamics, the safety of the system in $\mathscr{X}$ can be evaluated by only sampling a finite number of points in $\mathscr{X}$. Inspired by [4], we will show that we can adaptively sample the state space without losing safety guarantees. Similar to *Lemma 4* in [4], it can be shown that $h(x)$ and $\dot{h}(x)$ are Lipschitz continuous in x with Lipschitz constants L_h and $L_{\dot{h}}$, respectively.

Let $\mathscr{X}_\tau \subset \mathscr{X}$ be a discretization of the state space $\mathscr{X}$. The closest point in $\mathscr{X}_\tau$ to $x \in \mathscr{X}$ is denoted as $[x]_\tau$, where $\|x-[x]_\tau\| \leq \frac{\tau}{2}$.

Lemma 3.1: If the following holds for all $x \in \mathscr{X}_\tau$,

$$\begin{aligned} &\max_{u\in\mathscr{U}} \left\{ \frac{\partial h}{\partial x} g(x)u \right\} + \frac{\partial h}{\partial x} m(x) - k_\delta \left| \frac{\partial h}{\partial x} \right| \sigma(x) \\ &\qquad + \frac{\partial h}{\partial x} f(x) + \gamma h(x) \geq (L_{\dot{h}} + \gamma L_h)\tau, \end{aligned} \tag{9}$$

then the safety barrier constraint

$$\max_{u\in\mathscr{U}} \min_{d\in\mathscr{D}(x)} \left\{ \frac{\partial h}{\partial x}(f(x)+g(x)u+d)+\gamma h(x) \right\} \geq 0 \tag{10}$$

is satisfied for all $x \in \mathscr{X}$ with probability $(1-\delta)$, $\delta \in (0,1)$.

Proof: See [20]. ■

With the discretization of the state space, we only need to sample a finite number of points to validate the barrier certificates. However, the number of required sampling points is still very large. The following adaptive sampling strategy further reduces the number of sampling points required.

Proposition 3.2: If the following is satisfied at $x \in \mathscr{X}$,

$$\begin{aligned} &\max_{u\in\mathscr{U}} \left\{ \frac{\partial h}{\partial x} g(x)u \right\} + \frac{\partial h}{\partial x} m(x) - k_\delta \left| \frac{\partial h}{\partial x} \right| \sigma(x) \\ &\qquad + \frac{\partial h}{\partial x} f(x) + \gamma h(x) \geq (L_{\dot{h}} + \gamma L_h)k_\tau\tau, \end{aligned} \tag{11}$$

with $k_\tau \geq 0$, then the safety barrier constraint (10) is satisfied for all $y \in \mathscr{X}$ such that $\|x-y\| \leq k_\tau\tau$.

Proof: The proof is similar to *lemma* 3.1. ■

Leveraging the Lipschitz continuity of the barrier certificates, we can adaptively sample the state space without losing safety guarantees. Sparse sampling is performed at places with large safety margin, while dense sampling is only required at places with small safety margin.

B. Parameterization of the Barrier Certificates

Because maximizing the volume of $\mathscr{C}$ is a non-convex problem in general, we can parameterize the barrier certificate $h_\mu(x)$ with μ to simplify the optimization problem. For example, $h_\mu(x)$ can be formulated as $1-Z(x)^T\mu Z(x)$, where $Z(x)$ is the vector of monomials, and μ is a positive semi-definite matrix. Then maximizing $\mathrm{vol}(\mathscr{C})$ is equivalent to minimize the trace of μ. Further simplification can be made to fix the shape of $\mathscr{C}$ (by optimizing only with the known dynamics) and enlarge the level set of barrier certificates.

With the shape parameterization and adaptive sampling technique, the barrier certificate maximization problem (8) can be written as

$$
\begin{aligned}
&\max_{\mu} \quad \mathrm{vol}(\mathscr{C}) \\
&\text{s.t.} \quad \max_{u \in \mathscr{U}} \left\{ \frac{\partial h_\mu}{\partial x} g(x) u \right\} + \frac{\partial h_\mu}{\partial x} m(x) - k_\delta \left| \frac{\partial h_\mu}{\partial x} \right| \sigma(x) \\
&\qquad + \frac{\partial h_\mu}{\partial x} f(x) + \gamma h_\mu(x) \geq (L_{\dot{h}} + \gamma L_h)\tau, \forall x \in \mathscr{C} \cap \mathscr{X}_\tau .
\end{aligned} \tag{12}
$$

In order to increase the learning efficiency during the exploration phase, the most uncertain state in $\mathscr{C}$ is sampled,

$$
x_{\text{next}} = \underset{x \in \mathscr{C} \cap \mathscr{X}_\tau}{\operatorname{argmax}} \quad \sigma(x). \tag{13}
$$

It is assumed that a nominal exploration controller $\hat{u}$ can be designed to drive the system from the current state x to x_{next}, i.e., $\hat{u} = GoTo(x, x_{\text{next}})$. Then the safety barrier certificates are enforced through a QP-based controller to "rectify" the nominal control such that the system is always safe,

$$
\begin{aligned}
u^* = &\underset{u \in \mathscr{U}}{\operatorname{argmin}} \quad J(u) = \|u - \hat{u}\|^2 \\
&\text{s.t.} \qquad \frac{\partial h}{\partial x} g(x) u + \frac{\partial h}{\partial x} m(x) - k_\delta \left| \frac{\partial h}{\partial x} \right| \sigma(x) \\
&\qquad\qquad + \frac{\partial h}{\partial x} f(x) + \gamma h(x) \geq 0.
\end{aligned} \tag{14}
$$

Therefore, the actual exploration controller u^* tries to stay as close as possible to the desired controller $\hat{u}$, while always honoring the safety requirements. The exploration phase ends when the safe region $\mathscr{C}$ does not grow any more.

C. *Overview of the Safe Learning Algorithm*

An overview of the barrier certificates based safe learning algorithm is provided in **Algorithm 1**. At the beginning, a conservative barrier certified safe region $\mathscr{C}_0$ is provided. The most uncertain state x_{next} is computed based on the current GP model. Then, the QP based controller (14) is used to ensure that the system is driven to x_{next} without ever leaving $\mathscr{C}_n$. After updating the GP model with the sampled data at x_{next}, the barrier certificate optimization problem (12) is solved. The adaptive sampling technique (11) is adopted here to reduce the number of states to be sampled. This process is repeated until the safe region $\mathscr{C}_n$ stops growing.

Algorithm 1 Barrier Certificates based Safe Learning

Input: Initial safe set $\mathscr{C}_0 \subseteq \mathscr{X}$, GP model $\mathscr{GP}(0, k(x, x'))$, discretization $\mathscr{X}_\tau$, tolerance ε
Output: Final safe set $\mathscr{C}_n$
Initialization : $n = 0, x = x_0$

repeat
  $n = n + 1$
  Find x_{next} with (13)
  Design nominal controller $\hat{u} = GoTo(x, x_{\text{next}})$
  Drive to x_{next} with (14)
  Sample x_{next}, update GP
  Expand vol($\mathscr{C}_n$) with (12)
until vol($\mathscr{C}_n$)-vol($\mathscr{C}_{n-1}$)$\leq \varepsilon$
return $\mathscr{C}_n$

IV. Online Learning of Quadrotor Dynamics

The safe learning approach developed in Section III relies on a learning controller that drives the system to explore interested states. The challenge of designing this learning controller is that the 3D quadrotor system considered in this paper is highly nonlinear and unstable. In this section, we will present a recursive learning controller based on GP to learn the complex quadrotor dynamics online.

A. *Differential Flatness of 3D Quadrotor Dynamics*

The quadrotor coordinate frames and Euler angles (roll ϕ, pitch θ, and yaw ψ) are illustrated in Fig. 3 [22]. The world, body, and intermediate frames (after yaw angle rotation) are denoted by the subscripts w, b, and c, respectively.

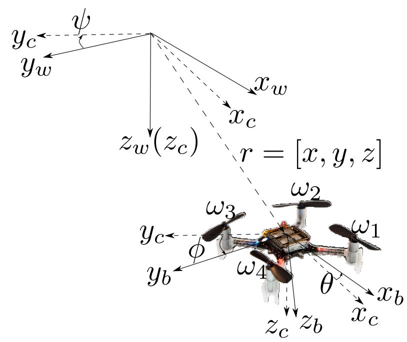


Fig. 3: Quadrotor coordinate frames.

Here, we adopt the quadrotor model used in [9] to describe the 3D nonlinear quadrotor dynamics,

$$
\begin{cases}
\ddot{r} = g z_w + \frac{1}{m} R z_w f_z, \\
\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & s\phi t\theta & c\phi t\theta \\ 0 & c\phi & -s\phi \\ 0 & s\phi sc\theta & c\phi sc\theta \end{bmatrix} \omega,
\end{cases} \tag{15}
$$

where $z_w = [0 \ 0 \ 1]^T$, and $r = [x, y, z]^T$, m, and g are the position of the center of mass, the mass, and the gravitational acceleration of the quadrotor, respectively. $s\theta$, $c\theta$, $t\theta$, and $sc\theta$ are short for $\sin\theta$, $\cos\theta$, $\tan\theta$ and $\sec\theta$. The control inputs of the quadrotor are the body rotational rates ($\omega = [\omega_x, \omega_y, \omega_z]^T$) and the thrust ($f$). R is the rotation matrix from the body frame to the world frame with the ZYX convention.

Similar to [22], the dynamics in (15) is differentially flat with the flat output chosen as $\eta = [r^T, \psi^T]^T$. With the differential flatness property, quadrotor trajectory planning can be simplified as smooth parametric curves. Given a desired trajectory $\eta_d(t) \in C^3$ that is three times differentiable, the feed forward control $u_{FF} = [f_{FF}, \omega_{FF}^T]$ can be derived by inverting the dynamics in (15). In addition, the unknown model error and tracking error need to be handled by a feedback control u_{FB}. The actual control applied to the quadrotor is $u = u_{FF} + u_{FB}$ as shown in [20].

B. *Learning based Control Using Gaussian Process*

The model in (15) neglects parameter inaccuracy and the uncertain effects of damping, drag force, and wind disturbances. Here, we will use GP models to learn the unmodeled dynamics. The unmodeled dynamics can be captured with six GPs along each dimension in the state space, i.e.,

$$\begin{cases} \ddot{r} \quad = g z_w + \frac{1}{m} R z_w f_z + \begin{bmatrix} \mathscr{GP}_1(0, k(q,q')) \\ \mathscr{GP}_2(0, k(q,q')) \\ \mathscr{GP}_3(0, k(q,q')) \end{bmatrix}, \\ \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & s\phi t\theta & c\phi t\theta \\ 0 & c\phi & -s\phi \\ 0 & s\phi sc\theta & c\phi sc\theta \end{bmatrix} \omega + \begin{bmatrix} \mathscr{GP}_4(0, k(q,q')) \\ \mathscr{GP}_5(0, k(q,q')) \\ \mathscr{GP}_6(0, k(q,q')) \end{bmatrix}, \end{cases}$$

where the input to the GPs is $q = [r^T, \dot{r}^T, \theta, \phi, \psi]^T$, and the observations for the GPs are $s = [\ddot{r}^T, \dot{\phi}, \dot{\theta}, \dot{\psi}]^T$, respectively. At a new query point q_*, the mean $m_i(q_*)$ and variance $\sigma_i^2(q_*)$ of the unknown dynamics can be inferred with (5). Based on the learned dynamics, a differential flatness based feed forward controller can be derived as,

$$\begin{cases} f_{FF} \;\; = \;\; -m\|\ddot{r}_d - [m_1(q), m_2(q), m_3(q)]^T - g z_w\|, \\ \omega_{FF} \;\; = \;\; \begin{bmatrix} 1 & 0 & -s\theta_d \\ 0 & c\phi_d & s\phi_d c\theta_d \\ 0 & -s\phi_d & c\phi_d c\theta_d \end{bmatrix} \begin{bmatrix} \dot{\phi}_d - m_4(q) \\ \dot{\theta}_d - m_5(q) \\ \dot{\psi}_d - m_6(q) \end{bmatrix}, \end{cases}$$

where $\theta_d = \operatorname{atan2}(\bar{\beta}_a, \bar{\beta}_b)$, $\phi_d = \operatorname{atan2}(\bar{\beta}_c, \sqrt{\bar{\beta}_a^2 + \bar{\beta}_b^2})$, $\bar{\beta}_a = -(\ddot{x}_d - m_1(q)) \cos\psi_d - (\ddot{y}_d - m_2(q)) \sin\psi_d$, $\bar{\beta}_b = -(\ddot{z}_d - m_3(q)) + g$, and $\bar{\beta}_c = -(\ddot{x}_d - m_1(q)) \sin\psi_d + (\ddot{y}_d - m_2(q)) \cos\psi_d$.

C. Recursive Online GP Learning

One issue with the GP regression is that the time complexity of GP inference is $O(N^3)$, where N is the number of data points. The majority of the time is used to compute the inverse of the kernel matrix K. Here, we propose a recursive online GP Learning method to compute the exact GP inference online. As the quadrotor moves forward, we will actively add relevant data points and remove least relevant data points into the kernel matrix. The recursive data addition and deletion operations are described as following.

1) Adding Multiple New Data to the Kernel Matrix: Let the kernel matrix at the ith time step be K_i, we can save the matrix inverse result from the previous step as $L_i = (K_i + \sigma_n^2 I)^{-1}$. Denote the number of new data to be added as M.

With the new data y_{i+1} and kernal vector k_{i+1}, we have

$$\begin{aligned} L_{i+1} &= \begin{bmatrix} L_i^{-1} & k_{i+1} \\ k_{i+1}^T & c_{i+1} + \sigma_n^2 I \end{bmatrix}^{-1} \\ &= \begin{bmatrix} L_i + L_i k_{i+1} (c_{i+1} + \sigma_n^2 I - k_{i+1}^T L_i k_{i+1})^{-1} k_{i+1}^T L_i & L_i k_{i+1} (c_{i+1} + \sigma_n^2 I - k_{i+1}^T L_i k_{i+1})^{-1} \\ -(c_{i+1} + \sigma_n^2 I - k_{i+1}^T L_i k_{i+1})^{-1} k_{i+1}^T L_i & (c_{i+1} + \sigma_n^2 I - k_{i+1}^T L_i k_{i+1}) \end{bmatrix}. \end{aligned}$$

Notice that inversion operation is only required be performed on a $M \times M$ matrix rather than a large $N \times N$ matrix.

2) Deleting Multiple Old Data from the Kernel Matrix: After deleting M data points from the old Kernel matrix inversion $L_i = (K_i + \sigma_n^2 I)^{-1}$, the new inverse of the kernel matrix becomes $\bar{L}_i = (\bar{K}_i + \sigma_n^2 I)^{-1}$. The data to be deleted is permuted to the bottom of the kernel matrix with a permutation matrix P_π, where $\pi : \mathbb{N} \to \mathbb{N}$ is a permutation of N elements. The permuted kernel matrix and its inversion are $K_i^P = P_\pi K_i P_\pi^T$ and $L_i^P = P_\pi L_i P_\pi^T$, respectively. The M least relevant data points at the bottom of the kernel matrix are then removed by comparing block matrices as shown in [20].

With the recursive data addition and deletion method, the GP inference can be obtained efficiently online.

V. SIMULATION RESULTS

The GP based learning algorithm is validated on a simulated quadrotor in two examples, i.e., online learning quadrotor dynamics and learning safety barrier certificates. The actual weight of the quadrotor is 1.4 times the weight used in the computation. In addition, an unknown constant wind of $0.1g$ is applied in the environment as shown in Fig. 4.

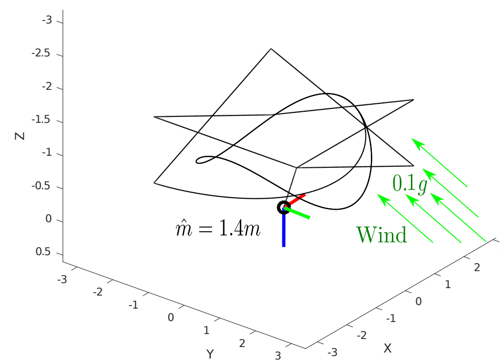
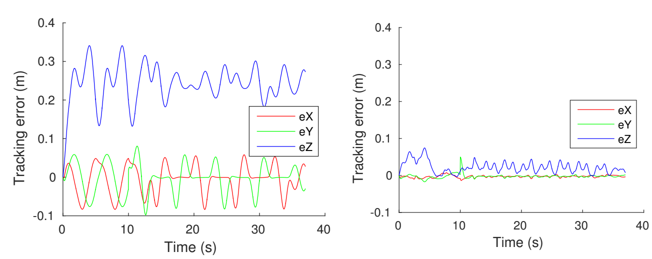


Fig. 4: A simulated quadrotor flies in an unknown wind field with an inaccurate model.

A. Online Learning of Quadrotor Dynamics

The quadrotor is commanded to track a nominal trajectory using a differential flatness based controller with the inaccurate model. The sample size of the recursive GP model is fixed at 300 data points by removing the most irrelevant data points online. The data relevance is decided by the kernel function $k(q, q^*)$. It can observed that the tracking error of the learning based controller is significantly smaller than the tracking error without GP inference, as shown in Fig. 5.



(a) Tracking error without GP (b) Tracking error with GP

Fig. 5: Tracking error of the differential flatness based flight controller with and without GP inference.

With the recursive learning strategy, the GP inference time is always kept below $20ms$, which is very suitable for online learning of quadrotor dynamics.

B. Learning Safety Barrier Certificates

In this example, the motion of the quadrotor is constrained within an ellipsoid safe region, i.e.,

$$\frac{x^2}{0.16}+\frac{y^2}{0.16}+\frac{(z+0.8)^2}{0.36}\leq 1.$$

The quadrotor is controlled to fly back and forth on a vertical path inside the ellipsoid. The goal is to learn how aggressively the quadrotor can fly in the z direction with an inaccurate model and limited thrust.

The barrier certificates are parameterized as

$$\begin{aligned} h_\mu(r) &= 1-\frac{(z+0.8)^2}{0.36}-\mu\dot{z}^2 \\ &\quad -\frac{x^2}{0.16}-\frac{y^2}{0.16}-\frac{\dot{x}^2}{0.25}-\frac{\dot{y}^2}{0.25}\geq 0, \end{aligned}$$

where μ is the barrier parameter to regulate how fast the quadrotor can fly in the z direction. The learning objective is to minimize μ and get more aggressive flight behavior.

The adaptive sampling strategy in Section III-A was adopted to reduce the required sample points. An illustrative example of this strategy is in Fig. 6. The places closer to the boundary of the safe region ($z=-1.2$ and $z=-0.2$) are sampled much denser than the central region ($z=0$).

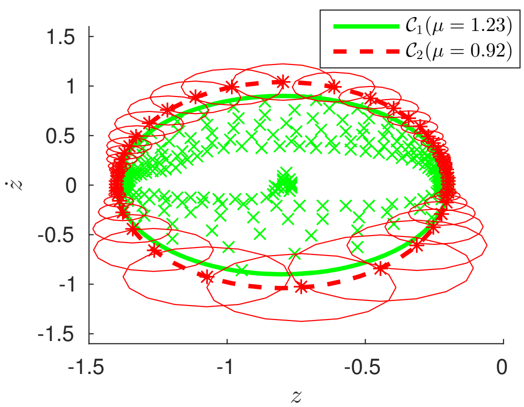


Fig. 6: Adaptive sampling of the state space. The region enclosed by the solid green ellipse $\mathscr{C}_1$ is the current safe region, while the region enclosed by the dashed red ellipse $\mathscr{C}_2$ is the optimized next safe region. The green cross markers and red asterisk markers are the data points already sampled and to be sampled, respectively. The red circles centered at those sample points are the confident safe regions. All the unexplored region between $\mathscr{C}_1$ and $\mathscr{C}_2$ are covered by the circular confident safe region.

A conservative barrier certificate ($\mu=6.3$) is provided at the beginning of the learning process. The quadrotor gradually explores the safe region $\mathscr{C}_0$ and eventually expands it to $\mathscr{C}_n$ ($\mu=0.6$).

VI. Conclusions

A safe learning algorithm based on barrier certificates was developed in this paper. The learning controller is regulated by the barrier certificates, such that the system never enters the unsafe region. The unmodel dynamics of the system was approximated with a Gaussian Process, from which a high probability safety guarantee for the dynamical system was derived. This safe learning technique was validated on a quadrotor system with 3D nonlinear dynamics.

References

[1] A. K. Akametalu, J. F. Fisac, J. H. Gillula, S. Kaynama, M. N. Zeilinger, and C. J. Tomlin. Reachability-based safe learning with gaussian processes. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 1424–1431. IEEE, 2014.

[2] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada. Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control*, 62(8):3861–3876, 2017.

[3] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin. Provably safe and robust learning-based model predictive control. *Automatica*, 49(5):1216–1226, 2013.

[4] F. Berkenkamp, R. Moriconi, A. P. Schoellig, and A. Krause. Safe learning of regions of attraction for uncertain, nonlinear systems with gaussian processes. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 4661–4666. IEEE, 2016.

[5] F. Berkenkamp, M. Turchetta, A. P. Schoellig, and A. Krause. Safe model-based reinforcement learning with stability guarantees. In *Proc. of Neural Information Processing Systems (NIPS)*, 2017.

[6] G. Chesi. *Domain of attraction: analysis and control via SOS programming*, volume 415. Springer Science & Business Media, 2011.

[7] M. P. Deisenroth, D. Fox, and C. E. Rasmussen. Gaussian processes for data-efficient learning in robotics and control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(2):408–423, 2015.

[8] H. Fukushima, T.-H. Kim, and T. Sugie. Adaptive model predictive control for a class of constrained linear systems based on the comparison model. *Automatica*, 43(2):301–308, 2007.

[9] M. Hehn and R. DAndrea. Real-time trajectory generation for quadrocopters. *IEEE Transactions on Robotics*, 31(4):877–892, 2015.

[10] Y. Ito, K. Fujimoto, and Y. Tadokoro. Second-order bounds of gaussian kernel-based functions and its application to nonlinear optimal control with stability. *arXiv preprint arXiv:1707.06240*, 2017.

[11] J. Lofberg. Yalmip: A toolbox for modeling and optimization in matlab. In *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*, pages 284–289. IEEE, 2005.

[12] Q. Nguyen, A. Hereid, J. W. Grizzle, A. D. Ames, and K. Sreenath. 3d dynamic walking on stepping stones with control barrier functions. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 827–834. IEEE, 2016.

[13] Y. Pan and E. Theodorou. Probabilistic differential dynamic programming. In *Advances in Neural Information Processing Systems*, pages 1907–1915, 2014.

[14] S. Prajna, A. Jadbabaie, and G. J. Pappas. Stochastic safety verification using barrier certificates. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, volume 1, pages 929–934. IEEE, 2004.

[15] C. E. Rasmussen. Gaussian processes for machine learning. 2006.

[16] H. Ravanbakhsh and S. Sankaranarayanan. Learning lyapunov (potential) functions from counterexamples and demonstrations. *arXiv preprint arXiv:1705.09619*, 2017.

[17] J. Vinogradska, B. Bischoff, D. Nguyen-Tuong, H. Schmidt, A. Romer, and J. Peters. Stability of controllers for gaussian process forward models. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning*, pages 545–554, 2016.

[18] L. Wang, A. D. Ames, and M. Egerstedt. Multi-objective compositions for collision-free connectivity maintenance in teams of mobile robots. In *Decisions and Control Conference (CDC)*, pages 2659–2664, 2016.

[19] L. Wang, A. D. Ames, and M. Egerstedt. Safe certificate-based maneuvers for teams of quadrotors using differential flatness. In *IEEE International Conference on Robotics and Automation*, pages 3293–3298, 2017.

[20] L. Wang, E. A. Theodorou, and M. Egerstedt. Safe learning of quadrotor dynamics using barrier certificates. *arXiv preprint arXiv:1710.05472*, 2017.

[21] X. Xu, P. Tabuada, J. W. Grizzle, and A. D. Ames. Robustness of control barrier functions for safety critical contro. In *Analysis and Design of Hybrid Systems, 2015 IFAC Conference on*. IEEE, Oct 2015.

[22] D. Zhou and M. Schwager. Vector field following for quadrotors using differential flatness. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6567–6572. IEEE, 2014.