Proceedings of Machine Learning Research vol 120:1–12, 2020 2nd Annual Conference on Learning for Dynamics and Control

Probabilistic Safety Constraints for Learned High Relative Degree System Dynamics

Mohammad Javad Khojasteh* MJKHOJAS@CALTECH.EDU
Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, 91125.

Vikas Dhiman* VDHIMAN@UCSD.EDU

Massimo Franceschetti MFRANCESCHETTI@UCSD.EDU

Nikolay Atanasov NATANASOV@UCSD.EDU
Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093

Editors: A. Bayen, A. Jadbabaie, G. J. Pappas, P. Parrilo, B. Recht, C. Tomlin, M. Zeilinger

Abstract

This paper focuses on learning a model of system dynamics online while satisfying safety constraints. Our motivation is to avoid offline system identification or hand-specified dynamics models and allow a system to safely and autonomously estimate and adapt its own model during online operation. Given streaming observations of the system state, we use Bayesian learning to obtain a distribution over the system dynamics. In turn, the distribution is used to optimize the system behavior and ensure safety with high probability, by specifying a chance constraint over a control barrier function.
Keywords: Gaussian Process, high relative-degree system safety, control barrier function

1. Introduction

Unmanned vehicles promise to transform many aspects of our lives, including transportation, agriculture, mining, and construction. Successful use of autonomous robots in these areas critically depends on the ability of robots to safely adapt to changing operational conditions. Existing systems, however, rely on brittle hand-designed dynamics models and safety rules that often fail to account for both the complexity and uncertainty of real-world operation. Recent work (Deisenroth and Rasmussen, 2011; Dean et al., 2019; Sarkar et al., 2019; Coulson et al., 2019; Chen et al., 2018; Khojasteh et al., 2018; Liu et al., 2019; Umlauft and Hirche, 2019; Fan et al., 2020; Chowdhary et al., 2014) has demonstrated that learning-based system identification and control techniques may be successful at complex tasks and control objectives. However, two critical considerations for applying these techniques onboard autonomous systems remain unattended: learning *online*, relying on streaming data, and guaranteeing *safe* operation, despite the uncertainty inherent to learning algorithms.

Motivated by the utility of Lyapunov functions for certifying stability properties, (Ames et al., 2016; Xu et al., 2017; Xu et al., 2015; Prajna et al., 2007; Ames et al., 2019) proposed *Control Barrier Functions* (CBFs) as a tool for characterizing the long-term safety of dynamical systems. A CBF certifies whether a control policy achieves forward invariance of a *safe set* C by evaluating if the system trajectory remains away from the boundary of C. Most of the literature on CBFs considers systems with known dynamics, low relative degree, no disturbances, and time-triggered

* indicates equal contribution.

control, in which the control inputs are recalculated at a fixed and sufficiently small period. This is limiting because, low control frequency in a time-triggered setting may lead to safety constraint violation in-between sampling times. On the other hand, high control frequency leads to inefficient use of computational resources and actuators. (Yang et al., 2019) extend the CBF framework to a self-triggered setup in which the longest time until a control input needs to be recomputed to guarantee safety is provided. CBF techniques handle nonlinear control-affine systems but many existing results apply only to relative-degree-one systems, in which the first time derivative of the CBF depends on the control input. This requirement is violated by many underactuated robot systems and motivated extensions to relative-degree-two systems, such as bipedal and car-like robots. (Hsu et al., 2015; Nguyen and Sreenath, 2016b). (Nguyen and Sreenath, 2016a) generalized these ideas by designing an exponential control barrier function (ECBF) capable of handling control-affine systems with any relative degree.

Providing safety guarantees for learning-based control techniques has lately been the focus of research. (Koller et al., 2018; Berkenkamp et al., 2016; Fisac et al., 2018; Bastani, 2019; Wabersich and Zeilinger, 2018; Biyik et al., 2019). In particular, the CBF framework have been extend to systems with unknown dynamics. For example, techniques for handling additive disturbances have been proposed in (Clark, 2019; Santoyo et al., 2019), while CBF conditions for systems with uncertain dynamics have been proposed in (Fan et al., 2019; Wang et al., 2018; Taylor and Ames, 2019; Cheng et al., 2019; Salehi et al., 2019). Furthermore, (Fan et al., 2019) study time-triggered CBF-based controllers for control-affine systems with relative degree one, where the input gain part of the dynamics is known and invertible. Bayesian learning is used in (Fan et al., 2019) to determine a distribution over the drift term of the dynamics. In particular, (Fan et al., 2019) compared the performances of Gaussian Process regression (Williams and Rasmussen, 2006), Dropout neural networks (Gal and Ghahramani, 2016), and ALPaCA (Harrison et al., 2018) in simulations. (Wang et al., 2018), (Cheng et al., 2019), and (Taylor and Ames, 2019) have studied time-triggered CBF-based control relative-degree-one systems in presence of additive uncertainty in the drift part of the dynamics. In (Wang et al., 2018), GP regression is used to approximate the unknown part of the 3D nonlinear dynamics of a quadrotor. (Cheng et al., 2019) proposed a two-layers control design architecture that integrates CBF-based controllers with model-free reinforcement learning. (Taylor and Ames, 2019) proposed adaptive CBFs to deal with parameter uncertainty. (Salehi et al., 2019) studies nonlinear systems only with drift terms and uses Extreme Learning Machines to approximate the dynamics.

Our work proposes a learning approach for estimating posterior distribution of robot dynamics from online data to design a control policy that guarantees safe operation. We make the following **contributions**. First, we develop a matrix variate Gaussian Process (GP) regression approach with efficient covariance factorization to learn the *drift term* and *input gain* terms of a nonlinear control-affine system. Second, we use the GP posterior to specify a probabilistic safety constraint and determine the longest time until a control input needs to be recomputed to guarantee safety with high probability. Finally, we extend our formulation to dynamical systems with arbitrary relative degree and show that a safety constraint can be specified only in terms of the mean and variance of the Lie derivatives of the CBF. **Notation, proofs, and additional remarks are available in the appendix at arXiv (Khojasteh et al., 2019)**.

2. Background

Consider a control-affine nonlinear system:

$$\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})\mathbf{u} = \begin{bmatrix} f(\mathbf{x}) & g(\mathbf{x}) \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix} =: F(\mathbf{x})\underline{\mathbf{u}} \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ and $\mathbf{u}(t) \in \mathbb{R}^m$ are the system state and control input, respectively, at time t. Assume that the *drift term* $f : \mathbb{R}^n \to \mathbb{R}^n$ and the *input gain* $g : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ are locally Lipschitz. We study the problem of enforcing probabilistic safety properties via CBF when f and g are unknown. We first review key results on CBF-based safety for *known dynamics* (Ames et al., 2019).

2.1. Known Dynamics: Control Barrier Functions for Safety

Let $\mathcal{C} \subset \mathcal{D} \subset \mathbb{R}^n$ be a *safe set* of system states. Assume $\mathcal{C} = \{\mathbf{x} \in \mathcal{D} \mid h(\mathbf{x}) \geq 0\}$ is specified as the superlevel set of $h \in \mathcal{C}^1(\mathcal{D}, \mathbb{R})$, a continuously differentiable function $\mathcal{D} \to \mathbb{R}$, such that $\nabla_{\mathbf{x}} h(\mathbf{x}) \neq 0$ for all $\mathbf{x}$ when $h(\mathbf{x}) = 0$. For any initial condition $\mathbf{x}(0)$, there exists a maximum time interval $I(\mathbf{x}(0)) = [0, \bar{t})$ with $\bar{t} \in \mathbb{R} \cup \{\infty\}$ such that $\mathbf{x}(t)$ is a unique solution to (1) (Khalil, 2002). System (1) is *safe* with respect to set $\mathcal{C}$ if $\mathcal{C}$ is *forward invariant*, i.e., for any $\mathbf{x}(0) \in \mathcal{C}$, $\mathbf{x}(t)$ remains in $\mathcal{C}$ for all t in $I(\mathbf{x}(0))$. System safety may be asserted as follows.

Definition 1 *A function $h \in \mathcal{C}^1(\mathcal{D}, \mathbb{R})$ is a control barrier function (CBF) for the system in (1) if the control barrier condition (CBC), $\sup_{\mathbf{u}} CBC(\mathbf{x}, \mathbf{u}) \geq 0$, is satisfied for all $\mathbf{x} \in \mathcal{D}$; where $CBC(\mathbf{x}, \mathbf{u}) := \mathcal{L}_f h(\mathbf{x}) + \mathcal{L}_g h(\mathbf{x})\mathbf{u} + \alpha(h(\mathbf{x}))$, α is any extended class K_∞ function and $\mathcal{L}_f h(\mathbf{x})$ and $\mathcal{L}_g h(\mathbf{x})$ are the Lie derivatives of h along f and g, respectively.*

Theorem 1 (Sufficient Condition for Safety (Ames et al., 2019)) *Consider a safe set $\mathcal{C}$ with associated function $h \in \mathcal{C}^1(\mathcal{D}, \mathbb{R})$. If $\nabla_{\mathbf{x}} h(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in \partial\mathcal{C}$, then any Lipschitz continuous control policy $\pi(\mathbf{x}) \in \{\mathbf{u} \in \mathcal{U} \mid CBC(\mathbf{x}, \mathbf{u}) \geq 0\}$ renders the system in (1) safe.*

Ames et al. (2019) also provide a necessary condition for safety allowing a concise charaterization:

$$\text{(1) is safe with respect to } \mathcal{C} \;\; \Leftrightarrow \;\; \exists\, \mathbf{u} = \pi(\mathbf{x}) \;\text{ s.t. }\; \text{CBC}(\mathbf{x}, \mathbf{u}) \geq 0 \;\; \forall \mathbf{x} \in \mathcal{D}. \tag{2}$$

2.2. Known Dynamics: Optimization-based Safe Control

The results in Sec. 2.1 allow designing a control policy $\pi(\mathbf{x})$ that guarantees system safety as long as $\text{CBC}(\mathbf{x}, \pi(\mathbf{x}))$ remains positive at all times. In practice, this is achieved by solving a quadratic program (QP) repeatedly at triggering times $t_k = k\tau$ for $k \in \mathbb{N}$ and $\tau > 0$:

$$\min_{\mathbf{u}_k} \; \mathbf{u}_k^\top Q \mathbf{u}_k \qquad \text{s.t. } \; \text{CBC}(\mathbf{x}_k, \mathbf{u}_k) \geq 0, \tag{3}$$

where $Q \succ 0$, $\mathbf{x}_k := \mathbf{x}(t_k)$, $\mathbf{u}_k := \mathbf{u}(t_k)$. While the QP above cannot be solved infinitely fast, Theorem 3 of Ames et al. (2016) shows that if f, g, and $\alpha \circ h$ are locally Lipschitz, then $\mathbf{u}_k(\mathbf{x})$ and $\text{CBC}(\mathbf{x}, \mathbf{u}_k(\mathbf{x}))$ are locally Lipschitz. Thus, for sufficiently small τ, solving (3) at $\{t_k\}_{k \in \mathbb{N}}$ ensures safety during the inter-triggering times as well.

3. Problem Statement

Consider a control-affine nonlinear system (1), where $F : \mathbb{R}^n \to \mathbb{R}^{n\times(m+1)}$ is *unknown*. Our objective is to estimate $F(\mathbf{x})$ from online observations of the system state and control trajectory and ensure that (1) remains safe with respect to a set $\mathcal{C}$.

Problem 1 *Given a prior Gaussian Process distribution* $vec(F(\mathbf{x}))$[1] $\sim \mathcal{GP}\left(vec(\mathbf{M}_0(\mathbf{x})), \mathbf{K}_0(\mathbf{x},\mathbf{x}')\right)$ *on the unknown system dynamics and a training set* $\mathbf{X}_{1:k} := [\mathbf{x}(t_1), \dots, \mathbf{x}(t_k)]$, $\mathbf{U}_{1:k} := [\mathbf{u}(t_1), \dots, \mathbf{u}(t_k)]$, $\dot{\mathbf{X}}_{1:k} = [\dot{\mathbf{x}}(t_1), \dots, \dot{\mathbf{x}}(t_k)]$[2], *compute the posterior Gaussian Process distribution* $\mathcal{GP}\left(vec(\mathbf{M}_k(\mathbf{x})), \mathbf{K}_k(\mathbf{x},\mathbf{x}')\right)$ *of* $vec(F(\mathbf{x}))$ *conditioned on* $(\mathbf{X}_{1:k}, \mathbf{U}_{1:k}, \dot{\mathbf{X}}_{1:k})$.

Problem 2 *Given a safe set* $\mathcal{C}$, *and a safe system state* $\mathbf{x}_k := \mathbf{x}(t_k) \in \mathcal{C}$, *and the distribution* $\mathcal{GP}(vec(\mathbf{M}_k(\mathbf{x})), \mathbf{K}_k(\mathbf{x},\mathbf{x}'))$ *of* $vec(F(\mathbf{x}))$ *at time* t_k, *choose a control input* $\mathbf{u}_k$ *and triggering period* τ_k *such that:*

$$\mathbb{P}(CBC(\mathbf{x}(t), \mathbf{u}_k) \geq 0) \geq p_k \quad \textit{for} \quad \mathbf{u}(t) \equiv \mathbf{u}_k \quad \textit{and} \quad t \in [t_k, t_k + \tau_k) \tag{4}$$

where $\mathbf{x}(t)$ *follows the dynamics in (1), and* $p_k \in (0,1)$ *is a user-specified risk tolerance.*

4. Matrix Variate Gaussian Process Regression of System Dynamics

We propose an efficient Gaussian Process (GP) regression approach to estimate a posterior distribution over the dynamics $F(\mathbf{x})$ of the nonlinear control-affine systems (1). The posterior will be used to determine the distribution of $\text{CBC}(\mathbf{x},\mathbf{u})$ in Sec. 5[3]. Since $F(\mathbf{x})$ is matrix-valued, we define a GP over its columnwise vectorization, $vec(F(\mathbf{x})) \sim \mathcal{GP}(vec(\mathbf{M}_0(\mathbf{x})), \mathbf{K}_0(\mathbf{x},\mathbf{x}'))$. The controller can observe $\mathbf{X}_{1:k}$ and $\mathbf{U}_{1:k}$ without noise, but the measurements $\dot{\mathbf{X}}_{1:k}$ might be noisy. As the controller observes $f(\mathbf{x})$ and $g(\mathbf{x})$ together via $\dot{\mathbf{X}}_{1:k}$, there may be a correlation between their different components. Thus, we develop an efficient factorization of $\mathbf{K}_0(\mathbf{x},\mathbf{x}')$ based on the Matrix Variate Gaussian distribution (Sun et al., 2017; Louizos and Welling, 2016) to learn $f(\mathbf{x})$ and $g(\mathbf{x})$ together. We provide definition and properties of the MVG distribution in Appendix B.1. Two alternative approaches to infer a posterior over $F(\mathbf{x})$ and their drawbacks are also discussed in Appendix B.1.

Note that if $\mathbf{X} \sim \mathcal{MN}(\mathbf{M}, \mathbf{A}, \mathbf{B})$, then $vec(\mathbf{X}) \sim \mathcal{N}(vec(\mathbf{M}), \mathbf{B} \otimes \mathbf{A})$. Based on this observation, we propose the following GP parameterization for the vector-valued functions $vec(F(\mathbf{x}))$:

$$vec(F(\mathbf{x})) \sim \mathcal{GP}(vec(\mathbf{M}_0(\mathbf{x})), \mathbf{B}_0(\mathbf{x},\mathbf{x}') \otimes \mathbf{A}) \tag{5}$$

The above parameterization is efficient as compared to learning the full covariance $\mathbf{K}_0(.,.) \in \mathbb{R}^{n(m+1)\times(m+1)n}$, because we need to learn smaller matrices, $\mathbf{B}_0(\mathbf{x},\mathbf{x}') \in \mathbb{R}^{(m+1)\times(m+1)}$ and $\mathbf{A} \in \mathbb{R}^{n\times n}$. Fortunately, this parameterization also preserves its structure on inference.

1. $vec(F(\mathbf{x})) \in \mathbb{R}^{n(m+1)}$ is a vector obtained by stacking the columns of $F(\mathbf{x})$
2. If not available, the derivatives may be approximated via $\dot{\mathbf{X}}_{1:k-1} := \left[\frac{\mathbf{x}(t_2)-\mathbf{x}(t_1)}{t_2-t_1}^\top, \dots, \frac{\mathbf{x}(t_k)-\mathbf{x}(t_{k-1})}{t_k-t_{k-1}}^\top\right]^\top$ provided that the inter-triggering times $\{\tau_k\}$ are sufficiently small.
3. We only consider epistemic but no aleatoric uncertainty. Namely, while $F(\mathbf{x})$ is sampled from a GP, no additive disturbances are considered for the dynamics (1).

Consider the training set $(\mathbf{X}_{1:k}, \mathbf{U}_{1:k}, \dot{\mathbf{X}}_{1:k})$ and a query test point $\mathbf{x}_*$. The train and test data are jointly Gaussian:

$$\begin{bmatrix} \dot{\mathbf{x}}_1 \\ \vdots \\ \dot{\mathbf{x}}_k \\ vec(F(\mathbf{x}_*)) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{M}_0(\mathbf{x}_1)\underline{\mathbf{u}}_1 \\ \vdots \\ \mathbf{M}_0(\mathbf{x}_k)\underline{\mathbf{u}}_k \\ vec(\mathbf{M}_0(\mathbf{x}_*)) \end{bmatrix}, \begin{bmatrix} \underline{\mathbf{u}}_1^\top \mathbf{B}_0(\mathbf{x}_1,\mathbf{x}_1)\underline{\mathbf{u}}_1 & \cdots & \underline{\mathbf{u}}_1^\top \mathbf{B}_0(\mathbf{x}_1,\mathbf{x}_k)\underline{\mathbf{u}}_k & \underline{\mathbf{u}}_1^\top \mathbf{B}_0(\mathbf{x}_1,\mathbf{x}_*) \\ \vdots & \ddots & \vdots & \vdots \\ \underline{\mathbf{u}}_k^\top \mathbf{B}_0(\mathbf{x}_k,\mathbf{x}_1)\underline{\mathbf{u}}_1 & \cdots & \underline{\mathbf{u}}_k^\top \mathbf{B}_0(\mathbf{x}_k,\mathbf{x}_k)\underline{\mathbf{u}}_k & \underline{\mathbf{u}}_k^\top \mathbf{B}_0(\mathbf{x}_k,\mathbf{x}_*) \\ \mathbf{B}_0(\mathbf{x}_*,\mathbf{x}_1)\underline{\mathbf{u}}_1 & \cdots & \mathbf{B}_0(\mathbf{x}_*,\mathbf{x}_k)\underline{\mathbf{u}}_k & \mathbf{B}_0(\mathbf{x}_*,\mathbf{x}_*) \end{bmatrix} \otimes \mathbf{A} \right).$$

In the above formulation, the resulting posterior is independent of query control input, $\mathbf{u}_*$, which allows us to use this posterior in Sec. 5 to efficiently compute a safe control input. To simplify notation, let $\mathbf{B}_0(\mathbf{X}_{1:k}, \mathbf{X}_{1:k}) \in \mathbb{R}^{k(m+1)\times k(m+1)}$ be a matrix with elements $[\mathbf{B}_0(\mathbf{X}_{1:k}, \mathbf{X}_{1:k})]_{ij} := \mathbf{B}_0(\mathbf{x}_i, \mathbf{x}_j)$ and define $\boldsymbol{\mathcal{M}}_{1:k} := \begin{bmatrix}\mathbf{M}_0(\mathbf{x}_1) & \cdots & \mathbf{M}_0(\mathbf{x}_k)\end{bmatrix} \in \mathbb{R}^{n\times k(m+1)}$ and $\underline{\boldsymbol{\mathcal{U}}}_{1:k} := \mathbf{diag}(\underline{\mathbf{u}}_1, \ldots, \underline{\mathbf{u}}_k) \in \mathbb{R}^{k(m+1)\times k}$. Applying a Schur complement, we can derive the posterior distribution of $vec(F(\mathbf{x}_*))$ conditioned on $(\mathbf{X}_{1:k}, \mathbf{U}_{1:k}, \dot{\mathbf{X}}_{1:k})$ as a Gaussian Process $\mathcal{GP}(vec(\mathbf{M}_k(\mathbf{x}_*)), \mathbf{B}_k(\mathbf{x}_*, \mathbf{x}'_*) \otimes \mathbf{A})$ with parameters:

$$\mathbf{M}_k(\mathbf{x}_*) := \mathbf{M}_0(\mathbf{x}_*) + \left(\dot{\mathbf{X}}_{1:k} - \boldsymbol{\mathcal{M}}_{1:k}\underline{\boldsymbol{\mathcal{U}}}_{1:k}\right)\left(\underline{\boldsymbol{\mathcal{U}}}_{1:k}^\top \mathbf{B}_0(\mathbf{X}_{1:k}, \mathbf{X}_{1:k})\underline{\boldsymbol{\mathcal{U}}}_{1:k}\right)^{-1} \underline{\boldsymbol{\mathcal{U}}}_{1:k}^\top \mathbf{B}_0(\mathbf{X}_{1:k}, \mathbf{x}_*)$$

$$\mathbf{B}_k(\mathbf{x}_*, \mathbf{x}'_*) := \mathbf{B}_0(\mathbf{x}_*, \mathbf{x}'_*) + \mathbf{B}_0(\mathbf{x}_*, \mathbf{X}_{1:k})\underline{\boldsymbol{\mathcal{U}}}_{1:k}\left(\underline{\boldsymbol{\mathcal{U}}}_{1:k}^\top \mathbf{B}_0(\mathbf{X}_{1:k}, \mathbf{X}_{1:k})\underline{\boldsymbol{\mathcal{U}}}_{1:k}\right)^{-1} \underline{\boldsymbol{\mathcal{U}}}_{1:k}^\top \mathbf{B}_0(\mathbf{X}_{1:k}, \mathbf{x}'_*)$$

This inference has a computation complexity of $O((1+m)^3k^2) + O(k^3)$ while the same for independent GP is $O((1+m)k^2) + O(k^3)$. Since $k >> m$ is common, the proposed model has almost same inference cost as independent GP. Step by step details are provided in Appendix C.1.2. For a given query control input $\mathbf{u}_*$, the posterior of $F(\mathbf{x}_*)\underline{\mathbf{u}}_*$ is:

$$F(\mathbf{x}_*)\underline{\mathbf{u}}_* = f(\mathbf{x}_*) + g(\mathbf{x}_*)\mathbf{u}_* \sim \mathcal{GP}(\mathbf{M}_k(\mathbf{x}_*)\underline{\mathbf{u}}_*, \underline{\mathbf{u}}_*^\top \mathbf{B}_k(\mathbf{x}_*, \mathbf{x}'_*)\underline{\mathbf{u}}_* \otimes \mathbf{A}). \tag{6}$$

5. Self-triggered Control with Probabilistic Safety Constraints

Sec. 4 addressed Problem 1 by proposing an efficient Gaussian Process inference algorithm for nonlinear control-affine systems. Now, we consider Problem (2). As discussed in Sec. 2.1 if f and g are locally Lipschitz, then system (1) has a unique solution for any $\mathbf{x}(0)$ for all time t in $I(\mathbf{x}(0))$. We assume the sample paths of the GP used to model the dynamics (1) are locally Lipschitz with high probability. Similar smoothness assumption has been made previously in Srinivas et al. (2010). As mentioned in Problem (2), we use a zero-order hold (ZOH) control mechanism in inter-triggering time, i.e., $\mathbf{u}(t) \equiv \mathbf{u}_k$ for $t \in [t_k, t_k + \tau_k)$. In detail, we assume that for any $L_k > 0$, $\mathbf{u}_k$, and triggering time t_k, there exists a constant $b_k > 0$, such that,

$$\mathbb{P}\left(\sup_{s\in[0,\tau_k)} \|F(\mathbf{x}(t_k+s))\underline{\mathbf{u}}_k - F(\mathbf{x}_k)\underline{\mathbf{u}}_k\| \le L_k\|\mathbf{x}(t_k+s) - \mathbf{x}_k\|\right) \ge q_k := 1 - e^{-b_k L_k}. \tag{7}$$

This assumption is valid for a large class of GPs, e.g., those with stationary kernels that are four times differentiable, such as squared exponential and some Matérn kernels (Ghosal et al., 2006; Shekhar et al., 2018). However, it may not hold for GPs with highly erratic sample paths.

The posterior of $F(\mathbf{x})\mathbf{u}$ in (6) induces a distribution over $\mathrm{CBC}(\mathbf{x}, \mathbf{u})$. To ensure that safety in the sense of (4) is preserved over a period of time $[t_k, t_k + \tau_k)$, we enforce a tighter constraint at time

t_k and determine the time τ_k for which it remains valid. In detail, we solve a chance-constrained version of (3) at time t_k,

$$\min_{\mathbf{u}_k} \mathbf{u}_k^\top Q \mathbf{u}_k \qquad \text{s.t. } \mathbb{P}(\text{CBC}(\mathbf{x}_k, \mathbf{u}_k) \geq \zeta | \mathbf{x}_k, \mathbf{u}_k) \geq \tilde{p}_k, \tag{8}$$

where $\tilde{p}_k = p_k / q_k$. The choice of ζ and its effect on τ_k is discussed next.

Lemma 2 *Consider the dynamics in* (1) *with posterior distribution in* (6)*. Given* $\mathbf{x}_k$ *and* $\mathbf{u}_k$*, CBC*$_k$:= *CBC*$(\mathbf{x}_k, \mathbf{u}_k)$ *is a Gaussian random variable with the following parameters:*

$$\mathbb{E}[\mathit{CBC}_k] = \nabla_{\mathbf{x}} h(\mathbf{x}_k)^\top \mathbf{M}_k(\mathbf{x}_k) \underline{\mathbf{u}}_k + \alpha(h(\mathbf{x}_k)), \tag{9}$$

$$\mathit{Var}[\mathit{CBC}_k] = \underline{\mathbf{u}}_k^\top \mathbf{B}_k(\mathbf{x}_k, \mathbf{x}_k) \underline{\mathbf{u}}_k \nabla_{\mathbf{x}} h(\mathbf{x}_k)^\top \mathbf{A} \nabla_{\mathbf{x}} h(\mathbf{x}_k) \tag{10}$$

Using Lemma 2, we can rewrite the safety constraint as

$$\mathbb{P}(\text{CBC}_k \geq \zeta | \mathbf{x}_k, \mathbf{u}_k) = 1 - \Phi\left(\frac{\zeta - \mathbb{E}[\text{CBC}_k]}{\sqrt{\mathit{Var}[\text{CBC}_k]}}\right) \geq \tilde{p}_k, \tag{11}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard Gaussian. Note that if the control input is chosen so that $\zeta - \mathbb{E}[\text{CBC}_k] < 0$, as the posterior variance of CBC_k tends to zero, the probability $\mathbb{P}(\text{CBC}_k \geq \zeta | \mathbf{x}_k, \mathbf{u}_k)$ tends to one. Namely, as the uncertainty about the system dynamics tends to zero, our results reduce to the setting of Sec. 2.1, and safety can be ensured with probability one. Noting that $\Phi^{-1}(1 - \tilde{p}_k) = \sqrt{2}\text{erf}^{-1}(1 - 2\tilde{p}_k)$, controller (8) can be rewritten as

$$\min_{\mathbf{u}_k} \mathbf{u}_k^\top Q \mathbf{u}_k \quad \text{s.t. } \mathbb{E}[\text{CBC}_k] - \zeta \geq 0 \text{ and } (\mathbb{E}[\text{CBC}_k] - \zeta)^2 \geq 2\, \mathit{Var}[\text{CBC}_k]\, (\text{erf}^{-1}(1 - 2\tilde{p}_k))^2. \tag{12}$$

The program (12) provides a probabilistic safety constraints at the triggering times $\{t_k\}_{k \in \mathbb{N}}$. Next, we will extend our analysis to inter-triggering times $\{\tau_k\}$. We continue by re-writing the Proposition 1 of (Yang et al., 2019) for our setup.

Proposition 1 *Consider the system in* (1) *with zero-order hold control in inter-triggering times. If the event* (7) *occurs at the kth triggering time, then for all* $s \in [0, \tau_k)$ *we have*

$$\|\mathbf{x}(t_k + s) - \mathbf{x}_k\| \leq \overline{r}_k(s) := \frac{1}{L_k} \|\dot{\mathbf{x}}_k\| \left(e^{L_k s} - 1\right). \tag{13}$$

Recall from Sec. 2.1 that h is a continuously differentiable function. Thus using Proposition 1, we notice for any inter-triggering time τ_k, there exist a constant $\chi_k > 0$ such that

$$\sup_{s \in [0, \tau_k)} \|\nabla h(\mathbf{x}(t_k + s))\| \leq \chi_k. \tag{14}$$

This is used in the next theorem which concerns Problem 2.

Theorem 3 *Consider the system in* (1) *with safe set* $\mathcal{C}$*. Assume the program* (8) *has a solution at triggering time* t_k*, event* (7) *occurs at least with probability* q_k*,* $\|\dot{\mathbf{x}}_k\| \neq 0$*, and for all* $s \in [0, \tau_k)$*,* $\alpha \circ h$ *satisfies the following Lipschitz property*

$$|\alpha \circ h(\mathbf{x}(t_k + s)) - \alpha \circ h(\mathbf{x}_k)| \leq L_{\alpha \circ h} \|\mathbf{x}(t_k + s) - \mathbf{x}_k\|. \tag{15}$$

Then (4) *is valid for* $p_k = \tilde{p}_k q_k$*, and* $\tau_k \leq \frac{1}{L_k} \ln\left(1 + \frac{L_k \zeta}{(\chi_k L_k + L_{\alpha \circ h})\|\dot{\mathbf{x}}_k\|}\right)$*, where* χ_k *is given in* (14)*.*

Remark 4 Assuming $\|\dot{x}(t_k)\| \neq 0$ in Theorem (3) is not restricting our results. Since, if the state of the system is safe and it does not change it remains safe.

6. Extension to Higher Relative-degree Systems

Next, we extend the probabilistic safety constraint formulation for systems with arbitrary relative degree, using an exponential control barrier function (ECBF) (Nguyen and Sreenath, 2016a; Ames et al., 2019) [4]

Let $r \geq 1$ be the relative degree of $h(\mathbf{x})$, that is, $\mathcal{L}_g\mathcal{L}_f^{(r-1)}h(\mathbf{x}) \neq 0$ and $\mathcal{L}_g\mathcal{L}_f^{(k-1)}h(\mathbf{x}) = 0$, $\forall k \in \{1, \dots, r-2\}$. Define traverse dynamics with traverse vector $\eta(\mathbf{x})$,

$$\dot{\eta}(\mathbf{x}) = \mathcal{F}\eta(\mathbf{x}) + \mathcal{G}\mathbf{u}, \qquad h(\mathbf{x}) = C\eta(\mathbf{x}) \tag{16}$$

where $C = [1, 0, \dots, 0]^\top \in \mathbb{R}^r$. Also, $\eta(\mathbf{x})$, $\mathcal{F}$, and $\mathcal{G}$ are defined in Appendix A.

Definition 2 *A function $h \in \mathcal{C}^r(\mathcal{D}, \mathbb{R})$ is an exponential control barrier function (ECBF) for the system in (1) if there exists a row vector $K_\alpha \in \mathbb{R}^r$ such that the rth order condition $CBC^{(r)}(\mathbf{x}, \mathbf{u}) := \mathcal{L}_f^{(r)}h(\mathbf{x}) + \mathcal{L}_g\mathcal{L}_f^{(r-1)}h(\mathbf{x})\mathbf{u} + K_\alpha\eta(\mathbf{x})$ satisfies $\sup_{\mathbf{u}} CBC^{(r)}(\mathbf{x}, \mathbf{u}) \geq 0$ for all $\mathbf{x} \in \mathcal{D}$, which results in $h(\mathbf{x}(t)) \geq C\eta(\mathbf{x}_0)e^{(\mathcal{F}-\mathcal{G}K_\alpha)t} \geq 0$, whenever $h(\mathbf{x}_0) \geq 0$.*

If K_α is chosen appropriately (see Appendix B.2), a control policy $\mathbf{u} = \pi(\mathbf{x})$ that ensures $\text{CBC}^{(r)} \geq 0$, renders the dynamics (1) safe with respect to set $\mathcal{C}$. Thus, as in (8), we are interested in solving

$$\min_{\mathbf{u}_k} \ \mathbf{u}_k^\top Q\mathbf{u}_k \qquad \text{s.t. } \mathbb{P}(\text{CBC}_k^{(r)} \geq \zeta | \mathbf{x}_k, \mathbf{u}_k) \geq \tilde{p}_k. \tag{17}$$

Proposition 2 *For a control-affine system of relative degree r, the expectation $\mathbb{E}[CBC_k^{(r)}]$ is affine in $\mathbf{u}$ and $Var[CBC_k^{(r)}]$ is quadratic in $\mathbf{u}$.*

Proposition 3 *For a control-affine system of relative degree r, as defined in (1), the system stays in the safe set $\mathcal{C}$ with ECBF h if the control is determined from the following Quadratically Constrained Quadratic Program (QCQP),*

$$\min_{\mathbf{u}_k^\top} \ \mathbf{u}_k^\top Q\mathbf{u}_k \quad \textit{s.t.} \quad \mathbb{E}[CBC_k^{(r)}] - \zeta \geq 0 \quad \textit{and} \quad (\mathbb{E}[CBC_k^{(r)}] - \zeta)^2 \geq \frac{\tilde{p}_k}{1-\tilde{p}_k} Var[CBC_k^{(r)}] \tag{18}$$

Solving the program (18) requires the knowledge of the mean and variance of $\text{CBC}_k^{(r)}$ (see Thm. 8 in Appendix C.3.1 for $\text{CBC}^{(2)}$). In general, Monte Carlo sampling could be used to estimate these quantities. The chance constraint in (18) can be interpreted the standard deviation of $\text{CBC}_k^{(r)}$ should be smaller than the mean by a factor of $\sqrt{\tilde{p}_k/(1-\tilde{p}_k)}$.

7. Simulations

We evaluate the proposed approach on a pendulum with mass m and length l with state $\mathbf{x} = [\theta, \omega]$ and control-affine dynamics $f(\mathbf{x}) = [\omega, -\frac{g}{l}\sin(\theta)]$ and $g(\mathbf{x}) = [0, \frac{1}{ml}]$ as depicted in Fig 1. A safe set is chosen as the complement of a radial region $[\theta_c - \Delta_{col}, \theta_c + \Delta_{col}]$ that needs to be avoided.

4. The motivation for assuming known relative degree and CBF but unknown dynamics comes from robotics applications. Commonly, the class of the system is known but the parameters (e.g., mass, the moment of inertia) and high-order interactions (e.g., jerk, snap) of the dynamics are unknown. Finding the relative degree and a proper CBF is left open for future work (cf. (Akella et al., 2020; Robey et al., 2020)).

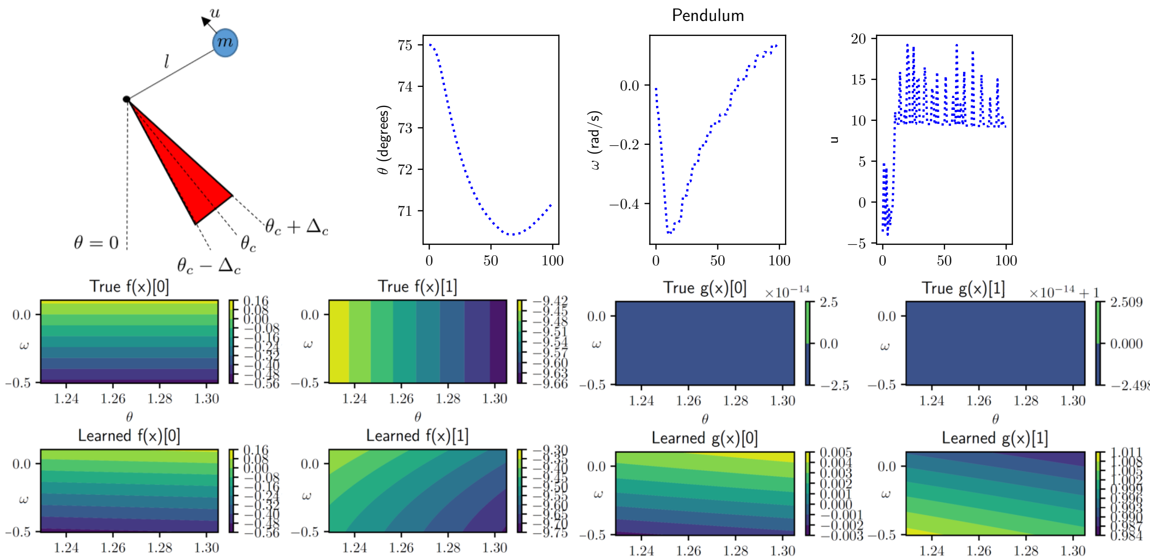

Figure 1: **Top left**: Pendulum simulation (left) with an unsafe (red) region. **Top right**: The pendulum trajectory (middle) resulting from the application of safe control inputs (right) is shown. **Bottom row**: Learned vs true pendulum dynamics using matrix variate Gaussian Process regression

The controller knows a priori that the system is control-affine with relative degree two, but it is not aware of f and g. The control barrier function is thus $h(\mathbf{x}) = \cos(\Delta_{col}) - \cos(\theta - \theta_c)$. We formulate a quadratically constrained quadratic program as in (18) for $r = 2$. We specify a task requiring the pendulum to track a reference control signal $\mathbf{u}_0$ and specify the optimization objective as $(\mathbf{u}_k - \mathbf{u}_0)^\top Q(\mathbf{u}_k - \mathbf{u}_0)$. We initialize the system with parameters $\theta_0 = 75°$, $\omega_0 = -0.01$, $\tau = 0.01$, $m = 1$, $g = 10$, $l = 1$, $\theta_c = 45$, $\Delta_{col} = 22.5$. The system dynamics are approximated accurately (see Fig. 1) while the system remains in the safe region (see Fig. 1). An ϵ-greedy exploration strategy is used to sample $\mathbf{u}_0 \in [-20, 20]$. We use an exponentially decreasing ϵ-greedy scheme going from 1 to 0.01 in 100 steps. Negative control inputs get rejected by the CBF-based constraint, while positive inputs allow the pendulum to bounce back from the unsafe region.

8. Conclusion

Allowing artificial systems to safely adapt their own models during online operation will have significant implications for their successful use in unstructured, changing real-world environments. This paper developed a Bayesian inference approach to approximate system dynamics and their uncertainty from online observations. The posterior distribution over the dynamics may be used to enforce probabilistic constraints that guarantee safe online operation with high probability. Our results offer a promising approach for controlling complex systems in challenging environments. Future work will focus on extending the self-triggering time analysis to systems with higher relative degree and on applications of the proposed approach to real robot systems.

Acknowledgments

We gratefully acknowledge support from NSF awards CNS-1446891 and ECCS-1917177, and support from ARL DCIST CRA W911NF-17-2-0181.

References

Prithvi Akella, Mohamadreza Ahmadi, Richard M Murray, and Aaron D Ames. Formal test synthesis for safety-critical autonomous systems based on control barrier functions. *arXiv preprint arXiv:2004.04227*, 2020.

Mauricio A Alvarez, Lorenzo Rosasco, and Neil D Lawrence. Kernels for vector-valued functions: A review. *Foundations and Trends in Machine Learning*, 4(3):195–266, 2012.

A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada. Control barrier functions: Theory and applications. In *2019 18th European Control Conference (ECC)*, pages 3420–3431, June 2019. doi: 10.23919/ECC.2019.8796030.

Aaron D Ames, Xiangru Xu, Jessy W Grizzle, and Paulo Tabuada. Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control*, 62(8): 3861–3876, 2016.

Osbert Bastani. Safe planning via model predictive shielding. *arXiv preprint arXiv:1905.10691*, 2019.

Felix Berkenkamp, Angela P. Schoellig, and Andreas Krause. Safe controller optimization for quadrotors with Gaussian processes. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 493–496, 2016.

Erdem Biyik, Jonathan Margoliash, Shahrouz Ryan Alimo, and Dorsa Sadigh. Efficient and safe exploration in deterministic markov decision processes with unknown transition models. In *2019 American Control Conference (ACC)*, pages 1792–1799. IEEE, 2019.

Steven Chen, Kelsey Saulnier, Nikolay Atanasov, Daniel D Lee, Vijay Kumar, George J Pappas, and Manfred Morari. Approximating explicit model predictive control using constrained neural networks. In *2018 Annual American Control Conference (ACC)*, pages 1520–1527. IEEE, 2018.

Richard Cheng, Gábor Orosz, Richard M Murray, and Joel W Burdick. End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 3387–3395, 2019.

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.

Andrew Clark. Control barrier functions for complete and incomplete information stochastic systems. In *2019 American Control Conference (ACC)*, pages 2928–2935. IEEE, 2019.

Jeremy Coulson, John Lygeros, and Florian Dörfler. Data-enabled predictive control: in the shallows of the deepc. In *2019 18th European Control Conference (ECC)*, pages 307–312. IEEE, 2019.

Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. On the sample complexity of the linear quadratic regulator. *Foundations of Computational Mathematics*, Aug 2019.

Marc Deisenroth and Carl E Rasmussen. PILCO: A model-based and data-efficient approach to policy search. In *Proceedings of the 28th International Conference on machine learning (ICML-11)*, pages 465–472, 2011.

David D Fan, Jennifer Nguyen, Rohan Thakker, Nikhilesh Alatur, Ali-akbar Agha-mohammadi, and Evangelos A Theodorou. Bayesian learning-based adaptive control for safety critical systems. *arXiv preprint arXiv:1910.02325*, 2019.

David D Fan, Ali-akbar Agha-mohammadi, and Evangelos A Theodorou. Deep learning tubes for tube MPC. *arXiv preprint arXiv:2002.01587*, 2020.

Jaime F Fisac, Anayo K Akametalu, Melanie N Zeilinger, Shahab Kaynama, Jeremy Gillula, and Claire J Tomlin. A general safety framework for learning-based control in uncertain robotic systems. *IEEE Transactions on Automatic Control*, 2018.

Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059, 2016.

Subhashis Ghosal, Anindya Roy, et al. Posterior consistency of Gaussian process prior for nonparametric binary regression. *The Annals of Statistics*, 34(5):2413–2429, 2006.

James Harrison, Apoorva Sharma, and Marco Pavone. Meta-learning priors for efficient online Bayesian regression. *arXiv preprint arXiv:1807.08912*, 2018.

Shao-Chen Hsu, Xiangru Xu, and Aaron D Ames. Control barrier function based quadratic programs with application to bipedal robotic walking. In *2015 American Control Conference (ACC)*, pages 4542–4548. IEEE, 2015.

Hassan K Khalil. *Nonlinear systems; 3rd ed.* Prentice-Hall, Upper Saddle River, NJ, 2002.

Mohammad Javad Khojasteh, Anatoly Khina, Massimo Franceschetti, and Tara Javidi. Learning-based attacks in cyber-physical systems. *arXiv preprint arXiv:1809.06023*, 2018.

Mohammad Javad Khojasteh, Vikas Dhiman, Massimo Franceschetti, and Nikolay Atanasov. Probabilistic safety constraints for learned high relative degree system dynamics. *arXiv preprint arXiv: 1912.10116*, 2019.

Torsten Koller, Felix Berkenkamp, Matteo Turchetta, and Andreas Krause. Learning-based model predictive control for safe exploration. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 6059–6066. IEEE, 2018.

Anqi Liu, Guanya Shi, Soon-Jo Chung, Anima Anandkumar, and Yisong Yue. Robust regression for safe exploration in control. *arXiv preprint arXiv:1906.05819*, 2019.

Christos Louizos and Max Welling. Structured and efficient variational deep learning with matrix Gaussian posteriors. In *International Conference on Machine Learning*, pages 1708–1716, 2016.

Quan Nguyen and Koushil Sreenath. Exponential control barrier functions for enforcing high relative-degree safety-critical constraints. In *2016 American Control Conference (ACC)*, pages 322–328. IEEE, 2016a.

Quan Nguyen and Koushil Sreenath. Optimal robust control for constrained nonlinear hybrid systems with application to bipedal locomotion. In *2016 American Control Conference (ACC)*, pages 4807–4813. IEEE, 2016b.

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, 2007.

Alexander Robey, Haimin Hu, Lars Lindemann, Hanwen Zhang, Dimos V Dimarogonas, Stephen Tu, and Nikolai Matni. Learning control barrier functions from expert demonstrations. *arXiv preprint arXiv:2004.03315*, 2020.

Iman Salehi, Gang Yao, and Ashwin P Dani. Active sampling based safe identification of dynamical systems using extreme learning machines and barrier certificates. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 22–28. IEEE, 2019.

Cesar Santoyo, Maxence Dutreix, and Samuel Coogan. A barrier function approach to finite-time stochastic system verification and control. *arXiv preprint arXiv:1909.05109*, 2019.

Tuhin Sarkar, Alexander Rakhlin, and Munther A Dahleh. Finite-time system identification for partially observed LTI systems of unknown order. *arXiv preprint arXiv:1902.01848*, 2019.

Shayle R Searle and Marvin HJ Gruber. *Linear models*. John Wiley & Sons, 1971.

Shubhanshu Shekhar, Tara Javidi, et al. Gaussian process bandits with adaptive discretization. *Electronic Journal of Statistics*, 12(2):3829–3874, 2018.

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In *In International Conference on Machine Learning*, 2010.

Shengyang Sun, Changyou Chen, and Lawrence Carin. Learning Structured Weight Uncertainty in Bayesian Neural Networks. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 1283–1292, 2017.

Andrew J Taylor and Aaron D Ames. Adaptive safety with control barrier functions. *arXiv preprint arXiv:1910.00555*, 2019.

Jonas Umlauft and Sandra Hirche. Feedback linearization based on Gaussian processes with event-triggered online learning. *IEEE Transactions on Automatic Control*, 2019.

Kim P Wabersich and Melanie N Zeilinger. Safe exploration of nonlinear dynamical systems: A predictive safety filter for reinforcement learning. *arXiv preprint arXiv:1812.05506*, 2018.

Li Wang, Evangelos A Theodorou, and Magnus Egerstedt. Safe learning of quadrotor dynamics using barrier certificates. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2460–2465. IEEE, 2018.

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

X. Xu, T. Waters, D. Pickem, P. Glotfelter, M. Egerstedt, P. Tabuada, J. W. Grizzle, and A. D. Ames. Realizing simultaneous lane keeping and adaptive speed regulation on accessible mobile robot testbeds. In *IEEE Conference on Control Technology and Applications (CCTA)*, pages 1769–1775, 2017.

Xiangru Xu, Paulo Tabuada, Jessy W Grizzle, and Aaron D Ames. Robustness of control barrier functions for safety critical control. *IFAC-PapersOnLine*, 48(27):54–61, 2015.

Guang Yang, Calin Belta, and Roberto Tron. Self-triggered control for safety critical systems using control barrier functions. In *2019 American Control Conference (ACC)*, pages 4454–4459. IEEE, 2019.