

# Optimized Control Invariance Conditions for Uncertain Input-Constrained Nonlinear Control Systems

Lukas Brunke, Siqi Zhou, Mingxuan Che, and Angela P. Schoellig

**Abstract**—Providing safety guarantees for learning-based controllers is important for real-world applications. One approach to realizing safety for arbitrary control policies is safety filtering. If necessary, the filter modifies control inputs to ensure that the trajectories of a closed-loop system stay within a given state constraint set for all future time, referred to as the set being positive invariant or the system being safe. Under the assumption of fully known dynamics, safety can be certified using control barrier functions (CBFs). However, the dynamics model is often either unknown or only partially known in practice. Learning-based methods have been proposed to approximate the CBF condition for unknown or uncertain systems from data; however, these techniques do not account for input constraints and, as a result, may not yield a valid CBF condition to render the safe set invariant. In this work, we study conditions that guarantee control invariance of the system under input constraints and propose an optimization problem to reduce the conservativeness of CBF-based safety filters. Building on these theoretical insights, we further develop a probabilistic learning approach that allows us to build a safety filter that guarantees safety for uncertain, input-constrained systems with high probability. We demonstrate the efficacy of our proposed approach in simulation and real-world experiments on a quadrotor and show that we can achieve safe closed-loop behavior for a learned system while satisfying state and input constraints.

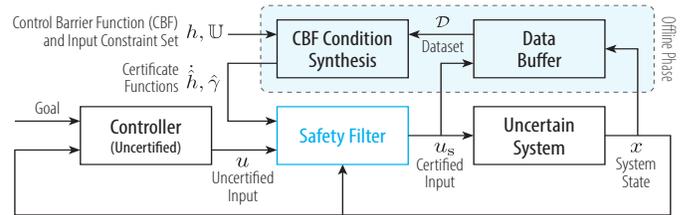
**Index Terms**—Constrained control, machine learning, robotics, uncertain systems

## I. INTRODUCTION

IN recent years, learning-based control techniques have enabled robots to perform complex tasks in uncertain environments. However, providing rigorous safety guarantees of learning-based controllers is typically challenging in such settings, which hinders their real-world deployment [1]. One approach to providing safety guarantees for learning-based controllers is to use a safety filter, which modifies the input to the system when it is deemed unsafe.

Control barrier function (CBF)-based safety filters [2] are an example of safety filters for continuous-time systems. A CBF-based safety filter requires a CBF which defines the safe set. The synthesis of a CBF for nonlinear systems typically necessitates considerable offline computation [3]. However, once the CBF has been computed and under the assumption of control affine dynamics, safety can be efficiently determined

The authors are with the [Learning Systems and Robotics Lab](#) at the Technical University of Munich, Germany and the Munich Institute of Robotics and Machine Intelligence (MIRMI). LB and APS are also affiliated with the University of Toronto Institute for Aerospace Studies, the University of Toronto Robotics Institute, and the Vector Institute for Artificial Intelligence. Emails: {lukas.brunke, siqi.zhou, mingxuan.che, angela.schoellig}@tum.de



**Fig. 1.** A block diagram of our proposed control barrier function safety filter approach for uncertain, input-constrained systems. Using the theoretical conditions we derived, we can verify that our proposed safety filter satisfies safety under input constraints. We can augment any unsafe controller with our proposed safety filter to guarantee safety with high probability.

online by solving a quadratic program (QP) [4], for which adequate solvers exist. Motivated by the goal of achieving safe real-world operation of resource-limited hardware systems, for example, miniature quadrotors, we focus on CBF-based safety filters in this paper due to their reduced online computation.

CBFs provide a scalar condition to determine whether a proposed control input keeps the system in the safe set [4]. The CBF condition is derived from Nagumo’s theorem [5], which intuitively defines safety by guaranteeing that for any state on the safe set’s boundary, the system will not leave the safe set. CBFs can be synthesized for systems with fully known dynamics [3], [6]–[8] or for unknown dynamics [9].

Real-world systems typically have actuation limits. For safety, these input constraints must be accounted for. Otherwise, the assumption of additional control authority can lead to safety violations. For systems with fully known dynamics, the authors in [10], [11] propose adapting the CBF condition’s lower bound online to achieve feasibility under control input constraints. If control invariance cannot be satisfied due to input constraints, the authors in [12] propose a method for determining control invariant subsets of the original safe set. In [8], the authors study robust control invariance under input constraints for uncertain systems.

In recent literature, learning-based CBF safety filters have been proposed to guarantee safety for uncertain systems [1]. Given a CBF, these approaches either directly learn the Lie derivative from data [13], [14] or indirectly learn it by first identifying a dynamics model of the uncertain system [15], [16]. In [14], [15], [17], probabilistic learning techniques for CBFs have been proposed to guarantee the system’s safety with high probability. While effective, these works do not account for control input constraints, which may not lead to valid CBF conditions to render the system safe on the original state constraint set. Recently, the authors in [18] have considered a data-based approach to synthesize a CBF-based safety controller for an unknown system with a discrete input set. In our work, we learn the system’s Lie derivative

that satisfies the CBF condition under general compact input constraints with high probability for an uncertain system, where the dynamics are (partially) unknown.

Our contributions are as follows: (i) derivation of control invariance conditions for generic control-affine nonlinear systems under input constraints, (ii) optimization-based design of CBF conditions to increase the set of feasible inputs while guaranteeing control invariance under input constraints, (iii) a probabilistic learning approach to obtain a CBF safety filter that provides probabilistic safety guarantees for uncertain, input-constrained systems, and (iv) simulation and real-world results demonstrating the efficacy of our approach.

## II. PROBLEM FORMULATION

In this work, we consider the control architecture in Fig. 1 and a continuous-time nonlinear control system represented in the following control-affine form:

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

where  $x \in \mathbb{X} \subset \mathbb{R}^n$  is the state of the system with  $\mathbb{X}$  being the set of admissible states,  $u \in \mathbb{U} \subset \mathbb{R}^m$  is the input of the system with  $\mathbb{U}$  being the set of admissible inputs, and  $f: \mathbb{R}^n \mapsto \mathbb{R}^n$  and  $g: \mathbb{R}^n \mapsto \mathbb{R}^{n \times m}$  are locally Lipschitz continuous functions. We assume that  $\mathbb{X}$  and  $\mathbb{U}$  are known compact sets, and  $f$  and  $g$  are (partially) unknown functions.

The safe set  $\mathbb{C} \subseteq \mathbb{X}$  is assumed to be given and is defined as the zero-superlevel set of a continuously differentiable function  $h: \mathbb{R}^n \rightarrow \mathbb{R}$ :  $\mathbb{C} = \{x \in \mathbb{X} \mid h(x) \geq 0\}$  where the boundary of the safe set is  $\partial\mathbb{C} = \{x \in \mathbb{X} \mid h(x) = 0\}$  with  $\partial h(x)/\partial x \neq 0$  for all  $x \in \partial\mathbb{C}$ , and the interior is  $\text{Int}(\mathbb{C}) = \{x \in \mathbb{X} \mid h(x) > 0\}$ . The corresponding unsafe set is defined as  $\bar{\mathbb{C}} = \{x \in \mathbb{X} \mid h(x) < 0\}$ . Our goal is to augment a given, potentially unsafe state-feedback controller  $\pi(x)$  with a safety filter such that the uncertain system is safe (i.e., the system's state  $x$  stays inside a safe set  $\mathbb{C}$  if it starts inside of  $\mathbb{C}$ ) under the input constraints  $\mathbb{U}$ .

## III. BACKGROUND

To facilitate our discussion, we introduce relevant definitions and background on CBF safety filters in this section.

*Definition 1 (Extended class- $\mathcal{K}_e$  function [4]):* A function  $\gamma: \mathbb{R} \rightarrow \mathbb{R}$  is said to be of class- $\mathcal{K}_e$  if it is continuous,  $\gamma(0) = 0$ , and strictly increasing.

*Definition 2 (Positively control invariant set):* Let  $\mathcal{U}$  be the set of all bounded controllers  $\nu: \mathbb{R}_{\geq 0} \rightarrow \mathbb{U}$ . A set  $\mathbb{C} \subseteq \mathbb{X}$  is a positively control invariant set for the control system in (1) if  $\forall x_0 \in \mathbb{C}, \exists \nu \in \mathcal{U}, \forall t \in \mathbb{T}_{x_0}^+, \phi(t, x_0, \nu) \in \mathbb{C}$ , where  $\phi(t, x_0, \nu)$  is the system's phase flow starting at  $x_0$  under the controller  $\nu$ , and  $\mathbb{T}_{x_0}^+$  is the maximum time interval.

*Definition 3 (CBF [4]):* Let  $\mathbb{C} \subseteq \mathbb{X}$  be the superlevel set of a continuously differentiable function  $h: \mathbb{X} \rightarrow \mathbb{R}$ , then  $h$  is a CBF if there exists a class- $\mathcal{K}_e$  function  $\gamma$  such that for all  $x \in \mathbb{X}$  the control system in (1) satisfies

$$\max_{u \in \mathbb{U}} [L_f h(x) + L_g h(x)u] \geq -\gamma(h(x)), \quad (2)$$

where  $L_f h(x)$  and  $L_g h(x)$  are the Lie derivatives of  $h$  along  $f$  and  $g$ , respectively. In the following, we write  $\dot{h}(x, u) = L_f h(x) + L_g h(x)u$  for simplicity.

Using a CBF, we can define an input set

$$\mathbb{U}_{\text{cbf}}(x) = \{u \in \mathbb{U} \mid \dot{h}(x, u) \geq -\gamma(h(x))\} \quad (3)$$

that renders the system safe [4]. This requires choosing a class- $\mathcal{K}_e$  function  $\gamma$  that yields  $\mathbb{U}_{\text{cbf}}(x) \neq \emptyset$  for all  $x \in \mathbb{C}$ . For a controller  $\pi(x)$  that is not initially designed to be safe, one can formulate a QP to modify the control input such that the system is guaranteed to be safe [2]

$$u_s(x) = \underset{u \in \mathbb{U}}{\text{argmin}} \quad \frac{1}{2} \|u - \pi(x)\|_2^2 \quad (4a)$$

$$\text{s.t.} \quad \dot{h}(x, u) \geq -\gamma(h(x)). \quad (4b)$$

Intuitively, the optimization problem in (4) finds an input in  $\mathbb{U}_{\text{cbf}}(x)$  that best matches  $\pi(x)$ , where the closeness of the inputs is specified with respect to a chosen distance measure (e.g., the Euclidean norm in (4a)).

## IV. CONTROL INVARIANCE UNDER INPUT CONSTRAINTS

In order for a continuously differentiable function  $h$  to be a valid CBF, we need to find a  $\gamma \in \mathcal{K}_e$  such that (2) holds. In general, finding such a class- $\mathcal{K}_e$  function under input constraints is not trivial. In [11], the authors proposed a method that scales a pre-selected class- $\mathcal{K}_e$  function online to guarantee that the safety filter optimization problem in (4) is feasible to solve. However, as noted by the authors, this does not guarantee control invariance of the safe set but rather only point-wise feasibility inside the safe set.

In this section, we assume fully known functions  $f$  and  $g$  to derive conditions for control invariance of the CBF's superlevel sets under compact input constraints  $\mathbb{U}$ . Based on these conditions, we further propose a method to increase the size of feasible control input sets. We will then extend the discussion to uncertain systems in Sec. V.

### A. Valid Class- $\mathcal{K}_e$ Functions Under Input Constraints

In this subsection, we devise a systematic approach to determine a class- $\mathcal{K}_e$  function for a given  $h$  that satisfies (2) under input constraints  $\mathbb{U}$ .

Before we start, we first note that, in order for the problem to be well-posed, we need to check whether there exists an input  $u$  to make the Lie derivative  $\dot{h}(x, u)$  non-negative for all  $x$  on the boundary of the safe set  $\mathbb{C}$ :

$$\forall x \in \partial\mathbb{C}, \exists u \in \mathbb{U}, \text{ s.t. } \dot{h}(x, u) \geq 0. \quad (5)$$

If the condition in (5) is not satisfied for some  $x \in \partial\mathbb{C}$ , then  $\dot{h}$  will be negative at the particular  $x$ , which implies that the system will leave the safe set. In this case, we cannot render the system safe, and the problem of finding a class- $\mathcal{K}_e$  function satisfying (2) will be infeasible.

In addition to the constraint on the Lie derivative at the boundary, (2) requires the Lie derivative to be lower-bounded by a negative class- $\mathcal{K}_e$  function  $-\gamma(h(x))$  for all  $x \in \mathbb{X}$ . The class- $\mathcal{K}_e$  function needs to be chosen such that there exists a  $u \in \mathbb{U}$  such that the constraint in (2) is feasible for all  $x \in \mathbb{X}$ . In order to facilitate the design of the class- $\mathcal{K}_e$  function, we derive a condition that ensures the feasibility of (2).

Note that, given the compact input set  $\mathbb{U}$ , we can determine a maximum realizable Lie derivative  $\dot{h}$  for each  $x \in \mathbb{X}$ .

We define the lower bound on the maximum realizable Lie derivative for each level set  $\mathbb{C}_c = \{x \in \mathbb{R}^n \mid h(x) = c\}$  as

$$\underline{\gamma}(c) = - \min_{x \in \mathbb{C}_c} \max_{u \in \mathbb{U}} \dot{h}(x, u), \quad (6)$$

where  $c \in [h_{\min}, h_{\max}]$  with  $h_{\min}$  and  $h_{\max}$  corresponding to the largest and the smallest level sets of  $h$  contained in the compact set  $\mathbb{X}$ , respectively. The following theorem then specifies a condition on  $\gamma \in \mathcal{K}_e$  that guarantees the satisfaction of (2) for a given  $h$  over the largest superlevel set of  $h$  contained in  $\mathbb{X}$  (i.e.,  $\mathbb{X}_{\max} = \{x \in \mathbb{X} \mid h(x) \geq h_{\min}\}$ ).

*Theorem 1:* Consider the CBF candidate  $h(x)$  and the dynamics defined in (1) with control input  $u \in \mathbb{U}$ . Let  $\gamma$  be a class- $\mathcal{K}_e$  function and  $\underline{\gamma}(c)$  be the lower bound on the maximum realizable Lie derivative for a level set  $\mathbb{C}_c$  (see (6)). If  $\gamma$  satisfies  $\underline{\gamma}(c) \leq \gamma(c)$  for all  $c \in [h_{\min}, h_{\max}]$ , then (2) is satisfied for all  $x \in \mathbb{X}_{\max}$ .

*Proof:* We first consider a level set  $\mathbb{C}_c = \{x \in \mathbb{R}^n \mid h(x) = c\}$  with  $c \in [h_{\min}, h_{\max}]$ . For all  $x \in \mathbb{C}_c$ , we have  $\max_{u \in \mathbb{U}} \dot{h}(x, u) \geq \min_{x \in \mathbb{C}_c} \max_{u \in \mathbb{U}} \dot{h}(x, u) = -\underline{\gamma}(c) = -\underline{\gamma}(h(x))$ . It follows that, if  $\underline{\gamma}(c) \leq \gamma(c)$ , then  $-\underline{\gamma}(h(x)) \geq -\gamma(h(x))$  and  $\max_{u \in \mathbb{U}} \dot{h}(x, u) \geq -\underline{\gamma}(h(x)) \geq -\gamma(h(x))$  for all  $x \in \mathbb{C}_c$ . Since  $h$  is continuously differentiable,  $\mathbb{X}$  is compact, and  $c \in [h_{\min}, h_{\max}]$ , the condition  $\max_{u \in \mathbb{U}} \dot{h}(x, u) \geq -\gamma(h(x))$  holds for all  $x \in \mathbb{X}_{\max}$ . ■

The definition of  $\underline{\gamma}$  also allows us to identify control invariant superlevel sets of  $h(x)$ . They are given as follows:

*Corollary 1:* Consider the CBF candidate  $h(x)$ , the dynamics defined in (1) with control input  $u \in \mathbb{U}$ , and  $\underline{\gamma}$  as defined in (6). Then, for  $c \in [h_{\min}, h_{\max}]$ ,  $\underline{\gamma}(c) \geq 0$  implies control invariance of  $\tilde{\mathbb{C}}_c = \{x \in \mathbb{X} \mid h_c(x) = h(x) - c \geq 0\}$ .

*Proof:* According to (6),  $\underline{\gamma}(c) \geq 0$  implies that, for all states in the level set  $\mathbb{C}_c$ , there exists  $u \in \mathbb{U}$  such that the Lie derivative  $\dot{h}(x, u)$  is non-negative. Then, control invariance follows from standard CBF arguments [4]. ■

Alternatively, for a known system, a subset  $\hat{\mathbb{C}} \subset \mathbb{C}$  that is control invariant under the control input constraint set  $\mathbb{U}$  can be found using the approach proposed in [12].

In practice, to compute  $\underline{\gamma}(c)$  for any  $c \in [h_{\min}, h_{\max}]$ , we can rewrite the maximization problem in (6) in the epigraph form [19], which yields (6)  $-\underline{\gamma}(c) = \min_{\rho \in \mathbb{R}, x \in \mathbb{C}_c} \rho$ , s.t.,  $\dot{h}(x, u) \leq \rho, \forall u \in \mathbb{U}$ . This optimization is a nonlinear program that is typically hard to solve. It is possible to reduce its complexity by removing the level set constraint and approximating the optimization in two steps: First, determine  $\rho_c^{j,*} = \min_{\rho \in \mathbb{R}} \rho$ , s.t.,  $\dot{h}(x^j, u) \leq \rho, \forall u \in \mathbb{U}$ , for a set of  $N > 0$  samples  $\{x^j\}_{j=1}^N$  with  $x^j \in \mathbb{C}_c$ . Then approximate  $-\underline{\gamma}(c) \approx \min_j \rho_c^{j,*}$ . Choosing  $N$  sufficiently large yields the desired accuracy of the approximation. In the special case where  $\mathbb{U}$  is a convex polytopic set, we can leverage the control-affine structure in (1) and replace  $\forall u \in \mathbb{U}$  by  $\forall u^i \in \text{vertices}(\mathbb{U})$  (i.e., enforcing the constraint on all the vertices of  $\mathbb{U}$ ). The above procedure applies to a single level set  $\mathbb{C}_c$ . To approximate  $-\underline{\gamma}(c)$  for all  $c \in [h_{\min}, h_{\max}]$ , we can repeat this procedure for a discrete set of  $c^k \in [h_{\min}, h_{\max}]$ , where  $k \in \{1, \dots, K\}$ . Choosing  $K$  sufficiently large yields the desired accuracy over all level sets. The computational complexity of this procedure scales exponentially with the number of state and input dimensions.

## B. Increasing Safe Control Input Set Size

In Sec. IV-A, we presented conditions for the existence of  $u \in \mathbb{U}$  to render the system positively invariant on  $\mathbb{C}$ . Now we turn to the problem of finding a class- $\mathcal{K}_e$  function that enlarges the size of the feasible control input set  $\mathbb{U}_{\text{cbf}}(x)$ . This is motivated by using the CBF condition as a constraint in a minimally invasive safety filter. Ideally, the CBF safety filter should not modify a control input  $u$  unless control invariance cannot be guaranteed. Formally, an uncertified control input  $u$  will not be modified if  $u \in \mathbb{U}_{\text{cbf}}(x)$  for  $x \in \mathbb{X}$ . We note that the set  $\mathbb{U}_{\text{cbf}}(x)$  is dependent on the choice of the class- $\mathcal{K}_e$  function. Therefore, it is desirable to use a class- $\mathcal{K}_e$  function in the CBF condition that allows the largest possible set of safe control inputs at each state  $\mathbb{U}_{\text{cbf}}(x)$ . An enlargement of the admissible control input set can be achieved by any class- $\mathcal{K}_e$  function that upper bounds  $\underline{\gamma}$ . Especially, any class- $\mathcal{K}_e$  function that approaches  $+\infty$  for all  $c \in (0, h_{\max}]$  will satisfy the maximal feasible control input set. However, for practical purposes, it is desirable to design a class- $\mathcal{K}_e$  function that reflects the input constraints. Towards this goal, we propose a target  $\gamma_{\text{target}}$  for  $\gamma \in \mathcal{K}_e$  to match in addition to satisfying the constraint in Thm. 1

$$\gamma_{\text{target}}(c) = - \min_{x \in \mathbb{C}_c, u \in \mathbb{U}} \dot{h}(x, u) \quad (7)$$

for all  $c \in [h_{\min}, h_{\max}]$ . The value of  $-\gamma_{\text{target}}$  determines the minimum value of  $\dot{h}(x, u)$  that can be achieved by at least one state in the level set  $\mathbb{C}_c$  and one control input  $u \in \mathbb{U}$ . The constraint on  $\gamma$  and the proposed target  $\gamma_{\text{target}}$  can help us design a class- $\mathcal{K}_e$  function that optimally accounts for the input constraints. To make the design problem tractable, we consider a parameterized class- $\mathcal{K}_e$  function  $\gamma_{\theta}$ , where  $\theta \in \Theta$  are the function's parameters. The class- $\mathcal{K}_e$  function design problem is then formulated as follows:

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{argmin}} \int_{h_{\min}}^{h_{\max}} \|\gamma_{\theta}(c) - \gamma_{\text{target}}(c)\|_2^2 dc \quad (8a)$$

$$\text{s.t. } \gamma_{\theta}(c) \geq \underline{\gamma}(c), \forall c \in [h_{\min}, h_{\max}]. \quad (8b)$$

In the case of a linear parameterization, we recover an offline version of the presented approaches in [10], [11]. However, our approach aims to certify a CBF candidate function  $h$  as a CBF with a corresponding class- $\mathcal{K}_e$  function *a priori* such that online adaptation for the feasibility of the CBF condition under input constraints is not required. We emphasize that skipping the verification of a CBF under input constraints and solely relying on an online adaptive approach can lead to safety violations. This is due to (5) being independent of the choice of the class- $\mathcal{K}_e$  function. Then, adapting its steepness has no impact. Therefore, the adaptive approach can only be safely applied if the safe set  $\mathbb{C}$  has previously been verified as control invariant under  $\mathbb{U}$ . We demonstrate the insufficiency of the adaptive approach in a simulation example in Sec. VI-A. Furthermore, the softening of the CBF condition (as in [10], [11]) may not be desirable in practice, as the class- $\mathcal{K}_e$  function also impacts the implementation of the CBF safety filter in a discrete-time setting. In discrete time, it is undesirable to approach the boundary too quickly to avoid an overshooting of the safe set boundary. Additional constraints on the slope around the origin can be incorporated for  $\gamma_{\theta}$  in (8).

## V. LEARNING PROBABILISTIC CONTROL INVARIANCE CONDITIONS UNDER INPUT CONSTRAINTS

In this section, we address the issue of guaranteeing control invariance of the safe set  $\mathbb{C}$  under control input constraints when the dynamics are uncertain.

### A. Constructing Probabilistic Control Invariance Conditions Under Input Constraints

We formulate a probabilistic learning approach to approximate the Lie derivative  $\dot{h}(x, u)$ . The probabilistic model could, for example, be Gaussian processes (GPs) or ensembles of deep neural networks [1].

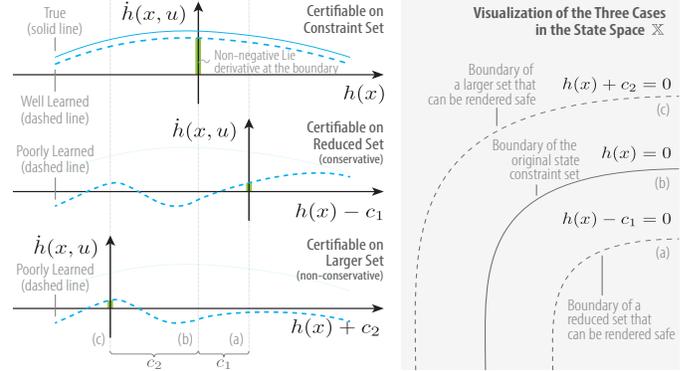
*Assumption 1:* We assume that the probabilistic model for approximating  $\dot{h}(x, u)$  is well-calibrated (i.e., has reliable uncertainty bounds) with  $\mu(x, u) = \bar{\alpha}(x) + \bar{\beta}(x)u$  and  $\sigma^2(x, u)$  denoting the posterior mean and variance functions of the probabilistic model approximating the Lie derivative  $\dot{h}(x, u)$ , respectively. We further assume that, for any  $s > 0$ , there exists a  $\delta \in (0, 1]$  such that  $\Pr(\dot{h}(x, u) \geq \mu(x, u) - s\sigma(x, u)) \geq 1 - \delta$  holds for all  $(x, u) \in \mathbb{X} \times \mathbb{U}$ .

We note that *Asm. 1* is not restrictive in practice. If the distribution's probability density function (PDF) is known, we can generally derive a tight confidence bound  $(1 - \delta)$  for a given  $s > 0$  by directly integrating the PDF. For a distribution with an unknown PDF but with a finite mean  $\mu$  and a finite non-zero standard deviation  $\sigma$ , we can relate  $\delta$  and  $s$  as  $\delta(s) = 1/s^2$  (for  $s > 1$  cases) by applying Chebyshev's inequality [20]. As an example, with Chebyshev's inequality, we can rewrite the confidence bound in *Asm. 1* as  $\Pr(\dot{h}(x, u) \geq \mu(x, u) - s\sigma(x, u)) \geq 1 - 1/s^2, \forall (x, u) \in \mathbb{X} \times \mathbb{U}$ . Using *Asm. 1*, we can construct different probabilistic lower bounds of  $\dot{h}$  by scaling the standard deviation with an appropriate  $s > 0$ . In *Sec. VI-B*, we outline steps to verify this assumption empirically.

Following the discussion in *Sec. IV*, a feasible control invariance condition with probability of at least  $(1 - \delta)$  requires that, for all  $x \in \partial\mathbb{C}$ , there exists an input  $u \in \mathbb{U}$  such that

$$\bar{\alpha}(x) + \bar{\beta}(x)u - s\sigma(x, u) \geq 0 \quad (9)$$

for some  $s > 0$  (see the well-learned case in *Fig. 2*). Intuitively,  $s$  represents the number of standard deviations away from the mean such that the CBF condition holds. If the standard deviation  $\sigma$  at the boundary is large, then we require a smaller number of standard deviations  $s$  to ensure that there exists a control input  $u$  satisfying (9) with probability  $(1 - \delta)$  at any time instance. If no  $s > 0$  satisfies (9) or the largest  $s$  satisfying (9) yields an undesirably low probability  $(1 - \delta)$ , then additional data needs to be collected to reduce the predictive uncertainty of the learned Lie derivative model. Once the desired predictive accuracy is achieved, the next step is determining a class- $\mathcal{K}_e$  function  $\hat{\gamma}$  for the estimated Lie derivative under input constraints. As the true Lie derivative is unknown, we use probabilistic lower bounds of the Lie derivative to provide safety guarantees. We can leverage the approach for finding class- $\mathcal{K}_e$  functions under input constraints presented in *Sec. IV*. Substituting  $\mu(x, u) - s\sigma(x, u)$  for  $\dot{h}(x, u)$  when finding  $\gamma$  and  $\gamma_{\text{target}}$  leads to an optimization problem as formulated in (8) for  $\hat{\gamma}$ . Given the mean  $\mu$  and the standard deviation  $\sigma$ , we provide a sufficient condition for



*Fig. 2.* Illustration showing which superlevel sets of  $h(x)$  can be rendered control invariant. The solid blue line represents the true Lie derivative  $\dot{h}(x, u)$ , and the dashed blue line represents a learned lower bound  $\mu(x, u) - s\sigma(x, u)$ . If there exists at least one control input such that a non-negative Lie derivative lower bound can be achieved for all states in a level set  $\mathbb{C}_c$ , then we can guarantee control invariance with high probability.

satisfying the CBF condition for the true system on  $\mathbb{C}$  with high probability.

*Theorem 2:* Consider the CBF candidate  $h(x)$ , the input constraint set  $\mathbb{U}$ , and a chosen  $s > 0$ . Under *Asm. 1*, if the estimated lower bound of the true Lie derivative  $\mu(x, u) - s\sigma(x, u)$  is non-negative for  $x \in \partial\mathbb{C}$ , then there exists  $\hat{\gamma} \in \mathcal{K}_e$  such that

$$\max_{u \in \mathbb{U}} [\mu(x, u) - s\sigma(x, u)] \geq -\hat{\gamma}(h(x)), \forall x \in \mathbb{C} \quad (10)$$

holds. Moreover, we can guarantee that the CBF condition  $\dot{h}(x, u) \geq -\hat{\gamma}(h(x))$  is satisfied by (1) for some input  $u \in \mathbb{U}$  at any time instance with a probability of at least  $(1 - \delta)$ .

*Proof:* Let  $\hat{\gamma}(c) = -\min_{x \in \mathbb{C}_c} \max_{u \in \mathbb{U}} \mu(x, u) - s\sigma(x, u)$  similar to (6). Due to the non-negativity of  $\mu(x, u) - s\sigma(x, u)$  for all  $x \in \partial\mathbb{C}$ , we can find a  $\hat{\gamma} \in \mathcal{K}_e$  such that  $\hat{\gamma}(c) \leq \hat{\gamma}(c)$  for all  $c \in [0, h_{\max}]$ . Following a similar argument as in the proof for *Thm. 1*, one can show that (10) is feasible for all  $x \in \mathbb{C}$ . Furthermore, by *Asm. 1*,  $\Pr(\dot{h}(x, u) \geq \mu(x, u) - s\sigma(x, u)) \geq 1 - \delta$ . For all  $x \in \mathbb{C}$ , the CBF condition  $\dot{h}(x, u) \geq -\hat{\gamma}(h(x))$  can be then satisfied by the true system for some  $u \in \mathbb{U}$  with a probability of at least  $(1 - \delta)$ . ■

We can use this result to formulate a safety filtering optimization for the uncertain system similar to that defined in (4). For this, we substitute the true Lie derivative  $\dot{h}(x, u)$  in constraint (4b) with the estimated lower bound on the Lie derivative  $\mu(x, u) - s\sigma(x, u)$ , which yields

$$u_s(x) = \underset{u \in \mathbb{U}}{\operatorname{argmin}} \frac{1}{2} \|u - \pi(x)\|_2^2 \quad (11a)$$

$$\text{s.t. } \mu(x, u) - s\sigma(x, u) \geq -\hat{\gamma}(h(x)). \quad (11b)$$

Extension to control invariance guarantees over a finite time interval can be considered in future work [21].

*Remark 1:* In the case where the non-negativity of  $\mu(x, u) - s\sigma(x, u)$  is not satisfied for all  $x \in \partial\mathbb{C}$ , one may potentially render a subset of  $\mathbb{C}$  safe if there exists a  $c \in (0, h_{\max}]$  such that  $\hat{\gamma}(c) \geq 0$  holds (see the second case in *Fig. 2* and *Cor. 1*). If such  $c \in (0, h_{\max}]$  does not exist, then one may only be able to render a larger set safe (see the third case in *Fig. 2*); in this case, we are no longer able to provide closed-loop safety guarantees, and the probabilistic model needs to be retrained using a larger dataset.

## B. Learning Probabilistic Lie Derivative Models

Suppose we are given sampled trajectory data as a training dataset. Then, we can determine a probabilistic model of  $h$  based on data using standard supervised learning methodology. The inputs to the probabilistic model are the state  $x$  and the control input  $u$ , and the output of the model is an estimate of the Lie derivative  $\dot{h}(x, u)$ . In practice, such a dataset could be collected through expert demonstrations.

In the following, we consider an ensemble of  $M > 1$  function approximators as our probabilistic model. For example, this could be an ensemble of deep neural networks with different weight initializations trained on the same training data set  $\mathcal{D}_{\text{train}}$ . We can retain the control affine structure using an ensemble of deep neural networks. For details on retaining the control affine structure with GPs, see [22]. We choose each function approximator as  $\hat{h}_i(x, u) = \alpha_i(x) + \beta_i(x)u$  with  $\alpha_i(x)$  and  $\beta_i(x)$  learned jointly using supervised learning. Then, the mean function is  $\mu(x, u) = \bar{\alpha}(x) + \bar{\beta}(x)u = \frac{1}{M} \sum_{i=1}^M \hat{h}_i(x, u)$  and the variance function is  $\sigma^2(x, u) = \frac{1}{M-1} \sum_{i=1}^M (\hat{h}_i(x, u) - \mu(x, u))^2$ .

The variance function is no longer affine in  $u$  and solving (11) generally requires nonlinear programming. In the special case where the set  $\mathbb{U}$  is convex and polytopic, we can reformulate the constrained optimization problem in (11) as a second-order cone program (SOCP) with a linear objective and two second-order cone constraints as in [22]. We have  $\sigma(x, u) = \|L(x)u + l(x)\|_2$ , where  $L(x)$  is the decomposition of a positive semidefinite matrix  $S(x) = L^T(x)L(x) = \frac{1}{M-1} \sum_{i=1}^M (\beta_i(x) - \bar{\beta}(x))^T (\beta_i(x) - \bar{\beta}(x)) \succeq 0$  and  $l(x) = \sqrt{\frac{1}{M-1} \sum_{i=1}^M (\alpha_i(x) - \bar{\alpha}(x))^2}$ . Plugging this into (11b) yields a second-order cone constraint. The quadratic objective in (11) can be replaced by the second-order cone constraint  $\frac{1}{\sqrt{2}}\|u - \pi(x)\|_2 \leq j$  and by minimizing the linear objective  $j$ .

*Remark 2:* For uncertain systems, the Lie derivative needs to be estimated numerically [13]. The resulting estimation error can be accounted for in the probabilistic model [14].

## VI. EVALUATION RESULTS

### A. Simulation Results

We evaluated the feasibility under input constraints of our proposed offline approach compared to the online adaptive approach presented in [10], [11]. We consider the damped inverted pendulum with the nonlinear control affine dynamics  $\dot{x}_1 = x_2$ ,  $\dot{x}_2 = 1/(ml^2)(-mgl \sin(x_1) - bx_2 + u)$ , with input constraints  $\mathbb{U} = \{u \in \mathbb{R} \mid -15 \text{ Nm} \leq u \leq 15 \text{ Nm}\}$ , mass  $m = 1 \text{ kg}$ , length  $l = 0.5 \text{ m}$ , damping factor  $b = 0.1 \frac{\text{N}}{\text{m}}$ , and the gravitational constant  $g = 9.81 \frac{\text{m}}{\text{s}^2}$ . We consider a CBF candidate  $h(x)$  that is a valid CBF for the pendulum system without input constraints. We assume full model knowledge and the uncertified control policy is  $\pi(x) = 0$ .

Our approach assesses feasibility under input constraints offline before employing the safety filter online. Based on our proposed approach and the remarks on the implementation for convex polytopic input sets  $\mathbb{U}$  in Sec. IV, we verify that the superlevel set of  $h(x) = c = 0.005$  as  $\tilde{\mathbb{C}}_c = \{x \in \mathbb{X} \mid h(x) - c \geq 0\}$  (dark gray in Fig. 3) can be rendered control invariant under the constraint  $\mathbb{U}$ . In contrast, the online adaptive approach employs a strategy that relies

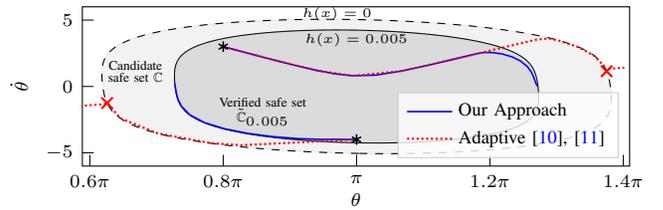


Fig. 3. Simulation of an inverted pendulum in closed-loop with safety filters using either our proposed offline input constraint verification (solid blue trajectories) and the online adaptive modification [10], [11] (dotted red trajectories) given a CBF candidate  $h$ . We use our approach to verify the control invariance of a superlevel set  $\tilde{\mathbb{C}}_{0.005}$  (dark gray) of the CBF candidate  $h$  (zero superlevel set  $\mathbb{C}$  in light gray). Starting from the same initial conditions (indicated by stars) inside the verified safe set, our approach guarantees control invariance, while the online adaptive approach cannot guarantee safety and results in the system leaving the candidate safe set  $\mathbb{C}$  as indicated by the red crosses.

on adapting the steepness of the lower bound in the CBF condition online. However, using this strategy for a candidate safe set  $\mathbb{C}$  (light gray in Fig. 3) that has not been verified *a priori* does not guarantee control invariance. On the boundary of an uncertifiable safe set, the lower bound on the CBF condition cannot be adapted and is zero by definition. In such a case, an online adaptive approach is insufficient for safety. In simulation, we find that for initial conditions starting in our verified control invariant safe set  $\tilde{\mathbb{C}}_{0.005}$ , the adaptive approach may leave this verified safe set and can even leave the larger uncertified candidate safe set  $\mathbb{C}$ , see Fig. 3.

### B. Quadrotor Experiments

We validated our proposed method through physical experiments on a miniature quadrotor, the Crazyflie. 2.1. In our experiments, our proposed safety filter for an uncertain input-constrained system prevents the falling quadrotor from crashing into the floor. In our experiment, we constrain the motion of the quadrotor to its  $z$ -axis by stabilizing its attitude. This gives a system representation with state  $x = [z_{\text{pos}} \ z_{\text{vel}}]^T \in \mathbb{X} \subseteq \mathbb{R}^2$  and the collective delta thrust  $\Delta u \in \mathbb{U} \subseteq \mathbb{R}$  as the input. The delta thrust is defined as  $\Delta u = u - u_0 \in [-0.086mg, +0.086mg]$  where  $u_0$  is the collective thrust to achieve hover,  $m = 0.03 \text{ kg}$  is the quadrotor's mass, and  $g = 9.8 \text{ m/s}^2$  is the gravitational constant. The quadrotor position  $z_{\text{pos}}$  and velocity  $z_{\text{vel}}$  are estimated by a Luenberger observer using position measurements from a motion capture system. We collect state and input data by sending sinusoidal inputs with constant amplitudes and frequencies. The input constraint set  $\mathbb{U}$  is the closed interval of the minimum and maximum applied inputs during the data collection. We use a nominal double integrator linear time-invariant (LTI) model and, from the collected data, fit an ensemble of five neural networks (NN) to the Lie derivative residual to augment the safety filter. Given a CBF candidate  $h(x) = 1 - (x - x_c)^T P (x - x_c)$  with  $P$  being symmetric and positive definite and  $x_c$  the center of the safe set, we apply our approach from Sec. V to verify the zero-superlevel set  $\mathbb{C}$  of the CBF candidate using the learned Lie derivative residual with  $s = 3$  under input constraints  $\mathbb{U}$ . Using Chebyshev's inequality, this yields  $\delta(s) = \frac{1}{s^2} = \frac{1}{9}$ . During the experiments, we start from hovering with an approximate initial condition  $x_0 = [1.4, 0.0]^T$ . Then, an uncertified control policy  $\pi(x) = -0.03mg$  is applied with the SOCP-based safety filter (11). After each execution, we verified the satisfaction of Asm. 1 for our proposed approach by determining

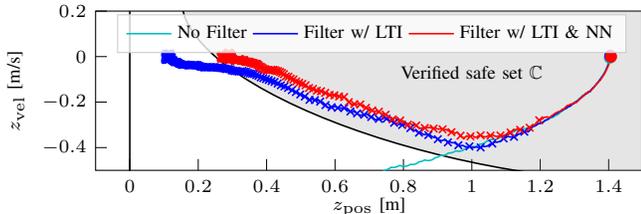


Fig. 4. Quadrotor state trajectories under unsafe dropping input (cyan) and CBF-augmented input using an LTI system (blue) and an LTI system with learned Lie derivative residual (red). The safe set with its interior (light grey) and its boundary (black) is defined as an ellipse. Solid circles indicate the initial states. Crosses indicate the states where the safety filter is active.

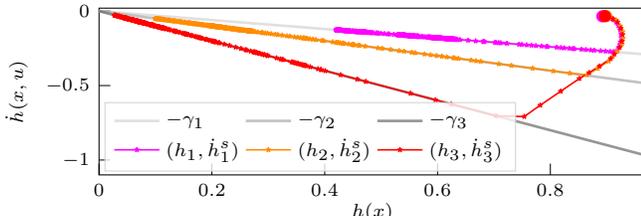


Fig. 5. Real-world trajectories of  $h$  and  $\hat{h}^s = \hat{h}_{\text{nominal}}(x, u) + \mu(x, u) - \sigma(x, u)$ , where  $\hat{h}_{\text{nominal}}$  is the nominal Lie derivative and  $s = 3$ , for three different linear extended class- $\mathcal{K}_e$  functions (in different shades of gray). Solid circles indicate the initial states.

the percentage of predicted Lie derivative residuals inside the confidence interval given by  $s = 3$ . The experiment results are presented in Fig. 4 and Fig. 5. A video of this experiment can be found here: <http://tiny.cc/CBF>.

The state trajectories without safety filter and with safety filter for the LTI system and the learned Lie derivative are shown in Fig. 4. Without the safety filter, the quadrotor leaves the safe set and hits the floor. With the safety filter that leverages the learned Lie derivative and the nominal LTI system, the system strictly stays inside the safe set, while the safety filter solely relying on the LTI system fails. The learned Lie derivative using the ensemble of neural networks allows us to explicitly account for uncertainty in the system, including model mismatch and disturbances in real-world systems. The effect of different linear class- $\mathcal{K}_e$  functions can be seen in Fig. 5. Class- $\mathcal{K}_e$  functions define the tolerable Lie derivative  $\hat{h}$  to guarantee safety when the system approaches the safe set's boundary. The quadrotor approaches the boundary of the safe set (and therefore the ground) faster for less conservative class- $\mathcal{K}_e$  like  $\gamma_3$ . A less steep function (e.g., as given by  $\gamma_1$ ) achieves a more conservative behavior. In this case, the optimization of the class- $\mathcal{K}_e$  is not restrictive, as the maximum realizable Lie derivative is positive for all level sets  $h(x) \in [0.0, 0.7]$ . However, the class- $\mathcal{K}_e$  must be carefully designed to achieve safe operation at discrete time steps.

## VII. CONCLUSION

In this paper, we proposed a systematic approach for learning valid CBF conditions for uncertain systems subject to input constraints. Based on known dynamics, we first derived a sufficient condition on the class- $\mathcal{K}_e$  function for validating a CBF safety filter design under input constraints. We proposed an approach to determine optimal class- $\mathcal{K}_e$  functions for a given CBF. Building on these results, we developed a learning-based method to construct probabilistic control invariance conditions for input-constrained uncertain systems. Given the

probabilistic CBF condition, we can design a CBF safety filter to guarantee the system's safety with high probability. In simulation and real-world robotic experiments, we demonstrated the efficacy of the learned safety filters and showed that optimized class- $\mathcal{K}_e$  functions can be less conservative and are significant for designing safe real-world operation.

## REFERENCES

- [1] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 411–444, 2022.
- [2] A. D. Ames, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs with application to adaptive cruise control," *Proc. of the IEEE Conf. on Decision and Control (CDC)*, pp. 6271–6278, 2014.
- [3] H. Dai and F. Permenter, "Convex synthesis and verification of control-Lyapunov and barrier functions with input constraints," in *Proc. of the IEEE American Control Conf. (ACC)*, 2023, accepted.
- [4] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *Proc. of the European Control Conf. (ECC)*, 2019, pp. 3420–3431.
- [5] M. Nagumo, "Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen," *Proc. of the Physico-Mathematical Society of Japan, 3rd Series*, vol. 24, pp. 551–559, 1942.
- [6] I. Mitchell, A. Bayen, and C. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *IEEE Trans. on Automatic Control*, vol. 50(7), pp. 947–957, 2005.
- [7] H. Wang, K. Margellos, and A. Papachristodoulou, "Safety verification and controller synthesis for systems with input constraints," *IFAC-PapersOnLine*, 2023, 22nd IFAC World Congress, accepted.
- [8] T. Wei, S. Kang, W. Zhao, and C. Liu, "Persistently feasible robust safe control by safety index synthesis and convex semi-infinite programming," *IEEE Control Systems Letters*, vol. 7, pp. 1213–1218, 2023.
- [9] Z. Qin, D. Sun, and C. Fan, "Sablas: Learning safe control for black-box dynamical systems," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1928–1935, 2022.
- [10] W. Xiao, C. Belta, and C. G. Cassandras, "Adaptive control barrier functions," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2267–2281, 2022.
- [11] J. Zeng, B. Zhang, Z. Li, and K. Sreenath, "Safety-critical control using optimal-decay control barrier function with guaranteed point-wise feasibility," in *Proc. of the IEEE American Control Conf. (ACC)*, 2021, pp. 3856–3863.
- [12] D. R. Agrawal and D. Panagou, "Safe control synthesis via input constrained control barrier functions," in *Proc. of the IEEE Conf. on Decision and Control (CDC)*, 2021, pp. 6113–6118.
- [13] A. Taylor, A. Singletary, Y. Yue, and A. Ames, "Learning for safety-critical control with control barrier functions," in *Proc. of the Learning for Dynamics and Control Conf. (LADC)*, vol. 120, 2020, pp. 708–717.
- [14] L. Brunke, S. Zhou, and A. P. Schoellig, "Barrier Bayesian linear regression: Online learning of control barrier conditions for safety-critical control of uncertain systems," in *Proc. of the Learning for Dynamics and Control Conf. (LADC)*, vol. 168, 2022, pp. 881–892.
- [15] L. Wang, E. A. Theodorou, and M. Egerstedt, "Safe learning of quadrotor dynamics using barrier certificates," in *Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2018, pp. 2460–2465.
- [16] M. Ohnishi, L. Wang, G. Notomista, and M. Egerstedt, "Barrier-certified adaptive reinforcement learning with applications to brushbot navigation," *IEEE Trans. on Rob.*, vol. 35(5), pp. 1186–1205, 2019.
- [17] F. Castañeda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, "Pointwise feasibility of gaussian process-based safety-critical control under model uncertainty," in *Proc. of the IEEE Conf. on Decision and Control (CDC)*, 2021, pp. 6762–6769.
- [18] A. Nejati and M. Zamani, "Data-driven synthesis of safety controllers via multiple control barrier certificates," *IEEE Control Systems Letters*, vol. 7, pp. 2497–2502, 2023.
- [19] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [20] P. Billingsley, *Probability and Measure*, ser. Wiley Series in Probability and Statistics. Wiley, 2012.
- [21] T. Lew, A. Sharma, J. Harrison, E. Schmerling, and M. Pavone, "On the problem of reformulating systems with uncertain dynamics as a stochastic differential equation," 2021.
- [22] F. Castañeda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, "Gaussian process-based min-norm stabilizing controller for control-affine systems with uncertain input effects and dynamics," in *2021 American Control Conference (ACC)*, 2021, pp. 3683–3690.