Robust Data-Driven Control Barrier Functions for Unknown Continuous Control Affine Systems

Zeyuan Jin, *Graduate Student Member, IEEE*, Mohammad Khajenejad, *Member, IEEE*, and Sze Zheng Yong, *Member, IEEE*

***Abstract*—In this letter, we introduce robust data-driven control barrier functions (CBF-DDs) to guarantee robust safety of unknown continuous control affine systems despite worst-case realizations of generalization errors from prior data under various continuity assumptions. To achieve this, we leverage non-parametric data-driven approaches for learning guaranteed upper and lower bounds of an unknown function from the data set to formulate/obtain a safe input set for a given state. By incorporating the safe input set into an optimization-based controller, the safety of the system can be ensured. Moreover, we present several complexity reduction approaches including providing subproblems that can be solved in parallel and downsampling strategies to improve computational performance.**

***Index Terms*—Constrained control, identification for control, optimization.**

I. Introduction

IN RECENT years, various control approaches have been developed to guarantee system safety, in addition to stability, especially for safety-critical applications such as bipedal robots, multi-agent systems, and adaptive cruise control systems. Specifically, a huge body of research proposed set invariance-based techniques, including designing barrier certificates, e.g., [1], and control barrier functions (CBFs), e.g., [2], to encode the forward controlled invariance of safety sets. However, disturbances, noise, and parameter uncertainties are inevitable and the assumption that known (parametric) mathematical models of the system dynamics is accessible is not always justifiable in real-world applications. Hence, addressing safety for systems with (partially) unknown dynamical models, including with the CBF framework, is an interesting and challenging problem.

Manuscript received 16 September 2022; revised 27 November 2022; accepted 20 December 2022. Date of publication 11 January 2023; date of current version 25 January 2023. This work was supported in part by NSF under Grant CNS-1943545 and Grant CNS-1932066. Recommended by Senior Editor S. Tarbouriech. *(Corresponding author: Sze Zheng Yong.)*

Zeyuan Jin is with the School for Engineering, Matter and Transport, Arizona State University, Tempe, AZ 85287 USA (e-mail: zjin43@asu.edu).

Mohammad Khajenejad is with the Department of Mechanical and Aerospace Engineering, University of California at San Diego, San Diego, CA 92122 USA (e-mail: mkhajenejad@ucsd.edu).

Sze Zheng Yong is with the Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: s.yong@northeastern.edu).

Digital Object Identifier 10.1109/LCSYS.2023.3235958

Literature review: In the context of systems with uncertain/noisy but known (parametric) models, the work in [3], [4] studied the robustness of CBFs to additive perturbations and model uncertainty. Further, adaptive/robust adaptive CBFs were introduced in [5], [6] to estimate the unknown parameters by means of adaptation laws, while [7] additionally investigated the fixed-time convergence of parameter estimates under persistence of excitation. On the other hand, robust CBFs were proposed, e.g., in [8], for ensuring safety of all possible values of parametric uncertainties, with the price of potentially being conservative. Further, the work in [9] considered robustness to both additive disturbances and state estimation errors.

On the other hand, when models of systems are unknown but input-output data sets are instead available, data-driven methods also need to be developed for maintaining system safety. The majority of data-driven CBF approaches rely on first identifying or learning the unknown dynamical models or their deviations from some assumed nominal models from data, e.g., using neural networks and reinforcement learning [10], [11], polynomial [12] or Gaussian process regression [13], [14], [15], and using these data-driven models to ensure safety. However, these methods often only provide probabilistic safety guarantees instead of being robust to worst-case generalization errors, i.e., the out-of-sample difference between the true system and learned models, that, although unlikely, may still lead to safety violations.

An interesting recent work [16] explicitly considered robustness to worst-case generalization errors using control certificate functions by assuming Lipschitz continuity of the deviation between the unknown dynamics and an assumed nominal model, but when incorporated into an optimization-based controller, the resulting second-order cone problems (SOCPs) are often not amenable to run-time implementation. This approach generally falls under the umbrella of non-parametric learning approaches, where no functional forms or features (e.g., linear, control-affine, polynomial or Gaussian) are assumed and the number of parameters are not fixed but grows with the size of data sets. More importantly, many non-parametric learning approaches enable the worst-case generalization errors to be bounded (independently of the number of data points) under some continuity assumptions such as Lipschitz or Hölder continuity or bounded Jacobians, e.g., [17], [18], [19], which we will leverage in this letter to obtain robust data-driven CBFs that are more computationally amenable.

Contributions: This letter proposes robust data-driven control barrier functions (CBF-DD) that can guarantee (global) robust safety despite worst-case generalization errors when only data sets of (safe/unsafe) demonstrations of unknown dynamical systems are available. In contrast to state-of-the-art techniques, e.g., [16], our proposed CBF-DD is a novel end-to-end method that directly learns sufficient conditions for making a CBF candidate controlled invariant, instead of first learning the system dynamics or (worst-case) deviations from assumed nominal models before enforcing safety via CBFs. Moreover, we consider various continuity assumptions when developing our robust CBF-DD by leveraging our prior experience with computing robust data-driven models [19], [20]. The corresponding sufficient conditions for data-driven robust controlled invariance can also be encoded within an optimization-based controller and are often less computationally intensive than solving SOCPs in [16]. Further, since our approach falls under the umbrella of non-parametric learning that are known to not scale well with the size of the data sets, we also propose several computational complexity reduction approaches, namely parallelization and downsampling strategies. The effectiveness of our approach is demonstrated on various examples, including for adaptive cruise control.

II. Preliminaries and Problem Formulation

A. Notations and Definitions

$\mathbb{R}^n, \mathbb{R}_+, \mathbb{R}_{\geq 0}$ and $\mathbb{Z}_n^+$ represent a set of real numbers of dimension n, a set of non-negative and positive real numbers, and a set of positive integers up to n, respectively. I_n represents an identity matrix of size n and $\mathbf{0}_{m\times n}$ represents a matrix of zeros of size $m \times n$. Moreover, for a vector/matrix M, $M^+ \triangleq \max\{M, 0\}$ and $M^- \triangleq M^+ - M$ with element-wise maximum, and vector inequalities are also element-wise. Further, a continuous function α: $[0, a) \to \mathbb{R}_+$ with $a \in \mathbb{R}_+$ is a class $\mathcal{K}$ function if it is strictly monotonically increasing and $\alpha(0) = 0$, and is a class $\mathcal{K}_\infty$ function if it is a class $\mathcal{K}$ function with $a = \infty$ and $\lim_{x\to\infty} \alpha(x) = \infty$.

We will rely on the following proposition (that slightly generalizes [21, Lemma 1]) to later prove Theorem 2.

Proposition 1: Let $\underline{x} \leq x \leq \overline{x} \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$ be decomposed as $A = A_u - A_l$ for any $A_u, A_l \geq 0$. Then,

$$\begin{aligned} A_u\underline{x} - A_l\overline{x} &\leq A^+\underline{x} - A^-\overline{x} \leq Ax \\ &\leq A^+\overline{x} - A^-\underline{x} \leq A_u\overline{x} - A_l\underline{x}. \end{aligned}$$

Proof: It follows from [21, Lemma 1] and the fact that $A_u, A_l \geq 0$, $A_u\underline{x} \leq A_u x \leq A_u\overline{x}$ and $A_l\underline{x} \leq Ax \leq A_l\overline{x} \Leftrightarrow -A_l\overline{x} \leq -Ax \leq -A_l\underline{x}$. It then follows from summing up these two inequalities that $A_u\underline{x} - A_l\overline{x} \leq Ax \leq A_u\overline{x} - A_l\underline{x}$. Further, again by [21, Lemma 1], it follows that $A^+\underline{x} - A^-\overline{x} \leq Ax \leq A^+\overline{x} - A^-\underline{x}$. Moreover, it has been shown in [22, Lemma 1] that these bounds are tighter than any other possible bounds. All these together return the results. ■

B. Control Barrier Function Fundamentals

Our work builds upon the concept of control barrier functions in the literature that enforce *safety* using controlled forward invariance of a safety set. Hence, in the following, we first recap the fundamentals of control barrier functions for *known* control affine systems, given by

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

where $x \in \mathcal{X} \subset \mathbb{R}^n$ and $u \in \mathcal{U} \subset \mathbb{R}^m$ are state and input vectors, respectively, with *known* functions f and g and the set $\mathcal{U}$ can be used to capture (convex) input constraints.

Definition 1 (Control Barrier Functions [2]): Let $\mathcal{S} \subseteq \mathcal{X}$ be the superlevel set of a continuously differentiable function $h : \mathcal{X} \to \mathbb{R}$, i.e., $\mathcal{S} \triangleq \{x \in \mathcal{X} \mid \exists u \in \mathcal{U} \text{ s.t. } h(x) \geq 0\}$. Then, h is a control barrier function (CBF) if there exists an extended class $\mathcal{K}_\infty$ function α such that for the control system (1),

$$\sup_{u\in\mathcal{U}} \dot{h}(x, u) \geq -\alpha(h(x)), \tag{2}$$

for all $x \in \mathcal{X}$, where $\dot{h}(x, u) = \nabla h(x)^\top f(x) + \nabla h(x)^\top g(x)u$.

Theorem 1 (Safety With CBF [2]): For an known control affine system (1) with a CBF $h : \mathcal{X} \to \mathbb{R}$ defined in Definition 1 and the associated safety set $\mathcal{S}$, any Lipschitz continuous controller $u(x)$ that satisfies (2) for the system (1) renders the set $\mathcal{S}$ controlled invariant (thus, safe).

Furthermore, given any (legacy) feedback controller $u = k(x) \in \mathcal{U}$ for the nonlinear control affine system (1), in order to minimally modify the controller while guaranteeing safety, the following optimization problem can be considered [2]:

$$u(x) = \arg\min_{u\in\mathcal{U}} \frac{1}{2}\|u - k(x)\|_2^2 \tag{3a}$$

$$s.t. \quad \dot{h}(x, u) \geq -\alpha(h(x)), \tag{3b}$$

where (3b) is derived from (2).

C. Problem Formulation

In this letter, we consider an *unknown* nonlinear control affine system that is similarly given by (1) with an input constraint set $\mathcal{U}$, but with *unknown* continuous functions f and g, which is common in the real-world setting since accurate mathematical models are often unavailable or hard to obtain. Instead, we assume that a candidate CBF h and prior data of the system operation are available, as follows.

Assumption 1: A continuously differentiable candidate CBF $h : \mathcal{X} \to \mathbb{R}$ and an a priori collected input-output data set $\mathcal{D} = \{(\dot{h}(x_i, u_i), x_i, u_i)\}_1^N$ with $x_i \in \mathcal{X}$, $u_i \in \mathcal{U}$, $i = 1, \ldots, N$, from the unknown system (1) are available. For convenience, we denote $\dot{h}(x_i, u_i)$ as $\dot{h}_i$.

Remark 1: The assumption of measured $\dot{h}(x_i, u_i)$ is mainly for ease of exposition. In practice, $\dot{h}(x_i, u_i) = \frac{\partial h}{\partial x}(x_i)(f(x_i) + g(x_i)u_i)$ is unknown even when $h(x_i)$ and $\frac{\partial h}{\partial x}(x_i)$ is known since f and g are unknown, but can be obtained/computed from time series/trajectory data of $h(t)$ and/or $x(t)$, e.g., using finite difference, and if needed, the errors introduced by applying finite difference can be considered as noise and incorporated in the learning process. Interested readers are referred to [19], [20] for more details.

Note that the candidate CBF h can be any desired safety condition, e.g., lane keeping with desired maximum acceleration [2]; thus, it is reasonable and realistic to assume that h is known even when the system dynamics is not. Further, note that we do not assume knowledge of the *function* $\dot{h}(x, u)$, even when $\{\dot{h}(x_i, u_i)\}$ is measured or approximated. Thus, the constraints in (2) and (3b) cannot be directly enforced. In light of the above, instead of first learning the functions $\dot{x}$ from data (as is commonly done in the literature), this letter proposes to directly learn the function $\dot{h}(x, u)$ from data (cf. Assumption 1). The benefit of doing this end-to-end process is that only a 1-dimensional function $\dot{h}(x, u)$ needs to be learned instead of an n-dimensional function $\dot{x}$; thus, the resulting robust data-driven CBF is simpler and more computationally efficient.

In particular, we wish to learn an over-approximation of the function $\dot{h}(x, u)$ such that its generalization errors can be characterized. For this, we will employ non-parametric learning approaches under some continuity assumptions.

Assumption 2 (Several Continuity Cases): The function $\dot{h} : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ is

I globally Lipschitz continuous with known Lipschitz constants $L_x \in \mathbb{R}_{\geq 0}$ and $L_u \in \mathbb{R}_+$, i.e., $|\dot{h}(x', u') - \dot{h}(x'', , u'')| \leq L_x\|x' - x''\|_p + L_u\|u' - u''\|_p$, with $p \in \{1, 2, \dots, \infty\}$,

II globally componentwise Lipschitz continuous with known componentwise Lipschitz constants $L_x \in \mathbb{R}^n_{\geq 0}$ and $L_u \in \mathbb{R}^m_+$, i.e., $|\dot{h}(x', u') - \dot{h}(x'', u'')| \leq L_x^\top|x' - x''| + L_u^\top|u' - u''|$, and/or

III differentiable with respect to x and u with globally bounded Jacobians with known finite-valued matrices $\overline{J}, \underline{J} \in \mathbb{R}^{n+m}_{\geq 0}$, i.e., $\underline{J} \leq \nabla\dot{h}(x, u) \leq \overline{J}$,

for all $x', x'' \in \mathcal{X} \subseteq \mathbb{R}^n$ and $u', u'' \in \mathcal{U} \subseteq \mathbb{R}^m$, where all inequalities and absolute values are componentwise.

A sufficient condition for the above continuity assumptions to hold for $\dot{h}$ is that they also hold for h, f and g. Moreover, when $p = 1$, Assumptions 2-*II)* and 2-*III)* are less restrictive than 2-*I)* (cf. Proposition 3).

Then, the problem we seek to address in this letter can be formally stated as follows.

Problem 1 (Robust Data-Driven CBF (CBF-DD)): Given a nonlinear unknown control affine system (1) as well as a continuously differentiable CBF candidate $h : \mathcal{R}^n \to \mathcal{R}$ satisfying one of the continuity assumptions in Assumption 2 and an *a priori* data set satisfying Assumption 1, find sufficient conditions for the robust controlled invariance of the safe set $\mathcal{S} \triangleq \{x \in \mathcal{X} \mid \exists u \in \mathcal{U} \text{ s.t. } h(x) \geq 0\}$ (with state feedback), i.e., the controlled invariance of $\mathcal{S}$ despite worst-case generalization errors.

A CBF candidate h that satisfies the robust controlled invariance condition using (only) data is called a robust data-driven control barrier function (CBF-DD).

III. Main Results

In this section, we derive CBF-DDs under some continuity assumptions (cf. Assumption 2), and provide and categorize the corresponding optimization problems for optimization-based control in terms of their computational complexity.

A. Data-Driven Control Barrier Functions (CBF-DDs)

First, we define CBF-DDs for the various continuity cases.

1) Case I (Lipschitz Continuous): We begin with the globally Lipschitz continuous case in Assumption 2-I). From definition, we can directly obtain a CBF-DD for this case using Lipschitz interpolation [20].

Definition 2 (CBF-DD-L): For an unknown control affine system (1) with input-output data set $\mathcal{D}$ satisfying Assumption 1, a continuously differentiable function $h : \mathcal{X} \to \mathbb{R}$ whose time-derivative $\dot{h}$ satisfies Assumption 2-I) is a *Lipschitz* robust data-driven control barrier function (CBF-DD-L) for the safe set $\mathcal{S} \triangleq \{x \in \mathcal{X} \mid \exists u \in \mathcal{U} \text{ s.t. } h(x) \geq 0\}$, if there exists a class $\mathcal{K}_\infty$ function $\alpha(\cdot)$ such that

$$\sup_{u\in\mathcal{U}} \max_{i\in\mathbb{Z}_N^+} \dot{h}_i - L_x\|x - x_i\|_p - L_u\|u - u_i\|_p \geq -\alpha(h(x)),$$

for all $x \in \mathcal{X}$ and $t \geq 0$. Moreover, for any $x \in \mathcal{S}$, we define the corresponding safe input set:

$$K_{\mathcal{S}}^L(x) = \{u \in \cup_{i=1}^N \mathcal{U}_i^L(x)\}, \tag{4}$$

where $\mathcal{U}_i^L(x)$ is defined as follows:

$$\mathcal{U}_i^L(x) \triangleq \{u|\dot{h}_i - L_x\|x - x_i\|_p - L_u\|u - u_i\|_p \geq -\alpha(h(x))\}.$$

2) Case II (Componentwise Lipschitz Continuous): Next, we consider the globally componentwise Lipschitz continuous case in Assumption 2-II). From definition, we can directly obtain a CBF-DD for this case using componentwise Lipschitz interpolation [19].

Definition 3 (CBF-DD-CL): For an unknown control affine system (1) with input-output data set $\mathcal{D}$ satisfying Assumption 1, a continuously differentiable function $h : \mathcal{X} \to \mathbb{R}$ whose time-derivative $\dot{h}$ satisfies Assumption 2-II) is a *componentwise Lipschitz* robust data-driven control barrier function (CBF-DD-CL) for the safety set $\mathcal{S} \triangleq \{x \in \mathcal{X} \mid \exists u \in \mathcal{U} \text{ s.t. } h(x) \geq 0\}$, if there exists a class $\mathcal{K}_\infty$ function $\alpha(\cdot)$ such that

$$\sup_{u\in\mathcal{U}} \max_{i\in\mathbb{Z}_N^+} \dot{h}_i - L_x^\top|x - x_i| - L_u^\top|u - u_i| \geq -\alpha(h(x)),$$

for all $x \in \mathcal{X}$ and $t \geq 0$. Moreover, for any $x \in \mathcal{S}$, we define the corresponding safe input set:

$$K_{\mathcal{S}}^{CL}(x) = \{u \in \cup_{i=1}^N \mathcal{U}_i^{CL}(x)\}, \tag{5}$$

where $\mathcal{U}_i^{CL}(x)$ is defined as follows:

$$\mathcal{U}_i^{CL}(x) \triangleq \{u|\dot{h}_i - L_x^\top|x - x_i| - L_u^\top|u - u_i| \geq -\alpha(h(x))\}.$$

3) Case III (Bounded Jacobians): Finally, we consider the globally bounded Jacobian case in Assumption 2-III). From definition and by leveraging Proposition 1, we can obtain two CBF-DD variants for this case using results from [19].

Definition 4 (CBF-DD-J1): For an unknown control affine system (1) with input-output data set $\mathcal{D}$ satisfying Assumption 1, a continuously differentiable function $h : \mathcal{X} \to \mathbb{R}$ whose time-derivative $\dot{h}$ satisfies Assumption 2-III) is a *bounded Jacobian* robust data-driven control barrier function (CBF-DD-J1) for the safety set $\mathcal{S} \triangleq \{x \in \mathcal{X} \mid \exists u \in \mathcal{U} \text{ s.t. } h(x) \geq 0\}$, if there exists a class $\mathcal{K}_\infty$ function $\alpha(\cdot)$ such that

$$\sup_{u\in\mathcal{U}} \max_{i\in\mathbb{Z}_N^+} \begin{matrix} \dot{h}_i + \underline{J}_x\Delta x_i^+ - \overline{J}_x\Delta x_i^- \\ +\underline{J}_u\Delta u_i^+ - \overline{J}_u\Delta u_i^- \end{matrix} \geq -\alpha(h(x)),$$

for all $x \in \mathcal{X}$ and $t \geq 0$, where $\Delta x_i \triangleq x - x_i$ and $\Delta u_i \triangleq u - u_i$. Moreover, for any $x \in \mathcal{S}$, we define the corresponding safe input set:

$$K_{\mathcal{S}}^{J1}(x) = \{u \in \cup_{i=1}^N \mathcal{U}_i^{J1}(x)\}, \tag{6}$$

where $\mathcal{U}_i^{J1}(x)$ is defined as follows:

$$\mathcal{U}_i^{J1}(x) \triangleq \{u|\dot{h}_i + \underline{J}_x\Delta x_i^+ - \overline{J}_x\Delta x_i^- + \underline{J}_u\Delta u_i^+ - \overline{J}_u\Delta u_i^- \geq -\alpha(h(x))\}. \tag{7}$$

Definition 5 (CBF-DD-J2): For an unknown control affine system (1) with input-output data set $\mathcal{D}$ satisfying Assumption 1, a continuously differentiable function $h : \mathcal{X} \to \mathbb{R}$ whose time-derivative $\dot{h}$ satisfies Assumption 2-III) is a *bounded Jacobian* robust data-driven control barrier function (CBF-DD-J2) for the safety set $\mathcal{S} \triangleq \{x \in \mathcal{X} \mid \exists u \in \mathcal{U} \text{ s.t. } h(x) \geq 0\}$, if there exists a class $\mathcal{K}_\infty$ function $\alpha(\cdot)$ and auxiliary (decision) variables $u_i^\oplus$, $u_i^\ominus$ such that

$$\begin{array}{l}\sup_{u\in\mathcal{U}, u_i^\oplus, u_i^\ominus} \max_{i\in\mathbb{Z}_N^+} \dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u u_i^\oplus - \overline{J}_u u_i^\ominus \\ \text{s.t. } u_i^\oplus \geq 0, u_i^\ominus \geq 0, u_i^\oplus - u_i^\ominus = u - u_i \end{array} \geq -\alpha(h(x)),$$

for all $x \in \mathcal{X}$ and $t \geq 0$, where $\Delta x_i \triangleq x - x_i$. Moreover, for any $x \in \mathcal{S}$, we define the corresponding safe input set:

$$K_\mathcal{S}^{J2}(x) = \{u \in \cup_{i=1}^N \mathcal{U}_i^{J2}(x)\}, \quad (8)$$

where $\mathcal{U}_i^{J2}(x)$ is defined as follows:

$$\begin{aligned}\mathcal{U}_i^{J2}(x) \triangleq \{u | &\dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u u_i^\oplus - \overline{J}_u u_i^\ominus \\ &\geq -\alpha(h(x)), u_i^\oplus \geq 0, u_i^\ominus \geq 0, u_i^\oplus - u_i^\ominus = u - u_i\}. \quad (9)\end{aligned}$$

These CBF-DDs are then shown to guarantee robust safety.

Theorem 2 (Robust Safety with CBF-DD-ϕ): For an unknown control affine system (1) with input-output data set $\mathcal{D}$ satisfying Assumption 1, a CBF-DD-ϕ with $\phi \in \{L, CL, J1, J2\}$, $h : \mathcal{X} \to \mathbb{R}$, as defined in Definitions 2–5, respectively, and their associated safety sets $\mathcal{S}$, any Lipschitz continuous controller $u(x) \in K_\mathcal{S}^\phi(x)$ for the system (1) render the set $\mathcal{S}$ robustly controlled invariant.

Proof: We prove the results for each ϕ:

I $\phi = L$: Lipschitz interpolation approaches, e.g., [20], ensure that despite worst-case generalization errors,

$$\dot{h}(x, u) \geq \max_{i\in\mathbb{Z}_N^+} \dot{h}_i - L_x\|x - x_i\|_p - L_u\|u - u_i\|_p.$$

II $\phi = CL$: From the definition of componentwise Lipschitz continuity (cf. Assumption 2-II)), the following holds despite worst-case generalization errors:

$$\dot{h}(x, u) \geq \max_{i\in\mathbb{Z}_N^+} \dot{h}_i - L_x^\top |x - x_i| - L_u|u - u_i|.$$

III $\phi = J1$ or $J2$: From the learning approach for bounded Jacobian functions in [19] and Proposition 1, despite worst-case generalization errors, we have

$$\begin{aligned}\dot{h}(x, u) &\geq \max_{i\in\mathbb{Z}_N^+} \dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u \Delta u_i^+ - \overline{J}_u \Delta u_i^- \\ &\geq \max_{i\in\mathbb{Z}_N^+} \dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u u_i^\oplus - \overline{J}_u u_i^\ominus, \quad (10)\end{aligned}$$

where $\Delta x_i \triangleq x - x_i$, $\Delta u_i \triangleq u - u_i$, $u_i^\oplus, u_i^\ominus \geq 0$ and $u^\oplus - u^\ominus = u - u_i = \Delta u_i$.

Thus, for all cases, the constraint in (2) holds and robust controlled invariance follows from Theorem 1. ■

The generalization error for each of the above cases can be defined as the difference between the left and right hand sides of the inequalities in the proof. As the data size increases (to more densely cover the entire $\mathcal{X}$ and $\mathcal{U}$), the generalization error decreases. More importantly, since we bound the right hand side by $-\alpha(h(x))$, $\dot{h}(x, u)$ is *always* robustly bounded, independently of the data size or sampling procedure.

B. Optimization-Based Controller and Problem Classes

Next, equipped with the various CBF-DDs in the previous section, we can directly obtain an optimization-based controller that to slightly perturb any (legacy) feedback controller to guarantee safety in the exact same way as (3), except that the constraint in (3b) is replaced by $u \in K_\mathcal{S}^\phi(x)$, $\phi \in \{L, CL, J1, J2\}$ from (4), (5), (6) or (8).

Moreover, we show that despite the controlled invariance condition for CBF-DD-J2 being more conservative than the one for CBF-DD-J1, as seen in (10), there is no loss of optimality for an optimization-based controller (3) with (3b) using $u \in K_\mathcal{S}^{J2}(x)$ when compared to using $u \in K_\mathcal{S}^{J1}(x)$.

Proposition 2: The solution to the optimization-based controller (3) with (3b) replaced by $u \in K_\mathcal{S}^{J1}(x)$ (cf. (6)) is equivalent to the solution to the optimization-based controller (3) with (3b) replaced by $u \in K_\mathcal{S}^{J2}(x)$ (cf. (8)).

Proof: This can be proved by showing that the $\mathcal{U}_i^{J1}(x)$ defined in (7) is same as the $\mathcal{U}_i^{J2}(x)$ constrained by (9).

First, we prove $\forall u \in \mathcal{U}_i^{J1}(x)$, $u \in \mathcal{U}_i^{J2}(x)$. For any $u \in \mathcal{U}_i^{J1}(x)$, $u^\oplus$ and $u^\ominus$ can be selected as $(u - u_i)^+$ and $(u - u_i)^-$ respectively, so that u satisfies (9).

Next, we prove $\forall u \notin \mathcal{U}_i^{J1}(x)$, $u \notin \mathcal{U}_i^{J2}(x)$. Since $u \notin \mathcal{U}_i^{J1}(x)$, $\dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u \Delta u_i^+ - \overline{J}_u \Delta u_i^- < -\alpha(h(x))$. According to Proposition 1, $\dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u u^\oplus - \overline{J}_u u^\ominus \leq \dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u \Delta u_i^+ - \overline{J}_u \Delta u_i^- < -\alpha(h(x))$, which means that there does not exist any $u^\oplus$ and $u^\ominus$ such that (9) can be satisfied for $u \notin \mathcal{U}_i^{J1}(x)$, i.e., $u \notin \mathcal{U}_i^{J2}(x)$. ■

In addition, we derive the relationship between the safe input sets for different CBF-DDs given the same data set.

Proposition 3: Given the same data point $(\dot{h}(x_i, u_i), x_i, u_i)$, the same comparison function $\alpha(\cdot)$ and p is chosen to be 1 for $\mathcal{U}_i^L(x)$, then $\mathcal{U}_i^L(x) \subseteq \mathcal{U}_i^{CL}(x) \subseteq \mathcal{U}_i^{J1}(x) = \mathcal{U}_i^{J2}(x)$.

Consequently, given the same data set $\mathcal{D}$ and if $p = 1$ in $K_\mathcal{S}^L(x)$, then $K_\mathcal{S}^L(x) \subseteq K_\mathcal{S}^{CL}(x) \subseteq K_\mathcal{S}^{J1}(x) = K_\mathcal{S}^{J2}(x)$.

Proof: From [19, Th. 2], we have $\dot{h}_i - L_x\|x - x_i\|_1 - L_u\|u - u_i\|_1 \leq \dot{h}_i - L_x^\top |x - x_i| - L_u^\top |u - u_i| \leq \dot{h}_i + \underline{J}_x \Delta x_i^+ - \overline{J}_x \Delta x_i^- + \underline{J}_u \Delta u_i^+ - \overline{J}_u \Delta u_i^-$, and thus the above result. ■

Further, note that the "choice" of the CBF-DDs corresponding to different continuity assumption leads to different classes of optimization problems with different corresponding computational complexity (see, e.g., [23] and references therein, for a full discussion on complexity analysis for each class of problems). This is an important point since one of the biggest appeals of the standard CBF is its simplicity with the need to only solve a quadratic program (QP) at run time. By contrast, the data-driven approach in [16] involves a second-order cone program (SOCP) that may not be as amenable for real-time implementation. In particular, if $p = 2$ for Case I, then the associated optimization-based controller with CBF-DD-L is a mixed-integer quadratically constrained quadratic program (MIQCQP) and if $p = 1$ or $p = \infty$, it becomes a mixed-integer quadratic program (MIQP). The reason for the presence of integers is because union operations are often encoded using integer constraints. Similarly, in Case II, the associated optimization problem with CBF-DD-CL is an MIQP, where the absolute value operators incur additional integer constraints. Finally, in Case III, the optimization-based controllers with CBF-DD-J1 and CBF-DD-J2 are both MIQPs but the variant with DDJ1 incurs additional integer constraints associated with the maximum operator in the definition of Δu_i^+ and Δu_i^-.

Mixed-integer convex optimization problems can be quite computationally expensive; hence, in the next section, we will describe complexity reduction techniques, including one to reduce the mixed-integer convex problems to a (small) set of convex optimization problems by specifically targeting the union operator that incurs integer constraints. With this additional step, it can then be shown that using CBF-DD-L with

$p = 2$ only involves quadratically constrained quadratic programs (QCQPs) and using CBF-DD-L with $p = 1$ or $p = \infty$ reduces to a QP. Further, using CBF-DD-J1 also still involves MIQPs due to maximum operators, but with CBF-DD-J2, it only involves QPs. Finally, the use of CBF-DD-CL still leads to MIQPs due to the presence of absolute value operators, but since they can be seen as a special case of Case III (including CBF-DD-J2) with $\overline{J} = -\underline{J} = L$, it can be solved using QPs similar to CBF-DD-J2.

Note that the observation that CBF-DD-J2 only involves QPs coupled with Proposition 2 implies that the use of CBF-DD-J2 is superior to using CBF-DD-J1 in terms of computational complexity and they are equivalent in terms of the obtained solution. Moreover, the use of CBF-DD-J2, CBF-DD-CL and CBF-DD-L with $p = 1$ or $p = \infty$ that only involves QPs potentially also require less computational time than the approach in [16] that involves SOCPs.

Remark 2: The Lipschitz constants/Jacobian bounds are not required to be minimal or tight, but smaller/tighter Lipschitz constants/Jacobian bounds can improve the performance of robust data-driven CBF methods. Moreover, if unknown, the Lipschitz constants/Jacobian bounds can be estimated from data with high probability from a sufficiently large data set by leveraging probably approximately correct (PAC) learning theory. More details on the estimation approach and the application of PAC learning can be found in [18], [19], [20].

Remark 3: For a continuously differentiable robust control Lyapunov function V for the unknown control affine system (1), if $\dot{V}(x, u)$ is as assumed in Assumption 2 and a data set $\{(\dot{V}(x_i, u_i), x_i, u_i)\}_1^N$ with $x_i \in \mathcal{X}$, $i = 1, \dots, N$ is available, similar techniques to the ones proposed in this letter can also be applied to ensure robust stabilization using robust data-driven control Lyapunov functions.

IV. Complexity Reduction Strategies

Solving the resulting optimization-based controllers using CBF-DDs can often be computationally expensive; hence, this section will introduce two strategies that can significantly reduce the complexity of the resulting optimization problems.

A. Simplified Implementation

First, we propose a method to break down the optimization problem in (3) with data-driven CBFs into multiple subproblems that can be solved in parallel (if desired) and to recover the optimal control input from the subproblems' solutions.

Specifically, we consider the following subproblems.

Definition 6 (CBF-DD-sub): Consider a data point $(\dot{h}_i, x_i, u_i)$ in the data set $\mathcal{D} = \{(\dot{h}_i, x_i, u_i)\}_1^N$ that satisfies Assumption 1, a continuously differentiable function $h : \mathcal{X} \to \mathbb{R}$, a class $\mathcal{K}_\infty$ function $\alpha(\cdot)$ and any controller $u = k(x)$, we can find the u_i that is closest to the u in the safe input set $\mathcal{U}_i(x)$ by solving the following optimization problem:

$$u_i^*(x) = \arg\min_u \ \frac{1}{2}\|u - k(x)\|_2^2 \qquad (11a)$$

$$s.t. \ \ u \in \mathcal{U}_i^\phi(x). \qquad (11b)$$

with $\mathcal{U}_i^\phi(x)$, $\phi \in \{L, CL, J1, J2\}$ based on the given continuity case in Assumption 2.

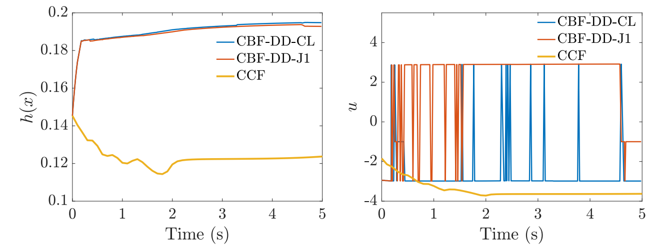

Fig. 1. CBF $h(x)$ (left) and input u (right) trajectories for the inverted pendulum example. The CCF result is obtaining by applying the approach in [16].

Since all u_i^* are independent to each other, they can be computed in parallel. After obtaining u_i^* for $i = 1, \dots, N$ by solving (11), the $u(x)$ can be selected by:

$$u(x) = \arg\min_{u \in \{u_1^*(x), \dots, u_N^*(x)\}} \frac{1}{2}\|u - k(x)\|_2^2. \qquad (12)$$

In practice, (11) need not be feasible for all data points and u_i^* can be set to be ∞ when $\mathcal{U}_i^\star(x) = \emptyset$.

B. Downsampling Method

Finally, we can rely on the following proposition to further reduce the complexity of the problem.

Proposition 4 (Monotonicity): The safe input sets in Definitions 2–5 satisfy monotonicity, in the sense that given two data sets $\mathcal{D}$ and $\mathcal{D}'$ and their corresponding safe input sets $K_\mathcal{S}(x)$ and $K'_\mathcal{S}(x)$, $\mathcal{D}' \subseteq \mathcal{D}$ implies that $K'_\mathcal{S}(x) \subseteq K_\mathcal{S}(x)$.

Proof: Let I and I' be the set of indices corresponding to data points included in $\mathcal{D}$ and $\mathcal{D}'$. Then $\mathcal{D}' \subseteq \mathcal{D} \Rightarrow I' \subseteq I \Rightarrow K'_\mathcal{S}(x) = \{u \in \cup_{i \in I'} \mathcal{U}_i\} \subseteq K_\mathcal{S}(x) = \{u \in \cup_{i \in I} \mathcal{U}_i\}$. ■

By Proposition 4, we can apply downsampling methods, e.g., kNN and clustering method (cf. [20] for examples), to reduce the number of constraints in the original optimization problem (3) with (3b) replaced by CBF-DD-ϕ, $\phi \in \{L, CL, J1, J2\}$ (cf. Theorem 2) or the number of corresponding subproblems (11), thus reducing computation without losing robust safety of the CBF-DDs.

V. Simulation Examples

We evaluate the proposed methods on two examples that are implemented in MATLAB and solved using Gurobi on a 2.2 GHz Intel Core i7 CPU with 16 GB RAM.

A. Inverted Pendulum

We first consider an inverted pendulum whose true system dynamics (with $g = 10$, $l = 0.7$, $m = 0.7$), CBF ($h(x) = 0.2 - x^\top P x$ with $P = \begin{bmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{bmatrix}$) and comparison function ($\alpha(r) = 0.268r$) are the same as described in [16]. We also adopt the same controller (with Lyapunov function $V(x) = x^\top P x$) and data set (with $\theta \in [0, 1]$, $\dot{\theta} \in [-0.25, 0.25]$ with grid sizes 0.025). The system is simulated with an initial condition of $x_0 = [0.1, 0.1]^\top$ for 10 seconds with control inputs given at 100 Hz. For the CBF-DD-L method, $p = 1$.

From Figure 1, we observe that safety is maintained with both CBF-DD-CL and CBF-DD-J2 (while the states are stabilized), albeit with very slight input chattering that presumably is due to the non-smooth nature (due to max operators) of CBF-DD conditions. In comparison, the approach in [16]

TABLE I
CPU TIME COMPARISON FOR DIFFERENT METHODS

Method	L	CL	J1	J2	SOCP [16]
CPU time (s)	3029	3054	3154	2724	2.84×10^5

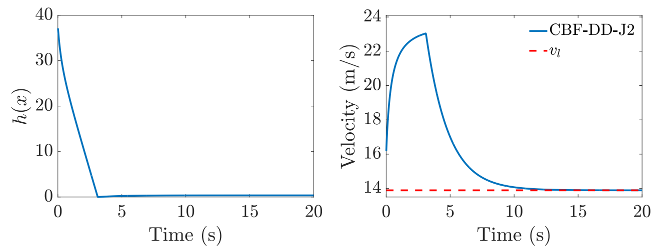

Fig. 2. CBF $h(x)$ (left) and velocity v (right) trajectories for the ACC-TTC example with fixed $v_l = 13.89$.

appeared to allow the states to go closer to the safety boundaries. Further, we compared the computation/CPU time of the various proposed methods for the entire trajectories. From Table I, solving CBF-DD-J2 (involving QP subproblems) was faster than other CBF-DDs that solve MIQP subproblems, and they are all approximately 100 times faster than the SOCP-based approach in [16].

B. Adaptive Cruise Control With Safe Time-to-Collision

We consider the adaptive cruise control of self-driving cars with similar dynamics as [5] with system parameters from [2], but with a given lead vehicle velocity $v_l = 13.89$ and a CBF, $h(D, v) = D - T_c(v - v_l)$, which is derived from *time-to-collision* (TTC) safety constraint with $T_c = 1.8$ as the time-to-collision. Moreover, we consider a control Lyapunov function $V = (v - v_d)^2$ to enable tracking of a desired velocity $v_d = 40$, when possible. The data set is obtained by gridding the state and input space (i.e., $v \in [11.5, 24]$, $D \in [0, 100]$ and $u \in [-10, 10]$) with 10 grids per dimension.

We tested both the CBF-DD-CL and CBF-DD-J2 methods using an initial condition of $x_0 = [16, 50]^\top$. CBF-DD-CL (that is more conservative than CBF-DD-J2, cf. Proposition 3) was infeasible, thus, incapable of guaranteeing safety, with the same amount of data (hence, not depicted), while CBF-DD-J2 was able to guarantee safety, as shown in Figure 2.

VI. CONCLUSION

In this letter, we proposed robust data-driven control barrier functions to guarantee robust safety of unknown control affine systems under several continuity assumptions. We leveraged non-parametric data-driven approaches to provide guaranteed lower bounds of the robust controlled invariance condition despite worst-case generalization errors. By incorporating the corresponding safe input set into an optimization-based controller, the safety of the unknown system can be ensured. We additionally presented several complexity reduction approaches (i.e., parallel subproblems and downsampling strategies) to improve computational performance. As expected, our proposed approaches are observed to guarantee robust safety in all our simulation examples, including for adaptive cruise control of (semi-)autonomous vehicles.

REFERENCES

[1] S. Prajna, A. Jadbabaie, and G. J. Pappas, "A framework for worst-case and stochastic safety verification using barrier certificates," *IEEE Trans. Autom. Control*, vol. 52, no. 8, pp. 1415–1428, Aug. 2007.
[2] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 3861–3876, Aug. 2017.
[3] X. Xu, P. Tabuada, J. W. Grizzle, and A. D. Ames, "Robustness of control barrier functions for safety critical control," *IFAC-PapersOnLine*, vol. 48, no. 27, pp. 54–61, 2015.
[4] S. Kolathaya and A. D. Ames, "Input-to-state safety with control barrier functions," *IEEE Control Syst. Lett.*, vol. 3, no. 1, pp. 108–113, Jan. 2019.
[5] A. J. Taylor and A. D. Ames, "Adaptive safety with control barrier functions," in *Proc. Amer. Control Conf.*, 2020, pp. 1399–1405.
[6] B. T. Lopez, J.-J. E. Slotine, and J. P. How, "Robust adaptive control barrier functions: An adaptive and data-driven approach to safety," *IEEE Control Syst. Lett.*, vol. 5, no. 3, pp. 1031–1036, Jul. 2021.
[7] M. Black, E. Arabi, and D. Panagou, "A fixed-time stable adaptation law for safety-critical control under parametric uncertainty," in *Proc. Eur. Control Conf. (ECC)*, 2021, pp. 1323–1328.
[8] M. Jankovic, "Robust control barrier functions for constrained stabilization of nonlinear systems," *Automatica*, vol. 96, pp. 359–367, Oct. 2018.
[9] K. Garg and D. Panagou, "Robust control barrier and control Lyapunov functions with fixed-time convergence guarantees," in *Proc. Amer. Control Conf. (ACC)*, 2021, pp. 2292–2297.
[10] J. Choi, F. Castaneda, C. J. Tomlin, and K. Sreenath, "Reinforcement learning for safety-critical control under model uncertainty, using control Lyapunov functions and control barrier functions," in *Proc. Robot. Sci. Syst. (RSS)*, 2020, pp. 1–9.
[11] A. J. Taylor, A. Singletary, Y. Yue, and A. D. Ames, "Learning for safety-critical control with control barrier functions," in *Proc. Learn. Dyn. Control*, 2020, pp. 708–717.
[12] A. Nejati, B. Zhong, M. Caccamo, and M. Zamani, "Data-driven controller synthesis of unknown nonlinear polynomial systems via control barrier certificates," in *Proc. Learn. Dyn. Control Conf.*, 2022, pp. 763–776.
[13] P. Jagtap, G. J. Pappas, and M. Zamani, "Control barrier functions for unknown nonlinear systems using Gaussian processes," in *Proc. IEEE Conf. Decis. Control (CDC)*, 2020, pp. 3699–3704.
[14] V. Dhiman, M. J. Khojasteh, M. Franceschetti, and N. Atanasov, "Control barriers in Bayesian learning of system dynamics," *IEEE Trans. Autom. Control*, vol. 68, no. 1, pp. 214–229, Jan. 2023.
[15] 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. IEEE Conf. Decis. Control (CDC)*, 2021, pp. 6762–6769.
[16] A. J. Taylor, V. D. Dorobantu, S. Dean, B. Recht, Y. Yue, and A. D. Ames, "Towards robust data-driven control synthesis for nonlinear systems with actuation uncertainty," in *Proc. IEEE Conf. Decis. Control (CDC)*, 2021, pp. 6469–6476.
[17] Z. Zabinsky, R. Smith, and B. Kristinsdottir, "Optimal estimation of univariate black-box Lipschitz functions with upper and lower error bounds," *Comput. Oper. Res.*, vol. 30, no. 10, pp. 1539–1553, 2003.
[18] J. Calliess, "Conservative decision-making and inference in uncertain dynamical systems," Ph.D. dissertation, Dept. Eng. Sci., Univ. Oxford, Oxford, U.K., 2014.
[19] Z. Jin, M. Khajenejad, and S. Z. Yong, "Data-driven abstraction and model invalidation for unknown systems with bounded Jacobians," *IEEE Control Syst. Lett.*, vol. 6, pp. 3421–3426, 2022.
[20] Z. Jin, M. Khajenejad, and S. Z. Yong, "Data-driven model invalidation for unknown Lipschitz continuous systems via abstraction," in *Proc. Amer. Control Conf. (ACC)*, 2020, pp. 2975–2980.
[21] D. Efimov, T. Raïssi, S. Chebotarev, and A. Zolghadri, "Interval state observer for nonlinear time varying systems," *Automatica*, vol. 49, no. 1, pp. 200–205, 2013.
[22] M. Khajenejad and S. Z. Yong, "Simultaneous input and state interval observers for nonlinear systems with rank-deficient direct feedthrough," in *Proc. Eur. Control Conf.*, 2021, pp. 2311–2316.
[23] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: SIAM, 2001.