Control Barrier Functions and Input-to-State Safety with Application to Automated Vehicles

Anil Alan[1], Andrew J. Taylor[2], Chaozhe R. He[1,3], Aaron D. Ames[2], and Gábor Orosz[1,4]

Abstract—**Balancing safety and performance is one of the predominant challenges in modern control system design. Moreover, it is crucial to robustly ensure safety without inducing unnecessary conservativeness that degrades performance. In this work we present a constructive approach for safety-critical control synthesis via *Control Barrier Functions* (CBF). By filtering a hand-designed controller via a CBF, we are able to attain performant behavior while providing rigorous guarantees of safety. In the face of disturbances, robust safety and performance are simultaneously achieved through the notion of *Input-to-State Safety* (ISSf). We take a tutorial approach by developing the CBF-design methodology in parallel with an inverted pendulum example, making the challenges and sensitivities in the design process concrete. To establish the capability of the proposed approach, we consider the practical setting of safety-critical design via CBFs for a *connected automated vehicle* (CAV) in the form of a class-8 truck without a trailer. Through experimentation we see the impact of unmodeled disturbances in the truck's actuation system on the safety guarantees provided by CBFs. We characterize these disturbances and using ISSf, produce a robust controller that achieves safety without conceding performance. We evaluate our design both in simulation, and for the first time on an automotive system, experimentally.**

Index Terms—**Robust safety-critical control, control barrier functions, input-to-state safety, connected automated vehicles.**

I. Introduction

Safety is an ever more pressing requirement for modern control systems as they are deployed into increasingly complex real-world environments. Simultaneously, meeting performance requirements is a major driving factor in control system design. As these two objectives may naturally oppose each other, it is necessary to consider an *active* approach for enforcing safety that impacts performance only when it is critical for the safety of the system [1], [2]. *Control Barrier Functions* (CBFs) have been demonstrated to be a powerful tool for constructively synthesizing controllers that yield strong performance and intervene only when safety is at risk of being compromised [3]–[5]. The utility of CBFs has been confirmed by their experimental application on real-world control systems, including mobile robots [1], [6], robotic swarms [7], autonomous aerial vehicles [8], robotic arms [9], robotic manipulators [10], quadrupedal robots [11], and bipedal robots [12], as well as simulation results on automotive systems [3], autonomous naval vehicles [13], and spacecraft [14]. The variety in this collection of results indicates that CBFs capture fundamental concepts underlying the notion of safety, irrespective of a specific domain, and suggests that CBFs are a valuable tool to consider in the process of modern control system design.

One of the appealing features of the CBF-based methodology for safety-critical control synthesis is the relatively intuitive nature of the theoretical safety guarantees they endow a system with. The study of *set invariance*, or the state of a system remaining within a prescribed set, has long been of interest in the study of dynamic systems [15] and control [16]. The foundational work in [17] proposed the notion of a *barrier function* as a tool for checking the invariance of a set given a model of the system dynamics. In simple terms[1], a barrier function takes positive values for states inside a set, and is zero on the boundary of the set. If the time derivative of the barrier function is positive on the boundary of the set, the value of barrier must grow and the system thus must remain in the set. This idea was quickly adapted to the context of control synthesis, yielding CBFs and a means to constructively achieve set invariance. Synthesis was first proposed through structured controllers [19], but was later expanded using convex optimization to produce *safety-filters* that minimally modify a hand-designed controller to ensure safety [1], [3], [4]. The combination of intuitive theoretical concepts with relatively simple control synthesis techniques promoted rapid development of CBFs, including formulations for higher-order systems [20]–[22] and discrete-time systems [23], as well as constructive tools for synthesizing CBFs [24]–[26] and methods for sets with complex geometries [27], [28].

Inherent in the theoretical safety guarantees provided by CBFs is a dependence on the model of the system dynamics, thus raising subsequent questions of robustness. Resulting works have explored robustness to disturbances [29]–[35], measurement errors [36], unmodeled dynamics [37], and sector-bounded uncertainties [38]. The early work in [29] noticed a robustness to disturbances inherent in CBFs, which drawing inspiration from the notion of Input-to-State Stability frequently seen when considering robust stabilization of nonlinear systems [39], was formalized into the idea of *Input-to-State Safety* (ISSf) in [32]. Instead of trying to keep a specific

This research is supported in part by the National Science Foundation, CPS Award #1932091.

[1] A. Alan, C. R. He, and G. Orosz are with the Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, USA {anilalan, hchaozhe, orosz}@umich.edu

[2] A. J. Taylor and A. D. Ames are with the Department of Computing & Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA {ajtaylor, ames}@caltech.edu

[3] C. R. He is also with Plus.ai Inc., Cupertino, CA 95014, USA chaozhe.he@plus.ai

[4] G. Orosz is also with the Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA

[1] We recommend the reader to [18] for a comprehensive mathematical study of the connections between barrier functions and set invariance.

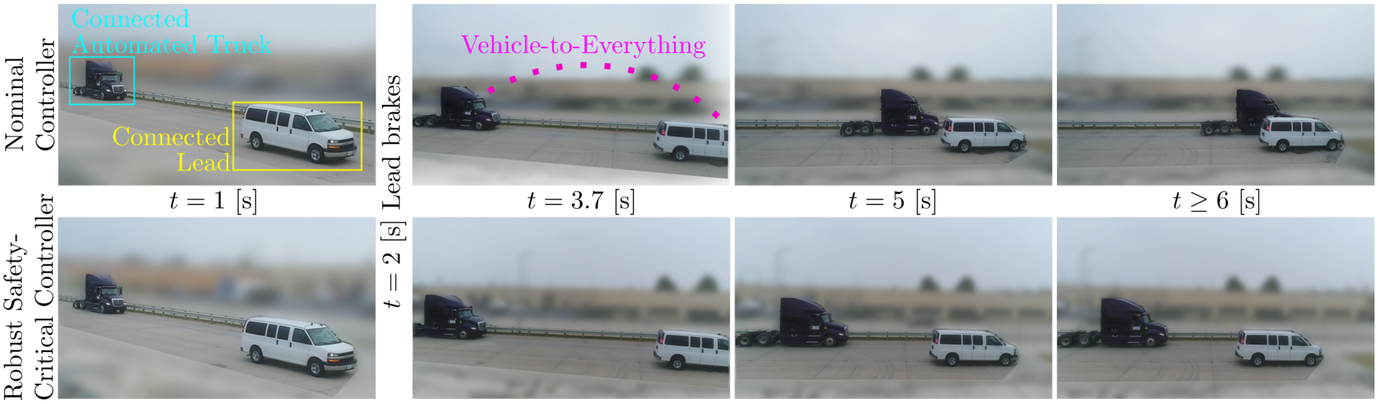

Fig. 1. Experimental configuration for heavy-duty CAV problem. (Top) Controller design without robustifying element yields safety violation and collision. (Bottom) Robust safety-critical controller ensures CAV brakes early and aggressively enough to maintain safe distance.

set invariant in the presence of disturbances as in [30], [31], which may induce conservativeness and degrade the performance of a controller, ISSf quantifies how the set kept invariant grows in the presence of disturbances. Moreover, it provides a simple modification for CBF-based controllers to control this growth, which was extended in [34] to permit greater performance while maintaining meaningful safety guarantees. As we demonstrate in this work, this paradigm for robust safety naturally lends itself to the design-test-redesign process, as the growth of the invariant set can be tuned to satisfy safety requirements while meeting performance metrics.

Despite the fact that CBFs were initially presented as a tool for safety-critical control synthesis for automotive systems [3], [29], they have yet to be experimentally realized on them. A primary challenge in using CBFs to ensure safety for a complex system such as a full-scale *connected automated vehicle* (CAV) lies in addressing discrepancies between the system model and the real-world system. In the context of a heavy-duty CAV, a significant portion of these discrepancies arise due to simplified models of the complex interactions within the CAV braking elements [40], and manifest as disturbances in the input applied to the system. Accounting for these complicated interactions in the controller design may greatly increase the intricacy of the resulting controller, but completely ignoring them may yield safety violations under critical conditions such as a harsh brake from a preceding vehicle as seen in Fig. 1 (top). Thus, balancing the complexity of the model used in design with the need to satisfy safety requirements is a challenging yet appropriate setting to deploy robust CBF-based control design.

There are two main contributions in this paper. The first is a tutorial presentation of a robust safety-critical design methodology using CBFs and ISSf. Concepts are introduced in parallel with an inverted pendulum example, thus providing a concrete context for readers to quickly establish an understanding of the relevant details in safety-critical control synthesis. We provide an appropriate level of theoretical discussion to clearly state the theoretical safety guarantees achieved with this control paradigm, but focus predominantly on the practical challenges and trade-offs encountered in safety-critical control design. Compared to the original works [3], [4] and overview work [5] on CBFs, we believe that this presentation provides a more approachable introduction to the topic of safety-critical control synthesis for practitioners. Moreover, all details necessary to exactly recreate the simulation results in the inverted pendulum example are provided.

The second contribution of this work is a more practical application of the presented safety-critical control design methodology that considers a heavy-duty CAV, seen in Figure 1. We highlight the entire process of safety-critical control design including system modeling, specification of safety requirements via a CBF, nominal performance-based controller design, simulation, and experimental testing on a full-scale automated class-8 tractor. The impacts of unmodeled disturbances seen in experimental results are quantified and used to robustify the safety-critical controller, which is subsequently implemented in simulation and experimentally. We believe that combined tutorial presentation and the proposed design-test-redesign process on a challenging real-world system is precisely the approach necessary to advance CBF-based control design from the academic setting to a tool useful for the practicing control engineer.

The organization of this paper is as follows. In Section II we present the safety-critical control problem, review CBFs, and explore how a nominal controller may be modified via CBFs to endow a system with theoretical safety guarantees. In Section III we introduce disturbances into the input to the system, and explore how these impact theoretical safety guarantees through the lens of ISSf. Moreover, we present a simple framework for robustly modifying a CBF and the resulting controller design to provide a measure of control over how these safety guarantees degrade. In Section IV the connected automated vehicle problem is presented considering an automated heavy-duty vehicle. A CBF specified to encode safety for the CAV and a hand-designed nominal controller are incorporated into a safety-critical controller that is evaluated in simulation and verified to ensure safety. In Section V we deploy the controller experimentally, and see how unmodeled disturbances lead to degradation of safety guarantees. We characterize these disturbances and robustify the controller design, and lastly verify the ability of this controller to meet safety requirements both in simulation and experiments.

II. Safety-Critical Control

In this section we provide a review of safety and Control Barrier Functions (CBFs). These definitions will be used in the formulation of the safety-critical control problem. To make these concepts more concrete, we apply them to an inverted pendulum system.

A. Control Barrier Functions

Consider the nonlinear control affine system:

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

with state $\mathbf{x} \in \mathbb{R}^n$, input $\mathbf{u} \in \mathbb{R}^m$, and continuous functions $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^{n\times m}$. Systems described by such equations often appear in robotics, aerospace, power electronics, and automotive systems.

Example 1. Consider a control system for an inverted pendulum as depicted in Fig. 2, and described by the model:

$$\frac{\mathrm{d}}{\mathrm{dt}}\begin{bmatrix}\theta \\ \dot{\theta}\end{bmatrix} = \underbrace{\begin{bmatrix}\dot{\theta} \\ \frac{g}{l}\sin\theta\end{bmatrix}}_{\mathbf{f}(\mathbf{x})} + \underbrace{\begin{bmatrix}0 \\ \frac{1}{ml^2}\end{bmatrix}}_{\mathbf{g}(\mathbf{x})} u, \tag{2}$$

with pendulum angle $\theta \in \mathbb{R}$ and angular velocity $\dot{\theta} \in \mathbb{R}$ defining the state $\mathbf{x} = [\theta, \dot{\theta}]^\top$, and parameters given by the mass m, length l, and gravitational acceleration constant g. In this example we will use the parameter values $m = 2$ [kg], $l = 1$ [m] and $g = 10$ [m/s^2]. The single input $u \in \mathbb{R}$ is a torque applied at the pendulum base.

The input $\mathbf{u}$ is often specified via a state-feedback controller $\mathbf{k} : \mathbb{R}^n \to \mathbb{R}^m$, yielding the closed-loop system dynamics:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{k}(\mathbf{x}). \tag{3}$$

We assume that for any initial condition $\mathbf{x}_0 \triangleq \mathbf{x}(0) \in \mathbb{R}^n$, there exists a unique solution $\mathbf{x}(t)$ to (3) for $t \geq 0$, such that the system is forward complete [41]. The notion of safety is formalized by specifying a *safe set* in the state space which the state of the system must remain in to be considered safe. This can be utilized in many practical applications such as distance-keeping [3], lane-keeping [6], and collision avoidance [7] of automated vehicles. In particular, consider a set $\mathcal{C} \subset \mathbb{R}^n$ defined as the 0-superlevel set of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$, yielding:

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}, \tag{4}$$

$$\partial\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = 0\}, \tag{5}$$

$$\mathrm{Int}(\mathcal{C}) = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) > 0\}, \tag{6}$$

where $\partial\mathcal{C}$ and $\mathrm{Int}(\mathcal{C})$ are the *boundary* and *interior*, respectively, of the set $\mathcal{C}$. We refer to $\mathcal{C}$ as the *safe set*. This construction motivates the following definitions of forward invariance and safety:

Definition 1 (*Forward Invariance & Safety*)**.** A set $\mathcal{C} \subset \mathbb{R}^n$ is *forward invariant* if for every $\mathbf{x}_0 \in \mathcal{C}$, the solution $\mathbf{x}(t)$ to (3) satisfies $\mathbf{x}(t) \in \mathcal{C}$ for all $t \geq 0$. The system (3) is *safe* with respect to the set $\mathcal{C}$ if the set $\mathcal{C}$ is forward invariant.

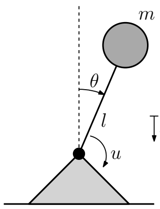

Fig. 2. Schematic of an inverted pendulum control system.

Example 2. A set $\mathcal{C}$ that we wish to keep safe for the inverted pendulum that restricts the angular position and velocity is given by the 0-superlevel set of the function:

$$h(\theta, \dot{\theta}) = 1 - \frac{\theta^2}{a^2} - \frac{\dot{\theta}^2}{b^2} - \frac{\theta\dot{\theta}}{ab}, \tag{7}$$

with parameters $a, b > 0$. In this example we will use the parameter values $a = 0.25$ [rad] and $b = 0.5$ [rad/s]. The resulting set:

$$\mathcal{C} = \left\{ \begin{bmatrix}\theta \\ \dot{\theta}\end{bmatrix} \in \mathbb{R}^2 \,\middle|\, 1 - \frac{\theta^2}{a^2} - \frac{\dot{\theta}^2}{b^2} - \frac{\theta\dot{\theta}}{ab} \geq 0 \right\}, \tag{8}$$

is an ellipse as depicted in Fig. 3 by a gold region.

Before defining Control Barrier Functions, we review the following definitions [5], [42]. We denote a continuous function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ as *class* $\mathcal{K}_\infty$ ($\alpha \in \mathcal{K}_\infty$) if $\alpha(0) = 0$, α is strictly increasing and $\lim_{r\to\infty} \alpha(r) = \infty$. As an example, any function in the form $\alpha(r) = r^c$ where $c > 0$ is class $\mathcal{K}_\infty$. Note that differentiability is not required for a class $\mathcal{K}_\infty$ function. Similarly, a continuous function $\alpha : \mathbb{R} \to \mathbb{R}$ is said to belong to *extended class* $\mathcal{K}_\infty$ ($\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$) if $\alpha(0) = 0$, α is strictly increasing, and $\lim_{r\to\infty} \alpha(r) = \infty$ and $\lim_{r\to-\infty} \alpha(r) = -\infty$. The previous example of $\alpha(r) = r^c$ is class $\mathcal{K}_\infty^{\mathrm{e}}$ for $c = 1, 3, 5, \ldots$ The inverses of class $\mathcal{K}_\infty$ and class $\mathcal{K}_\infty^{\mathrm{e}}$ functions belong to class $\mathcal{K}_\infty$ and class $\mathcal{K}_\infty^{\mathrm{e}}$, respectively. Examples of these functions and their inverses are depicted in Fig. 4. We may use these functions to define Control Barrier Functions.

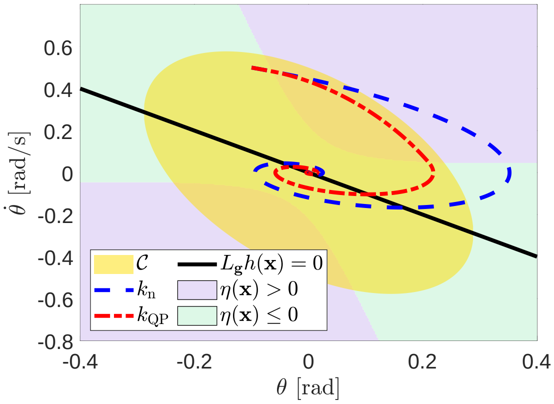

Fig. 3. Simulation results for the inverted pendulum system. The gold ellipse is the safe set $\mathcal{C}$ as defined in (8). The black line is the set where $L_{\mathbf{g}}h(\mathbf{x}) = 0$ as defined in (11). The dashed blue line is the trajectory of the system evolving under k_{n} as defined in (16), which leaves the safe set $\mathcal{C}$. The green and purple regions indicate where the controller k_{n} meets and fails to meet the CBF condition, respectively. The dashed red line is the trajectory of the system evolving under k_{QP} as defined in (21)-(22), which remains inside the safe set $\mathcal{C}$.

Definition 2 (*Control Barrier Function*, [4])**.** Let $\mathcal{C} \subset \mathbb{R}^n$ be the 0-superlevel set of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$. The function h is a *Control Barrier Function* (CBF) for the system (1) on $\mathcal{C}$ if there exists $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$ such that for all $\mathbf{x} \in \mathbb{R}^n$:

$$\sup_{\mathbf{u}\in\mathbb{R}^m} \Bigg[\overbrace{\underbrace{\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x})\mathbf{f}(\mathbf{x})}_{L_{\mathbf{f}}h(\mathbf{x})} + \underbrace{\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x})\mathbf{g}(\mathbf{x})}_{L_{\mathbf{g}}h(\mathbf{x})}\mathbf{u}}^{\dot{h}(\mathbf{x},\mathbf{u})} \Bigg] > -\alpha(h(\mathbf{x})). \tag{9}$$

An equivalent way to express (9) is given in [30] as:

$$L_{\mathbf{g}}h(\mathbf{x}) = \mathbf{0} \implies L_{\mathbf{f}}h(\mathbf{x}) + \alpha(h(\mathbf{x})) > 0. \tag{10}$$

This expression is often an easier condition to evaluate in certifying that a given function is a CBF.

Example 3. The function h given as in (7) is a CBF for the inverted pendulum system (2) on $\mathcal{C}$. To see this, consider a function $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$ defined as $\alpha(r) = \alpha_{\mathrm{c}} r$ with $\alpha_{\mathrm{c}} > 0$ satisfying $\alpha_{\mathrm{c}} \leq b/a$. In this example we will take the parameter value $\alpha_{\mathrm{c}} = 0.2$ [1/s]. Checking the CBF condition defined in (10), we see that:

$$L_{\mathbf{g}}h(\theta_0, \dot{\theta}_0) = 0 \implies \dot{\theta}_0 = -\frac{b}{2a}\theta_0. \tag{11}$$

This equation defines a line as depicted in Fig. 3 by a solid black line. We have that on this line:

$$L_{\mathbf{f}}h(\theta_0, \dot{\theta}_0) + \alpha(h(\theta_0, \dot{\theta}_0)) = \alpha_{\mathrm{c}} + \frac{3}{4a^2}\left(\frac{b}{a} - \alpha_{\mathrm{c}}\right)\theta_0^2 > 0,$$

such that the condition (10) is met for our choice of α_{c}.

We note that if we consider a set, denoted by $\tilde{\mathcal{C}}$ and defined as the 0-superlevel set of a function $\tilde{h} : \mathbb{R}^2 \to \mathbb{R}$ given by:

$$\tilde{h}(\theta, \dot{\theta}) = 1 - \frac{\theta^2}{a^2} - \frac{\dot{\theta}^2}{b^2}, \tag{12}$$

which does not include the term $\theta\dot{\theta}/ab$, then the function $\tilde{h}$ is not a CBF for the system (2) on $\tilde{\mathcal{C}}$. To see this, note that:

$$L_{\mathbf{g}}\tilde{h}(\theta_0, \dot{\theta}_0) = 0 \implies \dot{\theta}_0 = 0. \tag{13}$$

In turn, we have that for any $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$:

$$L_{\mathbf{f}}\tilde{h}(\theta_0, \dot{\theta}_0) + \alpha(\tilde{h}(\theta_0, \dot{\theta}_0)) = \alpha\left(1 - \theta_0^2/a^2\right). \tag{14}$$

The condition (10) is not satisfied for $|\theta_0| \geq a$ (including $|\theta_0| = a$, which would be in $\partial\tilde{\mathcal{C}}$ and thus in the safe set). Thus it is important to choose the safe set and design the CBF to be compatible with the system dynamics, eliminating points where the CBF condition is not met.

Given a CBF h for (1) on $\mathcal{C}$ and a corresponding function $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$, we can consider the point-wise set of all control values that satisfy (9):

$$K_{\mathrm{CBF}}(\mathbf{x}) = \left\{ \mathbf{u} \in \mathbb{R}^m \;\middle|\; \dot{h}(\mathbf{x}, \mathbf{u}) \geq -\alpha(h(\mathbf{x})) \right\}. \tag{15}$$

One of the main theoretical result for CBFs relates controllers taking values in the set K_{CBF} to the safety of (3) on $\mathcal{C}$:

Theorem 1 ([4], [18])**.** *Let $\mathcal{C} \subset \mathbb{R}^n$ be the 0-superlevel set of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$. If h is a CBF for* (1) *on $\mathcal{C}$, then the set $K_{\mathrm{CBF}}(\mathbf{x})$ is non-empty for each $\mathbf{x} \in \mathbb{R}^n$, and for any continuous controller $\mathbf{k} : \mathbb{R}^n \to \mathbb{R}^m$ such that $\mathbf{k}(\mathbf{x}) \in K_{\mathrm{CBF}}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, the system* (3) *is safe with respect to the set $\mathcal{C}$.*

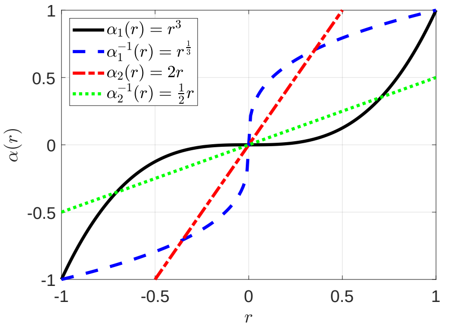

Fig. 4. Visualization of class $\mathcal{K}_\infty^{\mathrm{e}}$ functions and their inverses.

Proofs of Theorem 1 may be found in [4], [18]. We note the distinction of a strict inequality in the CBF condition in (9) and a non-strict inequality in (15). As studied in [30], satisfaction of the strict inequality in (9) and (10) is a property of the function h and the dynamics $\mathbf{f}$ and $\mathbf{g}$, but not does depend on a specific controller. This property is useful in establishing regularity properties of controllers synthesized with the CBF h. In particular, it imposes requirements on the function h when $L_{\mathbf{g}}h(\mathbf{x}) = \mathbf{0}$, as seen in the preceding example. In contrast, enforcing safety via Theorem 1 only require that the inputs specified by a given controller meet the non-strict inequality in (15) (such an input's existence is implied by the CBF condition in (9)).

B. Safety-Critical Controller

It is often possible to design a controller that achieves a desired degree of performance, but for which it is difficult to verify necessary safety requirements are met.

Example 4. Consider a continuous controller $k_{\mathrm{n}} : \mathbb{R}^2 \to \mathbb{R}$ for the inverted pendulum model (2) that stabilizes the pendulum to an upright position, given by the feedback linearization [43] or computed torque controller [44] of the form:

$$k_{\mathrm{n}}(\theta, \dot{\theta}) = ml^2\left(-\frac{g}{l}\sin\theta - K_{\mathrm{p}}\theta - K_{\mathrm{d}}\dot{\theta}\right), \tag{16}$$

with controller gains $K_{\mathrm{p}}, K_{\mathrm{d}} > 0$. This controller yields the closed-loop system:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\theta \\ \dot{\theta}\end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -K_{\mathrm{p}} & -K_{\mathrm{d}} \end{bmatrix}\begin{bmatrix}\theta \\ \dot{\theta}\end{bmatrix}, \tag{17}$$

such that the upright equilibrium $\mathbf{x}^* = \begin{bmatrix}0, 0\end{bmatrix}^\top$ is exponentially stable. In this example we will use the parameter values $K_{\mathrm{p}} = 0.6$ [1/s^2] and $K_{\mathrm{d}} = 0.6$ [1/s]. We use numerical integration to determine a solution trajectory from the initial condition $\mathbf{x}(0) = [-0.1,\ 0.5]^\top \in \mathcal{C}$. This trajectory is depicted in Fig. 3 by a dashed blue curve. Although the controller k_{n} stabilizes the system to the upright position, in doing so it causes the state of the system to leave the safe set $\mathcal{C}$.

CBFs provide a means for modifying a controller to ensure it explicitly enforces the safety of the system. Suppose that we have a continuous controller $\mathbf{k}_\mathrm{n} : \mathbb{R}^n \to \mathbb{R}^m$, referred to as the *nominal controller*, that does not necessarily ensure the closed-loop system (3) is safe with respect to the set $\mathcal{C}$, but achieves a desired degree of performance. Furthermore, suppose that we have a CBF h for (1) on $\mathcal{C}$ with corresponding function $\alpha \in \mathcal{K}_\infty^\mathrm{e}$. The goal of maintaining the performance of the nominal controller $\mathbf{k}_\mathrm{n}$ while ensuring the safety of the system (3) with respect to the set $\mathcal{C}$ motivates an optimization-based safety-critical controller $\mathbf{k}_\mathrm{QP} : \mathbb{R}^n \to \mathbb{R}^m$ defined as:

$$\begin{aligned} \mathbf{k}_\mathrm{QP}(\mathbf{x}) = \ & \underset{\mathbf{u}\in\mathbb{R}^m}{\operatorname{argmin}} && \frac{1}{2}\|\mathbf{u} - \mathbf{k}_\mathrm{n}(\mathbf{x})\|_2^2 \\ & \ \ \text{s.t.} && L_\mathbf{f}h(\mathbf{x}) + L_\mathbf{g}h(\mathbf{x})\mathbf{u} \geq -\alpha(h(\mathbf{x})). \end{aligned} \tag{18}$$

This controller takes the same value as the nominal controller if the nominal controller meets the requirements for safety specified by the CBF h, i.e., $\mathbf{k}_\mathrm{QP}(\mathbf{x}) = \mathbf{k}_\mathrm{n}(\mathbf{x})$ if $\mathbf{k}_\mathrm{n}(\mathbf{x}) \in K_\mathrm{CBF}(\mathbf{x})$. If the nominal controller does not meet the safety requirements, i.e., $\mathbf{k}_\mathrm{n}(\mathbf{x}) \notin K_\mathrm{CBF}(\mathbf{x})$, the input is chosen to meet the safety requirement with the smallest deviation from the value of $\mathbf{k}_\mathrm{n}$. The following theorem describes the feasibility of the optimization problem defining this controller, and provides a closed-form solution for the optimization problem:

Theorem 2. *Let $\mathcal{C}$ be the 0-superlevel set of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$, and let $\mathbf{k}_\mathrm{n} : \mathbb{R}^n \to \mathbb{R}^m$ be a continuous controller. If h is a CBF for* (1) *on the set $\mathcal{C}$ with corresponding function $\alpha \in \mathcal{K}_\infty^\mathrm{e}$, then the optimization problem in* (18) *is feasible for any $\mathbf{x} \in \mathbb{R}^n$ and has a closed-form solution given by:*

$$\mathbf{k}_\mathrm{QP}(\mathbf{x}) = \mathbf{k}_\mathrm{n}(\mathbf{x}) + \max\{0, \eta(\mathbf{x})\}\, L_\mathbf{g}h(\mathbf{x})^\top, \tag{19}$$

where the function $\eta : \mathbb{R}^n \to \mathbb{R}$ is defined as:

$$\eta(\mathbf{x}) = \begin{cases} -\frac{L_\mathbf{f}h(\mathbf{x}) + L_\mathbf{g}h(\mathbf{x})\mathbf{k}_\mathrm{n}(\mathbf{x}) + \alpha(h(\mathbf{x}))}{\|L_\mathbf{g}h(\mathbf{x})\|_2^2} & \text{if } L_\mathbf{g}h(\mathbf{x}) \neq \mathbf{0}, \\ 0 & \text{if } L_\mathbf{g}h(\mathbf{x}) = \mathbf{0}. \end{cases} \tag{20}$$

Furthermore, $\mathbf{k}_\mathrm{QP}$ is continuous and $\mathbf{k}_\mathrm{QP}(\mathbf{x}) \in K_\mathrm{CBF}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$.

A proof of this theorem is provided in the appendix. The function η only takes positive values ($\eta(\mathbf{x}) > 0$) when the nominal controller does not meet safety requirements:

$$L_\mathbf{f}h(\mathbf{x}) + L_\mathbf{g}h(\mathbf{x})\mathbf{k}_\mathrm{n}(\mathbf{x}) + \alpha(h(\mathbf{x})) < 0,$$

and thus the nominal controller $\mathbf{k}_\mathrm{n}$ is only modified when it does not satisfy safety requirements. The second case in the definition of the function η is presented to resolve the singularity that occurs at $L_\mathbf{g}h(\mathbf{x}) = \mathbf{0}$ when the closed-form solution (19) is implemented. We note that, as stated in Theorem 2, the controller $\mathbf{k}_\mathrm{QP}$ is continuous, and thus, this singularity does not produce a large jump in the input. It may even be ignored if the controller is implemented as the optimization problem in (18) and numerically solved.

Remark 1. For a single input ($m = 1$), if $L_\mathbf{g}h(\mathbf{x}) > 0$ for a particular $\mathbf{x} \in \mathbb{R}^n$, the controller (19) can be expressed as:

$$k_\mathrm{QP}(\mathbf{x}) = \max\left\{k_\mathrm{n}(\mathbf{x}), -\frac{L_\mathbf{f}h(\mathbf{x}) + \alpha(h(\mathbf{x}))}{L_\mathbf{g}h(\mathbf{x})}\right\}. \tag{21}$$

Similarly, if $L_\mathbf{g}h(\mathbf{x}) < 0$ for a particular $\mathbf{x} \in \mathbb{R}^n$, the controller (19) reduces to:

$$k_\mathrm{QP}(\mathbf{x}) = \min\left\{k_\mathrm{n}(\mathbf{x}), -\frac{L_\mathbf{f}h(\mathbf{x}) + \alpha(h(\mathbf{x}))}{L_\mathbf{g}h(\mathbf{x})}\right\}. \tag{22}$$

These controllers can be switched between based on the sign of $L_\mathbf{g}h(\mathbf{x})$, with $k_\mathrm{QP}(\mathbf{x}) = k_\mathrm{n}(\mathbf{x})$ when $L_\mathbf{g}h(\mathbf{x}) = 0$.

Example 5. We deploy the switching controller $k_\mathrm{QP} : \mathbb{R}^2 \to \mathbb{R}$ defined in (21)-(22) for the inverted pendulum system using the nominal controller $k_\mathrm{n} : \mathbb{R}^2 \to \mathbb{R}$ defined in (16). We use numerical integration to determine a solution trajectory from the initial condition $\mathbf{x}(0) = [-0.1,\ 0.5]^\top \in \mathcal{C}$. This trajectory is depicted in Fig. 3 by a dashed red curve. We see that the controller k_QP ensures that the solution trajectory remains within the safe set $\mathcal{C}$ by deviating from the nominal controller in the purple region as specified by (21)-(22).

III. Robustness to Disturbance

A challenge frequently encountered when deploying model-based controllers onto real-world systems is a mismatch between the commanded input and the input actually received by the system. This mismatch can arise due to actuator dynamics, actuator delays, input quantization, input saturation, or noise. In the case when a state feedback controller $\mathbf{k}$ is utilized, any error in state measurements can cause further variation from the ideal control effort. These imperfections in how control inputs affect the system can lead to degradation in the safety guarantees attained by the safety-critical controller (19).

In this part we consider a system with an input disturbance:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})(\mathbf{u} + \mathbf{d}(t)), \tag{23}$$

where $\mathbf{d} : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ reflects a time varying disturbance modifying the input $\mathbf{u}$ (such that the input the system actually receives is $\mathbf{u} + \mathbf{d}(t)$). We assume that the disturbance is bounded and piecewise continuous[2] in time. This is a practical assumption, and determining such bounds on the disturbance is an important step of the control design. This assumption also allows us to define:

$$\|\mathbf{d}\|_\infty = \sup_{t\geq 0}\|\mathbf{d}(t)\|_2 < \infty. \tag{24}$$

Given a continuous controller $\mathbf{k} : \mathbb{R}^n \to \mathbb{R}^m$, we may also introduce the notion of a disturbed closed-loop system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})(\mathbf{k}(\mathbf{x}) + \mathbf{d}(t)). \tag{25}$$

As before, we assume that for any initial condition $\mathbf{x}_0 \triangleq \mathbf{x}(0) \in \mathbb{R}^n$ and any bounded and piecewise continuous disturbance signal $\mathbf{d} : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$, there exists a unique solution $\mathbf{x}(t)$ to (25) for $t \geq 0$.

Example 6. We will consider an example disturbance signal for the inverted pendulum specified as:

$$d(t) = M(1 - s(t-5) - s(t-10) + s(t-15)) \tag{26}$$

[2] We take this definition as in [45], with the existence of one-sided limits.

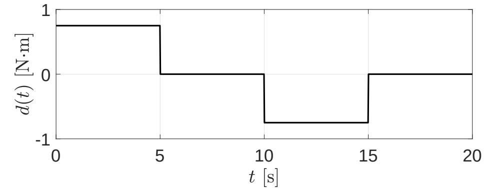

Fig. 5. Disturbance signal for the inverted pendulum system example as defined in (26).

where $M \geq 0$ and $s : \mathbb{R} \to \mathbb{R}$ is the *heaviside function*:

$$s(\tau) = \begin{cases} 0 & \text{if } \tau < 0, \\ 1 & \text{if } \tau \geq 0. \end{cases} \tag{27}$$

With this disturbance we have $\|d\|_\infty = M$. In this example we use the parameter value $M = 0.75$ [N·m] and the corresponding disturbance signal is depicted in Fig. 5.

A. Input-to-State Safety

In the presence of an input disturbance, *Input-to-State Safe Control Barrier Functions (ISSf-CBFs)* provide a tool for designing controllers with a formal safety guarantee [32], [34]. First, we present the notion of input-to-state safety (ISSf) which captures the intuition that it may no longer be possible to render the set $\mathcal{C}$ forward invariant (and thus safe) in the presence of disturbances. Instead, a larger set that scales proportionally with the disturbance may instead be rendered forward invariant. Specifically, consider the set $\mathcal{C}_\delta \subset \mathbb{R}^n$ defined as:

$$\mathcal{C}_\delta = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) + \gamma(h(\mathbf{x}), \delta) \geq 0\}, \tag{28}$$

$$\partial\mathcal{C}_\delta = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) + \gamma(h(\mathbf{x}), \delta) = 0\}, \tag{29}$$

$$\text{Int}(\mathcal{C}_\delta) = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) + \gamma(h(\mathbf{x}), \delta) > 0\}, \tag{30}$$

with $\gamma : \mathbb{R} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ satisfying $\gamma(a, \cdot) \in \mathcal{K}_\infty$ for all $a \in \mathbb{R}$. This implies $\mathcal{C}_\delta = \mathcal{C}$ when $\delta = 0$. We also require $\gamma(\cdot, b)$ to be continuously differentiable for all $b \in \mathbb{R}_{\geq 0}$. We have that $\partial\mathcal{C}_\delta$ and $\text{Int}(\mathcal{C}_\delta)$ are the *boundary* and *interior*, respectively, of the set $\mathcal{C}_\delta$. With this construction in mind, we have the following definition:

Definition 3 (*Input-to-State Safety (ISSf)*)**.** Let $\mathcal{C} \subset \mathbb{R}^n$ be the 0-superlevel set of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$. The system (25) is *input-to-state safe* (ISSf) with respect to the set $\mathcal{C}$ if there exists $\gamma : \mathbb{R} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ satisfying $\gamma(a, \cdot) \in \mathcal{K}_\infty$ for all $a \in \mathbb{R}$ and $\gamma(\cdot, b)$ continuously differentiable for all $b \in \mathbb{R}_{\geq 0}$ such that for all $\delta \geq 0$ and $\mathbf{d} : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ satisfying $\|\mathbf{d}\|_\infty \leq \delta$, the set $\mathcal{C}_\delta$ defined by (28)-(30) is forward invariant. If the system (25) is input-to-state safe with respect to the set $\mathcal{C}$, the set $\mathcal{C}$ is referred to as an *input-to-state safe set (ISSf set)*.

Similar to how Control Barrier Functions were defined in Sec. II, we now define *Input-to-State Safe Control Barrier Functions* as a tool for robust safety-critical control synthesis:

Definition 4 (*Input-to-State Safe Control Barrier Function (ISSf-CBF)*)**.** Let $\mathcal{C} \subset \mathbb{R}^n$ be the 0-superlevel set of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$. The function h is an *Input-to-State Safe Control Barrier Function (ISSf-CBF)* for (23) on $\mathcal{C}$ if there exists an $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$ and a continuously differentiable function $\epsilon : \mathbb{R} \to \mathbb{R}_{>0}$ such that for all $\mathbf{x} \in \mathbb{R}^n$:

$$\sup_{\mathbf{u} \in \mathbb{R}^m} [L_\mathbf{f} h(\mathbf{x}) + L_\mathbf{g} h(\mathbf{x})\mathbf{u}] > -\alpha(h(\mathbf{x})) + \frac{\|L_\mathbf{g} h(\mathbf{x})\|_2^2}{\epsilon(h(\mathbf{x}))}. \tag{31}$$

Given an ISSf-CBF h for (23) and corresponding functions $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$ and $\epsilon : \mathbb{R} \to \mathbb{R}_{>0}$, we can consider the point-wise set of all control values that satisfy (31):

$$K_{\text{ISSf}}(\mathbf{x}) = \left\{ \mathbf{u} \in \mathbb{R}^m \;\middle|\; \dot{h}(\mathbf{x}, \mathbf{u}) \geq -\alpha(h(\mathbf{x})) + \frac{\|L_\mathbf{g} h(\mathbf{x})\|_2^2}{\epsilon(h(\mathbf{x}))} \right\}. \tag{32}$$

The main theoretical result in [34] relates properties of the function ϵ and controllers synthesized via an ISSf-CBF to the input-to-state safety of the set $\mathcal{C}$:

Theorem 3 ([34])**.** *Let $\mathcal{C} \subset \mathbb{R}^n$ be the 0-superlevel set of a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}$. Let h be an ISSf-CBF for* (23) *on $\mathcal{C}$ with corresponding functions $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$ and $\epsilon : \mathbb{R} \to \mathbb{R}_{>0}$ such that ϵ and $\alpha^{-1} \in \mathcal{K}_\infty^{\mathrm{e}}$ are continuously differentiable and ϵ satisfies:*

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}r}(h(\mathbf{x})) \geq 0, \tag{33}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Then the set $K_{\text{ISSf}}(\mathbf{x})$ is non-empty for each $\mathbf{x} \in \mathbb{R}^n$, and if a continuous controller $\mathbf{k} : \mathbb{R}^n \to \mathbb{R}^m$ satisfies $\mathbf{k}(\mathbf{x}) \in K_{\text{ISSf}}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, then for any $\delta \geq 0$, the system (25) *is safe with respect to the set $\mathcal{C}_\delta$ defined as in* (28)-(30) *with γ defined as:*

$$\gamma(h(\mathbf{x}), \delta) \triangleq -\alpha^{-1}\left(-\frac{\epsilon(h(\mathbf{x}))\delta^2}{4}\right), \tag{34}$$

for all $\mathbf{d}$ satisfying $\|\mathbf{d}\|_\infty \leq \delta$. This implies $\mathcal{C}$ is an ISSf set.

Remark 2. The original definition of ISSf presented in [32] differs from Definition 3 in the function γ. We allow γ to be a function of h in addition to δ. This leads to a generalization of the ISSf-CBF definition in [32], which reduces to the definition given in [32] if $\epsilon(r) = c > 0$ for all $r \in \mathbb{R}$. The definitions presented here provide a factor of flexibility in controller design as detailed in [34].

The boundary of the set $\mathcal{C}_\delta$ that is rendered forward invariant is defined as a level-set of the ISSf-CBF h as in (29). Given a $\delta \geq 0$, the value of h on this level set, denoted as $h^* \leq 0$, can be found by solving the equation:

$$h^* \underbrace{-\alpha^{-1}\left(-\frac{\epsilon(h^*)\delta^2}{4}\right)}_{\gamma(h^*, \delta)} = 0. \tag{35}$$

By definition, $\gamma(h^*, \delta)$ must be strictly positive for $\delta > 0$, implying that $h^* < 0$ in the presence of disturbances. The safety-critical controllers designed in the next section will guarantee $h(\mathbf{x}(t)) \geq h^*$. Moreover, as δ increases, h^* must get more negative, implying that the boundary of $\mathcal{C}_\delta$ falls farther from the boundary of $\mathcal{C}$. Control over this degradation in safety can be achieved by modifying the function ϵ to yield different values of h^* as specified in (35). Various functions that satisfy the necessary conditions for ϵ can be seen in Fig. 6.

Example 7. Given our choice of α for the inverted pendulum

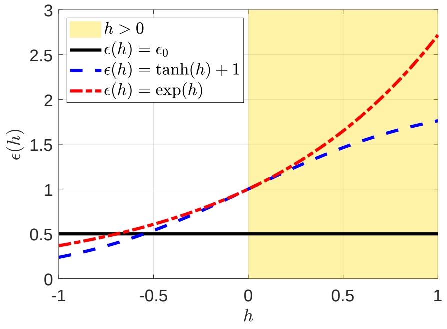

Fig. 6. Examples of the function ϵ that satisfy the condition in (33).

system, we have that:

$$\gamma(h(\theta,\dot{\theta}),\delta) = \frac{\epsilon(h(\theta,\dot{\theta}))\delta^2}{4\alpha_{\mathrm{c}}}. \tag{36}$$

As our disturbance signal is bounded by M, determining the set kept forward invariant is done by considering $\delta = M$. Thus, we use the parameter value $\delta = 0.75$ [N·m]. We choose the exponential function:

$$\epsilon(r) = \epsilon_0 \mathrm{e}^{\lambda r}, \tag{37}$$

with parameters $\epsilon_0 > 0$ and $\lambda \geq 0$. With this choice we have that (35) reduces to:

$$h^* + \frac{\epsilon_0 \mathrm{e}^{\lambda h^*}\delta^2}{4\alpha_{\mathrm{c}}} = 0. \tag{38}$$

Once ϵ_0 and λ are specified, (38) can be solved for h^* to find the value of h that corresponds to the boundary $\partial\mathcal{C}_\delta$. Fig. 7 (left) shows the value of h^* for the different choices of ϵ_0 and λ specified in Table I. The boundary $\partial\mathcal{C}_\delta$ corresponding to each set of parameters is shown in Fig. 7 (right) using the same color code. The black and red parameter sets return the same value of h^*, and thus the produce the same boundary $\partial\mathcal{C}_\delta$. In contrast, the green parameter set yields a larger set $\mathcal{C}_\delta$ as indicated by the smaller value of h^* in Table I.

Color	Black	Red	Green
$\epsilon_0\ [\frac{1}{\mathrm{N}^2\mathrm{m}^2\mathrm{s}}]$	0.15	0.5	4
λ	0	12	3
h^*	-0.1	-0.1	-0.55

TABLE I. Parameter sets for (37) in the inverted pendulum example.

B. Robust Safety-Critical Controller

As we saw with CBFs, it is possible to use an ISSf-CBF to synthesize controllers that render a set $\mathcal{C}_\delta$ forward invariant, thus rendering the set $\mathcal{C}$ ISSf. Suppose that we have an ISSf-CBF h for (23) on $\mathcal{C}$ with corresponding functions $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$ and $\epsilon : \mathbb{R} \to \mathbb{R}_{>0}$ that meet the requirements of Theorem 3. Furthermore, suppose we have a continuous nominal controller $\mathbf{k}_{\mathrm{n}} : \mathbb{R}^n \to \mathbb{R}^m$ that is not necessarily safe. Motivated by the optimization-based safety-critical formulation given in (18), we define a controller $\overline{\mathbf{k}}_{\mathrm{QP}} : \mathbb{R}^n \to \mathbb{R}^m$ as:

$$\begin{aligned} \overline{\mathbf{k}}_{\mathrm{QP}}(\mathbf{x}) = \underset{\mathbf{u}\in\mathbb{R}^m}{\operatorname{argmin}} \quad & \frac{1}{2}\|\mathbf{u} - \mathbf{k}_{\mathrm{n}}(\mathbf{x})\|_2^2 \\ \text{s.t.} \quad & \dot{h}(\mathbf{x},\mathbf{u}) \geq -\alpha(h(\mathbf{x})) + \frac{\|L_{\mathbf{g}}h(\mathbf{x})\|_2^2}{\epsilon(h(\mathbf{x}))}. \end{aligned} \tag{39}$$

The following theorem provides a closed-form solution to the optimization problem defining this controller and specify the continuity and safety properties of the resulting controller:

Theorem 4. *Let $\mathcal{C}$ be the 0-superlevel set of a function $h : \mathbb{R}^n \to \mathbb{R}$, and let $\mathbf{k}_{\mathrm{n}} : \mathbb{R}^n \to \mathbb{R}^m$ be a continuous controller. If h is an ISSf-CBF for* (23) *on the set $\mathcal{C}$ with corresponding functions $\alpha \in \mathcal{K}_\infty^{\mathrm{e}}$ with continuously differentiable inverse $\alpha^{-1} \in \mathcal{K}_\infty^{\mathrm{e}}$, and continuously differentiable $\epsilon : \mathbb{R} \to \mathbb{R}_{>0}$ satisfying* (33)*, then the optimization problem in* (39) *is feasible for any $\mathbf{x} \in \mathbb{R}^n$ and has a closed-form solution given by:*

$$\overline{\mathbf{k}}_{\mathrm{QP}}(\mathbf{x}) = \mathbf{k}_{\mathrm{n}}(\mathbf{x}) + \max\{0, \overline{\eta}(\mathbf{x})\} L_{\mathbf{g}}h(\mathbf{x})^\top, \tag{40}$$

where the function $\overline{\eta} : \mathbb{R}^n \to \mathbb{R}$ is defined as:

$$\overline{\eta}(\mathbf{x}) = \begin{cases} -\frac{\dot{h}(\mathbf{x},\mathbf{k}_{\mathrm{n}}(\mathbf{x})) + \alpha(h(\mathbf{x}))}{\|L_{\mathbf{g}}h(\mathbf{x})\|_2^2} + \frac{1}{\epsilon(h(\mathbf{x}))} & \text{if } L_{\mathbf{g}}h(\mathbf{x}) \neq \mathbf{0}, \\ 0 & \text{if } L_{\mathbf{g}}h(\mathbf{x}) = \mathbf{0}. \end{cases} \tag{41}$$

Furthermore, $\overline{\mathbf{k}}_{\mathrm{QP}}$ is continuous and $\overline{\mathbf{k}}_{\mathrm{QP}}(\mathbf{x}) \in K_{\mathrm{ISSf}}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$.

The proof of this theorem is performed similarly to the proof of Theorem 2 with simple modifications for the introduction of ϵ, and thus, it is omitted.

Remark 3. For a single input ($m = 1$), if $L_{\mathbf{g}}h(\mathbf{x}) > 0$ for a particular $\mathbf{x} \in \mathbb{R}^n$, the controller (40) can be expressed as:

$$\overline{k}_{\mathrm{QP}}(\mathbf{x}) = \max\left\{k_{\mathrm{n}}(\mathbf{x}), -\frac{L_{\mathbf{f}}h(\mathbf{x}) + \alpha(h(\mathbf{x}))}{L_{\mathbf{g}}h(\mathbf{x})} + \frac{L_{\mathbf{g}}h(\mathbf{x})}{\epsilon(h(\mathbf{x}))}\right\}. \tag{42}$$

Similarly, if $L_{\mathbf{g}}h(\mathbf{x}) < 0$ for a particular $\mathbf{x} \in \mathbb{R}^n$, the controller (40) reduces to:

$$\overline{k}_{\mathrm{QP}}(\mathbf{x}) = \min\left\{k_{\mathrm{n}}(\mathbf{x}), -\frac{L_{\mathbf{f}}h(\mathbf{x}) + \alpha(h(\mathbf{x}))}{L_{\mathbf{g}}h(\mathbf{x})} + \frac{L_{\mathbf{g}}h(\mathbf{x})}{\epsilon(h(\mathbf{x}))}\right\}. \tag{43}$$

These controllers can be switched between depending on the sign of $L_{\mathbf{g}}h(\mathbf{x})$, with $\overline{k}_{\mathrm{QP}}(\mathbf{x}) = k_{\mathrm{n}}(\mathbf{x})$ when $L_{\mathbf{g}}h(\mathbf{x}) = 0$.

Example 8. We deploy the safety-critical controller k_{QP} defined in (21)-(22) with the nominal controller k_{n} as in (16) to the inverted pendulum system without considering the disturbance d defined in (26). We see in the right panel of Fig. 7 that this controllers fails to keep the system in the safe set $\mathcal{C}$, and deviates from it significantly. We next deploy the safety-critical controller $\overline{k}_{\mathrm{QP}}$ defined in (42)-(43) with the nominal controller k_{n} as in (16). The exponential function given in (37) is utilized with the black and red parameter pairs as specified in Table I. We see in the right panel of Fig. 7 that for both parameter sets, the controller keep the trajectories within $\mathcal{C}$, and thus, within $\mathcal{C}_\delta$, that is, it guarantees $h(\mathbf{x}(t)) \geq h^*$. Despite having the same values of h^*, we see the red parameter sets allows the system to more closely approach the boundary of the safe set, while the black parameter set forces the system to the equilibrium more directly. A detailed discussion about the effect of the parameter λ on conservativeness of the controller is provided in [34].

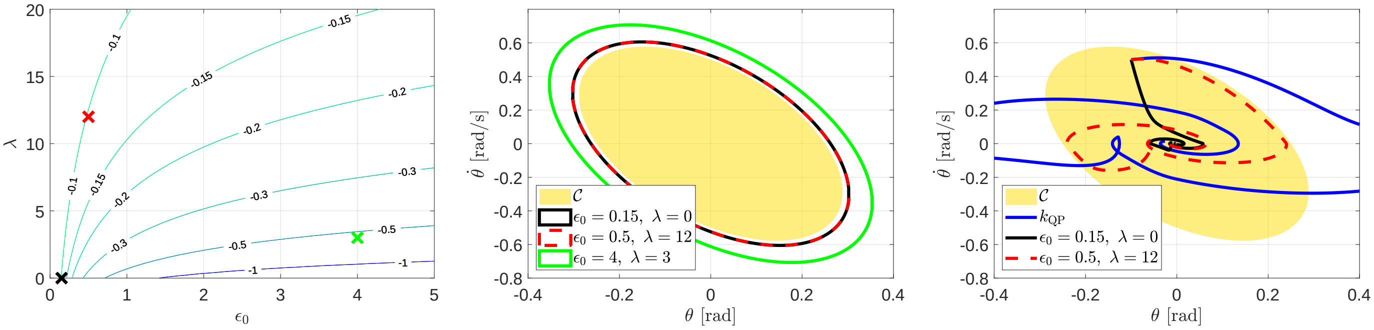

Fig. 7. (Left) Curves corresponding to the value of h^* solving (38) across the (ϵ_0, λ) parameter space for the inverted pendulum example. (Center) The boundary of the set $\mathcal{C}_\delta$ rendered forward invariant for different choices of the parameters ϵ_0 and λ for the inverted pendulum example. Note that the $\mathcal{C}_\delta$ contains the set $\mathcal{C}$ for each parameter set. (Right) Simulation results for the inverted pendulum system with disturbances. The gold ellipse is the safe set $\mathcal{C}$ defined in (8). The blue line is the trajectory of the system evolving under k_{QP} defined in (21)-(22), which is not robust to disturbances and leaves the safe set. The black and dashed red lines are the trajectory of the system evolving under $\overline{k}_{\mathrm{QP}}$ defined in (42)-(43) with different values for ϵ_0 and λ. While both parameter sets yield the same forward invariant set $\mathcal{C}_\delta$, the red parameter set is less conservative and allows the system to approach the boundary.

IV. Safety-Critical Controller Design for a Connected Automated Truck

In this section, we go through the process of designing a safety-critical longitudinal controller for a connected automated truck. We first introduce the physical system and define a safe set via a Control Barrier Function. We then present a nominal performance-based controller, and synthesize a safety-critical controller that modifies this nominal controller in a minimally invasive way while ensuring safety.

A. Modeling Longitudinal Dynamics

In this work we consider a rear-axle-driven truck without a trailer. Assuming the truck's tires roll without slipping and the truck travels on a flat road with no headwind, the longitudinal dynamics of the truck are described by the following model:

$$\dot{v} = \frac{T}{m_{\mathrm{eff}} R} - \frac{kv^2 + mg\gamma}{m_{\mathrm{eff}}}. \tag{44}$$

Here the state is given by the truck's longitudinal speed $v \in \mathbb{R}$, the input is the torque applied on the rear axle $T \in \mathbb{R}$, and the parameters in the model are the mass of the truck m, the mass moment of inertia of the rotating elements I, the tire radius R, the effective mass $m_{\mathrm{eff}} = m + \frac{I}{R^2}$, the air drag constant k, gravitational acceleration g, and rolling resistance coefficient γ. Note that the second term in (44) is dissipative in nature, and slows down the vehicle when it has a positive velocity. This term may be directly accounted for in the control design through via feedback linearization techniques [43], or may be ignored as its omission simply introduces a factor of conservativeness to the controller in terms of safety. The torque input commanded of the system is computed from a desired longitudinal acceleration command $u \in \mathbb{R}$ via feed-forward maps. This torque input command is provided by a drive-by-wire system to the low-level power generation systems that produce the actual torque T; see Fig. 8. With these feed-forward maps in mind, we may simplify the model of the longitudinal dynamics of the truck to:

$$\dot{v} = u. \tag{45}$$

Now let us consider the scenario when the truck follows a connected vehicle as depicted in Fig. 8. Using the truck model in (45), the dynamics of this connected system are given by:

$$\begin{aligned} \dot{D} &= v_{\mathrm{L}} - v, \\ \dot{v} &= u, \\ \dot{v}_{\mathrm{L}} &= a_{\mathrm{L}}, \end{aligned} \tag{46}$$

where $v_{\mathrm{L}}, a_{\mathrm{L}} \in \mathbb{R}$ are the speed and acceleration of the leading vehicle, respectively, and $D \in \mathbb{R}$ denotes the bumper-to-bumper headway distance between the truck and the lead vehicle, yielding the state $\mathbf{x} = [D, v, v_{\mathrm{L}}]^\top \in \mathbb{R}^3$. The truck and lead vehicle are outfitted with vehicle-to-vehicle (V2V) communication systems, permitting the truck to receive motion information from the lead vehicle such as its GPS position which yields the distance D, its speed v_{L}, and its acceleration a_{L}. We assume that the leader's behavior satisfies:

$$a_{\mathrm{L}} \in [-\underline{a}_{\mathrm{L}}, \overline{a}_{\mathrm{L}}], \qquad v_{\mathrm{L}} \in [0, \overline{v}_{\mathrm{L}}], \tag{47}$$

where the parameters $\underline{a}_{\mathrm{L}}, \overline{a}_{\mathrm{L}}, \overline{v}_{\mathrm{L}} > 0$ reflect a city-driving scenario; see Table II.

B. Safety and Control Barrier Function

The safety task for the truck is to maintain a safe distance behind the leader. This task motivates a Control Barrier

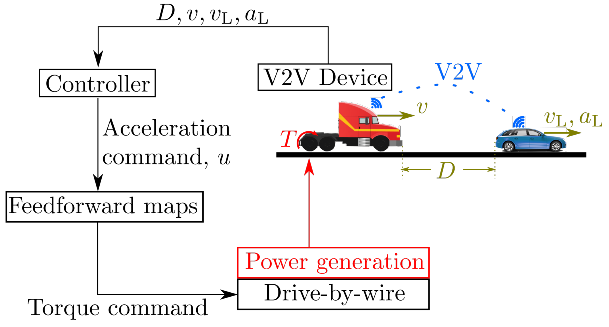
$D, v, v_{\mathrm{L}}, a_{\mathrm{L}}$
Controller
V2V Device
V2V
v
$v_{\mathrm{L}}, a_{\mathrm{L}}$
Acceleration command, u
T
D
Feedforward maps
Power generation
Drive-by-wire
Torque command

Fig. 8. A connected automated truck following a connected vehicle.

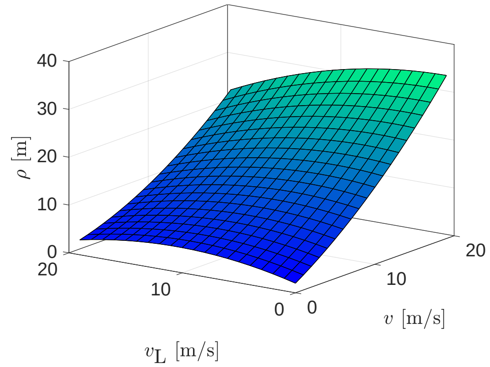

Fig. 9. The value of the function ρ defined in (49), which defines the minimum safe following distance as a function of the leader's velocity $v_{\rm L}$ and truck velocity v.

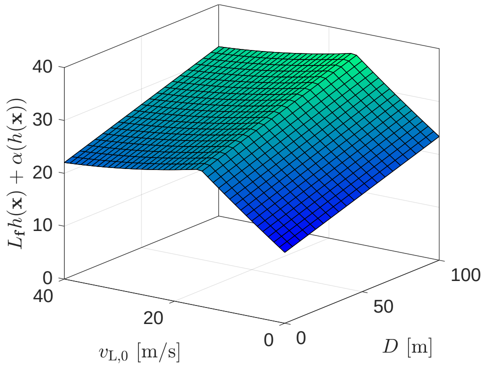

Fig. 10. The value of $L_{\mathbf{f}}h(\mathbf{x}) + \alpha(h(\mathbf{x}))$ as defined in (52) when $L_{\mathbf{g}}h(\mathbf{x}) = 0$ for various distances. As the function is strictly positive over the domain of interest, h is a CBF on $\mathcal{C}$ for (46).

Function of the form:

$$h(D, v, v_{\rm L}) = D - \rho(v, v_{\rm L}), \tag{48}$$

where the headway function $\rho : \mathbb{R}^2 \to \mathbb{R}$ describes the minimum safe distance between the vehicles given their current velocities, v and $v_{\rm L}$. Motivated by [4] and [46], we define the headway function as:

$$\rho(v, v_{\rm L}) = c_0 + c_1 v + c_2 v_{\rm L} + c_3 v^2 + c_4 v v_{\rm L} + c_5 v_{\rm L}^2, \tag{49}$$

with parameters $c_i \in \mathbb{R}$ for $i = 0, \ldots, 5$; see Table II. The value of the function ρ is visualized in the left panel of Fig. 9. The corresponding safe set defined by h is given by:

$$\mathcal{C} = \left\{ \begin{bmatrix} D \\ v \\ v_{\rm L} \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; D \geq \rho(v, v_{\rm L}) \right\}. \tag{50}$$

To verify that the function h is a CBF for (46), observe that:

$$L_{\mathbf{g}}h(D_0, v_0, v_{\rm L,0}) = 0 \implies c_1 + 2c_3 v_0 + c_4 v_{\rm L,0} = 0, \tag{51}$$

which describes a line in $(v, v_{\rm L})$ space where the condition (10) must be met for all $D \in \mathbb{R}$. We consider $\alpha(r) = \alpha_{\rm c} r$ with $\alpha_{\rm c} > 0$, yielding:

$$\begin{aligned} L_{\mathbf{f}}h(D, v_0, v_{\rm L,0}) &+ \alpha(h(D, v_0, v_{\rm L,0})) \\ &= v_{\rm L,0} - v_0 - a_{\rm L}(c_2 + c_4 v_0 + 2c_5 v_{\rm L,0}) \\ &\quad + \alpha_{\rm c}(D - \rho(v_0, v_{\rm L,0})). \end{aligned} \tag{52}$$

Since checking the condition (10) analytically may be cumbersome using (52), we graphically evaluate it over a range of D and $v_{\rm L,0}$ (and v_0 defined implicitly through (51)) while taking the worst case value of $a_{\rm L}$ making (52) as negative as possible; see the right panel in Fig. 9. This shows that for $\alpha_{\rm c} = 0.1$ [1/s], the value of (52) is strictly positive, ensuring the condition (10) is satisfied.

$\overline{a}_{\rm L} = 5$ [m/s^2]	$c_0 = 2$ [m]	$\kappa = 0.8$ [1/s]
$\underline{a}_{\rm L} = 10$ [m/s^2]	$c_1 = 1.1$ [s]	$\alpha_{\rm c} = 0.1$ [1/s]
$\overline{v}_{\rm L} = 20$ [m/s]	$c_2 = 0.6$ [s]	$D_{\rm st} = 5$ [m]
$\delta = 4.5$ [m/s^2]	$c_3 = 0.03$ [s^2/m]	$D_{\rm go} = 30$ [m]
$\epsilon_0 = 0.5$ [s^3/m]	$c_4 = -0.03$ [s^2/m]	$A = 0.4$ [1/s]
$\lambda = 0.4$ [1/m]	$c_5 = -0.03$ [s^2/m]	$B = 0.5$ [1/s]

TABLE II. Parameter values used in controller design.

C. Controller Design

Beyond the task of safety, we wish for our controller to maximize other performance criteria such as ride comfort, fuel economy, and string stability [47], [48]. To accomplish this goal, we first design a nominal controller that prioritizes performance. In particular, we will design a connected cruise controller (CCC) for the truck that utilizes information about the lead vehicle available through V2V connectivity. We propose the following controller structure:

$$k_{\rm n}(D, v, v_{\rm L}) = A(V(D) - v) + B(W(v_{\rm L}) - v), \tag{53}$$

with parameters $A, B > 0$, functions $V : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, and $W : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. The first term in (53) specifies the distance based speed error with the range policy:

$$V(D) = \begin{cases} 0 & \text{if} \quad D < D_{\rm st}, \\ \kappa(D - D_{\rm st}) & \text{if} \quad D_{\rm st} \leq D \leq D_{\rm go}, \\ \overline{v}_{\rm L} & \text{if} \quad D > D_{\rm go}, \end{cases} \tag{54}$$

depicted in the top panel of Fig. 11, producing a desired speed based on the distance D. Here $D_{\rm st} > 0$ is the desired stopping distance, $1/\kappa > 0$ is the desired time headway, and $D_{\rm go} = \overline{v}_{\rm L}/\kappa + D_{\rm st}$. The second term in (53) specifies the error related to the relative speed with the speed policy:

$$W(v_{\rm L}) = \begin{cases} v_{\rm L} & \text{if} \quad v_{\rm L} \leq \overline{v}_{\rm L}, \\ \overline{v}_{\rm L} & \text{if} \quad v_{\rm L} > \overline{v}_{\rm L}, \end{cases} \tag{55}$$

depicted in the bottom panel of Fig. 11, which bounds the speed error if the lead vehicle violates $v_{\rm L} \leq \overline{v}_{\rm L}$.

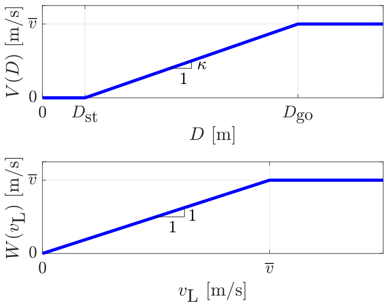

Fig. 11. (Top) Distance based range policy V defined in (54). (Bottom) Speed policy W defined in (55).

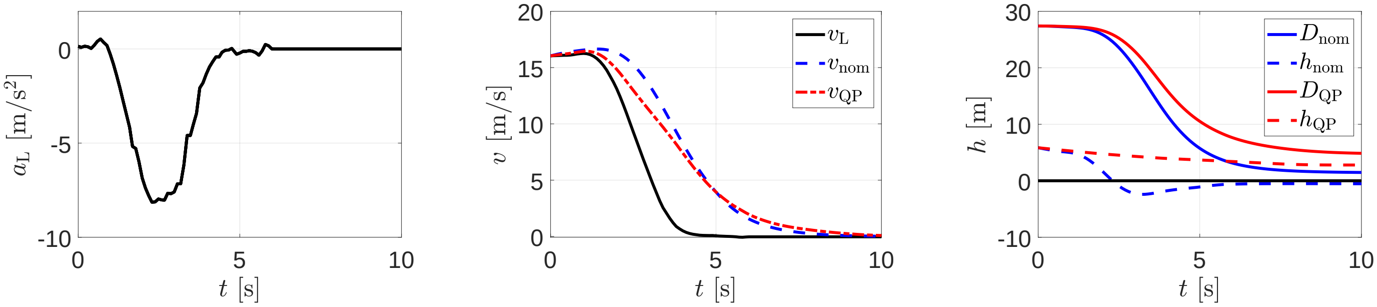

Fig. 12. (Left) Example profile for acceleration a_L of lead vehicle used in numerical simulation. (Center) Velocity v_L of lead vehicle (black) and velocity of the truck using the nominal controller (53) (blue) and safety-critical controller (57) (red). (Right) Following distance D and value of CBF h using the nominal controller and safety-critical controller.

Having designed the CBF h and the performance based nominal controller k_n, we can unify them via the safety-critical controller formulation for single input systems given in (21)-(22). Here we have:

$$\begin{aligned} L_\mathbf{f}h(D,v,v_\mathrm{L}) &= v_\mathrm{L} - v - a_\mathrm{L}\left(c_2 + c_4 v + 2c_5 v_\mathrm{L}\right), \\ L_\mathbf{g}h(D,v,v_\mathrm{L}) &= -c_1 - 2c_3 v - c_4 v_\mathrm{L}, \end{aligned} \tag{56}$$

where $L_\mathbf{g}h(D,v,v_\mathrm{L}) < 0$ for $v \geq 0$ and $v_\mathrm{L} \in [0, \overline{v}_\mathrm{L}]$. Then one may utilize the switch structure (22):

$$k_\mathrm{QP}(D,v,v_\mathrm{L}) = \min\left\{k_\mathrm{n}(D,v,v_\mathrm{L}), k_\mathrm{s}(D,v,v_\mathrm{L})\right\}, \tag{57}$$

where the second term is defined as:

$$k_\mathrm{s}(D,v,v_\mathrm{L}) = -\frac{L_\mathbf{f}h(D,v,v_\mathrm{L}) + \alpha_\mathrm{c} h(D,v,v_\mathrm{L})}{L_\mathbf{g}h(D,v,v_\mathrm{L})}. \tag{58}$$

This controller utilizes the nominal controller k_n to optimize the performance when it is safe. Otherwise, the provably safe controller k_s becomes smaller than k_n and intervenes to ensure safety. Note that $L_\mathbf{g}h(D,v,v_\mathrm{L}) > 0$ for sufficiently large $v_\mathrm{L} > \overline{v}_\mathrm{L}$ (as c_4 is negative) as well as sufficiently negative $v < 0$, yielding the switch structure (21), but this is outside the domain of interest in this application.

We simulate both the nominal controller and safety-critical controller via numerical integration of the model (46) from the initial condition $\mathbf{x}(0) = [27.4, 16, 16]^\top \in \mathcal{C}$. We use parameter values as specified in Table II. The acceleration a_L of the lead vehicle is given by a time profile reflecting a hard braking event, as seen in Fig. 12 (left). The velocity of the truck converges to zero and a crash does not occur for both controllers, but only the safety-critical controller ensures the truck maintains a safe distance (indicated by $h_\mathrm{QP}(\mathbf{x}(t)) \geq 0$) as seen in Fig. 12 (center, right). We see that the nominal controller brakes less aggressively than the safety-critical controller, and thus does not react quickly enough to avoid violating the safe following distance requirement.

V. Experimental Results & Robust Design

In this section we provide a description of the automated truck experimental configuration and present results using the nominal and safety-critical controllers. Furthermore, we deploy the method of robust control design using ISSf developed in Section III, and demonstrate its advantages experimentally.

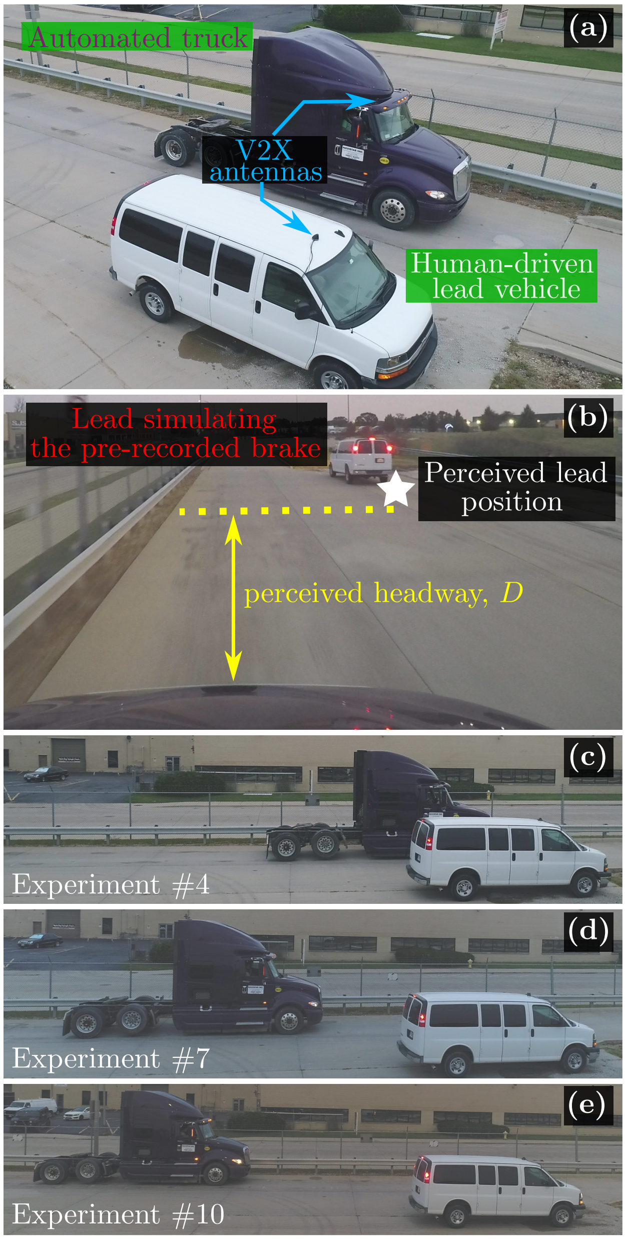

Fig. 13. (a) Vehicles used in experiments. (b) Image from the dashboard of the truck during an experimental run. (c,d,e) Final configurations of separate experiments. See [49] for a video.

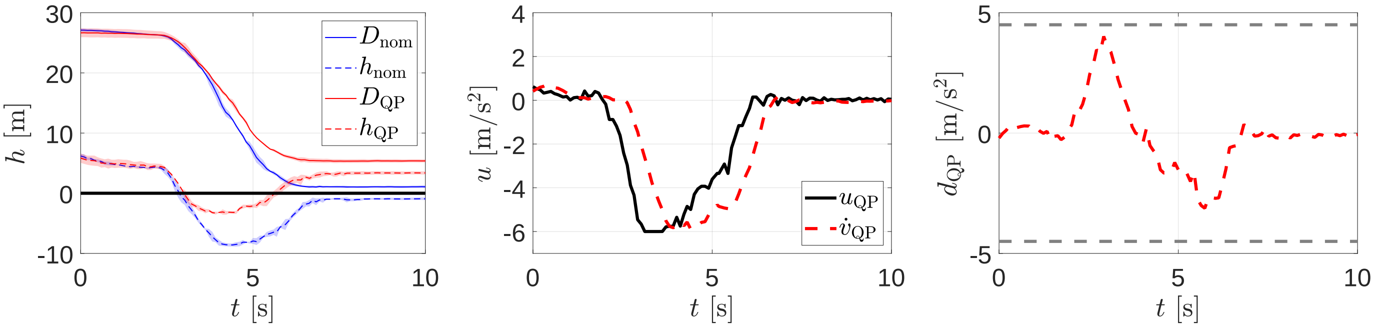

Fig. 14. (Left) Mean value (lines) and standard deviations (fills) of the distance D and the CBF h when using the nominal controller defined in (53) (blue) and the safety-critical controller defined in (57) (red) in the truck experiment. The repeated experiments with these controllers are highly consistent. (Center) Discrepancy between acceleration commanded by safety-critical controller (black) and actual acceleration of the automated truck (red). (Right) Disturbance signal in input seen by the truck used to define the model (59).

A. *Experimental Setup and Procedure*

The automated truck used in our experiments is an International ProStar+ Class-8 truck developed by the Navistar [50]; see Fig. 13(a). Both the automated truck and the lead vehicle are equipped with a V2X Onboard Unit (OBU) developed by Commsignia [51]. These units are equipped with an accelerometer, gyroscope, magnetometer, and GPS unit. Furthermore, these OBUs support peer-to-peer communication such that the automated truck may receive position, velocity, and acceleration data from the lead vehicle through V2X antennas shown in Fig. 13(a). The automated truck is additionally equipped with a Mobile Real-Time Targeting Machine developed by Speedgoat [52], which interfaces with the V2X OBU and the truck's Engine Controller Unit (ECU) through a Controller Area Network (CAN) bus. The Speedgoat runs the controller for the system given a measurement stream of values for D, v, v_{L}, and a_{L} coming from the V2X OBUs. It computes a desired acceleration input and converts it to a corresponding torque value through a feed-forward map. A drive-by-wire system on the truck controls the engine and the brake torques accordingly. The steering of the truck is done manually by a human driver in the experiments.

In an effort to evaluate the repeatability of our experiments, it is necessary to eliminate variation in the lead vehicle's behavior, which is being driven by a human. To achieve this, we use a pre-recorded time profile of position, velocity, and acceleration of the lead vehicle while it performs a hard braking event. This profile for a_{L} and v_{L} is seen in the left and center panels of Fig. 12, and was used to produce our simulation results. We stream this data to the truck controller as the *perceived lead vehicle* in our experiments. Experiments also include a physical lead vehicle simulating the pre-recorded motion for visualization purposes; see Fig. 13(b). Importantly, the evaluation of safety is derived from evaluating the CBF using the recorded time profiles rather than simply detecting collisions such as Fig. 13(c). A video of the experiments are available online [49].

B. *Input Disturbances*

We deploy both the nominal and safety-critical controller on the automated truck with results as seen in the left panel of Fig. 14. We see that not only does the nominal controller consistently fails to meet the safety requirements imposed by the CBF h, but the safety-critical controller also consistently fails to meet the safety requirements. The top row in Fig. 1 illustrates an experimental run with the nominal controller.

To understand why the safety-critical controller fails, we examine the discrepancy between the commanded acceleration and actual acceleration of the automated truck, as seen in the center panel of Fig. 14. One may observe a delay between the commanded acceleration and the achieved acceleration. This delay in acceleration is due to the fact that the power generation of the truck is a complex nonlinear dynamical system that has been imperfectly abstracted away by the feed-forward maps that allow the simplified model in (45). Rather than attempting to work with this complex nonlinear dynamic system and improving the feed-forward maps, we describe the discrepancy in commanded and actual acceleration as a disturbance in the simplified model:

$$\dot{v} = u + d(t), \tag{59}$$

where $d : \mathbb{R}_{\geq 0} \to \mathbb{R}$ reflects the difference between commanded acceleration and actual acceleration.

As the disturbance d is caused by the complicated interactions of the drive-by-wire system and power generation dynamics, it may be difficult to use model-based techniques to construct a meaningful bound δ for the worst-case disturbance. Instead, we estimate the worst-case disturbance empirically by comparing the actual acceleration $\dot{v}(t)$ to the commanded acceleration $u(t)$. In the right panel of Fig. 14, we see that the largest difference in the commanded and actual acceleration is around 4 [m/s^2]. Thus, we study the degradation of safety of the system taking a slightly larger value $\delta = 4.5$ [m/s^2].

C. *Robust Design*

To overcome this disturbance and improve the safe behavior of the truck, we deploy the tools of ISSf-CBFs described in Section III. As h satisfies the CBF condition (10), it also satisfies the ISSf-CBF condition (31), where we take:

$$\epsilon(r) = \epsilon_0 \mathrm{e}^{\lambda r}, \tag{60}$$

with $\epsilon_0 > 0$ and $\lambda \geq 0$. The parameter λ introduces a measure of flexibility by allowing one to require a greater degree of

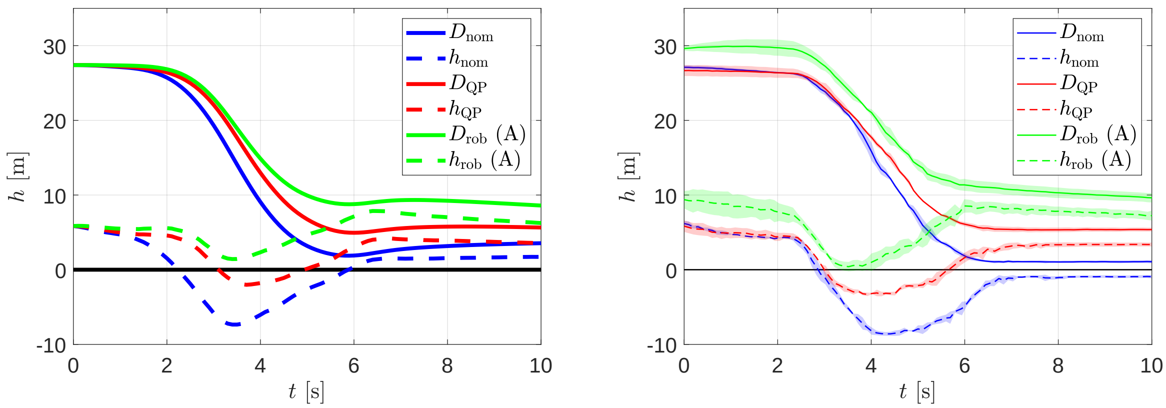

Fig. 15. (Left) Following distance and value of ISSf-CBF using the nominal controller (blue), safety-critical controller (red), and robust safety-critical controller (green) in the disturbed simulation. (Right) Mean value (lines) and standard deviations (fills) of the distance D and the ISSf-CBF h using the nominal controller (blue), safety-critical controller (red), and robust safety-critical controller (green) in experiment.

robustness when the truck is close to the leading vehicle, and less robustness when the distance is greater. Given (60), the forward invariant set is given by:

$$\mathcal{C}_\delta = \left\{ \begin{bmatrix} D \\ v \\ v_{\mathrm{L}} \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; h(D, v, v_{\mathrm{L}}) \geq -\frac{\epsilon_0 \mathrm{e}^{\lambda h(D,v,v_{\mathrm{L}})} \delta^2}{4\alpha_{\mathrm{c}}} \right\}. \tag{61}$$

As discussed in the inverted pendulum example, the set C_δ being forward invariant implies that $h(\mathbf{x}(t)) \geq h^*$, where h^* is the value of the ISSf-CBF h on the boundary of C_δ, which can be calculated by solving (35). The value of h^* for different choices of ϵ_0 and λ can be seen in Table III. We then construct an optimization-based controller giving the switch structure (43), since $L_{\mathbf{g}} h(D, v, v_{\mathrm{L}}) < 0$ for $v \geq 0$ and $v_{\mathrm{L}} \in [0, \overline{v}_{\mathrm{L}}]$ (cf. (56)). This results in:

$$k_{\mathrm{rob}}(D, v, v_{\mathrm{L}}) = \min\left\{ k_{\mathrm{n}}(D, v, v_{\mathrm{L}}), \overline{k}_{\mathrm{s}}(D, v, v_{\mathrm{L}}) \right\}, \tag{62}$$

where:

$$\overline{k}_{\mathrm{s}}(D, v, v_{\mathrm{L}}) = k_{\mathrm{s}}(D, v, v_{\mathrm{L}}) + \frac{L_{\mathbf{g}} h(D, v, v_{\mathrm{L}})}{\epsilon_0 \mathrm{e}^{\lambda h(D,v,v_{\mathrm{L}})}}, \tag{63}$$

and k_{n} and k_{s} are given by (53) and (58), respectively.

We simulate the nominal controller, safety-critical controller, and robust safety-critical controller via numerical integration of the model (46) from the initial condition $\mathbf{x}(0) = [27.4, 16, 16]^\top \in \mathcal{C}$ while disturbing the input using the signal in shown in the right panel of Fig. 14. We use parameter values as specified in Table II. We see in the left panel of Fig. 15 that introducing the disturbance signal into our simulation allows us to recreate the failures of the nominal controller and safety-critical controller that we saw experimentally in Fig. 14. Furthermore, we see that the robust safety-critical controller maintains the safety of the system even in the presence of the disturbance.

D. Robust Experimental Results

Here we show the results when the robust safety-critical controller is deployed on the connected automated truck. Sets of three experimental runs were conducted using each parameter pair ϵ_0 and λ shown in Table III. The experimental results using the parameter set $\epsilon_0 = 0.5$ [s^3/m] and $\lambda = 0.4$ [1/m] (labeled as parameter pair (A)) can be seen in the right panel of Fig. 15 and are visualized at the bottom row of Fig. 1. With these parameters the system is rendered safe, as the value of h does not drop below 0. Although the robust safety-critical controller displays a larger standard deviation across the three experimental runs compared to the nominal and safety-critical controllers, it consistently satisfies the original safety requirement.

When evaluating how the system behavior depends on the values of the parameters ϵ_0 and λ, we first consider whether the original safety requirement is met, i.e., whether or not the value of h remains positive. While the robust-safety critical controller does not provide a theoretical guarantee that h will remain non-negative (it only guarantees that $h(\mathbf{x}(t)) \geq h^*$), for certain values of ϵ_0 and λ the original safety requirement are still met, as seen in the inverted pendulum example as well as the connected automated truck experiments. The minimum value $h_{\min}$ of the barrier function, observed during the experimental runs, is shown in Table III. This is also visualized in the left panel of Fig. 16, where green markers indicate sets of parameter values for which the safety requirement is met, and red markers indicate those for which it is not met. We see that safety can be achieved using the original ISSf-CBF

Label	ϵ_0 [s^3/m]	λ [1/m]	h^* [m]	$h_{\min}$ [m]	$\tilde{D}_{\mathrm{ss}}$ [m]
(B)	0.8	0	−40.50	22.09	25.43
	3	0	−151.88	2.99	6.19
(D)	4	0	−202.50	1.02	4.69
	5	0	−253.13	−0.45	3.54
(A)	0.5	0.4	−4.38	0.35	4.70
	0.5	0.5	−3.80	−1.27	2.22
(C)	0.8	0.25	−7.01	0.78	4.34
	0.8	0.35	−5.64	−1.03	2.22
	1.0	0.25	−7.59	−0.86	3.07

TABLE III. Sets of parameter values used for the exponential function (60) in the automated truck experiments with theoretical safety guarantee h^*, minimum experimental value of the ISSf-CBF $h_{\min}$, and shift in the steady-state tracking distance $\tilde{D}_{\mathrm{ss}}$ by (64).

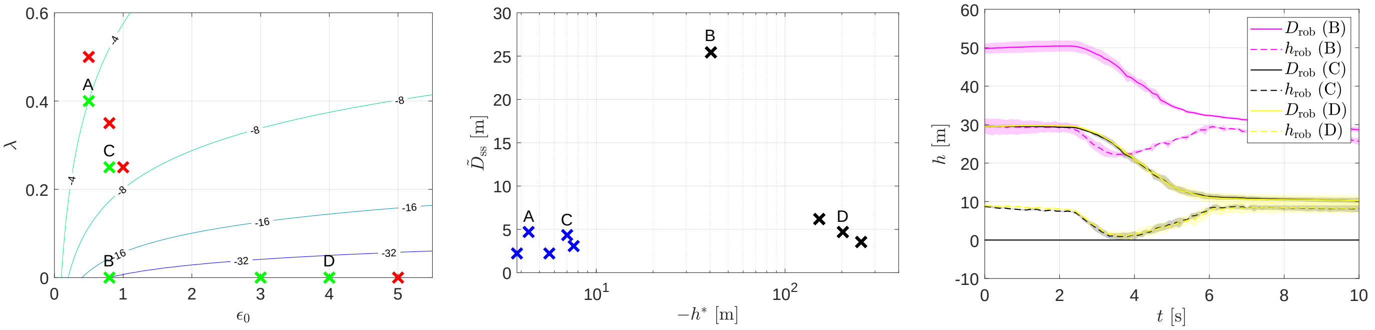

Fig. 16. (Left) Parameter values for ϵ_0 and λ used in the truck experiments, with contours showing theoretical values of h^*. Green markers denote parameter sets which achieve the original safety goal ($h \geq 0$), while red markers denote parameter sets for which the original safety goal is violated. (Center) Theoretical values of h^* and the shift in steady-state tracking distance, denoted by $\tilde{D}_{\rm ss}$, for the parameter sets used in the truck experiments. The blue markers denote parameter sets with $\lambda > 0$, while the black markers denote parameter sets with $\lambda = 0$. (Right) Experimental results for parameter pairs (B), (C) and (D) in Table III. Case (B) is highly conservative as indicated by the large steady-state tracking distance error. Cases (C) and (D) display nearly identical behavior, though case (C) possesses a much stronger theoretical guarantee.

formulation in [32] (where $\lambda = 0$) for sufficiently small values of ϵ_0, but may also be achieved using small values of λ.

We remark that when changing the controller from $k_{\rm QP}$ (cf. (57)) to $k_{\rm rob}$ (cf. (62)) the equilibrium of the system is shifted as can be noticed once comparing the runs on the right panel of Fig. 15. We characterize this by the shift in the steady-state tracking distance error defined as

$$\tilde{D}_{\rm ss} \triangleq D_{\rm ss}^{\rm exp} - D^*. \tag{64}$$

Here $D_{\rm ss}^{\rm exp}$ is the steady-state distance captured in experiments when the leader is moving with the steady-state speed $v^* \in (0, \overline{v}_{\rm L})$ before braking. The term $D^* = V^{-1}(v^*)$ captures the desired steady-state distance given by the inverse of the range policy (54). In the experiments we have $v^* = 16$ [m/s], yielding $D^* = 25$ [m]. The values of $\tilde{D}_{\rm ss}$ corresponding to different parameter pairs are given in Table III. In the right panel of Fig. 16 we visualize the theoretical values of h^* and the experimental values of $\tilde{D}_{\rm ss}$ for different parameter sets. The black markers indicate parameter sets with $\lambda = 0$, while the blue markers show parameter sets with $\lambda > 0$. With $\lambda = 0$, the theoretical guarantees are nearly meaningless (observe the large negative values of h^*), and improving them requires dramatically increasing $\tilde{D}_{\rm ss}$. In contrast, the parameter sets with $\lambda > 0$ allow us to obtain significantly (an order of magnitude) stronger theoretical guarantees without greatly increasing $\tilde{D}_{\rm ss}$, thereby also achieve good performance. In the right panel of Fig. 14 we give experimental results of three other parameter pairs labeled as (B), (C) and (D) in Table III. The poor performance of case (B) is indicated by the large value of $\tilde{D}_{\rm ss}$. The results for cases (C) and (D) nearly overlap, but the introduction of λ allows strong theoretical guarantee for case (C) which is missing for case (D).

VI. Conclusion

In conclusion, this work has developed a theoretically rigorous approach for safety-critical control synthesis through Control Barrier Functions (CBFs). The notion of Input-to-State Safety (ISSf) is utilized to capture the impact of disturbances in the input to the system. A simple parametric modification to CBFs enabled the formulation of ISSf-CBFs as a practical tool for achieving both performant behavior and meaningful theoretical safety guarantees. We provided a tutorial on these tools in the context of an inverted pendulum system, and carried out a practical design problem of a safety-critical controller for a connected automated truck. Moreover, we demonstrated the tangible benefits of the design using ISSf-CBFs by deploying this controller experimentally on an automated truck.

Appendix

A. Proof of Theorem 2

Proof. We first prove that the optimization problem in (18) has a closed-form solution given by (19) and (20), thereby proving it is feasible for any $\mathbf{x} \in \mathbb{R}^n$ and satisfies $\mathbf{k}_{\rm QP}(\mathbf{x}) \in K_{\rm CBF}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Then we prove that $\mathbf{k}_{\rm QP}$ is a continuous function.

Let us first consider an $\mathbf{x} \in \mathbb{R}^n$ such that $L_{\mathbf{g}}h(\mathbf{x}) = \mathbf{0}$. By assumption, the function h is a CBF for (1) on the set $\mathcal{C}$ with corresponding function $\alpha \in \mathcal{K}_\infty^{\rm e}$. Thus, we know from the condition in (10) that

$$L_{\mathbf{f}}h(\mathbf{x}) + \alpha(h(\mathbf{x})) > 0, \tag{65}$$

such that the inequality constraint in (18) is satisfied for any choice of $\mathbf{u}$. The definition of a norm requires that for any $\mathbf{y} \in \mathbb{R}^m$, we have $\|\mathbf{y}\|_2 \geq 0$ and $\|\mathbf{y}\|_2 = 0$ implies $\mathbf{y} = \mathbf{0}$. Thus, we may conclude that the minimizing choice of $\mathbf{u}$ is given by $\mathbf{u} = \mathbf{k}_{\rm n}(\mathbf{x})$, such that $\mathbf{k}_{\rm QP}(\mathbf{x}) = \mathbf{k}_{\rm n}(\mathbf{x})$ as required by the closed-form solution in (19) and (20).

Next let us consider an $\mathbf{x} \in \mathbb{R}^n$ such that $L_{\mathbf{g}}h(\mathbf{x}) \neq \mathbf{0}$. Observe that the cost function and constraint function defining (18) are both convex and continuously differentiable with respect to the decision variable $\mathbf{u}$. Thus the optimization problem is convex, and the *Karush-Kuhn Tucker* (KKT) conditions provide a necessary and sufficient[3] condition for optimality [53, §5.5.3]. More precisely, the KKT conditions state that for

[3] An additional *constraint qualification* is necessary for the KKT conditions to be necessary and sufficient conditions for optimality. One such qualification is *Slater's Condition* [53, §5.2.3], which is easily verified to hold in our setting.

an optimal solution $\mathbf{u}^* \in \mathbb{R}^m$ to (19), we must have a $\mu^* \in \mathbb{R}$ such that:

$$L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{u}^* + \alpha(h(\mathbf{x})) \geq 0, \tag{66}$$

$$\mu^* \geq 0, \tag{67}$$

$$\mu^*(L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{u}^* + \alpha(h(\mathbf{x}))) = 0, \tag{68}$$

$$\mathbf{u}^* - \mathbf{k}_{\rm n}(\mathbf{x}) - \mu^* L_{\mathbf{g}}h(\mathbf{x})^\top = 0. \tag{69}$$

The first and second conditions are referred to as *primal* and *dual feasibility*, respectively. The third condition is referred to as *complementary slackness*, and the fourth condition is referred to as *stationarity*.

Rearranging the stationarity condition (69) yields

$$\mathbf{u}^* = \mathbf{k}_{\rm n}(\mathbf{x}) + \mu^* L_{\mathbf{g}}h(\mathbf{x})^\top. \tag{70}$$

To solve for the value of μ^* (and consequently $\mathbf{u}^*$), we use the primal feasibility condition (66) and the complementary slackness condition (68). In particular, suppose that

$$L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{u}^* + \alpha(h(\mathbf{x})) > 0. \tag{71}$$

The complementary slackness condition (68) then implies $\mu^* = 0$, and thus, we have from (70) that $\mathbf{u}^* = \mathbf{k}_{\rm n}(\mathbf{x})$. Combining this with (71), we obtain

$$L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{k}_{\rm n}(\mathbf{x}) + \alpha(h(\mathbf{x})) > 0. \tag{72}$$

Next let us suppose that

$$L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{u}^* + \alpha(h(\mathbf{x})) = 0. \tag{73}$$

Using the expression for $\mathbf{u}^*$ in (70) yields

$$L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{k}_{\rm n}(\mathbf{x}) + \mu^* \|L_{\mathbf{g}}h(\mathbf{x})\|_2^2 + \alpha(h(\mathbf{x})) = 0, \tag{74}$$

which may be solved for μ^*, yielding

$$\mu^* = -\frac{L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{k}_{\rm n}(\mathbf{x}) + \alpha(h(\mathbf{x}))}{\|L_{\mathbf{g}}h(\mathbf{x})\|_2^2}. \tag{75}$$

Given this expression, the dual feasibility condition (67) requires that, if the equality in (73) holds, we must have

$$L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{k}_{\rm n}(\mathbf{x}) + \alpha(h(\mathbf{x})) \leq 0. \tag{76}$$

Substituting (75) into (70) yields

$$\mathbf{u}^* = \mathbf{k}_{\rm n}(\mathbf{x}) - \frac{L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{k}_{\rm n}(\mathbf{x}) + \alpha(h(\mathbf{x}))}{\|L_{\mathbf{g}}h(\mathbf{x})\|_2^2} L_{\mathbf{g}}h(\mathbf{x})^\top. \tag{77}$$

Noting that (71) and (72) are equivalent, and (73) and (76) are equivalent, we may combine these results with the preceding results obtained for $L_{\mathbf{g}}h(\mathbf{x}) = \mathbf{0}$ and conclude that

$$\mathbf{k}_{\rm QP}(\mathbf{x}) = \mathbf{k}_{\rm n}(\mathbf{x}) + \max\{0, \eta(\mathbf{x})\} L_{\mathbf{g}}h(\mathbf{x})^\top. \tag{78}$$

To show the function $\mathbf{k}_{\rm QP}$ is continuous, let us first define a function $\psi : \mathbb{R}^n \to \mathbb{R}$ as

$$\psi(\mathbf{x}) = L_{\mathbf{f}}h(\mathbf{x}) + L_{\mathbf{g}}h(\mathbf{x})\mathbf{k}_{\rm n}(\mathbf{x}) + \alpha(h(\mathbf{x})). \tag{79}$$

As h is continuously differentiable, and $\mathbf{f}, \mathbf{g}$, and α are continuous, we may conclude that the function ψ is continuous. Consider an arbitrary state $\mathbf{x} \in \mathbb{R}^n$ such that $\psi(\mathbf{x}) > 0$, noting that we may have $L_{\mathbf{g}}h(\mathbf{x}) = \mathbf{0}$. By (72), we have $\mathbf{k}_{\rm QP}(\mathbf{x}) = \mathbf{k}_{\rm n}(\mathbf{x})$. By continuity of ψ, we may conclude that there exists $\delta > 0$ such that $\psi(\mathbf{y}) > 0$ for all $\mathbf{y} \in B_\delta(\mathbf{x})$ (the open ball of radius δ centered at $\mathbf{x}$). By (72) we then have that $\mathbf{k}_{\rm QP}(\mathbf{y}) = \mathbf{k}_{\rm n}(\mathbf{y})$ for all $\mathbf{y} \in B_\delta(\mathbf{x})$. As $\mathbf{k}_{\rm n}$ is continuous, we may conclude that $\mathbf{k}_{\rm QP}$ is continuous at $\mathbf{x}$.

Next consider an arbitrary state $\mathbf{x} \in \mathbb{R}^n$ such that $\psi(\mathbf{x}) < 0$, noting that we must have $L_{\mathbf{g}}h(\mathbf{x}) \neq \mathbf{0}$ at this state. By (76) we have

$$\mathbf{k}_{\rm QP}(\mathbf{x}) = \mathbf{k}_{\rm n}(\mathbf{x}) - \frac{\psi(\mathbf{x})}{\|L_{\mathbf{g}}h(\mathbf{x})\|_2^2} L_{\mathbf{g}}h(\mathbf{x})^\top. \tag{80}$$

By the continuity of $L_{\mathbf{g}}h$ and ψ, we may conclude that there exists a $\delta > 0$ such that $L_{\mathbf{g}}h(\mathbf{y}) \neq \mathbf{0}$ and $\psi(\mathbf{y}) < 0$ for all $\mathbf{y} \in B_\delta(\mathbf{x})$. We then have

$$\mathbf{k}_{\rm QP}(\mathbf{y}) = \mathbf{k}_{\rm n}(\mathbf{y}) - \frac{\psi(\mathbf{y})}{\|L_{\mathbf{g}}h(\mathbf{y})\|_2^2} L_{\mathbf{g}}h(\mathbf{y})^\top. \tag{81}$$

for all $\mathbf{y} \in B_\delta(\mathbf{x})$. As $L_{\mathbf{g}}h$, $\mathbf{k}_{\rm n}$, and ψ are continuous and $L_{\mathbf{g}}h(\mathbf{x}) \neq \mathbf{0}$, we may conclude that $\mathbf{k}_{\rm QP}$ is continuous at $\mathbf{x}$.

Lastly, let us consider a state $\mathbf{x} \in \mathbb{R}^n$ such that $\psi(\mathbf{x}) = 0$, noting that we must have $L_{\mathbf{g}}h(\mathbf{x}) \neq \mathbf{0}$ at this state. By (76) and the fact $\psi(\mathbf{x}) = 0$, we have that $\mathbf{k}_{\rm QP}(\mathbf{x}) = \mathbf{k}_{\rm n}(\mathbf{x})$. Let $\epsilon > 0$ be arbitrary. By the continuity of $L_{\mathbf{g}}h$, we may conclude that there exists $\delta_1 > 0$ such that $L_{\mathbf{g}}h(\mathbf{y}) \neq \mathbf{0}$ for all $\mathbf{y} \in B_{\delta_1}(\mathbf{x})$. Let $\mathbf{y} \in B_{\delta_1}(\mathbf{x})$ be such that $\psi(\mathbf{y}) > 0$. As before, we then have $\mathbf{k}_{\rm QP}(\mathbf{y}) = \mathbf{k}_{\rm n}(\mathbf{y})$. By the continuity of $\mathbf{k}_{\rm n}$, there exists a $\delta_2 > 0$ with $\delta_2 < \delta_1$ such that if $\mathbf{y} \in B_{\delta_2}(\mathbf{x})$ and $\psi(\mathbf{y}) > 0$, then

$$\|\mathbf{k}_{\rm QP}(\mathbf{y}) - \mathbf{k}_{\rm QP}(\mathbf{x})\|_2 = \|\mathbf{k}_{\rm n}(\mathbf{y}) - \mathbf{k}_{\rm n}(\mathbf{x})\|_2 < \epsilon. \tag{82}$$

Next let $\mathbf{y} \in B_{\delta_1}$ be such that $\psi(\mathbf{y}) \leq 0$, such that $\mathbf{k}_{\rm QP}(\mathbf{y})$ is given in (81). Although $\mathbf{k}_{\rm QP}(\mathbf{x}) = \mathbf{k}_{\rm n}(\mathbf{x})$, we may use the fact that $\psi(\mathbf{x}) = 0$ to write $\mathbf{k}_{\rm QP}(\mathbf{x})$ as in (80). By the continuity of $\mathbf{k}_n$, ψ, and $L_{\mathbf{g}}h$, there exists a $\delta_3 > 0$ with $\delta_3 < \delta_1$ such that if $\mathbf{y} \in B_{\delta_3}(\mathbf{x})$ and $\psi(\mathbf{y}) \leq 0$, then

$$\|\mathbf{k}_{\rm n}(\mathbf{y}) - \mathbf{k}_{\rm n}(\mathbf{x})\|_2 < \frac{\epsilon}{2}, \tag{83}$$

and

$$\left\| \frac{\psi(\mathbf{y})}{\|L_{\mathbf{g}}h(\mathbf{y})\|_2^2} L_{\mathbf{g}}h(\mathbf{y})^\top - \frac{\psi(\mathbf{x})}{\|L_{\mathbf{g}}h(\mathbf{x})\|_2^2} L_{\mathbf{g}}h(\mathbf{x})^\top \right\|_2 < \frac{\epsilon}{2}. \tag{84}$$

Therefore we have that if $\mathbf{y} \in B_{\delta_3}(\mathbf{x})$ and $\psi(\mathbf{y}) \leq 0$, then:

$$\|\mathbf{k}_{\rm QP}(\mathbf{y}) - \mathbf{k}_{\rm QP}(\mathbf{x})\|_2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \tag{85}$$

Taking $\delta = \min\{\delta_2, \delta_3\}$, we have that $\mathbf{y} \in B_\delta(\mathbf{x})$ implies:

$$\|\mathbf{k}_{\rm QP}(\mathbf{y}) - \mathbf{k}_{\rm QP}(\mathbf{x})\|_2 < \epsilon, \tag{86}$$

proving $\mathbf{k}_{\rm QP}$ is continuous at $\mathbf{x}$. As we considered the three cases that $\psi(\mathbf{x}) > 0$, $\psi(\mathbf{x})$ <0, and $\psi(\mathbf{x}) = 0$, we have shown the function $\mathbf{k}_{\rm QP}$ is continuous.

□

References

[1] T. Gurriet, M. Mote, A. D. Ames, and E. Feron, "An online approach to active set invariance," in *Conference on Decision & Control (CDC)*. IEEE, 2018, pp. 3592–3599.

[2] K. Hobbs, M. Mote, M. Abate, S. Coogan, and E. Feron, "Run time assurance for safety-critical systems: An introduction to safety filtering approaches for complex control systems," *arXiv preprint arXiv:2110.03506*, 2021.

[3] A. Ames, J. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs with application to adaptive cruise control," in *Conference on Decision & Control (CDC)*. IEEE, 2014, pp. 6271–6278.

[4] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2017.

[5] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *European Control Conference (ECC)*. IEEE, 2019, pp. 3420–3431.

[6] 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 *Conference on Control Technology and Applications (CCTA)*. IEEE, 2017, pp. 1769–1775.

[7] L. Wang, A. D. Ames, and M. Egerstedt, "Safety barrier certificates for collisions-free multirobot systems," *Transactions on Robotics*, vol. 33, no. 3, pp. 661–674, 2017.

[8] L. Wang, E. A. Theodorou, and M. Egerstedt, "Safe learning of quadrotor dynamics using barrier certificates," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2460–2465.

[9] A. Singletary, W. Guffey, T. G. Molnar, R. Sinnet, and A. D. Ames, "Safety-critical manipulation for collision-free food preparation," *arXiv preprint arXiv:2205.01026*, 2022.

[10] W. S. Cortez, D. Oetomo, C. Manzie, and P. Choong, "Control barrier functions for mechanical systems: Theory and application to robotic grasping," *Transactions on Control Systems Technology*, 2019.

[11] R. Grandia, A. J. Taylor, A. D. Ames, and M. Hutter, "Multi-layered safety for legged robots via control barrier functions and model predictive control," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 8352–8358.

[12] N. Csomay-Shanklin, R. K. Cosner, M. Dai, A. J. Taylor, and A. D. Ames, "Episodic learning for safe bipedal locomotion with control barrier functions and projection-to-state safety," *Proceedings of Machine Learning Research (PMLR)*, vol. 144, pp. 1041–1053, 2021.

[13] E. H. Thyri, E. A. Basso, M. Breivik, K. Y. Pettersen, R. Skjetne, and A. M. Lekkas, "Reactive collision avoidance for asvs based on control barrier functions," in *Conference on Control Technology and Applications (CCTA)*. IEEE, 2020, pp. 380–387.

[14] M. L. Mote, "Optimization-based approaches to safety-critical control with applications to space systems," Ph.D. dissertation, Georgia Institute of Technology, 2021.

[15] M. Nagumo, "Über die lage der integralkurven gewöhnlicher differentialgleichungen," *Proceedings of the Physico-Mathematical Society of Japan, 3rd Series*, vol. 24, pp. 551–559, 1942.

[16] F. Blanchini and S. Miani, *Set-theoretic methods in control*. Springer, 2008.

[17] S. Prajna and A. Jadbabaie, "Safety verification of hybrid systems using barrier certificates," in *International Workshop on Hybrid Systems: Computation and Control (HSCC)*. Springer, 2004, pp. 477–492.

[18] R. Konda, A. D. Ames, and S. Coogan, "Characterizing safety: Minimal control barrier functions from scalar comparison systems," *Control Systems Letters*, vol. 5, no. 2, pp. 523–528, 2020.

[19] P. Wieland and F. Allgöwer, "Constructive safety using control barrier functions," *IFAC Proceedings Volumes*, vol. 40, no. 12, pp. 462–467, 2007.

[20] Q. Nguyen and K. Sreenath, "Exponential control barrier functions for enforcing high relative-degree safety-critical constraints," in *American Control Conference (ACC)*. IEEE, 2016, pp. 322–328.

[21] W. Xiao and C. Belta, "Control barrier functions for systems with high relative degree," in *Conference on Decision & Control (CDC)*. IEEE, 2019, pp. 474–479.

[22] X. Tan, W. S. Cortez, and D. V. Dimarogonas, "High-order barrier functions: Robustness, safety and performance-critical control," *IEEE Transactions on Automatic Control*, 2021.

[23] A. Agrawal and K. Sreenath, "Discrete control barrier functions for safety-critical control of discrete systems with application to bipedal robot navigation." in *Robotics: Science and Systems (RSS)*, vol. 13. Cambridge, MA, USA, 2017.

[24] L. Wang, D. Han, and M. Egerstedt, "Permissive barrier certificates for safe stabilization using sum-of-squares," in *American Control Conference (ACC)*, 2018, pp. 585–590.

[25] A. Robey, H. Hu, L. Lindemann, H. Zhang, D. V. Dimarogonas, S. Tu, and N. Matni, "Learning control barrier functions from expert demonstrations," in *Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 3717–3724.

[26] A. Clark, "Verification and synthesis of control barrier functions," in *Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 6105–6112.

[27] P. Glotfelter, J. Cortés, and M. Egerstedt, "Boolean composability of constraints and control synthesis for multi-robot systems via nonsmooth control barrier functions," in *Conference on Control Technology and Applications (CCTA)*. IEEE, 2018, pp. 897–902.

[28] G. Notomista and M. Saveriano, "Safety of dynamical systems with multiple non-convex unsafe sets using control barrier functions," *IEEE Control Systems Letters*, 2021.

[29] 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.

[30] M. Jankovic, "Robust control barrier functions for constrained stabilization of nonlinear systems," *Automatica*, vol. 96, pp. 359–367, 2018.

[31] R. Takano and M. Yamakita, "Robust constrained stabilization control using control Lyapunov and control barrier function in the presence of measurement noises," in *Conference on Control Technology and Applications (CCTA)*. IEEE, 2018, pp. 300–305.

[32] S. Kolathaya and A. D. Ames, "Input-to-state safety with control barrier functions," *Control Systems Letters*, vol. 3, no. 1, pp. 108–113, 2018.

[33] K. Garg and D. Panagou, "Robust control barrier and control Lyapunov functions with fixed-time convergence guarantees," in *American Control Conference (ACC)*, 2021.

[34] A. Alan, A. J. Taylor, C. R. He, G. Orosz, and A. D. Ames, "Safe controller synthesis with tunable input-to-state safe control barrier functions," *Control Systems Letters*, vol. 6, pp. 908–913, 2022.

[35] J. J. Choi, D. Lee, K. Sreenath, C. J. Tomlin, and S. L. Herbert, "Robust control barrier–value functions for safety-critical control," in *Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 6814–6821.

[36] S. Dean, A. J. Taylor, R. K. Cosner, B. Recht, and A. D. Ames, "Guaranteeing safety of learned perception modules via measurement-robust control barrier functions," in *Conference on Robotics Learning (CoRL)*, 2020.

[37] P. Seiler, M. Jankovic, and E. Hellstrom, "Control barrier functions with unmodeled input dynamics using integral quadratic constraints," *Control Systems Letters*, vol. 6, pp. 1664–1669, 2021.

[38] J. Buch, S.-C. Liao, and P. Seiler, "Robust control barrier functions with sector-bounded uncertainties," *IEEE Control Systems Letters*, vol. 6, pp. 1994–1999, 2021.

[39] E. D. Sontag, "Input to state stability: Basic concepts and results," in *Nonlinear and Optimal Control Theory*. Springer, 2008, pp. 163–220.

[40] N. Sridhar, K. B. Devika, S. C. Subramanian, G. Vivekanandan, and S. Sivaram, "Antilock brake algorithm for heavy commercial road vehicles with delay compensation and chattering mitigation," *Vehicle system dynamics*, vol. 59, no. 4, pp. 526–546, 2021.

[41] L. Perko, *Differential equations and dynamical systems*. Springer Science & Business Media, 2013, vol. 7.

[42] C. M. Kellett, "A compendium of comparison function results," *Mathematics of Control, Signals, and Systems*, vol. 26, no. 3, pp. 339–374, 2014.

[43] S. S. Sastry, *Nonlinear Systems: Analysis, Stability and Control*. NY: Springer, 1999.

[44] R. M. Murray, Z. Li, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC press, 1994.

[45] H. K. Khalil and J. W. Grizzle, *Nonlinear Systems*. Prentice Hall, 2002.

[46] C. R. He and G. Orosz, "Safety guaranteed connected cruise control," in *International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 549–554.

[47] G. Orosz, "Connected cruise control: modelling, delay effects, and nonlinear behaviour," *Vehicle System Dynamics*, vol. 54, no. 8, pp. 1147–1176, 2016.

[48] J. I. Ge, S. S. Avedisov, C. R. He, W. B. Qin, M. Sadeghpour, and G. Orosz, "Experimental validation of connected automated vehicle design among human-driven vehicles," *Transportation Research part C*, vol. 91, pp. 335–352, 2018.

[49] Video Supplement for "Control Barrier Functions and Input-to-State Safety with Application to Automated Vehicles". [Online]. Available: https://youtu.be/9dJtC1TCBbA

[50] International-Trucks, "Prostar+ truck," https://www.internationaltrucks.com/trucks/prostar.

[51] Commsignia, "V2X onboard unit," https://www.commsignia.com/products/obu/.

[52] Speedgoat, "Mobile real-time target machine," https://www.speedgoat.com/products-services/real-time-target-machines/mobile.

[53] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

Anil Alan received the BSc degree in mechanical engineering from Middle East Technical University, Turkey, in 2012 and the MSc degree in Bilkent University, Turkey, in 2017. He is currently pursuing the PhD degree in mechanical engineering with the University of Michigan, Ann Arbor, MI, USA. His current research interests include control of connected autonomous vehicles, safety-critical control, nonlinear control, vehicle dynamics.

Andrew J. Taylor received the B.S. and M.S. degrees in aerospace engineering from the University of Michigan, Ann Arbor, in 2016 and 2017, respectively. He is currently pursuing a Ph.D. degree at California Institute of Technology in Control and Dynamical Systems. His research interests include safety-critical control for robotic systems and data-driven control techniques for nonlinear systems.

Chaozhe R. He received the BSc degree in applied mathematics from the Beijing University of Aeronautics and Astronautics in 2012, the MSc and PhD in Mechanical Engineering from the University of Michigan, Ann Arbor, USA, in 2015 and 2018 respectively. Dr. He is with Plus.ai Inc. and is working on planning and control algorithm development. His research interests include dynamics and control of connected automated vehicles, optimal and nonlinear control theory, and data-driven control.

Aaron D. Ames is the Bren Professor of Mechanical and Civil Engineering and Control and Dynamical Systems at Caltech. Prior to joining Caltech in 2017, he was an Associate Professor at Georgia Tech in the Woodruff School of Mechanical Engineering and the School of Electrical & Computer Engineering. He received a B.S. in Mechanical Engineering and a B.A. in Mathematics from the University of St. Thomas in 2001, and he received a M.A. in Mathematics and a Ph.D. in Electrical Engineering and Computer Sciences from UC Berkeley in 2006. He served as a Postdoctoral Scholar in Control and Dynamical Systems at Caltech from 2006 to 2008, and began his faculty career at Texas A&M University in 2008. At UC Berkeley, he was the recipient of the 2005 Leon O. Chua Award for achievement in nonlinear science and the 2006 Bernard Friedman Memorial Prize in Applied Mathematics, and he received the NSF CAREER award in 2010, the 2015 Donald P. Eckman Award, and the 2019 IEEE CSS Antonio Ruberti Young Researcher Prize. His research interests span the areas of robotics, nonlinear, safety-critical control and hybrid systems, with a special focus on applications to dynamic robots -— both formally and through experimental validation.

Gábor Orosz received the M.Sc. degree in Engineering Physics from the Budapest University of Technology, Hungary, in 2002 and the Ph.D. degree in Engineering Mathematics from University of Bristol, UK, in 2006. He held postdoctoral positions at the University of Exeter, UK, and at the University of California, Santa Barbara. In 2010, he joined the University of Michigan, Ann Arbor where he is currently an Associate Professor in Mechanical Engineering and in Civil and Environmental Engineering. His research interests include nonlinear dynamics and control, time delay systems, and reinforcement learning with applications to connected and automated vehicles, traffic flow, and biological networks.