Unmatched Control Barrier Functions: 
Certainty Equivalence Adaptive Safety

Brett T. Lopez¹ and Jean-Jacques E. Slotine²

Abstract—This work applies universal adaptive control to control barrier functions to achieve safe control of dynamical systems with parametric model uncertainties. The proposed approach utilizes the certainty equivalence principle to methodically select a model-parameterized control barrier function and corresponding safety controller from an allowable set with instantaneous parameter estimates. While such a combination does not necessarily yield forward invariance without additional requirements on the barrier function, we show that safety can indeed be established by simply adjusting the adaptation gain online. Simulation results demonstrate the approach.

I. INTRODUCTION

Safety-critical controllers rely on precise model knowledge to ensure forward invariance of a safe set. Although these controllers often possess some inherent robustness [1], techniques that guarantee safety while effectively compensating for model uncertainties with minimal conservatism have only recently been proposed. One such framework is based on the notion of adaptive safety — a paradigm that achieves forward invariance of a safe set using results from adaptive control and model estimation theory. The work by [2] introduced the adaptive control barrier function (aCBF), and showed that subsets of a safe set were forward invariant when an aCBF was used to construct controllers. The conservatism of an aCBF was addressed by [3] through the so-called robust adaptive control barrier function (RaCBF). There are two key differences between an aCBF and RaCBF. Firstly, an RaCBF yields less conservative controllers as the system is allowed to approach the boundary of a tightened safe set. Secondly, an RaCBF can be combined with model estimation to further reduce conservatism if monotonic reduction in the model uncertainty can be established. Due to its effectiveness and strong theoretical guarantees, the adaptive safety paradigm has seen several recent extensions (see, e.g., [4]–[7]).

¹Verifiable and Control-Theoretic Robotics Laboratory, University of California, Los Angeles, Los Angeles, CA, btlepez@ucla.edu
²Nonlinear Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA, jjs@mit.edu

A fundamental limitation of current adaptive safety approaches is the inability to employ the certainty equivalence principle when the barrier function depends on uncertain model parameters. This dependency typically arises when the system possess unmatched model uncertainties, i.e., unknown dynamics outside the span of the control input matrix, but can also occur when the uncertainties are matched. The certainty equivalency principle is a design philosophy that entails constructing barrier functions as if the model were known. The uncertain parameters are then simply replaced with their online estimates. This philosophy works seamlessly when the barrier function does not depend on uncertain parameters. Conversely, it is very difficult to establish forward invariance of a parameter-dependent barrier function as sign-indefinite terms appear when trying to prove forward invariance. To cancel out the sign-indefinite terms, [2], [3] construct a barrier function for a modified system that depends on its own (unknown) barrier function. It will be shown that generating such a barrier function is more difficult than just dealing with the problematic terms directly.

The main contribution of this work is an adaptive safety algorithm which employs the certainty equivalence principle to achieve set invariance through online parameter adaptation. Unlike previous works, one just needs to construct a family of control barrier functions for all possible models — a much simpler procedure than that originally proposed by [2], [3]. It is shown that forward invariance of a safe set is achieved with a novel direct adaptation law that systematically adjusts the adaptation gain online; a technique recently developed for adaptive control with unmatched uncertainties [8] that has also found uses in other areas of control, e.g., direct adaptive optimal control [9]. Furthermore, the direct adaptation law can be combined with model learning to improve parameter adaptation transients and reduce conservatism. Simulation results of adaptive cruise control with uncertain model parameters demonstrates the effectiveness of the approach.

Notation: The set of positive and strictly-positive scalars...
will be denoted as \( \mathbb{R}_+ \) and \( \mathbb{R}_{>0} \), respectively. The shorthand notation for a function \( T \) parameterized by a vector \( a \) with vector argument \( s \) will be \( T_a(s) \equiv T(s;a) \). The partial differentiation with respect to variable \( x \in \mathbb{R}^n \) of function \( N(x,y) \) will be \( \nabla_x N(x,y) = \partial N/\partial x \in \mathbb{R}^n \). The subscript for \( \nabla \) will be omitted when it is clear which variable the differentiation is with respect to.

**II. Problem Formulation**

Consider the uncertain nonlinear system

\[
\dot{x} = f(x) - \Delta(x)^T \theta + g(x)u,
\]

with state \( x \in \mathbb{R}^n \), control input \( u \in \mathcal{U} \subseteq \mathbb{R}^m \), known dynamics \( f : \mathbb{R}^n \to \mathbb{R}^n \), unknown parameters \( \theta \in \Theta \subseteq \mathbb{R}^p \) with known regressor \( \Delta : \mathbb{R}^n \to \mathbb{R}^{p \times n} \), and known control input matrix \( g : \mathbb{R}^n \to \mathbb{R}^{n \times m} \). In this work we derive an adaptive safety controller that ensures \( x(t) \in \mathcal{C} \) for all time where \( \mathcal{C} \) is a set of safe states. The following assumption is made on the unknown parameters \( \theta \).

**Assumption 1.** The unknown parameters \( \theta \) belong to a known closed and bounded set \( \Theta \subseteq \mathbb{R}^p \).

An immediate consequence of Assumption 1 is that the parameter estimation error \( \hat{\theta} \triangleq \hat{\theta} - \theta \) must also belong to a known closed and bounded set, i.e., \( \hat{\Theta} \subseteq \Theta \). Moreover, each element must then have a finite maximum where \( \hat{\theta} \equiv \max_{\hat{\theta} \in \hat{\Theta}} \hat{\theta}_i \) for \( i = 1, \ldots, p \). Note that \( \hat{\theta} \) is equivalent to the maximum possible parameter estimation error.

**III. Background: Adaptive Safety**

The following definitions are stated for completeness, more detail can be found in [10], [11] and references therein. See [2], [3] for the first works on adaptive safety.

Let the set \( \mathcal{C} \subseteq \mathbb{R}^n \) be a 0-superlevel set of a continuously differentiable function \( h : \mathbb{R}^n \to \mathbb{R} \) where

\[
\mathcal{C} = \{ x \in \mathbb{R}^n : h(x) \geq 0 \}, \quad \partial \mathcal{C} = \{ x \in \mathbb{R}^n : h(x) = 0 \}, \quad \text{Int} (\mathcal{C}) = \{ x \in \mathbb{R}^n : h(x) > 0 \}.
\]

The following definitions assume the dynamics (1) are Lipschitz (at least locally) so that there exists a unique solution \( x(t) \) for \( t \geq t_0 \) with initial condition \( x_0 \triangleq x(t_0) \).

**Definition 1.** The set \( \mathcal{C} \) is forward invariant if for every \( x_0 \in \mathcal{C} \), \( x(t) \in \mathcal{C} \) for all \( t \geq t_0 \).

**Definition 2.** A system is safe with respect to set \( \mathcal{C} \) if the set \( \mathcal{C} \) is forward invariant.

**Definition 3.** A continuous function \( \alpha : \mathbb{R} \to \mathbb{R} \) is an extended class \( \mathcal{K}_\infty \) function if it is strictly increasing, \( \alpha(0) = 0 \), and is defined on the entire real line.

When model uncertainty is present, as is the case in (1), it is challenging or infeasible to derive a controller that renders a safe set forward invariant. Conceptually, adaptive safety is a framework that uses tools from adaptive control theory to systematically compute a safe controller via online parameter adaptation. Central to adaptive safety is the notion of model-parameterized safe sets \( \mathcal{C}_0 \) which are shown to be forward invariant with an aCBF [2] or RaCBF [3]. Due to the similarities between [3] and this work, only the core results from [3] are summarized below.

**Definition 4 (Robust Adaptive CBF [3]).** Let \( \mathcal{C}_0 \) be a family of 0-superlevel sets parameterized by \( \theta \) for a continuously differentiable function \( h_\theta^0 : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R} \). The function \( h_\theta^0(x) \) is a robust adaptive control barrier function (RaCBF) if there exists a controller \( u \in \mathcal{U} \) and extended class \( \mathcal{K}_\infty \) function \( \alpha(\cdot) \) such that for every \( \theta \in \Theta \)

\[
\sup_{u \in \mathcal{U}} \left\{ \nabla_x h_\theta^0^\top [f - \Delta^\top \Lambda_\theta + gu] \right\} \geq -\alpha(h_\theta^0),
\]

where \( \Lambda_\theta(x) \triangleq \theta - \gamma \nabla_\theta h_\theta^0(x) \) and \( \gamma \) is an admissible adaptation gain.

Note that an RaCBF is designed for a modified system that depends upon its own RaCBF via \( \Lambda_\theta(x) \). This is the main drawback of using an RaCBF (and an aCBF as the same modification is employed). Fundamentally, the modified dynamics are a byproduct of attempting to use the certainty equivalence principle with model-parameterized safe sets. It can be shown that the extra term is actually related to the adaptation law derived for an RaCBF [3]. Therefore, an RaCBF is constructed to account for parameter adaptation transients and hence represents a departure from the certainty equivalence principle philosophy. In this work we will show that a true certainty equivalence adaptive safety framework is possible when the adaptation gain is adjusted online.

**IV. Main Results**

**A. Overview**

This section contains the main results of this work. First, it is shown that set invariance is achieved by combining the so-called unmatched control barrier function and direct adap-
tive control with online adaptation gain adjustment. Then, two extensions, namely, the leakage modification and life-long model estimation, that improve transients and reduce conservatism are presented. Finally, a unified adaptive safety and tracking controller is discussed.

In the sequel we will make use of a special class functions called scaling functions.

**Definition 5 (Scaling Function).** A scaling function \( v : \mathbb{R} \to \mathbb{R} \) for \( \zeta > 1 \) and \( \rho \in \mathbb{R} \subseteq \mathbb{R} \) satisfies the conditions
\[
1 \leq v(\rho) \leq \zeta < \infty,
\]
\[
\nabla v(\rho) > 0.
\]

**Remark 1.** One example of a suitable scaling function is \( v(\rho) = \arctan(\rho) + 1 \) where \( \rho \in [0, 10] \).

**B. Direct Adaptive Safety**

We will consider two safe sets defined by a continuously differentiable function \( h_{\theta} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R} \), namely \( \mathcal{C}_\theta \triangleq \{ x \in \mathbb{R}^n, \theta \in \Theta : h_{\theta}(x) \geq 0 \} \) and \( \mathcal{C}_\theta^0 \triangleq \{ x \in \mathbb{R}^n, \theta \in \Theta : h_{\theta}(x) \geq \frac{1}{\gamma} \hat{\theta}^T \hat{\theta} \} \) where \( \mathcal{C}_\theta^0 \subseteq \mathcal{C}_\theta \).

**Definition 6 (Unmatched CBF).** Let \( \mathcal{C}_\theta \) be a family of 0-superlevel sets parameterized by \( \theta \) for a continuously differentiable function \( h_{\theta} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R} \). The function \( h_{\theta}(x) \) is an unmatched control barrier function (uCBF) if there exists a controller \( u \in \mathbb{U} \) and extended class \( \mathcal{K}_\infty \) function \( \alpha(\cdot) \) such that for every \( \theta \in \Theta \)
\[
\sup_{u \in \mathbb{U}} \left\{ \nabla_x h_{\theta}^T \left[ f - \Delta^T \theta + gu \right] \right\} \geq -\alpha(h_{\theta}).
\]

**Remark 2.** In addition to \( \alpha(\cdot) \in \mathcal{K}_\infty \), we require that \( \alpha(\alpha(r)) \leq \alpha(cr) \) for \( c \geq 1 \). This property is not restrictive as it satisfied by many common choices for \( \alpha(\cdot) \).

Fundamentally, condition (3) states that there exists a controller that renders \( \mathcal{C}_\theta \) invariant for every \( \theta \in \Theta \). Or, put another way, the uncertain system (1) can be made safe for every \( \theta \in \Theta \). This is analogous to an uncertain system being stabilizable for every \( \theta \in \Theta \) in the context of adaptive control, as discussed in [8]. Note that (3) is an invariance condition for \( \mathcal{C}_\theta \) with the actual dynamics, as opposed to the modified dynamics used in [2], [3]. This distinction has both theoretical and practical implications. In particular, safety is now formulated as an inherent property of the system since the actual dynamics are being evaluated for safety. Moreover, in terms of constructing an uCBF, the safety condition (3) is much easier to verify as it preserves bilinearity of \( h_{\theta}(x) \) and \( u \) [11]. Conversely, the condition for an aCBF or RaCBF is nonconvex so systematically constructing either barrier function is more difficult.

The following theorem establishes forward invariance of a parameter-dependent safe set \( \mathcal{C}_\theta \) when Definition 6 is combined with direct parameter adaptation and online adjustment of the adaptation gain.

**Theorem 1.** Let \( \mathcal{C}_\theta \) be a 0-superlevel set of a continuously differentiable function \( h_{\theta} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R} \). If \( h_{\theta}(x) \) is an uCBF on \( \mathcal{C}_\theta^0 \triangleq \{ x \in \mathbb{R}^n, \theta \in \Theta : h_{\theta}(x) \geq \frac{1}{\gamma} \hat{\theta}^T \hat{\theta} \} \subseteq \mathcal{C}_\theta \) then any locally Lipschitz continuous controller satisfying
\[
\sup_{u \in \mathbb{U}} \left\{ \nabla_x h_{\theta}^T \left[ f - \Delta^T \theta + gu \right] \right\} \geq -\alpha \left( h_{\theta} - \frac{1}{\gamma} \hat{\theta}^T \hat{\theta} \right),
\]
renders \( \mathcal{C}_\theta \) safe with the direct adaptation law
\[
\dot{\theta} = \gamma v(\rho) \Delta(x) \nabla_x h_{\theta}(x)
\]
\[
\dot{\rho} = -\frac{v(\rho)}{\nabla v(\rho)} (v(\rho) h_{\theta}(x) + \eta) \nabla v(\rho) h_{\theta}(x)^T \hat{\theta}
\]
where \( \gamma \) is an admissible adaptation gain, \( v(\rho) \) is a scaling function, and \( \eta \in \mathbb{R}_{>0} \) is a design parameter.

**Proof.** Consider the candidate barrier function
\[
h(t) = v(\rho) \left( h_{\theta}(x) + \eta \right) - \frac{1}{\gamma} \hat{\theta}^T \hat{\theta},
\]
where \( \eta > 0 \). We will show that the adaptation law (5) yields \( h_{\theta}(x(t)) \geq 0 \) for all \( t \) which is equivalent to \( \mathcal{C}_\theta \) being forward invariant. Differentiating \( h(t) \) and applying (5) yields
\[
\dot{h}(t) = v(\rho) \nabla_x h_{\theta}(x)^T \left[ f(x) - \Delta(x)^T \theta + g(x)u \right]
\]
\[
+ v(\rho) \nabla \hat{\theta} h_{\theta}(x)^T \hat{\theta} + \rho \nabla v(\rho) (h_{\theta}(x) + \eta) - \frac{1}{\gamma} \hat{\theta}^T \hat{\theta}
\]
\[
= v(\rho) \nabla_x h_{\theta}(x)^T \left[ f(x) - \Delta(x)^T \theta + g(x)u \right]
\]
\[
\geq -\alpha \left( v(\rho) h_{\theta}(x) - v(\rho) \frac{1}{\gamma} \hat{\theta}^T \hat{\theta} \right),
\]
where the inequality arises from (4) and the property \( c\alpha(r) \leq \alpha(cr) \) for \( c \geq 1 \). Since \( |\theta| \leq \hat{\theta} \) and \( v(\rho) \geq 1 \) then
\[
h(t) \geq v(\rho) \left( h_{\theta}(x) + \eta \right) - \frac{1}{\gamma} \hat{\theta}^T \hat{\theta}
\]
\[
\geq v(\rho) \left( h_{\theta}(x) + \eta - \frac{1}{\gamma} \hat{\theta}^T \hat{\theta} \right),
\]
yielding \( \dot{h}(t) \geq -\alpha (h(t) - v(\rho) \eta) \) which implies \( h(t) \geq v(\rho) \eta > 0 \) for all \( t \geq 0 \) if \( h(0) \geq v(\rho(0)) \eta \). Since \( v(\rho) \eta \leq h(t) \leq v(\rho) (h_{\theta}(x) + \eta) \), then \( h_{\theta}(x(t)) \geq 0 \) for all \( t \). Therefore, the controller (4) and direct adaptation law (5) render the set \( \mathcal{C}_\theta \) forward invariant. \( \square \)

**Remark 3.** The adaptation law (5a) updates \( \hat{\theta} \) toward values
that improve safety. For instance, if $\nabla_x h_\theta(x) \Delta(x)^T > 0$ then $\dot{\hat{\theta}} > 0$ eventually yielding $\tilde{\theta} \geq 0$. Then, $\nabla_x h_\theta(x) \Delta(x)^T \tilde{\theta} \geq 0$ which acts as an extra safety term in $\dot{h}_\theta(x)$. This is analogous to direct adaptive control where $\hat{\theta}$ is driven to values that reduce tracking error.

Remark 4. An immediate modification to (5a) is the projection operator $\text{Proj}_\Theta(\cdot)$ to enforce $\hat{\theta} \in \Theta$ without affecting safety [12], [13]. It can be shown that temporarily stopping adaptation when $\hat{\theta} \in \partial \Theta$ will only introduce positive terms in $\dot{h}(t)$ so the inequalities in Theorem 1 are unchanged.

It is instructive to analyze the online gain adjustment to develop an intuition about how the technique achieves forward invariance. In the case of no gain adjustment, i.e., $\dot{\rho} = 0$, then $\dot{h}(t)$ becomes

$$\dot{h}(t) \geq -\alpha(h(t) - v(\rho) \eta) + v(\rho) \nabla_\theta h_\theta(x)^T \dot{\hat{\theta}}.$$  

If $\nabla_\theta h_\theta(x)^T \dot{\hat{\theta}} \geq 0$ then safety is preserved as the same inequality used to prove Theorem 1 is obtained. However, if $\nabla_\theta h_\theta(x)^T \dot{\hat{\theta}} < 0$ then safety might be compromised since this could lead to $h(t) < 0$ and subsequently $h_\theta(x) < 0$. From (5), we see that $\hat{\rho}$ will be of opposite sign of $\nabla_\theta h_\theta(x)^T \dot{\hat{\theta}}$. Hence, if the parameter adaptation transients is negative (unsafe), then $\rho$ increases resulting in a larger effective adaptation gain $\gamma v(\rho)$. Conversely, if the transients is positive (safe) then $\rho$ decreases yielding a smaller effective adaptation gain. In this scenario, one could also set $\dot{\rho} = 0$ without sacrificing safety. To summarize, the effective adaptation gain $\gamma v(\rho)$ will increase if the parameter adaptation transients would otherwise compromise forward invariance of $C_\theta$ while $\gamma v(\rho)$ will decrease or remain constant if the parameter adaptation transients preserves safety. The above analysis of adaptation gain adjustment is analogous to that in universal adaptive control [8] where the effective adaptation gain changes to achieve a stable closed-loop system.

Theorem 1 requires the adaptation gain $\gamma$ be admissible in order for $C_\theta$ to be forward invariant. The following corollary establishes a lower bound on $\gamma$ thereby making it admissible.

**Corollary 1.** An admissible adaptation gain $\gamma$ for the adaptation law in Theorem 1 satisfies the inequality

$$\gamma \geq \frac{\dot{\gamma}^T \dot{\hat{\theta}}}{2h_\theta(x_0)},$$

where $h_\theta(x_0) \triangleq h(x(0), \hat{\theta}(0))$.

*Proof.* Theorem 1 established $h_\theta(x(t)) \in C_\theta$ uniformly if $h(0) \geq v(\rho) \eta$. We will show this condition is satisfied with (6). Using the definition of $h(t)$,

$$h(0) = v(\rho_0) (h_\theta(x_0) + \eta) - \frac{1}{\sqrt{2}} \theta_0 \dot{\theta}_0,$$

$$\geq h_\theta(x_0) + \eta - \frac{1}{\sqrt{2}} \theta_0^T \dot{\theta}_0$$

where we have chosen $v(\rho_0) = 1$. Hence, $h(0) \geq v(\rho_0) \eta = \eta \iff \gamma \geq \frac{\dot{\gamma}^T \dot{\hat{\theta}}}{2h_\theta(x_0)}$, thus yielding (6).

The lower bound for the adaptation gain is identical to that obtained in [3] and similar to that in [2] (the initial parameter estimation error must be known in the latter). Essentially, (6) states that the closer $h_\theta(x_0)$ is to $\partial C_\theta$ the faster the adaptation has to be in order to render $C_\theta$ invariant [2], [3].

An interesting consequence of Theorem 1 is that the set $C_\theta^r = \{ x \in \mathbb{R}^n, \hat{\theta} \in \Theta : h_\theta(x) \geq \frac{1}{\sqrt{2}} \theta_0^T \dot{\theta}_0 \}$ is input-to-state safe [14] with the proposed adaptive safety controller, as shown in the following corollary.

**Corollary 2.** The set $C_\theta^r = \{ x \in \mathbb{R}^n, \hat{\theta} \in \Theta : h_\theta(x) \geq \frac{1}{\sqrt{2}} \theta_0^T \dot{\theta}_0 \}$ is input-to-state safe (ISSf) with the controller and adaptation law in Theorem 1.

*Proof.* Follows immediately from the definition of ISSf which states that a set is ISSf if it is a subset of a forward invariant set. Since the controller and adaptation law render $C_\theta$ invariant, and $C_\theta^r \subseteq C_\theta$, then $C_\theta^r$ is ISSf.

**Remark 5.** Future work will investigate strengthening Corollary 2 to show that the set $C_\theta^r$ is asymptotically stable, as is the case with an RaCBF (see the Appendix of [15]).

**C. Leakage Modification**

Depending on the choice of $v(\rho)$, one may need to modify (5) in order for $\rho$ to remain bounded. One possibility is to reset $\rho$ once it exceeds a certain threshold. Even though $u$, $\hat{\theta}$, and $h_\theta(x)$ remain continuous after a reset, a thorough analysis is required to ensure the closed-loop system remains safe despite the barrier-like function $h(t)$ decreasing after the reset. Alternatively, one could add damping to $\dot{\rho}$, thereby bounding $\rho$. The following corollary shows that this leakage modification renders $C_\theta$ ISSf.

**Corollary 3.** Let $C_\theta$ be a 0-superlevel set of a continuously differentiable function $h_\theta : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}$. If $h_\theta(x)$ is an uCBF on $C_\theta^r \triangleq \{ x \in \mathbb{R}^n, \hat{\theta} \in \Theta : h_\theta(x) \geq \frac{1}{\sqrt{2}} \theta_0^T \dot{\theta}_0 \} \subseteq C_\theta$ then any locally Lipschitz continuous controller satisfying

$$\sup_{u \in U} \left\{ \nabla_x h_\theta^T \left[ f - \Delta^T \dot{\theta} + gu \right] \right\} \geq -\alpha \left( h_\theta - \frac{1}{\sqrt{2}} \theta_0^T \dot{\theta}_0 \right),$$

(4)
renders $C_{\hat{\theta}}$ input-to-state safe with the adaptation law
\begin{align}
\dot{\hat{\theta}} &= \gamma v(\rho) \Delta(x) \nabla x h_{\hat{\theta}}(x), \\
\dot{\rho} &= v(\rho) \left( \frac{1}{\nabla v(\rho) h_{\hat{\theta}}(x)} + \eta \left[ -\sigma \rho + w_{\hat{\theta}}(x) \right] \right),
\end{align}
(7a) (7b)
where
\[ w_{\hat{\theta}}(x) = \begin{cases} 0 & \text{if } \nabla h_{\hat{\theta}}(x)^{\top} \hat{\theta} \geq 0 \\ -\zeta \nabla h_{\hat{\theta}}(x)^{\top} \left[ \gamma \Delta(x) \nabla x h_{\hat{\theta}}(x) \right] & \text{otherwise}, \end{cases} \]
(8)
and $\gamma$ is an admissible adaptation gain, $v(\rho)$ is a scaling function, and $\eta, \sigma \in \mathbb{R}_{>0}$ are design parameters.

Proof. We must first establish that $\rho$ is bounded from above and non-negative before showing $C_{\hat{\theta}}$ is ISSf. Let $d_{\hat{\theta}}(x, \rho) \triangleq \frac{\sigma v(\rho) 1}{\nabla v(\rho) h_{\hat{\theta}}(x) + \eta}$ which is strictly positive based on Definitions 5 and 6 and $\sigma, \eta \in \mathbb{R}_{>0}$. Since (7b) is a stable first order filter with a bounded input $w_{\hat{\theta}}(x)$ (under the premise $x, \hat{\theta}$ are bounded and $\Delta(x), h_{\hat{\theta}}(x)$ are continuously differentiable), then $\rho$ must remain bounded. First consider the simple case where $\nabla h_{\hat{\theta}}(x)^{\top} \hat{\theta} \geq 0$. As noted above, this scenario yields the same forward invariance inequality used in the proof of Theorem 1 so $w_{\hat{\theta}}(x)$ can be trivially set to zero. If $\rho > 0$ then $\rho \rightarrow 0$ exponentially with rate $d(x, \rho)$ since $w_{\hat{\theta}}(x) = 0$. Conversely, if $\nabla h_{\hat{\theta}}(x)^{\top} \hat{\theta} < 0$ then $w_{\hat{\theta}}(x) + \nabla h_{\hat{\theta}}(x)^{\top} \hat{\theta} > 0$ which implies that $w_{\hat{\theta}}(x) > 0$ and subsequently $\rho > 0$ if $\rho(0) > 0$. Hence, $\rho \geq 0$ and is bounded from above.

Following identical steps to those taken in the proof of Theorem 1, one obtains $\dot{h}(t) \geq -\alpha \left( h(t) - v(\rho) \rho \right) - \sigma v(\rho) \rho$ where $\sigma(v(\rho) \rho \geq 0$ and is bounded. This yields $h(t) \geq v(\rho) \rho - \alpha^{-1}(\sigma v(\rho) \rho) \Rightarrow h_{\hat{\theta}}(x) \geq \frac{1}{\alpha} \sigma^{-1}(\sigma v(\rho) \rho) \geq 0$.

Lemma 1. Let $C_{\hat{\theta}}$ be a $\delta$-superlevel set for a continuously differentiable function $h_{\hat{\theta}} : \mathbb{R}^{n} \times \mathbb{R}^{p} \rightarrow \mathbb{R}$, i.e., $C_{\hat{\theta}} \triangleq \{ x \in \mathbb{R}^{n}, \hat{\theta} \in \Theta : h_{\hat{\theta}}(x) \geq \delta \}$ where $\delta \geq 0$. If $h_{\hat{\theta}}(x)$ is an uCBF on $C_{\hat{\theta}}$ then it is also an uCBF on $C_{\hat{\theta}} \supseteq C_{\hat{\theta}}$ where $C_{\hat{\theta}}$ is the 0-superlevel set of $h_{\hat{\theta}}(x)$.

Proof. If $h_{\hat{\theta}}(x)$ is an uCBF on $C_{\hat{\theta}}$ then there exists a controller $u$ and extended class $K_{\infty}$ function $\alpha(\cdot)$ such that $h_{\hat{\theta}}(x) \geq -\alpha(h_{\hat{\theta}}(x) - \delta)$. Since $\delta \geq 0$ then $h_{\hat{\theta}}(x) \geq -\alpha(h_{\hat{\theta}}(x))$. Therefore, $h_{\hat{\theta}}(x)$ is also a valid uCBF on $C_{\hat{\theta}}$.

Theorem 2. Let $C_{\hat{\theta}}$ be a 0-superlevel set of a continuously differentiable function $h_{\hat{\theta}} : \mathbb{R}^{n} \times \mathbb{R}^{p} \rightarrow \mathbb{R}$. If the model uncertainty $\hat{\theta}(t)$ monotonically decreases via a suitable model estimator and $h_{\hat{\theta}}(x)$ is an uCBF on $C_{\hat{\theta}} \triangleq \{ x \in \mathbb{R}^{n}, \hat{\theta} \in \Theta(0) : h_{\hat{\theta}}(x) \geq \frac{1}{\gamma_{\hat{\theta}}} \hat{\theta}(0)^{\top} \hat{\theta}(0) \} \subseteq C_{\hat{\theta}}$ then any locally Lipschitz continuous controller satisfying
\[
\sup_{u \in U} \left\{ \nabla x h_{\hat{\theta}}^{\top} \left[ f - \Delta^{\top} \hat{\theta} + gu \right] \right\} \geq -\alpha \left( h_{\hat{\theta}} - \frac{1}{\gamma_{\hat{\theta}}} \hat{\theta}(t)^{\top} \hat{\theta}(t) \right),
\]
renders $C_{\hat{\theta}}$ safe with (5) and the suitable model estimator.

Proof. Let $C_{\hat{\theta}}(t) \triangleq \{ x \in \mathbb{R}^{n}, \hat{\theta} \in \Theta(0) : h_{\hat{\theta}}(x) \geq \frac{1}{\gamma_{\hat{\theta}}} \hat{\theta}(0)^{\top} \hat{\theta}(0) \} \subseteq C_{\hat{\theta}}(t)$ for all $t > 0$. If $h_{\hat{\theta}}(x)$ is an uCBF on $C_{\hat{\theta}}$ then it is also an uCBF on $C_{\hat{\theta}}(t)$ by Lemma 1. Forward invariance of $C_{\hat{\theta}}$ then follows the arguments in Theorem 1.
E. Safe Tracking Control

An unmatched CBF can be immediately combined with an unmatched CLF [8] for safe stabilizing controller \( \kappa : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{U} \) that depends on parameter estimates \( \hat{\phi} \) and \( \hat{\theta} \) computed for tracking and safety, respectively. Either quadratic program (QP) formulation discussed in [11] can be used to compute \( \kappa \) online with \( \hat{\phi} \) and \( \hat{\theta} \) being generated by their respective adaptation laws.

V. SIMULATION: ADAPTIVE CRUISE CONTROL

The developed method was tested on the adaptive cruise control (ACC) problem with uncertain dynamics given by

\[
\dot{d} = v_r - v \\
\dot{v} = -\theta_1 - \theta_2 v - \theta_3 v^2 + u,
\]

where \( d \) is the distance between two cars, \( v_r \) is the speed of the lead car (uncontrolled), \( v \) is the speed of the following car (controlled), and \( \theta = [\theta_1 \ \theta_2 \ \theta_3]^T \) is unknown but belongs to a closed and bounded set \( \Theta \). An intuitive choice for a safe set is \( C_\theta = \{[d \ v]^T \in \mathbb{R}^2, \theta \in \Theta : d - d^*_\theta(v) - d_b \geq 0 \} \) where \( d^*_\theta(v) \) is the stopping distance of the follower, which depends on the unknown parameters despite the uncertainty being matched (a phenomenon also seen in adaptive optimal control [9]), and \( d_b \) is a small distance buffer. Practically, the safety constraint requires the distance between the two cars be greater or equal to the stopping distance of the follower plus the additional distance buffer. The considered scenario is as follows. The follower is commanded to travel at a speed \( v_d > v_r \) and the leader initially has a non-zero speed \( v_r > 0 \) up until a point where the leader stops \( v_r = 0 \). The goal is to ensure the follower never collides with the leader. At each time step, the stopping distance \( d^*_\theta(v) \) is computed numerically by forward simulating a constant deceleration profile \( u = -U_{in} \). A safe tracking controller is computed via the safety filter QP formulation where the adaptive safety controller modifies a simple velocity tracking controller. All simulation parameters can be found in the Appendix.

Fig. 1 shows the performance of three safety controllers: the proposed adaptive safety controller without (blue) and with (green) model estimation (see Appendix) and a safety controller that uses “nominal” parameter values (black), i.e., no online adaptation. Both adaptive safety controllers use the leakage modification so the effective adaptation gain \( \gamma v(\rho) \) returns to its nominal value after transients. The \( u \)CBF \( h^u(x) \) is shown in Fig. 1a where both adaptive safety controllers ensure \( \partial C_{\theta} \) (red) is never breached despite the presence of model uncertainties. The lack of conservatism of the safety controller with model estimation is noteworthy, and further highlights the benefit of combining adaptive control and model estimation to maximize performance as originally shown in [3]. Conversely, the controller with no adaptation not only violates the safety constraint, but actually causes the two vehicles to collide; as seen in Fig. 1b where \( d \) become negative. Both adaptive safety controllers maintain a safe distance albeit with different levels of conservatism dictated by how confident the controller is in the unknown parameters. The scaling function \( v(\rho) \) is also shown in Fig. 1c where a sharp increase is seen at the beginning of the simulation, translating to a higher effective adaptation gain \( \gamma v(\rho) \). From the discussion in Section IV, this indicates the initial model learning transients would have negatively affected safety if the adaptation gain was not adjusted. Following the initial increase, \( v(\rho) \to 1 \) per the leakage modification so \( \gamma v(\rho) \to \gamma \) as desired.

\( ^2 \)As in [3], we utilize the technique known as set membership identification [19], [20] to estimate the bounds on the uncertain parameters.
VI. CONCLUDING REMARKS

A new adaptive safety framework was presented that permits the use of the certainty equivalence principle for systematic online selection of a controller that renders a safe set forward invariant despite the presence of unmatched parametric uncertainties. The safe combination of policy selection and direct parameter adaptation was achieved by online adjustment of the adaptation gain inspired by [8]. The ability to employ the certainty equivalence principle significantly reduces the complexities associated with existing adaptive safety approaches without sacrificing strong theoretical guarantees.

VII. APPENDIX

A. Simulation Parameters

| Parameter | Value |
|-----------|-------|
| $\Theta$ | $[-4, 4] \text{ m/s}^2 \times [0.5, 1.2] / \text{s} \times [0.001, 0.03] / \text{m}$ |
| $\theta$ | $[-3 \text{ m/s}^2, 0.8 / \text{s}, 0.004 / \text{m}]$ |
| $v_0$ | $30 \text{ m/s}$ |
| $v_1$ | $20 \text{ m/s}$ |
| $d_i$ | $100 \text{ m}$ |
| $U_{max}$ | $30 \text{ m/s}^2$ |
| $\alpha(r)$ | $r$ |
| $\eta(r)$ | $\arctan(r) + 1$ |
| $\gamma$ | $1$ |
| $\sigma$ | $100 \text{ rad/s}$ |
| $\Delta t$ | $0.001 \text{ s}$ |

B. Set Membership Identification (SMID)

Set membership identification estimates the interval bounds for each parameter by constructing a set of unfalsified parameters values from a history of observations. The method can be formulated as a linear program, and has been shown to yield interval bounds that monotonically approach the true parameter value even if disturbances or noise are present. See, e.g., [19]–[23] and references therein.

REFERENCES

[1] 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.
[2] A. J. Taylor and A. D. Ames, “Adaptive safety with control barrier functions,” in 2020 American Control Conference (ACC), pp. 1399–1405, IEEE, 2020.
[3] B. T. Lopez, J.-J. E. Slotine, and J. P. How, “Robust adaptive control barrier functions: An adaptive and data-driven approach to safety,” IEEE Control Systems Letters, vol. 5, no. 3, pp. 1031–1036, 2020.
[4] M. Maghenem, A. J. Taylor, A. D. Ames, and R. G. Sanfelice, “Adaptive safety using control barrier functions and hybrid adaptation,” in 2021 American Control Conference (ACC), pp. 2418–2423, IEEE, 2021.
[5] A. Isaly, O. S. Patil, R. G. Sanfelice, and W. E. Dixon, “Adaptive safety with multiple barrier functions using integral concurrent learning,” in 2021 American Control Conference (ACC), pp. 3719–3724, IEEE, 2021.
[6] M. Black, E. Arabi, and D. Panagou, “A fixed-time stable adaptation law for safety-critical control under parametric uncertainty,” in 2021 European Control Conference (ECC), pp. 1328–1333, IEEE, 2021.
[7] M. H. Cohen and C. Delta, “High order robust adaptive control barrier functions and exponentially stabilizing adaptive control lyapunov functions,” arXiv:2203.01999, 2022.
[8] B. T. Lopez and J.-J. E. Slotine, “Universal adaptive control of nonlinear systems,” IEEE Control Systems Letters, vol. 6, no. 2, pp. 1826–1830, 2021.
[9] B. Lopez and J.-J. Slotine, “Adaptive variants of optimal feedback policies,” in Learning for Dynamics and Control Conference, pp. 1125–1136, PMLR, 2022.
[10] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, “Control barrier function based quadratic programs for safety critical systems,” IEEE Transactions on Automatic Control, vol. 62, no. 8, pp. 3861–3876, 2016.
[11] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, “Control barrier functions: Theory and applications,” in 2019 18th European Control Conference (ECC), pp. 3420–3431, IEEE, 2019.
[12] J.-J. E. Slotine and W. Li, Applied nonlinear control. Prentice Hall, 1991.
[13] P. A. Ioannou and J. Sun, Robust adaptive control. Courier Corporation, 2012.
[14] S. Kolathaya and A. D. Ames, “Input-to-state safety with control barrier functions,” IEEE Control Systems Letters, vol. 3, no. 1, pp. 108–113, 2018.
[15] B. T. Lopez and J.-J. Slotine, “Unmatched control barrier functions: Certainty equivalence adaptive safety,” arXiv:2207.13873, 2022.
[16] P. A. Ioannou and P. V. Kokotovic, “Instability analysis and improvement of robustness of adaptive control,” Automatica, vol. 20, no. 5, pp. 583–594, 1984.
[17] P. Ioannou and K. Tsakalis, “A robust direct adaptive controller,” IEEE Transactions on Automatic Control, vol. 31, no. 11, pp. 1033–1043, 1986.
[18] K. S. Narendra and A. M. Annaswamy, Stable adaptive systems. Courier Corporation, 2012.
[19] R. L. Kosut, M. K. Lau, and S. P. Boyd, “Set-membership identification of systems with parametric and nonparametric uncertainty,” IEEE Transactions on Automatic Control, vol. 37, no. 7, pp. 929–941, 1992.
[20] M. Tanaskovic, L. Fagiano, R. Smith, and M. Morari, “Adaptive receding horizon control for constrained mimo systems,” Automatica, vol. 50, no. 12, pp. 3019–3029, 2014.
[21] M. Lorenzen, F. Allgöwer, and M. Cannon, “Adaptive model predictive control with robust constraint satisfaction,” IFAC-PapersOnLine, vol. 50, no. 1, pp. 3313–3318, 2017.
[22] B. T. Lopez, Adaptive robust model predictive control for nonlinear systems. PhD thesis, Massachusetts Institute of Technology, 2019.
[23] B. Bujarbaruah, X. Zhang, M. Tanaskovic, and F. Borrelli, “Adaptive stochastic mpc under time-varying uncertainty,” IEEE Transactions on Automatic Control, vol. 66, no. 6, pp. 2840–2845, 2020.