An Error-Based Safety Buffer for Safe Adaptive Control (Extended Version)

Peter A. Fisher, Johannes Autenrieb, and Anuradha M. Annaswamy

October 28, 2025

Abstract

We consider the problem of adaptive control of a class of feedback linearizable plants with matched parametric uncertainties whose states are accessible, subject to state constraints, which often arise due to safety considerations. In this paper, we combine adaptation and control barrier functions into a real-time control architecture that guarantees stability, ensures control performance, and remains safe even with the parametric uncertainties. Two problems are considered, differing in the nature of the parametric uncertainties. In both cases, the control barrier function is assumed to have an arbitrary relative degree. In addition to guaranteeing stability, it is proved that both the control objective and safety objective are met with near-zero conservatism. No excitation conditions are imposed on the command signal. Simulation results demonstrate the non-conservatism of all of the theoretical developments.

I Introduction

Real time control of dynamic systems with parametric uncertainties so as to ensure high performance and safety is a challenging task. The field of adaptive control has focused on providing real-time inputs for dynamic systems through parameter learning and control design using a stability framework [1–6]. While these works have accommodated parametric uncertainties, with extensions to magnitude and rate constraints on the input, consideration of constraints on the state, often necessitated by safety considerations, has been absent for the most part. In settings where the plant is fully known, Control Barrier Functions (CBFs) [7–11] have gained prominence as a minimally-invasive tool for enforcing state constraints that ensure safety. This paper proposes two new approaches for simultaneously realizing both a control objective and a safety objective in the presence of parametric uncertainties. The class of dynamic systems we focus on is feedback-linearizable and time-invariant, with states accessible, and our approaches combine an adaptive control architecture with a new CBF-based controller.

The main challenges in realizing the two objectives in a combined manner are to ensure that state constraints are satisfied during transients in the closed-loop adaptive system, and that the control objectives are fully met in steady-state. The inherent nonlinearity in the transients is the main difficulty in the first challenge; stringent conditions needed to reduce the effect of the parametric uncertainty cause the second challenge. Unlike the current literature, which address these two challenges only with compromises and conservatism, our approaches provably ensure that both of these challenges are fully overcome and constitute a clear advance of the state of the art in safe and stable adaptive control.

The state of the art in addressing problems where both parametric uncertainties and state constraints are present can be found in [12–30]. Robust control designs have been proposed in [12–15] which offset the state constraints with buffer windows according to worst-case bounds on modeling uncertainties. The papers [16–20] have sought to learn from data to reduce conservatism. Disturbance observer-based approaches were proposed in [21–24] which treat the total effect of the parametric uncertainties on the dynamics as a state-dependent disturbance.

More recently, [25–28] have explored solutions which take an online, adaptive approach to safety. In [25], adaptive Control Barrier Functions (aCBFs) are proposed based on Control Lyapunov functions. Robust Adaptive Control Barrier Functions (RaCBFs) are proposed in [26, 27], where the aCBF approach is altered to remove jitter and enable online reduction in conservatism through data-driven methods. The results of [28] extend aCBFs to include input certainties. The advantage of these methods over others mentioned in [12–24] is that they can be less conservative, require no offline training, and avoid the need for high gains as in disturbance observers. Conservatism still remains, however, in these three papers, which can be removed only with persistent excitation.

The approach we propose in this paper avoids conservatism, does not require availability of data or offline training time, avoids any gain requirements, and allows both the control and safety objectives to be met. This approach builds on our past work in [29, 30] both of which constructed a CBF-based calibration of a reference input with constraints stemming from a reference model that is linear and time-invariant. This reference input introduced an error-based relaxation term that had the potential for ensuring safety during the adaptive transient period. With this calibrated reference input, stability and safety were guaranteed in [29, 30] assuming that the trajectories were within the safe set while the transients were present. In this paper, we introduce two new approaches that utilize two different error-based components for calibration of the reference input as well as the constraints. These approaches remove the assumptions in our earlier work, guarantee stability of all trajectories that start in the safe set, and ensure that the safety objective and control objective are met without any conservatism.

The contributions of this paper are the following: two approaches for ensuring simultaneous realization of safety and control objectives for linear time-invariant plants with parametric uncertainties are proposed whose states are accessible, one based on an Error-Based Safety Buffer (EBSB) and another based on an Error-Based Safety Filter (EBSF) with the latter applicable when input uncertainties are present as well. In all cases, stability and safety are guaranteed with no persistent conservatism in the safety constraints and without any need for high gains or persistent excitation. Furthermore, our designs accommodate CBFs with relative degree $r \geqslant 1$.

In Section II, we provide preliminaries and two problem statements. Section III presents EBSB, which addresses Problem 1, and Section IV presents EBSF, which addresses Problem 2. Section V discusses how to augment both EBSB and EBSF to take advantage of excitation if it is present, and Section VI demonstrates the efficacy of EBSB and EBSF in simulation. Proofs and supporting material can be found in the appendices.

II Preliminaries and Statement of the Problem

II.A Definitions and Notation

A few definitions are proposed in this section, pertaining to safety in dynamical systems of the form

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

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, and $\mathbf{f}(\mathbf{x})$ and $G(\mathbf{x})$ are locally Lipschitz. Consider a closed set $S \subset \mathbb{R}^n$ described by a smooth function $h(\mathbf{x})$ as follows:

$$S = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geqslant 0\}, \tag{2a}$$
$$\partial S = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = 0\}, \tag{2b}$$
$$\mathrm{int}(S) = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) > 0\} \tag{2c}$$

where ∂S and $\mathrm{int}(S)$ denote the boundary and interior of S respectively.

The following definitions from existing literature will be used in this paper:

Definition 1 ($\mathcal{K}_{\infty,e}$ [7]). *A continuous function $\alpha : \mathbb{R} \to \mathbb{R}$ is an extended class-$\mathcal{K}_\infty$ function (denoted $\alpha \in \mathcal{K}_{\infty,e}$) if it is strictly increasing, $\alpha(0) = 0$, $\lim_{z\to\infty} \alpha(z) = \infty$, and $\lim_{z\to-\infty} \alpha(z) = -\infty$.*

Definition 2 (Control Barrier Function (CBF) [7]). *Consider a smooth function $h : \mathbb{R}^n \to \mathbb{R}$ describing a set S as in (2a)-(2c). Then, h is a control barrier function for the dynamics in (1) on S if there exists an $\alpha \in \mathcal{K}_{\infty,e}$ such that, for every $\mathbf{x} \in S$, there exists a $\mathbf{u} \in \mathbb{R}^m$ satisfying*

$$\mathcal{L}_{\mathbf{f}} h(\mathbf{x}) + \mathcal{L}_G h(\mathbf{x})\mathbf{u} \geqslant -\alpha(h(\mathbf{x})), \tag{3}$$

where $\mathcal{L}_{\mathbf{f}} h$ and $\mathcal{L}_G h$ represent the Lie derivatives of h with respect to $\mathbf{f}$ and G respectively.

Definition 3 (Forward-Invariance [7]). *A control policy $\mathbf{u} = \mathbf{k}(\mathbf{x}, t)$ renders the dynamics in (1) forward-invariant in S if, for any $\mathbf{x}(0) \in S$, the policy results in $\mathbf{x}(t) \in S$ $\forall t \geqslant 0$.*

Definition 4 (Safety [7]). *A control policy $\mathbf{u} = \mathbf{k}(\mathbf{x}, t)$ renders the dynamics in (1) safe with respect to S if the policy renders the dynamics forward-invariant in S.*

Definition 5 (Relative Degree [31]). *Consider a function $h : \mathbb{R}^n \to \mathbb{R}$ describing a set S as in (2a)-(2c). Then, h has relative degree $r \geqslant 1$ with respect to the dynamics in (1) if $\mathcal{L}_G h(\mathbf{x}) = \mathcal{L}_G \mathcal{L}_f^1 h(\mathbf{x}) = \cdots = \mathcal{L}_G \mathcal{L}_f^{r-2} h(\mathbf{x}) = 0$ $\forall \mathbf{x} \in \mathbb{R}^n$, and $\mathcal{L}_G \mathcal{L}_f^{r-1} h(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ except on a measure-zero subset of $\mathbb{R}^n$.*

In addition, the following definition is useful:

Definition 6 (Tracking). *Given two time-varying signals $\mathbf{y}(t), \mathbf{z}(t) \in \mathbb{R}^d$, $\mathbf{y}$ is said to asymptotically track $\mathbf{z}$ with error M if there exists a finite $M \geqslant 0$ such that $\limsup_{t\to\infty} \|\mathbf{y}(t) - \mathbf{z}(t)\| = M$.*

In addition to standard notations, the $d \times d$ identity matrix is denoted as I_d, and a $p \times q$ zero matrix is denoted as $0_{p\times q}$. For a vector $\mathbf{a} \in \mathbb{R}^d$, $\mathrm{diag}(\mathbf{a}) \in \mathbb{R}^{d\times d}$ is the diagonal matrix with the elements of $\mathbf{a}$ on the diagonal. Finally, we denote the orthogonal projection of a vector $\mathbf{a} \in \mathbb{R}^n$ to a set $\mathcal{C} \subset \mathbb{R}^n$ as $\mathrm{proj}_{\mathcal{C}}[\mathbf{a}] = \arg\min_{\mathbf{b}\in\mathcal{C}} \|\mathbf{b} - \mathbf{a}\|$.

II.B Tangent Cones

In this section, we briefly introduce tangent cones, will be useful in our adaptive control design later on. The following is a property of closed, convex sets:

Lemma 1. *Any closed, convex set $\mathcal{C} \subset \mathbb{R}^n$ with nonzero n-dimensional volume can be expressed as $\mathcal{C} = \{\mathbf{v} \in \mathbb{R}^n : \min\{g_1(\mathbf{v}), g_2(\mathbf{v}), \dots\} \geqslant 0\}$, where each $g_k : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable with $\nabla g_k(\mathbf{v}) \neq 0$ wherever $g_k(\mathbf{v}) = 0$.*

Proof. See Appendix I.A. □

We now use Lemma 1 to define the tangent cone of a closed, convex set $\mathcal{C}$ as follows:

Definition 7 (Tangent Cone). *For a closed, convex set $\mathcal{C} \subset \mathbb{R}^n$ with nonzero n-dimensional volume, the tangent cone of $\mathcal{C}$ at a vector $\mathbf{v} \in \mathbb{R}^n$ is given by*

$$T_{\mathcal{C}}(\mathbf{v}) = \begin{cases} \mathbb{R}^n, & \mathbf{v} \in \operatorname{int}(\mathcal{C}) \\ T_{\partial\mathcal{C}}(\mathbf{v}), & \mathbf{v} \in \partial\mathcal{C} \\ \varnothing, & \mathbf{v} \notin \mathcal{C} \end{cases} \tag{4}$$

where

$$T_{\partial\mathcal{C}}(\mathbf{v}) = \left\{\mathbf{u} \in \mathbb{R}^n : \nabla g_k(\mathbf{v})^\top \mathbf{u} \geqslant 0 \ \forall k \text{ s.t. } g_k(\mathbf{v}) = 0\right\}. \tag{5}$$

and $g_1, g_2, \dots$ are the functions describing $\mathcal{C}$ as in Lemma 1.

Remark 1. *Definition 7 is equivalent to Definition 4 in [29] for a closed, convex set with nonzero n-dimensional volume.*

Finally, the following is a useful property of tangent cones:

Lemma 2. *Define $T_{\mathcal{C}}(\mathbf{v})$, the tangent cone of a closed, convex set $\mathcal{C} \subset \mathbb{R}^n$ with nonzero n-dimensional volume, as in Definition 7. Then, for any $\mathbf{v}, \mathbf{w} \in \mathcal{C}$ and any $\mathbf{z} \in \mathbb{R}^n$, we have $(\mathbf{v}-\mathbf{w})^\top \operatorname{proj}_{T_{\mathcal{C}}(\mathbf{v})}[\mathbf{z}] \leqslant (\mathbf{v}-\mathbf{w})^\top \mathbf{z}$.*

Proof. See Appendix I.B. □

II.C Statement of the Problem

The class of dynamic systems we consider in this paper is of the form

$$\dot{\mathbf{x}} = A\mathbf{x} + B(\Lambda\mathbf{u} - F(\mathbf{x})\theta_*) \tag{6}$$

where the state $\mathbf{x} \in \mathbb{R}^n$ is available for measurement, $\mathbf{u} \in \mathbb{R}^m$ is the input, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $F : \mathbb{R}^n \to \mathbb{R}^{m\times p}$ are known, F is bounded for bounded $\mathbf{x}$, and (A, B) is controllable. The system in (6) also includes parametric uncertainties θ_* and Λ, where $\theta_* \in \mathbb{R}^p$ and $\Lambda = \operatorname{diag}(\lambda_*)$ for an uncertain $\lambda_* \in \mathbb{R}^m$. In addition, a closed set S as in (2a)-(2c) and smooth function $h : \mathbb{R}^n \to \mathbb{R}$ with relative degree $r \geqslant 1$ with respect to (6) are assumed to be given. The problem addressed in this paper is the design of $\mathbf{u}$ in (6) so that a safety objective and control objective are met, which are defined as follows: the *safety objective* is to ensure that $\mathbf{x}(t) \in S \ \forall t \geqslant 0$. For the control objective, we define a desired trajectory $\mathbf{x}_*(t) \in \mathbb{R}^n$ as the solution to the dynamical system

$$\dot{\mathbf{x}}_* = A_m\mathbf{x}_* + B\mathbf{r}_* \tag{7}$$

for any bounded reference input $\mathbf{r}_*(t) \in \mathbb{R}^m$, where A_m is Hurwitz and $A_m = A + BK$ for some $K \in \mathbb{R}^{m\times n}$. Then, the *control objective* is to ensure that $\|\mathbf{x}(t) - \mathbf{x}_*(t)\|$ is bounded and that $\mathbf{x}$ asymptotically tracks $\mathbf{x}_*$ with an error M. The overall problem is therefore the design of $\mathbf{u}$ in (6) so as to meet the safety and control objectives simultaneously.

Our particular focus is on the following two problems that will be addressed in Sections III and IV respectively:

Problem 1 (Uncertain Unforced Dynamics). *Design an adaptive controller that meets the safety and control objectives for the plant in (6) with Λ known and θ_* unknown.*

Problem 2 (Uncertain Input Matrix). *Design an adaptive controller that meets the safety and control objectives for the plant in (6) with both Λ and θ_* unknown.*

III Safe Adaptive Control for Uncertain Unforced Dynamics

In this section, we address Problem 1. Without loss of generality, we assume that $\Lambda = I_m$, so that the plant in (6) becomes

$$\dot{\mathbf{x}} = A\mathbf{x} + B(\mathbf{u} - F(\mathbf{x})\theta_*). \tag{8}$$

In order to meet both the control and safety objectives, we introduce a reference model of the form

$$\dot{\mathbf{x}}_m = A_m\mathbf{x}_m + B\mathbf{r}_s \tag{9}$$

where A_m is defined in (7) and $\mathbf{r}_s$ is a locally Lipschitz reference signal. We will first design $\mathbf{u}$ as an adaptive controller so that $\mathbf{x}$ asymptotically tracks $\mathbf{x}_m$. This first step is done in Section III.A. Then, we will design $\mathbf{r}_s$ using h such that it renders both the reference model in (9) and the overall adaptive system safe with respect to S, and such that $\mathbf{x}_m$ asymptotically tracks $\mathbf{x}_*$. This second step in done in Section III.B.

We introduce the following assumption on the nature of the parametric uncertainty:

Assumption 1. *$\theta_* \in \Theta$ for a known compact, convex set $\Theta \subset \mathbb{R}^p$ with nonzero p-dimensional volume.*

III.A Adaptive Control Design

Recalling that $A_m = A + BK$, the adaptive controller is given by

$$\mathbf{u} = K\mathbf{x} + \mathbf{r}_s + F(\mathbf{x})\widehat{\theta}. \tag{10}$$

Combining (8)-(10) and defining $\mathbf{e}_x := \mathbf{x} - \mathbf{x}_m$ and $\widetilde{\theta} := \widehat{\theta} - \theta_*$, we obtain a standard error model [1]:

$$\dot{\mathbf{e}}_x = A_m\mathbf{e}_x + BF(\mathbf{x})\widetilde{\theta}. \tag{11}$$

Finally, the adaptive law for updating $\widehat{\theta}$ is given by

$$\dot{\widehat{\theta}} = \mathrm{proj}_{T_\Theta(\widehat{\theta})}[-\gamma F(\mathbf{x})^\top B^\top P\mathbf{e}_x] \tag{12}$$

where $T_\Theta(\widehat{\theta})$ is the tangent cone of Θ at $\widehat{\theta}$, $\gamma > 0$ is any adaptive gain, and P is the symmetric positive-definite solution to the Lyapunov equation

$$A_m^\top P + PA_m = -Q \tag{13}$$

for any symmetric positive-definite matrix Q.

We establish key properties of the adaptive control design in the following theorem, which follows from standard arguments in adaptive control [1]:

Theorem 1. *The closed-loop adaptive system consisting of* (8)-(12) *under Assumption 1 with $\widehat{\theta}(0) \in \Theta$ has the following properties:*

(a) $\widehat{\theta}(t) \in \Theta\ \forall t \geqslant 0$, *and*

(b) $\|\mathbf{e}_x(t)\| \leqslant E_1\ \forall t \geqslant 0$,

where E_1 is any positive constant satisfying $E_1^2 \geqslant \frac{\lambda_{max}(P)}{\lambda_{min}(P)}\|\mathbf{e}_x(0)\|^2 + \frac{1}{\gamma}\sup_{\theta\in\Theta}\|\widehat{\theta}(0) - \theta\|^2)$. Additionally, if $\mathbf{r}_s(t)$ is bounded, then

(c) $\mathbf{x}(t)$ *and* $\mathbf{x}_m(t)$ *are bounded, and*

(d) $\lim_{t\to\infty}\|\mathbf{e}_x(t)\| = 0$.

Proof. See Appendix II.A. □

III.B Error-Based Safety Buffer

We now design $\mathbf{r}_s$ so that (i) the reference model is safe with a suitable choice of $\mathbf{r}_s$ (Proposition 1), and (ii) an appropriately designed safety buffer Δ_{ebsb} ensures that the safety objective and the control objectives are met (Theorem 2). The following assumption is useful:

Assumption 2. *There exist constants $\kappa, c > 0$ such that $\|\frac{\partial h}{\partial \mathbf{x}}\| \leqslant \kappa \ \forall \mathbf{x} \in \mathcal{C} \subset \mathbb{R}^n$, where $\mathcal{C}$ is the convex hull of $\{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geqslant -c\}$.*

As h has relative degree $r \geqslant 1$ with respect to (8), we will first construct a sequence of high-order CBFs [31], as well as a sequence of matrices that encode the relative degree property of each high-order CBF. These matrices will be used to define additional assumptions, and will be used in our solution to Problem 1. The high-order CBF sequence is defined as follows:

$$h_1(\mathbf{x}) = h(\mathbf{x}) \tag{14a}$$

$$h_i(\mathbf{x}) = \frac{\partial h_{i-1}}{\partial \mathbf{x}}\Big|_{\mathbf{x}} A_m \mathbf{x} + \alpha_{i-1} h_{i-1}(\mathbf{x}) \ \forall i \in [2, r] \tag{14b}$$

for any constants $\alpha_1, \ldots, \alpha_{r-1} > 0$. Define $S_i = \{\mathbf{x} \in \mathbb{R}^n : h_i(\mathbf{x}) \geqslant 0\}$, and note that $S_1 = S$. Additionally, define a subspace $\mathbb{I}_h \subseteq \mathbb{R}^n$ as

$$\mathbb{I}_h = \left\{ \mathbf{v} \in \mathbb{R}^n : \exists \mathbf{x} \in \mathbb{R}^n \text{ where } \frac{\partial h}{\partial \mathbf{x}} \parallel \mathbf{v}^\top \right\} \tag{15}$$

and suppose without loss of generality that $\mathbb{I}_h$ is spanned by the orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_q\}$ for some $q \in [1, n]$. Then, we will make use of the following matrix recursion:

$$C_h = \sum_{j=1}^{q} \mathbf{w}_j \mathbf{w}_j^\top \tag{16a}$$

$$W_0 = \kappa^2 C_h^\top C_h, \tag{16b}$$

$$W_i = (\alpha_1 + \alpha_i) W_{i-1} + A_m^\top W_{i-1} + W_{i-1} A_m \ \forall i \in [1, r]. \tag{16c}$$

Finally, we establish key properties of the matrices in (16a)-(16c) in the following lemmas:

Lemma 3. *Assumption 2 implies that $|h(\mathbf{x}_1) - h(\mathbf{x}_2)| \leqslant \sqrt{(\mathbf{x}_1 - \mathbf{x}_2)^\top W_0 (\mathbf{x}_1 - \mathbf{x}_2)} \ \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$.*

Proof. See Appendix II.B. □

Lemma 4. *If $r > 1$, then the matrices W_i in (16b)-(16c) satisfy $W_i B = 0 \ \forall i < r - 1$.*

Proof. See Appendix II.C. □

With the definitions and lemmas above established, we propose the following choice of $\mathbf{r}_s$:

$$\begin{gathered} \mathbf{r}_s = \arg\min_{\mathbf{r} \in \mathbb{R}^m} \|\mathbf{r} - \mathbf{r}_*\|^2 \text{ s.t.} \\ \frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}_m} (A_m \mathbf{x}_m + B\mathbf{r}) \geqslant -\alpha_r h_r(\mathbf{x}_m) + \delta + \Delta_{ebsb}(\mathbf{x}_m, \mathbf{e}_x, \widehat{\theta}) \end{gathered} \tag{17}$$

for any constants $\alpha_r, \delta > 0$ and an Error-Based Safety Buffer (EBSB) denoted by $\Delta_{ebsb} : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$. The form of Δ_{ebsb} differs slightly depending on the relative degree: for $r = 1$,

$$\Delta_{ebsb}(\mathbf{x}_m, \mathbf{e}_x, \widehat{\theta}) = \max \left\{ \frac{\mathbf{e}_x^\top W_1 \mathbf{e}_x + 2\Psi(\widehat{\theta}) \|F(\mathbf{x})^\top B^\top W_0 \mathbf{e}_x\|}{2h_1(\mathbf{x}_m)}, 0 \right\}, \tag{18}$$

and for $r > 1$,

$$\Delta_{ebsb}(\mathbf{x}_m, \mathbf{e}_x, \widehat{\theta}) = \max \left\{ \frac{\mathbf{e}_x^\top W_r \mathbf{e}_x + 2\Psi(\widehat{\theta}) \|F(\mathbf{x})^\top B^\top W_{r-1} \mathbf{e}_x\|}{2h_1(\mathbf{x}_m)} - \frac{h_2(\mathbf{x}_m) h_r(\mathbf{x}_m)}{h_1(\mathbf{x}_m)}, 0 \right\}, \tag{19}$$

where

$$\Psi(\widehat{\theta}) = \sup_{\theta \in \Theta} \|\widehat{\theta} - \theta\|. \tag{20}$$

In order to satisfy the safety objective, we require the following assumption on the initial conditions of the closed-loop adaptive system:

Assumption 3. *The initial conditions of the closed-loop adaptive system satisfy:*

(a) $h_1(\mathbf{x}_m(0)) \geqslant \max\{\frac{\delta}{\prod_{k=1}^{r}\alpha_k}, \sqrt{\mathbf{e}_x(0)^\top W_0 \mathbf{e}_x(0)}\}$,

(b) $h_i(\mathbf{x}_m(0)) \geqslant \max\{\frac{\delta}{\prod_{k=i}^{r}\alpha_k}, \frac{\mathbf{e}_x(0)^\top W_{i-1}\mathbf{e}_x(0)}{2h_1(\mathbf{x}_m(0))}\}$
$\forall i \in [2, r]$,

(c) $h(\mathbf{x}(0)) \geqslant 0$, *and*

(d) $\hat{\theta}(0) \in \Theta$.

Proposition 1 shows that (17) renders the reference model in (9) safe with respect to a subset of S, and additionally shows that $h_1(\mathbf{x}_m(t)) > 0\ \forall t \geqslant 0$:

Proposition 1. *The choice of* $\mathbf{r}_s$ *in* (26) *under Assumption 3 ensures that, for every* $i \in [1, r]$*, we have* $h_i(\mathbf{x}_m(t)) \geqslant \frac{\delta}{\prod_{k=i}^{r}\alpha_k}\ \forall t \geqslant 0$.

Proof. See Appendix II.D. □

We now present our main result for Problem 1. First, in order to show that $\mathbf{r}_s$ remains bounded, including at locations where the gradient of h_r vanishes, we require the following additional technical assumptions:

Assumption 4. *The highest-order safe set* $S_r \subset \mathbb{R}^n$ *is bounded.*

Assumption 5. *There exists a constant* $d > 0$ *such that, at every* $\mathbf{x}_1 \in S_r$ *where* $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_1} B\| < d$*, we have* $\alpha_r h_r(\mathbf{x}_1) + \frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_1} A_m \mathbf{x}_1 \geqslant \delta + \overline{\Delta}_{ebsb}(\mathbf{x}_1)$*, where* $\overline{\Delta}_{ebsb}(\mathbf{x}_1) = \sup_{\theta \in \Theta, \mathbf{e} \in \mathbb{R}^n : \|\mathbf{e}\| \leqslant E_1} \Delta_{ebsb}(\mathbf{x}_1, \mathbf{e}, \theta)$ *and* E_1 *is defined in Theorem 1.*

Our main result for Problem 1 is summarized in Theorem 2:

Theorem 2. *For the closed-loop adaptive system consisting of* (8)-(20) *under Assumptions 1-5, the following results hold:*

(a) $h(\mathbf{x}(t)) \geqslant 0\ \forall t \geqslant 0$,

(b) all results of Theorem 1 hold,

(c) $\lim_{t\to\infty} \Delta_{ebsb}(\mathbf{x}_m(t), \mathbf{e}_x(t), \hat{\theta}(t)) = 0$, *and*

(d) the safety and control objectives are met with M *independent of the parametric uncertainty.*

Proof. See Appendix II.E. □

Remark 2. *There are two buffers in* (17) *which may keep* $\mathbf{x}_m$ *well within* S *even when* $\mathbf{x}_*$ *approaches or crosses* ∂S*: (i)* Δ_{ebsb}*, which varies with time, and (ii)* δ*, which is constant. As is shown in Theorem 2, the buffer* Δ_{ebsb} *vanishes as* $t \to \infty$*.* δ *does not vanish, but can be chosen to be as small as desired.*

Remark 3. *Assumption 3 makes results (c) and (d) of Theorem 1 into domain of attraction results, as they depend on the size of the initial conditions. This is necessary because of our problem statement that requires both the control and safety objectives to be met simultaneously. We guarantee that the reference input* $\mathbf{r}_s$ *is bounded if the plant and reference model are safe, and safety in turn depends on the initial conditions.*

Remark 4. *Assumption 4 can be relaxed by starting with an unbounded set* S_r *and extracting a bounded subset* $\overline{S}_r$ *using a modified function such as* $\overline{h}_r(\mathbf{x}) = \mathrm{softmin}(h_r(\mathbf{x}), D - \|\mathbf{x}\|^2) = -\ln(e^{-h_r(\mathbf{x})} + e^{\|\mathbf{x}\|^2 - D})$ *for some* $D > 0$*. Then,* $\overline{S}_r$ *as described by* $\overline{h}_r$ *as in* (2a)-(2c) *is bounded and is contained within* S_r*, and one may use* $\overline{h}_r$ *in lieu of* h_r *in* (17)-(19) *and Assumptions 3-5.*

Remark 5. *Assumption 5 states that wherever* $\frac{\partial h_r}{\partial \mathbf{x}} B$ *goes to zero within the safe set, the constraint in* (17) *is autonomously satisfied, i.e. satisfied by* $\mathbf{r}_s = 0$*. Without this assumption,* (17) *would require* $\mathbf{r}_s$ *of infinite magnitude whenever* $\mathbf{x}_m$ *passes through such a location. Such an assumption is required, implicitly or explicitly, in all CBF literature.*

Remark 6. *Reduced-order models have been explored in [32–35] for simplifying the design of safety filters. If a reduced-order model is available for the reference model in* (9)*, it may be used to simplify* (17)*. However, the full plant dynamics are still needed for the adaptive control design in Section III.A and for the EBSB design in* (18)-(19)*.*

Remark 7. *In this work, we consider only a single state constraint. However, extensions to box constraints on multiple outputs may be possible using the approach in [36], in which the KKT conditions are decoupled to obtain an explicit safe control solution.*

III.C Intuition Behind Δ_{ebsb}

Roughly speaking, Δ_{ebsb} maintains safety of both the plant and the reference model by keeping $\mathbf{x}_m$ farther away from ∂S than the error between $\mathbf{x}$ and $\mathbf{x}_m$. More concretely, Δ_{ebsb} ensures that $h(\mathbf{x}_m(t)) \geqslant |h(\mathbf{x}(t)) - h(\mathbf{x}_m(t))| \ \forall t \geqslant 0$. The specific forms of Δ_{ebsb} in (18) and (19) are derived from (56) and (59) respectively in the proof of Theorem 2, and are chosen to ensure that each final inequality holds.

The numerator of Δ_{ebsb} upper-bounds a combination of $\mathbf{e}_x$ and $\dot{\mathbf{e}}_x$, with an additional term in (19) due to the high relative degree. The presence of $h_1(\mathbf{x}_m)$ in the denominator can be thought of as varying the sensitivity to changes in $\mathbf{e}_x$ as the reference model moves towards or away from ∂S. Recall that the adaptive control design in Section III.A causes the plant to asymptotically track the reference model. Thus, as the reference model moves far away from ∂S ($h_1(\mathbf{x}_m)$ is large), we become less concerned that the plant might imminently leave the safe set, and Δ_{ebsb} can become less sensitive to changes in $\mathbf{e}_x$. Conversely, whenever the reference model is close to ∂S ($h_1(\mathbf{x}_m)$ is small), the safety objective requires us to closely watch for any deviations between the plant and reference model and quickly increase Δ_{ebsb} in response to changes in $\mathbf{e}_x$.

IV Safe Adaptive Control with an Uncertain Input Matrix

In this section, we address Problem 2. Accounting for the input uncertainty Λ in (6) adds further difficulty in meeting the safety objective. As will be show in Section IV.B, this added difficulty will be overcome using an Error-Based Safety Filter (EBSF) rather than an EBSB as in Section III.

The parametric uncertainty is assumed to satisfy the following assumption:

Assumption 6. *The uncertain parameters satisfy the following:*

(a) θ_* *satisfies Assumption 1, and*

(b) $\Lambda = \mathrm{diag}(\lambda_*)$, *where* $\lambda_* \in L := \{\lambda \in \mathbb{R}^m : \underline{\lambda}_i \leqslant \lambda_i \leqslant \overline{\lambda}_i \ \forall i\}$ *for known constants* $\overline{\lambda}_i > \underline{\lambda}_i > 0$.

IV.A Adaptive Control Design

The adaptive control input is designed in a straightforward manner as

$$\mathbf{u} = \mathrm{diag}(\widehat{\lambda})^{-1}(K\mathbf{x} + \mathbf{r}_s + F(\mathbf{x})\widehat{\theta}). \tag{21}$$

Combining (6), (9), and (21), and defining $\mathbf{e}_x := \mathbf{x} - \mathbf{x}_m$, $\widetilde{\theta} := \widehat{\theta} - \theta_*$ and $\widetilde{\lambda} := \widehat{\lambda} - \lambda_*$, we obtain a standard error model:

$$\dot{\mathbf{e}}_x = A_m\mathbf{e}_x + B(F(\mathbf{x})\widetilde{\theta} - \mathrm{diag}(\widetilde{\lambda})\mathbf{u}). \tag{22}$$

Finally, the adaptive laws for updating $\widehat{\theta}$ and $\widehat{\lambda}$ are given by

$$\dot{\widehat{\theta}} = \mathrm{proj}_{T_\Theta(\widehat{\theta})}[-\gamma_\theta F(\mathbf{x})^\top B^\top P\mathbf{e}_x], \tag{23}$$

$$\dot{\widehat{\lambda}} = \mathrm{proj}_{T_L(\widehat{\lambda})}[\gamma_\lambda \mathrm{diag}(\mathbf{u}) B^\top P\mathbf{e}_x] \tag{24}$$

where $T_\Theta(\widehat{\theta})$ and $T_L(\widehat{\lambda})$ are the tangent cones of Θ and L at $\widehat{\theta}$ and $\widehat{\lambda}$ respectively, $\gamma_\theta, \gamma_\lambda > 0$ are any adaptive gains, and P is the solution to (13) for any symmetric positive-definite matrix Q.

We establish key properties of the adaptive control design in the following theorem [1]:

Theorem 3. *The closed-loop adaptive system consisting of* (6), (9), *and* (21)-(24) *under Assumption 6 with* $\widehat{\theta}(0) \in \Theta$ *and* $\widehat{\lambda}(0) \in L$ *has the following properties:*

(a) $\widehat{\theta}(t) \in \Theta \ \forall t \geqslant 0$,

(b) $\widehat{\lambda}(t) \in L \ \forall t \geqslant 0$, *and*

(c) $\|\mathbf{e}_x(t)\| \leqslant E_2 \ \forall t \geqslant 0$,

where E_2 *is any positive constant satisfying* $E_2^2 \geqslant \frac{\lambda_{max}(P)}{\lambda_{min}(P)}\|\mathbf{e}_x(0)\|^2 + \frac{1}{\gamma_\theta}\sup_{\theta\in\Theta}\|\widehat{\theta}(0) - \theta\|^2) + \frac{1}{\gamma_\lambda}\sup_{\lambda\in L}\|\widehat{\lambda}(0) - \lambda\|^2)$. *Additionally, if* $\mathbf{r}_s(t)$ *is bounded, then*

(d) $\mathbf{x}(t)$ *and* $\mathbf{x}_m(t)$ *are bounded, and*

(e) $\lim_{t\to\infty}\|\mathbf{e}_x(t)\| = 0$.

Proof. See Appendix III.A. □

IV.B Error-Based Safety Filter

As in Section III.B, we now design $\mathbf{r}_s$ to simultaneously accomplish the control and safety objectives. As h has relative degree $r \geqslant 1$ with respect to (6), we require a sequence of high-order CBFs [31] as follows:

$$h_1(\mathbf{x}) = h(\mathbf{x}) \tag{25a}$$

$$h_i(\mathbf{x}) = \frac{\partial h_{i-1}}{\partial \mathbf{x}}\Big|_{\mathbf{x}} A\mathbf{x} + \alpha_{i-1} h_{i-1}(\mathbf{x}) \ \forall i \in [2, r] \tag{25b}$$

for any constants $\alpha_1, \ldots, \alpha_{r-1} > 0$. Define $S_i = \{\mathbf{x} \in \mathbb{R}^n : h_i(\mathbf{x}) \geqslant 0\}$, and note that $S_1 = S$. Then, we propose the following choice of $\mathbf{r}_s$, which we refer to as an Error-Based Safety Filter (EBSF):

$$\begin{gathered}\mathbf{r}_s = \arg\min_{\mathbf{r} \in \mathbb{R}^m} \|\mathbf{r} - \mathbf{r}_*\|^2 \text{ s.t.} \\ \min_{\theta \in \Theta, \lambda \in L} \frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}} \left[(1 - \beta_{ebsf}(\mathbf{x}, \mathbf{e}_x))\mathbf{z}_m(\mathbf{x}, \mathbf{r}) + \beta_{ebsf}(\mathbf{x}, \mathbf{e}_x)\mathbf{z}_p(\mathbf{x}, \mathbf{r}, \widehat{\theta}, \widehat{\lambda}, \theta, \lambda)\right] \geqslant -\alpha_r h_r(\mathbf{x}) + \delta\end{gathered} \tag{26}$$

for any constants $\alpha_r, \delta > 0$. $\mathbf{z}_m$ and $\mathbf{z}_p$ are defined, respectively, as

$$\mathbf{z}_m(\mathbf{x}, \mathbf{r}) = A_m \mathbf{x} + B\mathbf{r} \tag{27}$$

and

$$\mathbf{z}_p(\mathbf{x}, \mathbf{r}, \widehat{\theta}, \widehat{\lambda}, \theta, \lambda) = A\mathbf{x} + B\left(\operatorname{diag}(\lambda)\operatorname{diag}(\widehat{\lambda})^{-1}\left(K\mathbf{x} + \mathbf{r} + F(\mathbf{x})\widehat{\theta}\right) - F(\mathbf{x})\theta\right), \tag{28}$$

and $\beta_{ebsf} : \mathbb{R}^n \times \mathbb{R}^n \to [0, 1]$ is a state-dependent interpolation parameter given by

$$\beta_{ebsf}(\mathbf{x}, \mathbf{e}_x) = \operatorname{sat}\left(\frac{\max\{\alpha_{ebsf}(\|\mathbf{e}_x\|), \frac{2\delta}{3\alpha_r}\} - h_r(\mathbf{x})}{\max\{\alpha_{ebsf}(\|\mathbf{e}_x\|), \frac{2\delta}{3\alpha_r}\} - \frac{\delta}{3\alpha_r}}\right) \tag{29}$$

where sat is the saturation function given by

$$\operatorname{sat}(z) = \begin{cases} 0, & z \leqslant 0 \\ z, & 0 < z < 1 \\ 1, & z \geqslant 1 \end{cases}, \tag{30}$$

for any $\alpha_{ebsf} \in \mathcal{K}_{\infty,e}$ such that $\alpha_{ebsf}(E_2) < \sup_{\mathbf{x} \in S_r} h_r(\mathbf{x})$, and E_2 is defined as in Theorem 3. The goal of β_{ebsf} is to smoothly interpolate between rendering the pristine reference dynamics safe when $\beta_{ebsf} = 0$ and rendering the plant with the worst-case parameter errors safe when $\beta_{ebsf} = 1$, and it is chosen to have the following crucial properties:

(a) $\beta_{ebsf}(\mathbf{x}(t), \mathbf{e}_x(t))$ is continuous with time,

(b) $\beta_{ebsf} = 1$ whenever $h(\mathbf{x}) \approx 0$, ensuring safety for the uncertain plant whenever $\mathbf{x}$ is close to ∂S, and

(c) $\beta_{ebsf} = 0$ whenever $h_r(\mathbf{x}) \geqslant \frac{\delta}{\alpha_r}$ and $\mathbf{e}_x = 0$, ensuring that $\mathbf{r}_s$ becomes independent of the parametric uncertainty in the limit as $\mathbf{e}_x$ goes to zero.

We now present our main result for Problem 2. First, as in Section III.B, in order to ensure that the safety objective is satisfied at the initial time and that $\mathbf{r}_s$ remains bounded, including at locations where the gradient of h_r vanishes, we require the following additional assumptions:

Assumption 7. *The initial conditions of the closed-loop adaptive system satisfy:*

(a) $h_i(\mathbf{x}(0)) \geqslant 0 \ \forall i \in [1, r]$,

(b) $\widehat{\theta}(0) \in \Theta$, *and*

(c) $\widehat{\lambda}(0) \in L$.

Assumption 8. *The highest-order safe set* $S_r \subset \mathbb{R}^n$ *is bounded.*

Assumption 9. *There exists a constant* $d_1 > 0$ *such that, at every* $\mathbf{x}_1 \in S_r$ *where* $h_r(\mathbf{x}_1) < \max\{\alpha_{ebsf}(E_2), \frac{2\delta}{3\alpha_r}\}$, *we have* $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_1} B\| \geqslant d_1$.

Assumption 10. *There exists a constant* $d_2 > 0$ *such that, at every* $\mathbf{x}_2 \in S_r$ *where* $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_2} B\| < d_2$, *we have* $\alpha_r h_r(\mathbf{x}_2) + \frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_2} A_m \mathbf{x}_2 \geqslant \delta$.

Then, our main result for Problem 2 follows:

Theorem 4. *For the closed-loop adaptive system consisting of* (6)*,* (9)*, and* (21)*-*(24) *under Assumptions 6-10, the following results hold:*

(a) $h(\mathbf{x}(t)) \geqslant 0 \ \forall t \geqslant 0$*,*

(b) all results of Theorem 3 hold,

(c) there exists a finite time $T \geqslant 0$ *such that* $\beta_{ebsf}(\mathbf{x}(t), \mathbf{e}_x(t)) = 0 \ \forall t \geqslant T$*, and*

(d) the safety and control objectives are met with M *independent of the parametric uncertainty.*

Proof. See Appendix III.B. □

Remark 8. *It may not be immediately obvious how to implement the EBSF, as* (26) *is not in the form of a quadratic programming problem. However, as shown in Appendix Appendix IV, it is possible to rewrite* (26) *as a quadratic program with* $2^m - 1$ *constraints by leveraging Assumption 6.*

Remark 9. *It should be noted that the use of the word Filter is not in reference to the frequency response of* (26)*. Rather, we refer to* (26) *as an Error-Based Safety Filter as it tends to filter out unsafe reference inputs, in parity with the use of the term Safety Filter elsewhere in the CBF literature (e.g. [7]).*

IV.C Intuition Behind EBSF

EBSB ensures safety of the plant by keeping the reference model sufficiently far away from ∂S, thus ensuring safety of both the plant and reference model. In contrast, EBSF guarantees safety of the plant but not the reference model during the transient period while $\|\mathbf{e}_x\|$ is of significant magnitude. Due to the nature of the uncertainty, whenever $\mathbf{x}$ is near ∂S_r, one may need to apply $\mathbf{r}_s$ in a narrow range of directions to prevent $\mathbf{x}$ from crossing ∂S_r. If $\mathbf{e}_x \neq 0$, $\mathbf{x}_m$ may be at a different location near ∂S_r such that any $\mathbf{r}_s$ in that narrow range of directions would cause $\mathbf{x}_m$ to cross the boundary. It is difficult in general to avoid such transient situations where safety of $\mathbf{x}$ and $\mathbf{x}_m$ cannot be simultaneously ensured.

Instead of choosing $\mathbf{r}_s$ based on $\mathbf{x}_m$ and $\mathbf{e}_x$, EBSF considers two possible choices of $\mathbf{r}_s$. The first possibility assumes that the adaptive controller in (21) has accomplished its task and that the plant is exhibiting the desired closed-loop behavior, i.e. that $\dot{\mathbf{x}} = A_m\mathbf{x} + B\mathbf{r}_s$, and chooses an ideal $\mathbf{r}_s$ which will ensure $\frac{d}{dt}h_r(\mathbf{x}) \geqslant -\alpha_r h_r(\mathbf{x}) + \delta$ for the desired closed-loop plant dynamics. This first possibility is captured in $\mathbf{z}_m$ in (27). The second possibility for $\mathbf{r}_s$ makes no assumptions on the completeness of the adaptation, but rather uses prior knowledge of the parameter sets Θ and L to choose a conservative $\mathbf{r}_s$ which will ensure $\frac{d}{dt}h_r(\mathbf{x}) \geqslant -\alpha_r h_r(\mathbf{x}) + \delta$ for any $\theta_* \in \Theta$ and $\lambda_* \in L$. This second possibility is captured in $\mathbf{z}_p$ in (28). Given these two possible choices of $\mathbf{r}_s$, then, the EBSF approach is to apply the ideal $\mathbf{r}_s$ when $\mathbf{x}$ is far away from ∂S_r, to apply the worst-case $\mathbf{r}_s$ when $\mathbf{x}$ is close to ∂S_r, and to linearly interpolate between the two choices of $\mathbf{r}_s$ at intermediate distances from ∂S_r. Concretely, we establish an interpolation window around ∂S_r consisting of $h_r(\mathbf{x}) \in (\frac{\delta}{3\alpha_r}, \max\{\alpha_{ebsf}(\|\mathbf{e}_x\|), \frac{2\delta}{3\alpha_r}\})$. For $h_r(\mathbf{x}) \leqslant \frac{\delta}{3\alpha_r}$, we apply the worst-case choice of $\mathbf{r}_s$; for $h_r(\mathbf{x}) \geqslant \max\{\alpha_{ebsf}(\|\mathbf{e}_x\|), \frac{2\delta}{3\alpha_r}\}$, we apply the ideal choice of $\mathbf{r}_s$; and for $\mathbf{x}$ in the interpolation window, we linearly interpolate between the two choices of $\mathbf{r}_s$. α_{ebsf} is a hand-tuned function which controls how the width of the interpolation window varies with $\|\mathbf{e}_x\|$, and it ensures that in the limit as $\|\mathbf{e}_x\|$ goes to zero and the plant follows the reference model dynamics exactly, the width of the interpolation window shrinks to a value sufficiently small that the plant will never enter.

The results in Theorem 4 are independent of the choice of α_{ebsf} in (29). In fact, it turns out that one can remove the error dependence altogether by choosing $\alpha_{ebsf} = 0$, and Theorem 4 will still hold. However, for good tracking of $\mathbf{x}_*$, δ typically must be very small. Thus, an interpolation window given only by $h_r(\mathbf{x}) \in (\frac{\delta}{3\alpha_r}, \frac{2\delta}{3\alpha_r})$ would be very thin. If the plant were to enter such a thin interpolation window, β_{ebsf} could change from 0 to 1 rapidly, resulting in a rapid change in $\mathbf{r}_s$ that would tend to push the plant back across the window and cause β_{ebsf} to rapidly change from 1 to 0. Therefore, while $\mathbf{r}_s(t)$ is never discontinuous with time, a thin interpolation window can result in a jitter-like behavior where $\mathbf{r}_s$ changes rapidly between the two aforementioned possibilities. The EBSF solution to this problem is to vary the width of the interpolation window with $\|\mathbf{e}_x\|$. Intuitively, while $\|\mathbf{e}_x\|$ is large, the plant is less predictable and more inclined to enter the window, and thus we need the window to be wide. Conversely, as $\|\mathbf{e}_x\|$ goes to zero, the plant becomes predictable, and we can shrink the window without fear of the plant entering. Thus, α_{ebsf} acts as as an error-based smoothing, helping to minimize jitter in $\mathbf{r}_s$. More principled ways to tune α_{ebsf} or otherwise avoid jitter in $\mathbf{r}_s$ are an important matter for future work. The simulation results in Section VI will further explore this jittering behavior, and will propose a way of largely mitigating it in practice.

V Data-Driven Performance Enhancement

In [26], Set-Membership Identification (SMID) was proposed as a way to reduce conservatism of the RaCBF approach, allowing $\mathbf{x}_m$ to approach $\mathbf{x}_*$ more closely in response to excitation. As shown in Theorems 2 and 4, EBSB and EBSF have conservatism which vanishes regardless of excitation. Nevertheless, in the event that excitation is present, we show in this section how to leverage SMID in EBSB and EBSF to further reduce conservatism in response to excitation. Without making assumptions on the amount of excitation available to learn from, the addition of SMID can neither strengthen nor weaken the theoretical performance guarantees. Thus, we prove here that all results of Theorems 1-4 hold under augmentation with SMID, and we will explore the practical benefits of SMID in simulation in Section VI. It should be noted that SMID cannot easily be extended to the setting where parameters are time-varying, and thus is only applicable when parameters are constant for all time.

V.A Set Membership Identification (SMID)

Suppose that Assumptions 1 and 6 are initially satisfied with parameter sets Θ_0 and L_0 respectively. For the purposes of this work, an SMID algorithm is any method of updating the parameter sets at times $t_1, t_2, \ldots, t_k, \ldots$ such that:

(a) $\Theta_{k+1} \subseteq \Theta_k \ \forall k \geqslant 0$,

(b) $L_{k+1} \subseteq L_k \ \forall k \geqslant 0$,

(c) $\theta_* \in \Theta_k \ \forall k \geqslant 0$,

(d) $\lambda_* \in L_k \ \forall k \geqslant 0$, and

(e) Θ_k and L_k are closed and convex $\forall k \geqslant 0$.

Then, $\Theta(t)$ and $L(t)$ are piecewise-constant known parameter sets containing θ_* and λ_* respectively, where $t_0 = 0$, $\Theta(t) = \Theta_k \ \forall t \in [t_k, t_{k+1})$, and $L(t) = L_k \ \forall t \in [t_k, t_{k+1})$.

V.B Adding SMID to EBSB

Consider the following modifications to the control design in Section III. First, replace the adaptive law in (12) with

$$\dot{\hat{\theta}} = \mathrm{proj}_{T_{\Theta(t)}(\hat{\theta})}[-\gamma F(\mathbf{x})^\top B^\top P \mathbf{e}_x] \tag{31}$$

with an additional modification at each SMID update time t_k given by

$$\hat{\theta}(t_k^+) = \mathrm{proj}_{\Theta_k}[\hat{\theta}(t_k^-)]. \tag{32}$$

Second, simply replace Θ in (20) with $\Theta(t)$. Then, given these two modifications, it is straightforward to show that all results in Section III hold as follows:

Theorem 5. *For the closed-loop adaptive system consisting of* (8)-(19) *under Assumptions 1-5 with the modifications above and* $\Theta(t)$ *given by an SMID algorithm, all results in Theorems 1 and 2 hold.*

Proof. The extensions of the proofs of Theorems 1 and 2 are straightforward. For V in (43), $\dot{V}$ remains unchanged for all $t \notin \{t_1, t_2, \ldots\}$, and it is straightforward to verify that (32) along with $\Theta_{k+1} \subseteq \Theta_k \ \forall k$ ensures that $V(t_k^+) \leqslant V(t_k^-) \ \forall k$. Therefore, the results of Theorem 1 hold. Furthermore, (31) and (32) ensure that $\hat{\theta}(t) \in \Theta(t) \ \forall t \geqslant 0$, and thus Ψ in (20) with Θ replaced by $\Theta(t)$ has the property $\|\tilde{\theta}\| \leqslant \Psi(\hat{\theta})$. Therefore, the proof of Theorem 2 remains unchanged. □

V.C Adding SMID to EBSF

Consider the following modifications to the control design in Section IV. First, replace the adaptive laws in (23)-(24) with

$$\dot{\hat{\theta}} = \mathrm{proj}_{T_{\Theta(t)}(\hat{\theta})}[-\gamma_\theta F(\mathbf{x})^\top B^\top P \mathbf{e}_x], \tag{33}$$

$$\dot{\hat{\lambda}} = \mathrm{proj}_{T_{L(t)}(\hat{\lambda})}[\gamma_\lambda \mathrm{diag}(\mathbf{u}) B^\top P \mathbf{e}_x] \tag{34}$$

with an additional modification at each SMID update time t_k given by

$$\hat{\theta}(t_k^+) = \mathrm{proj}_{\Theta_k}[\hat{\theta}(t_k^-)], \tag{35}$$

$$\widehat{\lambda}(t_k^+) = \text{proj}_{L_k}[\widehat{\lambda}(t_k^-)]. \tag{36}$$

Second, simply replace Θ and L in (26) with $\Theta(t)$ and $L(t)$. Then, given these two modifications, it is straightforward to verify that all results in Section IV hold as follows:

Theorem 6. *For the closed-loop adaptive system consisting of* (6) *and* (21)-(29) *under Assumptions 6-10 with the modifications above and* $\Theta(t)$, $L(t)$ *given by an SMID algorithm, all results in Theorems 3 and 4 hold.*

Proof. The extensions of the proofs of Theorems 3 and 4 are straightforward. For V in (62), $\dot{V}$ remains unchanged for all $t \notin \{t_1, t_2, \dots\}$, and it is straightforward to verify that (35)-(36) along with $\Theta_{k+1} \subseteq \Theta_k$, $L_{k+1} \subseteq L_k$ $\forall k$ ensures that $V(t_k^+) \leqslant V(t_k^-)$ $\forall k$. Therefore, the results of Theorem 3 hold. Furthermore, (33)-(36) ensure that $\widehat{\theta}(t) \in \Theta(t)$ $\forall t \geqslant 0$ and $\widehat{\lambda}(t) \in L(t)$ $\forall t \geqslant 0$. Therefore, the proof of Theorem 4 remains unchanged. □

VI Simulations and Discussion

To demonstrate the efficacy of our proposed approaches, we simulate a simple academic system consisting of a point mass m moving in the x-y plane, tethered to the origin by an ideal spring and damper with coefficients k and b. The control inputs are forces F_x and F_y. The control objective is to cause $\mathbf{p} = [x, y]^\top$ to track a moving reference point $\mathbf{p}_*(t) = [x_*(t), y_*(t)]^\top$, while the safety objective is to avoid a circular pillar centered at $(x_{\text{pillar}}, y_{\text{pillar}})$ with radius r_{pillar}. In Problem 1, only k and b are unknown, while in Problem 2, m is unknown as well. It is straightforward to show that the dynamics reduce to

$$\begin{gathered}
\dot{\mathbf{x}} = A\mathbf{x} + B(\lambda_*\mathbf{u} - F(\mathbf{x})\theta_*), \\
\mathbf{x} = [x, y, \dot{x}, \dot{y}]^\top, \quad \mathbf{u} = [F_x, F_y]^\top, \quad \lambda_* = \frac{1}{m}, \quad \theta_* = \left[\frac{k}{m}, \frac{b}{m}\right]^\top, \\
A = \begin{bmatrix} 0_{2\times2} & I_2 \\ 0_{2\times2} & 0_{2\times2} \end{bmatrix}, \quad B = \begin{bmatrix} 0_{2\times2} \\ I_2 \end{bmatrix}, \quad F(\mathbf{x}) = \begin{bmatrix} x & \frac{x(x\dot{x}+y\dot{y})}{x^2+y^2} \\ y & \frac{y(x\dot{x}+y\dot{y})}{x^2+y^2} \end{bmatrix}.
\end{gathered} \tag{37}$$

We use the following true parameter values and initial parameter estimates: $m = 1.1$kg, $k = 1$N/m, $b = 0.9$N/(m/s), $\widehat{k}(0) = 2$N/m, and $\widehat{b}(0) = 0.5$N/(m/s). In the adaptive control design, we assume that $k \in [\underline{k}, \overline{k}]$ and $b \in [\underline{b}, \overline{b}]$ where $\underline{k} = 0$N/m, $\overline{k} = 3$N/m, $\underline{b} = 0$N/(m/s), and $\overline{b} = 2$N/(m/s). Additionally, in problem 2, we use the initial parameter estimate $\widehat{m}(0) = 2$kg, and we assume that $m \in [\underline{m}, \overline{m}]$ with $\underline{m} = 0.5$kg, $\overline{m} = 2.5$kg. Initial estimates $\widehat{\theta}(0)$ and $\widehat{\lambda}(0)$ are derived from $\widehat{k}(0)$, $\widehat{b}(0)$, and $\widehat{m}(0)$, and element-wise bounds on θ_* and λ_* are derived from $\underline{k}$, $\overline{k}$, $\underline{b}$, $\overline{b}$, $\underline{m}$, and $\overline{m}$. We implement each controller with a zero-order hold at 100 Hz. The initial state is chosen together with the initial parameter estimate such that the initial controller will tend to push the point mass toward the pillar more quickly than expected. Furthermore, the desired trajectory $\mathbf{p}_*$ is chosen to drive the plant through the obstacle and to stop inside it, necessitating deviation from the desired trajectory to ensure safety.

See Appendix V.A for the ideal control solution if all parameters are known. It should be noted that $h(\mathbf{x})$ in (85) satisfies Assumption 2 with $\kappa = 1$, but that the high-order safe set S_2 resulting from h_2 in (86a) does not satisfy Assumptions 4 and 8. Nevertheless, EBSB and EBSF will be shown to work well.

VI.A Problem 1: Simulations and Discussion

For Problem 1 with k and b unknown and m known, we simulate and compare four approaches with formal guarantees of safety: aCBF [25], RaCBF [26], EBSB, and EBSF. See Appendices V.B and V.C respectively for aCBF and RaCBF implementation details. For comparison, we also simulate the hypothetical ideal safe controller in Appendix V.A. Set Membership Identification (SMID), proposed in [26] to improve RaCBF performance, is not used here – we defer discussion of SMID to Section VI.D. The control adaptive laws in all approaches use the adaptation rate $\gamma_\theta = \gamma_\lambda = 3$, and for the additional safety adaptive laws in the aCBF and RaCBF approaches, $\gamma_{\theta,s} = \gamma_{\lambda,s} = 10$ is used. For α_{ebsf} in EBSF, we use the intuitive choice

$$\alpha_{ebsf}(\|\mathbf{e}_x\|) = \kappa\|\mathbf{e}_x\|, \tag{38}$$

where $\kappa = 1$ is the bound on the gradient of h from Assumption 2.

The (x, y) trajectories of all simulated approaches are compared in Fig. 1, and the reference inputs $\mathbf{r}_s$ are compared in Fig. 2. Clearly, while both aCBF and RaCBF maintain safety, they maintain a distance from the pillar which depends on the choice of $\overline{k} - \underline{k}$ and $\overline{b} - \underline{b}$. In contrast, EBSB and EBSF are able to stay much closer to the pillar, and even permit the plant to start approaching the pillar as $\|\mathbf{e}_x\|$ tends toward zero near the end of the simulation. Furthermore, it is apparent in Figs. 1 and 2 that EBSF is able to be less conservative than EBSB at the cost of some jitter in $\mathbf{r}_s$ starting at roughly 12 seconds, while $\mathbf{r}_s$ produced by EBSB remains smooth.

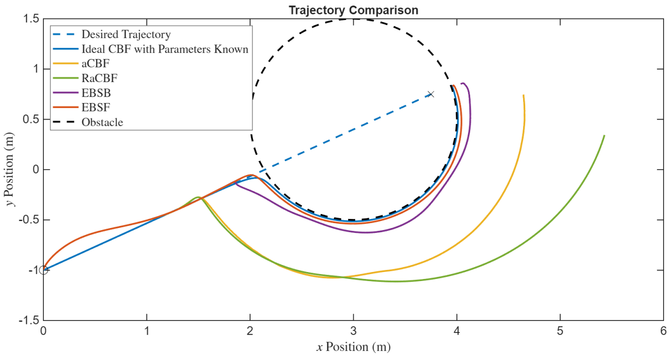

Figure 1: Trajectories resulting from aCBF [25], RaCBF [26], EBSB, and EBSF for Problem 1 compared to the ideal controller with all parameters known.

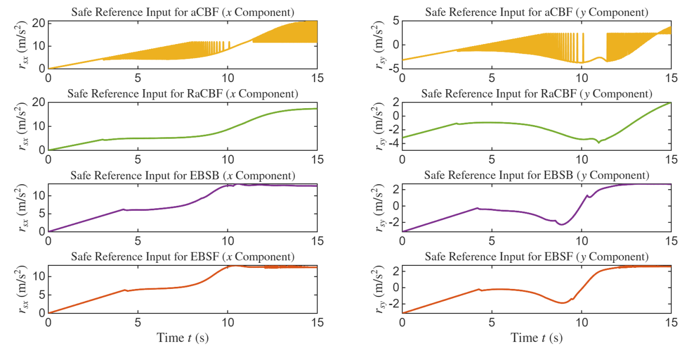

Figure 2: Reference inputs $\mathbf{r}_s$ from aCBF [25], RaCBF [26], EBSB, and EBSF for Problem 1.

VI.B Comparing EBSB and EBSF

Figs. 1 and 2 demonstrate the relative advantages and drawbacks of EBSB and EBSF when applied to Problem 1. EBSB has the advantages of requiring no hand-tuning beyond the standard parameter selection for adaptive control and CBFs, and of avoiding jitter in $\mathbf{r}_s$. However, it cannot be applied with an uncertain input matrix, and tends to be more conservative than EBSF. In contrast, EBSF applies to a more general problem statement and is able to be less conservative than EBSB, but its behavior depends greatly on hand-tuning of α_{ebsf}. If $\alpha_{ebsf}(\|\mathbf{e}_x\|)$ grows too quickly with $\|\mathbf{e}_x\|$, then the plant can veer wildly away from ∂S in response to small perturbations in $\mathbf{e}_x$. On the other hand, if $\alpha_{ebsf}(\|\mathbf{e}_x\|)$ grows too slowly with $\|\mathbf{e}_x\|$, then perturbations in $\mathbf{e}_x$ may cause the plant to enter a thin interpolation window as described in Section IV.C, resulting in the jitter observed in Fig. 2 starting at roughly 12 seconds. This jitter is an artifact of the zero-order hold used in our simulations. Under a zero-order hold, when $\alpha_{ebsf}(\|\mathbf{e}_x\|)$ becomes very small and adaptation is not yet complete, β_{ebsf} can go from zero to one or vice versa in the span of a single time step, causing the input produced by EBSF to change rapidly. When the time step is decreased, we find that the jitter shrinks in magnitude. However, any practical application of EBSF would be in a digital control system with a zero-order hold, and thus future work is needed to avoid jitter in digital control systems with EBSF.

VI.C Problem 2: Simulations and Discussion

For Problem 2 with k, b, and m unknown, we simulate and compare three approaches: aCBF [25], RaCBF [26], and EBSF. For comparison, we also simulate the hypothetical ideal safe controller in Appendix V.A. The true parameters, initial estimates, and upper/lower bounds are the same as in Section VI.A, with the addition of $\widehat{m}(0) = 2$kg and the assumption that $m \in [\underline{m}, \overline{m}]$ with $\underline{m} = 0.5$kg, $\overline{m} = 2.5$kg.

The (x, y) trajectories of all simulated approaches are compared in Fig. 3, and the reference inputs $\mathbf{r}_s$ are compared in Fig. 4. Again, aCBF and RaCBF remain safe but are even more conservative than in Problem 1 due to the increased uncertainty in the parameters. In contrast, EBSF is able to be nonconservative while maintaining safety even under input uncertainties. We do observe in Fig. 4, however, that the jitter observed in Problem 1 due to the zero-order hold is amplified in Problem 2 because of the additional uncertainty. In the following section, we explore the use of SMID to both further reduce the conservatism of EBSB and EBSF and reduce the jitter of EBSF in practice.

VI.D Performance Enhancement with Set Membership Identification

Set Membership Identification (SMID) was proposed in [26] as a method of reducing the conservatism of RaCBF in practice. In the preceeding simulations, SMID was omitted. In this section, we add SMID to RaCBF following the approach in [26], and we add SMID to EBSB and EBSF following the approaches in Section V. Details of our SMID implementation can be found in Appendix V.D.

Figs. 5 and 6 show the trajectories and reference inputs respectively of aCBF, RaCBF, EBSB, and EBSF augmented with SMID in the same simulation scenario as in Section VI.A. SMID greatly reduces the conservatism of RaCBF, and also further reduces the conservatism of EBSB and EBSF, enabling EBSB and EBSF to still provide superior reference tracking. Furthermore, because the bounds on the uncertainty are decreased online, the jitter in EBSF is greatly reduced in magnitude relative to Fig. 2.

Figs. 7 and 8 show the trajectories and reference inputs respectively of aCBF, RaCBF, EBSB, and EBSF augmented with SMID in the same simulation scenario as in Section VI.C. The same result is apparent: both RaCBF and EBSF

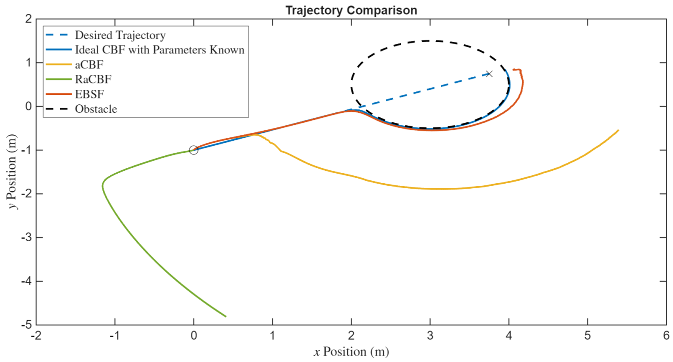

Figure 3: Trajectories resulting from aCBF [25], RaCBF [26], and EBSF for Problem 2 compared to the ideal controller with all parameters known.

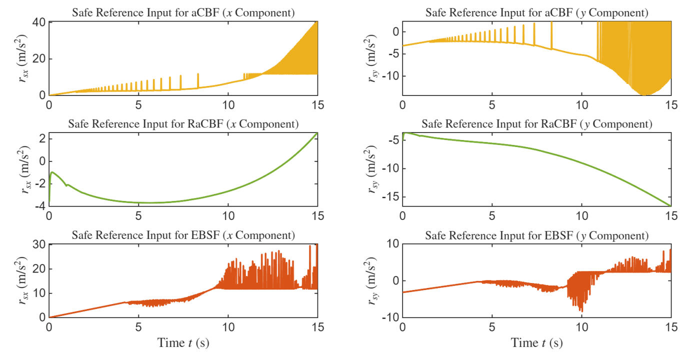

Figure 4: Reference inputs $\mathbf{r}_s$ from aCBF [25], RaCBF [26], and EBSF for Problem 2.

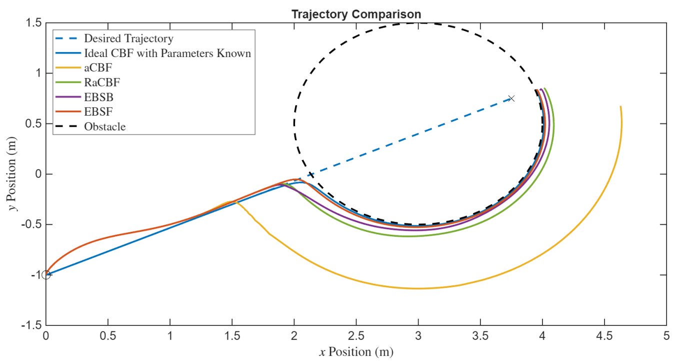

Figure 5: Trajectories resulting from aCBF [25], RaCBF [26], EBSB, and EBSF with SMID for Problem 1 compared to the ideal controller with all parameters known.

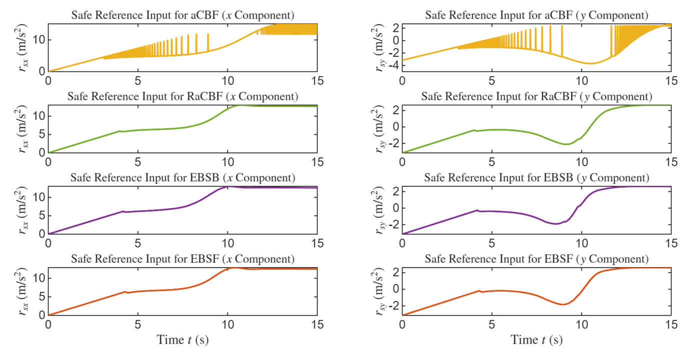

Figure 6: Reference inputs $\mathbf{r}_s$ from aCBF [25], RaCBF [26], EBSB, and EBSF with SMID for Problem 1.

are greatly reduced in conservatism, with EBSF still the least conservative approach. In this simulation, the jitter in EBSF is still present, but is again decreased in magnitude relative to Fig. 4.

VII Conclusions and Future Work

In this paper, we consider safety and stability of adaptive control of a class of feedback linearizable plants with matched parametric uncertainties whose states are accessible, subject to a state constraint. The adaptive control architectures proposed are shown to guarantee stability, ensure control performance, and maintain safety even with the parametric uncertainties. Two problems are considered, where the first includes parametric uncertainties in the system Jacobian, while the second assumes additional uncertainties in the input matrix. In both cases, the control barrier function is assumed to have an arbitrary relative degree. No excitation conditions are imposed on the command signal. Simulation results demonstrate the non-conservatism of all of the theoretical developments.

Considerable work remains in the direction of ensuring simultaneous satisfaction of control performance and safety. How the proposed methods can be extended to plants with output feedback, general nonlinear systems that are not feedback linearizable, and multi-inputs are important directions for future work. Accommodation of input constraints due to magnitude and rate limits needs to be addressed as well. More immediately, better tuning methods for α_{ebsf} so as to reduce jitter are an important topic for future work.

Appendix I

I.A Proof of Lemma 1

We can first say that the surface $\partial\mathcal{C}$ is continuous, as any discontinuities in $\partial\mathcal{C}$ would result in $\mathcal{C}$ being non-closed. Additionally, we can say that $\partial\mathcal{C}$ is piecewise smooth, as a set whose boundary is continuous but not piecewise smooth

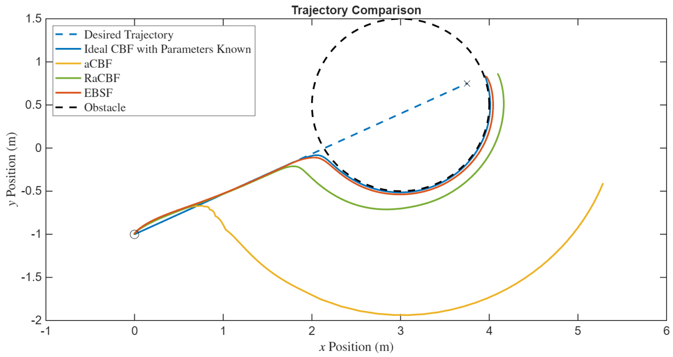

Figure 7: Trajectories resulting from aCBF [25], RaCBF [26], and EBSF with SMID for Problem 2 compared to the ideal controller with all parameters known.

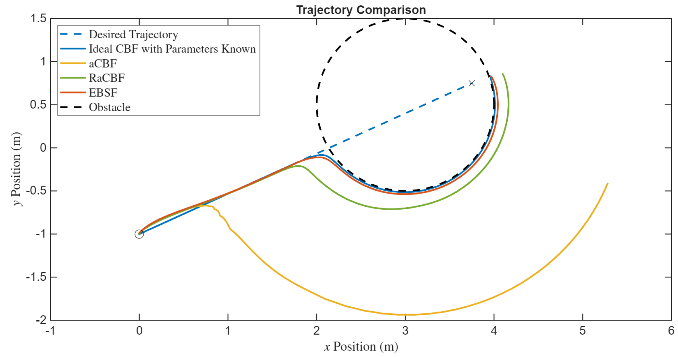

Figure 8: Trajectories resulting from aCBF [25], RaCBF [26], and EBSF with SMID for Problem 2 compared to the ideal controller with all parameters known.

must either be non-convex or have zero n-dimensional volume. Finally, it is easy to see that a closed, convex set $\mathcal{C} \subset \mathbb{R}^n$ whose boundary is continuous and piecewise smooth can be expressed as the intersection of zero-superlevel sets of smooth functions $g_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \ldots$, with each smooth piece of $\partial\mathcal{C}$ corresponding to a function g_k.

I.B Proof of Lemma 2

First, suppose that $\mathbf{v} \in \text{int}(\mathcal{C})$. Then, $\text{proj}_{T_\mathcal{C}(\mathbf{v})}[\mathbf{z}] = \mathbf{z}$, and the inequality is trivially satisfied.

Second, suppose that $\mathbf{v} \in \partial C$. Then, define $\mathbb{K} \subset \mathbb{N}$ such that $g_k(\mathbf{v}) = 0 \ \forall k \in \mathbb{K}$, and we have

$$\text{proj}_{T_\mathcal{C}(\mathbf{v})}[\mathbf{z}] = \underset{\mathbf{y}\in\mathbb{R}^n}{\arg\min} \|\mathbf{y} - \mathbf{z}\|^2 \text{ s.t. } \nabla g_k(\mathbf{v})^\top \mathbf{y} \geqslant 0 \ \forall k \in \mathbb{K}. \tag{39}$$

Forming the Lagrangian and applying the KKT conditions, the solution to (39) can be written as

$$\text{proj}_{T_\mathcal{C}(\mathbf{v})}[\mathbf{z}] = \mathbf{z} + \sum_{k\in\mathbb{K}} \lambda_k \nabla g_k(\mathbf{v}) \tag{40}$$

for some $\lambda_k \geqslant 0$. Finally, it follows from the fact that $\mathcal{C}$ is convex that

$$(\mathbf{w} - \mathbf{v})^\top \nabla g_k(\mathbf{v}) \geqslant 0 \ \forall \mathbf{w} \in \mathcal{C}, k \in \mathbb{K}, \tag{41}$$

and we obtain

$$(\mathbf{v} - \mathbf{w})^\top \text{proj}_{T_\mathcal{C}(\mathbf{v})}[\mathbf{z}] = (\mathbf{v} - \mathbf{w})^\top \mathbf{z} + \sum_{k\in\mathbb{K}} \lambda_k (\mathbf{v} - \mathbf{w})^\top \nabla g_k(\mathbf{v}) \leqslant (\mathbf{v} - \mathbf{w})^\top \mathbf{z}. \tag{42}$$

Appendix II

II.A Proof of Theorem 1

Consider the functions $g_{\Theta,1}, \ldots, g_{\Theta,N}$ describing Θ as in Lemma 1. Because $\dot{\widehat{\theta}}(t) \in T_\Theta(\widehat{\theta}(t)) \ \forall t \geqslant 0$, we know that whenever $g_{\Theta,k}(\widehat{\theta}) = 0$ for any k, we have $\frac{d}{dt} g_{\Theta,k}(\widehat{\theta}) = \nabla g_{\Theta,k}(\widehat{\theta})^\top \dot{\widehat{\theta}} \geqslant 0$. Result (a) is immediate from the fact that $\widehat{\theta}(0) \in \Theta$.

We propose the following Lyapunov function candidate, which is standard in adaptive control [1]:

$$V = \mathbf{e}_x^\top P \mathbf{e}_x + \frac{\|\widetilde{\theta}\|^2}{\gamma}. \tag{43}$$

Since θ_* is assumed constant, we have $\dot{\widetilde{\theta}} = \dot{\widehat{\theta}}$. Additionally, since $\widehat{\theta}, \theta_* \in \Theta$ and $\dot{\widehat{\theta}} \in T_\Theta(\widehat{\theta})$, from Lemma 2 and (12), we have

$$\widetilde{\theta}^\top \dot{\widehat{\theta}} \leqslant \widetilde{\theta}^\top(-\gamma F(\mathbf{x})^\top B^\top P \mathbf{e}_x). \tag{44}$$

Then, using (13), (11), and (44), one can show through straightforward algebra that

$$\dot{V} \leqslant -\mathbf{e}_x^\top Q \mathbf{e}_x \leqslant 0 \tag{45}$$

and thus that $V(t) \leqslant V(0)\ \forall t \geqslant 0$. Result (b) is immediate.

Now, if $\mathbf{r}_s(t)$ is bounded, then from (9) and Hurwitzness of A_m, we know that $\mathbf{x}_m(t)$ is bounded. It follows from boundedness of $\mathbf{e}_x(t)$ that $\mathbf{x}(t)$ is bounded, thus giving result (c). Finally, from (11), boundedness of $\mathbf{e}_x(t)$ and $\mathbf{x}(t)$ guarantees that $\dot{\mathbf{e}}_x(t)$ is bounded. Result (d) follows from Barbalat's lemma [1].

II.B Proof of Lemma 3

From the definition of $\mathbb{I}_h$ in (15), it follows immediately that for any $\mathbf{x} \in \mathbb{R}^n$, $(\frac{\partial h}{\partial \mathbf{x}})^\top$ can be written as a linear combination of the orthonormal vectors $\mathbf{w}_j$ spanning $\mathbb{I}_h$. Then, from (16a) and orthonormality of the vectors $\mathbf{w}_j$, we immediately obtain

$$\frac{\partial h}{\partial \mathbf{x}} C_h = \frac{\partial h}{\partial \mathbf{x}}\ \forall \mathbf{x} \in \mathbb{R}^n. \tag{46}$$

Now, consider a line given by $\mathbf{x}_2 + s(\mathbf{x}_1 - \mathbf{x}_2)$, as s varies from 0 to 1. Then, $\frac{dh}{ds} = \frac{\partial h}{\partial \mathbf{x}}|_{\mathbf{x}_2+s(\mathbf{x}_1-\mathbf{x}_2)}(\mathbf{x}_1 - \mathbf{x}_2)$. Finally, using (46), Assumption 2, the fact that $\mathcal{C}$ is convex, and (16b),

$$\begin{aligned} |h(\mathbf{x}_1) - h(\mathbf{x}_2)| &= \left| \int_0^1 \frac{\partial h}{\partial \mathbf{x}}\Big|_{\mathbf{x}_2+s(\mathbf{x}_1-\mathbf{x}_2)} (\mathbf{x}_1 - \mathbf{x}_2) \mathrm{d}s \right| \\ &= \left| \int_0^1 \frac{\partial h}{\partial \mathbf{x}}\Big|_{\mathbf{x}_2+s(\mathbf{x}_1-\mathbf{x}_2)} C_h(\mathbf{x}_1 - \mathbf{x}_2) \mathrm{d}s \right| \\ &\leqslant \|C_h(\mathbf{x}_1 - \mathbf{x}_2)\| \sup_{s\in[0,1]} \left\| \frac{\partial h}{\partial \mathbf{x}}\Big|_{\mathbf{x}_2+s(\mathbf{x}_1-\mathbf{x}_2)} \right\| \\ &\leqslant \kappa \|C_h(\mathbf{x}_1 - \mathbf{x}_2)\| \\ &= \sqrt{(\mathbf{x}_1 - \mathbf{x}_2) W_0 (\mathbf{x}_1 - \mathbf{x}_2)}. \end{aligned} \tag{47}$$

II.C Proof of Lemma 4

From the definition of relative degree r in Definition 5 applied to the reference model in (9), we obtain

$$\frac{\partial h}{\partial \mathbf{x}} B = \frac{\partial h}{\partial \mathbf{x}} A_m B = \cdots = \frac{\partial h}{\partial \mathbf{x}} A_m^{r-2} B = 0\ \forall \mathbf{x} \in \mathbb{R}^n. \tag{48}$$

Now, note that the definition of $\mathbb{I}_h$ in (15) implies that for any linear combination of the vectors $\mathbf{w}_j$, there exists $\mathbf{x} \in \mathbb{R}^n$ where $\frac{\partial h}{\partial \mathbf{x}}$ is parallel to that linear combination. Additionally, from (16a), we see that each row of C_h is a linear combination of the vectors $\mathbf{w}_j^\top$. From these two observations, we conclude that for each row of C_h, there exists $\mathbf{x} \in \mathbb{R}^n$ where $\frac{\partial h}{\partial \mathbf{x}}$ is parallel to that row. Finally, from this conclusion and (48), it immediately follows that

$$C_h B = C_h A_m B = \cdots = C_h A_m^{r-2} B = 0. \tag{49}$$

The claim follows from tracing the recursion in (16b)-(16c) up through W_{r-2}.

II.D Proof of Proposition 1

Define a collection of time-varying signals

$$H_{mi}(t) = h_i(\mathbf{x}_m(t)) - \frac{\delta}{\prod_{k=i}^r \alpha_k}\ \forall i \in [1, r] \tag{50}$$

with h_i defined in (14a)-(14b). Then, using (9),

$$\dot{H}_{mi} = \frac{\partial h_i}{\partial \mathbf{x}}\Big|_{\mathbf{x}_m} (A_m \mathbf{x}_m + B \mathbf{r}_s). \tag{51}$$

The definition of relative degree r implies that $\frac{\partial h_i}{\partial \mathbf{x}}|_{\mathbf{x}_m} B = 0\ \forall i < r$. Thus, using (14b) and (50), we have

$$\begin{aligned} \dot{H}_{mi} &= h_{i+1}(\mathbf{x}_m) - \alpha_i h_i(\mathbf{x}_m) \\ &= H_{m(i+1)} - \alpha_i H_{mi}\ \forall i \in [1, r-1], \end{aligned} \tag{52}$$

and using (17) with the fact that $\Delta_{ebsb} \geqslant 0$, we obtain

$$\dot{H}_{mr} \geqslant -\alpha_r h_r(\mathbf{x}_m) + \delta = -\alpha_r H_{mr}. \tag{53}$$

Assumption 3 implies that $H_{mi}(0) \geqslant 0\ \forall i \in [1, r]$. Therefore, it follows from Assumption 3 and (52)-(53) that $H_{mr}(t) \geqslant 0\ \forall t \geqslant 0$, and thus that $H_{m(r-1)}(t) \geqslant 0 \implies H_{m(r-2)}(t) \geqslant 0 \implies \cdots \implies H_{m1}(t) \geqslant 0\ \forall t \geqslant 0$. The claim follows from the definition of H_{mi} in (50).

II.E Proof of Theorem 2

Proof of Claim (a)

Define a collection of time-varying signals

$$H_{e1}(t) = h_1(\mathbf{x}_m(t))^2 - \mathbf{e}_x^\top W_0 \mathbf{e}_x \tag{54a}$$

$$H_{ei}(t) = 2h_1(\mathbf{x}_m(t))h_i(\mathbf{x}_m(t)) - \mathbf{e}_x^\top W_{i-1}\mathbf{e}_x \;\; \forall i \in [2, r] \tag{54b}$$

with h_i defined in (14a)-(14b). Then, using (9) and (11),

$$\dot{H}_{e1} = 2h_1(\mathbf{x}_m)\frac{\partial h_1}{\partial \mathbf{x}}\Big|_{\mathbf{x}_m}(A_m\mathbf{x}_m + B\mathbf{r}_s) - 2\mathbf{e}_x^\top W_0(A_m\mathbf{e}_x + BF(\mathbf{x})\widetilde{\theta}). \tag{55}$$

The remainder of the proof of claim (a) is split into two cases: $r = 1$ and $r > 1$.

Case 1: $r = 1$

In this case, using (17)-(18), the fact that $h_1(\mathbf{x}_m) > 0$ from Proposition 1, and the fact that $\|\widetilde{\theta}\| \leqslant \Psi(\widehat{\theta})$ from Assumption 1 and (20), (55) becomes

$$\begin{aligned}\dot{H}_{e1} &\geqslant 2h_1(\mathbf{x}_m)(-\alpha_1 h_1(\mathbf{x}_m) + \Delta_{ebsb}) - 2\mathbf{e}_x^\top W_0 A_m\mathbf{e}_x - 2\|\widetilde{\theta}\|\|F(\mathbf{x})^\top B^\top W_0\mathbf{e}_x\| \\ &= -2\alpha_1 H_{e1} + 2h_1(\mathbf{x}_m)\Delta_{ebsb} - \mathbf{e}_x^\top W_1\mathbf{e}_x - 2\|\widetilde{\theta}\|\|F(\mathbf{x})^\top B^\top W_0\mathbf{e}_x\| \\ &\geqslant -2\alpha_1 H_{e1}.\end{aligned} \tag{56}$$

Assumption 3 implies that $H_{e1}(0) \geqslant 0$. Therefore, it follows from Assumption 3 and (56) that $H_{e1}(t) \geqslant 0 \;\forall t \geqslant 0$.

Case 1: $r > 1$

Here, the definition of relative degree r implies that $\frac{\partial h_1}{\partial \mathbf{x}}|_{\mathbf{x}_m}B = 0$, and from Lemma 4, we also have $W_0 B = 0$. Then, using (14b) and (54a)-(54b), (55) becomes

$$\begin{aligned}\dot{H}_{e1} &= 2h_1(\mathbf{x}_m)(h_2(\mathbf{x}_m) - \alpha_1 h_1(\mathbf{x}_m)) - 2\mathbf{e}_x^\top W_0 A_m\mathbf{e}_x \\ &= 2h_1(\mathbf{x}_m)h_2(\mathbf{x}_m) - 2\alpha_1 H_{e1} - \mathbf{e}_x^\top W_1\mathbf{e}_x \\ &= H_{e2} - 2\alpha_1 H_{e1}.\end{aligned} \tag{57}$$

Furthermore, for $i > 1$, the definition of relative degree r implies that $\frac{\partial h_i}{\partial \mathbf{x}}|_{\mathbf{x}_m}B = 0 \;\forall i < r$, and from Lemma 4, we also have $W_i B = 0 \;\forall i < r-1$. Thus, using (9), (11), (14b), and (54b), for each $i \in [2, r-1]$, we have

$$\begin{aligned}\dot{H}_{ei} &= 2(h_2(\mathbf{x}_m) - \alpha_1 h_1(\mathbf{x}_m))h_i(\mathbf{x}_m) + 2h_1(\mathbf{x}_m)(h_{i+1}(\mathbf{x}_m) - \alpha_i h_i(\mathbf{x}_m)) - 2\mathbf{e}_x^\top W_{i-1}A_m\mathbf{e}_x \\ &= 2h_2(\mathbf{x}_m)h_i(\mathbf{x}_m) + 2h_1(\mathbf{x}_m)h_{i+1}(\mathbf{x}_m) - (\alpha_1 + \alpha_i)H_{ei} - \mathbf{e}_x^\top W_i\mathbf{e}_x \\ &= 2h_2(\mathbf{x}_m)h_i(\mathbf{x}_m) + H_{e(i+1)} - (\alpha_1 + \alpha_i)H_{ei}.\end{aligned} \tag{58}$$

Finally, using (54b), (9), (11), (17), (19), the fact that $h_1(\mathbf{x}_m) > 0$ from Proposition 1, and the fact that $\|\widetilde{\theta}\| \leqslant \Psi(\widehat{\theta})$ from Assumption 1 and (20), we obtain

$$\begin{aligned}\dot{H}_{er} &= 2(h_2(\mathbf{x}_m) - \alpha_1 h_1(\mathbf{x}_m))h_r(\mathbf{x}_m) + 2h_1(\mathbf{x}_m)\frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}_m}(A_m\mathbf{x}_m + B\mathbf{r}_s) - 2\mathbf{e}_x^\top W_{r-1}(A_m\mathbf{e}_x + BF(\mathbf{x})\widetilde{\theta}) \\ &\geqslant 2(h_2(\mathbf{x}_m) - \alpha_1 h_1(\mathbf{x}_m))h_r(\mathbf{x}_m) + 2h_1(\mathbf{x}_m)(-\alpha_r h_r(\mathbf{x}_m) + \Delta_{ebsb}) - 2\mathbf{e}_x^\top W_{r-1}(A_m\mathbf{e}_x + BF(\mathbf{x})\widetilde{\theta}) \\ &\geqslant -(\alpha_1 + \alpha_r)H_{er} + 2h_2(\mathbf{x}_m)h_r(\mathbf{x}_m) + 2h_1(\mathbf{x}_m)\Delta_{ebsb} - \mathbf{e}_x^\top W_r\mathbf{e}_x - 2\|\widetilde{\theta}\|\|F(\mathbf{x})^\top B^\top W_{r-1}\mathbf{e}_x\| \\ &\geqslant -(\alpha_1 + \alpha_r)H_{er}.\end{aligned} \tag{59}$$

Assumption 3 implies that $H_{ei}(0) \geqslant 0 \;\forall i \in [1, r]$, and we know from Proposition 1 that for every $i \in [1, r]$, $h_i(\mathbf{x}_m) > 0 \;\forall t \geqslant 0$. Therefore, it follows from Assumption 3, Proposition 1, and (57)-(59) that $H_{er}(t) \geqslant 0 \;\forall t \geqslant 0$, and thus that $H_{e(r-1)}(t) \geqslant 0 \implies H_{e(r-2)}(t) \geqslant 0 \implies \cdots \implies H_{e1}(t) \geqslant 0 \;\forall t \geqslant 0$.

Combining Cases 1 and 2

In both cases, we conclude that $H_{e1}(t) \geqslant 0 \;\forall t \geqslant 0$. From Proposition 1 and Lemma 3, whenever $\mathbf{x} \in \mathcal{C}$ for $\mathcal{C}$ defined in Assumption 2, we then have

$$h(\mathbf{x}_m) = |h_1(\mathbf{x}_m)| \geqslant \sqrt{\mathbf{e}_x^\top W_0\mathbf{e}_x} \geqslant |h(\mathbf{x}) - h(\mathbf{x}_m)| \implies h(\mathbf{x}) \geqslant 0. \tag{60}$$

Claim (a) follows from proof by contradiction: suppose that there exists a time $t_1 > 0$ where $h(\mathbf{x}(t_1)) < 0$. Assumption 3 states that $h(\mathbf{x}(0)) \geqslant 0$, and we know that the signal $h(\mathbf{x}(t))$ must be continuous with time. Therefore, there must exist a time $t_2 \in (0, t_1]$ where $h(\mathbf{x}(t_2)) \in (-c, 0)$ for c defined in Assumption 2. However, this implies that $h(\mathbf{x}(t_2)) < 0$ and $\mathbf{x}(t_2) \in \mathcal{C}$, contradicting the conclusion in (60).

Proof of Claim (b)

From Theorem 1, in order to prove claim (b), we need only to prove that $\mathbf{r}_s(t)$ is bounded. As we know that $\mathbf{r}_*(t)$ is bounded by assumption, it suffices to find a feasible solution to (17) and to show that that feasible solution is bounded. We also know from Proposition 1 and Theorem 1 that $\mathbf{x}_m(t) \in S_r$ and $\|\mathbf{x}(t) - \mathbf{x}_m(t)\| \leqslant E\ \forall t \geqslant 0$. From Assumption 4, therefore, it follows that $\mathbf{x}_m(t)$ and $\mathbf{x}(t)$ are bounded. In what follows, we will consider two cases: $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m} B\| < d$ and $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m} B\| \geqslant d$ for d defined in Assumption 5.

Case 1: $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m} B\| < d$

In this case, it follows immediately from Assumption 5 and all previously-stated results that $\mathbf{r}_s = 0$ is a feasible solution to (17).

Case 2: $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m} B\| \geqslant d$

Here, we can obtain immediately from (17) that

$$\mathbf{r}_s = \left(\delta + \Delta_{ebsb}(\mathbf{x}_m, \mathbf{e}_x, \widehat{\theta}) - \alpha_r h_r(\mathbf{x}_m) - \frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}_m} A_m \mathbf{x}_m\right) \frac{\left(\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m} B\right)^\top}{\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m} B\|^2} \tag{61}$$

is a feasible solution to (17). As previously stated, we know that $\mathbf{x}_m(t)$, $\mathbf{e}_x(t)$, and $\widehat{\theta}(t)$ are bounded. Additionally, we know from Proposition 1 that $h_1(\mathbf{x}_m(t)) \geqslant \frac{\delta}{\prod_{k=1}^r \alpha_k} > 0\ \forall t \geqslant 0$. Thus, we know that $\Delta_{ebsb}(\mathbf{x}_m(t), \mathbf{e}_x(t), \widehat{\theta}(t))$ and all other terms in the numerator of (61) are bounded. Additionally, as we are specifically considering the case where $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m} B\| \geqslant d$, we know that the denominator in (61) has a positive lower bound. Therefore, the feasible solution in (61) is bounded.

Combining Cases 1 and 2

As the above two cases cover all possible scenarios, $\mathbf{r}_s(t)$ is bounded, and claim (b) follows.

Proof of Claim (c)

From Proposition 1, we know that $h_1(\mathbf{x}_m(t)) \geqslant \frac{\delta}{\prod_{k=1}^r \alpha_k} > 0\ \forall t \geqslant 0$, and from Theorem 1 and claim (b) of Theorem 2, we know that $\mathbf{x}(t)$ is bounded and $\lim_{t\to\infty} \|\mathbf{e}_x(t)\| = 0$. Therefore, claim (c) follows immediately from (18)-(19).

Proof of Claim (d)

Proof of the safety objective is immediate from claim (a) of Theorem 2. For the control objective, $\limsup_{t\to\infty} \|\mathbf{x}(t) - \mathbf{x}_*(t)\| \leqslant M$ for a constant $M > 0$ follows immediately from the facts that $\mathbf{x}(t)$ is bounded as previously shown, and $\mathbf{x}_*(t)$ is bounded due to boundedness of $\mathbf{r}_*$ and Hurwitzness of A_m. Furthermore, we have shown that $\lim_{t\to\infty} \|\mathbf{x}(t) - \mathbf{x}_m(t)\| = 0$, and the only way for the parametric uncertainty to influence $\mathbf{x}_m(t)$ is through Δ_{ebsb}. Therefore, the fact that $\lim_{t\to\infty} \Delta_{ebsb}(\mathbf{x}(t), \mathbf{x}_m(t)) = 0$ implies that M is independent of the parametric uncertainty.

Appendix III

III.A Proof of Theorem 3

Consider the functions $g_{\Theta,1}, \ldots, g_{\Theta,N_\Theta}$ and $g_{L,1}, \ldots, g_{L,N_L}$ describing Θ and L respectively as in Lemma 1. Because $\dot{\widehat{\theta}}(t) \in T_\Theta(\widehat{\theta}(t))$ and $\dot{\widehat{\lambda}}(t) \in T_L(\widehat{\lambda}(t))\ \forall t \geqslant 0$, we know that whenever $g_{\Theta,k}(\widehat{\theta}) = 0$ for any k, we have $\frac{d}{dt} g_{\Theta,k}(\widehat{\theta}) = \nabla g_{\Theta,k}(\widehat{\theta})^\top \dot{\widehat{\theta}} \geqslant 0$, and likewise that whenever $g_{L,k}(\widehat{\lambda}) = 0$ for any k, we have $\frac{d}{dt} g_{L,k}(\widehat{\lambda}) = \nabla g_{L,k}(\widehat{\lambda})^\top \dot{\widehat{\lambda}} \geqslant 0$. Results (a) and (b) is immediate from the fact that $\widehat{\theta}(0) \in \Theta$ and $\widehat{\lambda}(0) \in L$.

We propose the following Lyapunov function candidate, which is standard in adaptive control [1]:

$$V = \mathbf{e}_x^\top P \mathbf{e}_x + \frac{\|\widetilde{\theta}\|^2}{\gamma_\theta} + \frac{\|\widetilde{\lambda}\|^2}{\gamma_\lambda}. \tag{62}$$

Since θ_* and λ_* are assumed constant, we have $\dot{\tilde{\theta}} = \dot{\hat{\theta}}$ and $\dot{\tilde{\lambda}} = \dot{\hat{\lambda}}$. Additionally, note that $\text{diag}(\tilde{\lambda})\mathbf{u} = \text{diag}(\mathbf{u})\tilde{\lambda}$. Furthermore, since $\hat{\theta}, \theta_* \in \Theta$, $\hat{\lambda}, \lambda_* \in L$, $\dot{\hat{\theta}} \in T_\Theta(\hat{\theta})$, and $\dot{\hat{\lambda}} \in T_L(\hat{\lambda})$, from Lemma 2 and (23)-(24), we have

$$\tilde{\theta}^\top \dot{\hat{\theta}} \leqslant \tilde{\theta}^\top(-\gamma_\theta F(\mathbf{x})^\top B^\top P \mathbf{e}_x), \tag{63}$$

$$\tilde{\lambda}^\top \dot{\hat{\lambda}} \leqslant \tilde{\lambda}^\top(\gamma_\lambda \text{diag}(\mathbf{u}) B^\top P \mathbf{e}_x). \tag{64}$$

Then, using (13), (22), and (63)-(64), one can show through straightforward algebra that

$$\dot{V} \leqslant -\mathbf{e}_x^\top Q \mathbf{e}_x \leqslant 0 \tag{65}$$

and thus that $V(t) \leqslant V(0)\ \forall t \geqslant 0$. Result (c) is immediate.

Finally, if $\mathbf{r}_s(t)$ is bounded, then $\mathbf{x}_m(t)$ is bounded because A_m is Hurwitz. Then, by boundedness of $\mathbf{e}_x(t)$, we know also that $\mathbf{x}(t)$ is bounded, and thus from (22) that $\mathbf{u}(t)$ and $\dot{\mathbf{e}}_x(t)$ are bounded. Result (d) is immediate from Barbalat's lemma [1].

III.B Proof of Theorem 4

Proof of Claim (a)

Define

$$H_{pi}(t) = h_i(\mathbf{x}(t))\ \forall i \in [1, r] \tag{66}$$

with h_i defined in (25a)-(25b). Then, using (6), (21), and (28),

$$\dot{H}_{pi} = \frac{\partial h_i}{\partial \mathbf{x}}\Big|_{\mathbf{x}} \mathbf{z}_p(\mathbf{x}, \mathbf{r}_s, \hat{\theta}, \hat{\lambda}, \theta_*, \lambda_*). \tag{67}$$

The definition of relative degree r implies that $\frac{\partial h_i}{\partial \mathbf{x}}|_{\mathbf{x}} B = 0\ \forall i < r$. Thus, using (28), (25b) and (66), we have

$$\begin{aligned}\dot{H}_{pi} &= \frac{\partial h_i}{\partial \mathbf{x}}\Big|_{\mathbf{x}} A\mathbf{x} \\ &= h_{i+1}(\mathbf{x}) - \alpha_i h_i(\mathbf{x}) \\ &= H_{p(i+1)} - \alpha_i H_{pi}\ \forall i \in [1, r-1].\end{aligned} \tag{68}$$

Furthermore, whenever $H_{pr} \leqslant \frac{\delta}{3\alpha_r}$, we know from (29) that $\beta_{ebsf} = 1$ and thus that (26) becomes

$$\begin{aligned}\mathbf{r}_s = \arg\min_{\mathbf{r} \in \mathbb{R}^m} &\|\mathbf{r} - \mathbf{r}_*\|^2 \text{ s.t.} \\ \min_{\theta \in \Theta, \lambda \in L} \frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}} &\mathbf{z}_p(\mathbf{x}, \mathbf{r}, \hat{\theta}, \hat{\lambda}, \theta, \lambda) \geqslant -\alpha_r h_r(\mathbf{x}) + \delta\end{aligned} \tag{69}$$

Therefore, using (69) and Assumption 6, whenever $H_{pr} \leqslant \frac{\delta}{3\alpha_r}$, we have

$$\dot{H}_{pr} \geqslant -\alpha_r h_r(\mathbf{x}) + \delta \geqslant -\alpha_r H_{pr}. \tag{70}$$

Assumption 7 implies that $H_{pi}(0) \geqslant 0\ \forall i \in [1, r]$. Therefore, it follows from Assumption 7, (68), and (70) that $H_{pr}(t) \geqslant 0\ \forall t \geqslant 0$, and thus that $H_{p(r-1)}(t) \geqslant 0 \implies H_{p(r-2)}(t) \geqslant 0 \implies \cdots \implies H_{p1}(t) \geqslant 0\ \forall t \geqslant 0$. The claim follows from the definition of H_{pi} in (66).

Proof of Claim (b)

From Theorem 3, in order to prove claim (b), we need only to prove that $\mathbf{r}_s(t)$ is bounded. As in the proof of Theorem 2, since we know that $\mathbf{r}_*(t)$ is bounded by assumption, it suffices to find a feasible solution to (26) and to show that that feasible solution is bounded. We have just shown that $H_{pr}(t) \geqslant 0\ \forall t \geqslant 0$, and therefore that $\mathbf{x}(t) \in S_r\ \forall t \geqslant 0$. It follows immediately from Assumption 8 that $\mathbf{x}(t)$ is bounded. Define $d = \min\{d_1, d_2\}$ with $d_1, d_2 > 0$ defined in Assumptions 9-10. In what follows, we will consider two cases: $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}} B\| < d$ and $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}} B\| \geqslant d$.

Case 1: $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}} B\| < d$

In this case, it follows from Assumption 9, Theorem 1, and (29) that

$$h_r(\mathbf{x}) \geqslant \max\left\{\alpha_{ebsf}(E_2), \frac{2\delta}{3\alpha_r}\right\} \geqslant \max\left\{\alpha_{ebsf}(\|\mathbf{e}_x\|), \frac{2\delta}{3\alpha_r}\right\} \implies \beta_{ebsf} = 0. \tag{71}$$

Then, using (27), (26) is equivalent to

$$\begin{aligned} \mathbf{r}_s = \arg\min_{\mathbf{r}\in\mathbb{R}^m} \|\mathbf{r}-\mathbf{r}_*\|^2 \text{ s.t.} \\ \frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}} [A_m\mathbf{x}+B\mathbf{r}] \geqslant -\alpha_r h_r(\mathbf{x})+\delta. \end{aligned} \tag{72}$$

It then follows from Assumption 10 that $\mathbf{r}_s = 0$ is a feasible solution to (72).

Case 2: $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}}B\| \geqslant d$

Define the quantity

$$\begin{aligned} \xi(\mathbf{x},\mathbf{e}_x,\widehat{\theta},\widehat{\lambda}) = &-\alpha_r h_r(\mathbf{x}) + \delta - \frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}}(A_m - \beta_{ebsf}BK)\mathbf{x} + \beta_{ebsf}\max_{\theta\in\Theta}\frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}}BF(\mathbf{x})\theta \\ &-\beta_{ebsf}\min_{\lambda\in L}\frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}}B\mathrm{diag}(\lambda)\mathrm{diag}(\widehat{\lambda})^{-1}(K\mathbf{x}+F(\mathbf{x})\widehat{\theta}), \end{aligned} \tag{73}$$

dropping the dependence of β_{ebsf} on $\mathbf{x}$ and $\mathbf{e}_x$ for ease of exposition. Then, (26) with (27)-(28) can be rewritten as

$$\begin{aligned} \mathbf{r}_s = \arg\min_{\mathbf{r}\in\mathbb{R}^m} \|\mathbf{r}-\mathbf{r}_*\|^2 \text{ s.t.} \\ \min_{\lambda\in L}\frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}} B\left((1-\beta_{ebsf})\mathrm{diag}(\widehat{\lambda})+\beta_{ebsf}\mathrm{diag}(\lambda)\right)\mathrm{diag}(\widehat{\lambda})^{-1}\mathbf{r} \geqslant \xi(\mathbf{x},\mathbf{e}_x,\widehat{\theta},\widehat{\lambda}). \end{aligned} \tag{74}$$

Now note that $\widehat{\lambda}(t) \in L \ \forall t \geqslant 0$ from Theorem 3, and note that from Assumption 6, $\mathrm{diag}(\lambda)$ is symmetric positive-definite for any $\lambda \in L$ with eigenvalues lower-bounded by $\underline{\lambda}_1,\ldots,\underline{\lambda}_m$. Then, because $\beta_{ebsf} \in [0,1]$, it follows that $(1-\beta_{ebsf})\mathrm{diag}(\widehat{\lambda}) + \beta_{ebsf}\mathrm{diag}(\lambda)$ is symmetric positive-definite for any $\lambda \in L$ with eigenvalues lower-bounded by $\underline{\lambda}_1,\ldots,\underline{\lambda}_m$. Define $\underline{\lambda} = [\underline{\lambda}_1,\ldots,\underline{\lambda}_m]^\top$. Then, whenever $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}}B\| \geqslant d$, we obtain in a straightforward manner from (74) that

$$\mathbf{r}_s = \frac{\max\{\xi(\mathbf{x},\mathbf{e}_x,\widehat{\theta},\widehat{\lambda}),0\}}{\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}}B\mathrm{diag}(\underline{\lambda})(\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}}B)^\top}\left(\frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}}B\mathrm{diag}(\widehat{\lambda})\right)^\top \tag{75}$$

is a feasible solution to (26). As previously stated, we know that $\mathbf{x}(t)$ is bounded. Additionally, we know from Theorem 3 that $\widehat{\theta}(t) \in \Theta$ and $\widehat{\lambda}(t) \in L \ \forall t \geqslant 0$, and that $\mathbf{e}_x(t)$ is bounded. Thus, boundedness of $\mathbf{x}(t)$, $\mathbf{e}_x(t)$, $\widehat{\theta}(t)$, and $\widehat{\lambda}(t)$, along with the guarantee that $\widehat{\lambda}(t)$ is invertible $\forall t \geqslant 0$ from the definition of L in Assumption 6, implies that $\xi(\mathbf{x}(t),\mathbf{e}_x(t),\widehat{\theta}(t),\widehat{\lambda}(t))$ is bounded. As we are specifically considering the case where $\|\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}_m}B\| \geqslant d$ and we know that $\mathrm{diag}(\underline{\lambda})$ is symmetric positive-definite, we also know that the denominator in (75) is lower-bounded. Therefore, the feasible solution in (75) is bounded.

Combining Cases 1 and 2

As the above cases cover all possible scenarios, $\mathbf{r}_s(t)$ is bounded, and claim (b) follows.

Proof of Claim (c)

Define $\mathbf{g}(\mathbf{x}) = (\frac{\partial h_r}{\partial \mathbf{x}})^\top$, and define the error signals $e_h(t) = h(\mathbf{x}(t)) - h(\mathbf{x}_m(t))$ and $\mathbf{e}_g(t) = g(\mathbf{x}(t)) - g(\mathbf{x}_m(t))$. Choosing $\mathbf{r}_s$ according to (26) implies that

$$\min_{\theta\in\Theta,\lambda\in L}\mathbf{g}(\mathbf{x})^\top\left[(1-\beta_{ebsf}(\mathbf{x},\mathbf{e}_x))\mathbf{z}_m(\mathbf{x},\mathbf{r}_s)+\beta_{ebsf}(\mathbf{x},\mathbf{e}_x)\mathbf{z}_p(\mathbf{x},\mathbf{r}_s,\widehat{\theta},\widehat{\lambda},\theta,\lambda)\right] \geqslant -\alpha_r h_r(\mathbf{x})+\delta \tag{76}$$

In particular, we know from Theorem 3 that $\widehat{\theta}(t) \in \Theta$ and $\widehat{\lambda}(t) \in L \ \forall t \geqslant 0$. Thus, (76) holds with $\theta = \widehat{\theta}$ and $\lambda = \widehat{\lambda}$. Substituting (27) and (28) with $\theta = \widehat{\theta}$ and $\lambda = \widehat{\lambda}$ and simplifying, we find that (76) implies that

$$\mathbf{g}(\mathbf{x})^\top(A_m\mathbf{x}+B\mathbf{r}_s) \geqslant -\alpha_r h_r(\mathbf{x})+\delta. \tag{77}$$

Now, consider again the time-varying signals $H_{mi}(t)$, $i \in [1,r]$ defined in (50) and first used in the proof of Proposition 1. For $i < r$, $\dot{H}_{mi}$ is given in (52) as before, and for $i = r$, we can apply (50) and (77) to obtain

$$\begin{aligned} \dot{H}_{mr} &= \mathbf{g}(\mathbf{x}_m)^\top(A_m\mathbf{x}_m+B\mathbf{r}_s) \\ &= \mathbf{g}(\mathbf{x})^\top(A_m\mathbf{x}+B\mathbf{r}_s) - \mathbf{e}_g^\top(A_m\mathbf{x}_m+B\mathbf{r}_s) - \mathbf{g}(\mathbf{x})^\top A_m\mathbf{e}_x \\ &\geqslant -\alpha_r h_r(\mathbf{x})+\delta - \mathbf{e}_g^\top(A_m\mathbf{x}_m+B\mathbf{r}_s) - \mathbf{g}(\mathbf{x})^\top A_m\mathbf{e}_x \end{aligned}$$

$$= -\alpha_r H_{mr} - \alpha_r e_h - \mathbf{e}_g^\top(A_m\mathbf{x}_m + B\mathbf{r}_s) - \mathbf{g}(\mathbf{x})^\top A_m\mathbf{e}_x. \tag{78}$$

Finally, from Theorem 3 and claims (a) and (b), we know that $\lim_{t\to\infty}\|\mathbf{e}_x(t)\| = 0$ and that $\mathbf{x}(t)$, $\mathbf{x}_m(t)$, and $\mathbf{r}_s(t)$ are all bounded. By smoothness of h, we also know that $\lim_{t\to\infty}|e_h(t)| = \lim_{t\to\infty}\|\mathbf{e}_g(t)\| = 0$. It follows immediately that the following three finite times must exist:

(a) $\exists T_1 \geqslant 0$ such that $|\alpha_r e_h(t) + \mathbf{e}_g(t)^\top(A_m\mathbf{x}_m(t) + B\mathbf{r}_s(t)) + \mathbf{g}(\mathbf{x}(t))^\top A_m\mathbf{e}_x(t)| \leqslant \frac{\delta}{12}\ \forall t \geqslant T_1$,

(b) $\exists T_2 \geqslant 0$ such that $|e_h(t)| \leqslant \frac{\delta}{6\alpha_r}\ \forall t \geqslant T_2$, and

(c) $\exists T_3 \geqslant 0$ such that $\alpha_{ebsf}(\|\mathbf{e}_x(t)\|) \leqslant \frac{2\delta}{3\alpha_r}\ \forall t \geqslant T_3$.

Then, for all $t \geqslant T_1$ such that $H_{mr}(t) \leqslant -\frac{\delta}{6\alpha_r}$, we have $\dot{H}_{mr} \geqslant \frac{\delta}{12}$, implying that there exists a finite time $T_4 \geqslant T_1$ such that $H_{mr}(t) \geqslant -\frac{\delta}{6\alpha_r} \implies h_r(\mathbf{x}_m(t)) \geqslant \frac{5\delta}{6\alpha_r}\ \forall t \geqslant T_4 \implies h(\mathbf{x}(t)) \geqslant \frac{2\delta}{3\alpha_r}\ \forall t \geqslant \max\{T_2, T_4\}$. It follows from the definition of β_{ebsf} in (29) that $\beta_{ebsf}(\mathbf{x}(t), \mathbf{e}_x(t)) = 0\ \forall t \geqslant \max\{T_2, T_3, T_4\} := T$.

Proof of Claim (d)

Proof of the safety objective is immediate from claim (a) of Theorem 4. For the control objective, $\limsup_{t\to\infty}\|\mathbf{x}(t) - \mathbf{x}_*(t)\| \leqslant M$ for a constant $M > 0$ follows immediately from the facts that $\mathbf{x}(t)$ is bounded as previously shown, and $\mathbf{x}_*(t)$ is bounded due to boundedness of $\mathbf{r}_*$ and Hurwitzness of A_m. Furthermore, whenever $\beta_{ebsf} = 0$, (26) becomes equivalent to (72), meaning that for T from claim (c), $\mathbf{r}_s(t)$ and therefore $\mathbf{x}_m(t)$ are independent of the parametric uncertainty for all $t \geqslant T$ except for initial conditions at time T. Independence of M from the parametric uncertainty follows from the fact that $\lim_{t\to\infty}\|\mathbf{x}(t) - \mathbf{x}_m(t)\| = 0$.

Appendix IV

In this appendix, we show how the EBSF in (26) can be rewritten as a quadratic programming problem. Define $\xi(\mathbf{x}, \mathbf{e}_x, \widehat{\theta}, \widehat{\lambda})$ as in (73), and further define $\mathbf{w}(\mathbf{x}) = (\frac{\partial h_r}{\partial \mathbf{x}}|_{\mathbf{x}} B)^\top$ and $\mathbf{z}_* = \mathrm{diag}(\widehat{\lambda})^{-1}\mathbf{r}_*$. Then, dropping dependencies for ease of exposition, (26) is equivalent to

$$\mathbf{r}_s = \mathrm{diag}(\widehat{\lambda})\mathbf{z}_s, \tag{79a}$$

$$\begin{aligned} \mathbf{z}_s = \arg\min_{\mathbf{z}\in\mathbb{R}^m} & (\mathbf{z} - \mathbf{z}_*)^\top \mathrm{diag}(\widehat{\lambda})^2(\mathbf{z} - \mathbf{z}_*) \text{ s.t.} \\ \min_{\lambda\in L} & \mathbf{w}^\top \left((1 - \beta_{ebsf})\mathrm{diag}(\widehat{\lambda}) + \beta_{ebsf}\mathrm{diag}(\lambda)\right)\mathbf{z} \geqslant \xi. \end{aligned} \tag{79b}$$

Then, using the index j to refer to the jth element of a vector, (79b) can be further rewritten as

$$\begin{aligned} \mathbf{z}_s = \arg\min_{\mathbf{z}\in\mathbb{R}^m} & (\mathbf{z} - \mathbf{z}_*)^\top \mathrm{diag}(\widehat{\lambda})^2(\mathbf{z} - \mathbf{z}_*) \text{ s.t.} \\ \min_{\lambda\in L} & \sum_{j=1}^m ((1 - \beta_{ebsf})\widehat{\lambda}_j + \beta_{ebsf}\lambda_j) w_j z_j \geqslant \xi, \end{aligned} \tag{80}$$

and using Assumption 6 and the fact that $\beta_{ebsf} \in [0, 1]$, (80) can be finally rewritten as

$$\begin{aligned} \mathbf{z}_s = \arg\min_{\mathbf{z}\in\mathbb{R}^m} & (\mathbf{z} - \mathbf{z}_*)^\top \mathrm{diag}(\widehat{\lambda})^2(\mathbf{z} - \mathbf{z}_*) \text{ s.t.} \\ \sum_{j=1}^m & \left((1 - \beta_{ebsf})\widehat{\lambda}_j + \beta_{ebsf}\begin{Bmatrix}\overline{\lambda}_j, & w_j z_j < 0 \\ \underline{\lambda}_j, & w_j z_j \geqslant 0\end{Bmatrix}\right) w_j z_j \geqslant \xi. \end{aligned} \tag{81}$$

Now, we examine how the optimization problem in (81) behaves when $\mathrm{diag}(\mathbf{w})\mathbf{z}$ is in each orthant of $\mathbb{R}^m$. Consider the set of indices $\mathbb{J} \subseteq \{1, \ldots, m\}$ such that $w_j z_j \geqslant 0\ \forall j \in \mathbb{J}$ and $w_j z_j < 0\ \forall j \in \{1, \ldots, m\}\backslash\mathbb{J}$. Define the vector $\ell_{\mathbb{J}} \in \mathbb{R}^m$ whose jth element is $(1 - \beta_{ebsf})\widehat{\lambda}_j + \beta_{ebsf}\begin{Bmatrix}\overline{\lambda}_j, & j\notin\mathbb{J} \\ \underline{\lambda}_j, & j\in\mathbb{J}\end{Bmatrix}$. There are 2^m unique sets $\mathbb{J}$ corresponding to $\mathrm{diag}(\mathbf{w})\mathbf{z}$ being in each orthant of $\mathbb{R}^m$, and thus $\ell_{\mathbb{J}}$ can point in 2^m directions which are linearly independent unless $\beta_{ebsf} = 0$, in which case all 2^m possibilities for $\ell_{\mathbb{J}}$ coincide. As the constraint in (81) should be continuous on the boundaries between orthants and the vectors $\ell_{\mathbb{J}}$ are linearly independent in general, one would expect that each of the 2^m vectors $\ell_{\mathbb{J}}$ would produce its own half-space constraint and that the solution to (81) should lie in the intersection of all of the 2^m resulting half-spaces.

There are two orthants which are exceptions to this reasoning: the first orthant, where $\mathbb{J} = \{1, \ldots, m\}$, and the last orthant, where $\mathbb{J} = \varnothing$. First, consider the case where $\xi \leqslant 0$. In this case, as all elements of $\ell_{\mathbb{J}}$ are positive for every possible $\mathbb{J}$, any $\operatorname{diag}(\mathbf{w})\mathbf{z}$ lying in the first orthant is a feasible solution to (81), and this orthant does not produce a half-space constraint. Instead, one can see through simple geometric reasoning that the half-space constraint produced by the first orthant permits $\operatorname{diag}(\mathbf{w})\mathbf{z}$ to lie in the first orthant, and thus the one constraint is sufficient for both orthants. Second, consider the case where $\xi > 0$. In this case, any $\operatorname{diag}(\mathbf{w})\mathbf{z}$ lying in the last orthant is infeasible, and this orthant does not produce a half-space constraint because no feasible solution can lie in it. Instead, one can see through simple geometric reasoning that the half-space constraint produced by the first orthant prevents $\operatorname{diag}(\mathbf{w})\mathbf{z}$ from lying in the last orthant, and thus the one constraint is sufficient for both orthants.

In summary, each of the $2^m - 2$ possibilities for $\mathbb{J}$ such that $\mathbb{J} \neq \varnothing$ and $\mathbb{J} \neq \{1, \ldots, m\}$ produces its own half-space constraint which must be considered separately from the others. Index the $2^m - 2$ such possible sets as $\mathbb{J}^i$, $i \in \{1, \ldots, 2^m - 2\}$. Additionally, when $\xi \leqslant 0$, $\mathbb{J} = \varnothing$ produces a half-space constraint and $\mathbb{J} = \{1, \ldots, m\}$ does not, while when $\xi > 0$, $\mathbb{J} = \{1, \ldots, m\}$ produces a half-space constraint and $\mathbb{J} = \varnothing$ does not. Thus, (81) can be rewritten with only half-space constraints, and is equivalent to the quadratic programming problem

$$\begin{aligned}
\mathbf{z}_s = \arg\min_{\mathbf{z} \in \mathbb{R}^m} & (\mathbf{z} - \mathbf{z}_*)^\top \operatorname{diag}(\hat{\lambda})^2 (\mathbf{z} - \mathbf{z}_*) \text{ s.t.} \\
& \ell_{\mathbb{J}^i}^\top \operatorname{diag}(\mathbf{w})\mathbf{z} \geqslant \xi \ \forall i \in \{1, \ldots, 2^m - 2\}, \\
& \begin{cases} \ell_{\varnothing}^\top \operatorname{diag}(\mathbf{w})\mathbf{z} \geqslant \xi, & \xi \leqslant 0 \\ \ell_{\{1,\ldots,m\}}^\top \operatorname{diag}(\mathbf{w})\mathbf{z} \geqslant \xi, & \xi > 0 \end{cases}
\end{aligned} \tag{82}$$

which has $2^m - 1$ constraints.

Appendix V

V.A Ideal Control Design if All Parameters Are Known

For the simulation setting in Section VI, if all parameters were known, one could satisfy the control and safety objectives as follows. Choose $\mathbf{u}$ as

$$\mathbf{u} = \frac{1}{\lambda_*}(K\mathbf{x} + \mathbf{r}_s + F(\mathbf{x})\theta_*) \tag{83}$$

where K is chosen by LQR on the dynamics (A, B). Define $A_m = A + BK$. Then, choose the nominal tracking reference input as

$$\mathbf{r}_* = T(s)^{-1}\mathbf{p}_*, \quad T(s) = \begin{bmatrix} I_2 & 0_{2\times 2} \end{bmatrix} (sI_4 - A_m)^{-1} B, \tag{84}$$

where $T(s)$ is the transfer function from $\mathbf{r}_s$ to $\mathbf{p}$. Finally, describe the safe set S as in (2a)-(2c) using the function

$$h(\mathbf{x}) = \sqrt{(x - x_{\text{pillar}})^2 + (y - y_{\text{pillar}})^2} - r_{\text{pillar}}. \tag{85}$$

Then, as h has relative degree 2 with respect to the dynamics in (37), choose $\mathbf{r}_s$ according to

$$h_1(\mathbf{x}) = h(\mathbf{x}), \quad h_2(\mathbf{x}) = \frac{\partial h_1}{\partial \mathbf{x}}\Big|_{\mathbf{x}} A_m\mathbf{x} + \alpha_1 h_1(\mathbf{x}), \tag{86a}$$

$$\begin{aligned} \mathbf{r}_s = \arg\min\nolimits_{\mathbf{r} \in \mathbb{R}^2} & \|\mathbf{r} - \mathbf{r}_*\|^2 \text{ s.t.} \\ \frac{\partial h_2}{\partial \mathbf{x}}\Big|_{\mathbf{x}} & (A_m\mathbf{x} + B\mathbf{r}) \geqslant -\alpha_2 h(\mathbf{x}) + \delta \end{aligned} \tag{86b}$$

with $\alpha_1, \alpha_2, \delta > 0$.

V.B aCBF Implementation

In our simulations, for simplicity and for ease of comparison, we integrate adaptive control with the aCBF approach in a way which mirrors the structure of our EBSB and EBSF control designs, differing from the aCLF-aCBF-QP approach proposed in [25]. In this section, we consider Problem 2 with the plant given in (6), and treat Problem 1 as merely a special case. Thus, we also extend the approach in [25] to handle uncertainties in the input matrix. We first require Assumptions 1-6, and design a standard model-reference adaptive controller using the safe reference model in (9) and the control design in (21)-(24). Then, substituting (21) into (6) and simplifying, we obtain

$$\dot{\mathbf{x}} = A_m\mathbf{x} + B(\mathbf{r}_s + F(\mathbf{x})(\hat{\theta} - \theta_*) - \operatorname{diag}(\mathbf{u})(\hat{\lambda} - \lambda_*)). \tag{87}$$

Now, to ensure safety of the plant, we construct a high-order CBF as in (25a)-(25b) for any $\alpha_1, \ldots, \alpha_{r-1} > 0$, and following [25], we design the adaptive control barrier function as

$$h_a(\mathbf{x}) = \begin{cases} \Delta_{acbf}^2, & h_r(\mathbf{x}) \geqslant \Delta_{acbf}, \\ \Delta_{acbf}^2 - (h_r(\mathbf{x}) - \Delta_{acbf})^2, & h_r(\mathbf{x}) < \Delta_{acbf} \end{cases} \tag{88}$$

for a constant $\Delta_{acbf} > 0$ to be specified later. Then, we choose $\mathbf{r}_s$ according to

$$\begin{aligned} \mathbf{r}_s &= \arg\min_{\mathbf{r} \in \mathbb{R}^m} \|\mathbf{r} - \mathbf{r}_*\|^2 \text{ s.t.} \\ \frac{\partial h_a}{\partial \mathbf{x}}\Big|_{\mathbf{x}} & \left(A_m \mathbf{x} + B \left(\mathbf{r} + F(\mathbf{x})(\widehat{\theta} - \widehat{\theta}_s) - \operatorname{diag}(\mathbf{u})(\widehat{\lambda} - \widehat{\lambda}_s) \right) \right) \geqslant 0 \end{aligned} \tag{89}$$

which is a quadratic programming problem in $\mathbf{r}$ by substituting (21) for $\mathbf{u}$. $\widehat{\theta}_s$ and $\widehat{\lambda}_s$ are auxiliary parameter estimates with adaptive laws given by

$$\dot{\widehat{\theta}}_s = \operatorname{proj}_{T_\Theta(\widehat{\theta}_s)} \left[\gamma_{\theta,s} \left(\frac{\partial h_a}{\partial \mathbf{x}}\Big|_{\mathbf{x}} B F(\mathbf{x}) \right)^\top \right], \tag{90}$$

$$\dot{\widehat{\lambda}}_s = \operatorname{proj}_{T_L(\widehat{\lambda}_s)} \left[-\gamma_{\lambda,s} \left(\frac{\partial h_a}{\partial \mathbf{x}}\Big|_{\mathbf{x}} B \operatorname{diag}(\mathbf{u}) \right)^\top \right] \tag{91}$$

for any adaptive gains $\gamma_{\theta,s}, \gamma_{\lambda,s} > 0$. Finally, as in [25], defining $\widetilde{\theta}_s := \widehat{\theta}_s - \theta_*$ and $\widetilde{\lambda}_s := \widehat{\lambda}_s - \lambda_*$, we require $\Delta_{acbf} \geqslant \sqrt{\frac{1}{2\gamma_{\theta,s}} \|\widetilde{\theta}_s(0)\|^2 + \frac{1}{2\gamma_{\lambda,s}} \|\widetilde{\lambda}_s(0)\|^2}$. The proof of safety follows the proof of Theorem 3 in [25].

V.C RaCBF Implementation

In our simulations, for simplicity and for ease of comparison, we integrate adaptive control with the RaCBF approach in a way which mirrors the structure of our EBSB and EBSF control designs, differing from the approach proposed in [26]. In this section, we consider Problem 2 with the plant given in (6), and treat Problem 1 as merely a special case. Thus, we also extend the approach in [26] to handle uncertainties in the input matrix. We first require Assumptions 1-6, and design a standard model-reference adaptive controller using the safe reference model in (9) and the control design in (21)-(24). Then, substituting (21) into (6) and simplifying, we obtain the closed-loop plant in (87).

Now, to ensure safety of the plant, we construct a high-order CBF as in (25a)-(25b) for any $\alpha_1, \ldots, \alpha_{r-1} > 0$, and we choose $h_r(\mathbf{x})$ as the robust adaptive control barrier function. Then, we choose $\mathbf{r}_s$ according to

$$\begin{aligned} \mathbf{r}_s &= \arg\min_{\mathbf{r} \in \mathbb{R}^m} \|\mathbf{r} - \mathbf{r}_*\|^2 \text{ s.t.} \\ \frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}} & \left(A_m \mathbf{x} + B \left(\mathbf{r} + F(\mathbf{x})(\widehat{\theta} - \widehat{\theta}_s) - \operatorname{diag}(\mathbf{u})(\widehat{\lambda} - \widehat{\lambda}_s) \right) \right) \geqslant -\alpha_r h_r(\mathbf{x}) + \Delta_{racbf}, \end{aligned} \tag{92}$$

which is a quadratic programming problem in $\mathbf{r}$ by substituting (21) for $\mathbf{u}$, for a constant $\Delta_{racbf} > 0$ to be specified later. $\widehat{\theta}_s$ and $\widehat{\lambda}_s$ are auxiliary parameter estimates with adaptive laws given by

$$\dot{\widehat{\theta}}_s = \operatorname{proj}_{T_\Theta(\widehat{\theta}_s)} \left[\gamma_{\theta,s} \left(\frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}} B F(\mathbf{x}) \right)^\top \right], \tag{93}$$

$$\dot{\widehat{\lambda}}_s = \operatorname{proj}_{T_L(\widehat{\lambda}_s)} \left[-\gamma_{\lambda,s} \left(\frac{\partial h_r}{\partial \mathbf{x}}\Big|_{\mathbf{x}} B \operatorname{diag}(\mathbf{u}) \right)^\top \right] \tag{94}$$

for any adaptive gains $\gamma_{\theta,s}, \gamma_{\lambda,s} > 0$. Finally, as in [26], defining $\widetilde{\theta}_s := \widehat{\theta}_s - \theta_*$ and $\widetilde{\lambda}_s := \widehat{\lambda}_s - \lambda_*$, we require $\Delta_{racbf} \geqslant \sup_{t \geqslant 0} \frac{\alpha_r}{2} (\frac{1}{\gamma_{\theta,s}} \|\widetilde{\theta}_s(t)\|^2 + \frac{1}{\gamma_{\lambda,s}} \|\widetilde{\lambda}_s(t)\|^2)$. The proof of safety follows the proof of Theorem 2 in [26].

V.D Set Membership Identification Implementation

Our implementation of Set Membership Identification maintains high-probability upper and lower bounds on each uncertain parameter, and updating the bounds in response to input-output data. We will describe our SMID approach for Problem 2 here, and the approach for Problem 2 is simply a special case. To begin, we Euler-discretize the dynamics in (6) with a time step $\Delta t > 0$, resulting in the discrete-time dynamics

$$\mathbf{x}_{k+1} = (I_n + A\Delta t)\mathbf{x}_k + B\Delta t(\operatorname{diag}(\lambda_*)\mathbf{u}_k - F(\mathbf{x}_k)\theta_*) + \mathbf{w}_{k+1} \tag{95}$$

where $\mathbf{w}_{k+1}$ represents the error in the discretization. Define the quantities $\mathbf{y}_{k+1} = (B\Delta t)^\dagger(\mathbf{x}_{k+1} - (I_n + A\Delta t)\mathbf{x}_k)$, $\Phi_k = [-F(\mathbf{x}_k), \mathrm{diag}(\mathbf{u}_k)]$, $\xi_* = [\theta_*^\top, \lambda_*^\top]^\top$, and $\eta_{k+1} = (B\Delta t)^\dagger \mathbf{w}_{k+1}$, where $\dagger$ represents the Moore-Penrose pseudo-inverse. We then make the following modeling assumptions:

Assumption 11. *B has linearly independent columns, so that $B^\dagger = (B^\top B)^{-1}B^\top$.*

Assumption 12. *At each time step $k \geqslant 0$, the scaled discretization error η_{k+1} is drawn from a zero-mean Gaussian distribution with symmetric positive-definite covariance Σ_{k+1}.*

Then, (95) is equivalent to

$$\mathbf{y}_{k+1} = \Phi_k \zeta_* + \eta_{k+1} \tag{96}$$

with $\Phi_k \in \mathbb{R}^{m\times(p+m)}$ and $\mathbf{y}_{k+1} \in \mathbb{R}^m$ measured, $\zeta_* \in \mathbb{R}^{p+m}$ the unknown parameter vector, and $\eta_{k+1} \sim \mathcal{N}(0, \Sigma_{k+1})\ \forall k \geqslant 0$.

One can show using Bayes' rule that, if the probability distribution of ζ_* at time step k is characterized by the mean and covariance $\widehat{\zeta}_k$ and P_k, then the updated mean and covariance in response to the data $(\Phi_k, \mathbf{y}_{k+1})$ are given by

$$\widehat{\zeta}_{k+1} = \widehat{\zeta}_k + P_k\Phi_k^\top(\Sigma_{k+1} + \Phi_k P_k \Phi_k^\top)^{-1}(\mathbf{y}_{k+1} - \Phi_k\widehat{\zeta}_k), \tag{97a}$$

$$P_{k+1} = P_k - P_k\Phi_k^\top(\Sigma_{k+1} + \Phi_k P_k \Phi_k^\top)^{-1}\Phi_k P_k. \tag{97b}$$

Now, given a mean and covariance for the probability distribution of ζ_* at each time step, we can calculate high-probability upper and lower bounds for each element of ζ_* as follows.

At time step k, we have $\zeta_* \sim \mathcal{N}(\widehat{\zeta}_k, P_k)$. Then, given any $\delta > 0$, we want to find $\underline{\zeta}_{ki}$ and $\overline{\zeta}_{ki}$ such that $\mathbb{P}[\zeta_{*i} \in [\underline{\zeta}_{ki}, \overline{\zeta}_{ki}]\ \forall i \in [1, p+m]] \geqslant 1-\delta$, where $\mathbb{P}[X]$ denotes the probability of event X. First, we diagonalize the covariance matrix using SVD to obtain

$$P_k = V_k D_k V_k^\top, \tag{98}$$

where V_k is orthogonal and $D_k = \mathrm{diag}(\mathbf{d}_k)$. Then, define the vectors

$$\mathbf{v}_* = V_k^\top \zeta_*, \quad \widehat{\mathbf{v}}_k = V_k^\top \widehat{\zeta}_k. \tag{99}$$

It is straightforward to show that $\mathbf{v}_* \sim \mathcal{N}(\widehat{\mathbf{v}}_k, D_k)$, or in other words, that $v_{*i} \sim \mathcal{N}(\widehat{v}_{ki}, d_{ki})$ for each $i \in [1, p+m]$. Given that the elements v_{*i} are independent for all i, we can straightforwardly produce high-probability bounds using their cumulative distribution functions: defining

$$\widetilde{v}_{ki} = \sqrt{2d_{ki}}\,\mathrm{erf}^{-1}\left(1 - \frac{\delta}{p+m}\right) \tag{100}$$

where erf^{-1} denotes the inverse error function, it is easy to show that $\mathbb{P}[v_{*i} \in [\widehat{v}_{ki} - \widetilde{v}_{ki}, \widehat{v}_{ki} + \widetilde{v}_{ki}]] \geqslant 1 - \frac{\delta}{p+m}$ for each i. Finally, denoting the (i,j)th element of V_k as v_{kij}, for each $j \in [1, p+m]$, with probability at least $1 - \frac{\delta}{p+m}$, we have

$$\zeta_{*j} = \sum_{j=1}^{p+m} v_{kij} v_{*j} \geqslant \sum_{j=1}^{p+m} v_{kij} \begin{Bmatrix} \widehat{v}_{kj} - \widetilde{v}_{kj}, & v_{kij} \geqslant 0 \\ \widehat{v}_{kj} + \widetilde{v}_{kj}, & v_{kij} < 0 \end{Bmatrix} = \sum_{j=1}^{p+m} (v_{kij}\widehat{v}_{kj} - |v_{kij}|\widetilde{v}_{kj}), \tag{101a}$$

$$\zeta_{*j} = \sum_{j=1}^{p+m} v_{kij} v_{*j} \leqslant \sum_{j=1}^{p+m} v_{kij} \begin{Bmatrix} \widehat{v}_{kj} + \widetilde{v}_{kj}, & v_{kij} \geqslant 0 \\ \widehat{v}_{kj} - \widetilde{v}_{kj}, & v_{kij} < 0 \end{Bmatrix} = \sum_{j=1}^{p+m} (v_{kij}\widehat{v}_{kj} + |v_{kij}|\widetilde{v}_{kj}). \tag{101b}$$

Then, defining $\underline{\zeta}_{ki} = \sum_{j=1}^{p+m}(v_{kij}\widehat{v}_{kj} - |v_{kij}|\widetilde{v}_{kj})$ and $\overline{\zeta}_{ki} = \sum_{j=1}^{p+m}(v_{kij}\widehat{v}_{kj} + |v_{kij}|\widetilde{v}_{kj})$, a simple union bound shows that $\mathbb{P}[\zeta_{*i} \in [\underline{\zeta}_{ki}, \overline{\zeta}_{ki}]\ \forall i \in [1, p+m]] \geqslant 1-\delta$.

As a final step before implementation, we must initialize $\widehat{\zeta}_0$ and P_0. We can simply choose $\widehat{\zeta}_0 = [\widehat{\theta}(0), \widehat{\lambda}(0)]^\top$, using the initializations from the adaptive laws in (23)-(24). Then, to initialize P_0, we require Assumptions 1-6, and choose P_0 such that the element-wise upper and lower bounds on ζ_* resulting from (98)-(101b) fully enclose Θ and L.

Acknowledgment

This work was supported by the Boeing Strategic University Initiative and by the Air Force Research Laboratory.

References

[1] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*. NJ: Dover Publications, 2005, (original publication by Prentice-Hall Inc., 1989).

[2] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Prentice-Hall, 1996.

[3] S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence and Robustness*. Prentice-Hall, 1989.

[4] J.-J. E. Slotine and W. Li, *Applied nonlinear control*. Prentice hall Englewood Cliffs, NJ, 1991.

[5] M. Krstic, P. V. Kokotovic, and I. Kanellakopoulos, *Nonlinear and adaptive control design*. John Wiley & Sons, Inc., 1995.

[6] K. J. Åström and B. Wittenmark, *Adaptive control*. Courier Corporation, 2013.

[7] 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)*, 2019, pp. 3420–3431.

[8] Q. Nguyen, A. Hereid, J. W. Grizzle, A. D. Ames, and K. Sreenath, “3d dynamic walking on stepping stones with control barrier functions,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 827–834.

[9] G. Gunter, M. Nice, M. Bunting, J. Sprinkle, and D. B. Work, “Experimental testing of a control barrier function on an automated vehicle in live multi-lane traffic,” in *2022 2nd Workshop on Data-Driven and Intelligent Cyber-Physical Systems for Smart Cities Workshop (DI-CPS)*, 2022, pp. 31–35.

[10] A. Alan, A. J. Taylor, C. R. He, A. D. Ames, and G. Orosz, “Control barrier functions and input-to-state safety with application to automated vehicles,” *IEEE Transactions on Control Systems Technology*, vol. 31, no. 6, pp. 2744–2759, 2023.

[11] S. Zhang, K. Garg, and C. Fan, “Neural graph control barrier functions guided distributed collision-avoidance multi-agent control,” in *Proceedings of The 7th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, J. Tan, M. Toussaint, and K. Darvish, Eds., vol. 229. PMLR, 06–09 Nov 2023, pp. 2373–2392. [Online]. Available: https://proceedings.mlr.press/v229/zhang23h.html

[12] T. Gurriet, A. Singletary, J. Reher, L. Ciarletta, E. Feron, and A. Ames, “Towards a framework for realizable safety critical control through active set invariance,” in *2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS)*, 2018, pp. 98–106.

[13] 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, analysis and Design of Hybrid Systems ADHS. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896315024106

[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, 2019.

[15] A. Alan, T. G. Molnar, A. D. Ames, and G. Orosz, “Parameterized barrier functions to guarantee safety under uncertainty,” *IEEE Control Systems Letters*, vol. 7, pp. 2077–2082, 2023.

[16] M. Ohnishi, L. Wang, G. Notomista, and M. Egerstedt, “Barrier-certified adaptive reinforcement learning with applications to brushbot navigation,” *IEEE Transactions on Robotics*, vol. 35, no. 5, pp. 1186–1205, 2019.

[17] R. Cheng, G. Orosz, R. M. Murray, and J. W. Burdick, “End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 3387–3395, Jul. 2019. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/4213

[18] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, “A general safety framework for learning-based control in uncertain robotic systems,” *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2737–2752, 2019.

[19] D. D. Fan, J. Nguyen, R. Thakker, N. Alatur, A.-a. Agha-mohammadi, and E. A. Theodorou, “Bayesian learning-based adaptive control for safety critical systems,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 4093–4099.

[20] A. Taylor, A. Singletary, Y. Yue, and A. Ames, "Learning for safety-critical control with control barrier functions," in *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, ser. Proceedings of Machine Learning Research, A. M. Bayen, A. Jadbabaie, G. Pappas, P. A. Parrilo, B. Recht, C. Tomlin, and M. Zeilinger, Eds., vol. 120. PMLR, 10–11 Jun 2020, pp. 708–717. [Online]. Available: https://proceedings.mlr.press/v120/taylor20a.html

[21] A. Alan, T. G. Molnar, E. Daş, A. D. Ames, and G. Orosz, "Disturbance observers for robust safety-critical control with control barrier functions," *IEEE Control Systems Letters*, vol. 7, pp. 1123–1128, 2023.

[22] E. Daş and J. W. Burdick, "Robust control barrier functions using uncertainty estimation with application to mobile robots," *IEEE Transactions on Automatic Control*, vol. 70, no. 7, pp. 4766–4773, 2025.

[23] A. Isaly, O. S. Patil, H. M. Sweatland, R. G. Sanfelice, and W. E. Dixon, "Adaptive safety with a rise-based disturbance observer," *IEEE Transactions on Automatic Control*, vol. 69, no. 7, pp. 4883–4890, 2024.

[24] J. Sun, J. Yang, and Z. Zeng, "Safety-critical control with control barrier function based on disturbance observer," *IEEE Transactions on Automatic Control*, vol. 69, no. 7, pp. 4750–4756, 2024.

[25] A. J. Taylor and A. D. Ames, "Adaptive safety with control barrier functions," in *2020 American Control Conference (ACC)*, 2020, pp. 1399–1405.

[26] 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, 2021.

[27] B. T. Lopez and J.-J. E. Slotine, "Unmatched control barrier functions: Certainty equivalence adaptive safety," in *2023 American Control Conference (ACC)*, 2023, pp. 3662–3668.

[28] Y. Wang and X. Xu, "Adaptive safety-critical control for a class of nonlinear systems with parametric uncertainties: A control barrier function approach," *Systems & Control Letters*, vol. 188, p. 105798, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167691124000860

[29] J. Autenrieb and A. Annaswamy, "Safe and stable adaptive control for a class of dynamic systems," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, 2023, pp. 5059–5066.

[30] J. A. Solano-Castellanos, P. A. Fisher, and A. Annaswamy, "Safe and stable formation control with autonomous multi-agents using adaptive control (extended version)," 2025. [Online]. Available: https://arxiv.org/abs/2403.15674

[31] W. Xiao and C. Belta, "Control barrier functions for systems with high relative degree," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 474–479.

[32] T. G. Molnar, R. K. Cosner, A. W. Singletary, W. Ubellacker, and A. D. Ames, "Model-free safety-critical control for robotic systems," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 944–951, 2022.

[33] T. G. Molnar and A. D. Ames, "Safety-critical control with bounded inputs via reduced order models," in *2023 American Control Conference (ACC)*, 2023, pp. 1414–1421.

[34] M. H. Cohen, T. G. Molnar, and A. D. Ames, "Safety-critical control for autonomous systems: Control barrier functions via reduced-order models," *Annual Reviews in Control*, vol. 57, p. 100947, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578824000166

[35] M. H. Cohen, N. Csomay-Shanklin, W. D. Compton, T. G. Molnar, and A. D. Ames, "Safety-critical controller synthesis with reduced-order models," in *2025 American Control Conference (ACC)*, 2025, pp. 5216–5221.

[36] E. Lavretsky and M. Menner, "Servo-controllers for linear time-invariant systems with operational constraints," in *2025 American Control Conference (ACC)*, 2025, pp. 4909–4916.