2024 International Joint Conference on Neural Networks (IJCNN) | 979-8-3503-5931-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/IJCNN60899.2024.10650058

Robust Barrier-Certified Safe Learning-based Adaptive Control for Multi-Agent Systems in Presence of Uncertain Environments

Shawon Dey and Hao Xu

Abstract— **This paper develops a decentralized safe learning-based adaptive control for multi-agent systems operating in uncertain environments. Due to the fact that the safe set of the local agent depends on the other agents' states, the uncertainty of these external systems leads to an uncertain safe set. As a result, the safe control design of the local agent system in a multi-agent setting becomes intractable. To address this challenge, a neural network (NN) based adaptive observer is developed to estimate the state of the unknown external agents. Based on the state estimation of external agents, an adaptive interplay control barrier function (AI-CBF) is formulated. The AI-CBF is designed by considering both the local agent's state and the NN-based estimated states of external agents. Notably, the limitation of forward invariance for the approximated safe set without guaranteeing the same for the actual safe set is acknowledged in AI-CBF design. The AI-CBF incorporates the bounds on state estimation errors of external agents to guarantee the strict safety requirements of the local agent while learning external agent dynamics. Then, a control framework is formulated using a quadratic programming (QP) method that integrates the safety and stability of the system.**

I. Introduction

Ensuring safety stands as a paramount consideration in real-world control design especially for automated systems such as unmanned aerial vehicles (UAVs) [1], [2] and autonomous ground vehicles [3]. With the potential risks associated with these technologies, it becomes crucial to emphasize the development of robust safety measures. Since the safety concept has been introduced to real-time system design in [4], a significant body of research has been conducted in the fields of safe control [5] and safe learning [6], [7] systems. The certification process in safety relies on the robust positive control invariant safe sets. Two distinct approaches for state constraint set certification are the Control Barrier Function (CBF) [8] and the Hamilton-Jacobi reachability [9] analysis. The CBF-based method [10] utilizes a CBF that enforces constraints on the behavior of the system. This CBF approach is successfully adopted in various applications, such as collision avoidance [11], safe lane change maneuvers [12], adaptive cruise control [13], and so on. Most of these existing techniques rely on precise agent state information to guarantee safety. However, implementing CBF to ensure the forward invariance of a safe set becomes challenging when uncertainty exists in the environment. Specifically, when a local agent shares its environment with other unknown agents, achieving safety for the local agent becomes challenging due to the environment uncertainty associated with the safe set. A recent study by Marvi et al. [10] proposed a safe control design for systems operating in shared environments. They designed a control strategy to ensure the forward invariance of the intersection of the actual safe set and the approximated safe set. The actual safe set is formed by using the external agents' states. However, in practice, the external agents' states cannot be measured directly and only the system's input and output are measurable. Lack of full state information on external systems, achieving the intersection of the actual and approximated safe set is very difficult and even impossible. This assumption of the actual safe set may violate the safety of the local agent when the exact full state information of the external agents is not available. In this paper, a novel neural network-based adaptive observer is designed to estimate external agents' states and further used for generating safety sets for local agents. However, the external agents' systems are unknown and uncertain. So the known model-based observer design can not be directly implemented in this system. This problem has been addressed in [14] which combined observer design with neural network-based system identification. This study used a static approximation of the gradient by assuming the system state to be constant and remain unchanged over time. However, this presumption lacks practicality and efficiency in practice. In the real world, external agents' states are usually time-varying and the local agent has no authority over these external agents' actions, i.e. control inputs. To address these issues, an online neural network (NN) based adaptive observer has been designed to learn the unknown dynamics as well as states of external agents. A modified objective function for the gradient descent weight update mechanism is designed while addressing the limitation of static gradient approximation. Additionally, a mathematical proof of the stability of the observer-based identification of the external system is provided. Then, the existing control barrier function is reformulated using the local agents' system state and observed external system states. Using the reformulated adaptive interplay control barrier function (AI-CBF), the safety criteria are designed as a function based on the local agents' own state and also the state of the external agents obtained from the observer. As the actual safe set remains unknown at the initial learning stage, and the local agent relies solely on the approximated safe set, an error bound is introduced for the estimated external agent state. This error bound is then incorporated in the AI-CBF to avoid the violation of strict safety and ensure the learning is safe. Later, the estimated external agents' states

The authors are with the Department of Electrical and Biomedical Engineering, University of Nevada, Reno, NV, 89557 USA. E-mail: sdey@unr.edu; haoxu@unr.edu. This work was supported by the National Science Foundation under Grant 2144646.

will converge to actual states along with the convergence of learning. Moreover, the approximated safe set can gradually converge toward the actual safe set. The contributions are:

- A decentralized safe learning-based control is developed for local agents navigating in an unknown multi-agent shared environment. The developed control design incorporates both the local agent and observer-based external agent estimated states. The developed method ensures robust safety using AI-CBF while learning the external system.
- A novel multi-NN-based adaptive observer is designed to estimate the state of the unknown external agents.

II. Problem Formulation and Background

Consider a nonlinear affine system of local agent $\mathcal{A}$ in multi-agent systems given by the following dynamics

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

where $x(t) \in \mathbb{R}^n$ is the state and $u(t) \in \mathbb{R}^m$ is the control input of the agent. Also, $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are the intrinsic dynamics of the system. It is assumed that $f(0) = 0$ and $f(x) + g(x)u$ is bounded by the Lipschitz constant. The objective of the agent $\mathcal{A}$ is to reach a predefined destination safely by avoiding collision in a multi-agent environment. The dynamic of an external agent i is:

$$\begin{aligned}\dot{z}_i(t) &= f_a(z_i(t), u_i)\\ y_i(t) &= Cz_i(t)\end{aligned} \tag{2}$$

where $z_i(t) \in \mathbb{R}^n$ is the state and $y_i(t) \in \mathbb{R}^p$ is the output and u_i is the control input of the external agent i. Also, f_a is an unknown nonlinear function that captures the effect of an external agent. It is assumed that the external multi-agent systems are observable. The dynamic of the external agent:

$$\begin{aligned}\dot{z}_i(t) &= Az_i(t) + F(z_i(t), u_i)\\ y_i(t) &= Cz_i(t)\end{aligned} \tag{3}$$

with, $F(z_i(t), u_i) = f_a(z_i(t), u_i) - Az_i(t)$ and A represents Hurwitz matrix. Also, x_d represents the desired destination that the agent is required to achieve. Then the error can be defined as $e = x - x_d$ with the error dynamic given as:

$$\dot{e} = f_a(e) + g_a(e)u \tag{4}$$

where $f_a(e) = f(e+x_d)$ and $g_a(e) = g(e+x_d)$. The agent's safety cannot be solely determined by its own control inputs and characteristics but also relies on the interplay with other external systems in an uncertain multi-agent environment. Thus, to maintain the safety of local agents, it is needed to consider both local agent's behavior as well as interaction with other agents. However, in real-world scenarios, the states of the external agents are not available while only the outputs of the external systems are measurable. Moreover, the dynamics of the external systems are also unknown. Therefore, the objective of this research is to ensure the safety and stability of the agent $\mathcal{A}$ in a shared multi-agent environment where the states of the external agents are not directly measurable and the dynamics of these agents are unknown. The objectives are:

1) Design a feedback controller for the agent $\mathcal{A}$ in a decentralized manner, which guarantees the trajectory of the agent stays inside a safe set in the multi-agent environment while achieving a desired destination x_d that satisfies

$$h_i(x(t), z_i(t)) \geq 0 \quad \text{for} \quad t \geq 0 \tag{5}$$

with $h_i(x(t), z_i(t))$ being a continuously differentiable function. The safe set for the agent $\mathcal{A}$ is defined as the intersection of the sets associated with all other external agents, i.e.,

$$S = S_1 \cap S_2 \ ... \ \cap S_N \tag{6}$$

where N is the total number of external agents in the environment.

2) Design a multiple NN-based adaptive observer to estimate the state of all the other external agents to guarantee the safety of local agent $\mathcal{A}$.

3) Guaranteeing the stability of local agent $\mathcal{A}$ and the observer-based external agent's state estimation.

Before proceeding, a concise introduction to the control barrier function (CBF) is provided.

Control Barrier Function (CBF): The safety framework [8] is characterized by the invariance of a set, known as the actual safe set. This set, denoted as $\mathcal{S}$, is defined as a super level set of a smooth function $h : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$. Then, for a given dynamical system $\dot{x} = f(x) + g(x)u$, the safe set can be defined as:

$$\begin{aligned}\mathcal{S} &= \{x \in \mathcal{X} \subset \mathbb{R}^n : h(x) \geq 0\},\\ \partial\mathcal{S} &= \{x \in \mathcal{X} \subset \mathbb{R}^n : h(x) = 0\},\\ \text{Int}(\mathcal{S}) &= \{x \in \mathcal{X} \subset \mathbb{R}^n : h(x) > 0\}.\end{aligned}$$

Here, $\partial\mathcal{S}$ represents the boundary and $\text{Int}(\mathcal{S})$ denotes the interior of the set $\mathcal{S}$. Now, h can be referred to as a control barrier function if there exists an extended class $\mathcal{K}_\alpha$ function α such that the given dynamical system satisfies:

$$\sup_{u\in U}[L_f h(x) + L_g h(x)u] \geq -\alpha(h(x)) \tag{7}$$

for all $x \in \mathcal{X}$. where $L_f = \frac{\partial h}{\partial x} f(x)$ and $L_g = \frac{\partial h}{\partial x} g(x)$ are Lie derivatives of $h(x)$ along f and g, respectively. Then, the extended $\mathcal{K}_\alpha$ function can be defined as follows:

Definition 1: A function [8] $\alpha : \mathbb{R} \to \mathbb{R}$ is a extended class $\mathcal{K}_\alpha$ function that is strictly increasing and $\alpha(0) = 0$.

Then, the set of control inputs that satisfy (7) and render $\mathcal{S}$ safe:

$$K_{\text{cbf}} = \{u \in U : L_f h(x) + L_g h(x)u + \alpha(h(x)) \geq 0\} \tag{8}$$

III. Neural Network Based Adaptive Observer with Adaptive Interplay Control Barrier Function Design

In this section, an adaptation has been introduced to the CBF within the context of a shared environment comprising multiple agents. These agents possess unknown dynamics, and their states are not accessible for direct online measurement. The barrier function is now formulated as a function of both x and z_i, where i is the index for the external agent. Specifically, an NN-based adaptive observer has been developed to estimate the states of the external agents first. Then, employing the estimated states of external agents from the observer, the CBF is reformulated to AI-CBF using both the state x of agent $\mathcal{A}$ and estimated state $\hat{z}_i$ of any external agent i. In this subsection, a multiple NN-based observer is designed to estimate the states of external agents. Here, NN are used to identify the unknown dynamics of the external agents, and observers are used along with NNs to estimate the states of the external agents. The observer model can be described as:

$$\begin{aligned}\dot{\hat{z}}_i(t) &= A\hat{z}_i(t) + F(\hat{z}_i(t), u_i) + H(y_i(t) - C\hat{z}_i(t)) \\ \hat{y}_i(t) &= C\hat{z}_i(t)\end{aligned} \tag{9}$$

where $\hat{z}_i$ and $\hat{y}_i$ represent the state and output of the observer for external agent i. The selection of the observer gain $H \in \mathbb{R}^{n\times p}$ ensures $A - HC$ is a Hurwitz matrix. Please note that the observability of pair (C, A) depends on the selection of the matrix A. By selecting A properly, the existence of the gain H is ensured. According to universal approximation theory [15], the unknown function of the external agent is represented as:

$$F(z_i(t), u_i) = W_i^T\phi(z_i) + \varepsilon_f \tag{10}$$

with $W \in \mathbb{R}^{l\times n}$ is the ideal weight of the NN and l is hidden layer neuron numbers. The activation function is bounded by $\|\phi(z_i)\| \leq \phi_M$ and the ideal weight is bounded as $\|W\| \leq W_M$. Now, the unknown function is approximated as

$$\hat{F}(\hat{z}_i(t), u_i) = \hat{W}_i^T\hat{\phi}(\hat{z}_i) \tag{11}$$

where $\hat{W} \in \mathbb{R}^{l\times n}$ is the estimated weight. Using the unknown function estimation, the observer model can be represented as

$$\begin{aligned}\dot{\hat{z}}_i(t) &= A\hat{z}_i(t) + \hat{W}_i^T\hat{\phi}(\hat{z}_i) + H(y_i(t) - C\hat{z}_i(t)) \\ \hat{y}_i(t) &= C\hat{z}_i(t)\end{aligned} \tag{12}$$

Then, the state and output estimation error of the observer is defined as $\tilde{z}_i = z_i - \hat{z}_i$ and $\tilde{y}_i = y_i - \hat{y}_i$. Next, the state and output error dynamics is evaluated using Eq. (12) and (10) as

$$\begin{aligned}\dot{\tilde{z}}_i(t) &= \dot{z}_i(t) - \dot{\hat{z}}_i(t) \\ &= (A - HC)\tilde{z}_i(t) + W_i^T\phi(z_i) - \hat{W}_i^T\hat{\phi}(\hat{z}_i) + \varepsilon_f\end{aligned} \tag{13}$$

And, $\tilde{y}_i(t) = y_i(t) - \hat{y}_i(t) = C\tilde{z}_i(t)$. Then, the weight estimation error is defined as $\tilde{W}_i = W_i - \hat{W}_i$ and the activation function approximation error is $\tilde{\phi}(\tilde{z}_i) = \phi(z_i) - \hat{\phi}(\hat{z}_i)$. Also, $A - HC = A_o$. The Eq. (13) can be rewritten as

$$\begin{aligned}\dot{\tilde{z}}_i(t) &= A_o\tilde{z}_i(t) + W^T\phi(z_i) - W^T\hat{\phi}(\hat{z}_i) + W^T\hat{\phi}(\hat{z}_i) \\ &\quad - \hat{W}^T\hat{\phi}(\hat{z}_i) + \varepsilon_f \\ &= A_o\tilde{z}_i(t) + W^T\tilde{\phi}(\tilde{z}_i) + \tilde{W}^T\hat{\phi}(\hat{z}_i) + \varepsilon_f\end{aligned} \tag{14}$$

Assumption 1: The external system activation function approximation error is Lipschitz continuous implies the existence of the Lipschitz function L_ϕ, that satisfies $\|\tilde{\phi}(\tilde{z}_i)\| \leq L_\phi\|\tilde{z}_i\|$.

In this section, a modified weight update law has been designed to guarantee the stability of the NN-based observer. Now. taking the first derivative of the approximated output of the external agent:

$$\begin{aligned}&\dot{\tilde{y}}_i(t) = C\dot{\tilde{z}}_i(t) \\ &= CA_o(C^TC)^{-1}C^T\tilde{y}_i + CW^T\tilde{\phi}(\tilde{z}_i) + C\tilde{W}^T\hat{\phi}(\hat{z}_i) + C\varepsilon_f\end{aligned} \tag{15}$$

where $(C^TC)^{-1}C^T = C^+$ is the pseudo-inverse of the matrix C. The Eq. (15) can be rewritten as

$$\dot{\tilde{y}}_i(t) - CA_oC^+\tilde{y}_i = CW^T\tilde{\phi}(\tilde{z}_i) + C\tilde{W}^T\hat{\phi}(\hat{z}_i) + C\varepsilon_f \tag{16}$$

Now, the objective function can be defined as

$$J_i = \frac{1}{2}(\dot{\tilde{y}}_i(t) - CA_oC^+\tilde{y}_i)^T(\dot{\tilde{y}}_i(t) - CA_oC^+\tilde{y}_i) \tag{17}$$

The gradient descent-based update law is defined as follows:

$$\begin{aligned}\dot{\hat{W}}_i &= -\alpha(\dot{\tilde{y}}_i(t) - CA_oC^+\tilde{y}_i)^T\frac{\partial(\dot{\tilde{y}}_i(t) - CA_oC^+\tilde{y}_i)}{\partial\hat{W}_i} \\ &= \alpha C\hat{\phi}(\hat{z}_i)\big[CW^T\tilde{\phi}(\tilde{z}_i) + C\tilde{W}^T\hat{\phi}(\hat{z}_i) + C\varepsilon_f\big]^T\end{aligned} \tag{18}$$

where α is the learning rate of the neural network. The weight approximation error dynamic can be defined as:

$$\dot{\tilde{W}}_i = -\alpha C\hat{\phi}(\hat{z}_i)\big[CW^T\tilde{\phi}(\tilde{z}_i) + C\tilde{W}^T\hat{\phi}(\hat{z}_i) + C\varepsilon_f\big]^T \tag{19}$$

A. AI-CBF Design for Observer-Based State Estimation of External Agents

The characterization of the safety framework necessitates the positive invariance of a safe set. While the local agent is operating within a shared environment, the safe set is defined as the intersection of all sets associated with the different external systems present in the environment. This approach ensures that the local agent remains within the boundaries of safety. However, as stated earlier, only the output of the external agents states are available and the dynamics as well as full state information of the external agents are unknown. Since there is no accurate state information available for the external agents, the actual safe set of agent $\mathcal{A}$ is not available. Therefore, local agent $\mathcal{A}$ depends on the estimated safe set for safe action on the environment. In this regard, the control barrier function is reformulated to incorporate the state of agent $\mathcal{A}$ and external agents. Besides that, a state approximation error bound is considered for the worst-case scenario to ensure strict safety even if the actual safe set is not available to the local agent. In the initial stages of the NN-based observer's training, the estimation error $\tilde{z}$ tends to be relatively large. Consequently, this leads to a substantial bound on the approximation error, resulting in a larger unsafe region for the local agent. However, while the NN-based observer is well-trained, the approximation error decreases significantly. Consequently, agents have more flexibility and a larger safe maneuvering space as the unsafe region diminishes in size. Now, let the output measurement error $\tilde{y}_i$ belong to a sector [16] that can be defined as

$$\gamma\|\tilde{z}_i\|^2 \leq \|\tilde{y}_i\|^2 \leq \beta\|\tilde{z}_i\|^2 \tag{20}$$

Here, $\tilde{z}_i$ is the state approximation error of the external agent i. And γ and β are real numbers that satisfies $\beta \geq \gamma$. Using this sector-bounded condition, the upper bound of the state approximation error can be defined as follows:

$$\|\tilde{z}_i\|^2 \leq \frac{1}{\gamma}\|\tilde{y}_i\|^2 \Longrightarrow \|\tilde{z}_i\| \leq \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\| \tag{21}$$

Now, for the given dynamical system of agent $\mathcal{A}$ in the Eq. (1) and external agent dynamic in (2), the approximated safe set associated with external agent i can be defined as

$$\hat{\mathcal{S}} = \{x \in \mathbb{R}^n : h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) \geq 0\},$$
$$\partial\hat{\mathcal{S}} = \{x \in \mathbb{R}^n : h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) = 0\},$$
$$\text{Int}(\hat{\mathcal{S}}) = \{x \in \mathbb{R}^n : h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) > 0\}.$$

The function $h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|)$ represents a smooth function that incorporates both the variables x and $\hat{z}_i$. Please note that $\hat{z}_i$ is the estimated state information of the external agent i which is available to the agent $\mathcal{A}$. Besides that, the bounded error is incorporated to ensure strict safety of the agent $\mathcal{A}$ for the observer-based external agent state estimation. Now, the safe set for the agent $\mathcal{A}$ is derived by taking the intersection of sets associated with all external agents in the shared environment. Then, the approximated safe set of the agent $\mathcal{A}$ is defined as:

$$\hat{\mathcal{S}}(x, \hat{z}) = \hat{\mathcal{S}}_1(x, \hat{z}_1) \cap \hat{\mathcal{S}}_2(x, \hat{z}_2) \ldots \cap \hat{\mathcal{S}}_N(x, \hat{z}_N) \tag{22}$$

where N is the total number of external systems. The function $h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|)$ is the adaptive interplay control barrier function (AI-CBF). If there exists an extended class $\mathcal{K}_\alpha$ function $\alpha : \mathbb{R} \to \mathbb{R}$ such that for given dynamical system in (1) and (2), the following conditions hold:

$$\sup_{u\in\mathcal{U}}[L_f h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) + L_g h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|)u +$$
$$L_F h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|)] \geq -\alpha h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) \tag{23}$$

where $L_f = \frac{\partial h_i}{\partial x} f(x)$, $L_g = \frac{\partial h_i}{\partial x} g(x)$ and $L_F = \frac{\partial h_i}{\partial \hat{z}_i} F(x)$ are Lie derivatives of $h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}})$ along f, g and F, respectively. Now, the set of control input that satisfies (23) can be defined as:

$$K_{\text{cbf}} = \{u \in \mathcal{U} : L_f h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) + L_g h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}$$
$$\|\tilde{y}_i\|)u + L_F h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) + \alpha h_i(x, \hat{z}_i \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) \geq 0\} \tag{24}$$

The safe set is defined here for both agent $\mathcal{A}$ and the external agents' systems. The establishment of the safe set is intrinsically tied to the external system, wherein the learning phase can only yield an approximate safe set. To ensure safety for the agent $\mathcal{A}$, the AI-CBF is formulated as a function of both the local agent and external agents and incorporates the worst-case scenario for the observer-based state estimation of the external agents. The formulated AI-CBF ensures the forward invariance of the approximated safe set $\hat{\mathcal{S}}$.

IV. Control Framework

A control input needs to be designed to guarantee safety and maintain the stability of the local system in a multi-agent environment. This requirement highlights the importance of integrating a Lyapunov function $\mathcal{V}_e(x, x_d)$. The Lyapunov function derivative constraint and AI-CBF constraint are unified to achieve robust safety and stability performance. Then, a nominal controller $\bar{u}$ for local agent $\mathcal{A}$ is given to guide it to the desired destination. Next, a quadratic programming (QP) [17] based method has been adopted. Building upon prior research efforts [8], [10], this QP-based controller unifies stability and safety constraints within an optimization framework. By leveraging quadratic programming, the controller facilitates continuous updates of the control actions.

$$u(x, \hat{z}_i) = \underset{(u,\delta)}{\arg\min} \quad \frac{1}{2}\|u - \bar{u}\|^2 + p\delta^2$$
$$\text{s.t.} \;\; L_f h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) + L_g h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|)u +$$
$$L_F h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) + \alpha h_i(x, \hat{z}_i, \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|) \geq 0$$
$$\dot{\mathcal{V}}_e(x, x_d) \leq \delta \tag{25}$$

where δ serves as a relaxation variable to guarantee quadratic program solvability and p represents a coefficient.

Lemma 1: There exists a control policy u for the dynamic equation given in (4) to guarantee the systems stability.

$$e^T\Big\{f_a(e(t)) + g_a(e(t))u\Big\} \leq -\gamma\|e\|^2 \tag{26}$$

Theorem 1: The NN weight is updated by the equation (18) and the learning rate α is a positive constant. Then, NN weight approximation error $\tilde{W}$, external agent state estimation error $\tilde{z}_i$, and the local agent $\mathcal{A}$ regulation error e are all ultimately uniformly bounded (UUB). Moreover, $\tilde{W}$, $\tilde{z}_i$, and e are asymptotically stable when the reconstruction error [18] and relaxation variable δ is zero.

Proof: Consider the Lyapunov candidate function

$$\mathcal{V}_{sys} = \mathcal{V}_e + \mathcal{V}_s \tag{27}$$

with $\mathcal{V}_e = \frac{1}{2}\text{tr}\{e^T(t)e(t)\}$ and $\mathcal{V}_s = \frac{1}{2}\tilde{z}_i^T P\tilde{z}_i + \frac{1}{2}\text{tr}\{\tilde{W}_i^T\tilde{W}_i\}$. Also, settling the relaxation variable δ to be equal to zero, the Eq. (27) can be written as:

$$\mathcal{V}_{sys} = \frac{1}{2}\text{tr}\{e^T(t)e(t)\} + \frac{1}{2}\tilde{z}_i^T P\tilde{z}_i + \frac{1}{2}\text{tr}\{\tilde{W}_i^T\tilde{W}_i\} \tag{28}$$

Taking the first derivative of Eq (27)

$$\dot{\mathcal{V}}_{sys} = \text{tr}\{e^T(t)\dot{e}(t)\} + \frac{1}{2}\dot{\tilde{z}}_i^T P\tilde{z}_i + \frac{1}{2}\tilde{z}_i^T P\dot{\tilde{z}}_i + \text{tr}\{\tilde{W}_i^T\dot{\tilde{W}}_i\} \tag{29}$$

Substituting *Lemma 1* and Eq. (29) is as:

$$\dot{\mathcal{V}}_{sys} \leq -\frac{1}{2}\gamma\|e\|^2 + \frac{2}{\gamma}g_l^2 L_u^2\|\tilde{z}_i\|^2 - [\frac{1}{2}\lambda_{min}(Q) - \|P\| \\ W_M L_\phi - \frac{1}{2\alpha^2}\|P\|^2 - \frac{1}{2}\|p\|^2 - \frac{1}{2}W_M^2 L_\phi^2]\|\tilde{z}_i\|^2 - [\alpha \\ \|C\|^2\|\phi_M\|^2 - \alpha^2\|C\|^4\phi_M^2 - \frac{1}{2}\alpha^2\phi_M^2]\|\tilde{W}_i\|^2 + \|\varepsilon_f\|^2 \tag{30}$$

with g_l being the Lipschitz constant of the function g_a. Also, there exists a Lipschitz constant L_u that satisfy the inequality $\|\tilde{u}(x, \tilde{z}_i)\| \leq L_u\|\tilde{z}_i\|$. Now, the Eq. (30) can be rewritten as:

$$\dot{\mathcal{V}}_{sys} \leq -\frac{1}{2}\gamma\|e\|^2 - \kappa_{i,zc}\|\tilde{z}_i\|^2 - \kappa_{i,Wc}\|\tilde{W}_i\|^2 + \|\varepsilon_f\|^2 \tag{31}$$

with,

$$\kappa_{i,zc} = \frac{1}{2}\lambda_{min}(Q) - \frac{2}{\gamma}g_l^2 L_u^2 - \|P\|W_M L_\phi - \frac{1}{2\alpha^2}\|P\|^2 \\ - \frac{1}{2}\|p\|^2 - \frac{1}{2}W_M^2 L_\phi^2$$

$$\kappa_{i,Wc} = \alpha\|C\|^2\|\phi_M\|^2 - \alpha^2\|C\|^4\phi_M^2 - \frac{1}{2}\alpha^2\phi_M^2\|\varepsilon_f\|^2$$

Now the first derivative of the Lyapunov function $\dot{\mathcal{V}}_{sys}$ is less than zero outside a compact set if

$$\|e\| > \sqrt{\frac{2}{\gamma}\|\varepsilon_f\|^2} \;\; ; \;\; \|\tilde{z}\| > \sqrt{\frac{1}{\kappa_{i,zc}}\|\varepsilon_f\|^2}$$

$$\|\tilde{W}_i\| > \sqrt{\frac{1}{\kappa_{i,Wc}}\|\varepsilon_f\|^2}$$

This completes the proof.

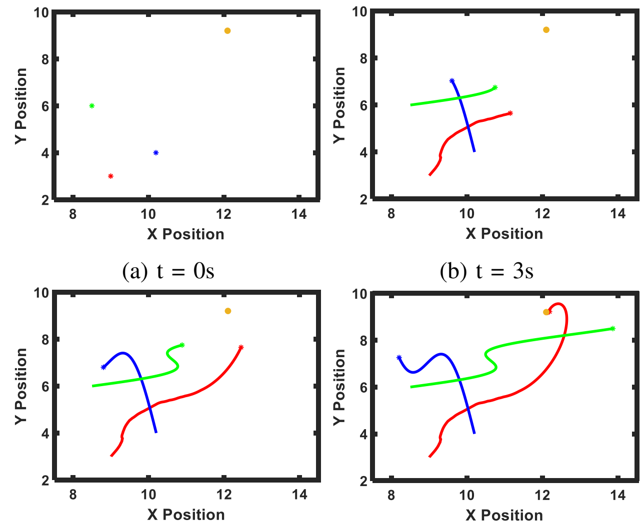
V. Simulation Result

In this section, we implement the developed algorithm into a multi-agent system to illustrate the secure maneuvering of a UAV within a shared airspace environment, alongside other external UAVs. There are a total of two external UAVs in the system. Let the initial state of the local UAV be selected as $x = [9 \quad 3 \quad 0 \quad 0]^T$ with its position and velocity. The predefined destination point of the UAV is $x = [12.2 \quad 9 \quad 0 \quad 0]^T$. Moreover, the initial state of the UAV-1 and UAV-2 is selected as $z_1 = [10.2 \;\; 4 \;\; 0 \;\; 0]^T$ and $z_2 = [8.5 \;\; 6 \;\; 0 \;\; 0]^T$. The intrinsic dynamic function of the local UAV is defined as:

$$f(x) = \begin{bmatrix} -x_1 + \frac{1}{2}x_2^2 \\ -0.4x_2^2 \\ x_2[cos(2x_1+1)^2 - 1] - x_1 \\ x_4[cos(2x_3+1)^2 - 1] - x_3 \end{bmatrix}$$

$$g(x) = \begin{bmatrix} 0 & 0 & cos(2x_1+1) & cos(2x_3+1) \end{bmatrix}^T$$

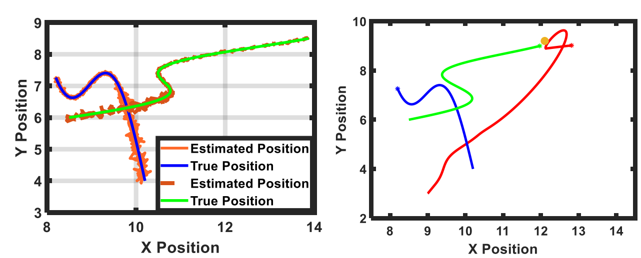
Moreover, the dynamic of the UAV-1 and UAV-2 is chosen as [19]. For the design of an NN-based observer, it is essential that the square matrix A be a Hurwitz matrix. The selection of matrices A ensures that both A and $A - HC$ have eigenvalues with strictly negative real parts, making them Hurwitz matrices. The activation function of each NN is selected as a hyperbolic tangent function, i.e. tanh(.). Also, the learning rate α of the NN is selected as 1×10^{-4}. Next, the AI-CBF is defined as $h_i(x, z_i) = \|x - z_i\|_Q^2 - \frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\| - r_{\min}$. Please note that the second term $\frac{1}{\sqrt{\gamma}}\|\tilde{y}_i\|$ is used to incorporate the estimation error bound of the external agents into the CBF design with $\gamma = 1.5$. Here, $r_{\min} = 0.2$ represents the radius of a closed circle. The relaxation factor coefficient is selected as $p = 0.5$. The differential equations are solved using MATLAB *ode45* and QP is solved using MATLAB *quadprog* function. In Figure 1, we illustrate the safe maneuvering of the local agent in a shared environment. We have plotted the trajectories of the UAVs in Figure 1 at various time points. It's important to note that the points at which the red curve (local UAV) intersects with the blue and green curves (external UAVs) do not represent simultaneous collisions but occur at different times. In Figure 2(a), we present the state estimation of external agents using the developed neural network (NN) adaptive observer. The true positions of UAV-



(a) t = 0s (b) t = 3s

(c) t = 7s (d) t = 10s

Fig. 1: The trajectory of all UAVs in the environment. The red curve represents local UAV. Also, blue and green curves are the trajectories of external UAVs in the environment.



(a) State estimation of external UAVs (b) Same final destination for local UAV and UAV-2.

Fig. 2: This figure illustrates both the estimated and actual positions of external agents, as well as a scenario demonstrating how the local UAV ensures strict safety.

1 and UAV-2 are represented by the blue and green curves, while the estimated positions are shown in orange and brown. Figure 2(a) effectively illustrates that as time progresses and the neural network observer is well-trained, the state estimation error decreases and approaches zero. In Figure 2(b), a scenario is presented to illustrate how the local UAV ensures strict safety. To prioritize safety, the local UAV places a strong emphasis on collision avoidance with UAV-2, even at the expense of reaching its desired destination. Figure 2(b) visually depicts the local UAV altering its course at the last moment to maintain a safe distance from UAV-2.

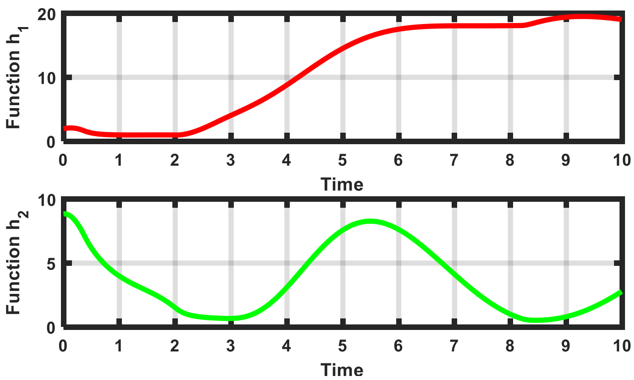


Fig. 3: AI-CBF for UAV-1 and UAV-2.

The adaptive interplay control barrier function (AI-CBF) for UAV-1 and UAV-2 is shown in Figure 3. In this figure, both barrier functions remain positive indicating that the local UAV always ensures safety. Initially, after the deployment of UAVs, the local UAV approaches UAV-1, causing its states to approach the boundary of the safe set. At that period of time, the value of the function is close to zero. Subsequently, the AI-CBF h_1 increases as the primary UAV moves away from UAV-1. Similarly, the AI-CBF h_2 experiences two decreases at different time points when UAV-2 comes closer to the local UAV on two occasions.

VI. CONCLUSION

This paper has developed a novel approach for safe control for local agents in a challenging multi-agent environment where the dynamics of external agents are both uncertain and uncontrollable, and accurate state information is unavailable. The developed method utilizes multiple neural network-based adaptive observers to estimate the states of these external agents. Through the integration of state information from both the external and local agents, an adaptive interplay control barrier function (AI-CBF) has been designed to ensure the local agent's safety. Notably, the AI-CBF guarantees the strict safety of the local agent by maintaining the forward invariance of an approximated safe set. Importantly, this algorithm has been proven to ensure system safety without the need for precise knowledge of the actual safe set. This AI-CBF along with the Lyapunov function is used for safe control development, which guides the local agent to a predefined destination point while guaranteeing safety and stability. The stability of the neural network-based observer design and the overall system stability have been demonstrated through Lyapunov stability analyses. Finally, numerical simulations have been provided to demonstrate the effectiveness of the developed algorithm.

REFERENCES

[1] S. Zhou, K. Guo, X. Yu, L. Guo, and L. Xie, "Fixed-time observer based safety control for a quadrotor uav," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 2815–2825, 2021.

[2] S. Dey and H. Xu, "Distributed adaptive flocking control for large-scale multiagent systems," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[3] P. Koopman and M. Wagner, "Autonomous vehicle safety: An interdisciplinary challenge," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 90–96, 2017.

[4] L. Lamport, "Proving the correctness of multiprocess programs," *IEEE transactions on software engineering*, no. 2, pp. 125–143, 1977.

[5] S. Dey and H. Xu, "Decentralized adaptive tracking control for large-scale multi-agent systems under unstructured environment," in *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2022, pp. 900–907.

[6] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 411–444, 2022.

[7] B. Thananjeyan, A. Balakrishna, S. Nair, M. Luo, K. Srinivasan, M. Hwang, J. E. Gonzalez, J. Ibarz, C. Finn, and K. Goldberg, "Recovery rl: Safe reinforcement learning with learned recovery zones," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4915–4922, 2021.

[8] 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)*. IEEE, 2019, pp. 3420–3431.

[9] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-jacobi reachability: A brief overview and recent advances," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 2242–2253.

[10] Z. Marvi and B. Kiumarsi, "Barrier-certified learning-enabled safe control design for systems operating in uncertain environments," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 3, pp. 437–449, 2021.

[11] R. Cheng, M. J. Khojasteh, A. D. Ames, and J. W. Burdick, "Safe multi-agent interaction through robust control barrier functions with learned uncertainties," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 777–783.

[12] S. He, J. Zeng, B. Zhang, and K. Sreenath, "Rule-based safety-critical control design using control barrier functions with application to autonomous lane change," in *2021 American Control Conference (ACC)*. IEEE, 2021, pp. 178–185.

[13] A. D. Ames, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs with application to adaptive cruise control," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 6271–6278.

[14] F. Abdollahi, H. A. Talebi, and R. V. Patel, "A stable neural network-based observer with application to flexible-joint manipulators," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 118–129, 2006.

[15] F. Scarselli and A. C. Tsoi, "Universal approximation using feedforward neural networks: A survey of some existing methods, and some new results," *Neural networks*, vol. 11, no. 1, pp. 15–37, 1998.

[16] H. K. Khalil, "Nonlinear control," *(No Title)*, 2015.

[17] J. Nocedal and S. J. Wright, "Quadratic programming," *Numerical optimization*, pp. 448–492, 2006.

[18] S. Dey and H. Xu, "Hierarchical game theoretical distributed adaptive control for large scale multi-group multi-agent system," *IET Control Theory & Applications*, vol. 17, no. 17, pp. 2332–2352, 2023.

[19] Z. Zhou and H. Xu, "A novel mean-field-game-type optimal control for very large-scale multiagent systems," *IEEE Transactions on Cybernetics*, vol. 52, no. 6, pp. 5197–5208, 2022.