Article

Critic Learning-Based Safe Optimal Control for Nonlinear Systems with Asymmetric Input Constraints and Unmatched Disturbances

Chunbin Qin [1], Kaijun Jiang [1], Jishi Zhang [2,*] and Tianzeng Zhu [1]

1   School of Artificial Intelligence, Henan University, Zhengzhou 450000, China; qcb@henu.edu.cn (C.Q.);
    1824100055@vip.henu.edu.cn (K.J.); kz520@henu.edu.cn (T.Z.)
2   School of Software, Henan University, Kaifeng 475000, China
*   Correspondence: 10250051@vip.henu.edu.cn

Abstract: In this paper, the safe optimal control method for continuous-time (CT) nonlinear safety-critical systems with asymmetric input constraints and unmatched disturbances based on the adaptive dynamic programming (ADP) is investigated. Initially, a new non-quadratic form function is implemented to effectively handle the asymmetric input constraints. Subsequently, the safe optimal control problem is transformed into a two-player zero-sum game (ZSG) problem to suppress the influence of unmatched disturbances, and a new Hamilton–Jacobi–Isaacs (HJI) equation is introduced by integrating the control barrier function (CBF) with the cost function to penalize unsafe behavior. Moreover, a damping factor is embedded in the CBF to balance safety and optimality. To obtain a safe optimal controller, only one critic neural network (CNN) is utilized to tackle the complex HJI equation, leading to a decreased computational load in contrast to the utilization of the conventional actor–critic network. Then, the system state and the parameters of the CNN are uniformly ultimately bounded (UUB) through the application of the Lyapunov stability method. Lastly, two examples are presented to confirm the efficacy of the presented approach.

Keywords: critic neural network; asymmetric input constraints; unmatched disturbances; safety; adaptive dynamic programming; nonlinear systems

1. Introduction

Safety-critical systems are those that, in case of accidents or failures, can result in significant consequences, including but not limited to injuries, loss of life, environmental harm, or financial losses. The emergence of safety-critical systems like unmanned aerial vehicles (UAVs) [1–3] and robots [4] has led to an increased focus on safety control design within the field of control systems [5,6]. Safety control designs entail control strategies that satisfy safety specifications imposed by environmental limitations or physical limitations of the system. Ignoring the detrimental impact of safety entails substantial risks to both the safety of belongings and personal security. To address the challenges of the safe controller design, researchers have provided some effective approaches [7–11]. The problem of safety in the presence of unmodeled dynamics or disturbances in drones has recently been addressed by designing the robust controller based on the nonlinear estimator in [9]. In ref. [10], the use of neural networks integrated with the Lyapunov theory was preliminarily treated with application in the automotive sector for critical situations, and this aspect was further addressed in an even more organic way. In ref. [11], the quadratic programming-based method was applied to develop a safe controller. Despite the fact that this method can guarantee safety at a local level for every time step, selecting a step size that is too small leads to redundant computations. In contrast, a step size that is too large causes unsafe behavior, making it challenging to ensure the safety of the system. Hence, it is crucial

to identify an appropriate control design method for CT safety-critical systems that can guarantee the safety of the systems.

Recently, the CBF technique has emerged as an effective approach for ensuring the security of safety-critical systems [12–14]. The underlying principle of the CBF is to insure the forward invariance of the safe set. In ref. [15], the safe-based reinforcement learning approach was demonstrated, where the CBF was merged into the cost function to assure both the safety and optimality of the system. Typically, the CBF component is contained within the primitive cost function to penalize behavior that violates safety constraints. Reference [16] incorporated damping factors into the CBF and intervened selectively only in the event of safety constraint violations, aiming to reduce disruptions to the optimal controller. Reference [17] introduced the utilization of the CBF and summarized the verification approach for safety-critical control systems. Nevertheless, the methods mentioned above do not take into account the presence of external disturbances, which served as a motivation for the research conducted in this paper.

As disturbances are present in almost all industrial systems and hurt control performance, it is necessary to consider external disturbances in actual projects. Recently, several methods have been proposed to address disturbances [18–24]. For instance, references [18,19] used $H\infty$ control to reduce external disturbances in nonlinear systems. Reference [23] combined ADP with sliding mode control for addressing optimal control problems of CT nonlinear systems considering uncertain disturbances. Reference [24] cast external disturbances as the ZSG problem, with the control strategy aiming at minimizing the cost function and the disturbance strategy striving towards maximizing it. It is well known that for the HJI equation of the ZSG, it is difficult to find its analytical solution. Fortunately, with the evolution of optimal control [25–27], the ADP approach [28] was employed to approximately tackle the ZSG problem. For example, in reference [29], a new database-based adaptive critic algorithm was presented to study the infinite-scale robust control for nonlinear systems. However, the aforementioned methods fail to consider the capability limitations of the system due to asymmetric input constraints.

Although symmetric input constraints have been widely investigated in prior research and tackled with various techniques, such as the control problem that concerns the uncertain impulse system that has input constraints, which was handled in [30], and the utilization of integral reinforcement learning with the actor–critic network to address the tracking control problem under input constraints in [31,32], there has been relatively little research on the treatment of asymmetric input constraints that frequently occur in practical systems. Several optimal control methods exist for addressing CT systems with nonlinear dynamics and input constraints that are asymmetric. Among them is one that used the cost function with adjustable upper and lower limits of integration [33–35]. Another proposed the switching function [36] to tackle the problem, but it is only applied to linear systems. However, none of these results considered the incorporation of CBF into the CT nonlinear safety-critical systems to study the safe and optimal control problem under asymmetric input constraints with external disturbances.

This study explores the safe and optimal control issue for safety-critical nonlinear systems to reject unmatched external disturbances under the condition of asymmetric input constraints. Unlike other works, a new non-quadratic form function for handling asymmetric input constraints is proposed in this paper. To tackle the challenge posed by unmatched disturbances, a two-player ZSG is put forward to formulate the optimization problem. The ZSG is then addressed by finding the Nash equilibrium point, which is obtained by addressing the HJI equation. However, since solving the HJI equation is challenging, an ADP technique similar to that used in references [37–39] is exploited to estimate the solution of the HJI equation. In addition, one single CNN is used instead of a dual actor–critic neural network to diminish the computational complexity in approximating the control policy. Consequently, the optimal control policy is obtained by considering the worst disturbance.

The contributions are outlined mainly as follows:

1.   Asymmetric input constraints are considered in the control problem of the CT nonlinear safety-critical systems. In addition, this paper proposes a new non-quadratic form function to address the issue of asymmetric input constraints. It is important to note that when applying this approach, the optimal control policy no longer remains at 0, even when the system state reaches the equilibrium point of $x = 0$ (see $u^*(x)$ in later Equation (15)).

2.   This paper adopts the CBF to construct safety constraints and proposes designing a damping coefficient within the CBF to balance the safety and optimality of safety-critical systems based on varying safety requirements in different applications.

3.   The safe optimal control problem is turned into the ZSG problem to address unmatched disturbances; then, the optimal control law is gained by tackling the HJI equation using one CNN. Moreover, the use of only one CNN to approximate the HJI equation is an effective way to reduce the computational burden compared to the actor–critic network and the system state, and CNN parameters are demonstrated to be UUB.

The following structure is adopted for this article. Section 2 provides the initial formulation of the problem. Section 3 presents a safe optimal control design for the two-player ZSG problem. Then, in Section 4, an adaptive CNN method for addressing the HJI equation using an online method is proposed, and its stability is verified. Section 5 introduces two examples to demonstrate that the presented approach is effective. Lastly, Section 6 gives conclusions.

2. Problem Statement

Consider the CT nonlinear safety-critical system as

$$\dot{x} = \mathcal{F}(x) + \mathcal{G}(x)u + \mathcal{P}(x)v, \tag{1}$$

where $x = [x_1, x_2, \ldots, x_n]^T \in \mathcal{C}_a \subseteq \mathcal{R}^n$ indicates the system state vector with n-dimensional parameters, $\mathcal{F}(x) \in \mathcal{R}^n$ represents the internal dynamics, $\mathcal{G}(x) \in \mathcal{R}^{n \times m}$ and $\mathcal{P}(x) \in \mathcal{R}^{n \times q}$ indicate control and disturbance coefficient matrices, respectively. Additionally, $u \in \mathcal{R}^m$ denotes an input variable with m-dimensional parameters denoted by $\complement_u = \{u | u_{max} \geq u \geq u_{min}\}$, where u_{max} and u_{min} stand for the upper and lower bounds, respectively. And $v \in \mathcal{R}^q$ is the unmatched disturbances. The paper assumes $\mathcal{F}(\cdot), \mathcal{G}(\cdot), \mathcal{P}(\cdot)$ are Lipschitz continuous and satisfy $\mathcal{F}(0) = 0$, and the safety-critical System (1) is stabilizable and controllable. Moreover, we assume there exist two constants $\mathcal{G}_M > 0$ and $\mathcal{P}_M > 0$. Both $\mathcal{G}(x)$ and $\mathcal{P}(x)$ have upper bounded values, i.e., $\mathcal{G}_M \geq \|\mathcal{G}(x)\|, \mathcal{P}_M \geq \|\mathcal{P}(x)\|$, for any $x \in \mathcal{R}^n$.

In addition, it is essential to emphasize that $\mathcal{C}_a$ represents a safe set for (1). $\mathcal{C}_a$ is derived from operational restrictions, such as the allowable states of the robot arm, which is mathematically determined by

$$\begin{cases} \mathcal{C}_a = \{x \in \mathcal{R}^n | z(x) \geq 0\}, \\ int(\mathcal{C}_a) = \{x \in \mathcal{R}^n | z(x) > 0\}, \\ \partial\mathcal{C}_a = \{x \in \mathcal{R}^n | z(x) = 0\}, \end{cases} \tag{2}$$

where $z(x)$ represents continuous concerning x. The set $int(\mathcal{C}_a)$ denotes the interior of $\mathcal{C}_a$, while $\partial\mathcal{C}_a$ represents the boundary of $\mathcal{C}_a$.

Subsequently, the representation of the infinite horizon cost function from $t = 0$ for System (1) is given by

$$V(x) = \int_0^\infty x(t)^T \mathcal{Q}x(t) + \mathfrak{U}(u) - \Upsilon^2 \|v\|^2 dt, \tag{3}$$

where $\mathcal{Q}$ represents a function with positive definite properties, $\|v\|^2 = v^T v$, $Y > 0$ represents a constant weight coefficient, $\mathfrak{U}(u)$ is a non-quadratic form function employed for handling the asymmetric input constraints determined by

$$
\begin{aligned}
\mathfrak{U}(u) &= 2 \int_{\mathfrak{I}}^{u} \Psi \tanh^{-1}\left(\frac{t - \mathfrak{I}}{\Psi}\right) dt \\
&= 2\Psi(u - \mathfrak{I})\tanh^{-1}\left(\frac{u - \mathfrak{I}}{\Psi}\right) + \Psi^2 \ln\left(1 - \frac{(u - \mathfrak{I})^2}{\Psi}\right),
\end{aligned}
\tag{4}
$$

with Ψ and $\mathfrak{I}$ defined as

$$
\Psi = \frac{1}{2}(u_{max} - u_{min}), \mathfrak{I} = \frac{1}{2}(u_{max} + u_{min}),
\tag{5}
$$

where $|u_{max}| \neq |u_{min}|$ and $\tanh(\mathfrak{z}) = (e^{\mathfrak{z}} - e^{-\mathfrak{z}})/(e^{\mathfrak{z}} + e^{-\mathfrak{z}})$ with $\mathfrak{z} \in \mathcal{R}$.

Remark 1. *Even though $\tanh(\mathfrak{z})$ is symmetric, $\mathfrak{U}(u)$ in (4) generates asymmetric constraints in the control signal $u^*(x)$ (see $u^*(x)$ in later (15)). This is due to the fact that $\mathfrak{I}$ is not equal to 0 in (4). This feature is different from studying the symmetric input constraints.*

Additionally, the ultimate objective of this paper is to devise the safe and optimal control input policy for (1), which involves the utilization of the CBF concept. In the upcoming section, this paper presents the concept of the CBF and proposes an ADP-based approach to design the safe and optimal controller.

3. Safe Optimal Control Design

This section presents a detailed explanation of the concept of the CBF. Then, the safe and optimal control problem is converted to the two-player ZSG to overcome the unmatched disturbances, and the CBF is integrated with the cost function without an intermediary to punish unsafe behavior.

3.1. Control Barrier Function

The utilization of the CBF provides a solution to address the safety constraint problem in safety-critical systems. The CBF is a function that is non-negative within the set $\mathcal{C}_a$ and exhibits divergence to infinity at the edge of $\mathcal{C}_a$. As the state x is about to reach the boundary of $\mathcal{C}_a$, the condition of negative derivative can bring the system state x back within $\mathcal{C}_a$, ensuring that the system state is always confined within $\mathcal{C}_a$. To better illustrate the properties of the CBF, the following assumption is given.

Assumption 1. *The CBF candidate $\mathcal{B}_r(x)$ meets the subsequent three characteristics [40,41]:*
(1) $\mathcal{B}_r(x) \geq 0, \forall x \in int(\mathcal{C}_a)$,
(2) $\mathcal{B}_r(x) \to \infty, \forall x \in \partial\mathcal{C}_a$,
(3) $\mathcal{B}_r(x)$ *is monotonically decreasing* $\forall x \in \mathcal{C}_a$.

Moreover, for all $x \in \mathcal{C}_a$, the CBF $\mathcal{B}_r(x)$ has the following properties:

$$
\frac{1}{\gamma_1(z(x))} \leq \mathcal{B}_r(x) \leq \frac{1}{\gamma_2(z(x))}, \dot{\mathcal{B}}_r(x) \leq \gamma_3(z(x)),
\tag{6}
$$

where $\gamma_1(\cdot), \gamma_2(\cdot)$, and $\gamma_3(\cdot)$ are class $\mathcal{K}$ functions.

Under the premise that Assumption 1 and Equation (6) both hold, a suitable choice for $\mathcal{B}_r(x)$ is $\rho y(x)/z(x)$, where $y(x)$ represents a special scheduling function determined by the user to allow for flexibility in selecting $\mathcal{B}_r(x)$. Specifically, $y(x)$ ensures that the CBF operates only when the system is close to the unsafe set. $\rho > 0$ is the damping factor used to balance safety and optimality.

Remark 2. *In contrast to the previous CBF [16], the ρ chosen here shows a positive correlation with the value of $\mathcal{B}_r(x)$. The larger the value of ρ, the faster the system state moves away from the unsafe set, and the smaller the value of ρ, the slower the state x moves away from the unsafe set. A smaller value of ρ emphasizes optimality and a larger value of ρ enforces safety.*

3.2. Safe and Optimal Control Approach

By augmenting the selected CBF $\mathcal{B}_r(x)$ to the cost function (3), a new refined cost function is obtained, that is,

$$V(x) = \int_0^\infty x(t)^T Qx(t) + \mathfrak{U}(u) - Y^2\|v\|^2 + \mathcal{B}_r(x)dt. \tag{7}$$

Remark 3. *To ensure the safety of the system, it is assumed that the original system state x is confined within the set $\mathcal{C}_a$. This is because the rapid increase in $\mathcal{B}_r(x)$ as the state x nears the boundary of $\mathcal{C}_a$ is the reason behind the penalization of state convergence behavior when the initial state is beyond $\mathcal{C}_a$. This prevents the system state from converging.*

The conventional control problems can be transformed into two-player ZSG problems. The Nash equilibrium point, i.e., the saddle point (u^*,v^*) can be obtained by addressing the special HJI equation. Then, the optimal cost function is defined by

$$V^*(x) = \min_u \max_v \int_0^\infty x(t)^T Qx(t) + \mathfrak{U}(u) - Y^2\|v\|^2 + \mathcal{B}_r(x)dt. \tag{8}$$

The purpose of the two-player ZSG problem is to identify a saddle point so that the following inequality can hold:

$$V^*(x, u^*, v) \le V^*(x, u^*, v^*) \le V^*(x, u, v^*). \tag{9}$$

Therefore, for the two-player ZSG problem, u^* is the optimal control input policy minimizing the cost function, and v^* represents the worst disturbance input policy maximizing the cost function.

Definition 1. *Input policy u is considered admissible in relation to (7) on $\mho \in \mathcal{R}^n$, denoted by $u \in \aleph(\mho)$, u stabilizes (1) on $\mho$ if u is continuous on $\mho$, and (7) is limited for any $x \in \mho$.*

For the admissible input policy $u \in \aleph(\mho)$, if Equation (7) is continuously differentiable, computing the gradient of $V(x)$ with respect to t on both sides of Equation (7) yields the nonlinear Lyapunov equation as

$$0 = \nabla V(x)^T(\mathcal{F}(x) + \mathcal{G}(x)u + \mathcal{P}(x)v) + x(t)^T Qx(t) + \mathfrak{U}(u) - Y^2\|v\|^2 + \mathcal{B}_r(x), \tag{10}$$

where $\nabla V(x)$ is the gradient of $V(x)$, $V(0) = 0$.

Based on the optimal control approach, the HJI equation for the two-player ZSG problem possesses an exclusive solution if there exists a saddle point, that is, if the following conditions hold:

$$0 = \min_u \max_v \mathcal{H}(x, u, v, \nabla V^*(x)) = \max_v \min_u \mathcal{H}(x, u, v, \nabla V^*(x)), \tag{11}$$

where $\mathcal{H}(x, u, v, \nabla V^*(x))$ refers to the Hamiltonian function of the safety-critical system (1), that is,

$$\begin{aligned}\mathcal{H}(x, u, v, \nabla V^*(x)) &= \nabla V^*(x)^T(\mathcal{F}(x) + \mathcal{G}(x)u + \mathcal{P}(x)v) \\ &\quad + x(t)^T Qx(t) + \mathfrak{U}(u) - Y^2\|v\|^2 + \mathcal{B}_r(x).\end{aligned} \tag{12}$$

By using Equations (11) and (12), the saddle point can be found by addressing two equations as

$$u^*(x) = arg \min_{u} \mathcal{H}(x, u, v, \nabla V^*(x)),$$ (13)

and

$$v^*(x) = arg \max_{v} \mathcal{H}(x, u, v, \nabla V^*(x)).$$ (14)

Thus, the saddle point (u^*, v^*) can be gained as

$$u^*(x) = -\Psi \, tanh(\frac{1}{2\Psi} \mathcal{G}(x)^T \nabla V^*(x)) + \psi_\Im,$$ (15)

and

$$v^*(x) = \frac{1}{2\Upsilon^2} \mathcal{P}(x)^T \nabla V^*(x),$$ (16)

where $\psi_\Im = [\Im, \Im, \dots, \Im]^T \in \mathcal{R}^m$ with $\Im$ given by Equation (5).

Remark 4. *Given that $\Im \neq 0$ from (5), it can be concluded that $u(0) = \Im \neq 0$. Therefore, in order to establish the equilibrium point of (1) at $x = 0$, the assumption of $\mathcal{G}(0) = 0$ is necessary.*

Substituting Equations (15) and (16) into Equation (11), the HJI equation can be redefined as

$$\begin{aligned} 0 = & \nabla V^*(x)^T (\mathcal{F}(x) + \mathcal{P}(x)v^*) + x(t)^T \mathcal{Q}x(t) + \mathfrak{U}(-\Psi \, tanh(T(x)) + \psi_\Im) \\ & - \Upsilon^2 \|v^*\|^2 - (\Psi \nabla V^*(x))^T \mathcal{G}(x) tanh(T(x)) + \nabla V^*(x)^T \mathcal{G}(x) \psi_\Im + \mathcal{B}_r(x), \end{aligned}$$ (17)

where $T(x) = 1/(2\Psi) \mathcal{G}(x)^T \nabla V^*(x)$ and $V^*(0) = 0$.

For the optimal safe control problem of the ZSG with unmatched external disturbances and asymmetric input constraints, it is necessary to obtain the value corresponding to the optimal cost Function (8) for achieving the optimal control input Policy (15) and the worst disturbance input Policy (16). Therefore, the solution of Equation (17) needs to be obtained. Nevertheless, since Equation (17) represents a nonlinear partial differential equation, it is challenging to find its analytical solution using conventional mathematical approaches. Hence, the solution of this equation is estimated by using the CNN in the next section.

4. Adaptive CNN Design

4.1. Solving the HJI Equation via the CNN

This section designs a CNN to estimate cost function $V^*(x)$ as

$$V^*(x) = W_c^T \delta(x) + \xi(x),$$ (18)

where $\xi(x)$ represents the estimation error about the CNN with $\xi(0) = 0$, $W_c \in \mathcal{R}^{\mathfrak{r}}$ represents the ideal weight vector of the CNN, $\delta(x) = [\delta_1(x); \delta_2(x); \dots; \delta_{\mathfrak{r}}(x)]$ represents activation function with $\delta_j(0) = 0, j = 1, 2, \dots, \mathfrak{r}$, $\mathfrak{r}$ is the number of neurons in the CNN.

The gradient of the approximate optimal cost function is

$$\nabla V^*(x) = \nabla \delta(x)^T W_c + \nabla \xi(x).$$ (19)

Substituting Equation (19) into Equation (15), $u^*(x)$ can be represented as

$$u^*(x) = -\Psi \, tanh(\bar{A}(x)) + \xi_{u^*}(x) + \psi_\Im,$$ (20)

where

$$\bar{A}(x) = \frac{1}{2\Psi} \mathcal{G}(x)^T \nabla \delta(x)^T W_c,$$ (21)

and

$$\xi_{u^*}(x) = -\frac{1}{2}(I_m - \Phi(\mathfrak{A}(x)))\mathcal{G}(x)^T\nabla\xi(x), \tag{22}$$

with $\Phi(\mathfrak{A}(x)) = diag\{tanh^2(\mathfrak{A}_l(x))\}(\mathfrak{l} = 1, 2, \dots, m)$ with $\mathfrak{A}_l(x) = [\mathfrak{A}_1(x); \mathfrak{A}_2(x); \dots; \mathfrak{A}_m(x)] \in \mathcal{R}^m$ being selected between $\bar{A}(x)$ and $T(x)$. Then, considering Equation (19), $v^*(x)$ in Equation (16) can be redefined as

$$v^*(x) = \frac{1}{2Y^2}\mathcal{P}(x)^T\nabla\delta(x)^T W_c + \xi_{v^*}(x), \tag{23}$$

where $\xi_{v^*}(x) = \frac{1}{2Y^2}\mathcal{P}(x)^T\nabla\xi(x)$.

Similarly, substituting Equation (19) into Equation (17), the HJI equation can be rewritten as

$$\begin{aligned}
0 =& W_c^T\nabla\delta(x)(\mathcal{F}(x) + \mathcal{P}(x)v^*) + x^TQx + \mathfrak{U}(-\Psi tanh(\bar{A}(x) + K(x) + \psi_\Im)) + \mathcal{B}_r(x) \\
&- \Psi W_c^T\nabla\delta(x)\mathcal{G}(x)tanh(\bar{A}(x) + K(x)) - \Psi\nabla\xi(x)^T\mathcal{G}(x)tanh(\bar{A}(x) + K(x)) \\
&+ \nabla\xi(x)^T(\mathcal{F}(x) + \mathcal{P}(x)v^*) - Y^2\|v^*\|^2 + (W_c^T\nabla\delta(x) + \nabla\xi(x)^T)\mathcal{G}(x)\psi_\Im,
\end{aligned} \tag{24}$$

where $K(x) = 1/(2\Psi)\mathcal{G}(x)^T\nabla\xi(x)$.

However, since the ideal CNN weight W_c in Equation (18) is unknown, it can not be used in the control procedure. Hence, the CNN is used to estimate the cost function and its gradient as

$$\hat{V}(x) = \hat{W}_c^T\delta(x), \tag{25}$$

$$\nabla\hat{V}(x) = \nabla\delta(x)^T\hat{W}_c, \tag{26}$$

where $\hat{W}_c$ represents the estimation of W_c.

Therefore, the approximate optimal input and the approximate worst disturbance input become

$$\hat{u}^*(x) = -\Psi tanh(\frac{1}{2\Psi}\mathcal{G}(x)^T\nabla\delta(x)^T\hat{W}_c) + \psi_\Im, \tag{27}$$

and

$$\hat{v}^*(x) = \frac{1}{2Y^2}\mathcal{G}(x)^T\nabla\delta(x)^T\hat{W}_c. \tag{28}$$

Subsequently, the approximated Hamilton function can be formulated by

$$\begin{aligned}
\hat{\mathcal{H}}(x, \hat{W}_c, \hat{v}^*) =& \hat{W}_c^T\eth + \hat{W}_c^T\nabla\delta(x)\mathcal{G}(x)\psi_\Im + \mathfrak{U}(-\Psi tanh(\Gamma(x)) + \psi_\Im) \\
&+ x^TQx - Y^2\|\hat{W}_c\|^2 - \Psi\hat{W}_c^T\nabla\delta(x)\mathcal{G}(x)tanh(\Gamma(x)) + \mathcal{B}_r(x),
\end{aligned} \tag{29}$$

where

$$\eth = \nabla\delta(x)(\mathcal{F}(x) + \mathcal{P}(x)\hat{v}^*) \tag{30}$$

and

$$\Gamma(x) = \frac{1}{2\Psi}\mathcal{G}(x)^T\nabla\delta(x)^T\hat{W}_c. \tag{31}$$

The CNN weight estimation error is denoted by

$$\tilde{W}_c = W_c - \hat{W}_c, \tag{32}$$

and the approximation error $\varrho_{\mathfrak{c}}$ of the Hamiltonian function is derived as

$$\begin{aligned}
\varrho_{\mathfrak{c}} &= \hat{\mathcal{H}}(x, \hat{W}_c, \hat{v}^*) - \mathcal{H}(x, u^*, v^*, \nabla V^*(x)) \\
&= \hat{\mathcal{H}}(x, \hat{W}_c, \hat{v}^*).
\end{aligned} \tag{33}$$

To achieve $\hat{W}_c \to W_c$, it is necessary to ensure that $\varrho_{\mathfrak{c}} \to 0$. Therefore, the chosen target function is denoted by $\mathrm{E} = \frac{1}{2}\varrho_{\mathfrak{c}}^T \varrho_{\mathfrak{c}}(1/(1+\eth^T\eth)^2)$, where $\mathcal{O} = 1 + \eth^T\eth$. Consequently, based on a normalized gradient descent algorithm, the weight vector $\hat{W}_c$ is defined by

$$\dot{\hat{W}}_c = -\frac{\alpha}{\mathcal{O}^2}\frac{\partial \mathrm{E}}{\partial \hat{W}_c} = -\frac{\alpha}{\mathcal{O}^2}\varrho_{\mathfrak{c}}, \tag{34}$$

with $\alpha > 0$ being the adjustable parameter and $\varrho_{\mathfrak{c}}$ defined as Equation (33).

Using Equations (32) and (34), the weight approximation error $\dot{\tilde{W}}_c$ can be expressed as

$$\dot{\tilde{W}}_c = \frac{\alpha\zeta}{\mathcal{O}}\xi_c - \alpha\zeta\zeta^T\tilde{W}_c, \tag{35}$$

where $\xi_c = -\nabla\xi(x)^T(\mathcal{F}(x) + \mathcal{P}(x)\hat{v}^*)$ is the residual error and $\zeta = \frac{\eth}{\mathcal{O}}$.

4.2. Stability Analysis

The UUB of both the state x and the CNN parameters in the closed-loop system is demonstrated by utilizing the Lyapunov stability analysis principle in this subsection. First, two assumptions that were also used in [28,42] are required, as

Assumption 2. *The ideal optimal CNN weight vector W_c is upper bounded, i.e., $\|W_c\| \le b_{W_c}$, where $b_{W_c} > 0$ is a constant. Moreover, for any $x \in \mho$, this paper assumes that there are two known constants $b_{\nabla\delta} > 0$, $b_\delta > 0$ so that $\|\nabla\delta(x)\| \le b_{\nabla\delta}, \|\delta(x)\| \le b_\delta$. Meanwhile, there exist $b_{\nabla\xi} > 0$ and $b_\xi > 0$ so that $\|\nabla\xi(x)\| \le b_{\nabla\xi}, \|\xi(x)\| \le b_\xi$ for any $x \in \mho$.*

Assumption 3. *We make $b_{\xi_{u^*}}, b_{\xi_{v^*}}, b_{\xi_c}$ be positive constants.*

(1) $b_{\xi_{u^*}} \ge \|\xi_{u^*}(x)\|$ *for any $x \in \mho$.*
(2) $b_{\xi_{v^*}} \ge \|\xi_{v^*}(x)\|$ *for any $x \in \mho$.*
(3) $b_{\xi_c} \ge \|\xi_c\|$ *for any $x \in \mho$.*

Theorem 1. *Assuming Assumptions 1–3 are met, we consider System (1) with the associated Control (27) and the update rule of CNN (34), ensuring all signals in the nonlinear system are UUB if the following condition holds:*

$$\alpha\Bbbk_{min}(\zeta\zeta^T) - (1/\mathrm{Y}^2)\Im_{\nabla\delta}^2\mathcal{P}_M^2 > 0. \tag{36}$$

Proof. We let the Lyapunov candidate function as the following (note: for convenience, $V^*(x)$ and $(1/2)\tilde{W}_c^T\tilde{W}_c$ are abbreviated as L_1 and L_2 below):

$$L(t) = \underbrace{V^*(x)}_{L_1} + \underbrace{(1/2)\tilde{W}_c^T\tilde{W}_c}_{L_2}. \tag{37}$$

Taking the derivation of L_1 in Equation (37) and using System (1), the derivation of L_1 can be expressed as

$$\begin{aligned}
\dot{L}_1 &= \frac{dV^*(x)}{dt} \\
&= \nabla V^*(x)^T(\mathcal{F}(x) + \mathcal{G}(x)\hat{u}^* + \mathcal{P}(x)\hat{v}^*) \\
&= \nabla V^*(x)^T(\mathcal{F}(x) + \mathcal{G}(x)u^* + \mathcal{P}(x)v^*) \\
&\quad + \nabla V^*(x)^T\mathcal{P}(x)(\hat{v}^* - v^*) + \nabla V^*(x)^T\mathcal{G}(x)(\hat{u}^* - u^*).
\end{aligned} \tag{38}$$

Then, using Equations (12) and (11), it can be derived as

$$\nabla V^*(x)^T(\mathcal{F}(x) + \mathcal{G}(x)u^* + \mathcal{P}(x)v^*) = -x^TQx - \mathfrak{U}(u^*) + \mathrm{Y}^2\|v^*\|^2 - \mathcal{B}_r(x). \tag{39}$$

Similarly, taking into account Equations (27) and (28), the derived results are

$$\nabla V^*(x)^T \mathcal{G}(x) = 2\Psi(tanh^{-1}((\psi_\Im - u^*)/(\Psi)))^T, \tag{40}$$

and

$$\nabla V^*(x)^T \mathcal{P}(x) = 2Y^2 v^{*T}. \tag{41}$$

According to Equations (38)–(41), Equation (38) can be rewritten as follows (note: for convenience, $\bar{\omega} - \mathfrak{U}(u^*)$ and $2Y^2 v^{*T}\hat{v}^* - Y^2\|v^*\|^2 - \mathcal{B}_r(x)$ are abbreviated as Λ_1 and Λ_2 below):

$$\dot{L}_1 = -x^T \mathcal{Q}x + \underbrace{\bar{\omega} - \mathfrak{U}(u^*)}_{\Lambda_1} + \underbrace{2Y^2 v^{*T}\hat{v}^* - Y^2\|v^*\|^2 - \mathcal{B}_r(x)}_{\Lambda_2}, \tag{42}$$

where

$$\bar{\omega} = 2\Psi(tanh^{-1}((\psi_\Im - u^*)/\Psi))(\hat{u}^* - u^*). \tag{43}$$

We apply Young's inequality to Equation (43). Additionally, considering Equations (19), (20), (27), (40) and (41), $\bar{\omega}$ can be formulated as

$$
\begin{aligned}
\bar{\omega} &\leq \left\| \Psi(tanh^{-1}((\psi_\Im - u^*)/\Psi)) \right\|^2 + \|\hat{u}^* - u^*\|^2 \\
&= \frac{1}{4}\left\| \mathcal{G}(x)^T \nabla V^*(x) \right\|^2 + \|\hat{u}^* - u^*\|^2 \\
&= \frac{1}{4}\left\| \mathcal{G}(x)^T (\nabla\delta(x)^T W_c + \nabla\xi(x)) \right\|^2 \\
&\quad + \left\| -\Psi tanh(\Gamma(x)) + \Psi tanh(\bar{A}(x)) - \xi_{u^*}(x) \right\|^2.
\end{aligned} \tag{44}
$$

Furthermore, utilizing Young's inequality, $\bar{\omega}$ in Equation (44) further yields

$$
\begin{aligned}
\bar{\omega} &\leq 2\left\| -\Psi tanh(\Gamma(x)) + \Psi tanh(\bar{A}(x)) \right\|^2 + 2\|\xi_{u^*}(x)\|^2 \\
&\quad + \frac{1}{2}\left(\left\| \mathcal{G}(x)^T \nabla\delta(x)^T W_c \right\|^2 \right) + \frac{1}{2}\left(\left\| \mathcal{G}(x)^T \nabla\xi(x) \right\|^2 \right) \\
&\leq 4\|\Psi tanh(\Gamma(x))\|^2 + \left\| \Psi tanh(\bar{A}(x)) \right\|^2 + 2\|\xi_{u^*}(x)\|^2 \\
&\quad + \frac{1}{2}\left(\left\| \mathcal{G}(x)^T \nabla\delta(x)^T W_c \right\|^2 \right) + \frac{1}{2}\left(\left\| \mathcal{G}(x)^T \nabla\xi(x) \right\|^2 \right).
\end{aligned} \tag{45}
$$

According to Equations (21) and (31), the following inequalities can be depicted as

$$\|tanh(\Gamma(x))\|^2 = tanh^2(\Gamma(x)) \leq m \tag{46}$$

and

$$\|tanh(\bar{A}(x))\|^2 = tanh^2(\bar{A}(x)) \leq m. \tag{47}$$

Based on Equation (46) and Assumptions 2 and 3, $\bar{\omega}$ can be expressed as

$$\bar{\omega} \leq 8\Psi^2 m + \frac{1}{2}\mathcal{G}_M^2(\Im_{\nabla\delta}^2 \Im_{\nabla W_c}^2 + \Im_{\nabla\xi}^2) + 2b_{\xi_{u^*}}^2. \tag{48}$$

By observing Equations (4) and (5), it can be concluded that $\mathfrak{U}(u^*) > 0$. Using Young's inequality and Equation (48), the expression of Λ_1 in Equation (42) can be rewritten as

$$\Lambda_1 \leq 8\Psi^2 m + \frac{1}{2}\mathcal{G}_M^2(\Im_{\nabla\delta}^2 \Im_{\nabla W_c}^2 + \Im_{\nabla\xi}^2) + 2b_{\xi_{u^*}}^2. \tag{49}$$

Similarly, Λ_2 in Equation (42) can be rewritten as follows (note: from Assumption 1, $\mathcal{B}_r(x) \geq 0$):

$$
\begin{aligned}
\Lambda_2 &= Y^2 \|\hat{v}^*\|^2 - \mathcal{B}_r(x) \\
&= -Y^2\|v^*\|^2 + Y^2\|v^*\|^2 + Y^2\|\hat{v}^*\|^2 - \mathcal{B}_r(x) \\
&\leq -Y^2\|v^*\|^2 + Y^2\|v^*\|^2 + Y^2\|\hat{v}^*\|^2 \\
&\leq -Y^2\|v^*\|^2 + Y^2(\|v^*\|^2 + \|\hat{v}^*\|^2) \\
&= (1/(4Y^2))\left\|\mathcal{P}(x)^T \nabla \delta(x)^T (W_c - \tilde{W}_c)\right\|^2.
\end{aligned}
\tag{50}
$$

Meanwhile, using Young's inequality and Assumptions 1 and 3, Λ_2 in Equation (50) further yields

$$
\begin{aligned}
\Lambda_2 &\leq (1/(4Y^2))\mathcal{P}_M^2 \mathfrak{I}_{\nabla\delta}^2 \|W_c - \tilde{W}_c\|^2 \\
&\leq (1/(2Y^2))\mathcal{P}_M^2 \mathfrak{I}_{\nabla\delta}^2 (\mathfrak{I}_{W_c}^2 + \|\tilde{W}_c\|^2).
\end{aligned}
\tag{51}
$$

Hence, by observing Equations (49) and (51), it can be inferred that $\dot{L}_1$ in Equation (42) satisfies

$$
\begin{aligned}
\dot{L}_1 &\leq -\Bbbk_{min}(\mathcal{Q})\|x\|^2 + (1/(2Y^2))\mathcal{P}_M^2 \mathfrak{I}_{\nabla\delta}^2 \mathfrak{I}_{W_c}^2 + 8\Psi^2 m + 2b_{\xi_{u^*}}^2 \\
&\quad + (1/2)\mathcal{G}_M^2(\mathfrak{I}_{\nabla\delta}^2 \mathfrak{I}_{W_c}^2 + \mathfrak{I}_{\nabla\xi}^2) + (1/(2Y^2))\mathcal{P}_M^2 \mathfrak{I}_{\nabla\delta}^2 \|\tilde{W}_c\|^2.
\end{aligned}
\tag{52}
$$

Then, the derivative of L_2 in Equation (37) along the solution of Equation (34) is as follows (note: $\alpha \tilde{W}_c^T(\zeta/\mathcal{O})\xi_c$ is abbreviated as Λ_3 below):

$$
\dot{L}_2 = \tilde{W}_c^T \dot{\tilde{W}}_c = \underbrace{\alpha \tilde{W}_c^T(\zeta/\mathcal{O})\xi_c}_{\Lambda_3} - \alpha \tilde{W}_c^T \zeta\zeta^T \tilde{W}_c.
\tag{53}
$$

Immediately after, using Young's inequality, Λ_3 can be depicted as

$$
\begin{aligned}
\Lambda_3 &\leq \frac{\alpha}{2\mathcal{O}}\left(\left\|\zeta^T \tilde{W}_c\right\|^2 + \|\xi_c\|^2\right) \\
&\leq \alpha\left(\frac{1}{2}\tilde{W}_c^T \zeta\zeta^T \tilde{W}_c + \frac{1}{2}\|\xi_c\|^2\right).
\end{aligned}
\tag{54}
$$

Additionally, with Assumption 3 holding, it can be deduced that $\dot{L}_2$ in Equation (53) satisfies

$$
\begin{aligned}
\dot{L}_2 &\leq \frac{1}{2}\left(-\alpha \tilde{W}_c^T \zeta\zeta^T \tilde{W}_c + \alpha\|\xi_c\|^2\right) \\
&\leq \frac{1}{2}\left(-\alpha\Bbbk_{min}(\zeta\zeta^T)\|\tilde{W}_c\|^2 + \alpha\mathfrak{I}_{\xi_c}^2\right).
\end{aligned}
\tag{55}
$$

Using Equations (37), (52) and (55), $\dot{L}$ can be depicted as

$$
\begin{aligned}
\dot{L} &\leq -\Bbbk_{min}(\mathcal{Q})\|x\|^2 + (1/(2Y^2)\mathcal{P}_M^2 \mathfrak{I}_{\nabla\delta}^2 \mathfrak{I}_W^2) + (1/2)\mathcal{G}_M^2(\mathfrak{I}_{\nabla\delta}^2 \mathfrak{I}_W^2 + \mathfrak{I}_{\nabla\xi}^2) \\
&\quad - (1/2)(\alpha\Bbbk_{min}\zeta\zeta^T - (1/Y^2)\mathcal{P}_M^2 \mathfrak{I}_{\nabla\delta}^2)\|\tilde{W}_c\|^2 + 8\Psi^2 m + 2b_{\xi_{u^*}}^2 + (\alpha/2)\mathfrak{I}_{\xi_c}^2.
\end{aligned}
\tag{56}
$$

Finally, $\dot{L} < 0$ is true if $x \notin \mho(x)$ or $\tilde{W}_c \notin \mho(\tilde{W}_c)$, and based on Equation (36), $\mho(x)$ and $\mho(\tilde{W}_c)$ can be respectively formulated as

$$
\mho(x) = \left\{\|x\| \leq \sqrt{\frac{\alpha\mathfrak{I}_{\xi_c}^2 + \Xi + A_1 \mathcal{P}_M^2}{2\Bbbk_{min}(\mathcal{Q})}}\right\},
\tag{57}
$$

and

$$\mho(\tilde{W}_c) = \left\{ \|\tilde{W}_c\| \leq \sqrt{\frac{\alpha \Im_{\tilde{\zeta}_c}^2 + \Xi + A_1 \mathcal{P}_M^2}{\alpha \Bbbk_{min}(\zeta\zeta^T) - A_2 \mathcal{P}_M^2}} \right\}, \tag{58}$$

where $\Xi = \mathcal{G}_M^2(\Im_{\nabla\delta}^2 \Im_{W_c}^2 + \Im_{\nabla\zeta}^2) + 16\Psi^2 m + 4b_{\zeta u^*}^2$, $A_1 = (1/Y^2)\Im_{\nabla\delta}^2 \Im_{W_c}^2$ and $A_2 = (1/Y^2)\Im_{\nabla\delta}^2$.

To summarize, the Lyapunov stability method has been used to demonstrate the state x of Equation (1) and $\tilde{W}_c$ are UUB, with Equations (57) and (58) representing their respective bounds. The proof is complete. $\square$

5. Simulation Study

Within this section, two examples are utilized to validate the efficacy of the proposed approach.

5.1. Example 1

Consider the F16 aircraft plant used in [28] as

$$\dot{x} = \mathcal{F}(x) + \mathcal{G}(x)u + \mathcal{P}(x)v, \tag{59}$$

where $x(t) = [x_1, x_2, x_3]^T \in \mathcal{R}^3$ with $x_0 = [1, -1, 1]^T$ represents the system state vector, where x_1, x_2 and x_3 represent the attack angle, the pitch rate, and the elevator deflection angle, respectively. u is control input, v is disturbance input. The internal dynamics, control, and disturbance coefficient matrices are expressed as

$$\mathcal{F}(x) = \begin{bmatrix} -1.01887x_1 + 0.90506x_2 - 0.00215x_3 \\ 0.82225x_1 - 1.07741x_2 - 0.17555x_3 \\ -x_3 \end{bmatrix}, \mathcal{G}(x) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \mathcal{P}(x) = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}.$$

The control input u is constrained to be greater than -1 and less than 2. Hence, $\Psi = 1.5$ and $\Im = 0.5$. And then, the danger region is described as a ball with a radius of 0.15 and a center at $[0.3, 0.05, -0.05]^T$. The $y(x)$ is chosen as

$$\frac{1.5\sqrt{(x_1 - 0.3)^2 + 0.1(x_2 - 0.05)^2 + 1.2(x_3 + 0.05)^2} - 0.15}{\sqrt{(x_1 - 0.3)^2 + (x_2 - 0.05)^2 + 25(x_3 + 0.05)^2} - 0.15}.$$

The $z(x)$ is chosen as

$$\sqrt{(x_1 - 0.3)^2 + (x_2 - 0.05)^2 + (x_3 + 0.05)^2} - 0.15.$$

In addition, substituting Ψ and $\Im$ into Equation (4), $\mathfrak{U}(u)$ can be expressed as

$$\mathfrak{U}(u) = 2\Psi(u - \Im)tanh^{-1}(\frac{u - \Im}{\Psi}) + \Psi^2 \ln(1 - \frac{(u - \Im)^2}{\Psi})$$

$$= 3(u - 0.5)tanh^{-1}(\frac{u - 0.5}{1.5}) + 2.25\ln(1 - \frac{(u - 0.5)^2}{1.5}). \tag{60}$$

Letting $\mathcal{Q} = I_3$ and $Y = 2$, the cost function for Equation (62) is formulated as

$$V(x) = \int_0^\infty x(t)^T \mathcal{Q}x(t) + \mathfrak{U}(u) - 2^2\|v\|^2 + \mathcal{B}_r(x)dt, \tag{61}$$

where $\mathcal{B}_r(x) = \rho\frac{y(x)}{z(x)}$ represents the CBF and $\rho = 2$.

The activation function is given as $\delta(x) = [x_1^2, x_1 x_2, x_1 x_3, x_2^2, x_2 x_3, x_3^2]^T$ and the CNN weight vector is $\hat{W}_c = [\hat{W}_{c1}, \hat{W}_{c2}, \hat{W}_{c3}, \hat{W}_{c4}, \hat{W}_{c5}, \hat{W}_{c6}]^T$. In addition, the adjustable parame-

ter α is 10, and the original parameters of the CNN are configured as 1. At last, the probing noise $exp(-0.1t)(0.001)(sin(t)^2cos(t) + sin(2t)^2cos(0.1t))$ is added to the control input policy for the initial 30 s in order to ensure the persistence of the excitation.

Through simulation experiments, Figures 1–7 are obtained. Figure 1 displays that $\hat{W}_c$ is convergent after the first 10 s, and can know the ideal vector $W_c^* = [16.4603, -6.5022, -4.3910, 4.8851, 3.7081, 11.6158]^T$. Figure 2 displays the convergence of the states x_1, x_2, and x_3. Figure 3 displays the danger region, which is represented by the ball, and the original states are in the danger area. However, the system states controlled by the safe optimal controller bypass this ball, and as the damping coefficient ρ increases, the distance between the system states and the dangerous region becomes larger and larger. Figure 3 shows that as states x_1, x_2, and x_3 gradually approach the danger zone, the convergence of x_3 is accelerated due to the CBF and cost function. Figure 4 presents the control input u with asymmetric input constraints. The plot reveals that the value of u remains within the specified range, bounded by $u_{max} = 2$ and $u_{min} = -1$, providing evidence that the asymmetric input constraints are implemented successfully. Figure 5 presents the disturbance input v. Figure 6 presents the cost function of the system. It can be seen that when the system states confront the danger area, the cost function changes significantly and eventually converges to zero. According to the principle of optimal control, when the cost function converges to zero, the following conclusion can be drawn: The cost function imposes a higher penalty on control actions that do not comply with the asymmetric input constraints and safety constraints. Therefore, when the cost function converges to zero, the system finds the optimal control actions that satisfy all the constraints.

In order to further show the efficiency of the presented method, Equation (4) is redefined as $u^T R u$ (where $R = I_1$), and the simulation results are illustrated in Figure 7. Subsequently, Figure 4 illustrates the control input, which is restricted to the limits of -1 to 2. This can be observed by comparing it with Figure 7, where the input is clearly outside this range.
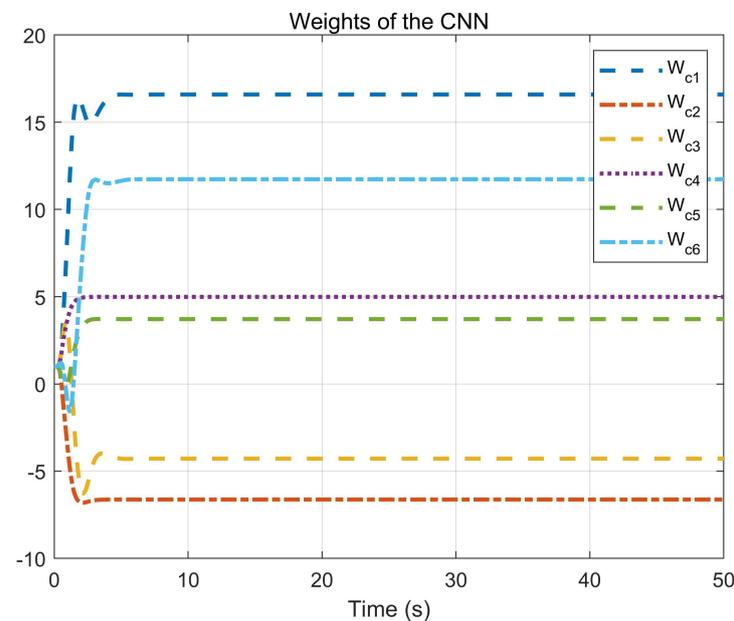
Figure 1. Convergence of the CNN weights.
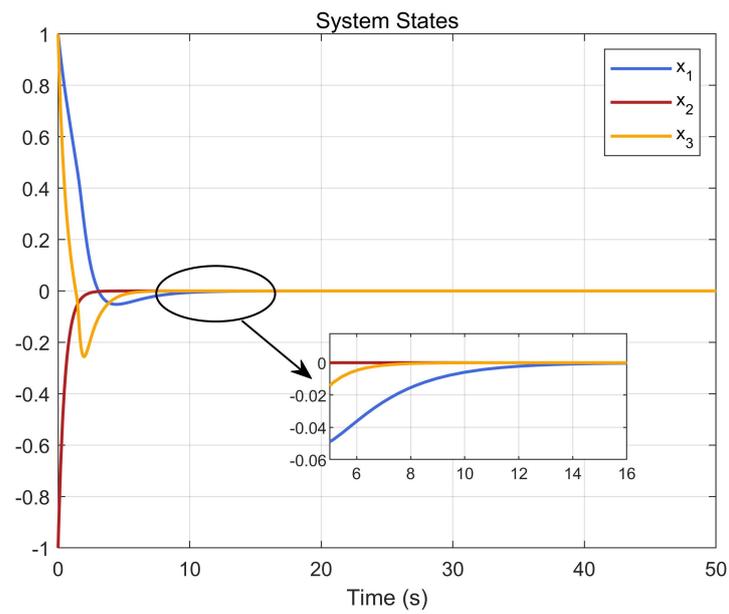
Figure 2. Convergence of system states x_1, x_2, and x_3.
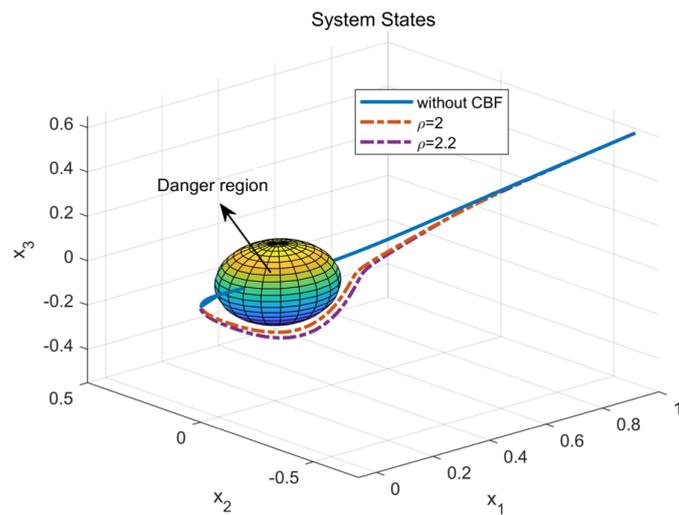
Figure 3. The comparison between the safe and unsafe states.
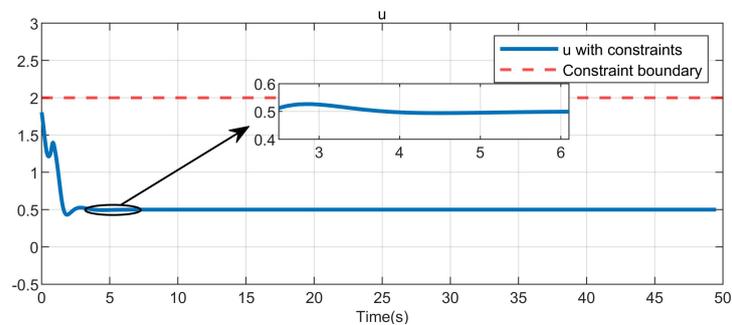
Figure 4. Control input in the system.
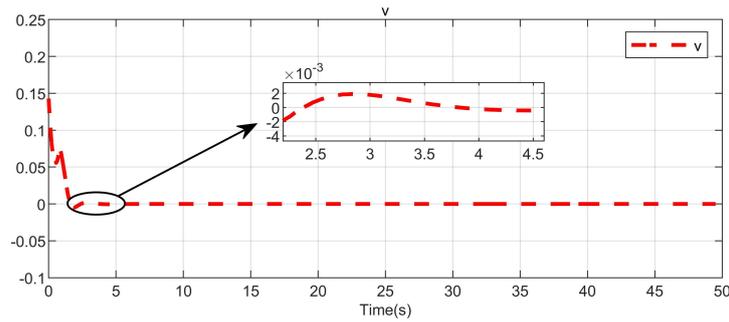
Figure 5. Disturbance input in the system.
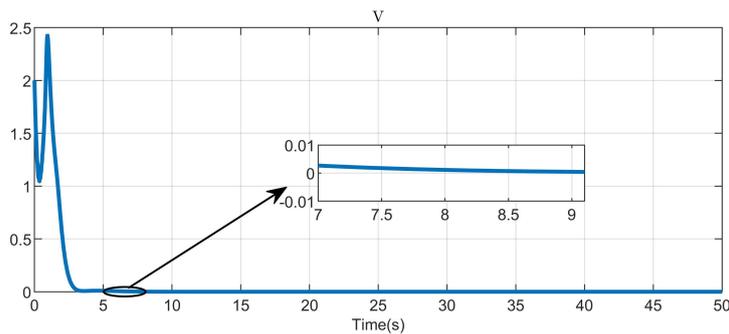
Figure 6. The cost function of the system.
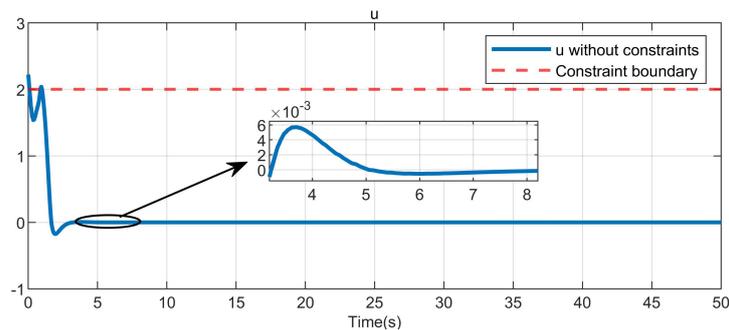
Figure 7. Control input without asymmetric input constraints.

5.2. Example 2

We consider the nonlinear system as

$$\dot{x} = \mathcal{F}(x) + \mathcal{G}(x)\mathbf{u} + \mathcal{P}(x)\mathbf{v}, \tag{62}$$

where $x(t) = [x_1, x_2]^T \in \mathcal{R}^2$ with $x_0 = [1, -1]^T$ represents the system state vector; the internal dynamics, control, and disturbance coefficient matrices are expressed as

$$\mathcal{F}(x) = \begin{bmatrix} -\frac{1}{2}x_1 + x_2 \\ -2x_2 cos(2x_1) \end{bmatrix}, \ \mathcal{G}(x) = \begin{bmatrix} 0 \\ -x_1 \end{bmatrix}, \ \mathcal{P}(x) = \begin{bmatrix} 0 \\ x_1 \end{bmatrix}.$$

Just like F16, the control input u is subject to an asymmetrical boundary, with a lower bound of -1 and an upper bound of 3, establishing its limits. Hence, $\Psi = 2$ and $\Im = 1$. And then, the danger region is described as a circle with radius $= 0.1$, and the center of the circle is $[0.19, -0.12]^T$. The $y(x)$ is chosen as

$$atan(\frac{1}{\sqrt{(x_1 - 0.19)^2 + (x_2 + 0.12)^2} - 0.1}).$$

The $z(x)$ is chosen as

$$\sqrt{(x_1 - 0.19)^2 + (x_2 + 0.12)^2} - 0.1.$$

In addition, substituting Ψ and $\Im$ into Equation (4), $\mathfrak{U}(u)$ can be expressed as

$$\mathfrak{U}(u) = 2\Psi(u - \Im)tanh^{-1}(\frac{u - \Im}{\Psi}) + \Psi^2 \ln(1 - \frac{(u - \Im)^2}{\Psi})$$

$$= 4(u - 1)tanh^{-1}(\frac{u - 1}{2}) + 4\ln(1 - \frac{(u - 1)^2}{2}). \tag{63}$$

Letting $\mathcal{Q} = I_2$ and $Y = 1.35$, the cost function for Equation (62) is formulated as

$$V(x) = \int_0^\infty x(t)^T \mathcal{Q}x(t) + \mathfrak{U}(u) - 1.35^2 \|v\|^2 + \mathcal{B}_r(x)dt, \tag{64}$$

where $\mathcal{B}_r(x) = \rho\frac{y(x)}{z(x)}$ represents the CBF and $\rho = 0.3$.

Then, the CNN presented as Equation (18) is applied to address the HJI equation for Equation (62). The activation function is given as $\delta(x) = [x_1^2, x_1x_2, x_2^2, x_1^4, x_1^3x_2, x_1^2x_2^2, x_1x_2^3, x_2^4]^T$ and the CNN weight vector is $\hat{W}_c = [\hat{W}_{c1}, \hat{W}_{c2}, \hat{W}_{c3}, \hat{W}_{c4}, \hat{W}_{c5}, \hat{W}_{c6}, \hat{W}_{c7}, \hat{W}_{c8}]^T$. In addition, the adjustable parameter α is 20, the original parameters of the CNN are configured as 1. At last, the probing noise $exp(-0.001t)(-0.1(sin(t)^2cos(t) + sin(t)^5 + sin(2t)^2cos(0.1t) + sin(-1.2t)^2cos(0.5t))$ is added to the control input policy for the initial 30 s.

Through simulation experiments, Figures 8–14 are obtained. Figure 8 displays that $\hat{W}_c$ is convergent after the first 10 s, and can know the ideal vector $W_c^* = [84.6487, -12.2017, 9.5269, 11.7425, -3.0924, 3.4273, -0.5533, 2.0591]^T$. Figure 9 displays the convergence of the states x_1 and x_2. Figure 10 illustrates the relationship between the system states and the dangerous area, revealing that increasing the damping factor ρ leads to a greater distance between the system states and the dangerous zone. Evidently, system states x_1 and x_2 with a safe and optimal controller take an alternate route to avoid the dangerous region, while the conventional optimal controller cannot circumvent the dangerous region. As can be seen from Figure 10, when states x_1 and x_2 gradually approach the danger zone, the convergence speed of x_2 is accelerated due to the influence of CBF and cost function and obtains an optimal trajectory around the danger zone again. Figure 11 shows input u with asymmetric input constraints. The plot reveals that the value of u remains within the specified range, bounded by $u_{max} = 3$ and $u_{min} = -1$, providing evidence that the asymmetric input constraints are implemented successfully. Figure 12 presents disturbance input v. Figure 13 presents the cost function of the system. It can be seen that the cost function eventually converges to zero. Similar to the linear system, when the cost function converges to zero, it can be concluded that the system finds the optimal control action that satisfies the asymmetric input constraints and safety constraints.

In this paper, asymmetric input constraints and unmatched disturbances are applied to nonlinear safety-critical systems for the first time, and Equation (4) is used to handle the asymmetric input constraints. To further demonstrate the efficacy of the presented algorithm, as in articles [14,16,28], (4) is redefined as u^TRu (where $R = I_1$) and the simulation results are shown in Figure 14. Subsequently, the control input in Figure 11 is constrained to fall within the limits of -1 to 3, as can be observed by comparing it with Figure 14, while the input in Figure 14 is clearly outside this range.
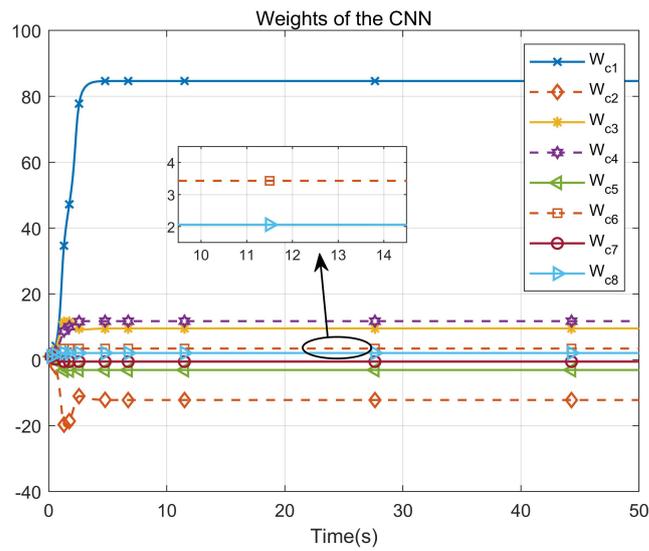
Figure 8. Convergence of the CNN weights.
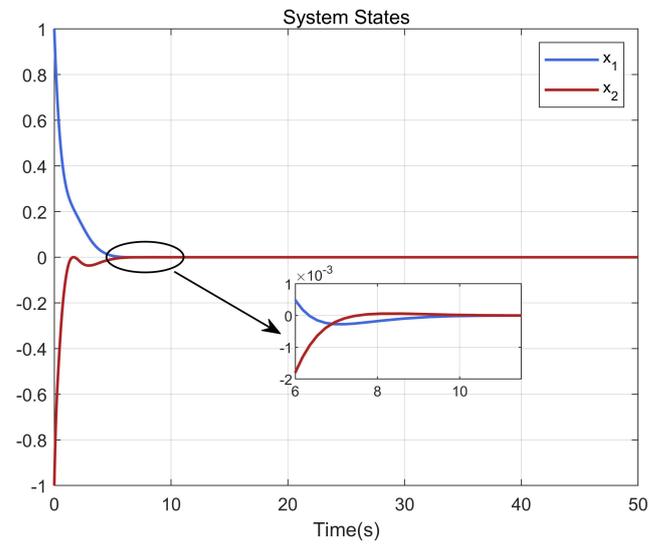
Figure 9. Convergence of system states x_1 and x_2.
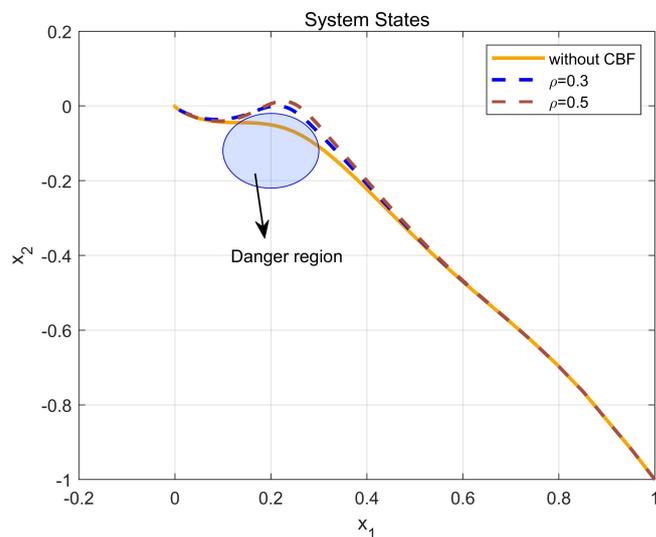
Figure 10. The comparison between the safe and unsafe states.
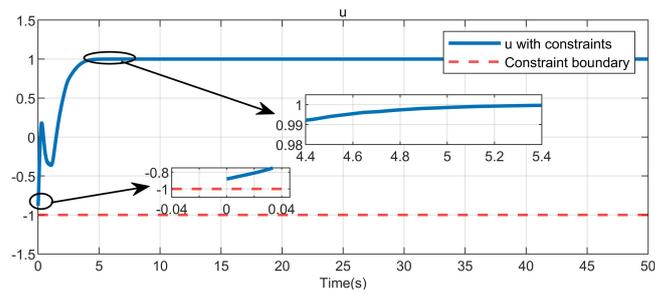
Figure 11. Control input in the system.
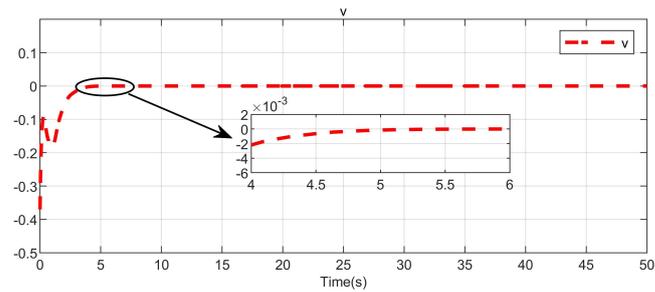
Figure 12. Disturbance input in the system.
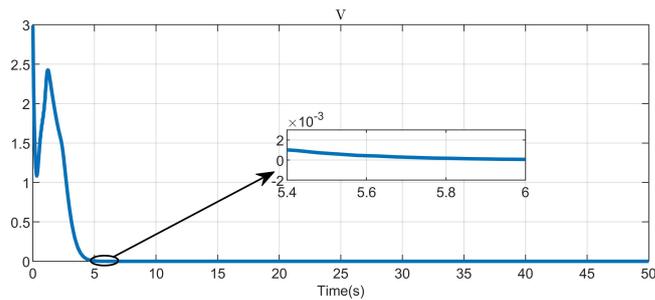
Figure 13. The cost function of the system.
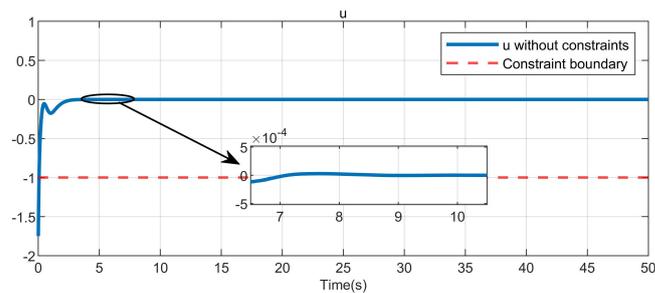
Figure 14. Control input without asymmetric input constraints.

6. Conclusions

The safe and optimal control problem of the nonlinear CT safety-critical systems with asymmetric input constraints and unmatched disturbances was addressed. Firstly, the new non-quadratic form function was considered for addressing the issue of asymmetric input constraints. Then, the control design was transformed into the two-player ZSG problem to handle unmatched disturbances. In order to obtain the optimal controller for safety, the combination of the CBF and cost function was directly used to penalize unsafe behavior. Moreover, the CNN was applied to reduce the computational complexity of dual actor–critic network. The effectiveness of the proposed method was validated by the simulation results.

References

1. Yi, X.; Luo, B.; Zhao, Y. Adaptive dynamic programming-based visual servoing control for quadrotor. *Neurocomputing* **2022**, *504*, 251–261. [CrossRef]
2. Liscouët, J.; Pollet, F.; Jézégou, J.; Budinger, M.; Delbecq, S.; Moschetta, J. A methodology to integrate reliability into the conceptual design of safety-critical multirotor unmanned aerial vehicles. *Aerosp. Sci. Technol.* **2022**, *127*, 107681. [CrossRef]
3. Dou, L.; Cai, S.; Zhang, X.; Su, X.; Zhang, R. Event-triggered-based adaptive dynamic programming for distributed formation control of multi-UAV. *J. Frankl. Inst.* **2022**, *359*, 3671–3691. [CrossRef]
4. Molnar, T.; Cosner, R.; Singletary, A.; Ubellacker, W.; Ames, A. Model-free safety-critical control for robotic systems. *IEEE Robot. Autom. Lett.* **2021**, *7*, 944–951. [CrossRef]
5. Nguyen, Q.; Sreenath, K. Robust safety-critical control for dynamic robotics. *IEEE Trans. Autom. Control* **2021**, *67*, 1073–1088. [CrossRef]
6. Liu, S.; Liu, L.; Yu, Z. Safe reinforcement learning for affine nonlinear systems with state constraints and input saturation using control barrier functions. *Neurocomputing* **2023**, *518*, 562–576. [CrossRef]
7. Han, J.; Liu, X.; Wei, X.; Sun, S. A dynamic proportional-integral observer-based nonlinear fault-tolerant controller design for nonlinear system with partially unknown dynamic. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *52*, 5092–5104. [CrossRef]
8. Ohnishi, M.; Wang, L.; Notomista, G.; Egerstedt, M. Barrier-certified adaptive reinforcement learning with applications to brushbot navigation. *IEEE Trans. Robot.* **2019**, *35*, 1186–1205. [CrossRef]
9. Bianchi, D.; Di Gennaro, S.; Di Ferdinando, M.; Acosta Lùa, C. Robust Control of UAV with Disturbances and Uncertainty Estimation. *Machines* **2023**, *11*, 352. [CrossRef]
10. Bianchi, D.; Borri, A.; Di Benedetto, M.; Di Gennaro, S. Active Attitude Control of Ground Vehicles with Partially Unknown Model. *IFAC-PapersOnLine* **2020**, *53*, 14420–14425. [CrossRef]
11. Ames, A.; Xu, X.; Grizzle, J.; Tabuada, P. Control barrier function based quadratic programs for safety critical systems. *IEEE Trans. Autom. Control* **2016**, *62*, 3861–3876. [CrossRef]
12. Wang, H.; Peng, J.; Zhang, F.; Zhang, H.; Wang, Y. High-order control barrier functions-based impedance control of a robotic manipulator with time-varying output constraints. *ISA Trans.* **2022**, *129*, 361–369. [CrossRef] [PubMed]
13. Liu, S.; Liu, L.; Yu, Z. Safe reinforcement learning for discrete-time fully cooperative games with partial state and control constraints using control barrier functions. *Neurocomputing* **2023**, *517*, 118–132. [CrossRef]
14. Qin, C.; Wang, J.; Zhu, H.; Zhang, J.; Hu, S.; Zhang, D. Neural network-based safe optimal robust control for affine nonlinear systems with unmatched disturbances. *Neurocomputing* **2022**, *506*, 228–239. [CrossRef]
15. Xu, X.; Tabuada, P.; Grizzle, J.W.; Ames, A. Robustness of control barrier functions for safety critical control. *IFAC-PapersOnLine* **2015**, *48*, 54–61. [CrossRef]
16. Marvi, Z.; Kiumarsi, B. Safe reinforcement learning: A control barrier function optimization approach. *Int. J. Robust Nonlinear Control* **2021**, *31*, 1923–1940. [CrossRef]
17. Xiao, W.; Belta, C.; Cassandras, C. Adaptive control barrier functions. *IEEE Trans. Autom. Control* **2021**, *67*, 2267–2281. [CrossRef]
18. Modares, H.; Lewis, F.; Sistani, M. Online solution of nonquadratic two-player zero-sum games arising in the H∞ control of constrained input systems. *Int. J. Adapt. Control Signal Process.* **2014**, *28*, 232–254. [CrossRef]
19. Qin, C.; Zhu, H.; Wang, J.; Xiao, Q.; Zhang, D. Event-triggered safe control for the zero-sum game of nonlinear safety-critical systems with input saturation. *IEEE Access* **2022**, *10*, 40324–40337. [CrossRef]
20. Song, R.; Zhu, L. Stable value iteration for two-player zero-sum game of discrete-time nonlinear systems based on adaptive dynamic programming. *Neurocomputing* **2019**, *340*, 180–195. [CrossRef]
21. Lu, W.; Li, Q.; Lu, K.; Lu, Y.; Guo, L.; Yan, W.; Xu, F. Load adaptive PMSM drive system based on an improved ADRC for manipulator joint. *IEEE Access* **2021**, *9*, 33369–33384. [CrossRef]
22. Qin, C.; Qiao, X.; Wang, J.; Zhang, D. Robust Trajectory Tracking Control for Continuous-Time Nonlinear Systems with State Constraints and Uncertain Disturbances. *Entropy* **2022**, *24*, 816. [CrossRef] [PubMed]
23. Fan, Q.; Yang, G. Adaptive actor—Critic design-based integral sliding-mode control for partially unknown nonlinear systems with input disturbances. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 165–177. [CrossRef]

24. Yang, X.; He, H. Event-driven *H*∞-constrained control using adaptive critic learning. *IEEE Trans. Cybern.* **2020**, *51*, 4860–4872. [CrossRef] [PubMed]

25. Lewis, F.; Vrabie, D.; Syrmos, V. *Optimal Control*; John Wiley & Sons: Hoboken, NJ, USA, 2012.

26. Kiumarsi, B.; Vamvoudakis, K.; Modares, H.; Lewis, F. Optimal and autonomous control using reinforcement learning: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 2042–2062. [CrossRef]

27. Liu, D.; Xue, S.; Zhao, B.; Luo, B.; Wei, Q. Adaptive dynamic programming for control: A survey and recent advances. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *51*, 142–160. [CrossRef]

28. Vamvoudakis, K.; Lewis, F. Online actor—Critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* **2010**, *46*, 878–888. [CrossRef]

29. Han, H.; Zhang, J.; Yang, H.; Hou, Y.; Qiao, J. Data-driven robust optimal control for nonlinear system with uncertain disturbances. *Inf. Sci.* **2023**, *621*, 248–264. [CrossRef]

30. Lou, X.; Zhang, X.; Ye, Q. Robust control for uncertain impulsive systems with input constraints and external disturbance. *Int. J. Robust Nonlinear Control* **2022**, *32*, 2330–2343. [CrossRef]

31. Wang, N.; Gao, Y.; Yang, C.; Zhang, X. Reinforcement learning-based finite-time tracking control of an unknown unmanned surface vehicle with input constraints. *Neurocomputing* **2022**, *484*, 26–37. [CrossRef]

32. Liu, C.; Zhang, H.; Xiao, G.; Sun, S. Integral reinforcement learning based decentralized optimal tracking control of unknown nonlinear large-scale interconnected systems with constrained-input. *Neurocomputing* **2019**, *323*, 1–11. [CrossRef]

33. Yang, X.; Zhao, B. Optimal neuro-control strategy for nonlinear systems with asymmetric input constraints. *IEEE/CAA J. Autom. Sin.* **2020**, *7*, 575–583. [CrossRef]

34. Tang, Y.; Yang, X. Robust tracking control with reinforcement learning for nonlinear-constrained systems. *Int. J. Robust Nonlinear Control* **2022**, *32*, 9902–9919. [CrossRef]

35. Zhou, W.; Liu, H.; He, H.; Yi, J.; Li, T. Neuro-optimal tracking control for continuous stirred tank reactor with input constraints. *IEEE Trans. Ind. Inform.* **2018**, *15*, 4516–4524. [CrossRef]

36. Kong, L.; He, W.; Dong, Y.; Cheng, L.; Yang, C.; Li, Z. Asymmetric bounded neural control for an uncertain robot by state feedback and output feedback. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *51*, 1735–1746. [CrossRef]

37. Zhang, Y.; Zhao, B.; Liu, D.; Zhang, S. Event-triggered control of discrete-time zero-sum games via deterministic policy gradient adaptive dynamic programming. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *52*, 4823–4835. [CrossRef]

38. Zhang, S.; Zhao, B.; Liu, D.; Zhang, Y. Observer-based event-triggered control for zero-sum games of input constrained multi-player nonlinear systems. *Neural Netw.* **2021**, *144*, 101–112. [CrossRef]

39. Wei, Q.; Liu, D.; Lin, Q.; Song, R. Adaptive dynamic programming for discrete-time zero-sum games. *IEEE Trans. Neural Networks Learn. Syst.* **2017**, *29*, 957–969. [CrossRef]

40. Perrusquía, A.; Yu, W. Continuous-time reinforcement learning for robust control under worst-case uncertainty. *Int. J. Syst. Sci.* **2021**, *52*, 770–784. [CrossRef]

41. Yang, Y.; Vamvoudakis, K.; Modares, H.; Yin, Y.; Wunsch, D. Safe intermittent reinforcement learning with static and dynamic event generators. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 5441–5455. [CrossRef]

42. Fu, Z.; Xie, W.; Rakheja, S.; Na, J. Observer-based adaptive optimal control for unknown singularly perturbed nonlinear systems with input constraints. *IEEE/CAA J. Autom. Sin.* **2017**, *4*, 48–57. [CrossRef]