



Article Reinforcement Learning-Based Decentralized Safety Control for Constrained Interconnected Nonlinear Safety-Critical Systems

Chunbin Qin ¹, Yinliang Wu ¹, Jishi Zhang ^{2,*} and Tianzeng Zhu ¹

- ¹ School of Artificial Intelligence, Henan University, Zhengzhou 450046, China; qcb@henu.edu.cn (C.Q.); 104754222776@henu.edu.cn (Y.W.); Kz520@henu.edu.cn (T.Z.)
- ² School of Software, Henan University, Kaifeng 475000, China

* Correspondence: 10250051@vip.henu.edu.cn

Abstract: This paper addresses the problem of decentralized safety control (DSC) of constrained interconnected nonlinear safety-critical systems under reinforcement learning strategies, where asymmetric input constraints and security constraints are considered. To begin with, improved performance functions associated with the actuator estimates for each auxiliary subsystem are constructed. Then, the decentralized control problem with security constraints and asymmetric input constraints is transformed into an equivalent decentralized control problem with asymmetric input constraints using the barrier function. This approach ensures that safety-critical systems operate and learn optimal DSC policies within their safe global domains. Then, the optimal control strategy is shown to ensure that the entire system is uniformly ultimately bounded (UUB). In addition, all signals in the closed-loop auxiliary subsystem, based on Lyapunov theory, are uniformly ultimately bounded, and the effectiveness of the designed method is verified by practical simulation.

Keywords: interconnected nonlinear safety-critical systems; barrier function; asymmetric input constraints; safety constraints; decentralized control



Citation: Qin, C.; Wu, Y.; Zhang, J.; Zhu, T. Reinforcement Learning-Based Decentralized Safety Control for Constrained Interconnected Nonlinear Safety-Critical Systems. *Entropy* **2023**, *25*, 1158. https:// doi.org/10.3390/e25081158

Academic Editor: António Lopes

Received: 30 May 2023 Revised: 21 June 2023 Accepted: 1 July 2023 Published: 2 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

Over the past few decades, safety has received increasing attention in autonomous driving [1], intelligent robots [2], robotic arms [3], adaptive cruise control [4], etc. The design of these systems and controllers require that the system state trajectories evolve within a set called the safe set, reflecting the inherent properties of the system [5]. In practice, many engineering systems must operate within a specific safety range, beyond which the controlled system may be at risk [6]. Safety-critical systems primarily refer to systems having control behaviors that prioritize safety. The designed control schemes aim to reduce the potential for severe consequences, such as personal injury and environmental pollution, which may arise due to system shutdown or operational errors [7]. To ensure the safety and reliability of the system, scholars developed many safety control schemes. The classical approach focused on extending and applying Naguma's theorem to safe sets defined by continuously differentiable functions [8]. In particular, barrier functions have become an effective tool for verifying security and have been widely used in [9-11]. They were used to convert a system with security constraints into an equivalency system that satisfies security requirements and then a security controller was designed to protect the system. In [9,10], penalty functions and BF-based state transitions were employed to merge states into a reinforcement learning framework to solve optimal control problems with full-state constraints. In [11], a safe non-strategic reinforcement learning method to solve secure nonlinear systems with dynamic uncertainty was proposed. In [12,13], a new secure reinforcement learning method was proposed to solve secure nonlinear systems with symmetric input constraints. However, the results in [9–13], mentioned

2 of 23

above, were mainly based on studying the optimal safety control in a single continuoustime/discrete-time nonlinear system. The security control of interconnected systems has not been fully resolved.

On the other hand, interconnected systems consist of multiple subsystems with interconnected characteristics, and designing controllers for them through a concept similar to that of a single-system approach is difficult [14]. To solve this problem [15-17], the decentralized control approach, based on local subsystem information, was proposed. This approach involved using multiple controllers to control the interconnected systems. In [18,19], the decentralized control approach differed by initially decomposing the entire system control problem into a series of subproblems that could be solved independently. The solutions to the subproblems (i.e., independent controllers) were then joined to form a decentralized controller to stabilize the entire system. In addition, implementing the decentralized control algorithm used only the local subsystem's knowledge, not the complete system's information. Recently, scholars have proposed many schemes or techniques for designing decentralized controllers, including quantization techniques [20], fuzzy techniques [21], and optimal control methods [22]. This paper develops decentralized control strategies from the optimal control perspective. Problems of optimal control are usually solved via the solution of the Hamilton-Jacobi-Bellman (HJB) partial differentiation equation [23,24]. However, the HJB equation is generally not solvable analytically due to its inherent nonlinearity [25,26]. Therefore, adaptive dynamic programming (ADP) and reinforcement learning (RL) algorithms were proposed to obtain numerical solutions to the HJB equation and were widely applied to nonlinear interconnected systems [27-30]. In [31,32], the two previously mentioned algorithms could be deemed closely related, as they exhibited similar characteristics in addressing optimal control problems. For example, in [27,28], the distributed optimal controller was designed using robust ADP for nonlinear interconnected systems with unknown dynamics and parameters. In [29], the optimal decentralized control problem for interconnected nonlinear systems subject to stochastic dynamics was solved by enhancing the performance function of the auxiliary subsystem and transforming the original control problem into a set of optimal control strategies sampled in periodic patterns. Furthermore, in [30], the identifier-critic network framework was used to solve the problem of decentralized event-triggered control based on sliding-mode surfaces, avoiding the need for knowledge of the system's internal dynamics. It is worth noting that the control results provided in [27–30] did not consider input constraints.

Control constraints are commonly encountered in industrial processes, where they are widespread and have a detrimental impact on the performance of systems [33,34]. Therefore, the study of constrained nonlinear systems is of practical importance. In [35,36], the RL-based decentralized algorithm was developed for tracking control of constrained interconnected nonlinear systems. In [37], the problem of decentralized optimal control of a constrained interconnected nonlinear system was solved by introducing a nonquadratic performance function to overcome the symmetric input constraint. The results in [35–37], mentioned above, mainly addressed the symmetric input constraint. However, the problem of asymmetric input constraints was identified in several project cases [38,39]. In [40], the optimal decentralized control problem with asymmetric input constraints was solved by designing a new non-quadratic performance function. In [41], a new performance function was proposed for interconnected nonlinear systems to successfully overcome the asymmetric input constraint and to solve the decentralized fault-tolerant control problem. However, none of the above studies considered the safety of the system. The optimal decentralized safety control (DSC) for constrained interconnected nonlinear safety-critical systems has not been thoroughly investigated thus far, which inspired our current study.

Motivated by previous discussions, this paper proposes an RL-based decentralized DSC strategy for constrained interconnected nonlinear safety-critical systems. The primary achievements are concluded below:

1. The reinforcement learning algorithm is used to solve the optimal DSC problem for restricted interconnected nonlinear safety-critical systems, and the asymmetric

input constraint is successfully solved. The method optimizes the control strategy by minimizing the performance function, ensuring the safety of the system's state, while considering the asymmetric input constraints.

- 2. Nonlinear interconnected safety-critical systems with asymmetric input constraints and safety constraints are converted to equivalent systems that satisfy user-defined safety constraints using barrier functions. Unlike the nonlinear safety-critical systems [3,9,10,13], this paper solves the security constraint problem of the interconnection term through the potential barrier function, which ensures the interconnected nonlinear safety-critical system satisfies the security constraint.
- 3. The asymmetric input constraints are solved by utilizing a single CNN architecture for online approximation of the performance function. Theoretical demonstrations show that the optimal DSC method can achieve uniformly ultimately bounded (UUB) system states and neural network weight estimation errors. In addition, a simulation example verified the feasibility and effectiveness of the developed DSC method.

The remainder of this article is structured as follows. In Section 2, the issue formulation and conversion are presented. In Section 3, the decentralized optimal safety DSC design scheme is presented. The design scheme for the critical neural network is presented in Section 4. In Section 5, the analyses of system stability are presented. In Section 6, the simulation sample demonstrates the effectiveness of the presented approach. Lastly, conclusions are given in Section 7.

2. Preliminaries

2.1. Problem Descriptions

Consider a constrained interconnected nonlinear safety-critical system composed of n subsystems and the formula below:

$$\begin{cases} \dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(t) + \triangle h_i(x(t)), \\ x_i(0) = x_{i0}, i = 1, 2, \dots, n, \end{cases}$$
(1)

where $x_i(t) \in \mathbb{R}^{n_i}$ is the *i*th subsystem's state vector and $x_i(0)$ represents the initial state, $x = [x_1^T, x_2^T, \dots, x_n^T] \in \mathbb{R}^{\sum_{i=1}^n n_i}$ represents the overall state vector of the constrained interconnected nonlinear safety-critical system, $u_i = [u_{i,1}, u_{i,2}, \dots, u_{i,j}]^T \in \mathfrak{t}_i$ represents the control input, and the set of asymmetric constraints is represented as $\mathfrak{t}_i = \{u_{i,m_i} \in \mathbb{R}^{m_i}, \mathfrak{h}_{i\,min} \leq |u_{i,j}| \leq \mathfrak{h}_{i\,max}, \mathfrak{j} = 1, 2, \dots, m_i\}$ with $\mathfrak{h}_{i\,min}$ and $\mathfrak{h}_{i\,max}$ being the asymmetric saturating minimum and maximum bounds, $f_i(\cdot) \in \mathbb{R}^{n_i}$ and $g_i(\cdot) \in \mathbb{R}^{n_i \times m_i}$ represent the drift system dynamics and input dynamics of the *i*th subsystem, respectively, and are Lipschitz continuous, and $\Delta h_i \in \mathbb{R}^{n_i}$ represents the unknown interconnected term.

To simplify the design of the controller, let us introduce some assumptions. For i = 1, 2, ..., n, we suppose the equilibrium of the *i*th subsystem's state is $x_i = 0$.

Assumption 1. For i = 1, 2, ..., n, the $\triangle h_i(x)$ satisfies the below unmatched condition:

$$riangle h_i = \eta_i(x_i)\mathcal{P}_i(x),$$

where $\eta_i(x_i)$ is a known function with $\eta_i(x_i) \in \mathbb{R}^{n_i \times q_i} \neq g_i(x_i)$, and $\mathcal{P}_i(x)$ is a bounded vector function that satisfies

$$\|\mathcal{P}_i(x)\| \le \sum_{j=1}^n b_{i,j}\beta_{i,j}(x_j),\tag{2}$$

where $b_{i,j} > 0$ is a constant, and $\beta_{i,j}(x_j)$ are normal definite functions. Furthermore, $\beta_{i,j}(0) = 0$ and $\mathcal{P}_i(0) = 0$. Then, assuming $\beta_j(x_j) = \max_{1 \le i \le n} \{\beta_{i,j}(x_j)\}$, the unequal Equation (2) is denoted as:

$$\|\mathcal{P}_i(x)\| \le \sum_{j=1}^n C_{i,j}\beta_j(x_j),\tag{3}$$

where $C_{i,j} \ge (b_{i,j}\beta_{i,j}(x_j)) / \beta_j(x_j)$ is a positive constant, and j = 1, 2, ..., n.

Remark 1. It is noted that constraints (2) and (3) specified by Assumption 1 are strict restrictions on specific interrelated nonlinear systems. Nevertheless, when we consider the function $\mathcal{P}_i(x)$ that satisfies no constraints (2) and (3), we discover that the calculational costs to address the stability of the closed-loop system are high. In fact, in real-world applications, constraints like inequalities (2) and (3) impose on the mismatched interconnection terms of the system (1) [40,42].

Assumption 2. For i = 1, 2, ..., n, the known function $g_i(x_i)$ is bounded as $||g_i(x_i)|| \le g_{i,m}$, where $g_{i,m}$ is a known constant. Furthermore, $rank(g_i(x_i)) = m_i$ and $g_i^T(x_i)\eta_i(x_i) = 0$.

Based on the *i*th subsystem (1) described, the *i*th auxiliary subsystem is designed as:

$$\dot{x}_i = f_i(x_i) + g_i(x_i)u_i + (I_{n_i} - g_i(x_i)g_i^+(x_i))\eta_i(x_i)v_i,$$
(4)

where $v_i \in \mathbb{R}^{q_i}$ is used to compensate for mismatched interconnections and stands for auxiliary control, $g_i^+(x_i) \in \mathbb{R}^{m_i \times n_i}$ is Moore–Penrose pseudo-reverse. According to Assumption 2, it can be found that the matrix $g_i^+(x_i) = (g_i^T(x_i)g_i(x_i))^{-1}g_i^T(x_i)$ and $g_i^+(x_i)\eta_i(x_i) = (g_i^T(x_i)g_i(x_i))^{-1}g_i^T(x_i)\eta_i(x_i) = 0$. Then, we rewrite the auxiliary subsystem (4) as:

$$\dot{x}_{i} = f_{i}(x_{i}) + g_{i}(x_{i})u_{i} + \eta_{i}(x_{i})v_{i}.$$
(5)

2.2. Security Conversion Issues

For the *i*th subsystem in the system (1), its state $x_i = [x_{i,1}, x_{i,2}, ..., x_{i,k}]^T$ satisfies the following security constraints:

$$\begin{cases} x_{i,1} \in (a_{i,1}, A_{i,1}), \\ x_{i,2} \in (a_{i,2}, A_{i,2}), \\ \vdots \\ \vdots \\ x_{i,k} \in (a_{i,k}, A_{i,k}). \end{cases}$$
(6)

For nonlinear interconnect safety-critical systems with asymmetric input constraints and security constraints, we need to define the performance function as:

$$J_i(x_i) = \int_t^\infty e^{-\alpha_i(\tau-t)} (\iota_i + \Theta(x_i, u_i, v_i)) d\tau,$$
(7)

where α_i is the discount factor, $\iota_i(x_i) = h_i \beta_j^2(x_i)$ and $\Theta(x_i, u_i, v_i) = x_i^T H_i x_i + \mathfrak{W}_i(u_i) + \xi_i v_i^T v_i$ with H_i and $\mathfrak{W}_i(u_i)$ are positive definite functions, where h_i and ξ_i are positive design parameters.

Remark 2. Due to accounting for safety constraints and asymmetric input constraints in (7), the optimal control law does not converge to zero while the system state achieves the stable phase [43]. The discount factor $\alpha_i = 0$, $J_i(x_i)$ may be unbounded, so it is necessary to consider the discount factor.

Problem 1. (Decentralized control problems with security constraints and asymmetric input constraints) Consider the safety-critical system (1) and find the policy $u_i(.)$ and auxiliary control strategy $v_i(.) : \mathbb{R}^{n_i} \to \mathbb{R}^{m_i}$ in the ith subsystem. The performance function is given by (7) with the ith subsystem state $x_i = [x_{i,1}, \ldots, x_{i,k}]^T$ and the control input u_i satisfying the following conditions:

$$u_{i,\min} \le u_{i,j} \le u_{i,\max}, \ |u_{i,\min}| \ne |u_{i,\max}|, \tag{8}$$

$$x_{i,k} \in (a_{i,k}, A_{i,k}), \ \forall k = 1, \dots, n_i.$$

$$(9)$$

Ensure that the security-critical system state is consistently within the security constraints. Further, the definitions of some barrier functions are given.

Definition 1 (Barrier function [9,10]). *The function* $B(\cdot) : \mathbb{R} \to \mathbb{R}$ *defined on interval (a, A) is referred to as the barrier function if*

$$B(z;a,A) = \log \frac{A(a-z)}{a(A-z)}, \quad \forall z \in (a,A),$$
(10)

where *a* and *A* are two constants satisfying a < A. Moreover, the potential function is invertible on the interval (a, A), i.e.,

$$B^{-1}(y;a,A) = \frac{aA\left(e^{\frac{y}{2}} - e^{-\frac{y}{2}}\right)}{ae^{\frac{y}{2}} - Ae^{-\frac{y}{2}}}, \quad \forall y \in \mathbb{R}.$$
 (11)

Furthermore, the derivative of (11) is

$$\frac{\mathrm{d}B^{-1}(y;a,A)}{\mathrm{d}y} = \frac{Aa^2 - aA^2}{a^2e^y - 2aA + A^2e^{-y}}.$$
 (12)

Based on Definition 1, we consider the state transition based on the potential barrier function as follows:

$$s_{i,k} = B(x_{i,k}; a_{i,k}, A_{i,k}),$$
 (13)

$$x_{i,k} = B^{-1}(s_{i,k}; a_{i,k}, A_{i,k}),$$
(14)

where $k = 1, 2, ..., n_i$. So, the $x_{i,k}$'s derivative concerning t is $\frac{dx_{i,k}}{dt} = \frac{dx_{i,k}}{ds_{i,k}} \frac{ds_{i,k}}{dt}$, and after using Definition 1, we obtain:

$$\begin{split} \dot{s}_{i,k} &= \frac{a_{i,k+1}A_{i,k+1}\left(e^{\frac{s_{i,k+1}}{2}} - e^{-\frac{s_{i,k+1}}{2}}\right)}{a_{i,k+1}e^{\frac{s_{i,k+1}}{2}} - A_{i,k+1}e^{-\frac{s_{i,k+1}}{2}}} \times \frac{A_{i,k}^2e^{-s_{i,k}} - 2a_{i,k}A_{i,k} + a_{i,k}^2e^{s_{i,k}}}{A_{i,k}a_{i,k}^2 - a_{i,k}A_{i,k}^2} \\ &= F_{ik}(s_{i,k}, s_{i,k+1}), \ k = 1, \ \dots, \ n_i - 1, \\ \dot{s}_{i,n_i} &= \dot{x}_i \times \frac{A_{i,k}^2e^{-s_{i,k}} - 2a_{i,k}A_{i,k} + a_{i,k}^2e^{s_{i,k}}}{A_{i,k}a_{i,k}^2 - a_{i,k}A_{i,k}^2} \\ &= F_{i,n_i}(s_{i,n_i}) + G_{i,n_i}(s_{i,n_i})u_{i,n_i} + Y_{i,n_i}(s_{i,n_i}), \end{split}$$

where

$$F_{i,n_i}(s_i) = \frac{a_{i,n_i}^2 e_{i,n_i}^{s_i} - 2a_{i,n_i} A_{i,n} + A_{i,n_i} e^{-s_{i,n_i}}}{A_{i,n_i} a_{i,n_i}^2 - a_{i,n_i} A_{i,n_i}^2} \times f_i([B_{i,1}^{-1}(s_{i,1}) \dots B_{i,n_i}^{-1}(s_{i,n_i})]),$$

$$G_{i,n_i}(s_i) = \frac{a_{i,n_i}^2 e_{i,n_i}^{s_i} - 2a_{i,n_i} A_{i,n_i} + A_{i,n_i} e^{-s_{i,n_i}}}{A_{i,n_i}^2 - a_{i,n_i} A_{i,n_i}^2} \times g_i([B_{i,1}^{-1}(s_{i,1}) \dots B_{i,n_i}^{-1}(s_{i,n_i})]),$$

and $Y_{i,n_i}(s_{i,n_i})$ is the interconnection term of the n_i th term in the *i*th subsystem.

Then, the interconnected nonlinear safety-critical system (1) can be rewritten as:

$$\dot{s}_i = F_i(s_i) + G_i(s_i)u_i(t) + Y_i(s_i),$$
(15)

where $F_i(s_i) = [F_{i1}(s_{i,1}, s_{i,2}), ..., F_{i,n_i}(s_i)]^T$, $G_i(s_i) = [0, ..., G_{i,n_i}(s_i)]^T$ and $Y_i(s_i)$ is the unknown interconnected term.

Based on Assumption 1, we define the unknown interconnection term after the system transformation as:

$$Y_i(s_i) = \wp_i(s_i) U_i(s), \tag{16}$$

where $\wp_i(s_i) = [\wp_{1,n_1}(s_1), 0, ..., 0]^T$, and

$$\wp_{1,n_1}(s_1) = \frac{a_{i,2}^2 e^{s_{i,2}} - 2a_{i,2}A_{i,2} + A_{i,2}e^{-s_{i,2}}}{A_{i,2}a_{i,2}^2 - a_2A_{i,2}^2} \times \eta_{1,n_1}(x_1),$$

and $U_i(s_i)$ is a bounded vector function that satisfies

$$\|U_i(s)\| \le \sum_{j=1}^n b_{i,j}\vartheta_{i,j}(s_j),$$
 (17)

where $\vartheta_{i,j}(s_j)$ is a positive definite function. Then, assuming $\vartheta_j(s_j) = \max_{1 \le i \le n} \{\vartheta_{i,j}(s_j)\}$ and $\vartheta_j(s_j) = [\vartheta_{j,1}(s_{j,1}, s_{j,2}), \dots, \vartheta_{j,n_i}(s_j)]^T$, where

$$\vartheta_{j,n_i}(s_j) = \frac{a_{j,n_i}^2 e_{j,n_i}^{s_j} - 2a_{j,n_i} A_{j,n_i} + A_{j,n_i} e^{-s_{j,n_i}}}{A_{j,n_i} a_{j,n_i}^2 - a_{j,n_j} A_{j,n_i}^2} \times \beta_j([B_{j,1}^{-1}(x_1) \dots B_{j,n_i}^{-1}(x_j)]).$$
(18)

According to (3) and (18), the inequality (17) is expressed as:

$$\|U_i(s)\| \le \sum_{j=1}^n S_{i,j}\vartheta_j(s_j),\tag{19}$$

where $S_{i,j} \ge (b_{i,j}\vartheta_{i,j}(s_j))/\vartheta_j(s_j)$ is a positive constant, and i, j = 1, 2, ..., n.

Assumption 3. $F_i(s_i)$ is Lipschitz continuous with $F_i(0) = 0$, $\mathcal{P}_i(0) = 0$, $G_i(s_i)$ and $\wp_i(s_i)$ are upper-bounded, then $||F_i(s_i)|| \leq f_{i,m_i}||s_i||$, $||G_i(s_i)|| \leq g_{i,m_i}$, and $||\wp_i(s_i)|| \leq \eta_{i,m_i}$, $||U_i(s_i)|| \leq \mathcal{P}_{i,m_i}||s_i||$, where f_{i,m_i} , g_{i,m_i} , η_{i,m_i} , \mathcal{P}_{i,m_i} are positive constants. rank $(G_i(s_i)) = m_i$ and $G_i^T(s_i)\wp_i(s_i) = 0$. Moreover, the modified system (15) is within the manageable range, and $s_i = 0$ is the balance point for (15).

Lemma 1 ([32]). $\forall (s_1, s_2) \in \mathbb{R}^2$, we have the following condition,

$$s_1s_2 \leq \frac{\varepsilon_1^{p_1}}{p_1}|s_1|^{p_1} + \frac{1}{p_2\varepsilon_1^{p_2}}|s_2|^{p_2},$$

where $\varepsilon_1 > 0$, $(p_1 - 1)(p_2 - 1) = 1$ and $p_1, p_2 > 1$.

Remark 3. The barrier function in Definition 1, which has the following characteristics, ensures that the safety-critical system (15) always satisfies the safety constraints [9,10].

1. The state s_i of the system is restricted to be bounded, so the system state x_i satisfies constraints (8) and (9), *i.e.*,

$$|B(z_i;a_i,A_i)| < +\infty, \quad \forall z_i \in (a_i,A_i)$$

2. When the system's state approaches the boundary of the safety area, the barrier function changes as follows:

$$\lim_{z_i\to a_i^+} B(z_i;a_i,A_i)=-\infty,$$

3. The barrier function fails to function when the system state reaches equilibrium, i.e.,

$$B(0; a_i, A_i) = 0, \qquad \forall a_i < A_i.$$

3. Decentralized Optimal DSC Design

This section consists of two main subsections to establish the decentralized optimal DSC method. First, the security constraint problem is dealt with through the systematic transformation of the barrier function and the HJB equation for the *i*th auxiliary subsystem without security constraints is developed by introducing the improved performance function. Finally, the decentralized safety controller is constructed by solving the HJB equation for the auxiliary subsystem.

3.1. Barrier Function Conversion

According to the *i*th subsystem (15) described, the *i*th auxiliary subsystem is designed as:

$$\dot{s}_{i} = F_{i}(s_{i}) + G_{i}(s_{i})u_{i} + (I_{n_{i}} - G_{i}(s_{i})G_{i}^{+}(s_{i}))\wp_{i}(s_{i})v_{i},$$
(20)

where $G_i^+(s_i) \in \mathbb{R}^{m_i \times n_i}$ is Moore–Penrose pseudo-reverse. According to Assumptions 2 and 3, the matrix if found to be $G_i^+(s_i) = (G_i^T(s_i)G_i(s_i))^{-1}G_i^T(s_i)$ and $G_i^+(s_i)\wp_i(s_i) = (G_i^T(s_i)G_i(s_i))^{-1}G_i^T(s_i)\wp_i(s_i) = 0$. Then, the auxiliary subsystem (20) is rewritten as:

$$\dot{s}_i = F_i(s_i) + G_i(s_i)u_i + \wp_i(s_i)v_i.$$
 (21)

Regarding the converted system (15), analogous to (7), the performance function below is introduced:

$$V_i(s_i) = \int_t^\infty e^{-\alpha_i(\tau-t)} (\pi_i + \gamma(s_i, u_i, v_i)) d\tau,$$
(22)

where $\pi_i(s_i) = h_i \vartheta_j^2(s_i)$ and $\gamma(s_i, u_i, v_i) = s_i^T Q_i s_i + \mathfrak{W}_i(u_i) + \xi_i v_i^T v_i$, Q_i is the positive definition matrix. Furthermore, $s_{i0} = s_i(0)$ denotes the initial state, and $\mathfrak{W}_i(u_i)$ is a non-quadratic utility function that solves the asymmetric input constraint. Then, $\mathfrak{W}_i(u_i)$ is defined in the following form:

$$\mathfrak{W}_{i}(u_{i}) = \sum_{j=1}^{m_{i}} 2\lambda_{i} \int_{c_{i}}^{u_{i,j}} \Psi^{-1}((v_{i} - c_{i})/\lambda_{i}) \mathrm{d}v_{i},$$
(23)

where $\lambda_i = (\mathfrak{h}_{i \max} - \mathfrak{h}_{i \min})/2$ and $c_i = (\mathfrak{h}_{i \max} + \mathfrak{h}_{i \min})/2$, and $\Psi_i(.)$ represent the monotonic odd function, where $\Psi_i(0) = 0$. In this paper, without sacrificing generality, $\Psi_i(s_i) = (e^{s_i} - e^{-s_i})/(e^{s_i} + e^{-s_i})$.

Remark 4. Unlike the traditional form of symmetric input constraints [35], this article considered asymmetric constraints on the controlling inputs [44]. The revised hyperbolic tangent function presented in (22) effectively transforms the asymmetric constrained control problem into an unconstrained control problem by devising different maximum and minimum bounds.

Problem 2. (Optimal decentralized control problems with asymmetric input constraints) Finding the control policy u_i and auxiliary control strategy v_i in the ith subsystem, the performance function becomes (22).

Based on the subsystem (21), as well as the performance function (22), the corresponding Hamiltonian is given by:

$$H(s_{i}, u_{i}, v_{i}, \nabla V_{i}(s_{i})) = (\nabla V_{i}(s_{i}))^{T} (F_{i}(s_{i}) + G_{i}(s_{i})u_{i}(t) + \wp_{i}(s_{i})v_{i})$$

$$+\pi_i + \gamma(s_i, u_i, v_i) - \alpha_i V_i, \tag{24}$$

with $\nabla V_i(s_i) = \frac{\partial V_i(s_i)}{\partial s_i}$. The optimal performance function is

$$V_i^*(s_i) = \min_{u_i, v_i \in \Psi(\Omega_i)} V_i(s_i),$$
(25)

where $\Psi(\Omega_i)$ is a collection of all acceptable control policies and auxiliary control strategies for Ω_i .

Based on Bellman's optimality principle [31], $V_i^*(s_i)$ in (25) satisfies the HJB

$$\min_{u_i,v_i\in\Psi(\Omega_i)} H(s_i, u_i, v_i, \nabla V_i^*(s_i)) = 0,$$
(26)

where $\nabla V_i^*(s_i) = \frac{\partial V_i^*(s_i)}{\partial s_i}$. Then, the optimal control policy and the auxiliary control policy can be derived as follows:

$$u_i^*(s_i) = -\lambda_i \tanh(\frac{1}{2\lambda_i}G_i^T(s_i)\nabla V_i^*(s_i)) + c_i,$$
(27)

$$v_{i}^{*}(s_{i}) = -\frac{1}{2\xi_{i}} \wp_{i}^{T} \nabla V_{i}^{*}(s_{i}),$$
(28)

where $c_i = [c_1, ..., c_{m_i}]$.

Substituting $u_i^*(s_i)$ and $v_i^*(s_i)$ into (26), the HJB equation is rewritten as:

$$(\nabla V_i^*(s_i))^T F_i(s_i) + (\nabla V_i^*(s_i))^T G_i(s_i) u_i^*(s_i) - \xi_i \|v_i^*(s_i)\|^2 -\alpha_i V_i^* + \pi_i(s_i) + s_i^T Q_i s_i + \mathfrak{W}_i(u_i^*(s_i)) = 0,$$
(29)

with $V_i^*(0) = 0$.

Through the BF-based system transformation, the decentralized control problem 1 with asymmetric input constraints and security constraints is transformed into an unconstrained optimization problem, i.e., the decentralized control problem 2. Next, the following lemma is discussed to ensure the equivalence between the decentralized control problems 1 and 2.

Lemma 2. Assume that Assumptions 1 to 3 are met and that control policy $u_i(\cdot)$ and auxiliary control strategy $v_i(\cdot)$ solve the decentralized control problem 2 of (21). It follows, then, that the below holds:

- If the initial state x_0 of the interconnected nonlinear safety-critical system (1) is in the range 1. $(a_{i,k}, A_{i,k}), \forall k = 1, 2, ..., n_i$, then the closed-loop system satisfies (6).
- 2. If the functions $H_i(x)$ and $Q_i(x)$ satisfies the condition $H_i(x_i) = Q_i(B_i(x_i)) = Q_i(s_i)$, the performance described in (22) is equivalent to the one in (7).

Proof. Both the performance function and Assumption 3 satisfy the observability of zero states, guaranteeing the presence of the safety-optimal performance function $V_i^*(s_i)$. From (24), we obtain $\nabla V_i^*(t) \leq 0$, which allows us to obtain $V_i^*(s_i(t)) \leq V_i^*(s_i(0))$ for all $t \geq 0$. Consequently, as stated in Remark 3, if the initial state $x_i(0)$ of the system (21) satisfies the security constraint (6), and $V_i^*(s_i(0))$ is bounded, then the $V_i^*(s_i(t))$ is also bounded. Finally, we obtain

$$x_{i,k}(t) \in (a_{i,k}, A_{i,k}), \ k = 1, 2, \dots, n_i.$$
 (30)

Therefore, the given u_i^* and v_i^* satisfy the constraints of the decentralized control problem 1.

Now, consider the state transition based on the barrier function described in (13) and Since x_i satisfies the constraints given in (8), each element of the state (14). $s_i = [B_{i,1}(x_{i,1}), \dots, B_{i,k}(x_{i,k})]^T$ is finite. By comparing the performance functions (7) and (22), the equivalence relation $J_i(x_i(0)) = V_i(s_i(0))$ is obtained, provided that $H_i(x_i) = Q_i(s_i)$. This completes the proof. \Box

3.2. Designing the Optimal DSC Strategy by Solving n HJB Equations

Throughout this section, we show that the optimal DSC strategies for interconnected nonlinear systems can be constructed by solving the n HJB equations.

Theorem 1. Consider *n* subsystems under Assumptions 1 to 3 with DSC policies $u_i^*(s_i)$ and auxiliary control strategies $v_i^*(s_i)$, having the corresponding conditions as below:

$$\|v_i^*(s_i)\|^2 < s_i^T Q_i s_i , t \ge t_0.$$
(31)

Next, consider n positive constants h_i^* , i = 1, 2, ..., n, so that for anything $h_i \ge h_i^*$, the optimal DSC policies $u_1^*(s_1)$, $u_2^*(s_2)$, ..., $u_n^*(s_n)$ guarantee that the interconnected nonlinear system (15) with security constraints is UUB.

Proof. The Lyapunov candidacy function $L_{i,1}(s)$ below was selected:

$$L_{i,1}(s) = \sum_{i=1}^{n} V_i^*(s_i),$$
(32)

where the $V_i^*(s_i)$ is defined in the same way as (22), and we denote the time derivative along the trajectory $\dot{s_i} = F_i(s_i) + G_i(s_i)u_i(t) + Y_i(s_i)$ as:

$$\dot{L}_{i,1}(s) = \sum_{i=1}^{n} (\nabla V_i^*)^T (G_i(s_i)u_i^* + F_i(s_i) + Y_i(s)).$$
(33)

By using (27) and (28), we obtain:

$$(\nabla V_i^*(s_i))^T G_i(s_i) = -2\lambda_i \tanh^{-T}\left(\frac{u_i^* - c_i}{\lambda_i}\right),\tag{34}$$

$$\nabla V_i^*(s_i))^T \wp_i(s_i) = -2\xi_i(v_i^*(s_i))^T.$$
(35)

 $(\nabla V_i^*(s_i))^T \wp_i(s_i) = -2$ Inserting (29), (34) and (35) into (33), we have

$$\dot{L}_{i,1}(s) = \sum_{i=1}^{n} [\alpha_i V_i^* - \pi_i(s_i) - s_i^T Q_i s_i - \mathfrak{W}_i(u_i^*) + \xi_i \|v_i^*(s_i)\|^2 - 2\xi_i (v_i^*(s_i))^T U_i(s)].$$
(36)

According to the optimal DSC policy (27), the term $\mathfrak{W}_i(u_i^*)$ becomes

$$\mathfrak{W}_{i}(u_{i}^{*}(s_{i})) = 2\lambda_{i} \sum_{j=1}^{m_{j}} \int_{0}^{u_{i,j}^{*}-c_{i}} \tanh^{-1}(\frac{u_{i}-c_{i}}{\lambda_{i}})d(u_{i}-c_{i}).$$
(37)

By appealing to the proof in [44], Equation (37) can be further reduced to

$$\mathfrak{W}_{i}(u_{i}^{*}(s_{i})) = \underbrace{\lambda_{i}^{2} \sum_{i=1}^{m_{i}} (\tanh^{-1}(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}}))}_{\beta_{1}}}_{\beta_{1}} - \underbrace{2\lambda_{i}^{2} \sum_{j=1}^{m_{i}} \int_{0}^{\tanh^{-1}(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}})} (u_{i} - c_{i}) \tanh^{2}(u_{i} - c_{i})d(u_{i} - c_{i})}_{\beta_{2}}}_{\beta_{2}}$$
(38)

replacing (38) into (36), one has

$$\dot{L}_{i,1}(s) \leq -\sum_{i=1}^{n} (2\xi_{i}(s_{i}^{T}Q_{i}s_{i} - \|v_{i}^{*}(s_{i})\|^{2})) - \sum_{i=1}^{n} (1 - 2\xi_{i})(s_{i}^{T}Q_{i}s_{i}) - \sum_{i=1}^{n} (\pi_{i}(s_{i})) - 2\xi_{i}\sum_{j=1}^{m_{i}} \|v_{i}^{*}(s_{i})\|b_{i,j}\vartheta_{i,j}(s_{j}) + \xi^{2}\|v_{i}^{*}(s_{i})\|^{2}) + \alpha_{i}V_{i}^{*} - \beta_{1} + \beta_{2}.$$
(39)

It is known from [45] that there is a positive constant $\delta_{i,M}$ such that $0 \le ||\nabla V_i^*(s_i)|| \le \delta_{i,M}$. Therefore, using Lemma 1, Assumption 1, (17), (19), and (27), we obtain

$$2\beta_{1} \leq 2\lambda_{i}^{2} \tanh^{-T}\left(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}}\right) \tanh^{-1}\left(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}}\right) \\ = \frac{1}{2} (\nabla V_{i}^{*}(s_{i}))^{T} G_{i}(s_{i}) G_{i}^{T}(s_{i}) (\nabla V_{i}^{*}(s_{i})) \\ \leq \frac{1}{2} G_{i,m}^{2} \delta_{i,m'}^{2}$$
(40)

Utilizing the integral median theorem [46] and the inequality (40), the β_2 (38) can be deduced as:

$$\beta_{2} = 2\lambda_{i}^{2} \sum_{j=1}^{m_{i}} \tanh^{-1}\left(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}}\right) \varpi_{i} \tanh^{-2} \varpi_{i}$$

$$\leq 2\lambda_{i}^{2} \sum_{j=1}^{m_{i}} \tanh^{-1}\left(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}}\right) \varpi_{i}$$

$$\leq 2\lambda_{i}^{2} \tanh^{-T}\left(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}}\right) \tanh^{-1}\left(\frac{u_{i,j}^{*} - c_{i}}{\lambda_{i}}\right)$$

$$\leq \frac{1}{2}G_{i,m}^{2}\delta_{i,m'}^{2}, \qquad (41)$$

where $\omega_i \in (0, \tanh^{-1}(\frac{u_{i,j}^* - c_i}{\lambda_i}))$.

From [27], we conclude that $\|\alpha_i V_i^*(s_i)\| \le \varrho_{i,m}$, where $\varrho_{i,m}$ is a positive constant. Then, plugging (40) and (41) into (39), and taking into consideration the conclusion mentioned above, we can rephrase inequality (39) as follows:

$$\dot{L}_{i,1}(s) \leq -\sum_{i=1}^{n} (2\xi_i (s_i^T Q_i s_i - \|v_i^*(s_i)\|^2)) - \sum_{i=1}^{n} (1 - 2\xi_i) (s_i^T Q_i s_i) -\sum_{i=1}^{n} (h_i \vartheta_i (s_j)^2 - 2\xi_i \sum_{j=1}^{m_i} \|v_i^*(s_i)\| b_{i,j} \vartheta_{i,j}(s_j) + \xi^2 \|v_i^*(s_i)\|^2) + \varrho_i + \frac{1}{4} \sum_{i=1}^{n} G_{i,m}^2 \delta_{i,m}^2,$$
(42)

by denoting $\Lambda = diag\{h_1, h_2, \dots, h_n\}$ and $Z = [\vartheta_1(s_1), \dots, \vartheta_n(s_n), \xi_1 || v_1^*(s_1) ||, \dots, \xi_n || v_n^*(s_n) ||]$. Let the condition (31) be satisfied, so we have

$$\dot{L}_{i,1}(s) \le -\sum_{i=1}^{n} (1 - 2\xi_i) (s_i^T Q_i s_i) - Z^T X Z + \varrho_i + \frac{1}{4} \sum_{i=1}^{n} G_{i,m}^2 \delta_{i,m}^2,$$
(43)

with $X = \begin{bmatrix} \Lambda & A^T \\ A & I_n \end{bmatrix}$ and $A = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}$.

From the matrix *X* expression, positive definiteness is maintained by choosing a sufficiently large Λ . In other words, there is $h_i^* > 0$, such that $h_i > h_i^*$, ensuring $Z^T X Z > 0$. Thus, the inequality (43) is further deduced as:

$$\dot{L}_{i,1}(s) \leq -\sum_{i=1}^{n} (1 - 2\xi_i) \lambda_{min}(Q_i) \|s_i\|^2 + \varrho_i + \frac{1}{4} \sum_{i=1}^{n} G_{i,m}^2 \delta_{i,m}^2.$$
(44)

The inequality (44) means that $\dot{L}_{i,1}(s) < 0$ whenever $s_i(t)$ lies outside the following set N_{s_i} :

$$N_{s_i} = \left\{ s_i : \|s_i\| \le \sqrt{\frac{\frac{1}{4}G_{i,M}^2 \delta_{i,M}^2 + \varrho_i}{\lambda_{min}(Q_i)(1 - 2\xi_i)}} \right\}.$$
(45)

Based on Lyapunov's extension theorem [47], it is shown that the optimal performance functions $V_i^*(s_i)$ guarantee that the interconnected nonlinear system (15) with asymmetric input constraints is UUB. Since the performance function (7) and (22) yield the same results, it can be shown that the optimal performance function $J_i^*(x_i)$ guarantees that the interconnected nonlinear safety-critical system (1) with security constraints and asymmetric input constraints is UUB.

4. Critic Network for Approximation

The critic neural network is introduced in this section, with the aim of approximating the optimal performance function. Then, the evaluation network of the auxiliary subsystem (21) is used to construct the estimated optimal control strategy. According to [48], $V_i^*(s_i)$ is expressed as:

$$V_i^*(s_i) = W_{\mathfrak{c}_i}^I \sigma_{\mathfrak{c}_i}(s_i) + \varepsilon_{\mathfrak{c}_i}(s_i), \tag{46}$$

where $\sigma_{\mathfrak{c}_i}(s_i) = [\sigma_{\mathfrak{c}_i,1}(s_i), \sigma_{\mathfrak{c}_i,2}(s_i), \ldots, \sigma_{\mathfrak{c}_i,N_i}(s_i)] \in \mathbb{R}^{N_i}$ denotes the activation function, $W_{\mathfrak{c}_i} \in \mathbb{R}^{N_i}$ denotes the ideal weight vector, N_i denotes the number of neurons, and $\varepsilon_{\mathfrak{c}_i}(s_i) \in \mathbb{R}^{N_i}$ is the reconstruction error of NN. The vector activation function $\sigma_{\mathfrak{c}_i,p}(s_i)$ is denoted as a continuously differentiable function, where $p = 1, 2, \ldots, N_i$. For $s_i \neq 0$, $\{\sigma_{\mathfrak{c}_i,p}(s_i)\}_{p=1}^{N_i}$ is linearly independent. Then, the derivative of $V_i^*(s_i)$ can be expressed as:

$$\nabla V_i^*(s_i) = \nabla \sigma_{\mathfrak{c}_i}^T(s_i) W_{\mathfrak{c}_i} + \nabla \varepsilon_{\mathfrak{c}_i}(s_i), \tag{47}$$

where $\nabla \sigma_{\mathfrak{c}_i}(s_i) = \frac{\partial \sigma_{\mathfrak{c}_i}(s_i)}{\partial s_i}$ and $\nabla \varepsilon_{\mathfrak{c}_i}(s_i) = \frac{\partial \varepsilon_{\mathfrak{c}_i}(s_i)}{\partial s_i}$.

From Equations (27), (28) and (47), the optimal safety control policy $u_i^*(s_i)$ and the auxiliary control strategy $v_i^*(s_i)$ are rephrased as:

$$u_i^*(s_i) = -\lambda_i \tanh(\frac{1}{2\lambda_i} G_i^T(s_i) \nabla \sigma_{\mathbf{c}_i}^T(s_i) W_{\mathbf{c}_i}) + c_{d_i} + \varepsilon_{u_i}(s_i),$$
(48)

$$v_i^*(s_i) = -\frac{1}{2\xi_i} \wp_i^T(s_i) \nabla \sigma_{\mathfrak{c}_i}^T(s_i) W_{\mathfrak{c}_i} + \varepsilon_{v_i}(s_i),$$
(49)

where

$$\begin{split} \varepsilon_{u_i}(s_i) &= -\frac{1}{2}(I_{m_i} - \tanh^2(\zeta))G_i^T(s_i)\nabla\varepsilon_{\mathfrak{c}_i}(s_i),\\ \varepsilon_{v_i}(s_i) &= -\frac{1}{2\xi_i}\wp_i^T(s_i)\nabla\varepsilon_{\mathfrak{c}_i}(s_i), \end{split}$$

with $I_{m_i} = [1, 1, ..., 1]^T \in \mathbb{R}^{m_i}$. The seclected value of ζ is between $\frac{1}{2\lambda_i}G_i^T(s_i)\nabla\sigma_{\mathfrak{c}_i}^T(s_i)W_{\mathfrak{c}_i}$ and $\frac{1}{2\lambda_i}G_i^T(s_i)(\nabla\sigma_{\mathfrak{c}_i}^T(s_i)W_{\mathfrak{c}_i} + \nabla\varepsilon_{\mathfrak{c}_i}(s_i))$. The ideal weight vector W_{c_i} is not available and the optimal control strategy $u_i^*(s_i)$ is not directly applicable. Therefore, the estimated weight vector \hat{W}_{c_i} is constructed to replace W_{c_i} as:

$$\hat{V}_i^*(s_i) = \hat{W}_{\mathbf{c}_i}^T \sigma_{\mathbf{c}_i}(s_i).$$
⁽⁵⁰⁾

The estimation error $\tilde{W}_{c_i} = W_{c_i} - \hat{W}_{c_i}$ is defined. Similarly, according to (50), the (49) and (48) are further developed as:

$$\hat{u}_i(s_i) = -\lambda_i \tanh(\frac{1}{2\lambda_i} G_i^T(s_i) \nabla \sigma_{\mathbf{c}_i}^T(s_i) \hat{W}_{\mathbf{c}_i}) + c_{d_i},$$
(51)

$$\hat{v}_i(s_i) = -\frac{1}{2\xi_i} \wp_i^T(s_i) \nabla \sigma_{\mathfrak{c}_i}^T(s_i) \hat{W}_{\mathfrak{c}_i}.$$
(52)

Combining (50), (51) and (52), the Hamiltonian is re-expressed as:

$$H(s_{i}, \hat{u}_{i}, \hat{v}_{i}, \nabla \hat{V}_{i}(s_{i})) = (\nabla \hat{V}_{i}(s_{i}))^{T} (G_{i}^{T}(s_{i})\hat{u}_{i} + F_{i}(s_{i}) + \wp_{i}(s_{i})\hat{v}_{i}) + \pi_{i}(s_{i}) + \gamma_{i}(s_{i}, \hat{u}_{i}, \hat{v}_{i}) - \alpha_{i}\hat{V}_{i}.$$
(53)

According to (53), the error of the Hamiltonian is given by:

$$e_{i} = H(s_{i}, \hat{u}_{i}, \hat{v}_{i}, \nabla \hat{V}_{i}(s_{i})) - H(s_{i}, u_{i}^{*}, v_{i}^{*}, \nabla V_{i}^{*}(s_{i}))$$

= $\pi_{i}(s_{i}) + s_{i}^{T}Q_{i}s_{i} + \mathfrak{W}_{i}(\hat{u}_{i}) + \xi_{i}\hat{v}_{i}^{T}\hat{v}_{i} + \hat{W}_{c_{i}}^{T}\varrho_{i},$ (54)

with $\varrho_i = \nabla \sigma_{c_i}(x_i)(G_i^T(s_i)\hat{u}_i + F_i(s_i) + \wp_i(s_i)\hat{v}_i) - \alpha_i\sigma_{c_i}(s_i)$. In order to make $u_i(s_i) \rightarrow u_i^*(s_i)$, the error e_i should be guaranteed to be sufficiently small. To solve this issue, a critic weight adjustment law \hat{W}_{c_i} is proposed to minimize the objective function $\phi_i = \frac{1}{2}e_i^T e_i$. Next, the critic updating law is developed as:

$$\hat{W}_{\mathbf{c}_i} = -\frac{\alpha_{\mathbf{c}_i} \varrho_i e_i}{(1 + \varrho_i^T \varrho_i)^2},\tag{55}$$

where the constant α_{c_i} is the positive learning rate.

Remark 5. To minimize the Hamiltonian error e_i , it is necessary to maintain the derivative of ϕ_i as $\dot{\phi}_i < 0$. Therefore, the critic weight adjustment law is derived by employing the normalization term $(1 + \varrho_i^T \varrho_i)^{-2}$ and applying the gradient descent method with respect to \hat{W}_{c_i} [49].

By considering the definition of \tilde{W}_{c_i} , we obtain

$$\dot{\tilde{W}}_{\mathfrak{c}_{i}} = -\alpha_{\mathfrak{c}_{i}}\ell_{i}\ell_{i}^{T}\tilde{W}_{\mathfrak{c}_{i}} + \frac{\alpha_{\mathfrak{c}_{i}}\ell_{i}e_{H_{i}}}{\iota_{i}},\tag{56}$$

where $\ell_i = \frac{\varrho_i}{1+\varrho_i^T \varrho_i}$ and $\iota_i = 1 + \varrho_i^T \varrho_i$. e_{H_i} denotes the residual error, defined as $e_{H_i} = \nabla \sigma_{\mathbf{c}_i}(x_i)(G_i^T(s_i)\hat{u}_i + F_i(s_i) + \wp_i(s_i)\hat{v}_i)$.

The proposed decentralized DSC strategy for the ith subsystem with a single critic-NN is illustrated in Figure 1.



Figure 1. The block diagram of the developed optimal DSC scheme.

5. Stability Analysis

This section focuses on the stability of the *n*-auxiliary subsystem for the given control scheme. We need to make some Assumptions to satisfy the theorem.

Assumption 4. For $s_i \in \Omega_i$, i = 1, ..., n, there exist some positive constants $D_{\varepsilon_{u_i}}, \eta_{i,M}, D_{\sigma_{\varepsilon_i}}, D_{\varepsilon_{v_i}}$ and $D_{e_{H_i}}$ satisfying $\|\varepsilon_{u_i}(s_i)\| \leq D_{\varepsilon_{u_i}}, \|\wp_i(s_i)\| \leq \wp_{i,M}, \|\nabla \sigma_{\varepsilon_i}(s_i)\| \leq D_{\sigma_{\varepsilon_i}}, \|\varepsilon_{v_i}(s_i)\| \leq D_{\varepsilon_{v_i}}$ and $\|e_{H_i}\| \leq D_{e_{H_i}}$.

Assumption 5. Consider the time period $[t, t + t_{\mathfrak{k}}]$ and $t_{\mathfrak{k}} > 0$. Then, the term $\ell_i \ell_i^T$ fulfills the following condition:

$$\epsilon_i I_{N_i} \le \ell_i \ell_i^1 \le \mathfrak{s}_i I_{N_i},\tag{57}$$

where ϵ_i and \mathfrak{s}_i are positive constants.

Theorem 2. For the nonlinear interconnected safety-critical system (15), we design the estimated optimal safety policies and auxiliary control strategies as (51) and (52), respectively. Assume that Assumptions 1–5 hold. If \hat{W}_{c_i} is updated by (55), then s_i and \hat{W}_{c_i} are UUB if α_{c_i} in (55) satisfies

$$\alpha_{\mathfrak{c}_{i}} > \frac{\wp_{i,M}^{2} D_{\sigma_{\mathfrak{c}_{i}}}^{2}}{\xi_{i} \lambda_{min}(\ell_{i} \ell_{i}^{T})}.$$
(58)

Proof. The candidate Lyapunov function is considered to be:

$$L_{i}(t) = \sum_{i=1}^{n} (V_{i}^{*}(s_{i}) + \frac{1}{2} \tilde{W}_{\mathfrak{c}_{i}}^{T} \tilde{W}_{\mathfrak{c}_{i}}).$$
(59)

Then, defining $L_{i,1}(t) = V_i^*(s_i)$ and $L_{i,2}(t) = \frac{1}{2} \tilde{W}_{c_i}^T \tilde{W}_{c_i}$, the time derivative by $L_{i,1}(t)$ is

$$\dot{L}_{i,1}(t) = (\nabla V_i^*(s_i))^T (G_i^T(s_i)\hat{u}_i + F_i(s_i) + \wp_i(s_i)\hat{v}_i)
= (\nabla V_i^*(s_i))^T (G_i^T(s_i)u_i^* + F_i(s_i) + \wp_i(s_i)v_i^*)
+ \underbrace{(\nabla V_i^*(s_i))^T G_i^T(s_i)(\hat{u}_i - u_i^*)}_{\beta_3} + \underbrace{(\nabla V_i^*(s_i))^T \wp_i(s_i)(\hat{v}_i - v_i^*)}_{\beta_4}.$$
(60)

Combining (29), (34) and (35). The (60) is further deduced as:

$$\dot{L}_{i,1}(t) = \alpha_i V_i^* - \pi_i(s_i) - s_i^T Q_i s_i - \mathfrak{W}_i(u_i^*) + \xi_i \|v_i^*(s_i)\|^2 + \beta_3 + \beta_4.$$
(61)

According to Lemma 1, and taking into account (40), (48), (51), we observe that the β_3 term in (61) is satisfied by

$$\beta_{3} \leq \lambda_{i}^{2} \left\| \tanh^{-1}\left(\frac{u_{i}^{*}(s_{i}) - c_{d_{i}}}{\lambda_{i}}\right) \right\| + \left\| \hat{u}_{i} - u_{i}^{*} \right\|^{2}$$

$$\leq \beta_{1} + \underbrace{\left\| \lambda_{i}(\tanh(\mathfrak{Y}_{i,1}(s_{i})) - \tanh(\mathfrak{Y}_{i,2}(s_{i}))) - \varepsilon_{u_{i}}(s_{i}) \right\|^{2}}_{\beta_{5}}$$

$$\leq \frac{1}{4} G_{i,M}^{2} \delta_{i,M}^{2} + \beta_{5}, \qquad (62)$$

where $\mathfrak{Y}_i(s_i) = \frac{1}{2\lambda_i} G_i^T(s_i) \nabla V_i^*(s_i)$. Then, based on the fact $\| \tanh(\mathfrak{Y}_{i,\mathfrak{k}}(s_i)) \| \leq \sqrt{m_i}$, $\mathfrak{k} = 1, 2$ in [44], according to Assumption 5, β_5 is derived as:

$$\beta_{5} \leq 2\lambda_{i}^{2} \| \tanh(\mathfrak{Y}_{i,1}(s_{i})) - \tanh(\mathfrak{Y}_{i,2}(s_{i})) \|^{2} + 2 \| \varepsilon_{u_{i}}(s_{i}) \|^{2}$$

$$\leq 4\lambda_{i}^{2} (\| \tanh(\mathfrak{Y}_{i,1}(s_{i})) \|^{2} + \| \tanh(\mathfrak{Y}_{i,2}(s_{i})) \|^{2}) + 2 \| \varepsilon_{u_{i}}(s_{i}) \|^{2}$$

$$\leq 8\lambda_{i}^{2} m_{i} + 2D_{\varepsilon_{u_{i}}}^{2}.$$
(63)

Similarly, the last term of (61) is deduced from (35), (49) and (52) as:

$$\begin{aligned} \beta_{4} &\leq -\xi_{i} \|v_{i}^{*}\|^{2} - \xi_{i} \|\hat{v}_{i} - v_{i}^{*}\|^{2} \\ &\leq -\xi_{i} \|v_{i}^{*}\|^{2} + 2\xi_{i} (\|\hat{v}_{i}\|^{2} - \|v_{i}^{*}\|^{2}) + 2\xi_{i} \|\varepsilon_{v_{i}}\|^{2} \\ &\leq -\xi_{i} \|v_{i}^{*}\|^{2} + \frac{1}{2\xi_{i}} \wp_{i,M}^{2} D_{\sigma_{\varepsilon_{i}}}^{2} \|\tilde{W}_{\varepsilon_{i}}\|^{2} + 2\xi_{i} D_{\varepsilon_{v_{i}}}^{2}. \end{aligned}$$

$$(64)$$

By using (38), (62)–(64) and the fact that $\|\alpha_i V_i^*(s_i)\| \le \varrho_{i,M}$ the following is derived:

$$\dot{L}_{i,1}(t) \le -\lambda_{\min}(Q_i) \|s_i\|^2 + \frac{1}{2\xi_i} \wp_{i,M}^2 D_{\sigma_{\mathfrak{c}_i}}^2 \|\tilde{W}_{\mathfrak{c}_i}\|^2 + \Theta_i,$$
(65)

with $\Theta_i = \varrho_{i,M} + \frac{1}{2}G_{i,M}^2\delta_{i,M}^2 + 8\lambda_i^2m_i + 2D_{\varepsilon_{u_i}}^2 + 2\xi_i D_{\varepsilon_{v_i}}^2$. The error weight update law \tilde{W}_{c_i} . $L_{i,1}(t)$ is considered with the time derivative

$$\dot{L}_{i,2}(t) = -\alpha_{\mathfrak{c}_i} \tilde{W}_{\mathfrak{c}_i}^T \ell_i \ell_i^T \tilde{W}_{\mathfrak{c}_i} + \alpha_{\mathfrak{c}_i} \frac{\tilde{W}_{\mathfrak{c}_i}^T \ell_i}{\iota_i} e_{H_i}.$$
(66)

Combining Lemma 1 and Assumption 4, the following conclusion is drawn:

$$\alpha_{\mathfrak{c}_{i}} \frac{\tilde{W}_{\mathfrak{c}_{i}}^{T} \ell_{i}}{\iota_{i}} e_{H_{i}} \leq \frac{\alpha_{\mathfrak{c}_{i}}}{2} \tilde{W}_{\mathfrak{c}_{i}}^{T} \ell_{i} \ell_{i}^{T} \tilde{W}_{\mathfrak{c}_{i}} + \frac{\alpha_{\mathfrak{c}_{i}}}{2} D_{e_{H_{i}}}^{2}.$$

$$(67)$$

Combining inequalities (66) and (67), we derive the following inequalities:

$$\dot{L}_{i,2}(t) \le -\frac{\alpha_{\mathfrak{c}_i}}{2} \lambda_{\min}(\ell_i \ell_i^T) \left\| \tilde{W}_{\mathfrak{c}_i} \right\|^2 + \frac{\alpha_{\mathfrak{c}_i}}{2} D_{e_{H_i}}^2.$$
(68)

Substituting (65) and (68) into (59), the following inequality is obtained:

$$\dot{L}_{i}(t) \leq \sum_{i=1}^{n} (-\lambda_{min}(Q_{i}) \|s_{i}\|^{2} - \mathfrak{r}_{i} \|\tilde{W}_{\mathfrak{c}_{i}}\|^{2} + \Theta_{i} + \frac{\alpha_{\mathfrak{c}_{i}}}{2} D_{e_{H_{i}}}^{2}),$$
(69)

where $\mathfrak{x}_i = \frac{\alpha_{\mathfrak{c}_i}}{2} \lambda_{min}(\ell_i \ell_i^T) - \frac{1}{2\xi_i} \wp_{i,M}^2 D_{\sigma_{\mathfrak{c}_i}}^2$, $\lambda_{min}(\ell_i \ell_i^T)$ means the minimum eigenvalue of $\ell_i \ell_i^T$.

Therefore, Equations (58) and (69) mean $\dot{L}_i(t) < 0$, provided that the parameters s_i and \tilde{W}_{c_i} are not in the set of

$$N_i \left\{ s_i : \|s_i\| \le \sqrt{\frac{2\Theta_i + D_{e_{H_i}}^2}{2\lambda_{min}(Q_i)}} \right\},\tag{70}$$

$$N_{\tilde{W}_{\mathfrak{c}_{i}}}\left\{\tilde{W}_{\mathfrak{c}_{i}}:\left\|\tilde{W}_{\mathfrak{c}_{i}}\right\|\leq\sqrt{\frac{2\Theta_{i}+D_{e_{H_{i}}}^{2}}{\mathfrak{x}_{i}}}\right\}.$$
(71)

Introducing Lyapunov's extension theorem, ref. [47], ensures the stability of the closed-loop system. This proof ensures that the weight estimation error \tilde{W}_{c_i} is UUB. At this point, this completes the proof process.

Remark 6. In contrast to techniques that aim to achieve input saturation [10,13], this article proposes an RL technique to solve the optimal DSC problem with safety constraints and asymmetric input constraints. This approach ensures not only the safety of the system but also minimizes the input constraints. Therefore, the developed reinforcement learning technique, based on security constraints and asymmetric input constraints, is better suited for some project applications, particularly for systems where the system state must be globally within the security settings.

6. Simulation Example

In this section, we provide a simulation example to verify the effectiveness of the proposed approach. The simulation involved a dual-linked robotic arm system [42]. The state space model of the system is defined by

$$\begin{split} \dot{x}_{1,1} &= x_{1,2}, \\ \dot{x}_{1,2} &= -\frac{M_1}{\tilde{G}_1} x_{1,2} - \frac{m_1 \tilde{g} \tilde{l}_1}{\tilde{G}_1} \sin(x_{1,1}) + \frac{1}{\tilde{G}_1} u_1 + \triangle h_1, \\ \dot{x}_{2,1} &= x_{2,2}, \\ \dot{x}_{2,2} &= -\frac{M_2}{\tilde{G}_2} x_{2,2} - \frac{m_2 \tilde{g} \tilde{l}_2}{\tilde{G}_2} \sin(x_{2,1}) + \frac{1}{\tilde{G}_2} u_2 + \triangle h_2, \end{split}$$
(72)

where $x_{i,1}$ and $x_{i,2}$ (i = 1, 2) indicate the angular location of the robot arm, u_i stands for control input, and the $\triangle h_i = \eta_i \mathcal{P}_i$ represents the interconnection terms. The other parameters of the robotic arm system (72) are depicted in Table 1. The initial system state was selected as $x_0 = [2, 2, 2, 2]^T$. We first defined the state variable $x_i = [x_{i,1}, x_{i,2}]^T$ and constructed the internal dynamics and input gain matrix as follows:

$$f_i(x_i) = \begin{bmatrix} x_{i,2} \\ -\frac{M_i}{\tilde{G}_i} x_{i,2} - \frac{m_i \tilde{g}\tilde{l}_i}{\tilde{G}_i} \sin(x_{i,1}) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{\tilde{G}_i} \end{bmatrix} u_i + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathcal{P}_i(x_i),$$

where $\mathcal{P}_1(x_1)$, $\mathcal{P}_2(x_2)$ denote the uncertain interconnection terms of subsystems 1 and 2, i.e.,

$$\mathcal{P}_1(x_1) = 0.1x_{1,1}\sin(x_{2,2}),$$

$$\mathcal{P}_2(x_2) = (x_{1,2} - 3\sin(0.1x_{2,1})).$$

Furthermore, the two robotic arm subsystems were in a state that satisfied the below security constraints:

$$\begin{aligned} x_{1,1} &\in (-0.5, 2.9), \quad x_{1,2} \in (-1.5, 2.5), \\ x_{2,1} &\in (-1, 2.5), \quad x_{2,2} \in (-3.5, 3). \end{aligned}$$

Therefore, to deal with the security constraint, the following system of transformations without security constraint was obtained, using the BF-based system transformation (13):

s. .

$$s_i = F_i(s_i) + G_i(s_i)u_i + \wp_i(s_i)U_i,$$
(74)

where

$$F_{i}(s_{i}) = \begin{bmatrix} \frac{a_{i,2}A_{i,2}(e^{\frac{s_{i,2}}{2}} - e^{-\frac{s_{i,2}}{2}})}{a_{i,1}e^{\frac{s_{i,1}}{2}} - A_{i,2}e^{-\frac{s_{i,2}}{2}}} & \frac{a_{i,1}e^{s_{i,1}} - 2a_{i,1}A_{i,1} + A_{i,1}e^{-s_{i,1}}}{A_{i,1}a_{i,1}^{2} - a_{1}A_{i,1}^{2}} \\ f_{i}(B^{-1}(s_{i})) \frac{a_{i,2}^{2}e^{\frac{s_{i,2}}{2}} - 2a_{i,2}A_{i,2} + A_{i,2}e^{-s_{i,2}}}{A_{i,2}a_{i,2}^{2} - a_{i,2}A_{i,2}^{2}} \end{bmatrix},$$

$$G_{i}(s_{i}) = \begin{bmatrix} 0 \\ \frac{1}{\tilde{G}} & \frac{a_{i,2}^{2}e^{s_{i,2}} - 2a_{i,2}A_{i,2} + A_{i,2}e^{-s_{i,2}}}{A_{i,2}a_{i,2}^{2} - a_{2}A_{i,2}^{2}} \end{bmatrix},$$

$$\wp_{i}(s_{i}) = \begin{bmatrix} \frac{a_{i,2}^{2}e^{s_{i,2}} - 2a_{i,2}A_{i,2} + A_{i,2}e^{-s_{i,2}}}{A_{i,2}a_{i,2}^{2} - a_{2}A_{i,2}^{2}} \end{bmatrix},$$

$$(75)$$

Table	1 . l	Meanings	and va	lues of	symbols	s used i	in robotic	arm sy	stems.

The <i>i</i> th Subsystem	Parameter	Meaning	Value
The first subsystem	$ \begin{array}{c} m_1 \\ M_1 \\ \tilde{l}_1 \\ \tilde{G}_1 \\ \tilde{g}_1 \end{array} $	Mass of payload Viscous friction Length of the arm Moment of inertia Acceleration of gravity	5 kg 2 N 0.5 m 10 kg 9.81 m/s
The second subsystem	m_2 M_2 \tilde{l}_2 \tilde{G}_2 \tilde{g}_2	Mass of payload Viscous friction Length of the arm Moment of inertia Acceleration of gravity	10 kg 2 N 1 m 10 kg 9.81 m/s

For the transformed dual-linked robotic arm system (74), the initial state was chosen by $s_{i,0} = [s_{i,0}(1), s_{i,0}(2)]^T = [B(x_{i,0}(1); a_{i,1}, A_{i,1}), B(x_{i,0}(2); a_{i,2}, A_{i,2})]^T$. The discount factors were chosen as $\alpha_1 = 1$ and $\alpha_2 = 0.1$. The matrices were designed as $Q_1 = 0.5I^2$ and $Q_2 = I^2$, $R_1 = 1$ and $R_2 = 1$. The upper and lower limits were allocated as below: $\mathfrak{h}_{1max} = 0.75$, $\mathfrak{h}_{1min} = -0.25$ and $\mathfrak{h}_{2max} = 1.5$, $\mathfrak{h}_{2min} = -0.5$. Let $\vartheta_1 = ||s_1||$ and $\vartheta_2 = ||s_2||$. Additional design factors were setup as below: $\xi_1 = 8$, $\xi_2 = 4$, $a_{\mathfrak{c}_1} = 2$, $a_{\mathfrak{c}_2} = 2$. Choose the activation functions $\sigma_{\mathfrak{c}_i}(s_i) = [s_{1,1}^2, s_{1,1}s_{1,2}, s_{1,2}^2]^T$ and $\sigma_{\mathfrak{c}_i}(s_i) = [s_{2,1}^2, s_{2,1}s_{2,2}, s_{2,2}^2]^T$.

The simulation outcomes are presented in Figures 2–13. The states of the system are depicted in Figures 2 and 8, and it can be observed that the closed-loop system stabilized after 20 s and 35 s, respectively. However, the system failed to meet the specified security constraints. Figures 3 and 9, shown in comparison with Figures 2 and 8, not only assured that the system states converged to zero, but also satisfied the given safety constraints. The evolving states $s_1(t)$ and $s_2(t)$ are presented in Figures 4 and 10, based on the safe control method with asymmetric input constraints. The optimal DSC policies are shown in Figures 5 and 11. We found that the optimal DSC policies were restricted to the asymmetric set [-0.25, 0.75] and [-0.5, 1.5]. Figures 6 and 12 represent the optimal auxiliary control strategies for subsystems 1 and 2, respectively. Figures 7 and 13 show the critic updated laws. It can be observed that the weights converged after 15 s. According to Theorem 3, we concluded that the proposed optimal safety control policy and the auxiliary control policy could stabilize the closed-loop nonlinear system and satisfy the safety constraints on the system state. Moreover, the optimal control policy eventually converged to a predefined set of constraints. Finally, the results of the simulation showed that the presented optimal

DSC solution for constrained interconnected nonlinear safety-critical systems, affected by system state constraints, is effective.



Figure 2. Evolution of state $x_1(t)$ without using the DSC method.



Figure 3. Evolution of state $x_1(t)$ using the DSC method.



Figure 4. Evolution of state $s_1(t)$ using the DSC method.



Figure 5. Control evolution of input u_1 .



Figure 6. Evolution of the auxiliary control input v_1 using the DSC method.



Figure 7. Evolution of the critic weight vector W_{c_1} using the DSC method.



Figure 8. Evolution of state $x_2(t)$ without using the DSC method.



Figure 9. Evolution of state $x_2(t)$ using the DSC method.



Figure 10. Evolution of state $s_2(t)$ using the DSC method.



Figure 11. Control evolution of input u_2 .



Figure 12. Evolution of the auxiliary control input v_2 using the DSC method.



Figure 13. Evolution of the critic weight vector W_{c_2} using the DSC method.

7. Conclusions

This article presents an RL-based DSC scheme for interconnected nonlinear safetycritical systems with security constraints and asymmetric input constraints. The proposed method transformed an interconnected nonlinear safety-critical system with security and asymmetric input constraints into an interconnected nonlinear safety-critical system with only asymmetric input constraints by using the barrier function. The non-quadratic utility function was added to the performance function to address the asymmetric input constraint. The critic network was also used to approach the optimal performance function and to establish the best security policy. Our control scheme stabilizes the closed-loop system and minimizes the improved performance function. In addition, the simulation results demonstrated the efficacy of the proposed distributed security solution. Future work will explore the optimal safety control of stochastic interconnected nonlinear systems with event triggering.

Author Contributions: C.Q. and Y.W. provided methodology, validation, and writing—original draft preparation; T.Z. provided conceptualization, writing—review; J.Z. provided supervision; C.Q. provided funding support. All authors read and agreed to the published version of the manuscript.

Funding: This work was supported by the science and technology research project of the Henan province 222102240014.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The authors can confirm that all relevant data are included in the article.

Conflicts of Interest: The authors declare that they have no conflict of interest. All authors have approved the manuscript and agreed with submission to this journal.

References

- Son, T.D.; Nguyen, Q. Safety-critical control for non-affine nonlinear systems with application on autonomous vehicle. In Proceedings of the 2019 IEEE 58th Conference on Decision and Control (CDC), Nice, France, 11–13 December 2019; pp. 7623–7628.
- Manjunath, A.; Nguyen, Q. Safe and robust motion planning for dynamic robotics via control barrier functions. In Proceedings of the 2021 60th IEEE Conference on Decision and Control (CDC), Austin, TX, USA, 14–17 December 2021; pp. 2122–2128.
- 3. Wang, J.; Qin, C.; Qiao, X.; Zhang, D.; Zhang, Z.; Shang, Z.; Zhu, H. Constrained optimal control for nonlinear multi-input safety-critical systems with time-varying safety constraints. *Mathematics* **2022**, *10*, 2744. [CrossRef]
- 4. Liu, Z.; Yuan, Q.; Nie, G.; Tian, Y. A multi-objective model predictive control for vehicle adaptive cruise control system based on a new safe distance model. *Int. J. Automot. Technol.* **2021**, *22*, 475–487. [CrossRef]
- 5. Ames, A.D.; Xu, X.; Grizzle, J.W.; Tabuada, P. Control barrier function based quadratic programs for safety critical systems. *IEEE Trans. Autom. Control* **2016**, *62*, 3861–3876. [CrossRef]

- 6. 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]
- 7. Qin, C.; Wang, J.; Zhu, H.; Xiao, Q.; Zhang, D. Safe adaptive learning algorithm with neural network implementation for *H*_∞ control of nonlinear safety-critical system. *Int. J. Robust Nonlinear Control* **2023**, *33*, 372–391. [CrossRef]
- Srinivasan, M.; Abate, M.; Nilsson, G.; Coogan, S. Extent-compatible control barrier functions. Syst. Control Lett. 2021, 150, 104895.
 [CrossRef]
- Yang, Y.; Yin, Y.; He, W.; Vamvoudakis, K.G.; Modares, H. Safety-aware reinforcement learning framework with an actor-criticbarrier structure. In Proceedings of the 2019 American Control Conference (ACC), Philadelphia, PA, USA, 10–12 July 2019; pp. 2352–2358.
- Yang, Y.; Vamvoudakis, K.G.; Modares, H.; Yin, Y.; Wunsch, D.C. Safe intermittent reinforcement learning with static and dynamic event generators. *IEEE Trans. Neural Netw. Learn. Syst.* 2020, 31, 5441–5455. [CrossRef]
- 11. Xu, J.; Wang, J.; Rao, J.; Zhong, Y.; Wang, H. Adaptive dynamic programming for optimal control of discrete-time nonlinear system with state constraints based on control barrier function. *Int. J. Robust Nonlinear Control* 2022, *32*, 3408–3424. [CrossRef]
- 12. Brunke, L.; Greeff, M.; Hall, A.W.; Yuan, Z.; Zhou, S.; Panerati, J.; Schoellig, A.P. Safe learning in robotics: From learning-based control to safe reinforcement learning. *Annu. Rev. Control Robot. Auton. Syst.* **2022**, *5*, 411–444. [CrossRef]
- 13. 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]
- 14. Bakule, L. Decentralized control: An overview. Annu. Rev. Control. 2008, 32, 87–98. [CrossRef]
- 15. Xu, L.X.; Wang, Y.L.; Wang, X.; Peng, C. Decentralized Event-Triggered Adaptive Control for Interconnected Nonlinear Systems With Actuator Failures. *IEEE Trans. Fuzzy Syst.* 2022, *31*, 148–159. [CrossRef]
- 16. Guo, B.; Dian, S.; Zhao, T. Robust NN-based decentralized optimal tracking control for interconnected nonlinear systems via adaptive dynamic programming. *Nonlinear Dyn.* **2022**, *110*, 3429–3446. [CrossRef]
- 17. Feng, Z.; Li, R.B.; Wu, L. Adaptive decentralized control for constrained strong interconnected nonlinear systems and its application to inverted pendulum. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–11. [CrossRef]
- Zouhri, A.; Boumhidi, I. Stability analysis of interconnected complex nonlinear systems using the Lyapunov and Finsler property. *Multimed. Tools Appl.* 2021, 80, 19971–19988. [CrossRef]
- Li, X.; Zhan, Y.; Tong, S. Adaptive neural network decentralized fault-tolerant control for nonlinear interconnected fractional-order systems. *Neurocomputing* 2022, 488, 14–22. [CrossRef]
- Tan, Y.; Yuan, Y.; Xie, X.; Tian, E.; Liu, J. Observer-based event-triggered control for interval type-2 fuzzy networked system with network attacks. *IEEE Trans. Fuzzy Syst.* 2023, 1–10. [CrossRef]
- Zhang, J.; Li, S.; Ahn, C.K.; Xiang, Z. Adaptive fuzzy decentralized dynamic surface control for switched large-scale nonlinear systems with full-state constraints. *IEEE Trans. Cybern.* 2021, 52, 10761–10772. [CrossRef]
- 22. Huo, X.; Karimi, H.R.; Zhao, X.; Wang, B.; Zong, G. Adaptive-critic design for decentralized event-triggered control of constrained nonlinear interconnected systems within an identifier-critic framework. *IEEE Trans. Cybern.* **2021**, *52*, 7478–7491. [CrossRef]
- 23. Bao, C.; Wang, P.; Tang, G. Data-Driven Based Model-Free Adaptive Optimal Control Method for Hypersonic Morphing Vehicle. *IEEE Trans. Aerosp. Electron. Syst.* 2022, 1–15. [CrossRef]
- Farzanegan, B.; Suratgar, A.A.; Menhaj, M.B.; Zamani, M. Distributed optimal control for continuous-time nonaffine nonlinear interconnected systems. *Int. J. Control* 2022, 95, 3462–3476. [CrossRef]
- 25. Heydari, M.H.; Razzaghi, M. A numerical approach for a class of nonlinear optimal control problems with piecewise fractional derivative. *Chaos Solitons Fractals* **2021**, *152*, 111465. [CrossRef]
- 26. Liu, S.; Niu, B.; Zong, G.; Zhao, X.; Xu, N. Data-driven-based event-triggered optimal control of unknown nonlinear systems with input constraints. *Nonlinear Dyn.* **2022**, *109*, 891–909. [CrossRef]
- 27. Niu, B.; Liu, J.; Wang, D.; Zhao, X.; Wang, H. Adaptive decentralized asymptotic tracking control for large-scale nonlinear systems with unknown strong interconnections. *IEEE/CAA J. Autom. Sin.* **2021**, *9*, 173–186. [CrossRef]
- 28. Zhao, B.; Luo, F.; Lin, H.; Liu, D. Particle swarm optimized neural networks based local tracking control scheme of unknown nonlinear interconnected systems. *Neural Netw.* **2021**, *134*, 54–63.
- 29. Zhao, Y.; Niu, B.; Zong, G.; Xu, N.; Ahmad, A.M. Event-triggered optimal decentralized control for stochastic interconnected nonlinear systems via adaptive dynamic programming. *Neurocomputing* **2023**, *539*, 126163. [CrossRef]
- Wang, T.; Wang, H.; Xu, N.; Zhang, L.; Alharbi, K.H. Sliding-mode surface-based decentralized event-triggered control of partially unknown interconnected nonlinear systems via reinforcement learning. *Inf. Sci.* 2023, 641, 119070. [CrossRef]
- Lewis, F.L.; Vrabie, D. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits Syst. Mag.* 2009, 9, 32–50. [CrossRef]
- 32. Tang, F.; Niu, B.; Zong, G.; Zhao, X.; Xu, N. Periodic event-triggered adaptive tracking control design for nonlinear discrete-time systems via reinforcement learning. *Neural Netw.* **2022**, 154, 43–55. [CrossRef]
- 33. Sun, J.; Liu, C. Backstepping-based adaptive dynamic programming for missile-target guidance systems with state and input constraints. *J. Frankl. Inst.* **2018**, *355*, 8412–8440. [CrossRef]
- Zhao, S.; Wang, J.; Wang, H.; Xu, H. Goal representation adaptive critic design for discrete-time uncertain systems subjected to input constraints: The event-triggered case. *Neurocomputing* 2022, 492, 676–688. [CrossRef]

- 35. 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]
- Sun, H.; Hou, L. Adaptive decentralized finite-time tracking control for uncertain interconnected nonlinear systems with input quantization. *Int. J. Robust Nonlinear Control* 2021, 31, 4491–4510. [CrossRef]
- Duan, D.; Liu, C. Finite-horizon optimal tracking control for constrained-input nonlinear interconnected system using aperiodic distributed nonzero-sum games. *IET Control Theory Appl.* 2021, 15, 1199–1213. [CrossRef]
- Li, Y.; Li, Y.-X.; Tong, S. Event-based finite-time control for nonlinear multi-agent systems with asymptotic tracking. *IEEE Trans. Autom. Control* 2023, 68, 3790–3797. [CrossRef]
- 39. Zhang, H.; Zhao, X.; Zong, G.; Xu, N. Fully distributed consensus of switched heterogeneous nonlinear multi-agent systems with bouc-wen hysteresis input. *IEEE Trans. Netw. Sci. Eng.* 2022, *9*, 4198–4208. [CrossRef]
- 40. Yang, X.; Zhou, Y.; Dong, N.; Wei, Q. Adaptive critics for decentralized stabilization of constrained-input nonlinear interconnected systems. *IEEE Trans. Syst. Man Cybern. Syst.* 2021, 52, 4187–4199. [CrossRef]
- Zhao, Y.; Wang, H.; Xu, N.; Zong, G.; Zhao, X. Reinforcement learning-based decentralized fault tolerant control for constrained interconnected nonlinear systems. *Chaos Solitons Fractals* 2023, 167, 113034. [CrossRef]
- Cui, L.; Zhang, Y.; Wang, X.; Xie, X. Event-triggered distributed self-learning robust tracking control for uncertain nonlinear interconnected systems. *Appl. Math. Comput.* 2021, 395, 125871. [CrossRef]
- Tang, Y.; Yang, X. Robust tracking control with reinforcement learning for nonlinear-constrained systems. Int. J. Robust Nonlinear Control 2022, 32, 9902–9919. [CrossRef]
- 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]
- 45. Beard, R.W.; Saridis, G.N.; Wen, J.T. Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation. *Automatica* **1997**, *33*, 2159–2177. [CrossRef]
- Liu, D.; Yang, X.; Wang, D.; Wei, Q. Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints. *IEEE Trans. Cybern.* 2015, 45, 1372–1385. [CrossRef]
- Pishro, A.; Shahrokhi, M.; Sadeghi, H. Fault-tolerant adaptive fractional controller design for incommensurate fractional-order nonlinear dynamic systems subject to input and output restrictions. *Chaos Solitons Fractals* 2022, 157, 111930. [CrossRef]
- 48. Zhang, L.; Zhao, X.; Zhao, N. Real-time reachable set control for neutral singular Markov jump systems with mixed delays. *IEEE Trans. Circuits Syst. II Express Briefs* **2021**, *69*, 1367–1371. [CrossRef]
- 49. Lakmesari, S.H.; Mahmoodabadi, M.J.; Ibrahim, M.Y. Fuzzy logic and gradient descent-based optimal adaptive robust controller with inverted pendulum verification. *Chaos Solitons Fractals* **2021**, *151*, 111257. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.