Event-Triggered Adaptive Dynamic Programming of Nonlinear System with Asymmetric Input Constraints and Uncertain Disturbances

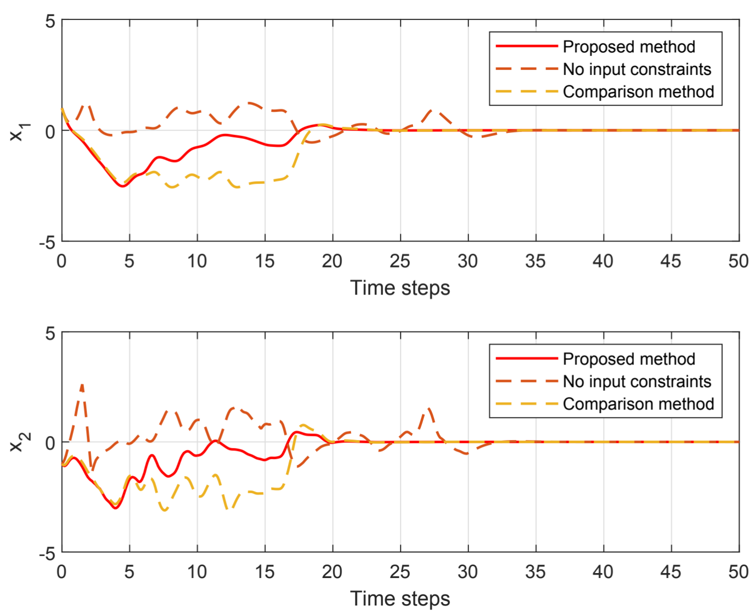
Qie Liu[1], Yan Liu[2], Guoqing Qiu[2], and Wenyao Gou[2]

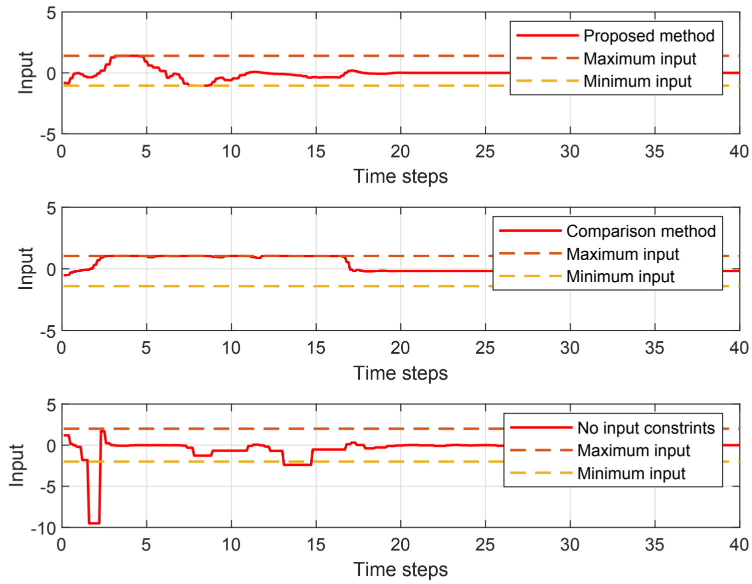
[1]Chongqing University School of Automation
[2]Chongqing University of Posts and Telecommunications

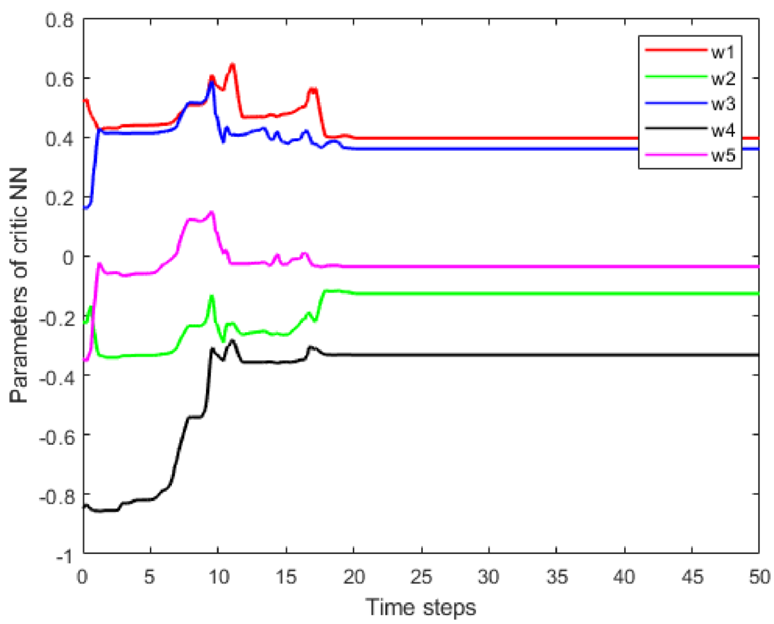
December 5, 2022

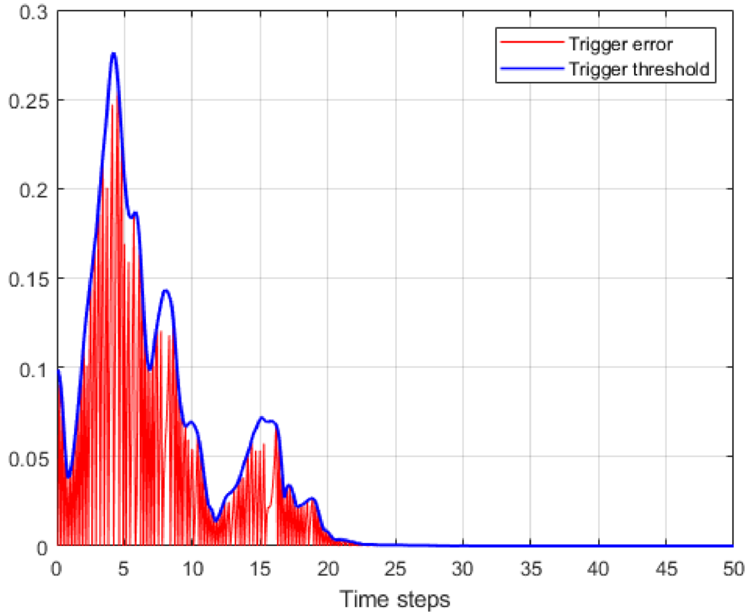
Abstract

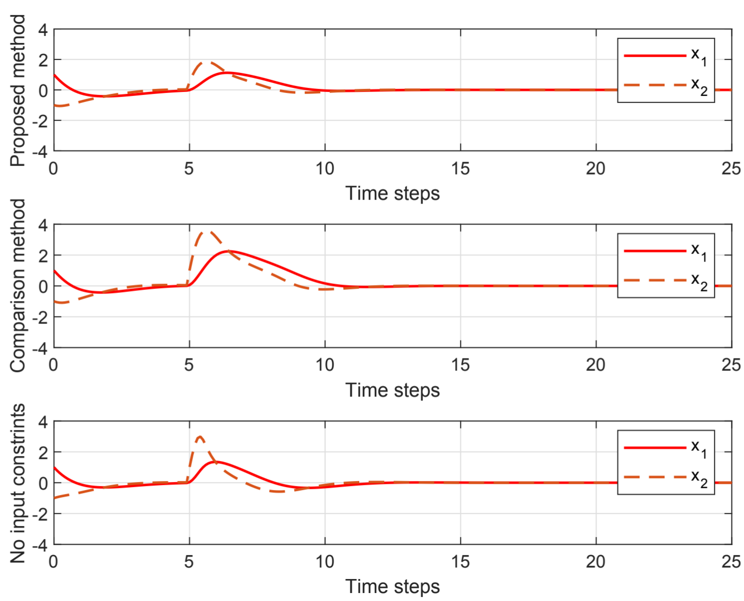
In this paper, an uncertain disturbance rejection control problem for the affine system in the presence of asymmetric input constraints is addressed using an event-triggered control method. The disturbance rejection control is converted to an H [?] optimal control problem, and a Zero-sum game-based method is proposed to solve this H [?] optimal control problem. To deal with the input constraints, a new cost function is proposed. The event-triggered controller is updated only when the triggering condition is satisfied, which can reduce the computational complexity.In order to obtain a controller that minimizes the performance index function in the worst-case disturbance, we use a critic-only network to solve the Hamilton-Jacobi-Isaacs(HJI) equation, and the critic network weight is tuned through a gradient descent method with the historical state data. The stability of the closed-loop system and the uniform ultimate boundedness of the critic network parameters are proved by the Lyapunov method. Two numerical examples are provided to verify the effectiveness of the proposed methods.

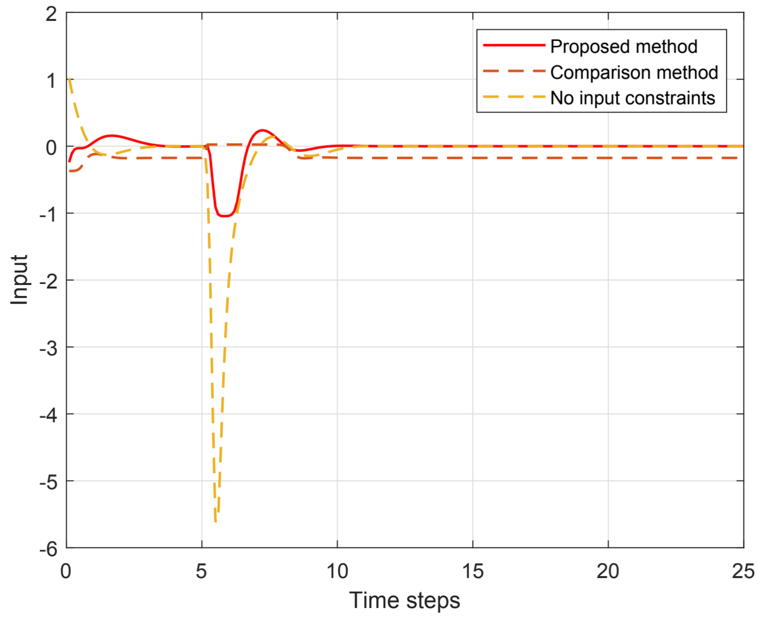


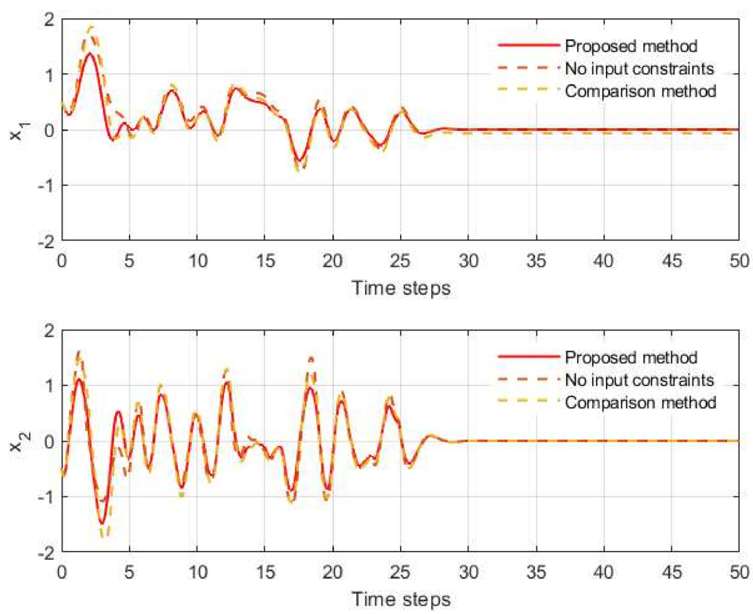

Input
Time steps
Proposed method
Maximum input
Minimum input
Comparison method
Maximum input
Minimum input
No input constrints
Maximum input
Minimum input

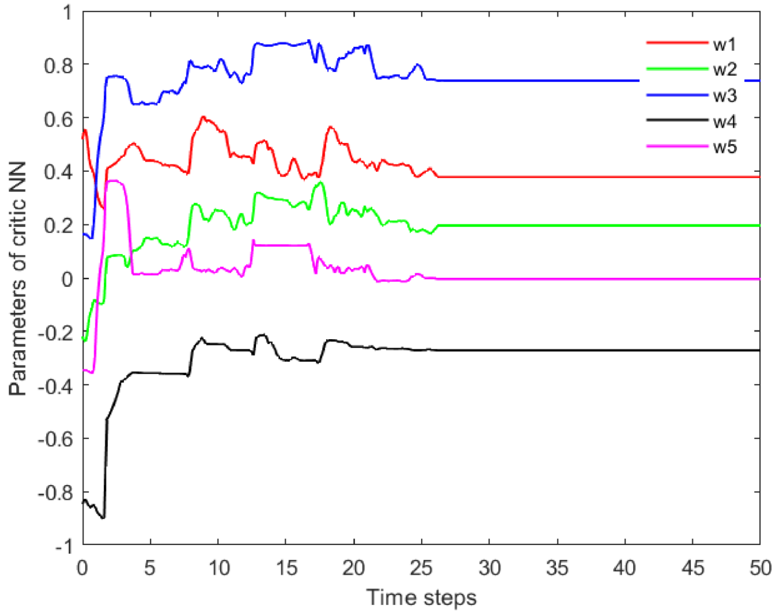

Parameters of critic NN
Time steps
w1
w2
w3
w4
w5

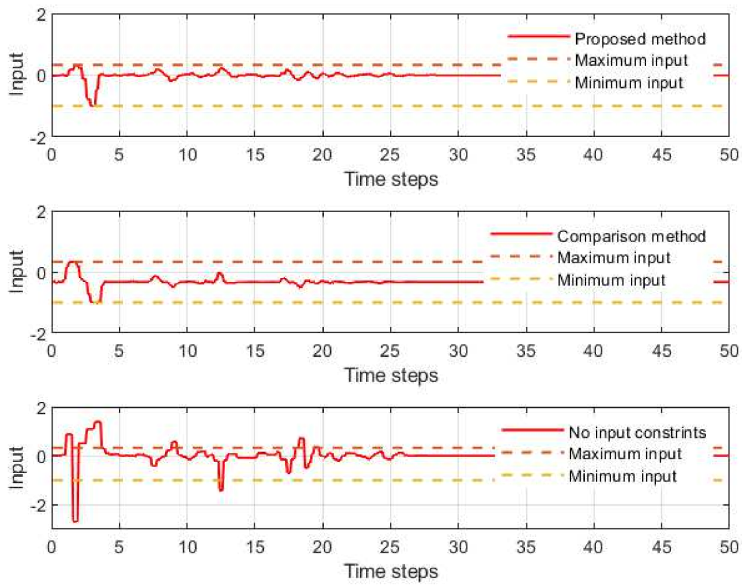

Trigger error
Trigger threshold
0.3
0.25
0.2
0.15
0.1
0.05
0
0 5 10 15 20 25 30 35 40 45 50
Time steps

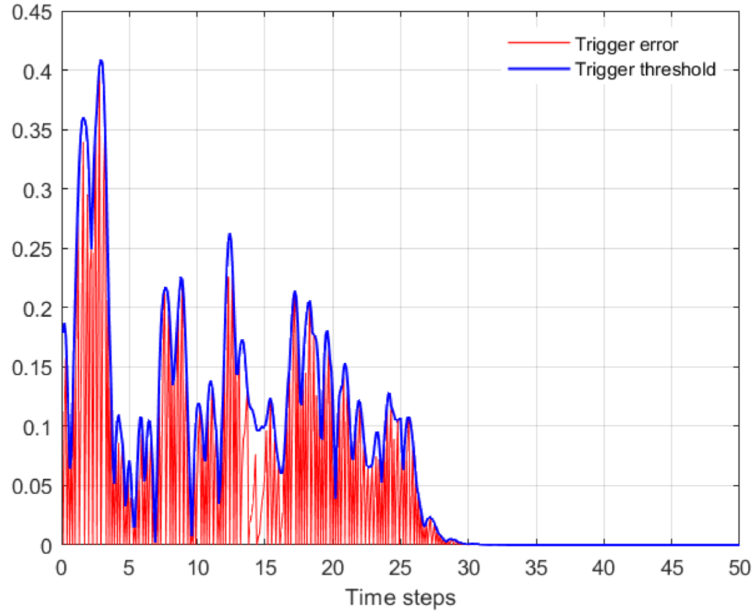

Proposed method
Comparison method
No input constrints
x_1
x_2
4
2
0
-2
-4
0 5 10 15 20 25
Time steps


Proposed method
Comparison method
No input constraints
Input
Time steps

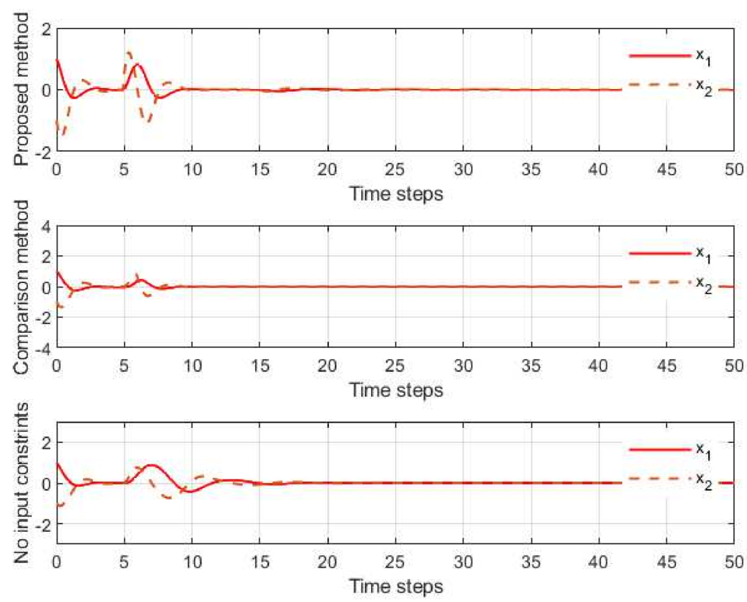

Proposed method
No input constraints
Comparison method
x_1
Time steps
Proposed method
No input constraints
Comparison method
x_2
Time steps


w1
w2
w3
w4
w5
Parameters of critic NN
Time steps

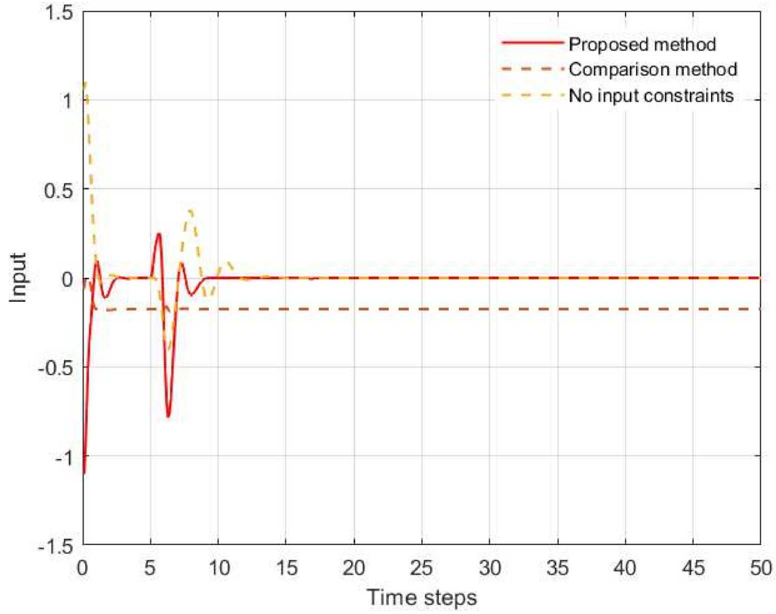

Proposed method
Maximum input
Minimum input
Input
Time steps
Comparison method
Maximum input
Minimum input
Input
Time steps
No input constrints
Maximum input
Minimum input
Input
Time steps



0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 5 10 15 20 25 30 35 40 45 50
Time steps
Trigger error
Trigger threshold



Proposed method
Comparison method
No input constrints
Time steps
x_1
x_2
0 5 10 15 20 25 30 35 40 45 50


Proposed method
Comparison method
No input constraints
1.5
1
0.5
0
-0.5
-1
-1.5
Input
0
5
10
15
20
25
30
35
40
45
50
Time steps

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

ARTICLE TYPE

Event-Triggered Adaptive Dynamic Programming of Nonlinear System with Asymmetric Input Constraints and Uncertain Disturbances

Yan Liu[1] | Guoqing Qiu[1] | Wenyao Gou[1] | Qie Liu*[2]

[1]School of Automation, Chongqing University of Posts and Telecommunications, Chongqing, China
[2]School of Automation, Chongqing University, Chongqing, China

Correspondence
*Qie Liu, Chongqing University, Chongqing 400044, China. Email: liuqiebuct@gmail.com

Present Address
School of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Summary

In this paper, an uncertain disturbance rejection control problem for the affine system in the presence of asymmetric input constraints is addressed using an event-triggered control method. The disturbance rejection control is converted to an H_∞ optimal control problem, and a Zero-sum game-based method is proposed to solve this H_∞ optimal control problem. To deal with the input constraints, a new cost function is proposed. The event-triggered controller is updated only when the triggering condition is satisfied, which can reduce the computational complexity.In order to obtain a controller that minimizes the performance index function in the worst-case disturbance, we use a critic-only network to solve the Hamilton-Jacobi-Isaacs(HJI) equation, and the critic network weight is tuned through a gradient descent method with the historical state data. The stability of the closed-loop system and the uniform ultimate boundedness of the critic network parameters are proved by the Lyapunov method. Two numerical examples are provided to verify the effectiveness of the proposed methods.

KEYWORDS:
Adaptive dynamic programming, event-triggered, input constraints, optimal control, neural networks.

1 | INTRODUCTION

Disturbance widely exists in almost all the industrial systems and has negative effects on the control performance. Hence, the disturbance rejection is almost the primary concern for all the control systems, which has attracted much attention in recent years[1]. This issue has been well studied for linear system, and many methods or strategies, such as PID[2,3], DOB[4,5], ADRC[6,7,8,9] et. al., have been proposed to deal with the disturbances. In addition to the complex external disturbances, the safety and physical restriction of the actuators, regarded as input constraints, are also commonly encountered for control system design. The existence of the input constraints limits the capability of the system, which makes it difficult for disturbance rejection, especially for the nonlinear system. Moreover, the disturbance form has a significant effect for the controller design. For the classical control strategies with fixed structure such as DOB[10,11], ADRC[12,13] et. al., only the deterministic disturbance with known form can be rejected in most cases. It is still an open problem for the rejection of the disturbances with complex form in the presence of input constraints, and much more works are needed to be done to solve this problem.

Optimal control is a main branch of modern control theory, focusing on the basic conditions and comprehensive methods to optimize the performance specification of the control systems. We focus on the optimal control for disturbance rejection in this paper. In the framework of optimal control, the controller is designed by solving an optimization problem. The adaptive dynamic programming (ADP), whose idea is to use approximate functional structures (such as neural networks (NNs)[14], fuzzy models[15,16,17], polynomials[18,19], etc.) to estimate the optimal value of the objective function, is one of the main optimization methods for optimal control. It serves as an effective way to achieve the optimal performance in a general case, and has been widely applied to controller design in recent years. Among the ADP-based optimal control, robust ADP is one of the typical methods for the disturbance rejection in the presence of input constraints[20]. In[21], an optimal control algorithm was proposed for the bounded robust controller design for finite-time-horizon nonlinear system. The algorithm was constructed by using the least squares method, while the closed-loop stability analysis was not fully discussed. In[22], a novel strategy was proposed to design the robust controller for a class of continuous-time nonlinear systems with uncertainties based on the online policy iteration algorithm. In[23], the robust optimal control of continuous-time affine nonlinear systems with matched uncertainties was investigated by using a data-driven integral policy iteration approach. However, these methods were proposed for the matched disturbances. In general, the methods for matching disturbances are not always suitable for the mismatched disturbances. In fact, the model uncertainty can be described by the mismatched disturbance. In[24], the stability analysis and fuzzy control of a class of discrete time fuzzy systems with implicit semi-Markov stochastic uncertainty are studied. The practical application are performed in[25], which researched the composite re-fined anti-disturbance rejection control problem of a two degree of freedom robot arm system modeled by a semi-Markov jump system with multiple disturbances. Therefore, it is of importance to study the control design in the presence of mismatched disturbance.

For the optimal control in the presence of input constraints, the commonly used strategy is to design a non-quadratic objective function, instead of the traditional quadratic index, where the control inputs are guaranteed to have certain bounds. The symmetric input constraints are the most commonly considered constraints in the literature, and a lot of methods are proposed to deal with such constraints. In fact, there exists many nonlinear plants subject to asymmetric input constraints, which presents challenges in deriving optimal control laws in this case, and only a few works focused on the control asymmetric input constraints.Yang and Zhao[26] proposed an optimal control scheme for continuous-time nonlinear systems with asymmetric input constraints, where the objective function with variable upper and lower bounds of integration were used to deal with the asymmetric input constraints. In[27], Yang and Wei proposed an event-triggered ADP subject to asymmetric input constraints. In[28], Kong proposed a switching function to tackle the asymmetric input constraints problem, but it was difficult to find such a switching function owing to the complexity of nonlinear system. All these results were proposed for the set-point tracking control, the disturbance rejection performance in the presence of asymmetric input constraints has not been well considered in literature.

Motivated by the limitations of the existing works, we focus on the control of the nonlinear system for disturbance rejection in the presence of input constraints. The mismatched disturbance and the asymmetric input constraint are considered here. A non-quadratic objective function for the input constraints problem is proposed. Different from the objective functions with variable upper and lower bounds of integration used in other works, a new objective function is proposed to deal with the asymmetric input constraints. In order to solve the problem of external disturbance in ADP, a Zero-sum game is proposed to establish the optimization problem, and the problem is converted to find the Nash equilibrium point by solving the Hamilton-Jacobi-Isaacs (HJI) equation. As a result, the optimal control law is obtained under the worst disturbance. To reduce the computational complexity, we use the event-triggered mechanism to determine the control update time.

The main contributions of this paper are summarized as follows:

1) In the ADP framework, a new objective function is proposed in this paper to deal with the asymmetric input constraints.

2) Considering the mismatched disturbance rejection, the control problem is transformed into a Zero-sum game problem, and the optimal control law in the presence of mismatched disturbance is obtained for the affine system by solving the HJI equation.

3) The event-triggered mechanism and a critic-only NN structure are used, then an event-triggered condition is developed which can not only ensure the stability of the system, but also ensure the convergence of neural network parameters.

The rest of this article is organized as follows. In Section 2, we introduce the performance index function under input constraints, and the event-triggered mechanism is addressed. In Section 3, we transform the disturbance rejection control problem into a Zero-sum game problem, and the HJI equation and the event-trigger condition are developed. In Section 4, we use a critic-only NN to solve the HJI equation, and the NN's weights are tuned by the gradient descent method. Section 5 gives the simulation studies of this paper, and Section 6 draws the final conclusions.

2 | PROBLEM STATEMENT

In this paper, a nonlinear continuous-time system with external disturbance is considered, which is given by,

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

where $x \in R^n$ is the n-dimensional measurable state variable, $f(x) \in R^n$, $g(x) \in R^{n\times m}$ and $k(x) \in R^{n\times p}$ denote the drift dynamics, the input dynamics and the disturbance dynamics of the system, respectively. $u \in R^m$ is the m-dimensional input variable, denoted by $\Pi_u = \{u|u \in R^m, u_{min} \le u_i \le u_{max}\}$, where u_{min} is the lower bound and u_{max} is the upper bound. The disturbance d is bounded as $||d|| \le d_m$,where d_m is the upper bound of the uncertain term. We make the following assumptions for the system[22,29].

Assumption1: System (1) is controllable.

Assumption2: $f(x) + g(x)u + k(x)d$ is Lipschitz continuous on $\Omega \in R^n$ and satisfies the following conditions,

(1) $f(0) = 0$ and $||f(x)|| \le L_f||x||$, where L_f is a Lipschitz constant and $L_f > 0$.

(2) $||g(x)|| \le g_M$ and the constant $g_M > 0$.

Assumption3: The state variable $x = 0$ is an equilibrium of the system.

Under the time-triggered mechanism, the controller is periodically updated at a fixed sampling interval, and the system state is periodically transmitted to the controller. This update method leads to a high computational cost and a high transmission cost. In order to avoid such a problem, an event-triggered ADP control is used here. In the event-triggered mechanism, we define a monotonically increasing time series $\{t_j\}_{j=0}^{\infty}$, where $t_j < t_{j+1}$ for $j = 0, 1, ..., \infty$, which is called as the trigger moment. The control law updates only at these non-periodic sampling instants, which are determined by the event triggering conditions. The event-triggered error and condition are defined as,

$$e_j(t) = \hat{x}_j - x(t), \tag{2}$$

$$e_j(t) \le e_T \tag{3}$$

where $t \in [t_j, t_{j+1}], j = 0, 1, ...\infty$ and e_T is defined as the triggering threshold. At a sampling time, if the event-triggered error $e_j(t)$ is greater than the trigger threshold e_T, the current moment is marked as the triggering moment, and the state $x(t)$ at the current moment is marked as a new triggering state $\hat{x}_j$ ($\hat{x}_j = x(t_j)$), which is transmitted to the controller to update the control law. If the trigger condition is not violated, the control law does not update, and the system updates with the last input stored in the zero-order holder(ZOH)[30]. Therefore, the control input for the event-triggered control can be expressed as,

$$u(x(t)) = \mu(\hat{x}_j). \tag{4}$$

Under the event-triggered condition, the system (1) can be rewritten as,

$$\dot{x}(t) = f(x) + g(x)\mu(x(t) + e_j(t)) + k(x)d(x) \tag{5}$$

The primary control objective in this paper is to find an event-triggered control input to minimize the following objective function,

$$V(x) = \int_t^{\infty} r(x, \mu, d)d\tau = \int_t^{\infty} x^TQx + W(\mu) - \gamma^2||d||^2d\tau \tag{6}$$

where Q is a symmetric positive definite matrix, $||d||^2 = d^Td$, γ is a prescribed positive constant, and $W(\mu)$ is defined as,

$$W(\mu) = 2\sum_{i=1}^{m}\int_0^{\mu_i} \varphi^{-1}(z_i/\bar{U})\bar{U}\,dz_i \tag{7}$$

where $\bar{U} = (u_{max} - u_{min})/2$. Considering that the input is limited, for symmetric input constraints, we always choose the hyperbolic tangent function $\varphi(\cdot) = tanh(\cdot)$, where the upper and lower bound norms are equivalent. For asymmetric input constraints, we choose a new function $\varphi(\cdot) = \frac{e^x - e^{-x}}{(1/u_{min})e^x - (1/u_{max})e^{-x}}$. Using an asymmetric function to deal with asymmetric input constraints, so as to avoid the problem that the input cannot return to zero caused by adding a fixed value to the input when using a symmetric function. Note that, when $u_{max} = 0$, let $u_{max} = \xi$; similarly when $u_{min} = 0$, let $u_{min} = \xi$, where ξ is an arbitrarily small number, and specify that $u_{max} > 0, u_{min} < 0$.

Based on the assumptions above, the objective of this paper is to find an event-triggered control law to stabilize system (5) and to minimize the performance index (6). The ADP-based method to solve this problem is proposed in the following context.

3 | EVENT-TRIGGERED OPTIMAL CONTROL

3.1 | Event-Triggered Control and the ETC-HJI equation

In this section, the optimal control problem is converted to a two-player Zero-sum game problem. By solving the HJI equation, we can obtain the saddle point (u^*, d^*), where u^* is the optimal control law and d^* is the worst-case disturbance[31].

For the time-triggered case, the value function is defined as Eq. (6). Assume the optimal objective function is given by,

$$V^*(x) = \min_u \max_d \int_t^\infty x^T Qx + W(u) - \gamma^2 ||d||^2 d\tau \tag{8}$$

Take the derivative of $V(x)$ with respected to t on the left and right sides of Eq. (6) and obtain

$$\nabla V^T (f(x) + g(x)u + k(x)d) + x^T Qx + W(u) - \gamma^2 ||d||^2 = 0 \tag{9}$$

where $\nabla V = \partial V/\partial x \in R^n, V(0) = 0$.

The Hamiltonian function is defined as,

$$H(x, \nabla V^*, u, d) = \nabla V^{*T} (f + gu + kd) + x^T Qx + W(u) - \gamma^2 ||d||^2 \tag{10}$$

The HJI equation can be written as,

$$\max_d \min_u H(x, \nabla V^*, u, d) = 0 \tag{11}$$

The saddle point (u^*, d^*) exists for the Zero-sum game, when the following condition holds[32],

$$\min_u \max_d H(x, \nabla V, u, d) = \max_d \min_u H(x, \nabla V, u, d) \tag{12}$$

The optimal policy and disturbance law corresponding to the solution of the HJI equation Eq.(11) are as follow,

$$u^*(x) = \arg\min_u H(x, \nabla V^*, u, d) \tag{13}$$

$$d^*(x) = \arg\max_d H(x, \nabla V^*, u, d) \tag{14}$$

By solving these two optimization problem, the optimal control law and the worst-case disturbance can be expressed as follow, respectively,

$$u^*(x) = \bar{U}\varphi(-\frac{1}{2\bar{U}} g(x)^T \nabla V^*(x)) \tag{15}$$

$$d^*(x) = \frac{1}{2\gamma^2} k^T \nabla V^*(x) \tag{16}$$

where $\varphi(\cdot) = \frac{e^x - e^{-x}}{(1/u_{min})e^x - (1/u_{max})e^{-x}}$, and the HJI equation is rewritten as,

$$\begin{aligned} 0 &= H(x, \nabla V^*, u^*, d^*) \\ &= \bar{U}\nabla V^{*T}(x) g(x) \varphi(-\frac{1}{2\bar{U}} g(x)^T \nabla V^*(x)) \\ &+ x^T Qx + \nabla V^{*T} f + W(\bar{U}\varphi(-\frac{1}{2\bar{U}} g(x)^T \nabla V^*(x))) \\ &+ \frac{1}{4\gamma^2} \nabla V(x)^{*T} k(x) k(x)^T \nabla V^*(x) \end{aligned} \tag{17}$$

Up to this point, we have established a time-triggered HJI equation. Based on this, we can obtain the event-triggered HJI equation.

In the event-triggered case, the control law is updated at the sampling state $\hat{x}_j$. Then, Eq. (15) can be rewritten as,

$$\mu^*(\hat{x}_j) = \bar{U}\varphi(-\frac{1}{2\bar{U}} g(\hat{x}_j)^T \nabla V^*(\hat{x}_j)) \tag{18}$$

Substituting the event-triggered control law with asymmetric control constraints into Eq. (17), we can obtain,

$$\begin{aligned} 0 = &\frac{1}{4\gamma^2} \nabla V(x)^{*T} k(x) k(x)^T \nabla V^*(x) \\ &\bar{U}\nabla V^{*T} g(x) \varphi(-\frac{1}{2\bar{U}} g(\hat{x}_j)^T \nabla V^*(\hat{x}_j)) \\ &+ x^T Qx + \nabla V^{*T} f(x) + W(\bar{U}\varphi(-\frac{1}{2\bar{U}} g(\hat{x}_j)^T \nabla V^*(\hat{x}_j))) \end{aligned} \tag{19}$$

For this HJI equation, we have the following assumption and theorem, Assumption 4[33,34]: Assume that $u^*(x)$ satisfies the Lipschitz condition, and there exists a Lipschitz constant L_{u^*} such that the control law and the triggering error satisfy the inequality,

$$||u^*(x) - \mu^*(\hat{x}_j)|| \leq L_{u^*}||x - \hat{x}_j|| = L_{u^*}||e_j|| \tag{20}$$

Theorem1: The state of system (1) is uniform ultimate boundedness (UUB), if the following event-triggered condition holds, $\forall t \in [t_j, t_{j+1}), j = 0, 1, ..., \infty$,

$$||e_j(t)|| \leq \frac{(1-2\theta)x^T Q x - \gamma^2 \cdot ||d^*||^2}{L_{u^*}^2} \tag{21}$$

where θ is a small positive parameter and $d^*(x)$ is given by Eq. (16).

Proof: We choose $V^*(x)$, the solution of HJI equation, as the Lyapunov function. Based on the expression obtained from the previous derivation of the control law and the disturbance law, the derivation of $V(x)$ with respected to t, $\dot{V}^*(x)$ is given by,

$$\begin{aligned} \dot{V}^*(x) &= (\frac{\partial V^*}{\partial x})^T (f(x) + g(x)\mu^*(\hat{x}_j) + k(x)d^*(x)) \\ &= (\frac{\partial V^*}{\partial x})^T (f(x) + g(x)u^*(x)) \\ &+ (\frac{\partial V^*}{\partial x})^T g(x)(\mu^*(\hat{x}_j) - u^*(x)) + (\frac{\partial V^*}{\partial x})^T k(x)d^*(x) \end{aligned} \tag{22}$$

Based on the Eqs.(15) and (16), we can obtain the following equations,

$$g^T \frac{\partial V^*}{\partial x} = -2\bar{U}\varphi^{-T}(\frac{u^*(x)}{\bar{U}}) \tag{23}$$

$$k^T \frac{\partial V^*}{\partial x} = 2\gamma^2 d^*(x) \tag{24}$$

Substituting the above expressions into Eq. (22), we can obtain,

$$\begin{aligned} \dot{V}^*(x) &= -W(u^*) - x^T Q x - 2\bar{U}\varphi^{-1}(\frac{u^*}{\bar{U}}) \\ &(\mu^*(\hat{x}_j) - u^*(x)) + \frac{1}{4\gamma^2}(\nabla V^*(x))^T k(x) k^T(x) \nabla V^*(x) \\ &= -W(u^*) - x^T Q x + 2\bar{U}\varphi^{-1}(\frac{u^*}{\bar{U}})(u^*(x) - \mu^*(\hat{x}_j)) \\ &+ \gamma^2 ||d^*(x)||^2 \end{aligned} \tag{25}$$

The above formula can be simplified as,

$$\begin{aligned} \dot{V}^*(x) &\leq -x^T Q x - W(u^*) + \gamma^2||d^*(x)||^2 + \\ &||u^*(x) - \mu^*(\hat{x}_j)||^2 + ||\bar{U}\varphi^{-T}(\frac{u^*}{\bar{U}})||^2 \end{aligned} \tag{26}$$

According to the assumption 4 we can obtain,

$$\begin{aligned} \dot{V}^*(x) &\leq -x^T Q x - W(u^*) + \gamma^2||d^*(x)||^2 + \\ &L_{u^*}^2 ||e_j(t)||^2 + ||\bar{U}\varphi^{-T}(\frac{u^*}{\bar{U}})||^2 \end{aligned} \tag{27}$$

Note that, $W(u^*)$ can be rewritten as,

$$W(u^*) = 2\sum_{i=1}^{m} \int_0^{u_i^*} \varphi^{-1}(\frac{z_i}{\bar{U}})\bar{U}\, dz_i \tag{28}$$

Let $\tau_i = \varphi^{-1}(\frac{z_i}{\bar{U}})$, and $z_i = \bar{U}\varphi(\tau_i)$ then

$$\begin{aligned} W(u^*) &= 2\bar{U}\sum_{i=1}^{m} \int_0^{\varphi^{-1}(\frac{u_i^*}{\bar{U}})} \tau_i d(\bar{U}\varphi(\tau_i)) \\ &= 2\bar{U}^2 \sum_{i=1}^{m} \int_0^{\varphi^{-1}(\frac{u_i^*}{\bar{U}})} \tau_i \frac{\partial \varphi(\tau_i)}{\partial \tau_i} d\tau_i \end{aligned} \tag{29}$$

The derivative of function $\varphi(\tau_i)$ in the Eq.(29) with respect to τ_i as follows,

$$\begin{aligned}\frac{\partial\varphi(\tau_i)}{\partial\tau_i} &= \frac{\frac{2}{u_{min}}-\frac{2}{u_{max}}}{(\frac{e^{x}}{u_{min}}-\frac{e^{-x}}{u_{max}})^2} \\ &= 1-(\frac{(\frac{e^{\tau_i}}{u_{min}}-\frac{e^{-\tau_i}}{u_{max}})^2-(\frac{2}{u_{min}}-\frac{2}{u_{max}})}{(\frac{e^{\tau_i}}{u_{min}}-\frac{e^{\tau_i}}{u_{max}})^2})\end{aligned} \tag{30}$$

Let

$$p=\frac{(\frac{e^{\tau_i}}{u_{min}}-\frac{e^{-\tau_i}}{u_{max}})^2-(\frac{2}{u_{min}}-\frac{2}{u_{max}})}{(\frac{e^{\tau_i}}{u_{min}}-\frac{e^{\tau_i}}{u_{max}})^2} \tag{31}$$

Mentioned that $u_{max}>0, u_{min}<0$ before, so $p<\zeta$ is a bounded positive number, and ζ is the maximum of p. Then Eq.(29)can be further simplified as follows,

$$\begin{aligned}W(u^*) &= 2\bar{U}^2\sum_{i=1}^{m}\int_0^{\varphi^{-1}(\frac{u_i^*}{\bar{U}})}\tau_i(1-p)d\tau_i \\ &= \bar{U}^2\sum_{i=1}^{m}(\varphi^{-1}(\frac{u_i^*}{\bar{U}}))^2-2\bar{U}^2\sum_{i=1}^{m}\int_0^{\varphi^{-1}(\frac{u_i^*}{\bar{U}})}\tau_i p d\tau_i\end{aligned} \tag{32}$$

Note that

$$\sum_{i=1}^{m}(\varphi^{-1}(\frac{u_i^*}{\bar{U}}))^2=\varphi^{-T}(\frac{u^*}{\bar{U}})\varphi^{-1}(\frac{u^*}{\bar{U}}) \tag{33}$$

Simplify the second term of the above equation by the mean value theorems for definite integrals as follow,

$$\begin{aligned}&2\bar{U}^2\sum_{i=1}^{m}\int_0^{\varphi^{-1}(\frac{u_i^*}{\bar{U}})}\tau_i p d\tau_i \\ &=2\bar{U}^2\sum_{i=1}^{m}\varphi^{-1}(\frac{u_i^*}{\bar{U}})\varpi_i p \\ &\le 2\bar{U}^2\sum_{i=1}^{m}(\varphi^{-1}(\frac{u_i^*}{\bar{U}}))^2 p \\ &\le \frac{1}{2}(\nabla V^*)^T g(x)g^T(x)\nabla V^*\zeta \\ &\le \frac{1}{2}\rho_1^2 g_M^2\zeta\end{aligned} \tag{34}$$

where ϖ_i is in the range of $[0,\varphi^{-1}(\frac{u_i^*}{\bar{U}})]$, and $||\nabla V^*||\le\rho_1$, where ρ_1 is a positive constant and ζ is a positive parameter. Based on the above derivation and Eq.(21), we can obtain,

$$\begin{aligned}\dot{V}^*(x) &\le -x^TQx+\frac{1}{2}\rho_1^2 g_M^2\zeta+\gamma^2||d^*(x)||^2+L_{u^*}^2||e_j(t)||^2 \\ &\le -2\theta\underline{\lambda}(Q)||x||^2-(1-2\theta)\underline{\lambda}(Q)||x||^2+\frac{1}{2}\rho_1^2 g_M^2\zeta \\ &+L_{u^*}^2||e_j(t)||^2+\gamma^2||d^*(x)||^2 \\ &\le -2\theta\underline{\lambda}(Q)||x||^2+\frac{1}{2}\rho_1^2 g_M^2\zeta\end{aligned} \tag{35}$$

where $\underline{\lambda}(Q)$ is the minimum eigenvalue of Q, and θ is a small positive constant. Therefore, the system state is UUB, if the state x is outside of set Ω_x,

$$\Omega_x=\{x:||x||\le\frac{1}{2}\sqrt{\frac{\rho_1^2 g_M^2\zeta}{\theta\underline{\lambda}(Q)}}\} \tag{36}$$

This shows that $V^*(x)$ is one of the Lyapunov function of system (1). Under the control policy $u^*(x)$, the system is UUB, and the state is out of range Ω_x.

Remark1: According to Eq.(32) ,we know that $-(1-2\theta)\underline{\lambda}(Q)||x||^2+L_{u^*}^2||e_j(t)||^2+\gamma^2||d^*(x)||^2<0$, then it is known that $(1-2\theta)\underline{\lambda}(Q)||x||^2>\gamma^2||d^*(x)||^2$. Therefore, the right side of Eq.(22) is a positive number.

3.2 | Zeno Behavior analysis

For control schemes with event triggering, we need to consider the minimal inter sample trigger time τ_{min} to avoid the infinite triggering behavior in the system. We have the following theorem for Zeno behavior for the proposed method.

Theorem 2: Given the system (5) and the event-triggered condition (21), the Zeno behavior will not happen when the minimal inter sample trigger time τ_{min} has the lower bound,

$$\tau_{min}\geq\frac{1}{K}ln(1+\Gamma_{j,min})>0 \tag{37}$$

where K is a positive constant, $\Gamma_{j,min}=\min\limits_{j\in N}(||e_T^{j+1}||/(||\hat{x}_j||+\chi))>0, e_T^{j+1}=||e_T(t_{j+1}^-)||=\lim\limits_{\iota\to 0}e_j(t_{j+1}-\iota)$, with $\iota\in(0,t_{j+1}-t_j)$ and χ is a small positive constant.

Proof: The derivative of event-triggered error is shown as

$$\dot{e}_j(t)=\dot{\hat{x}}_j-\dot{x}(t)=-\dot{x}(t) \tag{38}$$

Suppose the assumptions 2 and 3 are satisfied, the following conclusion is obtained as,

$$||\dot{x}||=||f(x)+g(x)\mu(\hat{x}_j)+k(x)d(x)||\leq K||x||+K\chi \tag{39}$$

and then we have,

$$||\dot{e}_j(t)||\leq K||e_j(t)||+K(||\hat{x}_j||+\chi) \tag{40}$$

$$||e_j(t)||\leq\int_{t_j}^{t}e^{K(t-s)}K(||\hat{x}||+\chi)ds=(||\hat{x}||+\chi)(e^{K(t-t_j)}-1) \tag{41}$$

Based on the known condition $||e_j(t_{j+1}^-)||=e_T^{j+1}$, we can obtain the lower bound on jth inter sample time, which can be expressed as,

$$\tau_j=t_{j+1}-t_j\geq\frac{1}{K}ln(1+||e_T^{j+1}||/(||\hat{x}_j||+\chi)) \tag{42}$$

Then we can obtain that $\tau_{min}=\min\limits_{j}\{\tau_j\}\geq\frac{1}{K}ln(1+\Gamma_{j,min})$, where $\Gamma_{j,min}=\min\limits_{j\in N}(||e_T^{j+1}||/(||\hat{x}_j||+\chi))>0$ and the lower bound of τ_{min} is a positive constant.

4 | NEURAL NETWORK FOR ONLINE POLICY ITERATION ALGORITHM

4.1 | Solving the event-triggered HJI equation via Single Critic Network

As we all know, neural network has the ability of universal approximation, we can fit the optimal value function with the neural network as,

$$V^*(x)=W_c^Th(x)+\epsilon(x) \tag{43}$$

where W_c is the ideal weight, which is usually unavailable, $h(x)$ is the activation function vector, and $\epsilon(x)$ is the error of the NN approximation.

Then we can obtain,

$$\nabla V^*(x)=\nabla h(x)^TW_c+\nabla\epsilon(x) \tag{44}$$

where $\nabla h(x)=\partial h(x)/\partial x, \partial h(0)=0$

Substituting Eq. (44) into Eq. (18),

$$\mu^*(\hat{x}_j)=\bar{U}\varphi(-\frac{1}{2\bar{U}}g(\hat{x}_j)^T\nabla h(\hat{x}_j)^TW_c)+\varepsilon_{\mu 1^*} \tag{45}$$

where $\varepsilon_{\mu 1^*}=-\frac{1}{2}g(\hat{x}_j)^T\nabla\epsilon(\hat{x}_j)\varphi'$. The estimated value of the optimal neural network is used, and the value function can be written as ,

$$\hat{V}(x)=\hat{W}_c^Th(x) \tag{46}$$

The estimated control input can be expressed as,

$$\hat{\mu}(\hat{x}_j)=\bar{U}\varphi(-\frac{1}{2\bar{U}}g(\hat{x}_j)^T\nabla h(\hat{x}_j)^T\hat{W}_c) \tag{47}$$

and the disturbance law can be expressed as,

$$\hat{d}(x) = \frac{1}{2\gamma^2} k^T \nabla h(x)^T \hat{W}_c \tag{48}$$

According to the control law and disturbance law, the estimated HJI equation can be expressed as,

$$H(x, \nabla \hat{V}(x), \hat{\mu}(\hat{x}_j), \hat{d}(x)) = r(x, \hat{\mu}(\hat{x}_j), \hat{d}(x)) + \hat{W}_c^T \varrho \tag{49}$$

where $r(x, \hat{\mu}(\hat{x}_j), \hat{d}(x)) = x^T Q x + W(\hat{\mu}(\hat{x}_j)) - \gamma^2 ||\hat{d}(x)||^2$, $\varrho = \nabla h(x)(f(x) + g(x)\hat{\mu}(\hat{x}_j) + k(x)\hat{d}(x))$.

Define the error between $H(x, \nabla V^*, u^*, d^*)$ and $H(x, \nabla \hat{V}(x), \hat{\mu}(\hat{x}_j), \hat{d}(x))$ as ε_{HJI}, namely,

$$\varepsilon_{HJI} = r(x, \hat{\mu}(\hat{(}x_j), \hat{d}(x)) + \hat{W}_c^T \varrho \tag{50}$$

Because of the use of event-triggered method, the error at time t_k can be expressed as,

$$\varepsilon_{HJI}(t_k) = r(t_k) + \hat{W}_c^T \varrho_k \tag{51}$$

where $r(t_k) = x^T(t_k) Q x(t_k) + W(\hat{\mu}(\hat{x}_j)) - \gamma^2 ||\hat{d}(x)||^2$ and $\varrho_k = \nabla h(x(t_k))(f(x(t_k)) + g(x(t_k))\hat{\mu}(\hat{x}_j) + k(x(t_k))\hat{d}(x))$, where $t_k \in [t_j, t_{j+1})$, $(k \in \{1, ..., p\})$. Let $rank([\varrho_1, ..., \varrho_p]) = n_c, (p > n_c)$, where n_c is the number of neurons in the critic network, and p is a large number of historical state data.

In order to achieve $\lim_{n_c \to \infty} \hat{W}_c = W_c$, our goal is to make $\varepsilon_{HJI} \to 0$. Define $E = (1/2)\varepsilon_{HJI}^T \varepsilon_{HJI}/(1 + \varrho^T \varrho)^2$ and $E_k = \sum_{k=1}^{p} (1/2)\varepsilon_{HJI}^T(t_k)\varepsilon_{HJI}(t_k)/(1 + \varrho_k^T \varrho_k)^2$, the gradient is given by,

$$\begin{aligned}
\dot{\hat{W}}_c &= -l_c \cdot \frac{\partial E}{\partial \hat{W}_c} - l_c \cdot \frac{\partial E_k}{\partial \hat{W}_c} \\
&= -\frac{l_c}{(1 + \varrho^T \varrho)^2} \varepsilon_{HJI} \cdot \varrho \\
&- \sum_{k=1}^{p} \frac{l_c}{(1 + \varrho_k^T \varrho_k)^2} \varepsilon_{HJI}(t_k) \cdot \varrho_k
\end{aligned} \tag{52}$$

where l_c is the learning rate. Define $\tilde{W}_c$ as the weight estimation error, namely $\tilde{W}_c = W_c - \hat{W}_c$. Through the above derivation, we known that $\varepsilon_{HJI} = -\tilde{W}_c \varrho - \nabla \epsilon(x)(f(x) + g(x)\mu(\hat{x}_j) + k(x)d(x))$ and $\varepsilon_{HJI}(t_k) = -\tilde{W}_c \varrho_k - \nabla \epsilon(x(t_k))(f(x(t_k)) + g(x(t_k))\mu(\hat{x}_j) + k(x(t_k))d(x))$, then we can obtain,

$$\begin{aligned}
\dot{\hat{W}}_c &= -lc\left(\frac{\varrho^2}{(1 + \varrho^T \varrho)^2} - \sum_{k=1}^{p} \frac{\varrho_k^2}{(1 + \varrho_k^T \varrho_k)^2}\right)\tilde{W}_c \\
&-lc\left(\frac{\varrho}{(1 + \varrho^T \varrho)^2} \varepsilon_a + \sum_{k=1}^{p} \frac{\varrho_k}{(1 + \varrho_k^T \varrho_k)^2} \varepsilon_b(t_k)\right)
\end{aligned} \tag{53}$$

where $\varepsilon_a = \nabla \epsilon(x)(f(x) + g(x)\mu(\hat{x}_j) + k(x)d(x))$ and $\varepsilon_b(t_k) = \nabla \epsilon(x(t_k))(f(x(t_k)) + g(x(t_k))\mu(\hat{x}_j) + k(x(t_k))d(x))$

4.2 | Stability Analysis

Therem 3: According to the designed controller (47) and the critic neural network update formula (52), the closed-loop system is stable and the critic network parameters is UUB under the event-triggered mechanism (21).

Proof:The Lyapunov function is designed as,

$$L(t) = \underbrace{V^*(x)}_{L_1(t)} + \underbrace{V^*(\hat{x}_j)}_{L_2(t)} + \underbrace{\frac{1}{2\gamma} \tilde{W}_c^T \tilde{W}_c}_{L_3(t)} \tag{54}$$

Case 1: The event-triggered condition is not violated. In this case, we have $\dot{L}_2(t) = 0$ and

$$\begin{aligned}
\dot{L}_1 &= \left(\frac{\partial V^*}{\partial x}\right)^T (f(x) + g(x)\hat{\mu}(\hat{x}_j) + k(x)\hat{d}(x)) \\
&= (\nabla V^*)^T (f(x) + g(x)u^*(x)) \\
&+ (\nabla V^*)^T g(x)(\hat{\mu}(\hat{x}_j) - u^*(x)) + (\nabla V^*)^T k(x)\hat{d}(x) \\
&= -x^T Q x - W(\hat{\mu}) - \gamma^2 d^{*T}(x) d^*(x) - \\
2\bar{U}\varphi^{-1}&\left(\frac{u^*}{\bar{U}}\right)^T (\hat{\mu}(\hat{x}_j) - u^*(x)) + 2\gamma^2 d^{*T}(x)\hat{d}(x)
\end{aligned} \tag{55}$$

Based on the Young's inequality, we can obtain,

$$\begin{aligned}\dot{L}_1(t) \leq & -x^T Qx - W(\hat{\mu}) - 2\bar{U}\varphi^{-1}(\frac{u^*}{\bar{U}})^T(\hat{\mu}(\hat{x}_j) \\ & - u^*(x)) + \gamma^2 \hat{d}(x)^T \hat{d}(x) \\ & \leq -x^T Qx - W(\hat{\mu}) + \gamma^2 \hat{d}(x)^T \hat{d}(x) \\ & + (\bar{U}\varphi^{-1}(\frac{u^*}{\bar{U}})^T)^2 + (||u^*(x) - \hat{\mu}(\hat{x}_j)||)^2 \\ & \leq -x^T Qx + L_{u^*}^2 ||e_j||^2 + \gamma^2 \hat{d}(x)^T \hat{d}(x) + \frac{1}{2}\rho_1^2\rho_2^2\zeta\end{aligned} \tag{56}$$

Based on this result, if Eqs. (21) and (36) hold, $\dot{L}_1(t) < 0$ holds.

Third part of the Eq. (54) is given by.

$$\begin{aligned}\dot{L}_3(t) = & \tilde{W_c}^T \dot{\hat{W}}_c \\ & = -\tilde{W_c}^T(\frac{\varrho^2}{(1+\varrho^T\varrho)^2} - \sum_{k=1}^{p}\frac{\varrho_k^2}{(1+\varrho_k^T\varrho_k)^2})\hat{W}_c \\ & - \tilde{W_c}^T(\frac{\varrho}{(1+\varrho^T\varrho)^2}\varepsilon_a + \sum_{k=1}^{p}\frac{\varrho_k}{(1+\varrho_k^T\varrho_k)^2}\varepsilon_b(t_k))\end{aligned} \tag{57}$$

Let $\eta = \frac{\varrho}{(1+\varrho^T\varrho)}$, from Young's inequality, and we can obtain,

$$\begin{aligned}\frac{\varrho}{(1+\varrho^T\varrho)}\tilde{W}_c^T\eta & \leq \frac{\varrho}{2(1+\varrho^T\varrho)}(\tilde{W}_c^T\eta\eta^T\tilde{W}_c^T + \varepsilon_a^T\varepsilon_a) \\ & \leq \frac{1}{2}(\tilde{W}_c^T\eta\eta^T\tilde{W}_c^T + \varepsilon_a^T\varepsilon_a)\end{aligned} \tag{58}$$

Then,

$$\begin{aligned}\dot{L}_3(t) \leq & -\frac{1}{2}\tilde{W}_c^T(\eta\eta^T - \sum_{k=1}^{p}\eta(t_k)\eta^T(t_k))\tilde{W}_c \\ & - \frac{1}{2}(\varepsilon_a^T\varepsilon_a + \sum_{k=1}^{p}\varepsilon_b^T(t_k)\varepsilon_b(t_k))\end{aligned} \tag{59}$$

Let $Z = \eta\eta^T - \sum_{k=1}^{p}\eta(t_k)\eta^T(t_k)$ and presume $||\varepsilon_a|| \cdot ||\varepsilon_b|| \leq q$, then,

$$\dot{L}_3(t) \leq -\frac{1}{2}(\underline{\lambda}(Z)||\tilde{W}_c||^2 + (1+p)q^2) \tag{60}$$

Therefore $\dot{L}_3(t) \leq 0$,when the following inequality holds,

$$||\tilde{W}_c||^2 > \Omega_{\tilde{W}_c} \tag{61}$$

where $\Omega_{\tilde{W}_c} = \frac{(1+p)q^2)}{\eta\eta^T - \sum_{k=1}^{p}\eta(t_k)\eta^T(t_k)}$.

Case 2: Event-triggered condition is violated, $t = t_{j+1}$ for $j = 1, 2, ..., \infty$.
Now we define the Lyapunov function as,

$$\begin{aligned}\triangle L(t_j) = & V^*(\hat{x}_{j+1}) - V^*(\hat{x}_j) + V^*(x(t_{j+1})) - V^*(x(t_{j+1}^-)) \\ & + \frac{1}{2\gamma}\tilde{W}_c^T(t_{j+1})\tilde{W}_c(t_{j+1}) - \frac{1}{2\gamma}\tilde{W}_c^T(t_{j+1}^-)\tilde{W}_c(t_{j+1}^-)\end{aligned} \tag{62}$$

Based on the conclusion of Case 1, we can obtain that $L_2(t)$ and $L_3(t)$ are decreasing strictly monotonically on the interval $[t_j, t_{j+1})$,

$$\begin{aligned}L_2(t_{j+1}) & < L_2(t_{j+1} - \delta) \\ & \leq \lim_{\delta \to 0^+} L_2(t_{j+1} - \delta) \\ & = L_2(t_{j+1}^-)\end{aligned} \tag{63}$$

Similarly, we can obtain $L_3(t_{j+1}) < L_3(t_{j+1}^-)$, where $\delta \in (0, t_{j+1} - t_j)$.

Based on the definitions of $L_1(t)$, $L_2(t)$ and $L_3(t)$ in Case 1, we can obtain,

$$\begin{aligned}& V^*(x(t_{j+1})) + \frac{1}{2\gamma}\tilde{W}_c^T(t_{j+1})\tilde{W}_c(t_{j+1}) \\ & \leq V^*(x(t_{j+1}^-)) + \frac{1}{2\gamma}\tilde{W}_c^T(t_{j+1}^-)\tilde{W}_c(t_{j+1}^-)\end{aligned} \tag{64}$$

and

$$V^*(\hat{x}_{j+1}) \le V^*(\hat{x}_j) \tag{65}$$

Then, we can obtain $\triangle L(t_j) < 0$, and the proof is completed.

5 | SIMULATION STUDIES

In this section, tow simulation examples are carried out to verify the effectiveness of the proposed methods.

5.1 | Example 1

A nonlinear system is described as

$$\dot{x} = \begin{bmatrix} -x_1 + x_2 \\ -0.5(x_1 + x_2) + 0.5x_2 \sin(x_1)^2 \end{bmatrix} + \begin{bmatrix} 0 \\ \sin(x_1) \end{bmatrix} u + \begin{bmatrix} 0 \\ \cos(x_1) \end{bmatrix} d \tag{66}$$

where $x = [x_1, x_2]^T$ and $x_0 = [1, -1]^T$. $L_{u^*} = 3$, $p = 10$, and Q is the determined matrix with the appropriate dimension. During network training, the exploration noise is added to satisfy the persistency of excitation condition as: $N(t) = -1.5e^{0.006t}\sin(t)^2\cos(t)\sin(-1.2t)^2\cos(0.5t) + \sin(t)^5 + \sin(1.12t)^2 + \cos(2.4t)\sin(2.4t)^3$.

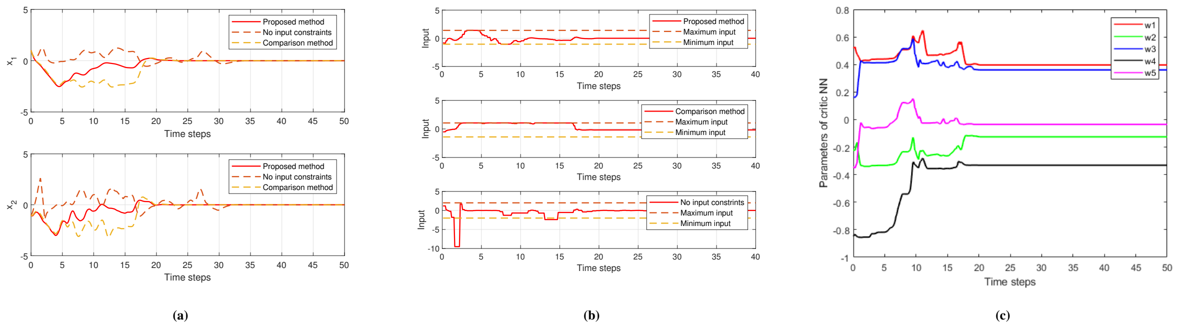


Figure 1 (a) State in Example 1. (b) Input in Example 1. (c) $\hat{W}_c$ in Example 1

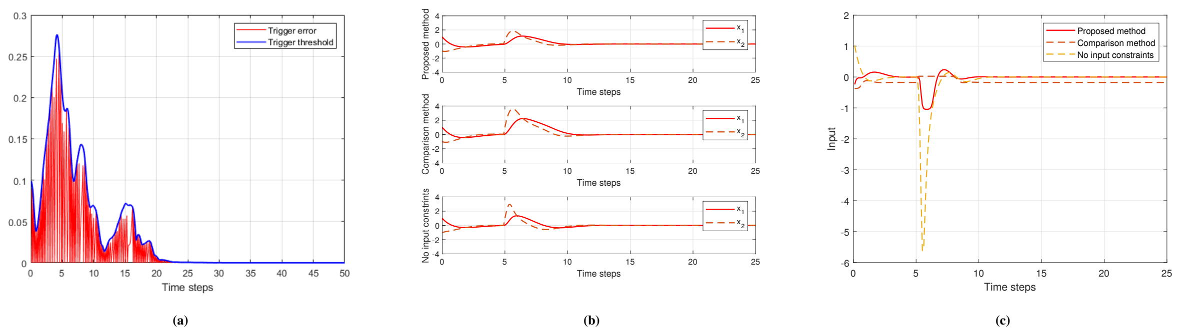


Figure 2 (a) Triggered error and triggered threshold in Example 1. (b) Input under d_1 in Example 1. (c) Input under d_1 in Example 1

Table 1 Integrated Absolute Error (IAE) with Different Disturbance.

Different Disturbance	Proposed Method	Comparison Method	No Input Constraint
$d_1 = 10e^{-(t-5)}\cos(t-5)$	3.0287	4.6358	5.6314
$d_2 = 12\sin(t-5)e^{-0.7(t-5)}$	2.4548	3.5553	3.4313
$d_3 = 13\sin(t-5)\cos(t-5)e^{-0.7(t-5)}$	1.9599	3.7300	4.5085

The neuron structure of the critic network is set as 2-5-1. The activation function of the critic network is selected as $h(x) = [x_1^2, x_1x_2, x_2^2, x_1^4, x_2^4]^T$ and the initial value of the weight is chosen as $\hat{W}_c = [0.4719, 0.06273, 0.8243, 0.2005, 0.31346]^T$.

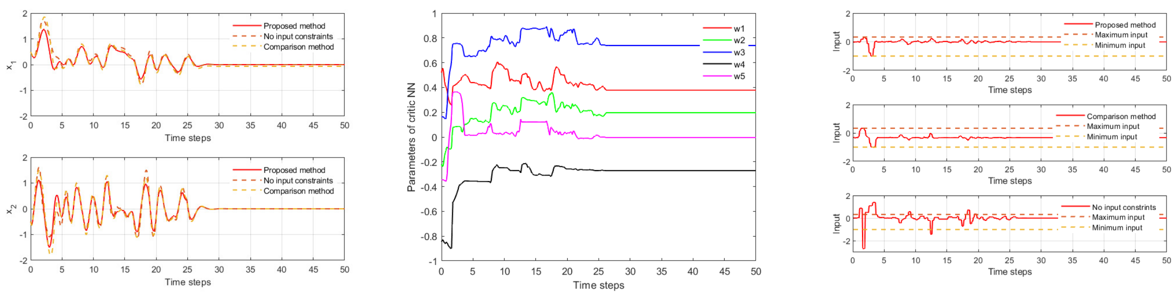
Fig. 1(a) shows the state convergence of the proposed method under asymmetric input constraints, and compares it with the general method under asymmetric input constraints [26] and the general method without input constraints [35]. The control policy of above three methods is shown in Fig. 1(b). It can be seen that the proposed method has smaller fluctuation and shorter transient process, and the control input converges to zero within the given constraint range $-1.4 \leq |u(t)| \leq 1.05$. The learning process of the function network weights is shown in Fig. 1(c). The relationship of the triggered error $||e_j||$ and the threshold e_T is shown in Fig. 2(a), which can be seen that the triggered error converges to zero. In this paper, we only verify whether the system can still return to zero stably under asymmetric input constraints. Therefore, a disturbance $d = 10e^{-(t-5)}\cos(t-5)$ is applied to the system when $t > 5s$. The system states and control input are shown in Fig. 2(b) and Fig. 2(c), respectively. In Table 1, the results of this method are compared with those of other two methods, where the different disturbances are defined as $d_1 = 10e^{-(t-5)}\cos(t-5)$, $d_2 = 12\sin(t-5)e^{-0.7(t-5)}$ and $d_3 = 13\sin(t-5)\cos(t-5)e^{-0.7(t-5)}$. It can be seen from the table that the proposed method has better performance on the disturbance suppression.

5.2 | Example 2

Consider the following single connected robot arm system,

$$\begin{cases} \frac{d\theta}{dt} = \omega \\ \frac{d\omega}{dt} = -\frac{MgH}{G}\sin(\theta) - \frac{D}{G}\omega + \frac{1}{G}u + \frac{1}{L}d \end{cases} \tag{67}$$

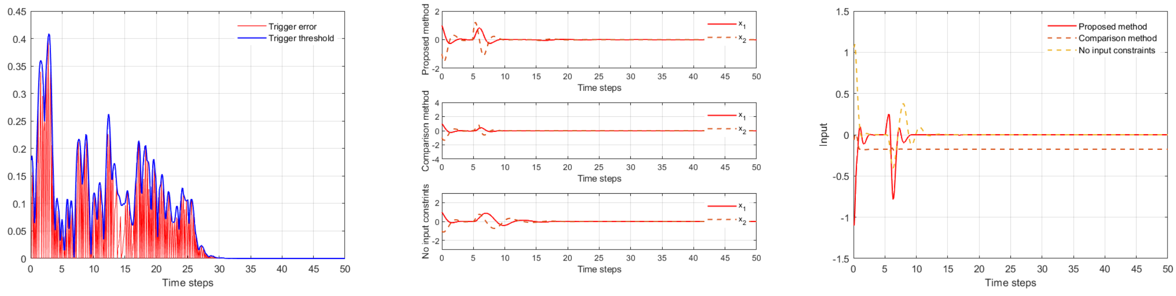
where $x_1 = \theta$ and $x_2 = \omega$ denote the angle and the angular velocity of robot arm; $D = 2$ is viscous friction; $H = 0.5$ is the length of the robot arm; $g = 9.8$ is gravitational acceleration; $M = 1$ is the mass of the payload; $G = 1$ is the moment of inertia;L=0.5; other parameters are consistent with those in Example 1.



(a) (b) (c)

Figure 3 (a) State in Example 2. (b) $\hat{W}_c$ in Example 2. (c) Input in Example 2

In Example 2, the proposed method is also compared with the conventional method and the no input constraint method. Fig. 3(a) shows that compared with the other two methods, the proposed method can bring the state back to equilibrium faster. Fig. 3(b) shows the convergence process of network weights. From Fig. 3(c), it can be observed that the asymmetric input is constrained to the range of $-1 \leq |u(t)| \leq 0.3$. The event-triggered error and the threshold converge to zero in Fig. 4(a). To verify the effectiveness of the proposed methods to the disturbance, we compare the proposed method with the other two methods and



(a) (b) (c)

Figure 4 (a) Triggered error and triggered threshold in Example 2. (b) State under d_2 in Example 2. (c) Input under d_1 in Example 2

find that the proposed method has better performance in dealing with d_2 in Fig. 4(b). In order to illustrate the benefits of input constraint, the method proposed in this paper is compared with the method without constraint. It can be seen from Fig. 4(c) that compared with the input with constraint, the input without constraint will produce a relatively violent fluctuation after external disturbance.

6 | CONCLUSION

We have proposed a disturbance rejection control strategy for nonlinear system with input constrains. To deal with the uncertain disturbance rejection, the control design is converted into a Zero-sum game problem, and an ADP-based method is proposed to solve this Zero-sum game problem. To deal with the asymmetric input constraints, a new objective function has proposed. We use a critic-only network to approximate the optimal performance index function, and use the event-triggered mechanism to reduce the computational pressure. The critic network weight vector is tuned through a modified gradient descent method with the historical state data. The stability of the closed-loop system and the UUB of the critic network parameters is proved by the Lyapunov method. The comparison studies show the effectiveness and merits of the proposed method to deal with the uncertain disturbance and asymmetric constraints.

References

1. W. H. Chen, J. Yang, L. Guo, and S. Li, "Disturbance-Observer-Based Control and Related Methods-An Overview," IEEE Transactions on Industrial Electronics, vol. 63, no. 2, pp. 1083–1095, 2016.
2. W. Chenlu, C. Zengqiang, S. Qinglin, and Z. Qing, "Design of PID and ADRC based quadrotor helicopter control system," in Control Decision Conference, pp. 5860–5865, 2016.
3. Z. Gao, S. Hu, and F. Jiang, "A novel motion control design approach based on active disturbance rejection," in Decision and Control, 2001. Proceedings of the 40th IEEE Conference on, 2001.
4. K. Ohnishi, "A new servo method in mechatronics," Trans of Japanese Society of Electrical Engineering D, vol. 177, 1987.
5. J. Yang, S. Li, X. Chen, and L. Qi, "Disturbance rejection of ball mill grinding circuits using DOB and MPC," Powder Technology, vol. 198, no. 2, pp. 219–228, 2010.
6. A. Rf, A. Sk, A. Mya, B. Mb, and C. Mba, "Active disturbance rejection control for robotic systems: A review,"
7. J. Han, "From PID to Active Disturbance Rejection Control," IEEE Transactions on Industrial Electronics, vol. 56, no. 3, pp. 900–906, 2009.
8. J. Han, "Auto-disturbances-rejection Controller and Its Applications," Control and Decision, 1998.

9. J. Q. Han, "Nonlinear design methods for control systems," IFAC Proceedings Volumes, vol. 32, no. 2, pp. 1531–1536, 1999.

10. Q. Liu, Q. Jin, B. Huang, and L. Min, "Iteration Tuning of Disturbance Observer-Based Control System Satisfying Robustness Index for FOPTD Processes," IEEE Transactions on Control Systems Technology, vol. 25, no. 6, pp. 1978–1988, 2016.

11. J. Yang, S. Li, X. Chen, and L. Qi, "Disturbance rejection of dead-time processes using disturbance observer and model predictive control," Chemical Engineering Research Design, vol. 89, no. 2, pp. 125–135, 2011.

12. Z. Huang, H. Zheng, S. Wang, Y. Liu, J. Ma, and Y. Liv, "A self-searching optimal ADRC for the pitch angle control of an underwater thermal glider in the vertical plane motion," Ocean Engineering, vol. 159, no. JUL.1, pp. 98–111, 2018.

13. W. Lu, Q. Li, K. Lu, Y. Lu, and X. Feng, "Load Adaptive PMSM Drive System Based on an Improved ADRC for Manipulator Joint," IEEE Access, vol. PP, no. 99, pp. 1–1, 2021.

14. Z. Ni, H. He, J. Wen, "Adaptive Learning in Tracking Control Based on the Dual Critic Network Design," IEEE Trans Neural Netw Learn Syst, no. 6, 2013.

15. X. Bu and Q. Qi, "Fuzzy optimal tracking control of hypersonic flight vehicles via single-network adaptive critic design," IEEE Transactions on Fuzzy Systems, vol. 30, no. 1, pp. 270-278, January 2022.

16. X. Bu, Y. Xiao, and H. Lei, "An adaptive critic design-based fuzzy neural controller for hypersonic vehicles: predefined behavioral nonaffine control," IEEE/ASME Transactions on Mechatronics, vol. 24, no. 4, pp. 1871-1881, AUGUST 2019.

17. X. Bu, B. Jiang, and H. Lei, "Non-fragile Quantitative Prescribed Performance Control of Waverider Vehicles With Actuator Saturation," IEEE Transactions on Aerospace and Electronic Systems, vol. 58, no. 4, pp. 3538-3548, August 2022.

18. Mu.C,Zhang.Y,Gao.Z, Sun. C, "ADP-Based Robust Tracking Control for a Class of Nonlinear Systems With Unmatched Uncertainties," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 50,2013.

19. Y. Xiong, H. Bo, X. N, "Approximate Dynamic Programming for Nonlinear-Constrained Optimizations," IEEE transactions on cybernetics, 2019.

20. Z. P. Jiang and Y. Jiang, "Robust adaptive dynamic programming for linear and nonlinear systems: An overview," European Journal of Control, vol. 19, no. 5, pp. 417–425, 2013.

21. I. N. Kar, D. M. Adhyaru, and M. Gopal, "Fixed final time optimal control approach for bounded robust controller design using Hamilton-Jacobi-Bellman solution," Iet Control Theory Applications, vol. 3, no. 9, pp. 1183–1195, 2009.

22. D. Wang, D. Liu, and H. Li, "Policy Iteration Algorithm for Online Design of Robust Control for a Class of Continuous-Time Nonlinear Systems," IEEE Transactions on Automation Science and Engineering, vol. 11, no. 2, pp. 627–632, 2014.

23. X. Yang, D. Liu, B. Luo, and C. Li, "Data-based robust adaptive control for a class of unknown nonlinear constrained-input systems via integral reinforcement learning," Information ences, pp. 731–747, 2016.

24. B. Cai, L.X. Zhang, "Control Synthesis of Hidden Semi-Markov Uncertain Fuzzy Systems via Observations of Hidden Modes," IEEE Transactions on Cybernetics, vol. 50, no. 8, pp. 3709–3718, 2020.

25. X. Yao, L. Zhang, and W. X. Zheng, "Uncertain Disturbance Rejection and Attenuation for Semi-Markov Jump Systems With Application to 2-Degree-Freedom Robot Arm," IEEE Trans. Circuits Syst, vol. 68, no. 9, pp. 3836–3845, 2021.

26. Xiong, Yang, Bo, and Zhao, "Optimal neuro-control strategy for nonlinear systems with asymmetric input constraints," IEEE/CAA Journal of Automatica Sinica, vol. v.7, no. 02, pp. 264–272, 2020.

27. X. Yang and Q. Wei, "Adaptive Critic Learning for Constrained Optimal Event-Triggered Control With Discounted Cost," IEEE Transactions on Neural Networks and Learning Systems, vol. PP, no. 99, pp. 1–14, 2020.

28. L. Kong, W. He, Y. Dong, L. Cheng, C. Yang, and Z. Li, "Asymmetric Bounded Neural Control for an Uncertain Robot by State Feedback and Output Feedback," IEEE Transactions on Systems, Man, and Cybernetics Systems, pp. 1–12, 2019.

29. F. L. Lewis and D. Liu, "Reinforcement Learning and Approximate Dynamic Programming for Feedback Control," pp. 452–473, 2012.

30. W. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in Decision and Control (CDC), 2012 IEEE 51st Annual Conference on, 2012.

31. T. Basar and P. Bernhard, H_∞-Optimal Control and Related Minimax Design Problems. SIAM Review, pp. 538–540, 2008.

32. M. Abu-Khalaf, F. L. Lewis, and J.Huang, "Policy Iterations on the Hamilton-Jacobi-Isaacs Equation for for H_∞ State Feedback Control With Input Saturation," IEEE Transactions on Automatic Control, vol. 51, no. 12, pp. p.1989–1995, 2006.

33. X. Yang,H. He, "Event-Triggered Robust Stabilization of Nonlinear Input-Constrained Systems Using Single Network Adaptive Critic Designs," IEEE Trans. Syst, vol. 50, no. 9, pp. p.3145–3157, 2020.

34. K. G. Vamvoudakis, "Event-Triggered Optimal Adaptive Control Algorithm for Continuous-Time Nonlinear Systems," IEEE/CAA J. of Autom. Sinica, vol. 001, no. 003, pp. 282–293, 2014.

35. Zhang. K, Zhang. H, Xiao. G, Su. H, "Tracking Control Optimization Scheme of Continuous-Time Nonlinear System via Online Single Network Adaptive Critic Design Method," Neurocomputing, vol. 251, pp. 127-135, 2017.

AUTHOR BIOGRAPHY

Yan Liu received the B.S. degree from Shandong University of Technology, ShanDong, China, in 2020. She is currently a graduate student in Chongqing University of Posts and Telecommunications.

Her research interests include adaptive dynamic programming and reinforcement learning.

GuoQing Qiu is currently an Associate Professor with the College of Automation, Chongqing University of Posts and Telecommunications.

His research interests include control theory and application, Internet of intelligent instrument, and electromechanical integration technology.

Wenyao Gou received his B.S. degree from Shandong Jianzhu University, ShanDong, China, in 2020. He is now studying in Chongqing University of Posts and Telecommunications as a postgraduate student.

His research interests include reinforcement learning and robot control.

Qie Liu (M'18) received his BSc and PhD degrees in Automatic Control from the Beijing University of Chemical Technology, Beijing, China, in 2011 and 2016, respectively. He was a visitingPhD student at the University of Alberta, Edmonton, AB, Canada from 2014 to 2015. From 2016 to 2018, he worked as a Post-Doctoral Fellow with the Department of Automation, Tsinghua University, Beijing. Since 2018, he has joined the School of Automation in Chongqing University, Chongqing, China. He is currently an associate professor at Chongqing University.

His main research interests include the artificial intelligence and its applications on the control and optimization of complex processes.