



Article Critic-Only Learning Based Tracking Control for Uncertain Nonlinear Systems with Prescribed Performance

Yanping Gao * D and Zuojun Liu

School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China; liuzuojun@hebut.edu.cn * Correspondence: zbkz_gao@163.com

Abstract: A critic-only learning-based tracking control with prescribed performance was proposed for a class of uncertain nonlinear systems. Based on an estimator and an optimal controller, a novel controller was designed to make tracking errors uniformly ultimately bounded and limited in a prescribed region. First, an unknown system dynamic estimator was employed online to approximate the uncertainty with an invariant manifold. Subsequently, by running a novel cost function, an optimal controller was derived by online learning with a critic-only neural network, which ensured that tracking errors can evolve within a prescribed area while minimizing the cost function. Specifically, weight update can be driven by weight estimation error, avoiding introducing an actor-critic architecture with a complicated law. At last, the stability of a closed-loop system was analyzed by Lyapunov theorem, and tracking errors evolved within prescribed performance with the optimal controller. The effectiveness of the proposed control can be demonstrated by two examples.

Keywords: prescribed performance control; tracking control; adaptive dynamic programming

1. Introduction

Nowadays, many plants in the real word can be formulated as uncertain nonlinear systems, such as autonomous underwater vehicles [1], mobile robots [2], and quadrotors [3]. Thus, tracking problems for uncertain nonlinear systems have become one of the issues, and various control methods have been developed and employed in many physical systems to achieve tracking performance.

To deal with the impact induced by uncertainties during controller design, there emerge a large amount of tools to approximate uncertainties and combined with other controllers, e.g., neural networks (NNs) [4,5], fuzzy logic systems [6–8], and disturbance observers [9–11]. For example, by integrating with a backstepping control, an adaptive tracking controller is designed in [5] for uncertain nonlinear systems, where an NN is utilized to estimate the uncertain term of the model. Due to an uncertain dynamical model in surface vessels, proportional derivative feedback combined with a fuzzy logic system is proposed in [7] with satisfying tracking performance and theoretical results. Taking the disturbance ability of an observer-based controller into consideration, the tracking performance can be guaranteed. For manipulators, a sliding-mode controller is investigated for tracking problems, where a nonlinear disturbance observer is implemented to predict and remove the effect of disturbance [10]. However, manually tunned parameters limit above controllers applied in practical applications. Recently, the unknown system dynamics estimator (USDE) is a novel estimation method to deal with uncertainty and disturbance of nonlinear systems [12], where the state can be an added filtered operation such that invariant manifold is conducted for accurate estimation. Different from NNs and fuzzy logic systems with repeated tuning, the estimator only uses the system state and the control input and can achieve rapid convergence of disturbance estimation by adjusting a parameter. For the motion control of robot systems, lumped disturbances are estimated through an improved unknown disturbance estimator [13]. In [14], by implementing a



Citation: Gao, Y.; Liu, Z. Critic-Only Learning Based Tracking Control for Uncertain Nonlinear Systems with Prescribed Performance. *Electronics* **2023**, *12*, 2545. https://doi.org/ 10.3390/electronics12112545

Academic Editor: Mahmut Reyhanoglu

Received: 25 April 2023 Revised: 29 May 2023 Accepted: 2 June 2023 Published: 5 June 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/). USDE to compensate the disturbance, a sliding-mode control is designed to obtain the performance with fast convergence and strong robustness. Since USDEs in the above literature are aimed at integral series systems, it is difficult to solve the strong coupling and multivariable functions of nonlinear systems. Moreover, it is worth mentioning that the above control methods with an inherent structure do not discuss the output constraints such as the convergence rate and the maximum steady-state error, which are important in engineering.

Combined with reinforcement learning (RL), which is a branch of machine learning, a mass of adaptive controllers for stability of closed-loop systems have been developed. By integrating it with dynamic programming, there emerge an optimal controller with a learning ability and the balancing between the tracking performance and the control cost, which are generalized as adaptive dynamic programming (ADP) without dimensionality. One of prevalent structures is actor-critic ADP to pursue the optimal control and the optimal value function [15-17]. However, tracking error may satisfy the preassigned convergence in practical engineering applications. To address this issue, the introduction of the RL algorithm into the prescribed performance control (PPC) has attracted attention [18], which significantly reduces tracking error and control input and improves performance [19–21]. In [19], a data-driven RL algorithm for performance specification was proposed to simultaneously pursue control methods for satisfying optimality and tracking errors to meet output constraints. Combined with the fault-tolerant control (FTC), nonlinear systems with output constraints are considered by RL algorithms [22–24]. It is noted that it is difficult to achieve the fault tolerant control by RL alone. In [23], for nonlinear systems with actuator faults, a model-free adaptive optimal control method with specified performance was designed, where an adaptive observer is employed to estimate faults, and the incremental system parameters are estimated by the recursive least squares identification method. By combining the PPC with ADP, the optimal control strategy is obtained, such that tracking error satisfies the specified performance. By introducing an intermediate controller, a controller, and a fault tolerant controller based on RL algorithms, a fault-tolerant dynamic surface control algorithm based on actor-critic ADP was proposed in [24] for nonlinear systems with unknown parameters and actuator faults, which avoids the difficulty of RL in fault tolerant control. For the manipulator, a robust motion control method with specified performance based on reinforcement learning was proposed in [25]. The measurement noise is eliminated by carefully adding an integral term and adopting a robust generalized proportional integral observer, and an optimal control strategy based on error transformation was designed with the actor-critic ADP method to ensure the stability of the system. For the high-order nonlinear multi-agent system containing uncertainty, the optimal consistency control problem with specified performance is considered in [26], where the stability of a closed-loop system and the convergence of consensus errors within a certain range are proved. Based on an actor-critic network, the optimal control is investigated for robot [27] and pure feedback system [28] separately. It is noted that the above controllers rely on multiple neural networks to ensure the stability and optimality of the system. The weight of the neural network is complicated due to the increasing number of nodes, and many parameters are difficult to adjust. In order to satisfy the optimal predetermined performance, the main idea is to transform the constrained tracking error into an unconstrained variable by constructing transformation function, and approximate optimal control is designed within actor-actor NN by minimizing the value function related to the unconstrained variable. Noting that the Hamilton-Jacobi-Isaacs (HJI) equation can be solved for deriving the optimal control in [29,30], it is deemed to be conservative for the worst case of disturbance with massive control inputs. Moreover, existing design and theoretical analysis of the preset performance optimal control are mostly combined with fault-tolerant control, robust control, and adaptive control. There is no relevant research for the USDE-based optimal control to assure the preassigned convergence rate.

Motivated by the above statements, we proposed an optimal control with asymmetric performance constraints under the framework of critic-only ADP by constructing a new value function. The highlights in this article can be expressed as follows:

(1) Compared with the existing observer-based controllers [9–11], where the disturbance is estimated with observers by adjusting manually multiple parameters, the USDE was employed to approximate the lumped disturbance of nonlinear uncertain systems with an invariant manifold principle. Moreover, differing from previous function approximator-based control schemes [12–14], by combining the RL technique into the optimal control design, the precise tracking performance and the low control cost can be achieved.

(2) Different from the optimal control derived by the actor-critic ADP framework in [24,26–28], a critic-only NN was designed for online learn optimal control without constructing an actor NN. In addition, to achieve the convergence rate with a preassigned region, a novel value function was minimized, such that tracking errors can evolve within the prescribed region with low control consumption. In contrast to the traditional gradient descent for weight update in actor-critic ADP, the weight law was designed ingeniously to update the weight of a critic NN, reduce the online training computation and accelerate the weight convergence.

The outline of the article is organized as follows. We first gave the definition of the optimal tracking control problem with prescribed performance in Section 2. In Section 3, the main result of the control design was discussed, in which a feedback controller and an optimal controller were proposed. The stability analysis is illustrated in Section 4. Section 5 provides the effectiveness of the proposed controller on two examples.

Throughout this paper, the vector or matrix is represented by bold fonts, which is different from the scalar. **0** represents the zero matrix and *I* denotes the identity matrix. diag(*x*) is the diagonal matrix conducted with vector *x*. $\lambda_{\min}(\cdot)$ is the minimized eigenvalue of the corresponding matrix.

2. Preliminaries

The following multiple-input-multiple-output system with disturbance is considered:

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

where $\mathbf{x} = [x_1, x_2, ..., x_n]^T$ denotes the measurable state vector and $\mathbf{u} = [u_1, u_2, ..., u_m]^T$ represents the control input vectors; $f(\mathbf{x})$ and d(t) are uncertain due to modelling errors and disturbance caused by the environment; matrix $g(\mathbf{x})$ is precisely known, which represents the input dynamics. Given bounded reference command \mathbf{x}_d , the dynamic system of tracking errors can be derived from System (1) as:

$$\dot{e} = \dot{x} - \dot{x}_d = f(x) + g(x)u + d(t) - \dot{x}_d$$
(2)

Due to the Lipschitz continuity of f(x), g(x), and d(t), System (2) can be stabilizable from [31].

The goal of the paper is to conduct the optimal tracking control, such that tracking errors are limited within a prescribed region while minimizing the novel value function. Specifically, the prescribed region is denoted by the following equality:

$$-\vartheta_{li}(t) < e_i(t) < \vartheta_{ui}(t), \text{ for } i = 1, \dots, n$$
(3)

where $\vartheta_{li}(t) > 0$ and $\vartheta_{ui}(t) > 0$ (i = 1, 2, ..., n) are predefined envelope functions with specific expressions:

$$\vartheta_{li}(t) = \underline{l}\vartheta_i(t) = \underline{l}[(\vartheta_{i0} - \vartheta_{i\infty})e_i^{-a_it} + \vartheta_{i\infty}]$$
(4)

$$\vartheta_{ui}(t) = \bar{l}\vartheta_i(t) = \bar{l}[(\vartheta_{i0} - \vartheta_{i\infty})e_i^{-a_it} + \vartheta_{i\infty}]$$
(5)

with $a_i > 0$ representing the lower bound on the rate of convergence and $\vartheta_{i0} \gg \vartheta_{i\infty} > 0$, $\underline{l} > 0$, $\overline{l} > 0$. It should be noted that uncertain f(x) and d(t) can be considered as lumped

disturbances $\delta = f(x) + d(t)$. On account of the initial state $x(0) = x_0$, the following assumptions are necessary to achieve the controller design.

Assumption 1. *There exists the positive constant* Δ *satisfying* $\delta \leq \Delta$ *.*

Assumption 2. The $\vartheta_{li}(0)$ and $\vartheta_{ui}(0)$ are chosen such that (3) holds.

Remark 1. There exist many plants, such as autonomous underwater vehicles [1] and quadrotors [3], which can be modeled as System (1) with disturbance. Unlike the investigated controllers in [1] and [3] regardless of control consumption, the proposed optimal tracking controller was designed to achieve the stability of a closed-loop system with an adequate control input.

3. Results

In this section, the whole controller consisting of a feedback controller with an estimator and an optimal controller is given. Firstly, the feedback controller is derived by a USDE, which is employed for estimating lumped disturbances. Based on it, the optimal tracking controller is designed within critic-only ADP while minimizing the novel value function considering prescribed functions (4) and (5). Therefore, the whole controller is constructed as:

$$u = u_d + u^* \tag{6}$$

with u_d is the feedback controller and u^* is the optimal controller. The specific diagram of the whole controller is detailed in Figure 1.



Figure 1. The diagram of the whole controller.

3.1. Feedback Controller Design

Inspired by [12], a USDE is employed on System (2) to approximate δ . By importing it into the filter operation of tracking error and control input, the following equations can be obtained:

$$\begin{cases} k\dot{x}^{f} + x^{f} = x, \dot{x}^{f}(0) = \mathbf{0} \\ k\dot{u}^{f} + u^{f} = u, u^{f}(0) = \mathbf{0} \end{cases}$$

$$\tag{7}$$

where k is positive and can be adjusted to generate filter variables x^{t} and u^{t} .

Lemma 1. The vector $\boldsymbol{\chi} = k^{-1}(\boldsymbol{x} - \boldsymbol{x}^f) - (\boldsymbol{g}\boldsymbol{u}^f + \boldsymbol{\delta})$ is bounded such that:

$$\lim_{k \to 0} \left\{ \lim_{t \to \infty} \left[\frac{1}{k} (\boldsymbol{x} - \boldsymbol{x}^{f}) - (\boldsymbol{g} \boldsymbol{u}^{f} + \boldsymbol{\delta}) \right] \right\} = \boldsymbol{0}$$
(8)

Noting that χ represents the map form filtered variables and lumped disturbances, there exists an invariant manifold χ , which further derives:

$$\hat{\delta} = \frac{1}{k} (\mathbf{x} - \mathbf{x}^f) - g \mathbf{u}^f \tag{9}$$

where $\hat{\delta}$ is the estimation of δ . The estimation error can converge to the neighborhood of zero.

Theorem 1. For nonlinear system (1) with uncertainty, by virtue of estimation from Equation (9), the estimation errors $\tilde{\delta} = \delta - \hat{\delta}$ can converge to the following region:

$$|\widetilde{\delta}|| \le \sqrt{\|\widetilde{\delta}(0)\|^2 e^{-t/k} + k^2 \Delta^2}$$
(10)

Proof of Theorem 1. By importing the first-order filter operation on nonlinear system (3), we can derive the following form with the Laplace operator:

$$\frac{1}{ks+1}\dot{x} = \frac{1}{ks+1}g(x)u + \frac{1}{ks+1}\delta$$
(11)

Based on nonlinear system (1) and filter operation (7), it can be induced that:

$$\dot{\mathbf{x}}^{f} = \frac{1}{k}(\mathbf{x} - \mathbf{x}^{f}) = \mathbf{g}(\mathbf{x})\mathbf{u}^{f} + \delta^{f}$$
(12)

where $\delta^f = \delta/(ks+1)$.

From Equation (9), one has $\delta^f = \hat{\delta}$. Therefore, the estimation error can be expressed as:

$$\widetilde{\delta} = \delta - \delta^f = \frac{ks}{ks+1}\delta \tag{13}$$

Combined with Equations (7) and (13), we deduce the time derivative of the estimation error as: $\dot{\tilde{\delta}} = \dot{\delta} - \dot{\delta}^{f}$

$$\widetilde{\delta} = \widetilde{\delta} - \widetilde{\delta}'
= \widetilde{\delta} - \frac{\widetilde{x} - \widetilde{x}^f + k\widetilde{u}^f}{k}
= \widetilde{\delta} - \frac{\widetilde{x} - (x - x^f)/k + k\widetilde{u}^f}{k}
= \widetilde{\delta} - \frac{g(x)u + \Delta - (x - x^f)/k + k\widetilde{u}^f}{k}
= \widetilde{\delta} - \frac{\delta - \delta f}{k}$$
(14)

Then, the estimator errors can be analyzed by the following candidate function:

$$V_1 = \frac{1}{2} \widetilde{\delta}^T \widetilde{\delta} \tag{15}$$

Taking the derivative respect to time and following (), one can have:

There exist the positive constant α and Young's inequality such that:

which can further obtain:

$$V_1(t) \le -\frac{1}{k} V_1(0) e^{-t/k} + \frac{k^2 \Delta^2}{2}$$
(18)

Thus, the estimation error can converge the following region:

$$\|\widetilde{\delta}\| = \sqrt{2V_1(t)} \le \sqrt{\|\widetilde{\delta}(0)\|^2 e^{-t/k} + k^2 \Delta^2}$$
(19)

According to above discussion, the estimation error can converge to zero with given positive k, when time goes to infinity. \Box

For eliminating the impression of lumped disturbances, the feedback controller is constructed as:

$$u_d = G[\dot{x}_d - \delta - Ke] \tag{20}$$

where *G* is a generalized inverse matrix and *K* is a positive control gain. $G = (g^T g)^T g$ induces the tracking errors as:

$$\dot{e} = -Ke + g(x)u^* + \hat{\delta}$$
⁽²¹⁾

In the next section, we will consider designing the controller by defining a novel value function that balances control consumption and tracking performance.

3.2. Optimal Controller Design

To stabilize Equation (21), a value function is predefined for balancing control cost and tracking performance. In addition, the tracking error should be required to satisfy the predefined function, namely prescribed performance control. The PPC can be regarded as the constraint of the asymptotic rate and the overshoot of the tracking error. Generally, by introducing an error transformation function, the constrained dynamic of the tracking error can be converted to an equivalent unconstrained model. Different from introducing a transformation function, we first consider the tracking error directly:

$$r_s = e^T Q e \tag{22}$$

which $Q \in \mathbb{R}^{n \times n}$ is a positive definite matrix. Noting that it incorporates an actual tracking error, it depicts the distance from a desired trajectory.

In addition, the corresponding term related to control u^* is added to the value function to achieve a tradeoff between tracking performance and control consumption in a quadratic form:

$$r_u = u^{*\,l} R u^* \tag{23}$$

where positive definite matrix $\mathbf{R} \in \mathbb{R}^{m \times m}$ denotes the weight of control expanse. It is common that Equations (22) and (23) are constructed in the utility function for controller design, which can be find in [15–17].

Last but not least, we take the PPC into consideration in the value function design. With prescribed performance, the evolvement of tracking error *e* must be within the range of predefined behavior. Furthermore, the maximum overshoot should be kept away from exceeding the predetermined performance, which may cause the damage to facilities. Thus, the tracking error can be constrained within the predetermined region and satisfy:

$$r_c = -e^T Q_c C e \tag{24}$$

where $Q_c \in \mathbb{R}^{n \times n}$ is a positive definite matrix and matrix *C* is denoted as:

$$C = \operatorname{diag}\left(\ln\frac{\vartheta_{ui}(t) - e_i(t)}{\vartheta_{li}(t) + e_i(t)}\right), i = 1, 2, \dots, n$$
(25)

In view of the above analysis, the value function is predefined by Equation (25), consisting of the corresponding characterizations of the tracking error, the control, and the PPC as:

$$r = r_s + r_u + r_c \tag{26}$$

The optimal control is designed by minimizing the following value function:

$$V(\boldsymbol{e}) = \int_{t}^{\infty} r(\boldsymbol{e}(\tau), \boldsymbol{u}^{*}(\tau)) d\tau$$
(27)

where *r* is defined in Equation (26). On the basis of Equation (27), the Hamilton function can be derived as:

$$H(\boldsymbol{e},\boldsymbol{u}^*,\boldsymbol{V}) = \boldsymbol{V}_{\boldsymbol{e}}^T \left[-\boldsymbol{K}\boldsymbol{e} + \boldsymbol{g}\boldsymbol{u}^* + \widetilde{\boldsymbol{\delta}} \right] + \boldsymbol{r}_s + \boldsymbol{r}_u + \boldsymbol{r}_c$$
(28)

where V_e is a partial derivative of V. By taking the derivatives of Equations (27) and (21), the HJB equation can be derived as:

$$H(\boldsymbol{e},\boldsymbol{u},\boldsymbol{V}) = \boldsymbol{V}_{\boldsymbol{e}}^{*T}[-\boldsymbol{K}\boldsymbol{e} + \boldsymbol{g}\boldsymbol{u}^{*} + \widetilde{\boldsymbol{\delta}}] + \boldsymbol{r}_{s} + \boldsymbol{u}^{*T}\boldsymbol{R}\boldsymbol{u}^{*} + \boldsymbol{r}_{c} = 0$$
(29)

In order to learn an optimal control strategy, one can derive the following equation from $\partial H / \partial u^* = 0$:

$$u^* = -\frac{1}{2}R^{-1}g^T V_e^*$$
(30)

By substituting Equation (30) into Equation (29), the Hamilton-Jacobian-Bellman (HJB) equation can be rewritten as:

$$\boldsymbol{V}_{e}^{*T}[-\boldsymbol{K}\boldsymbol{e}+\widetilde{\boldsymbol{\delta}}]+\boldsymbol{r}_{c}+\boldsymbol{r}_{s}-\frac{1}{4}\boldsymbol{V}_{e}^{*T}\boldsymbol{g}\boldsymbol{R}\boldsymbol{g}^{T}\boldsymbol{V}_{e}^{*}=0$$
(31)

Since Equation (31) is a nonlinear equation of the optimal value function, the optimal control cannot be obtained directly by Equation (30). Referring to the core of ADP, an NN can be introduced to approximate the optimal value function and its derivative:

$$V^*(\boldsymbol{e}) = \boldsymbol{W}^T \boldsymbol{\sigma}(\boldsymbol{e}) + \boldsymbol{\varepsilon}$$
(32)

$$V_e^*(e) = \nabla \sigma(e)^T \mathbf{W} + \nabla \varepsilon$$
(33)

where *W* and σ are the ideal weight and the activation function; ε is the estimation error; $\nabla \sigma(e)$ and $\nabla \varepsilon$ are gradients of the activation function and the estimation error.

Noting that the ideal weight is unknown, the approximate optimal value function can be estimated by:

$$\hat{V} = \hat{W}^T \sigma(e) + \varepsilon \tag{34}$$

In addition, the optimal control can be derived as:

$$\hat{\boldsymbol{u}} = -\frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{g}^{T}\nabla\boldsymbol{\sigma}(\boldsymbol{e})^{T}\hat{\boldsymbol{W}}$$
(35)

Supposing $\psi = r_s + r_u + r_c$ and $\Phi = \nabla \sigma (-Ke + gu_e)$, one can yield:

$$\psi = -\mathbf{\Phi}^T \mathbf{W} - \varepsilon_{HJB} \tag{36}$$

By denoting the matrices:

$$\begin{cases} \dot{N} = -\ell N + \boldsymbol{\Phi} \boldsymbol{\Phi}^{T}, \quad N(0) = 0\\ \dot{S} = -\ell S + \boldsymbol{\Phi} \psi, \qquad S(0) = 0 \end{cases}$$
(37)

the auxiliary matrix is introduced for designing the weight law:

$$\mathbf{P} = N\hat{W} + S \tag{38}$$

From Equation (37), there holds:

$$\begin{cases} \mathbf{N}(t) = \int_0^t e^{-\ell(t-r)} \mathbf{\Phi} \mathbf{\Phi}^T dr, & \mathbf{N}(0) = 0\\ \mathbf{S}(t) = \int_0^t e^{-\ell(t-r)} \mathbf{\Phi} \psi dr, & \mathbf{S}(0) = 0 \end{cases}$$
(39)

Then, we can derive the following equations with Equation (36):

$$S = -NW + \varsigma = -NW - \int_0^t e^{-c(t-r)} \Phi \varepsilon_{HJB} dr$$
(40)

Thus, one has:

$$P = N\hat{W} + S = -N\widetilde{W} + \varsigma \tag{41}$$

where \tilde{W} is the weight estimation error, inducing the weight law:

$$\hat{W} = -\Upsilon P \tag{42}$$

where Υ represents a positive gain. The convergence of weight can be derived by the following Lyapunov function:

$$V_W = \frac{1}{2} \widetilde{W}^T \Upsilon^{-1} \widetilde{W}$$
(43)

with derivation respect to time:

$$\dot{V}_{W} = \widetilde{W}^{T} \Upsilon^{-1} \widetilde{W} = \widetilde{W}^{T} (-N\widetilde{W} + \varsigma)
\leq \left(-\gamma \|\widetilde{W}\|^{2} + \|\varsigma\| \|\widetilde{W}\| \right)
\leq -\|\widetilde{W}\| \left(\gamma \|\widetilde{W}\| - \varepsilon_{\varsigma} \right)$$
(44)

where $\gamma = \lambda_{\min}(N)$ and $\|\boldsymbol{\varsigma}\| \leq \varepsilon_{\varsigma}$.

4. Stability Analysis

Due to the above-mentioned theoretical result of the disturbance approximation error and weight error, the stability of the closed-loop system can be proven with prescribed performance.

Theorem 2. For nonlinear uncertain system (1) with constraint (3), the optimal control is designed with weight update (42), the tracking error is uniform ultimate boundedness and evolves within predefined region.

Proof of Theorem 2. According to Equations (21) and (35), the system of the tracking error can be obtained:

$$\dot{e} = -Ke + g\hat{u} + \tilde{\delta} = -Ke - \frac{1}{2}gR^{-1}g^{T}\nabla\sigma^{T}\hat{W} + \frac{1}{2}gR^{-1}g^{T}(\nabla\sigma^{T}W + \nabla\varepsilon) + gu^{*} + \tilde{\delta} = -Ke + \frac{1}{2}gR^{-1}g^{T}\nabla\sigma^{T}\widetilde{W} + \frac{1}{2}gR^{-1}g^{T}\nabla\varepsilon + gu^{*} + \tilde{\delta}$$

$$(45)$$

$$L = \frac{1}{2} \widetilde{\boldsymbol{W}}^T \widetilde{\boldsymbol{\Upsilon}}_0^{-1} \widetilde{\boldsymbol{W}} + \Gamma_1 \boldsymbol{e}^T \boldsymbol{e} + \Gamma_2 \boldsymbol{V}^* + \Gamma_3 \boldsymbol{\zeta}^T \boldsymbol{\zeta}$$

= $L_1 + L_2 + L_3$ (46)

with *V*^{*} is the optimal value function form and $\Gamma_1, \Gamma_2, \Gamma_3 > 0$. \Box

Firstly, due to the weight convergence of the critic NN, one has:

$$\dot{L}_{1} = \widetilde{\boldsymbol{W}}^{T} \mathbf{\Upsilon}^{-1} \widetilde{\boldsymbol{W}} = \widetilde{\boldsymbol{W}}^{T} \left(-N\widetilde{\boldsymbol{W}} + \boldsymbol{\varsigma} \right)
\leq \left(-\gamma \|\widetilde{\boldsymbol{W}}\|^{2} + \|\boldsymbol{\varsigma}\| \|\widetilde{\boldsymbol{W}}\| \right)
\leq -\|\widetilde{\boldsymbol{W}}\| \left(\gamma \|\widetilde{\boldsymbol{W}}\| - \|\boldsymbol{\varsigma}\| \right)
\leq -\left[\left(\gamma - \frac{1}{2\tau\Gamma_{3}} \right) \|\widetilde{\boldsymbol{W}}\|^{2} - \frac{\tau\Gamma_{3}\|\boldsymbol{\varsigma}\|^{2}}{2} \right]$$
(47)

Then,

$$\dot{L}_{2} = 2\Gamma_{1}e^{T}\dot{e} + \Gamma_{2}\dot{V}^{*}
= 2\Gamma_{1}e^{T}\dot{e} + \Gamma_{2}(-e^{T}Qe - u^{*T}Ru^{*} - r_{c})
= 2\Gamma_{1}e^{T}\left[-Ke + \frac{1}{2}gR^{-1}g^{T}\nabla\sigma^{T}\widetilde{W} + \frac{1}{2}gR^{-1}g^{T}\nabla\varepsilon + gu^{*} + \widetilde{\delta}\right]
+\Gamma_{2}(-e^{T}Qe - u^{*T}Ru^{*} - r_{c})
\leq -\left[2\lambda_{\min}(K)\Gamma_{1} + \Gamma_{2}\lambda_{\min}(Q) - \left(\|gR^{-1}g^{T}\nabla\sigma^{T}\| + \|gR^{-1}g^{T}\| + 2\right)\Gamma_{1}\right]\|e\|^{2}
+ \frac{1}{4}\Gamma_{1}\|gR^{-1}g^{T}\|\|\nabla\varepsilon\|^{2} + \frac{1}{4}\Gamma_{1}\|gR^{-1}g^{T}\nabla\sigma^{T}\|\|\widetilde{W}\|^{2} + \Gamma_{1}\|\widetilde{\delta}\|^{2} - \Gamma_{2}r_{c}
- \left(\Gamma_{2}\lambda_{\min}(R) - \Gamma_{1}\|g\|^{2}\right)\|u^{*}\|^{2}$$
(48)

In addition,

$$L_{3} = 2\Gamma_{3}\varsigma^{T}\dot{\varsigma}$$

$$= 2\Gamma_{3}\varsigma^{T}(-\ell\varsigma + \Phi\varepsilon_{HJB})$$

$$= 2\Gamma_{3}\varsigma^{T}\left\{-\ell\varsigma + \Phi\left[\left(W^{T}\nabla\sigma + \nabla\varepsilon\right)^{T}\tilde{\delta} + \nabla\varepsilon^{T}\left(-\frac{1}{2}gR^{-1}g^{T}\nabla\sigma^{T}\hat{W}\right) - \nabla\varepsilon^{T}Ke\right]\right\}$$

$$= -\Gamma_{3}(2\ell - 3\tau)\|\varsigma\|^{2} + \frac{\Gamma_{3}}{\tau}\|\Phi\left(\nabla\sigma^{T}W + \nabla\varepsilon\right)^{T}\|^{2}\|\tilde{\delta}\|^{2} + \frac{\Gamma_{3}}{\tau}\|\Phi\nabla\varepsilon^{T}K\|^{2}\|e\|^{2}$$

$$+ \frac{\Gamma_{3}}{4\tau}\|\Phi\nabla\varepsilon^{T}gR^{-1}g^{T}\nabla\sigma^{T}\hat{W}\|^{2}$$
(49)

Lastly, the stability of the whole system can be derived as:

$$L = L_{1} + L_{2} + L_{3}$$

$$\leq -\left[\left(\gamma - \frac{1}{2\tau\Gamma_{3}}\right) - \frac{1}{4}\Gamma_{1}\|gR^{-1}g^{T}\nabla\sigma^{T}\|\right]\|\widetilde{W}\|^{2} - \left(\Gamma_{2}\lambda_{\min}(R) - \Gamma_{1}\|g\|^{2}\right)\|u_{e}^{*}\|^{2}$$

$$-\left[2\lambda_{\min}(K)\Gamma_{1} + \Gamma_{2}\lambda_{\min}(Q) - \left(\|gR^{-1}g^{T}\nabla\sigma^{T}\| + \|gR^{-1}g^{T}\| + 2\right)\Gamma_{1} - \frac{\Gamma_{3}}{\tau}\|\Phi\nabla\varepsilon^{T}K\|^{2}\right]\|e\|^{2}$$

$$-\Gamma_{3}\left(2\ell - \frac{7}{2}\tau\right)\|\varsigma\|^{2} + \frac{1}{4}\Gamma_{1}\|gR^{-1}g^{T}\|\|\nabla\varepsilon\|^{2} + \frac{\Gamma_{3}}{4\tau}\|\Phi\nabla\varepsilon^{T}gR^{-1}g^{T}\nabla\sigma^{T}\hat{W}\|^{2}$$

$$+ \left(\Gamma_{1} + \frac{\Gamma_{3}}{\tau}\|\Phi(W^{T}\nabla\sigma + \nabla\varepsilon)^{T}\|^{2}\right)\|\widetilde{\delta}\|^{2} - \Gamma_{2}r_{c}$$
(50)

If $\Gamma_2 \lambda_{\min}(\mathbf{R}) - \Gamma_1 \|\mathbf{g}\|^2 > 0$, Equation (50) can be written in the following form:

$$\dot{L} \le -c_1 \|\widetilde{\mathbf{W}}\|^2 - c_2 \|\mathbf{e}\|^2 - c_3 \|\mathbf{\varsigma}\|^2 + \rho$$
(51)

where

$$c_{1} = \gamma - \frac{1}{2\tau\Gamma_{3}} - \frac{1}{4}\Gamma_{1} \|gR^{-1}g^{T}\nabla\sigma^{T}\|$$

$$c_{2} = 2\lambda_{\min}(K)\Gamma_{1} + \Gamma_{2}\lambda_{\min}(Q) - \left(\|gR^{-1}g^{T}\nabla\sigma^{T}\| + \|gR^{-1}g^{T}\| + 2\right)\Gamma_{1} - \frac{\Gamma_{3}}{\tau}\|\Phi\nabla\varepsilon^{T}K\|^{2}$$

$$c_{3} = \Gamma_{3}\left(2\ell - \frac{7}{2}\tau\right)$$

$$\rho = \frac{1}{4}\Gamma_{1}\|gR^{-1}g^{T}\|\|\nabla\varepsilon\|^{2} + \frac{\Gamma_{3}}{4\tau}\|\Phi\nabla\varepsilon^{T}gR^{-1}g^{T}\nabla\sigma^{T}\hat{W}\|^{2} + \left(\Gamma_{1} + \frac{\Gamma_{3}}{\tau}\|\Phi(\nabla\sigma^{T}W + \nabla\varepsilon)\|^{2}\right)\|\tilde{\delta}\|^{2} - \Gamma_{2}r_{c}$$

Noting that r_c is bounded, the boundedness of ρ can be derived along the boundness of disturbance estimation errors, the weight estimation error, and the HJB error. Thus, the parameters of the controller design should satisfy:

$$\begin{split} &\Gamma_{1} < \frac{4\gamma}{\|gR^{-1}g^{T}\nabla\sigma^{T}\|} \\ &\tau > \max\left(\frac{1}{\Gamma_{3}(2\gamma - (1/2)\Gamma_{2}\|gR^{-1}g^{T}\nabla\sigma^{T}\|)}, \frac{\Gamma_{3}\|\Phi\nabla\varepsilon^{T}K\|^{2}}{2\lambda_{\min}(K)\Gamma_{2}}\right) \\ &\Gamma_{2} > \max\left(\frac{-2\lambda_{\min}(K)\Gamma_{1} + (\Gamma_{3}/\tau)\|\Phi\nabla\varepsilon K\|^{2} - (\|gR^{-1}g^{T}\nabla\sigma^{T}\| + \|gR^{-1}g^{T}\| + 2)\Gamma_{1}}{\lambda_{\min}(Q)}, \frac{\Gamma_{1}\|g\|^{2}}{\lambda_{\min}(R)}\right) \end{split}$$

$$\lambda_{\min}(\mathbf{K}) > \frac{\|\mathbf{g}\mathbf{R}^{-1}\mathbf{g}^T\| + \|\mathbf{g}\mathbf{R}^{-1}\mathbf{g}^T\nabla\sigma^T\| + 2}{2}$$

$$\ell > \frac{7}{4}\tau, \Gamma_3 > 0$$

such that $c_1, c_2, c_3 > 0$, resulting in the uniform ultimate boundedness of the tracking error. For arbitrary $|r_c| \le \varepsilon_r(\varepsilon_r > 0)$, there exists:

$$\left|\ln\frac{\vartheta_{ui}(t) - e_i(t)}{\vartheta_{li}(t) + e_i(t)}\right| \le \varepsilon_r' \tag{52}$$

which can indicate the PPC is be violated.

5. Simulations

In this section, we implemented the following examples to demonstrate the effectiveness and superiority of the investigated controller scheme. Without loss of generality, the sampling period was 5 ms, and solver ode4 was selected during the simulations.

Example 1. In order to verify the design of the optimal tracking controller with preset performance, a second-order nonlinear system is considered as:

$$\begin{cases} \dot{x}_1 = -x_1 + x_2 \\ \dot{x}_2 = -0.5x_1 - 0.5x_2(1 - (\cos(2x_1) + 2)^2) + (\cos(2x_1) + 2)u \end{cases}$$
(53)

The given track commands are $x_1^d = \sin(t)$ *and* $x_2^d = \cos(t) + \sin(t)$ *, and the tracking error should meet the following constraints:*

$$-\vartheta_{li}(t) < e_i(t) < \vartheta_{ui}(t), \text{ for } i = 1,2$$
(54)

where $\underline{l}^1 = 2$, $\overline{l}^1 = 3$, $\vartheta_{10} = 3$, $\vartheta_{1\infty} = 0.2$, $a_1 = 3$ and $\underline{l}^2 = 1.8$, $\overline{l}^2 = 1.5$, $\vartheta_{20} = 3$, $\vartheta_{2\infty} = 0.25$, $a_2 = 2$. The initial state is $\mathbf{x} = [3, -1]^T$. For verifying the devised controller, we assume that:

$$f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} -x_1 + x_2 \\ -0.5x_1 - 0.5x_2(1 - (\cos(2x_1) + 2)^2) \end{bmatrix}$$
(55)

which is uncertain, and disturbance $d(t) = [0; \sin(2t)]^T$ is added.

In order to achieve the control objective, the proposed optimal tracking controller is realized based on a feedback control and an optimal control. The feedback control is based on a USDE to compensate the influence of disturbance on the system. The optimal regulation law is to minimize value function (26) under the ADP framework, where Q, R, and Q_c are the identity matrices of the respective dimensions. In order to approximate the optimal value function, the activation function was selected as $\sigma = [e_1^2, e_1e_2, e_2^2]^T$, and other simulation parameters were selected as K = diag[0.5, 0.5], $\Gamma = 2I$, l = 5, and k = 0.01. The initial weight was $W_0 = [0, 0, 0]^T$.

The simulation results for a period of 20 s are shown from Figures 2–8. Figure 2 gives the simulation results of the system state and the reference command, which indicates that the state can accurately track the reference command within 4 s. The tracking errors enter the steady state quickly, which can be due to the estimation from the USDE, as shown in Figure 3. Lumped disturbances can be approximated precisely within 0.1 s.

The control inputs of the proposed controller are described in Figure 4, which include the feedback control and the optimal control. Figure 5 describes the simulation results of the tracking error and asymmetric output performance constraints. It can be found that the tracking errors converged to a specified asymmetric envelope, which indicated that the designed controller with an optimal preset performance can make tracking errors converge to zero within a predetermined convergence rate. The stability of the closed-loop system hinged on the weight of the critic NN, as illustrated in Figure 6, which showed the approximation of weight can converge to the ideal weight within 2 s.



Figure 2. The tracking performance.



Figure 3. The estimation errors of lumped disturbances.



Figure 4. The control inputs.



Figure 5. The tracking errors with the PPC.



Figure 6. The weight of the critic NN (Example 1).



Figure 7. The tracking errors with the WPPC.



Figure 8. Comparison of cost functions.

In order to illustrate the superiority of the proposed controller in achieving preset performance, it was compared with the optimal controller based on the traditional value function in [32], which is named WPPC. Noting that the controller in [32] is based on a fixed-time disturbance observer, we set the feedback controller based on a USDE for the sake of fair comparison. Figure 7 shows the simulation results of the tracking error and the PPC envelope. Although the system state tracked the reference command successfully, the tracking errors could not evolve within the envelope, and the transient performance could not be guaranteed to meet the output constraints. This indicates that the convergence rate of the tracking error was slower than that of the specified convergence rate and the predetermined performance could not be guaranteed. Figure 8 compares the value functions of the two controllers for 20 s. In contrast, the value function of the new controller decreased by 6% compared to the optimal controller without taking into preset performance account. Therefore, the new controller can ensure that the tracking error meets the output performance constraints while its value function is smaller.

Example 2. Noting that the trajectory tracking problem of the quadrotor can be affected by uncertain dynamic drift and disturbance induced by wind, the effectiveness of the proposed control is carried on a quadrotor [3,31]. Noting that position and attitude loops are considered in [3,31], the position dynamic of the quadrotor is considered as:

$$\begin{cases} \dot{\boldsymbol{p}} = \boldsymbol{v} \\ \dot{\boldsymbol{v}} = (-g\boldsymbol{h}_3 - \prod \boldsymbol{v})/\boldsymbol{m} + \boldsymbol{u} + \boldsymbol{d}_{\boldsymbol{v}} \end{cases}$$
(56)

where parameters of the model are listed in Table 1. The given reference trajectory is $p^d = [10(1 - \cos(0.1\pi t)), 5\sin(0.2\pi t), 9(1 - e^{-0.3t})]^T$. To carry out the proposed controller on Equation (53), we reformulate it as:

$$\dot{x} = f(x) + g(x)u + d(t) \tag{57}$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{p} \\ \mathbf{v} \end{bmatrix}, f(\mathbf{x}) = \begin{bmatrix} \mathbf{v} \\ (-g\mathbf{h}_3 - \prod \mathbf{v})/m \end{bmatrix}, g(\mathbf{x}) = \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}$$
(58)

For approximating the optimal control, the critic NN is structured with 15 neurons with [-5,5]. For the initial position and weight, we choose the following parameters: $p(0) = [-5,3,2]^T$ and $W_0 = 0$. The other controller parameters are listed in Table 2.

Table 1. Parameters of the quadrotor attitude dynamics.

Section	Values
Mass	2
Inertia moment matrix	diag([0.01, 0.01, 0.01])
Disturbances	$[\sin(4t) + \cos(2t) - \sin(t); \cos(4t) + \sin(2t) - \cos(t); \sin(3t)\cos(2t) - \cos(t)]^T$

Section	Values
USDE	k = 0.01
РРС	$\underline{l}^{px} = 2.5, \overline{l}^{px} = 1.5, \vartheta_0^{px} = 5, \vartheta_{\infty}^{px} = 0.2, a^{px} = 2$
	$\underline{l}^{py} = 2.5, \overline{l}^{py} = 3, \vartheta_0^{px} = 5, \vartheta_\infty^{px} = 0.2, a^{px} = 3$
	$\underline{l}^{pz} = 2.5, \overline{l}^{pz} = 2, \vartheta_0^{pz} = 5, \vartheta_\infty^{pz} = 0.2, a^{pz} = 3$
	$\underline{l}^{vx} = 2, \overline{l}^{vx} = 1.5, \vartheta_0^{vx} = 8, \vartheta_{\infty}^{vx} = 0.2, a^{vx} = 3$
	$\underline{l}^{vy} = 1.3, \overline{l}^{vy} = 1.5, \vartheta_0^{vy} = 10, \vartheta_{\infty}^{vy} = 0.2, a^{vy} = 3$
	$\underline{l}^{vz} = 0.6, \ \overline{l}^{vz} = 0.8, \ \vartheta_0^{vz} = 8, \ \vartheta_\infty^{vz} = 0.2, \ a^{vz} = 3$
Control gain	K = diag[2, 2, 2, 2, 2, 2]

Table 2. Parameters of the controllers.

The performance of the proposed controller was discussed as follows. Figures 9 and 10 show the tracking performances of the position and the velocity. It can be found that tracking errors evolved the predefined constraints. By running the novel value function, tracking errors converged to zero within the convergence rate. During the controller design, the weight of the critic NN is crucial for solving the HJB equation. Figure 11 gives the convergence of weight, which is essential for the stability of the closed-loop system. To verify the advantage of the novel value function, the optimal tracking controller in [32] was executed for System (54). In the position loop, the tracking errors could not evolve the predefined envelope by minimizing the value function without the PPC constraints. Therefore, the effectiveness of the proposed controller and the optimal tracking controller in [32] are based on simulation, we will further consider experimental validation based on physical systems to verify the feasibility of the proposed method, where complete response associated is impossible.



Figure 9. The tracking errors of the position.



Figure 10. The tracking errors of the velocity.



Figure 11. The weight of the critic NN (Example 2).



Figure 12. The tracking errors of the position [32].



Figure 13. The tracking errors of the velocity [32].

6. Conclusions

In this paper, an optimal tracking control with output constraints was proposed to ensure tracking errors with prescribed performance. The perturbation estimation error was introduced into critic-only ADP, and the novel value function was run within a critic-only ADP framework to derive the optimal control. It proved that the tracking error can be realized in the specified envelope of the PPC. The numerical simulation verified that the designed controller can achieve tracking errors with prescribed performance on the quadrotor trajectory tracking problem and a class of second-order nonlinear systems. However, actuator saturation may limit the control input, which makes it difficult to guarantee prescribed performance. We will focus on the optimal tracking controller for uncertain nonlinear systems with prescribed performance under actuator saturation.

Author Contributions: Conceptualization, Y.G. and Z.L.; software, Y.G.; validation, Y.G.; writing—original draft preparation, Y.G.; supervision, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the first author.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Vu, Q.V.; Dinh, T.A.; Nguyen, T.V.; Tran, H.V.; Le, H.X.; Pham, H.V.; Kim, T.D.; Nguyen, L. An adaptive hierarchical sliding mode controller for autonomous underwater vehicles. *Electronics* **2021**, *10*, 2316. [CrossRef]
- Liang, L.; Liu, H.; Li, X.; Zhu, X.; Lan, B.; Liu, Y.; Wang, X. Model-based coordinated trajectory tracking control of skid-steer mobile robot with timing-belt servo system. *Electronics* 2023, 12, 699. [CrossRef]
- 3. Shao, X.; Yue, X.; Li, J. Event-triggered robust control for quadrotors with preassigned time performance constraints. *Appl. Math. Comput.* **2021**, 392, 125667. [CrossRef]
- Li, S.; Ahn, C.K.; Guo, J.; Xiang, Z. Neural-network approximation-based adaptive periodic event-triggered output-feedback control of switched nonlinear systems. *IEEE Trans. Cybern.* 2020, *51*, 4011–4020. [CrossRef] [PubMed]
- 5. Li, Y.; Li, K.; Tong, S. Adaptive neural network finite-time control for multi-input and multi-output nonlinear systems with positive powers of odd rational numbers. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 2532–2543. [CrossRef]
- Tatlicioglu, E.; Yilmaz, B.M.; Savran, A.; Alci, M. Adaptive fuzzy logic with self-adjusting membership functions based tracking control of surface vessels. *Ocean. Eng.* 2022, 253, 11129. [CrossRef]
- Tang, J.; Dang, Z.; Deng, Z.; Li, C. Adaptive fuzzy nonlinear integral sliding mode control for unmanned underwater vehicles based on ESO. *Ocean. Eng.* 2022, 266, 113154. [CrossRef]
- Li, Y.; Li, K.; Tong, S. An observer-based fuzzy adaptive consensus control method for nonlinear multiagent systems. *IEEE Trans. Fuzzy Syst.* 2022, 30, 4667–4678. [CrossRef]

- Chen, M.; Xiong, S.; Wu, Q. Tracking flight control of quadrotor based on disturbance observer. *IEEE Trans. Syst. Man Cybern.* 2019, 51, 1414–1423. [CrossRef]
- Guo, K.; Shi, P.; Wang, P.; He, C.; Zhang, H. Non-singular terminal sliding mode controller with nonlinear disturbance observer for robotic manipulator. *Electronics* 2023, 12, 849. [CrossRef]
- Kukurowski, N.; Pazera, M.; Witczak, M. Fault-tolerant tracking control for a descriptor system under an unknown input disturbances. *Electronics* 2021, 10, 2247. [CrossRef]
- 12. Huang, Y.; Wu, J.; Na, J.; Han, S.; Gao, G. Unknown system dynamics estimator for active vehicle suspension control systems with time-varying delay. *IEEE Trans. Cybern.* 2021, 52, 8504–8514. [CrossRef] [PubMed]
- Na, J.; Jing, B.; Huang, Y.; Gao, G.; Zhang, C. Unknown system dynamics estimator for motion control of nonlinear robotic systems. *IEEE Trans. Ind. Electron.* 2019, 67, 3850–3859. [CrossRef]
- 14. Wang, S.; Tao, L.; Chen, Q.; Na, J.; Ren, X. USDE-based sliding mode control for servo mechanisms with unknown system dynamics. *IEEE/ASME Trans. Mech.* 2020, 25, 1056–1066. [CrossRef]
- 15. Khodamipour, G.; Khorashadizadeh, S.; Farshad, M. Adaptive formation control of leader-follower mobile robots using reinforcement learning and the fourier series expansion. *ISA Trans.* **2023**, *in press.* [CrossRef]
- 16. Bao, C.; Wang, P.; He, R.; Tang, G. Observer-based optimal control method combination with event-triggered strategy for hypersonic morphing vehicle. *Aerosp. Sci. Technol.* **2023**, *136*, 108219. [CrossRef]
- Hua, H.; Fang, Y. A novel reinforcement learning-based robust control strategy for a quadrotor. *IEEE Trans. Ind. Electron.* 2022, 70, 2812–2821. [CrossRef]
- 18. Bechlioulis, C.P.; Rovithakis, G.A. Robust adaptive control of feedback linearizable MIMO nonlinear systems with prescribed performance. *IEEE Trans. Autom. Control* 2008, 53, 2090–2099. [CrossRef]
- 19. Wang, N.; Gao, Y.; Zhang, X. Data-driven performance-prescribed reinforcement learning control of an unmanned surface vehicle. *IEEE Trans. Neural Netw. Learn. Syst.* 2021, 32, 5456–5467. [CrossRef]
- 20. Liu, H.; Cheng, Q.; Xiao, J.; Hao, L. Data-driven optimal tracking control for SMA actuated systems with prescribed performance via reinforcement learning. *Mech. Syst. Signal Process.* **2022**, *177*, 109191. [CrossRef]
- Chen, H.; Yan, H.; Wang, Y.; Xie, S.; Zhang, D. Reinforcement learning-based close formation control for underactuated surface vehicle with prescribed performance and time-varying state constraints. *Ocean. Eng.* 2022, 256, 111361. [CrossRef]
- Wang, X.; Wang, Q.; Sun, C. Prescribed performance fault-tolerant control for uncertain nonlinear MIMO system using actor–critic learning structure. *IEEE Trans. Neural Netw. Learn. Syst.* 2021, 33, 4479–4490. [CrossRef] [PubMed]
- Zhang, S.; Huang, C.; Ji, K.; Zhang, H. Prescribed performance incremental adaptive optimal fault-tolerant control for nonlinear systems with actuator faults. *ISA Trans.* 2022, 120, 99–109. [CrossRef] [PubMed]
- Li, D.; Dong, J. Performance-constrained fault-tolerant DSC based on reinforcement learning for nonlinear systems with uncertain parameters. *Appl. Math. Comput.* 2023, 443, 127759. [CrossRef]
- 25. Liu, G.; Sun, N.; Yang, T.; Fang, Y. Reinforcement learning-based prescribed performance motion control of pneumatic muscle actuated robotic arms with measurement noises. *IEEE Trans. Syst. Man Cybern.* **2022**, *in press.* [CrossRef]
- 26. Yan, L.; Liu, Z.; Chen, C.L.P.; Zhang, Y.; Wu, Z. Optimized adaptive consensus control for multi-agent systems with prescribed performance. *Inf. Sci.* **2022**, *613*, 649–666. [CrossRef]
- 27. Ouyang, Y.; Sun, C.; Dong, L. Actor-critic learning based coordinated control for a dual-arm robot with prescribed performance and unknown backlash-like hysteresis. *ISA Trans.* **2022**, *126*, 1–13. [CrossRef]
- Luo, A.; Xiao, W.; Li, X.M.; Yao, D.; Zhou, Q. Performance-guaranteed containment control for pure-feedback multi agent systems via reinforcement learning algorithm. *Int. J. Robust Nonlinear Control.* 2022, 32, 10180–10200. [CrossRef]
- 29. Peng, Z.; Ji, H.; Zou, C.; Kuang, Y.; Cheng, H.; Shi, K.; Ghosh, B. Optimal H tracking control of nonlinear systems with zero-equilibrium-free via novel adaptive critic designs. *Neural Netw.* **2023**, *164*, 105–114. [CrossRef]
- Huo, Y.; Wang, D.; Qiao, J.; Li, M. Adaptive critic design for nonlinear multi-player zero-sum games with unknown dynamics and control constraints. *Nonlinear Dyn.* 2023, 111, 11671–11683. [CrossRef]
- 31. Vamvoudakis, K.; Lewis, F. Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* **2010**, *46*, 878–888. [CrossRef]
- 32. Liu, H.; Li, B.; Xiao, B.; Ran, D.; Zhang, C. Reinforcement-learning-based tracking control for a quadrotor unmanned aerial vehicle under external disturbances. *Int. J. Robust Nonlinear Control.* **2022**, *in press*. [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.