

Research Article

Open Access



Decentralized tracking control design based on intelligent critic for an interconnected spring-mass-damper system

Wenqian Fan^{1,2,3,4}, Ao Liu^{1,2,3,4}, Ding Wang^{1,2,3,4}

¹Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China.

²Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, China.

³Beijing Institute of Artificial Intelligence, Beijing University of Technology, Beijing 100124, China.

⁴Beijing Laboratory of Smart Environmental Protection, Beijing University of Technology, Beijing 100124, China.

Correspondence to: Prof. Ding Wang, Faculty of Information Technology, Beijing University of Technology, 100 Pingleyuan, Chaoyang District, Beijing 100124, China. E-mail: dingwang@bjut.edu.cn

How to cite this article: Fan W, Liu A, Wang D. Decentralized tracking control design based on intelligent critic for an interconnected spring-mass-damper system. *Complex Eng Syst* 2023;3:5. <http://dx.doi.org/10.20517/ces.2023.04>

Received: 17 Feb 2023 **First Decision:** 6 Mar 2023 **Revised:** 14 Mar 2023 **Accepted:** 20 Mar 2023 **Published:** 30 Mar 2023

Academic Editor: Yurong Liu **Copy Editor:** Fanglin Lan **Production Editor:** Fanglin Lan

Abstract

In this paper, the decentralized tracking control (DTC) problem is investigated for a class of continuous-time nonlinear systems with external disturbances. First, the DTC problem is resolved by converting it into the optimal tracking controller design for augmented tracking isolated subsystems (ATISs). A cost function with a discount is taken into consideration. Then, in the case of external disturbances, the DTC scheme is effectively constructed via adding the appropriate feedback gain to each ATIS. In addition, utilizing the approximation property of the neural network, the critic network is constructed to solve the Hamilton-Jacobi-Isaacs equation, which can derive the optimal tracking control law and the worst disturbance law. Moreover, the updating rule is improved during the process of weight learning, which removes the requirement for initial admission control. Finally, through the interconnected spring-mass-damper system, a simulation example is given to verify the availability of the DTC scheme.

Keywords: Adaptive dynamic programming, discounted cost function, decentralized tracking control, disturbance rejection, interconnected spring-mass-damper systems, neural networks, optimal control



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



1. INTRODUCTION

For large-scale nonlinear interconnected systems, which are considered as nonlinear plants consisting of many interconnected subsystems, decentralized control has become a research hotspot in the last few decades^[1–4]. Compared with the centralized control, the decentralized control has the advantages of simplifying the structure and reducing the computation burden of the controller. Besides, the local controller only depends on the information of the local subsystem. Meanwhile, with the development of science and technology, interconnected engineering applications have become increasingly complex, such as robotic systems^[5] and power systems^[6,7]. In^[8–10], we found that the decentralized control of the large-scale system was connected with the optimal control of the isolated subsystems, which means the optimal control method^[11–14] can be adopted to achieve the design purpose of the decentralized controllers. However, the optimal control of the nonlinear system often needs to solve the Hamilton-Jacobi-Bellman (HJB) or Hamilton-Jacobi-Isaacs (HJI) equation, which can be solved by using the adaptive dynamic programming (ADP) method^[15,16]. Besides, in^[13], Wang *et al.* investigated the latest intelligent critic framework for advanced optimal control. In^[14], the optimal feedback stabilization problem was discussed with discounted guaranteed cost for nonlinear systems. It follows that the interconnection plays a significant role in designing the controller. Hence, it can be classified as decentralized and distributed control schemes. There is a certain distinction between decentralized control and distributed control. For decentralized control, each sub-controller only uses local information and the interconnection among subsystems can be assumed to be weak in nature. Compared with the decentralized control, the distributed control^[17–19] can be introduced to improve the performance of the subsystems when the interconnections among subsystems become strong. In^[20], the distributed optimal observer was devised to assess the nonlinear leader state for all followers. In^[21], the distributed control was developed by means of online reinforcement learning for interconnected systems with exploration.

It is worth mentioning that the ADP algorithm has been extensively employed for dealing with various optimal regulation problems and tracking problems^[22–24], which will achieve the goal, that is, the actual signal can track the reference signal under the noisy and the uncertain environment. In^[25], Ha *et al.* proposed a novel cost function to explore the evaluation framework of the optimal tracking control problem. Then, aimed at complicated control systems, it is necessary to consider decentralized tracking control (DTC) problems^[26–29]. The DTC systems can be transformed into the nominal augmented tracking isolated subsystems (ATISs), which are composed of the tracking error and the reference signal. In^[26], Qu *et al.* proposed a novel formulation consisting of a steady-state controller and a modified optimal feedback controller of the DTC strategy. Besides, the asymptotic DTC was realized by introducing two integral bounded functions in^[27]. In^[28], Liu *et al.* proposed a finite-time DTC method for a class of nonstrict feedback interconnected systems with disturbances. Moreover, the adaptive fuzzy output-feedback DTC design was investigated for switched large-scale systems in^[29].

Game theory is a discipline that implements corresponding strategies. It contains cooperative and noncooperative types, that is, zero-sum (ZS) games and non-ZS games. In particular, ZS games have been widely applied in many fields^[30–33]. The object of the ZS game is to derive the Nash equilibrium of nonlinear systems, which makes the cost function optimized. In^[31], the finite-horizon H-infinity state estimator design was studied for periodic neural networks over multiple fading channels. The noncooperative control problem was formulated as a two-player ZS game in^[32]. In^[33], Wang *et al.* investigated the stability of the general value iteration algorithm for ZS games. At the same time, we can also combine the ZS problem with the tracking problem to make the system more stable while achieving the trajectory tracking. In^[34], Zhang *et al.* developed an online model-free integral reinforcement learning algorithm for solving the H-infinity optimal tracking problem for completely unknown systems. In^[35], a general bounded L_2 gain tracking policy was introduced with a discounted function. In^[36], Hou *et al.* proposed an action-disturbance-critic neural network frame to realize the iterative dual heuristic programming algorithm.

As can be seen from the above, there are few studies that combine the DTC problem with the ZS game problem. It is necessary to take the related discounted cost function into account for the DTC system, which can transform the DTC problem into an optimal control problem with disturbances. In practice, the existence of disturbances will make an unpredictable impact on the plant. Hence, it is of vital importance to consider the stability of the DTC system. In the experimental simulation, it is a challenge to achieve the goal of effective online weight training, which is implemented under the tracking control law and the disturbance control law. Consequently, in this paper, we put forward a novel method in view of ADP to resolve the DTC problem with external disturbances for continuous-time (CT) nonlinear systems. More importantly, for the sake of overcoming the difficulty of selecting initial admissible control policies, an additional term is added during the weight updating process. Remarkably, in this paper, we introduce the discount factor for maximizing and minimizing the corresponding cost function.

The contributions of this paper are as follows: First, considering the disturbance input in the DTC system, the strategy feasibility and the system stability are discussed through theoretical proofs. It is worth noting that the discount factor is introduced to the cost function. Moreover, in the process of online weight training, we can make the DTC system reach a stable state without selecting the initial admissible control law. Additionally, we present the experimental process of the spring-mass-damper system. Besides, we derive the desired tracking error curves as well as control strategy curves, which demonstrates that they are uniformly ultimately bounded (UUB).

The whole paper is divided into six sections. The first section is the introduction of relevant background knowledges of the research content. The second section is the problem statement of basic problems about the two person ZS game and the DTC strategy. In the third section, we design the decentralized tracking controller by using the optimal control method through solving the HJI equations. Meanwhile, the relevant lemma and theorem are given to validate the establishment of the DTC strategy. In the fourth section, the design method in accordance with adaptive critic is elaborated. Most importantly, an improved critic learning rule is implemented via critic networks. In the fifth section, the practicability of this method is validated by an interconnected spring-mass-damper system. Finally, the sixth section displays conclusions and summarizes overall research content of the whole paper.

2. PROBLEM STATEMENT

Consider a CT nonlinear interconnected system with disturbances, which is composed of N interconnected subsystems. Its dynamic description can be expressed as

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t)) (\bar{u}_i(x_i(t)) + \bar{Z}_i(x(t))) + h_i(x_i(t))v_i(x_i(t)), \tag{1}$$

where $i = 1, 2, \dots, N$, $x_i(t) \in \mathbb{R}^{n_i}$ is the state vector of the i th subsystem and $x(t)$ denotes the partial interconnected state related to other subsystems of the large-scale system. $\bar{u}_i(x_i(t)) \in \mathbb{R}^{m_i}$ is the control input and $v_i(x_i(t)) \in \mathbb{R}^{q_i}$ is the external disturbance input. As for the i th subsystem, we denote $f_i(x_i(t))$, $g_i(x_i(t))$, $h_i(x_i(t))$, and $\bar{Z}_i(x(t))$ as the nonlinear internal dynamics, the input gain matrix, the disturbance gain matrix, and the interconnected item in sequence. Besides, $[x_1^T, x_2^T, \dots, x_N^T]^T \in \mathbb{R}^n$ denotes the whole state of the large-scale system Equation (1), where $n = \sum_{i=1}^N n_i$. Accordingly, x_1, x_2, \dots, x_N are named local states and $\bar{u}_1(x_1), \bar{u}_2(x_2), \dots, \bar{u}_N(x_N)$ are called local controllers. We let $R_i \in \mathbb{R}^{m_i \times m_i}$ be the symmetric positive definite matrix and denote $Z_i(x(t)) = R_i^{1/2} \bar{Z}_i(x(t))$. In addition, $Z_i(x(t)) \in \mathbb{R}^{m_i}$ is bounded as follows:

$$\|Z_i(x(t))\| \leq \sum_{j=1}^N \alpha_{ij} \theta_{ij}(x_j) \leq \sum_{j=1}^N \beta_{ij} \theta_j(x_j), \tag{2}$$

where $j = 1, 2, \dots, N$, α_{ij} is the nonnegative constant, $\theta_{ij}(x_j)$ is the positive semidefinite function. Besides, we define $\theta_j(x_j) = \max \{\theta_{1j}(x_j), \theta_{2j}(x_j), \dots, \theta_{Nj}(x_j)\}$ and the element of $\{\theta_{1j}(x_j), \theta_{2j}(x_j), \dots, \theta_{Nj}(x_j)\}$

will not reach zero at the same time. For this reason, $\beta_{ij} \geq \alpha_{ij}\theta_{ij}(x_j)/\theta_j(x_j)$ holds, where β_{ij} is also the nonnegative constant.

In this paper, considering the nonlinear system [Equation \(1\)](#), a reference system is introduced as follows:

$$\dot{r}_i(t) = \zeta_i(r_i(t)), \quad (3)$$

where $r_i(t) \in \mathbb{R}^{n_i}$ denotes the desired trajectory with $r_i(0) = r_{i0}$, the function ζ_i is locally Lipschitz continuous satisfying $\zeta_i(0) = 0$. For the i th subsystem, the trajectory tracking error can be defined as $e_i(t) = x_i(t) - r_i(t)$ with $e_i(0) = e_{i0}$. Thus, the dynamics of the tracking error is

$$\dot{e}_i(t) = f_i(x_i(t)) + g_i(x_i(t)) (\bar{u}_i(x_i(t)) + \bar{Z}_i(x_i(t))) + h_i(x_i(t))v_i(x_i(t)) - \zeta_i(r_i(t)). \quad (4)$$

Noticing $x_i(t) = e_i(t) + r_i(t)$, we define the augmented subsystem states as $y_i(t) = [e_i^\top(t), r_i^\top(t)]^\top \in \mathbb{R}^{2n_i}$ with $y_i(0) = y_{i0} = [e_{i0}^\top, r_{i0}^\top]^\top$. Hence, the dynamic of the i th ATIS based on [Equations \(1\) and \(3\)](#) can be formulated as a concise form

$$\dot{y}_i(t) = \mathcal{F}_i(y_i(t)) + \mathcal{G}_i(y_i(t)) (\bar{u}_i(y_i(t)) + \bar{Z}_i(y_i(t))) + \mathcal{H}_i(y_i(t))v_i(y_i(t)), \quad (5)$$

where $\mathcal{F}_i(y_i(t)) \in \mathbb{R}^{2n_i}$, $\mathcal{G}_i(y_i(t)) \in \mathbb{R}^{2n_i \times m_i}$, and $\mathcal{H}_i(y_i(t)) \in \mathbb{R}^{2n_i \times q_i}$ respectively. Specifically, they can be expressed as

$$\mathcal{F}_i(y_i(t)) = \begin{bmatrix} f_i(e_i(t) + r_i(t)) - \zeta_i(r_i(t)) \\ \zeta_i(r_i(t)) \end{bmatrix}, \quad (6)$$

$$\mathcal{G}_i(y_i(t)) = \begin{bmatrix} g_i(e_i(t) + r_i(t)) \\ 0_{n_i \times m_i} \end{bmatrix}, \quad (7)$$

$$\mathcal{H}_i(y_i(t)) = \begin{bmatrix} h_i(e_i(t) + r_i(t)) \\ 0_{n_i \times q_i} \end{bmatrix}. \quad (8)$$

We aim to design a pair of decentralized control policies $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_N$ to ensure that large-scale system [Equation \(1\)](#) can track the desired object while being restricted by external disturbances. It means that as $t \rightarrow +\infty$, $\|x_i(t) - r_i(t)\| \rightarrow 0$. Meanwhile, it is noteworthy that the control pair $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_N$ should be pointed out only as a corresponding controller with the local information. In what follows, it presents the DTC problem by transforming it into the optimal controller design of ATISs by considering an appropriate discounted cost function.

3. DTC DESIGN VIA OPTIMAL REGULATION

3.1. Optimal control and the HJI equations

In this section, the optimal DTC strategy of the ATIS with the disturbance rejection is elaborated. It is addressed by solving the HJI equation with a discounted cost function. Then, we consider the nominal part of the augmented system [Equation \(5\)](#) as

$$\dot{y}_i(t) = \mathcal{F}_i(y_i(t)) + \mathcal{G}_i(y_i(t))u_i(y_i(t)) + \mathcal{H}_i(y_i(t))v_i(y_i(t)). \quad (9)$$

We assume that $\mathcal{F}_i + \mathcal{G}_i u_i + \mathcal{H}_i v_i$ is Lipschitz continuous on a set $\Omega_i \subset \mathbb{R}^{2n_i}$, which is commonly used in the field of adaptive critic control to ensure the existence and uniqueness of the solution for the differential equation. Related to the i th ATIS, we manage to minimize and maximize the discounted cost function as

$$J_i(y_{i0}) = \int_0^\infty e^{-\lambda_i(\tau-t)} \left\{ y_i^\top(\tau) Q_i y_i(\tau) + u_i^\top(\tau) R_i u_i(\tau) - \varrho_i^2 v_i^\top(\tau) v_i(\tau) \right\} d\tau, \quad (10)$$

where $Q_i \in \mathbb{R}^{2n_i \times 2n_i}$, $R_i \in \mathbb{R}^{m_i \times m_i}$ are both positive definite matrices. Herein, we let $y_i^\top Q_i y_i - \varrho_i^2 v_i^\top(y_i) v_i(y_i) = \gamma_i^2(y_i)$ and $\theta_i(y_i) \leq \sqrt{\gamma_i^2(y_i) - \lambda_i J_i(y_i)}$, where $\gamma_i^2(y_i) > \lambda_i J_i(y_i)$. It is worth noting that this inequality is employed to prove the feasibility of Theorem 1. Then, Equation (10) can be equivalent to

$$J_i(y_{i0}) = \int_0^\infty e^{-\lambda_i(\tau-t)} \left\{ \gamma_i^2(y_i) + u_i^\top(y_i(\tau)) R_i u_i(y_i(\tau)) \right\} d\tau. \tag{11}$$

If Equation (11) is continuously differentiable, the nonlinear Lyapunov equation is the infinitely small form of Equation (11). The Lyapunov equation is as follows:

$$\gamma_i^2(y_i) + u_i^\top(y_i) R_i u_i(y_i) - \lambda_i J_i(y_i) + (\nabla J_i(y_i))^\top [\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i) u_i(y_i) + \mathcal{H}_i(y_i) v_i(y_i)] = 0. \tag{12}$$

Define the Hamiltonian of the i th ATIS for the optimization problem as

$$H_i(y_i, u_i, v_i, \nabla J_i(y_i)) = \gamma_i^2(y_i) + u_i^\top(y_i) R_i u_i(y_i) - \lambda_i J_i(y_i) + (\nabla J_i(y_i))^\top \times [\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i) u_i(y_i) + \mathcal{H}_i(y_i) v_i(y_i)]. \tag{13}$$

To acquire the saddle point solution $\{u_i^*, v_i^*\}$, the local optimal cost function need to satisfy the following Nash condition

$$J_i^*(y_{i0}) = \min_{u_i} \max_{v_i} J_i(y_{i0}). \tag{14}$$

Then, the optimal cost function $J_i^*(y_i)$ is derived via solving the local HJI equation in the following:

$$\min_{u_i} \max_{v_i} H_i(y_i, u_i, v_i, \nabla J_i^*(y_i)) = 0. \tag{15}$$

Due to the saddle point solution $\{u_i^*, v_i^*\}$ satisfies the extremum theorem, the optimal tracking control law and the worst disturbance law can be computed by

$$u_i^*(y_i) = -\frac{1}{2} R_i^{-1} \mathcal{G}_i^\top(y_i) \nabla J_i^*(y_i), \tag{16}$$

$$v_i^*(y_i) = \frac{1}{2\varrho_i^2} \mathcal{H}_i^\top(y_i) \nabla J_i^*(y_i). \tag{17}$$

Substituting the optimal tracking control strategy Equation (16) into Equation (15), the HJI equation for the i th ATIS becomes

$$(\nabla J_i^*(y_i))^\top [\mathcal{F}_i(y_i) + \mathcal{H}_i(y_i) v_i^*(y_i)] + \gamma_i^2(y_i) - \lambda_i J_i^*(y_i) - \frac{1}{4} (\nabla J_i^*(y_i))^\top \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^\top(y_i) \nabla J_i^*(y_i) = 0. \tag{18}$$

3.2. Establishment of the DTC strategy design

In the following, we present the DTC strategy by adding the feedback gain to the interconnected system Equation (5). Herein, the following lemma is given by

Lemma 1 Considering the ATIS Equation (9), the feedback control

$$\bar{u}_i(y_i) = k_i u_i^*(y_i) \tag{19}$$

can ensure the N ATISs are asymptotically stable as long as $k_i \geq 1/2$, which makes the tracking error approach to zero.

Proof. The lemma can be proved by showing $J_i^*(y_i)$ is a candidate Lyapunov function. We can find $J_i^*(y_i) \geq 0$ in Equation (11), which implies that $J_i^*(y_i)$ is a positive definite function. The derivative of $J_i^*(y_i)$ along with the i th ATIS is given by

$$\begin{aligned} \dot{J}_i^*(y_i) &= (\nabla J_i^*(y_i))^T \dot{y}_i \\ &= (\nabla J_i^*(y_i))^T [\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i)\bar{u}_i(y_i) + \mathcal{H}_i(y_i)v_i(y_i)]. \end{aligned} \quad (20)$$

Substituting Equations (18) and (19) into Equation (20), we can rewrite it as

$$\begin{aligned} \dot{J}_i^*(y_i) &= -\gamma_i^2(y_i) + \lambda_i J_i^*(y_i) + \frac{1}{4}(\nabla J_i^*(y_i))^T \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) \nabla J_i^*(y_i) \\ &\quad - \frac{1}{2} k_i (\nabla J_i^*(y_i))^T \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) \nabla J_i^*(y_i) \\ &= -(\gamma_i^2(y_i) - \lambda_i J_i^*(y_i)) - \left(\frac{1}{2} k_i - \frac{1}{4} \right) \left\| R_i^{-\frac{1}{2}} \mathcal{G}_i^T(y_i) \nabla J_i^*(y_i) \right\|^2. \end{aligned} \quad (21)$$

Observing Equation (21), we can obtain that $\dot{J}_i^*(y_i) < 0$ holds under the condition $\gamma_i^2(y_i) > \lambda_i J_i^*(y_i)$ for all $k_i \geq 1/2$ and $y_i \neq 0$. Thus, the conditions are satisfied for Lyapunov local stability theory and the actual state of each ATIS can realize desired tracking objectives under the feedback control strategy. The proof is completed.

Remark 1. It is worth mentioning that only when $k_i = 1$, the feedback control is optimal. Then, we will show the following theorem to verify the proposed control law can effectively establish the DTC strategy.

Theorem 1 Taking Equation (2) and the interconnected augmented tracking system Equation (5) into account, there exist N positive numbers k_i^* , such that, for any $k_i > k_i^*$, the feedback control policies given by Equation (19) guarantee that the interconnected tracking system can maintain the asymptotic stability. In other words, the control pair $\bar{u}_1(y_1), \bar{u}_2(y_2), \dots, \bar{u}_N(y_N)$ is the DTC strategy for the large-scale system.

Proof. Inspired by Lemma 1, we observe that $J_i^*(y_i)$ is the Lyapunov function. Therefore, a composite Lyapunov function of $J_i^*(y_i)$ is chosen as

$$\mathcal{L}(y) = \sum_{i=1}^N \mu_i J_i^*(y_i), \quad (22)$$

where μ_i is a random positive constant. Taking the time derivative of $\mathcal{L}(y)$, we have

$$\begin{aligned} \dot{\mathcal{L}}(y) &= \sum_{i=1}^N \mu_i \dot{J}_i^*(y_i) \\ &= \sum_{i=1}^N \mu_i \left\{ (\nabla J_i^*(y_i))^T [\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i)\bar{u}_i(y_i) + \mathcal{H}_i(y_i)v_i(y_i)] \right. \\ &\quad \left. + (\nabla J_i^*(y_i))^T \mathcal{G}_i(y_i) \bar{Z}_i(y) \right\}. \end{aligned} \quad (23)$$

Considering Equation (2), the mentioned inequality $\theta_i(y_i) \leq \sqrt{\gamma_i^2(y_i) - \lambda_i J_i(y_i)}$, where $\gamma_i^2(y_i) > \lambda_i J_i(y_i)$, and Equation (21), the upper formula can be converted to

$$\begin{aligned} \dot{\mathcal{L}}(y) &\leq - \sum_{i=1}^N \mu_i \left\{ \gamma_i^2(y_i) - \lambda_i J_i(y_i) + \left(\frac{1}{2} k_i - \frac{1}{4} \right) \left\| (\nabla J_i^*(y_i))^T \mathcal{G}_i(y_i) R_i^{-\frac{1}{2}} \right\|^2 \right. \\ &\quad \left. - \left\| (\nabla J_i^*(y_i))^T \mathcal{G}_i(y_i) R_i^{-\frac{1}{2}} \right\| \sum_{i=1}^N \beta_{ij} \sqrt{\gamma_i^2(y_i) - \lambda_i J_i(y_i)} \right\}. \end{aligned} \quad (24)$$

Herein, in order to transform Equation (24) to the compact form, we denote

$$M = \text{diag}\{\mu_1, \mu_2, \dots, \mu_N\}, \tag{25}$$

$$K = \text{diag}\left\{\frac{1}{2}k_1 - \frac{1}{4}, \frac{1}{2}k_2 - \frac{1}{4}, \dots, \frac{1}{2}k_N - \frac{1}{4}\right\}, \tag{26}$$

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1N} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1} & \beta_{N2} & \dots & \beta_{NN} \end{bmatrix}. \tag{27}$$

Therefore, we introduce a $2N$ -dimensional column vector ϑ , which consists of the N -dimensional column vector $\sqrt{\gamma_i^2(y_i) - \lambda_i J_i(y_i)}$ and the N -dimensional column vector $\left\|(\nabla J_i^*(y_i))^T \mathcal{G}_i(y_i) R_i^{-\frac{1}{2}}\right\|$. Its form is as follows:

$$\vartheta = \begin{bmatrix} \sqrt{\gamma_1^2(y_1) - \lambda_1 J_1(y_1)} \\ \sqrt{\gamma_2^2(y_2) - \lambda_2 J_2(y_2)} \\ \vdots \\ \sqrt{\gamma_N^2(y_N) - \lambda_N J_N(y_N)} \\ \hline \left\|(\nabla J_1^*(y_1))^T \mathcal{G}_1(y_1) R_1^{-\frac{1}{2}}\right\| \\ \left\|(\nabla J_2^*(y_2))^T \mathcal{G}_2(y_2) R_2^{-\frac{1}{2}}\right\| \\ \vdots \\ \left\|(\nabla J_N^*(y_N))^T \mathcal{G}_N(y_N) R_N^{-\frac{1}{2}}\right\| \end{bmatrix} \tag{28}$$

Next, Equation (24) can be transformed to the following compact form:

$$\begin{aligned} \dot{\mathcal{L}}(y) &\leq -\vartheta^T \begin{bmatrix} M & -\frac{1}{2}B^T M \\ -\frac{1}{2}MB & MK \end{bmatrix} \vartheta \\ &\triangleq -\vartheta^T \mathcal{A} \vartheta. \end{aligned} \tag{29}$$

According to Equation (29), it can be concluded that when k_i is sufficiently large, the matrix \mathcal{A} is positive definite, which means there exist k_i^* so that any $k_i > k_i^*$ sufficiently large to ensure the positive definite property of \mathcal{A} . Then, we get $\dot{\mathcal{L}}(y) < 0$. Consequently, the DTC strategy with external disturbances is constructed. The proof is completed.

Obviously, the key point of designing the DTC strategy is to obtain the optimal controller of the ATIS based on Theorem 1. Next, for the sake of getting hold of optimal controllers for the N ATISs by solving the HJI equations, in the following, we employ the ADP method to obtain the approximate optimal solutions by means of critic networks.

4. OPTIMAL DTC DESIGN VIA NEURAL NETWORKS

4.1. Implementation procedure via neural networks

In this section, we show the process of finding the approximate optimal solution by employing the ADP method based on neural networks. The critic networks have the capability of approximating nonlinear mapping, and the approximate cost function can be derived for the DTC system. Hence, $J_i^*(y_i)$ can be expressed as

$$J_i^*(y_i) = w_{ci}^T \sigma_{ci}(y_i) + \xi_{ci}(y_i), \quad (30)$$

where $w_{ci} \in \mathbb{R}^{l_{ci}}$ is the ideal weight vector, l_{ci} is the number of neurons in the hidden layer, $\sigma_{ci}(y_i) \in \mathbb{R}^{l_{ci}}$ is the activation function, and $\xi_{ci}(y_i)$ is the reconstruction error of the i th neural network. The gradient of $J_i^*(y_i)$ is formulated as

$$\nabla J_i^*(y_i) = (\nabla \sigma_{ci}(y_i))^T w_{ci} + \nabla \xi_{ci}(y_i), \quad (31)$$

Considering Equation (16), the optimal control policy for the i th ATIS is replaced by

$$u_i^*(y_i) = -\frac{1}{2} R_i^{-1} \mathcal{G}_i^T(y_i) \left((\nabla \sigma_{ci}(y_i))^T w_{ci} + \nabla \xi_{ci}(y_i) \right). \quad (32)$$

Utilizing Equations (31) and (32), the Hamiltonian associated with the i th ATIS is obtained as

$$\begin{aligned} H_i(y_i, v_i(y_i), w_{ci}) &= \gamma_i^2(y_i) - \lambda_i (w_{ci}^T \sigma_{ci}(y_i)) + w_{ci}^T (\nabla \sigma_{ci}(y_i)) [\mathcal{F}_i(y_i) + \mathcal{H}_i(y_i) v_i(y_i)] \\ &\quad - \frac{1}{4} w_{ci}^T \nabla \sigma_{ci}(y_i) \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) (\nabla \sigma_{ci}(y_i))^T w_{ci} + e_{chi} = 0, \end{aligned} \quad (33)$$

where e_{chi} is the residual error of the neural network. To avoid the unknown ideal weight vector, we construct N critic neural networks to approximate $J_i^*(y_i)$ as

$$\hat{J}_i^*(y_i) = \hat{w}_{ci}^T \sigma_{ci}(y_i), \quad (34)$$

where \hat{w}_{ci} is the estimated weight. Likewise, the derivative of $\hat{J}_i^*(y_i)$ is

$$\nabla \hat{J}_i^*(y_i) = (\nabla \sigma_{ci}(y_i))^T \hat{w}_{ci}. \quad (35)$$

Based on Equation (35), we obtain the estimated value of $u_i^*(y_i)$ and $v_i^*(y_i)$ as

$$\hat{u}_i^*(y_i) = -\frac{1}{2} R_i^{-1} \mathcal{G}_i^T(y_i) (\nabla \sigma_{ci}(y_i))^T \hat{w}_{ci}, \quad (36)$$

$$\hat{v}_i^*(y_i) = \frac{1}{2\varrho_i^2} \mathcal{H}_i^T(y_i) (\nabla \sigma_{ci}(y_i))^T \hat{w}_{ci}. \quad (37)$$

Considering Equations (34-36), the approximate Hamiltonian is expressed as

$$\begin{aligned} \hat{H}_i(y_i, \hat{v}_i^*(y_i), \hat{w}_{ci}) &= \gamma_i^2(y_i) - \lambda_i (\hat{w}_{ci}^T \sigma_{ci}(y_i)) + \hat{w}_{ci}^T (\nabla \sigma_{ci}(y_i)) [\mathcal{F}_i(y_i) + \mathcal{H}_i(y_i) \hat{v}_i^*(y_i)] \\ &\quad - \frac{1}{4} \hat{w}_{ci}^T \nabla \sigma_{ci}(y_i) \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) (\nabla \sigma_{ci}(y_i))^T \hat{w}_{ci} = e_{ci}. \end{aligned} \quad (38)$$

Then, we obtain an error function of the Hamiltonian, which is denoted as e_{ci} and is expressed by

$$\begin{aligned} e_{ci} &= \hat{H}_i(y_i, \hat{v}_i^*(y_i), \hat{w}_{ci}) - H_i(y_i, v_i(y_i), w_{ci}) \\ &= \lambda_i (\tilde{w}_{ci}^T \sigma_{ci}(y_i)) - \tilde{w}_{ci}^T (\nabla \sigma_{ci}(y_i)) [\mathcal{F}_i(y_i) + \mathcal{H}_i(y_i) v_i(y_i)] \\ &\quad - \frac{1}{4} \tilde{w}_{ci}^T \nabla \sigma_{ci}(y_i) \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) (\nabla \sigma_{ci}(y_i))^T \tilde{w}_{ci} \\ &\quad + \frac{1}{2} w_{ci}^T \nabla \sigma_{ci}(y_i) \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) (\nabla \sigma_{ci}(y_i))^T \tilde{w}_{ci} - e_{chi}, \end{aligned} \quad (39)$$

where $\tilde{w}_{ci} = w_{ci} - \hat{w}_{ci}$ is the weight error vector. At present, in order to minimize the objective function $E_{ci} = (1/2)e_{ci}^T e_{ci}$, the normalised steepest descent algorithm based on Equation (38) is employed as follows:

$$\dot{\hat{w}}_{ci} = -\eta_{ci} \frac{1}{(1 + \phi_i^T \phi_i)^2} \left(\frac{\partial E_{ci}}{\partial \hat{w}_{ci}} \right) = -\eta_{ci} \frac{\phi_i}{(1 + \phi_i^T \phi_i)^2} e_{ci}, \tag{40}$$

where $\eta_{ci} > 0$ represents the basic learning rate. Besides, $(1 + \phi_i^T \phi_i)^2$ is introduced for the normalization to simplify the critic error dynamics, and ϕ_i is derived as

$$\phi_i = \nabla \sigma_{ci}(y_i) [\mathcal{F}_i(y_i) + \mathcal{H}_i(y_i) \hat{v}_i^*(y_i)] - \lambda_i \sigma_{ci}(y_i). \tag{41}$$

Usually, in the traditional weight training process, it is often necessary to select the appropriate initial weight vector for effective training. To eliminate the initial admissible control law, an improved critic learning rule is presented in the following.

4.2. Improved critic learning rule via neural networks

Herein, an additional Lyapunov function is introduced for the purpose of improving the critic learning mechanism. Then, the following rational assumption is given.

Assumption 1 Consider the dynamic of the i th ATIS Equation (9) with the optimal cost function Equation (14) and the closed-loop optimal control policy Equation (32). We select $J_{si}(y_i)$ as a continuously differentiable Lyapunov function and have the following relation:

$$J_{si}(y_i) = (\nabla J_{si}(y_i))^T [\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i) u_i^*(y_i) + \mathcal{H}_i(y_i) v_i^*(y_i)] < 0. \tag{42}$$

In other words, there exists a positive definite matrix \mathcal{B} such that

$$(\nabla J_{si}(y_i))^T [\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i) u_i^*(y_i) + \mathcal{H}_i(y_i) v_i^*(y_i)] = -(\nabla J_{si}(y_i))^T \mathcal{B} \nabla J_{si}(y_i) \leq -s_{mi} \|\nabla J_{si}(y_i)\|^2, \tag{43}$$

where s_{mi} is the minimum eigenvalue of the matrix \mathcal{B} .

Remark 2. Herein, the motivation of selecting the cost function $J_{si}(y_i)$ is to obtain the optimal DTC strategy, which can minimize and maximize $J_{si}(y_i)$ under the optimal control law and the worst disturbance law. Moreover, we can discuss the stability of closed-loop systems by the constructed optimal cost function. Besides, just to be clear, $J_{si}(y_i)$ is derived by properly selecting the quadratic polynomial in terms of the state vector. We generally choose $J_{si}(y_i) = 0.5 y_i^T y_i$.

When the condition occurs, that is, $(\nabla J_{si}(y_i))^T [\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i) u_i^*(y_i) + \mathcal{H}_i(y_i) v_i^*(y_i)] > 0$, which means the system is in an unstable state under the optimal control law Equation (36). In this case, an additional term is introduced to ensure the system stability. Based on Equation (36), some processing is performed as follows:

$$\begin{aligned} & \frac{-\partial [(\nabla J_{si}(y_i))^T (\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i) u_i^*(y_i) + \mathcal{H}_i(y_i) v_i^*(y_i))]}{\partial \hat{w}_{ci}} \\ &= \left(\frac{\partial \hat{u}_i^*(y_i)}{\partial \hat{w}_{ci}} \right)^T \frac{-\partial [(\nabla J_{si}(y_i))^T (\mathcal{F}_i(y_i) + \mathcal{G}_i(y_i) u_i^*(y_i) + \mathcal{H}_i(y_i) v_i^*(y_i))]}{\partial \hat{u}_i^*(y_i)} \\ &= \frac{1}{2} \nabla \sigma_{ci}(y_i) \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) \nabla J_{si}(y_i). \end{aligned} \tag{44}$$

Thus, we describe the improved learning rule as

$$\dot{\hat{w}}_{ci} = -\eta_{ci} \frac{\phi_i}{(1 + \phi_i^T \phi_i)^2} e_{ci} + \frac{1}{2} \eta_{si} \Pi_i(y_i, \hat{u}_i^*, \hat{v}_i^*) \nabla \sigma_{ci}(y_i) \mathcal{G}_i(y_i) R_i^{-1} \mathcal{G}_i^T(y_i) \nabla J_{si}(y_i), \tag{45}$$

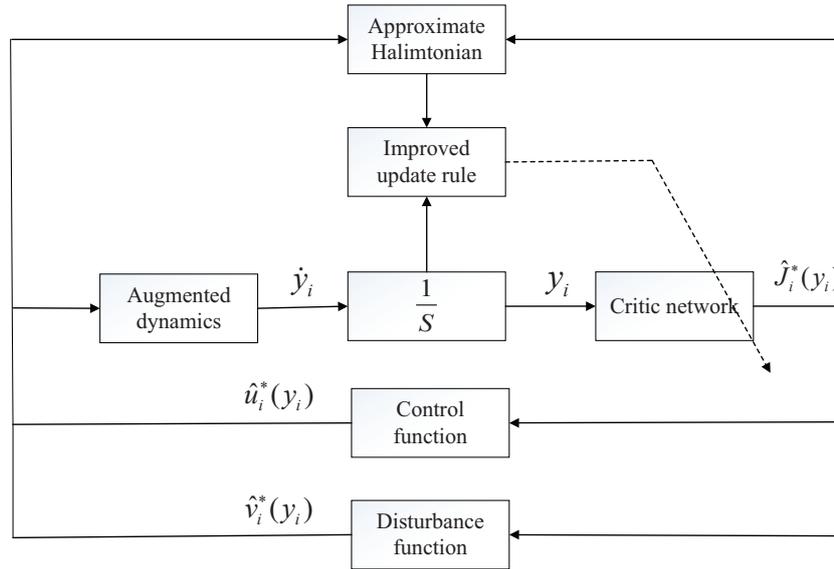


Figure 1. Control structure of the ATIS. ATIS: augmented tracking isolated subsystem.

where $\eta_{si} > 0$ represents the additional learning rate with respect to the stabilising term and $\Pi_i(y_i, \hat{u}_i^*, \hat{v}_i^*)$ stands for the adaptation parameter term that tests the stability of the ATIS. The definition of Π_i is as follows:

$$\Pi_i(y_i, \hat{u}_i^*, \hat{v}_i^*) = \begin{cases} 0, & \text{if } \dot{J}_{si}^*(y_i) < 0, \\ 1, & \text{else.} \end{cases} \tag{46}$$

It is found that when the derivative of $J_{si}(y_i)$ satisfies $\dot{J}_{si}(y_i) < 0$, the latter term of the weight update rule does not play its role so that the update mode is still the traditional normalized steepest descent algorithm. When $\dot{J}_{si}(y_i) > 0$, the latter term of the weight update rule starts to play its role of ensuring the stability, that is, the improved weight update method is adopted. It can be seen that the system can be adjusted to be stable under the improved weight updating criterion. Moreover, in order to clearly highlight that we have achieved the elimination of the initial admissible control law, herein, we set the initial weight vector to zero. Through the new critic learning rule, the structure of the proposed DTC strategy for ATIS is performed in [Figure 1](#).

In accordance to $\dot{\tilde{w}}_{ci} = -\dot{\hat{w}}_{ci}$ and [Equation \(39\)](#), the specific form of \tilde{w}_{ci} is derived. Then, we can convert the estimated weight \hat{w}_{ci} into the form of the weight vector w_{ci} and the error weight vector \tilde{w}_{ci} , which can be employed by proving the state y_i and the weight estimation error \tilde{w}_{ci} are UUB for the closed-loop system.

5. SIMULATION EXPERIMENT

In this section, we will introduce the common mechanical vibration system, that is, the spring-mass-damper system. The structural diagram of the mechanical system is shown in [Figure 2](#). From it, M_1 and M_2 denote the mass of two objects, K_1, K_2 , and K_3 represent the stiffness constants of three springs. C_1, C_2 , and C_3 stand for the damping, respectively.

In addition, let P_i, V_i, F_i , and f_{μ_i} be the position, the velocity, the force, and the friction applied to the object, where $i = 1, 2$. Hence, the system dynamics for M_1 and M_2 are as follows:

$$\dot{P}_1 = V_1, \tag{47}$$

$$M_1 \dot{V}_1 = -K_1 P_1 - C_1 V_1 + K_2 (P_2 - P_1) + C_2 (V_2 - V_1) + F_1 - f_{\mu_1}, \tag{48}$$

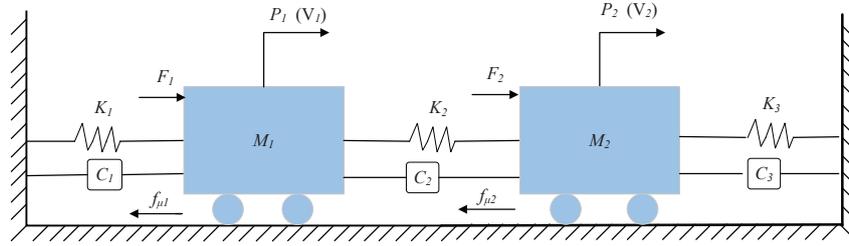


Figure 2. Simple diagram of the interconnected mass-spring-damper system.

$$\dot{P}_2 = V_2, \tag{49}$$

$$M_2 \dot{V}_2 = -K_3 P_2 - C_3 V_2 + K_2 (P_1 - P_2) + C_2 (V_1 - V_2) + F_2 - f_{\mu 2}. \tag{50}$$

For the object M_1 , we define $x_{11} = P_1, x_{12} = V_1, \bar{u}_1(x_1) = F_1$, and $v_1(x_1) = f_{\mu 1}$. In the same way, we let $x_{21} = P_2, x_{22} = V_2, \bar{u}_2(x_2) = F_2$, and $v_2(x_2) = f_{\mu 2}$ for the object M_2 . Next, the state-space of the spring-mass-damper system is written as

$$\dot{x}_1 = \begin{bmatrix} \dot{x}_{11} \\ \dot{x}_{12} \end{bmatrix} = \begin{bmatrix} x_{12} \\ -\frac{K_1}{M_1} x_{11} - \frac{C_1}{M_1} x_{12} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{M_1} \end{bmatrix} (\bar{u}_1(x_1) + Z_1(x)) + \begin{bmatrix} 0 \\ -\frac{1}{M_1} \end{bmatrix} v_1(x_1) \tag{51}$$

and

$$\dot{x}_2 = \begin{bmatrix} \dot{x}_{21} \\ \dot{x}_{22} \end{bmatrix} = \begin{bmatrix} x_{22} \\ -\frac{K_3}{M_2} x_{21} - \frac{C_3}{M_2} x_{22} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{M_2} \end{bmatrix} (\bar{u}_2(x_2) + Z_2(x)) + \begin{bmatrix} 0 \\ -\frac{1}{M_2} \end{bmatrix} v_2(x_2), \tag{52}$$

where $x_1 = [x_{11}, x_{12}]^T \in \mathbb{R}^2$ and $x_2 = [x_{21}, x_{22}]^T \in \mathbb{R}^2$ are system states. $\bar{u}_1(x_1) \in \mathbb{R}, \bar{u}_2(x_2) \in \mathbb{R}, v_1(x_1) \in \mathbb{R}$, and $v_2(x_2) \in \mathbb{R}$ are control inputs and disturbance inputs of the subsystem 1 and the subsystem 2, respectively. Simultaneously, $Z_1(x) = K_2(x_{21} - x_{11}) + C_2(x_{22} - x_{12})$ and $Z_2(x) = K_2(x_{11} - x_{21}) + C_2(x_{12} - x_{22})$, which indicates the spring K_2 and the damping C_2 play a connecting role for two subsystems. Herein, we let $\theta_1(x_1) = \|x_1\|$ and $\theta_2(x_2) = \|x_2\|$. Besides, we choose $\beta_{11} = \beta_{12} = 1, \beta_{21} = \beta_{22} = 1/2$, and $\mu_1 = \mu_2 = 1$. Moreover, we select $\lambda_1 = \lambda_2 = 0.6, \varrho_1 = \varrho_2 = 1, R_1 = R_2 = 2$, and $Q_1 = Q_2 = 2I_4$, where I_4 is the four-dimensional identity matrix. Above all, the desired reference trajectories r_1 and r_2 for two subsystems are generated by the following command system:

$$\dot{r}_i = \begin{bmatrix} \dot{r}_{i1} \\ \dot{r}_{i2} \end{bmatrix} = \begin{bmatrix} -0.5r_{i1} - 0.5r_{i2} \cos(r_{i1}) \\ \sin(r_{i1}) - 0.5r_{i2} \end{bmatrix}, \quad i = 1, 2, \tag{53}$$

where $r_1 = [r_{11}, r_{12}]^T \in \mathbb{R}^2$ and $r_2 = [r_{21}, r_{22}]^T \in \mathbb{R}^2$ are reference states. Then, we define the tracking errors as $e_{i1} = x_{i1} - r_{i1}$ and $e_{i2} = x_{i2} - r_{i2}$. Hence, the augmented state vector can be expressed as $y_i = [y_{i1}, y_{i2}, y_{i3}, y_{i4}]^T = [e_{i1}, e_{i2}, r_{i1}, r_{i2}]^T, i = 1, 2$. We set practical parameters as $M_1 = 1\text{kg}, K_1 = 3\text{N/m}$, and $C_1 = 0.5\text{Ns/m}$ for the subsystem 1. Similarly, we let $M_2 = 2\text{kg}, K_3 = 5\text{N/m}$, and $C_3 = 1\text{Ns/m}$ for the subsystem 2. Considering Equations (51-53), the augmented system dynamics \dot{y}_1 and \dot{y}_2 can be obtained in the following forms:

$$\dot{y}_1 = \begin{bmatrix} r_{12} + e_{12} + 0.5r_{11} + 0.5r_{12} \cos(r_{11}) \\ -3(r_{11} + e_{11}) - 0.5(r_{12} + e_{12}) - \sin(r_{11}) + 0.5r_{12} \\ -0.5r_{11} - 0.5r_{12} \cos(r_{11}) \\ \sin(r_{11}) - 0.5r_{12} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \bar{u}_1(y_1) + \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} v_1(y_1) \tag{54}$$

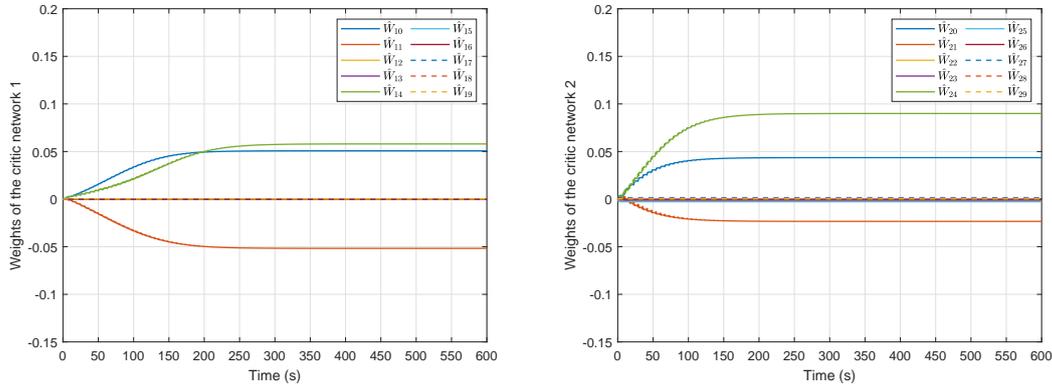


Figure 3. Weights convergence process of the critic network 1 and the critic network 2.

and

$$\dot{y}_2 = \begin{bmatrix} r_{22} + e_{22} + 0.5r_{21} + 0.5r_{22} \cos(r_{21}) \\ -2.5(r_{21} + e_{21}) - 0.5(r_{22} + e_{22}) - \sin(r_{21}) + 0.5r_{22} \\ -0.5r_{21} - 0.5r_{22} \cos(r_{21}) \\ \sin(r_{21}) - 0.5r_{22} \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0 \\ 0 \end{bmatrix} \bar{u}_2(y_2) + \begin{bmatrix} 0 \\ -0.5 \\ 0 \\ 0 \end{bmatrix} v_2(y_2). \quad (55)$$

Based on the online ADP algorithm, two critic networks are constructed as follows:

$$\hat{J}_1^*(y_1) = \hat{w}_{10}y_{11}^2 + \hat{w}_{11}y_{11}y_{12} + \hat{w}_{12}y_{11}y_{13} + \hat{w}_{13}y_{11}y_{14} + \hat{w}_{14}y_{12}^2 + \hat{w}_{15}y_{12}y_{13} + \hat{w}_{16}y_{12}y_{14} + \hat{w}_{17}y_{13}^2 + \hat{w}_{18}y_{13}y_{14} + \hat{w}_{19}y_{14}^2 \quad (56)$$

and

$$\hat{J}_2^*(y_2) = \hat{w}_{20}y_{21}^2 + \hat{w}_{21}y_{21}y_{22} + \hat{w}_{22}y_{21}y_{23} + \hat{w}_{23}y_{21}y_{24} + \hat{w}_{24}y_{22}^2 + \hat{w}_{25}y_{22}y_{23} + \hat{w}_{26}y_{22}y_{24} + \hat{w}_{27}y_{23}^2 + \hat{w}_{28}y_{23}y_{24} + \hat{w}_{29}y_{24}^2. \quad (57)$$

During the online learning process, we take basic learning rates and additional learning rates as $\eta_{c1} = 0.01$, $\eta_{c2} = 0.03$ as well as $\eta_{s1} = \eta_{s2} = 0.01$. Let initial system states and reference states be $x_{10} = [1.5, 0]^T$, $x_{20} = [1, -1]^T$, and $r_{10} = r_{20} = [0.5, -0.5]^T$, respectively. Therefore, initial states of the ATIS are $y_{10} = [1, 0.5, 0.5, -0.5]^T$ and $y_{20} = [0.5, -0.5, 0.5, -0.5]^T$.

Herein, two probing noises are added within the beginning 400 steps to keep the persistence of excitation condition of the ATIS. The weight convergence curves are shown in Figure 3. It can be seen that the weight has converged to a certain numerical value before turning off the excitation condition, which confirms the validity of the improved weight update algorithm. From it, we find the initial weights are selected as zero, which indicates the initial admissible control is eliminated.

Next, in order to make the system achieve the purpose of the optimal tracking, feedback gains are selected as $k_1 = k_2 = 1$. Then, the DTC strategy $\{k_1 \hat{u}_1^*(y_1), k_2 \hat{u}_2^*(y_2)\}$ can be derived from the obtained weight vector for the spring-mass-damper interconnected system. In addition, the evolution curves are shown in Figure 4, which displays the tracking control inputs and disturbance inputs for the subsystem 1 and the subsystem 2. Then, the obtained DTC strategy is applied to the controlled system for 50 s, and its tracking error trajectory curves are displayed in Figure 5. It is obvious that the tracking error curves are eventually enforced to the origin. Taken together, this simulation result verifies the effectiveness of the proposed DTC strategy.

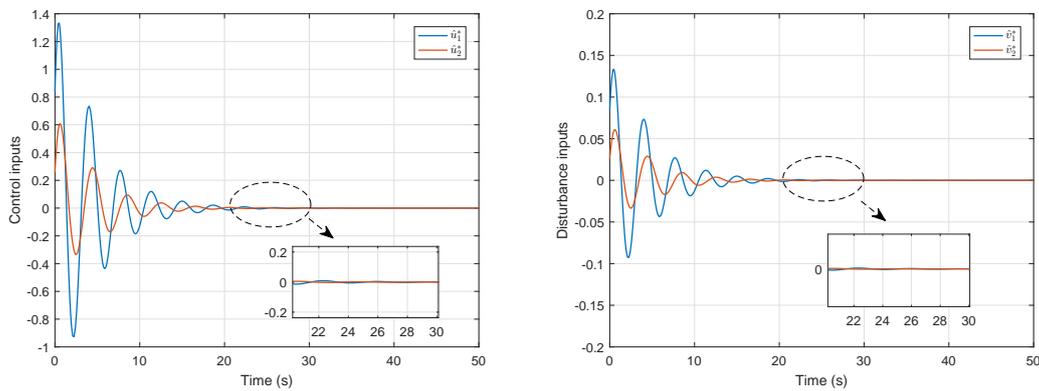


Figure 4. Tracking control inputs and disturbance inputs for subsystem 1 and subsystem 2

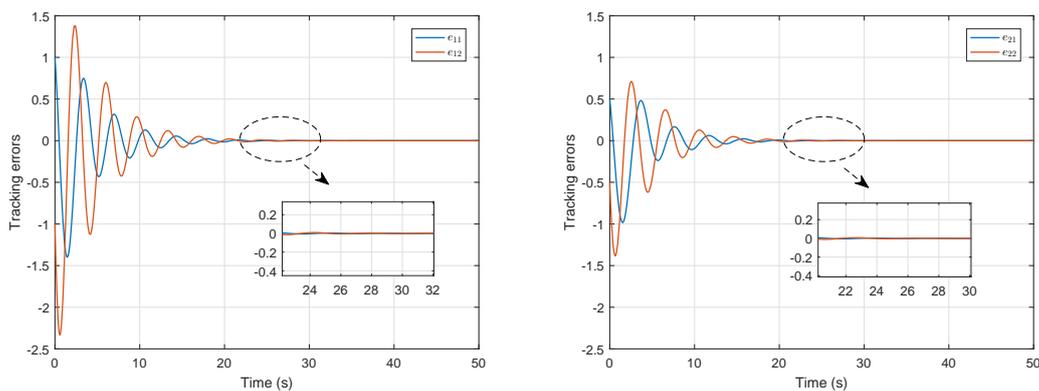


Figure 5. Tracking error trajectories for subsystem 1 and subsystem 2.

6. CONCLUSION

In this paper, the optimal DTC strategy for CT nonlinear large-scale systems with external disturbances is proposed by employing the ADP algorithm. The approximate optimal control law of the ATISs can achieve the trajectory tracking goal. Then, the establishment of the DTC strategy is derived by adding the appropriate feedback gain, whose feasibility has been proved via the Lyapunov theory. Note that all the above-mentioned results are investigated by considering a cost function with the discount. Then, only a series of single critic networks are employed to solve HJI equations of N ATISs, so that we acquire the approximate optimal control law and the worst disturbance law. In addition, the stability term added in the weight updating process avoids the selection of the initial stable control policy. Furthermore, the simulation results are displayed for the spring-mass-damper system to indicate the validity of the proposed DTC method. In the future, we will utilize more advanced methods to deal with the DTC problem for nonaffine systems. Besides, we can also consider the unmatched interconnected relationship for the DTC problem, which is a considerable direction of improved research.

DECLARATIONS

Authors' contributions

Made significant contributions to the conception and experiments: Fan W, Wang D

Made significant contributions to the writing: Fan W, Wang D

Made substantial contributions to the revision and translation: Liu A, Wang D

Availability of data and materials

Not applicable

Financial support and sponsorship

This work was supported in part by the National Natural Science Foundation of China (No. 62222301; No. 61890930-5 and No. 62021003); in part by the National Key Research and Development Program of China (No. 2021ZD0112302; No. 2021ZD0112301 and No. 2018YFC1900800-5); and in part by the Beijing Natural Science Foundation (No. JQ19013).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2023.

REFERENCES

1. Saberi A. On optimality of decentralized control for a class of nonlinear interconnected systems. *Automatica* 1988;24:1101–4. DOI
2. Mu CX, Sun CY, Wang D, Song AG, Qian CS. Decentralized adaptive optimal stabilization of nonlinear systems with matched interconnections. *Soft Comput* 2018;22:82705–15. DOI
3. Mehraeen S, Jagannathan S. Decentralized optimal control of a class of interconnected nonlinear discrete-time systems by using online Hamilton-Jacobi-Bellman formulation. *IEEE Trans Neural Netw* 2011;22:111715–96. DOI
4. Yang X, He HB. Adaptive dynamic programming for decentralized stabilization of uncertain nonlinear large-scale systems with mismatched interconnections. *IEEE Trans Syst Man Cybern Syst* 2020;50:82870–82. DOI
5. Karimi HR. How to deal with the complexity in robotic systems? *Complex Eng Syst* 2022;2:15. DOI
6. Xu Q, Yu C, Yuan X, Fu Z, Liu H. A distributed electricity energy trading strategy under energy shortage environment. *Complex Eng Syst* 2022;2:14. DOI
7. Bian T, Jiang Y, Jiang ZP. Decentralized adaptive optimal control of large-scale systems with application to power systems. *IEEE Trans Ind Electron* 2015;62:42439–47. DOI
8. Liu DR, Wang D, Li HL. Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach. *IEEE Trans Neural Netw Learn Syst* 2014;25:2418–28. DOI
9. Sun KK, Sui S, Tong SC. Fuzzy adaptive decentralized optimal control for strict feedback nonlinear large-scale systems. *IEEE Trans Cybern* 2018;48:41326–39. DOI
10. Wang XM, Feng ZG, Zhang GJ, Niu B, Yang D, et al. Adaptive decentralised control for large-scale non-linear non-strict-feedback interconnected systems with time-varying asymmetric output constraints and dead-zone inputs. *IET Control Theory & Appl* 2020;14:203417–27. DOI
11. Wei QL, Liu DR, Lin Q, Song RZ. Discrete-time optimal control via local policy iteration adaptive dynamic programming. *IEEE Trans Cybern* 2017;47:103367–79. DOI
12. Wang D, Ren J, Ha MM, Qiao JF. System stability of learning-based linear optimal control with general discounted value iteration. *IEEE Trans Neural Netw Learn Syst* 2022. DOI
13. Wang D, Ha MM, Zhao MM. The intelligent critic framework for advanced optimal control. *Artif Intell Rev* 2022;55:11–22. DOI
14. Wang D, Qiao JF, Cheng L. An approximate neuro-optimal solution of discounted guaranteed cost control design. *IEEE Trans Cybern* 2022;52:177–86. DOI
15. Li YM, Liu YJ, Tong SC. Observer-based neuro-adaptive optimized control of strict-feedback nonlinear systems with state constraints. *IEEE Trans Neural Netw Learn Syst* 2022;33:73131–45. DOI
16. Wang H, Yang CY, Liu XM, Zhou LN. Neural-network-based adaptive control of uncertain MIMO singularly perturbed systems with full-state constraints. *IEEE Trans Neural Netw Learn Syst* 2021. DOI
17. Zhang H, Hong QQ, Yan HC, Yang FW, Guo G. Event-based distributed H-infinity filtering networks of 2-DOF quarter-car suspension systems. *IEEE Trans Ind Inform* 2017;13:1312–21. DOI
18. Chen YG, Fei SM, Li YM. Robust stabilization for uncertain saturated time-delay systems: A distributed-delay-dependent polytopic

- approach. *IEEE Trans Automat Contr* 2017;62:73455–60. [DOI](#)
19. Chen YG, Wang ZD. Local stabilization for discrete-time systems with distributed state delay and fast-varying input delay under actuator saturations. *IEEE Trans Automat Contr* 2021;66:31337–44. [DOI](#)
 20. Fu H, Chen X, Wu M. Distributed optimal observer design of networked systems via adaptive critic design. *IEEE Trans Syst Man Cybern Syst* 2021;51:116976–85. [DOI](#)
 21. Narayanan V, Jagannathan S. Event-triggered distributed control of nonlinear interconnected systems using online reinforcement learning with exploration. *IEEE Trans Cybern* 2018;48:92510–9. [DOI](#)
 22. Wang D, Zhao MM, Ha MM, Qiao JF. Intelligent optimal tracking with application verifications via discounted generalized value iteration. *Acta Automatica Sinica* 2022;48:1182–93. [DOI](#)
 23. Zhang HG, Zhang K, Cai YL, Han J. Adaptive fuzzy fault-tolerant tracking control for partially unknown systems with actuator faults via integral reinforcement learning method. *IEEE Trans Fuzzy Syst* 2019;27:101986–98. [DOI](#)
 24. Modares H, Lewis FL. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica* 2014;50:71780–1792. [DOI](#)
 25. Ha MM, Wang D, Liu DR. Discounted iterative adaptive critic designs with novel stability analysis for tracking control. *IEEE/CAA Journal of Automat Sinica* 2022;9:71262–1272. [DOI](#)
 26. Qu QX, Zhang HG, Feng T, Jiang H. Decentralized adaptive tracking control scheme for nonlinear large-scale interconnected systems via adaptive dynamic programming. *Neurocomputing* 2017;225:1–10. [DOI](#)
 27. Niu B, Liu JD, Wang D, Zhao XD, Wang HQ. Adaptive decentralized asymptotic tracking control for large-scale nonlinear systems with unknown strong interconnections. *IEEE/CAA Journal of Automatica Sinica* 2022;9:1173–86. [DOI](#)
 28. Liu JD, Niu B, Kao YG, Zhao P, Yang D. Decentralized adaptive command filtered neural tracking control of large-scale nonlinear systems: An almost fast finite-time framework. *IEEE Trans Neural Netw Learn Syst* 2021;32:83621–2. [DOI](#)
 29. Tong SC, Zhang LL, Li YM. Observed-based adaptive fuzzy decentralized tracking control for switched uncertain nonlinear large-scale systems with dead zones. *IEEE Trans Syst Man Cybern Syst* 2016;46:137–47. [DOI](#)
 30. Wang D, Hu LZ, Zhao MM, Qiao JF. Dual event-triggered constrained control through adaptive critic for discrete-time zero-sum games. *IEEE Trans Syst Man Cybern Syst* 2023;53:31584–9. [DOI](#)
 31. Li XM, Zhang B, Li PS, Zhou Q, Lu RQ. Finite-horizon H-infinity state estimation for periodic neural networks over fading channels. *IEEE Trans Neural Netw Learn Syst* 2020;31:51450–60. [DOI](#)
 32. Duan JJ, Xu H, Liu WX, Peng JC, Jiang H. Zero-sum game based cooperative control for onboard pulsed power load accommodation. *IEEE Trans Ind Inform* 2020;16:1238–47. [DOI](#)
 33. Wang D, Zhao MM, Ha MM, Qiao JF. Stability and admissibility analysis for zero-sum games under general value iteration formulation. *IEEE Trans Neural Netw Learn Syst* 2022. [DOI](#)
 34. Zhang HG, Cui XH, Luo YH, Jiang H. Finite-horizon H-infinity tracking control for unknown nonlinear systems with saturating actuators. *IEEE Trans Neural Netw Learn Syst* 2018;29:41200–12. [DOI](#)
 35. Modares H, Lewis FL, Jiang ZP. H-infinity tracking control of completely unknown continuous-time systems via off-policy reinforcement learning. *IEEE Trans Neural Netw Learn Syst* 2015;26:102550–62. [DOI](#)
 36. Hou JX, Wang D, Liu DR, Zhang Y. Model-free H-infinity optimal tracking control of constrained nonlinear systems via an iterative adaptive learning algorithm. *IEEE Trans Syst Man Cybern Syst* 2020;50:114097–108. [DOI](#)