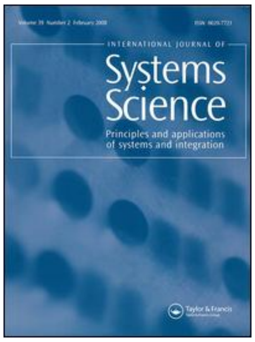

International Journal of Systems Science

ISSN: 0020-7721 (Print) 1464-5319 (Online) Journal homepage: http://www.tandfonline.com/loi/tsys20

Observer-critic structure-based adaptive dynamic programming for decentralised tracking control of unknown large-scale nonlinear systems

Bo Zhao, Derong Liu, Xiong Yang & Yuanchun Li

INTERNATIONAL JOURNAL OF SYSTEMS SCIENCE, 2017
VOL. 48, NO. 9, 1978–1989
http://dx.doi.org/10.1080/00207721.2017.1296982

Taylor & Francis
Taylor & Francis Group

Observer-critic structure-based adaptive dynamic programming for decentralised tracking control of unknown large-scale nonlinear systems

Bo Zhao [a], Derong Liu [b], Xiong Yang [c] and Yuanchun Li [d]

[a] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China; [b] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China; [c] School of Electrical Engineering and Automation, Tianjin University, Tianjin, China; [d] Department of Control Science and Engineering, Changchun University of Technology, Changchun, China

ABSTRACT

In this paper, a decentralised tracking control (DTC) scheme is developed for unknown large-scale nonlinear systems by using observer-critic structure-based adaptive dynamic programming. The control consists of local desired control, local tracking error control and a compensator. By introducing the local neural network observer, the subsystem dynamics can be identified. The identified subsystems can be used for the local desired control and the control input matrix, which is used in local tracking error control. Meanwhile, Hamiltonian-Jacobi-Bellman equation can be solved by constructing a critic neural network. Thus, the local tracking error control can be derived directly. To compensate the overall error caused by substitution, observation and approximation of the local tracking error control, an adaptive robustifying term is employed. Simulation examples are provided to demonstrate the effectiveness of the proposed DTC scheme.

ARTICLE HISTORY
Received 23 October 2016
Accepted 12 February 2017

KEYWORDS
Adaptive dynamic programming; decentralised tracking control; unknown large-scale nonlinear systems; observer-critic structure; neural networks

1. Introduction

The increasing demands of production efficiency and quality have led many practical systems to become large-scale and complex, such as power systems, smart grids, urban traffic systems and ecosystems. Generally speaking, the difficulty in designing many feedback loops for these systems inspires the development of decentralised control. The superiority of this approach lies in that it can reduce the design complexity by using only the local information of corresponding subsystems.

However, it is worthy pointing out that the major challenge in designing decentralised controllers for large-scale systems is how to deal with the interconnections, which affect the control performance, and even cause the system to be unstable. To solve this problem, considerable efforts have been made to the design of decentralised controllers for large-scale systems. Labibi, Lohmann, Sedigh, and Maralani (2002) provided sufficient conditions for minimising the weight sensitivity of the interconnections between the subsystems, and developed a constructive decentralised control scheme. Al-Tamimi, Lewis, and Abu-Khalaf (2008) addressed the global decentralised discrete-time sliding mode control of interconnected systems using only output information. They considered unmatched uncertainties, unknown interconnections and bounded time-varying delays for analysing system stability. Tong, Zhang, and Li (2016) investigated an observer-based adaptive fuzzy decentralised output-feedback tracking control for switched unknown systems with dead zones. Yi and Zheng (2016) proposed a decentralised proportional integral control scheme with delay dependent analysis to solve a constrained decentralised shaping control problem. Choi and Yoo (2016) proposed a decentralised approximation-free control design approach for interconnected nonlinear time-delay systems with unknown non-affine pure-feedback nonlinearities. By designing a reduced-order observer for estimating the unmeasured state variables, Hua, Zhang, and Guan (2015) presented a decentralised output feedback adaptive NN tracking control for time-delay stochastic interconnected systems with prescribed performance. Koo, Park, and Joo (2016) presented a decentralised sampled-data fuzzy observer to minimise the ratio of interconnection bound to attenuation degree. Each observer design problem was formulated as an optimisation problem with linear matrix inequality (LMI). Nowadays, the decentralised control methods have been utilised in robot manipulators (Henikl, Kemmetmller, Meurer, & Kugi, 2016; Zhao & Li, 2014), microgrids (Sadabadi, Karimi, & Karimi, 2015), power systems (Langarica & Ortega, 2015), spacecrafts (Zhang, Zhang, & Zhang, 2015), unmanned vehicles (Yang, Naeem, Irwin, & Li, 2014), and so on.

It is well known that optimal control problem of nonlinear systems can be addressed by solving the Hamilton-Jacobi-Bellman (HJB) equations. HJB equations can be solved by adaptive dynamic programming (ADP) (Werbos, 1992) to remove the 'curse of dimensionality' with approximators, such as NNs. There are many synonyms used for ADP, such as adaptive dynamic programming (Wang, Zhang, & Liu, 2009), approximate dynamic programming (Al-Tamimi et al., 2008), neuro-dynamic programming (Bertsekas & Tsitsiklis, 1995), adaptive critic designs (Prokhorov & Wunsch, 1997) and reinforcement learning (Kaelbling, Littman, & Moore, 1996). Werbos (1992) classified ADP approaches into heuristic dynamic programming (HDP), dual heuristic dynamic programming (DHP), action-dependent HDP (ADHDP) and action-dependent DHP (ADDHP). After that, two other approaches called globalised

CONTACT Bo Zhao zhaobo@ia.ac.cn

DHP (GDHP) and ADGDHP were proposed (Prokhorov & Wunsch, 1997). In recent few years, ADP algorithms were developed further to solve control problems of continuous-time systems (Liu, Li, Li, Wang, & Ma, 2015; Liu, Wang, & Li, 2014), discrete-time systems (Mehraeen & Jagannathan, 2011; Qin, Zhang, Luo, & Wang, 2014), external disturbances and uncertainties (Gao, Jiang, Jiang, & Chai, 2016; Wang, Liu, Mu, & Ma, 2016; Wang, Liu, Zhang, & Zhao, 2016), trajectory tracking (Enns & Si, 2003; Lin, Wei, & Liu, 2016; Mu, Sun, Song, & Yu, 2016; Wei & Liu, 2014; Zhang, Song, Wei, & Zhang, 2011; Zhang, Wei, & Luo, 2008), control input saturation (Yang, Liu, & Wang, 2014; Zhang, Luo, & Liu, 2009), fault tolerant (Zhao, Liu, & Li, 2016, 2017), time-delay (Zhang et al., 2011), zero-sum games (Fu, Fu, & Chai, 2015; Zhang, Cui, & Luo, 2013), event-driven systems (Wang, Mu, He, & Liu, 2016; Wang, Mu, Zhang, & Liu, 2016), etc.

In recent literature, ADP-based decentralised control problems have been tackled extensively. For linear interconnected systems, Jiang and Jiang (2012) and Bian, Jiang, and Jiang (2015) presented a decentralised control via robust ADP and policy iteration (PI) technique. Gao et al. (2016) developed a data-driven output-feedback control policy based on both PI and value iteration (VI) methods. Tlili and Braiek (2014) investigated a decentralised observation and control approach for linear interconnected systems with nonlinear interconnections. The control problem was formulated as an optimisation problem by LMI to compute the robust observation and control gain matrix simultaneously. Hioe, Hudon, and Bao (2014) utilised linear partial differential Hamilton-Jacobi equation to solve the robust nonlinear control problem which was transformed from the dissipativity shaping problem. Liu et al. (2014) constructed the cost functions for the isolated subsystems with the assumed known bounded interconnections. Then, a decentralised control strategy was developed to stabilise continuous-time nonlinear interconnected large-scale systems. Furthermore, Wang et al. (2016) considered the interconnected subsystems as a whole system and constructed a cost function for the overall plant. Then, they developed the decentralised guaranteed cost control by solving the modified HJB equation. For unknown nonlinear interconnected systems, Liu et al. (2015) established an online model-free integral PI algorithm based decentralised control scheme via actor-critic technique. Lu, Si, and Xie (2008) applied the direct HDP to address the coordinated control for large power systems with uncertainties. Yang et al. (2014) designed a tracking control with filtered tracking error by using direct HDP. For decentralised tracking control (DTC) problem, Mehraeen and Jagannathan (2011) proposed a decentralised nearly optimal controller using online tuned action NN and critic NN by assuming that the input gain matrix was known and the unknown interconnection was weak. From the aforementioned literature, the optimal tracking control commonly consists of the feedforward controller and the feedback controller (Park, Choi, & Lee, 1996). The feedforward controller requires a priori knowledge of the system dynamics, while the feedback controller can be derived by only utilising ADP methods. However, we can observe that the existing works on DTC via ADP mainly focused on systems with known dynamics. Since DTC for large-scale systems has wide potential in practice, only a few results based on ADP have been carried out, and it is still an open problem to be solved.

Motivated by Mehraeen and Jagannathan (2011), Zhang, Cui, Zhang, and Luo (2011) and Liu et al. (2014), in this paper, a DTC scheme via observer-critic structure-based ADP is proposed for unknown large-scale nonlinear systems. The developed decentralised control consists of local desired control, local tracking error control and an adaptive robustifying compensator. In order to remove the assumptions on boundedness and matched condition of interconnections, the desired trajectories of coupled subsystems are shared to substitute their actual ones. Then, the substituted subsystem dynamics is identified by establishing a local NN observer. It helps to derive the local desired control, as well as the control input matrix of the local tracking error control. Together with the solution to HJB equation based on the critic NN approximated value function, the local tracking error control can be obtained directly. The overall error, which contains the substitution error, observation error and approximation error of local tracking error control, is compensated by an adaptive robustifying term. The proposed DTC can guarantee the tracking error of the closed-loop system to be asymptotically stable via Lyapunov's direct method. Two simulation examples are provided to demonstrate the effectiveness of the proposed scheme.

The main contributions of this paper have the following four aspects:

(1) To the best of our knowledge, it is the first time to extend the ADP approach to solve the DTC problem for unknown large-scale nonlinear systems. This method establishes the local NN observer to identify the unknown subsystem dynamics, which helps derive not only the local desired control, but also the local tracking error control. The controllers have the similar structures, which are different from existing methods (Lin et al., 2016; Mehraeen & Jagannathan, 2011; Yang, Liu, Wei, & Wang, 2016).
(2) Unlike existing methods (Gao et al., 2016; Liu et al., 2014, 2015), the states of coupled subsystems are substituted by their desired states. Thus, the assumptions on the boundedness and matched condition of interconnections can be relaxed.
(3) The local tracking error control is derived by combining the identified control input matrix with the critic NN approximated value function. Therefore, the action NN, which is commonly adopted in existing methods, is not required anymore.
(4) The substitution error, observation error and approximation error can be compensated simultaneously by employing an adaptive robustifying term, and the tracking error of the closed-loop system can be guaranteed to be asymptotically stable.

The rest of the paper is organised as follows. In Section 2, the problem statement is presented, and the subsystem dynamics are transformed for simplifying the controller design. In Section 3, the detailed design procedure of the DTC that consists of the local desired control, the local tracking error controller and a compensator is given. Then, the stability analysis is provided via Lyapunov's direct method. In Section 4, two simulation examples are employed to verify the effectiveness of the developed scheme. In Section 5, the conclusion is drawn.

2. Problem statement

Consider unknown large-scale nonlinear systems that are composed of N interconnected subsystems, whose ith ($i = 1, 2, \ldots, N$) subsystem can be described by

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(x_i(t)) + h_i(x(t)), \quad (1)$$

where $x_i(t) = [x_{i1}(t), x_{i2}(t), \ldots, x_{i(n_i)}(t)]^{\mathsf{T}} \in \mathbb{R}^{n_i}$, $i = 1, \ldots, N$ and $u_i(x_i(t)) \in \mathbb{R}^{m_i}$ are the state vector and control input of the ith subsystem, respectively; $x(t) = [x_1^{\mathsf{T}}(t), \ldots, x_N^{\mathsf{T}}(t)]^{\mathsf{T}} \in \mathbb{R}^n$ is the overall system state vector with $n = \sum_{i=1}^{N} n_i$; $f_i(x_i(t))$, $g_i(x_i(t))$ and $h_i(x(t))$ are unknown nonlinear internal dynamics, input gain matrix and interconnection term, respectively.

Assumption 2.1: *The nonlinear functions $f_i(x_i(t))$, $g_i(x_i(t))$ and $h_i(x(t))$ are Lipschitz and continuous in their arguments with $f_i(0) = 0$, and the subsystem (1) is controllable.*

Unlike assuming $h_i(x)$ to be bounded and satisfying the matching conditions (Liu et al., 2014, 2015), the desired trajectories of the coupled subsystems are employed to substitute their actual states, so the interconnection term can be expressed as

$$h_i(x) = h_i(x_i, x_{jd}) + \Delta h_i(x, x_{jd}), \quad (2)$$

where x_{jd} denotes the desired trajectories of the coupled subsystems with $j = 1, \ldots, i-1, i+1, \ldots N$. $\Delta h_i(x, x_{jd}) = h_i(x) - h_i(x_i, x_{jd})$ denotes the substitution error. Thus, (1) becomes

$$\dot{x}_i = F_i(x_i, x_{jd}) + g_i(x_i)u_i(x_i) + \Delta h_i(x, x_{jd}), \quad (3)$$

where $F_i(x_i, x_{jd}) = f_i(x_i) + h_i(x_i, x_{jd})$, which is still Lipschitz continuous on a set $\Omega_i \in \mathbb{R}^{n_i}$ according to Assumption 2.1. Since the interconnection satisfies the global Lipschitz condition, which implies

$$\left\| \Delta h_i(x, x_{jd}) \right\| \le \sum_{j=1, j\neq i}^{n} d_{ij} E_j, \quad (4)$$

where $E_j = \|x_j - x_{jd}\|$, and $d_{ij} \ge 0$ is an unknown global Lipschitz constant.

The objective of this paper is to find a set of decentralised tracking control policies $u_1(x_1), \ldots, u_i(x_i), \ldots, u_N(x_N)$ such that the states of the overall unknown large-scale nonlinear system track the desired trajectories.

Remark 2.1: We notice that the interconnection term is approximated by signals from the local subsystem and the desired signals from coupled subsystems. It is worthy pointing out that the desired signals, which are decided according to the control objective, are shared to each subsystem before the system runs.

For ith subsystem, define the tracking error as

$$e_i = x_i - x_{id}, \quad (5)$$

where x_{id} is the predefined desired trajectory.

Combining (5) with (2), the tracking error dynamics can be expressed as

$$\dot{e}_i = \dot{x}_i - \dot{x}_{id}. \quad (6)$$

Thus, associated with the tracking error dynamics (6), the local tracking error control policy should minimise the following local infinite horizon value function

$$V_i(e_i(t)) = \int_t^{\infty} U_i\left(e_i(\tau), u_{ie}(\tau)\right) \mathrm{d}\tau, \quad (7)$$

where $U_i(e_i(t), u_{ie}(e_i)) = e_i^{\mathsf{T}}(t)Q_i e_i(t) + u_{ie}^{\mathsf{T}}(e_i)R_i u_{ie}(e_i)$ is the local utility function, $U_i(0, 0) = 0$, and $U_i(e_i, u_{ie}) \ge 0$ for all e_i and u_{ie}, in which $Q_i \in \mathbb{R}^{n_i \times n_i}$ and $R_i \in \mathbb{R}^{m_i}$ are positive definite matrices, $u_{ie} = u_i(x_i) - u_{id}(x_{id})$ is the local control input error, and $u_{id}(x_{id})$ is the local desired control input.

3. Decentralised tracking controller design

The detailed design procedure of DTC in the optimal manner for unknown large-scale nonlinear system is given in this section.

3.1. Decentralised controller design of systems with known dynamics

In order to achieve the control objective, the controller design procedure for unknown systems follows the strategy for systems with known dynamics. In this subsection, the DTC for systems with known dynamics is introduced.

The optimal tracking control problem should be solved for N isolated subsystems. For systems with known dynamics, the local desired control input can be obtain by (3) as

$$u_{id}(x_{id}) = g_i^{+}(x_{id})\left(\dot{x}_{id} - F_i(x_d) - \Delta h_i(x_d)\right), \quad (8)$$

where $x_d = [x_{1d}^{\mathsf{T}}, \ldots, x_{Nd}^{\mathsf{T}}]^{\mathsf{T}}$, and $g_i^{+}(\cdot)$ is the Moore-Penrose pseudo-inverse of $g_i(\cdot)$.

According to the optimal control theory, the designed tracking error control policy must not only ensure the tracking error converge to a small neighbourhood on Ω_i, but also guarantee the local value function (7) to be finite. In other words, the tracking error control policy should be admissible.

Definition 3.1: For local tracking error dynamics (6), a tracking error control policy $\mu_{ie}(e_i)$ is said to be admissible if $\mu_{ie}(e_i)$ is continuous on a set Ω_i with $\mu_{ie}(0) = 0$, $\mu_{ie}(e_i)$ ensures the convergence of the ith subsystem (1) on Ω_i, and $J_i(e_i(t))$ is finite for all $e_i \in \Omega_i$.

For any admissible control policy $\mu_i(e_i) \in \psi_i(\Omega_i)$ of subsystem (1), where $\psi_i(\Omega_i)$ is the set of admissible control, if the local value function

$$V_i(e_i(t)) = \int_t^{\infty} U_i\left(e_i(\tau), \mu_{ie}(\tau)\right) \mathrm{d}\tau \quad (9)$$

is continuously differentiable, then the infinitesimal version of (9) is the so-called Lyapunov equation

$$0 = U_i\left(e_i, \mu_{ie}\right) + \left(\nabla V_i(e_i)\right)^{\mathsf{T}} \dot{e}_i \tag{10}$$

with $V_i(0) = 0$, and the term $\nabla V_i(e_i)$ denotes the partial derivative of $V_i(e_i)$ with respect to the local tracking error e_i, i.e. $\nabla V_i(e_i) = \partial V_i(e_i)/\partial e_i$.

The Hamiltonian of the optimal control problem and the optimal value function can be formulated as

$$H_i\left(e_i, \mu_{ie}, \nabla V_i(e_i)\right) = U_i\left(e_i, \mu_{ie}\right) + \left(\nabla V_i(e_i)\right)^{\mathsf{T}} \dot{e}_i,$$

and

$$V_i^*(e_i) = \min_{\mu_{ie} \in \psi_i(e_i)} \int_t^{\infty} U_i\left(e_i(\tau), \mu_{ie}(\tau)\right) \mathrm{d}\tau.$$

Thus,

$$0 = \min_{\mu_{ie} \in \psi_i(e_i)} H_i\left(e_i, \mu_{ie}, \nabla V_i^*(e_i)\right),$$

where $\nabla J_i^*(e_i) = \partial J_i^*(e_i)/\partial e_i$. If the solution $V_i^*(e_i)$ exists and is continuously differentiable, the local desired optimal tracking error control can be described as

$$u_{ie}^*(e_i) = -\frac{1}{2} R_i^{-1} g_i^{\mathsf{T}}(x_i) \nabla V_i^*(e_i). \tag{11}$$

Therefore, the desired DTC can be expressed as

$$u_i(x_i) = u_{ie}^*(e_i) + u_{id}(x_{id}). \tag{12}$$

Remark 3.1: In existing results on dealing with trajectory tracking problems, ADP-based controllers as in (12) commonly contain two parts, namely desired control and desired tracking error control. Inspired by that, in this paper, we concern the DTC scheme in a similar way. That is, the DTC for systems with available dynamics in Section 3.1 provides a design strategy for that of large-scale system with unknown dynamics. The design of DTC for such systems will be detailed in the following subsections.

3.2. Neural network observer-based unknown subsystem identification

The purpose of designing a local NN observer is to identify dynamics of the unknown subsystems. With the help of identified subsystems, the DTC can be designed for unknown large-scale nonlinear systems in a similar control structure as in Section 3.1.

For the ith subsystem of the unknown large-scale nonlinear system (1), it can be identified by a local NN observer, which can be established as

$$\dot{\hat{x}}_i = \hat{F}_i(\hat{x}_i, x_{jd}) + \hat{g}_i(\hat{x}_i) u_i(x_i) + K_{io}(x_i - \hat{x}_i), \tag{13}$$

where $\hat{x}_i = [\hat{x}_{i1}, \ldots, \hat{x}_{i(n_i)}]^{\mathsf{T}} \in \mathbb{R}^{n_i}$ is the state vector of the developed observer, $\hat{F}_i(\hat{x}_i, x_{jd})$ and $\hat{g}_i(\hat{x}_i)$ are the observation of nonlinear dynamics $F_i(x_i, x_{jd})$ and $g_i(x_i)$, respectively; $K_{io} = \mathrm{diag}[k_{i1o}, k_{i2o}]$ is a positive definite observation gain matrix.

Define the observation error vector $e_{io} = x_i - \hat{x}_i$, combining (3) with (13), the observation error dynamics can be described as

$$\begin{aligned}\dot{e}_{io} = {} & F_i(x_i, x_{jd}) - \hat{F}_i(\hat{x}_i, x_{jd}) + \left(g_i(x_i) - \hat{g}_i(\hat{x}_i)\right) u_i(x_i) \\ & + \Delta h_i(x, x_{jd}) - K_{io} e_{io}.\end{aligned}$$

The nonlinear unknown terms $F_i(x_i, x_{jd})$ and $g_i(x_i)$ are approximated by two ideal radial basis function (RBF) NNs as

$$F_i(x_i, x_{jd}) = W_{if}^{\mathsf{T}} \sigma_{if}(x_i, x_{jd}) + \varepsilon_{if}, \ \left\|\varepsilon_{if}\right\| \le \varepsilon_{i1}, \tag{14}$$

$$g_i(x_i) = W_{ig}^{\mathsf{T}} \sigma_{ig}(x_i) + \varepsilon_{ig}, \ \left\|\varepsilon_{ig}\right\| \le \varepsilon_{i2}, \tag{15}$$

where W_{if} and W_{ig} are ideal weight vectors from the hidden layer to the output layer, $\sigma_{if}(x_i, x_{jd})$ and $\sigma_{ig}(x_i)$ are basis functions, ϵ_{if} and ϵ_{ig} are approximation errors, and ϵ_{i1} and ϵ_{i2} are unknown positive constants.

Let $\hat{W}_{if}$ and $\hat{W}_{ig}$ be the estimations of W_{if} and W_{ig}, respectively. We have

$$\hat{F}_i(\hat{x}_i, x_{jd}) = \hat{W}_{if}^{\mathsf{T}} \sigma_{if}(\hat{x}_i, x_{jd}), \tag{16}$$

$$\hat{g}_i(\hat{x}_i) = \hat{W}_{ig}^{\mathsf{T}} \sigma_{ig}(\hat{x}_i), \tag{17}$$

where $\hat{W}_{if}$ and $\hat{W}_{ig}$ can be updated by the adaptive laws as

$$\dot{\hat{W}}_{if} = \Gamma_{if} e_{io} \sigma_{if}(\hat{x}_i, x_{jd}), \tag{18}$$

$$\dot{\hat{W}}_{ig} = \Gamma_{ig} e_{io} \sigma_{ig}(\hat{x}_i) u_i, \tag{19}$$

where Γ_{if} and Γ_{ig} are positive constants.

Combining (14) and (16), (15) and (17), we have

$$\begin{aligned}F_i(x_i, x_{jd}) - \hat{F}_i(\hat{x}_i, x_{jd}) = {} & W_{if}^{\mathsf{T}} \tilde{\sigma}_{if}(x_i, \hat{x}_i, x_{jd}) \\ & + \tilde{W}_{if}^{\mathsf{T}} \sigma_{if}(\hat{x}_i, x_{jd}) + \varepsilon_{if},\end{aligned} \tag{20}$$

$$g_i(x_i) - \hat{g}_i(\hat{x}_i) = W_{ig}^{\mathsf{T}} \tilde{\sigma}_{ig}(x_i, \hat{x}_i) + \tilde{W}_{ig}^{\mathsf{T}} \sigma_{ig}(\hat{x}_i) + \varepsilon_{ig}, \tag{21}$$

where $\tilde{W}_{if} = W_{if} - \hat{W}_{if}$ and $\tilde{W}_{ig} = W_{ig} - \hat{W}_{ig}$ are the weight estimation errors, $\tilde{\sigma}_{if}(x_i, \hat{x}_i, x_{jd}) = \sigma_{if}(x_i, x_{jd}) - \hat{\sigma}_{if}(\hat{x}_i, x_{jd})$ and $\tilde{\sigma}_{ig}(x_i, \hat{x}_i) = \sigma_{ig}(x_i) - \sigma_{ig}(\hat{x}_i)$ are the estimation errors of RBFs, respectively.

Theorem 3.1: *For interconnected subsystem* (1), *the developed local NN observer can guarantee the observation error e_{io} to be uniformly ultimately bounded* (*UUB*) *with the updating laws* (18)–(19).

Proof: Select a Lyapunov function candidate as

$$L_{i1} = \frac{1}{2} e_{io}^{\mathsf{T}} e_{io} + \frac{1}{2} \tilde{W}_{if}^{\mathsf{T}} \Gamma_{if}^{-1} \tilde{W}_{if} + \frac{1}{2} \tilde{W}_{ig}^{\mathsf{T}} \Gamma_{ig}^{-1} \tilde{W}_{ig}. \tag{22}$$

The time derivative of (22) is

$$\begin{aligned}\dot{L}_{i1} &= e_{io}^{\mathsf{T}}\dot{e}_{io} - \tilde{W}_{if}^{\mathsf{T}}\Gamma_{if}^{-1}\dot{\hat{W}}_{if} - \tilde{W}_{ig}^{\mathsf{T}}\Gamma_{ig}^{-1}\dot{\hat{W}}_{ig} \\ &= e_{io}^{\mathsf{T}}(F_i(x_i, x_{jd}) - \hat{F}_i(\hat{x}_i, x_{jd}) + \left(g_i(x_i) - \hat{g}_i(\hat{x}_i)\right) u_i(x_i) \\ &\quad + \Delta h_i(x, x_{jd}) - K_{io}e_{io}) \\ &\quad - \tilde{W}_{if}^{\mathsf{T}}\Gamma_{if}^{-1}\dot{\hat{W}}_{if} - \tilde{W}_{ig}^{\mathsf{T}}\Gamma_{ig}^{-1}\dot{\hat{W}}_{ig}. \end{aligned} \tag{23}$$

Combining (23) with (20) and (21), we have

$$\begin{aligned}\dot{L}_{i1} &= e_{io}^{\mathsf{T}}\big(\tilde{W}_{if}^{\mathsf{T}}\sigma_{if}(\hat{x}_i, x_{jd}) + \tilde{W}_{ig}^{\mathsf{T}}\sigma_{ig}(\hat{x}_i)u_i(x_i) + \varpi_{i1} + \Delta h_i(x, x_{jd})\big) \\ &\quad - e_{io}^{\mathsf{T}}K_{io}e_{io} - \tilde{W}_{if}^{\mathsf{T}}\Gamma_{if}^{-1}\dot{\hat{W}}_{if} - \tilde{W}_{ig}^{\mathsf{T}}\Gamma_{ig}^{-1}\dot{\hat{W}}_{ig}, \end{aligned} \tag{24}$$

where $\varpi_{i1} = W_{if}^{\mathsf{T}}\tilde{\sigma}_{if}(x_i, \hat{x}_i, x_{jd}) + \varepsilon_{if} + (W_{ig}^{\mathsf{T}}\tilde{\sigma}_{ig}(x_i, \hat{x}_i) + \varepsilon_{ig})u_i$ denotes the overall NN approximation error.

Substituting (18) and (19) into (24), we have

$$\dot{L}_{i1} = e_{io}^{\mathsf{T}}\left(\varpi_{i1} + \Delta h_i(x, x_{jd})\right) - e_{io}^{\mathsf{T}}K_{io}e_{io}. \tag{25}$$

■

Assumption 3.1: *The defined approximation ϖ_{i1} is norm-bounded, i.e.$\|\varpi_{i1}\| \le \eta_{i1}$, where η_{i1} is an unknown positive constant.*

Letting $\eta_{i2} = \sum_{j=1, j\neq i}^{N} d_{ij}E_j$, according to (4), (25) becomes

$$\begin{aligned}\dot{L}_{i1} \le&\ \|e_{io}\|\left(\eta_{i1} + \eta_{i2}\right) - \lambda_{\min}(K_{io})\,\|e_{io}\|^2 \\ =& -\|e_{io}\|\left(\lambda_{\min}(K_{io})\,\|e_{io}\| - (\eta_{i1} + \eta_{i2})\right),\end{aligned}$$

where $\lambda_{\min}(K_{io})$ denotes the minimum eigenvalue of K_{io}. We can observe that $\dot{L}_{i1} \le 0$ when e_{io} lies outside of the compact set

$$\Omega_{e_{io}} = \left\{ e_{io} : \|e_{io}\| \le \frac{\eta_{i1} + \eta_{i2}}{\lambda_{\min}(K_{io})} \right\}.$$

Therefore, according to Lyapunov's direct method, the observation error e_{io} is UUB. This completes the proof.

Remark 3.2: It should be pointed out that controllers of unknown nonlinear systems were commonly designed by introducing an action NN in many previous works (Liu et al., 2014, 2015). Different from these, in our approach, the control law is designed in a similar control structure to model-based controller by using the identified input gain matrix via local NN observer. On the other hand, the key point for obtaining the desired trajectory tracking control for unknown systems is to find the system dynamics, which can be identified by the present local NN observer.

3.3. *Decentralised tracking controller design for unknown large-scale nonlinear systems*

From the present observer (7), the identifier of the ith subsystem should be expressed as

$$\dot{\hat{x}}_i = \hat{F}_i(\hat{x}_i, x_{jd}) + \hat{g}_i(\hat{x}_i)u_i(\hat{x}_i), \tag{26}$$

where $u_{io}(\hat{x}_i)$ is the control input of the identifier. Thus, the desired control input of identifier $u_{id}(x_{id})$ can be obtained by

$$u_{id}(x_{id}) = \hat{g}_i^{+}(x_{id})\big(\dot{x}_{id} - \hat{F}_i(x_d)\big). \tag{27}$$

Since the value function is highly nonlinear and nonanalytic, it can be approximated by NNs, which are powerful tools for approximating nonlinear functions. For the ith subsystem, a critic NN is employed to approximate the corresponding assumed continuous local value function on the compact set Ω_i as

$$V_i(e_i) = W_{ic}^{\mathsf{T}}\sigma_{ic}(e_i) + \varepsilon_{ic}(e_i), \tag{28}$$

where $W_{ic} \in \mathbb{R}^{l_i \times n_i}$ is the ideal weight vector, $\sigma_{ic}(e_i) \in \mathbb{R}^{l_i}$ is the activation function, l_i is the number of neurons in the hidden layer and $\epsilon_{ic}(e_i)$ is the approximation error of NN. Then, the gradient of $V_i(e_i)$ with respect to e_i is

$$\nabla V_i(e_i) = \left(\nabla\sigma_{ic}(e_i)\right)^{\mathsf{T}} W_{ic} + \nabla\varepsilon_{ic}(e_i), \tag{29}$$

where $\nabla\sigma_{ic}(e_i) = \partial\sigma_{ic}(e_i)/\partial e_i \in \mathbb{R}^{l_i}$ and $\nabla\epsilon_{ic}(e_i)$ are the gradients of the activation function and the approximation error, respectively.

Combining (10) with (29), we have

$$0 = U_i\left(e_i, \mu_{ie}\right) + \left(\left(\nabla\sigma_{ic}(e_i)\right)^{\mathsf{T}} W_{ic} + \nabla\varepsilon_{ic}(e_i)\right)^{\mathsf{T}} \dot{e}_i.$$

Therefore, the local Hamiltonian can be expressed as

$$\begin{aligned}H_i\left(e_i, \mu_{ie}, W_{ic}\right) &= U_i\left(e_i, \mu_{ie}\right) + W_{ic}^{\mathsf{T}}\nabla\sigma_i(e_i)\dot{e}_i \\ &= -\nabla\varepsilon_{ic}(x_i)\dot{e}_i = e_{icH},\end{aligned} \tag{30}$$

where e_{icH} is the residual error caused by NN approximation.

The critic NN (28) can be approximated as

$$\hat{V}_i(e_i) = \hat{W}_{ic}^{\mathsf{T}}\sigma_{ic}(e_i), \tag{31}$$

where $\hat{W}_{ic} \in \mathbb{R}^{l_i \times n_i}$ is the weight estimation.

Then, the gradient of (31) with respect to e_i is

$$\nabla\hat{V}_i(e_i) = \left(\nabla\sigma_{ic}(e_i)\right)^{\mathsf{T}} \hat{W}_{ic}.$$

Therefore, the approximate local Hamiltonian can be expressed as

$$H_i(e_i, \mu_{ie}, \hat{W}_{ic}) = U_i\left(e_i, \mu_{ie}\right) + \hat{W}_{ic}^{\mathsf{T}}\nabla\sigma_i(e_i)\dot{e}_i = e_{ic}. \tag{32}$$

Let $\theta_i = \nabla\sigma_i(e_i)\dot{e}_i$. From (30) and (32), we have

$$e_{ic} = e_{icH} - \tilde{W}_{ic}^{\mathsf{T}}\theta_i,$$

where $\tilde{W}_{ic} = W_{ic} - \hat{W}_{ic}$, and it can be updated as

$$\dot{\tilde{W}}_{ic} = -\dot{\hat{W}}_{ic} = l_{i1}\big(e_{icH} - \tilde{W}_{icH}^{\mathsf{T}}\theta_i\big)\theta_i, \tag{33}$$

where $l_{i1} > 0$ is the learning rate of the critic NN.

To obtain the updating rule of the critic NN weight vector $\hat{W}_{ic}$, with the steepest decent algorithm, the local objective function $E_{ic} = \frac{1}{2}e_{ic}^{\mathsf{T}}e_{ic}$ should be minimised as

$$\dot{\hat{W}}_{ic} = -\dot{\tilde{W}}_{ic} = -l_{i1}e_{ic}\theta_i. \tag{34}$$

Therefore, the ideal local optimal tracking error control can be derived as

$$\mu_{ie}(e_i) = -\frac{1}{2}R_i^{-1}g_i^{\mathsf{T}}(x_i)\left((\nabla\sigma_{ic}(e_i))^{\mathsf{T}}W_{ic} + \nabla\varepsilon_{ic}(e_i)\right).$$

Since the nonlinear system is unknown, consider the identified control input matrix (17) and the approximate critic NN (31), the local optimal tracking error control can be expressed as

$$\hat{\mu}_{ie}(e_i) = -\frac{1}{2}R_i^{-1}\hat{g}_i^{\mathsf{T}}(x_i)\,(\nabla\sigma_{ic}(e_i))^{\mathsf{T}}\,\hat{W}_{ic}. \tag{35}$$

Theorem 3.2: *For ith interconnected subsystem (1), the weight approximation error $\tilde{W}_{ic}$ can be guaranteed to be UUB as long as the weights of the critic NN are updated by (34).*

Proof: Select the Lyapunov function candidate as

$$L_{i2} = \frac{1}{2l_{i1}}\tilde{W}_{ic}^{\mathsf{T}}\tilde{W}_{ic}. \tag{36}$$

Along the solutions of (33), the time derivative of (36) is

$$\begin{aligned}\dot{L}_{i2} &= \frac{1}{l_{i1}}\tilde{W}_{ic}^{\mathsf{T}}\dot{\tilde{W}}_{ic}\\ &= \tilde{W}_{ic}^{\mathsf{T}}e_{icH}\theta_i - \left\|\tilde{W}_{ic}\theta_i\right\|^2\\ &\le \frac{1}{2}e_{icH}^2 - \frac{1}{2}\left\|\tilde{W}_{ic}\theta_i\right\|^2.\end{aligned}$$

Assume $\|\theta_i\| \le \theta_{iM}$. Hence, $\dot{L}_{i2} < 0$ whenever the approximation error of the critic NN $\tilde{W}_{ic}$ lies outside of the compact set

$$\Omega_{\tilde{W}_{ic}} = \left\{\tilde{W}_{ic} : \left\|\tilde{W}_{ic}\right\| \le \left\|\frac{e_{icH}}{\theta_{iM}}\right\|\right\}.$$

According to Lyapunov's direct method, the weight approximation error is UUB. This completes the proof. ■

Taking the difference between (8) and (27) as well as the approximation error between (11) and (35) into account, they may cause the system performance degradation or even destroy the system stability. Thus, they should be compensated by an adaptive robustifying term as

$$u_{ic} = -\hat{g}_i^{+}(x_{id})\mathrm{sgn}(e_i)\hat{\varpi}_i, \tag{37}$$

where $\mathrm{sgn}(e_i) = [\mathrm{sgn}(e_{i1}), \ldots, \mathrm{sgn}(e_{i(n_i)})]^{\mathsf{T}}$, $\hat{\varpi}_i$ is the estimation of overall error ϖ_i, which will be defined later. It can be updated by the following adaptive law

$$\dot{\hat{\varpi}}_i = \Gamma_{i\varpi}\sum_{k=1}^{n_i}|e_{ik}|, \tag{38}$$

where $\Gamma_{i\varpi}$ is a positive constant.

In summary, the overall DTC can be developed as

$$u_i = u_{id} + \hat{\mu}_{ie} + u_{ic}. \tag{39}$$

The control architecture of the proposed DTC for unknown large-scale nonlinear systems via observer-critic structure-based ADP is shown in Figure 1.

Remark 3.3: In local identifier design, RBFNN is employed to construct local NN observers, since the convergence rate is higher than that of back propagation (BP) NN. On the other hand, the local tracking error controller requires the partial derivative of local critic NN, which has heavy computational burden. To trade off between the convergence rate and computational burden, BPNN is selected for local critic NN. Thus, different structures are chosen for these two NNs.

Remark 3.4: Actually, the proposed DTC scheme is an online algorithm. On the one hand, the unknown dynamics of large-scale nonlinear systems can be identified by the developed local NN observer (13) in real time, which helps obtain the subsystem dynamics as in (26). Therefore, the local desired control can be derived in real time. On the other hand, the local tracking error control is obtained by employing local critic NN, which is also trained online. Therefore, the strategy of the proposed DTC is online.

3.4. Stability analysis

Theorem 3.3: *Consider the unknown large-scale nonlinear systems which are composed of N subsystems as in (1) with the local value function (7). The developed observer-critic structure-based DTC (39) can guarantee the tracking error of closed-loop system to converge to zero asymptotically.*

Proof: Select the Lyapunov function candidate as

$$L_{i3} = \frac{1}{2}e_i^{\mathsf{T}}e_i + V_i(e_i) + \Gamma_{i\varpi}^{-1}\tilde{\varpi}_i^2. \tag{40}$$

As $F_i(\cdot)$ is locally Lipschitz, there exists a positive constant η_{if} such that $\|F_i(x_i, x_{jd}) - F_i(x_d)\| \le \eta_{if}\|e_i\|$. Assuming that $\left\|\hat{g}_i(x_{id})\right\| \le \eta_{ig}$ and denoting $\hat{\mu}_{ie} = \mu_{ie} - \tilde{\mu}_{ie}$, the time derivative of (40) becomes

$$\begin{aligned}\dot{L}_{i3} &= e_i^{\mathsf{T}}\dot{e}_i + \nabla V_i(e_i)\dot{e}_i - \Gamma_{i\varpi}^{-1}\dot{\hat{\varpi}}_i\tilde{\varpi}_i\\ &= e_i^{\mathsf{T}}\big(F_i(x_i, x_{jd}) - F_i(x_d) + F_i(x_d) - \hat{F}_i(x_d)\\ &\quad + \Delta h_i(x, x_{jd})\big) - U_i(e_i, \mu_{ie}) - \Gamma_{i\varpi}^{-1}\dot{\hat{\varpi}}_i\tilde{\varpi}_i\\ &\quad + e_i^{\mathsf{T}}\left(\left(g_i(x_i) - \hat{g}_i(x_{id}) + \hat{g}_i(x_{id})\right)u_i(x_i) - \hat{g}_i(x_{id})u_{id}(x_{id})\right)\\ &\le \eta_{if}\left\|e_i\right\|^2 + e_i^{\mathsf{T}}\big(\tilde{F}_i(x_d) + \tilde{g}_i(x_{id})u_i(x_i)\\ &\quad + \Delta h_i(x, x_{jd}) - \hat{g}_i(x_{id})\tilde{\mu}_{ie}\big)\\ &\quad + \eta_{ig}\left\|e_i\right\|\left\|\mu_{ie}\right\| + e_i^{\mathsf{T}}\hat{g}_i(x_{id})u_{ic} - U_i(e_i, \mu_{ie}) - \Gamma_{i\varpi}^{-1}\dot{\hat{\varpi}}_i\tilde{\varpi}_i\\ &= \eta_{if}\left\|e_i\right\|^2 + e_i^{\mathsf{T}}\big(\tilde{F}_i(x_d) + \tilde{g}_i(x_{id})u_i(x_i)\\ &\quad + \Delta h_i(x, x_{jd}) - \hat{g}_i(x_{id})\tilde{\mu}_{ie}\big) + \frac{1}{2}\left\|e_i\right\|^2\\ &\quad + e_i^{\mathsf{T}}\hat{g}_i(x_{id})u_{ic} - \lambda_{\min}(Q_i)\left\|e_i\right\|^2\\ &\quad - \left(\lambda_{\min}(R_i) - \eta_{ig}^2\right)\left\|\mu_{ie}\right\|^2 - \Gamma_{i\varpi}^{-1}\tilde{\varpi}_i\dot{\hat{\varpi}}_i.\end{aligned} \tag{41}$$

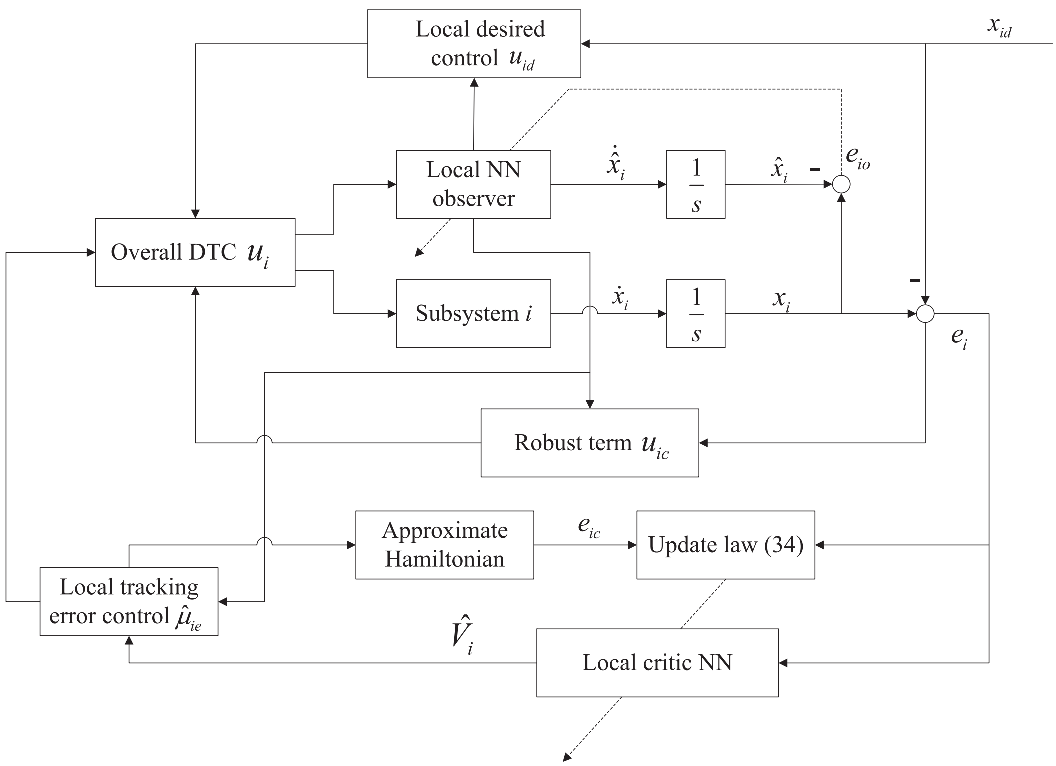

Figure 1. The control architecture of the proposed DTC.

Denoting $\delta_i = \tilde{F}_i(x_d) + \tilde{g}_i(x_i, x_{id})u_i(x_i) + \Delta h_i(x, x_{jd}) - \hat{g}_i(x_{id})\tilde{\mu}_{ie}$ as the overall error, where $\tilde{g}_i(x_i, x_{id}) = g_i(x_i) - \hat{g}_i(x_{id})$ and δ_i is assumed to be the upper bounded, i.e. $\|\delta_i\| \le \omega_i$, we have

$$\begin{aligned}
\dot{L}_{i3} &= \eta_{if}\,\|e_i\|^2 + |e_i|\,\omega_i + \frac{1}{2}\,\|e_i\|^2 + e_i^{\mathsf{T}}\hat{g}_i(x_{id})u_{ic} \\
&\quad - \lambda_{\min}(Q_i)\,\|e_i\|^2 - \left(\lambda_{\min}(R_i) - \eta_{ig}^2\right)\|\mu_{ie}\|^2 - \Gamma_{iw}^{-1}\dot{\hat{\omega}}_i\tilde{\omega}_i \\
&\le \eta_{if}\,\|e_i\|^2 + \sum_{k=1}^{n_i}|e_{ik}|\,\omega_i + \frac{1}{2}\,\|e_i\|^2 + e_i^{\mathsf{T}}\hat{g}_i(x_{id})u_{ic} \\
&\quad - \lambda_{\min}(Q_i)\,\|e_i\|^2 - \left(\lambda_{\min}(R_i) - \eta_{ig}^2\right)\|\mu_{ie}\|^2 - \Gamma_{iw}^{-1}\dot{\hat{\omega}}_i\tilde{\omega}_i.
\end{aligned} \tag{42}$$

Substituting (37) into (42), and combining with (38), we have

$$\begin{aligned}
\dot{L}_{i3} &= \eta_{if}\,\|e_i\|^2 + \sum_{k=1}^{n_i}|e_{ik}|\,\tilde{\omega}_i + \frac{1}{2}\,\|e_i\|^2 \\
&\quad - \lambda_{\min}(Q_i)\,\|e_i\|^2 - \left(\lambda_{\min}(R_i) - \eta_{if}^2\right)\|\mu_{ie}\|^2 \\
&\quad - \Gamma_{iw}^{-1}\dot{\hat{\omega}}_i\tilde{\omega}_i \\
&= -\left(\lambda_{\min}(Q_i) - \eta_{if} - \frac{1}{2}\right)\|e_i\|^2 \\
&\quad - \left(\lambda_{\min}(R_i) - \eta_{ig}^2\right)\|\mu_{ie}\|^2.
\end{aligned}$$

We can observe that $\dot{L}_{i3} \le 0$ whenever the following conditions hold

$$\begin{cases} \lambda_{\min}(Q_i) \ge \eta_{if} + \dfrac{1}{2}, \\ \lambda_{\min}(R_i) \ge \eta_{ig}^2. \end{cases}$$

It implies that the developed observer-critic structure ADP based DTC (39) ensures the tracking errors of the unknown large-scale closed-loop system converge to zero asymptotically. This completes the proof. ■

Remark 3.5: We can see that δ_i includes the substitution error, observation error and approximation error of the local tracking error control. They can be considered as overall error, which is compensated simultaneously by (37). Therefore, the tracking errors of the unknown large-scale nonlinear system can converge to zero asymptotically.

Remark 3.6: Some observer-based ADP methods have been studied for optimal control (He & Jagannathan, 2005; Yang, Liu, & Wang, 2014), but action NNs are always employed to approximate the control law. Different from them, the proposed DTC scheme is obtained by the critic NN only, the training of action NN is no longer required. It implies that the computational burden can be reduced.

4. Simulation studies

To show the effectiveness of the developed DTC scheme, two examples are given in this section.

Example 4.1: Consider a hard spring connected parallel inverted pendulum system (Hua, Li, Wang, & Guan, 2015), whose model can be expressed as

$$\begin{cases} m_1 l_1^2 \ddot{\theta}_1 - m_1 g l_1 \sin\theta_1 + b_1\dot{\theta}_1 - Fa_1\cos(\theta_1-\beta) = \delta_1 u_1, \\ m_2 l_2^2 \ddot{\theta}_2 - m_2 g l_2 \sin\theta_2 + b_2\dot{\theta}_2 - Fa_2\cos(\theta_2-\beta) = \delta_2 u_2, \end{cases} \tag{43}$$

where b_1 and b_2 are damping coefficients, and

$$F = k\left\{1 + A^2\left(l_k - l_0\right)^2\right\}\left(l_k - l_0\right),$$

$$\left|A\left(l_k - l_0\right)\right| < 1,$$

$$\beta = \arctan\left(\frac{a_1\cos\theta_1 - a_2\cos\theta_2}{l_0 - a_1\sin\theta_1 + a_2\sin\theta_2}\right),$$

$$l_k = \left\{(l_0 - a_1\sin\theta_1 + a_2\sin\theta_2)^2 + (a_1\cos\theta_1 - a_2\cos\theta_2)^2\right\}^2.$$

In this simulation, the parameters of the coupled inverted pendulums are chosen as: $\delta_1 = \delta_2 = 1$, $m_1 = m_2 = 1kg$, $l_1 = l_2 = 0.5m$, $l_0 = 1m$, $g = 9.8m/s^2$, $b_1 = b_2 = 0.009$, $k = 30$, $A = 0.1$, and the spring position $a_1 = a_2 = 0.1$.

Let $x_i = [x_{i1}, x_{i2}]^{\mathsf{T}} = \left[\theta_{i1}, \dot{\theta}_{i1}\right]^{\mathsf{T}} \in \mathbb{R}^2$. The modified model (43) can be expressed as

$$\dot{x}_i = f_i(x_i) + g_i(x_i)u_i + h_i(x),$$

where $f_i(x_i) = \left[\begin{smallmatrix} x_{i2} \\ 5.88\sin x_{i1} - 0.036x_{i2}\end{smallmatrix}\right]$, $g_i(x_i) = \left[\begin{smallmatrix} 0 \\ \delta_i\end{smallmatrix}\right]$, $h_i(x) = \left[\begin{smallmatrix} 0 \\ 4Fa_i\cos(x_{i1}-\beta)\end{smallmatrix}\right]$.

In this simulation, the desired trajectories of the two subsystems can be given as

$$\begin{cases} x_{11d} = 0.5\cos(0.5t), \\ x_{21d} = 0.8\sin(0.3t + \pi/6). \end{cases}$$

Denote $X_i = [x_i, x_{jd}]^{\mathsf{T}}$. The basis functions of RBFNNs in the observers are chosen as Gaussian style as

$$\sigma_{if}(X_i) = \exp\left(\frac{-(X_i - c_{if})^{\mathsf{T}}(X_i - c_{if})}{b_{if}^2}\right),$$

$$\sigma_{ig}(x_i) = \exp\left(\frac{-(x_i - c_{ig})^{\mathsf{T}}(x_i - c_{ig})}{b_{ig}^2}\right),$$

where the centres of the basis functions are

$$c_{if} = \begin{bmatrix} -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \end{bmatrix},$$

$$c_{ig} = \begin{bmatrix} -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \end{bmatrix},$$

and the widths of the basis functions are $b_{if} = b_{ig} = 0.5$.

Let the initial states of the subsystems be $x_{10} = x_{20} = [1, 0]^{\mathsf{T}}$, the initial states of the observers be $\hat{x}_{10} = [2, -1]^{\mathsf{T}}$, $\hat{x}_{20} = [1.5, -0.5]^{\mathsf{T}}$, the observer gain matrix be $K_{io} = \text{diag}[k_{i1o}, k_{i2o}]$ $= \text{diag}[400, 1200]$ and the RBFNN weights learning rate of the observer be $\Gamma_{if} = \Gamma_{ig} = 0.002$. The local value function (7) is approximated by critic NN, whose structure is chosen as 2–3–1 with two input neurons, three hidden neurons and one output neuron, and the weight vector as $\hat{W}_{ic} = [\hat{W}_{ic1}, \hat{W}_{ic2}, \hat{W}_{ic3}]^{\mathsf{T}}$ with initial values $\hat{W}_{1c} = [0.4, 1.8, 1.2]^{\mathsf{T}}$ and $\hat{W}_{2c} = [0.2, 0.4, 0.2]^{\mathsf{T}}$. The activation function of the critic NN is chosen as $\sigma_{ic}(e_i) = [e_{i1}^2, e_{i1}e_{i2}, e_{i2}^2]$. Let the weight learning rates of the critic NN be $l_{i1} = 0.1$, the gain of the compensator (38) be $\Gamma_{iw} = 15$, and $Q_i = 2I_2$, $R_1 = 0.001I$, $R_2 = 0.0001I$, where I_n denotes the identity matrix with appropriate dimensions.

The simulation results are shown as Figures 2–5. Figure 2 describes that the observation errors converge to a small region

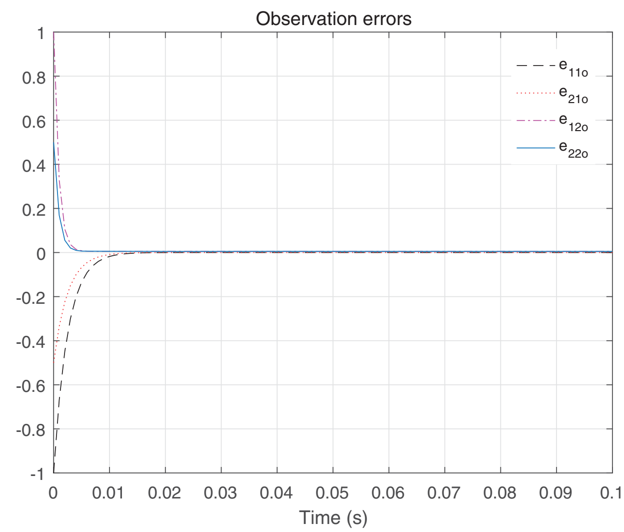

Figure 2. The observation errors by using the neural network observer of Example 4.1.

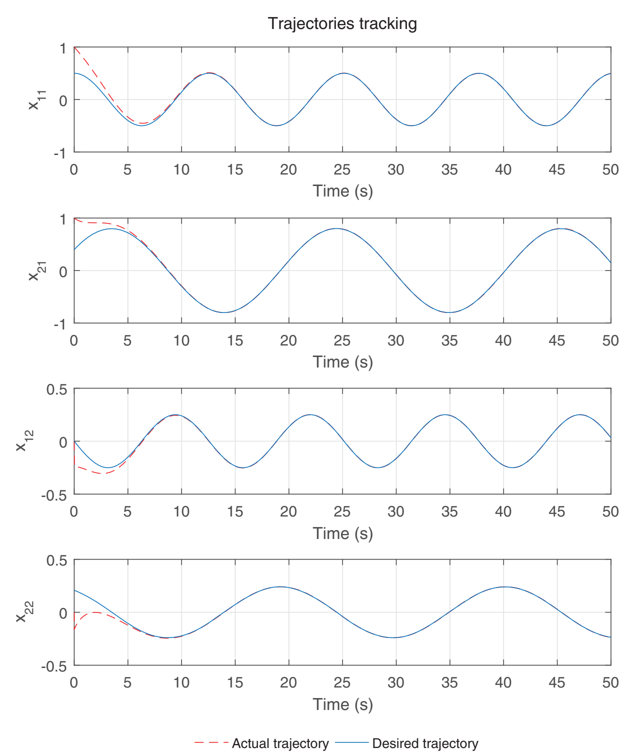

Figure 3. The trajectories tracking performance of Example 4.1.

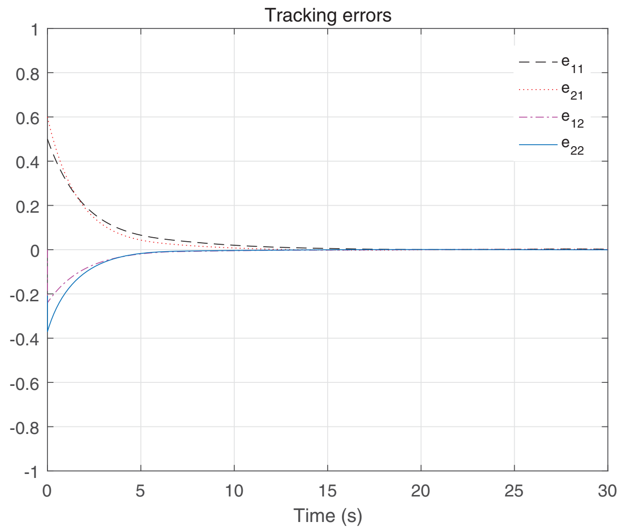

Figure 4. The tracking errors of Example 4.1.

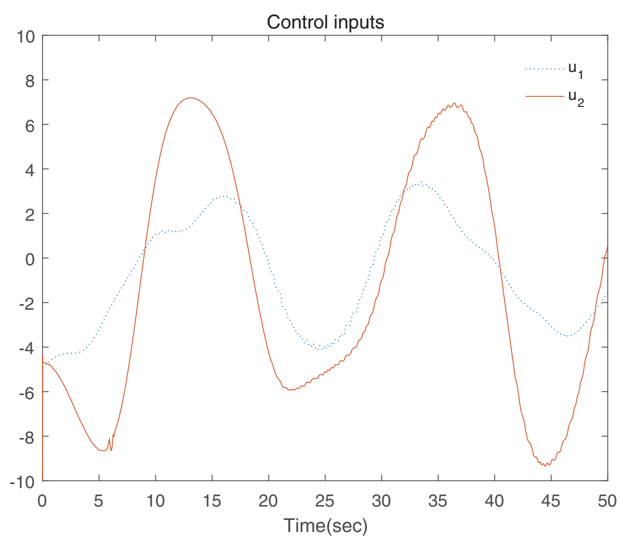

Figure 5. The control inputs of Example 4.1.

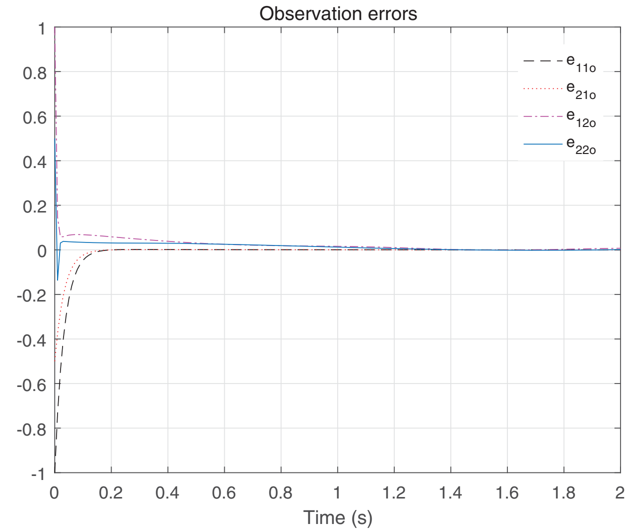

Figure 6. The observation errors by using the neural network observer of Example 4.2.

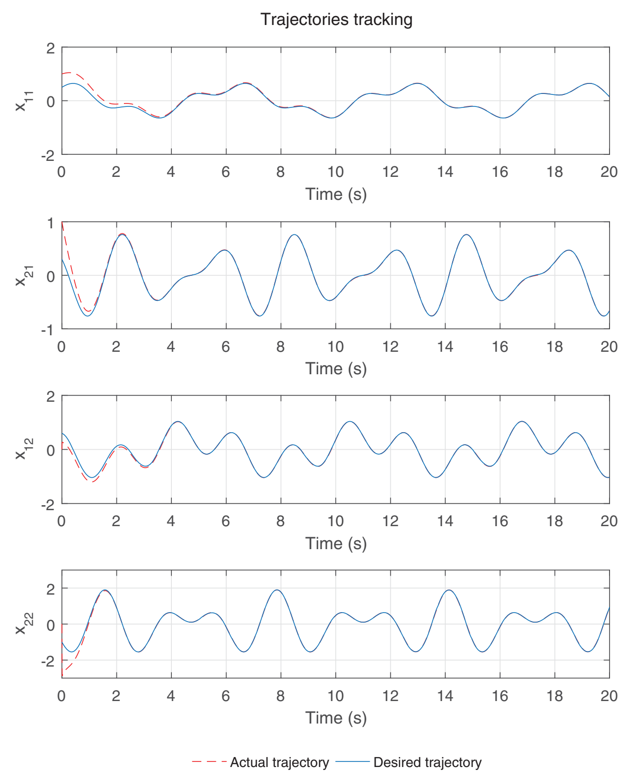

Figure 7. The trajectories tracking performance of Example 4.2.

by using the local NN observers, which ensure the observation errors to be UUB. It implies that the unknown subsystems are identified online successfully. The trajectories tracking curves are illustrated as Figure 3, we can see that the actual trajectories can follow their desired ones after the system runs for a short time by using the developed DTC (39). Figure 4 shows the tracking errors between the desired trajectories and actual trajectories, which give the same conclusion more intuitively. Figure 5 gives the curves of control inputs. From these figures, the closed-loop system can be guaranteed to be asymptotically stable. Therefore, the simulation results demonstrate the effectiveness of the proposed DTC scheme.

Example 4.2: In order to further test the effectiveness of the present observer-critic-based ADP for DTC method, a reconfigurable manipulator with 2-DOF (degree of freedom) is employed in our simulation (Zhao & Li, 2014).

Reconfigurable manipulators that consist of standard links and joint modules can be considered as a set of subsystems interconnected by coupling torques. In this simulation, the entire dynamics of reconfigurable manipulator can be expressed as

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) = u,$$

where $q \in \mathbb{R}^2$ is the vector of joint displacements, $M(q) \in \mathbb{R}^{2\times 2}$ is the inertia matrix, $C(q,\dot{q}) \in \mathbb{R}^2$ is the Coriolis and centripetal force, $G(q) \in \mathbb{R}^2$ is the gravity term and $u \in \mathbb{R}^2$ is the applied joint torque. The system matrices are

$$M(q) = \begin{bmatrix} 0.36\cos(q_2) + 0.6066 & 0.18\cos(q_2) + 0.1233 \\ 0.18\cos(q_2) + 0.1233 & 0.1233 \end{bmatrix},$$

$$C(q,\dot q)=\begin{bmatrix}-0.36\sin(q_2)\dot q_2 & -0.18\sin(q_2)\dot q_2\\ 0.18\sin(q_2)(\dot q_1-\dot q_2) & 0.18\sin(q_2)\dot q_1\end{bmatrix},$$
$$G(q)=\begin{bmatrix}-5.88\sin(q_1+q_2)-17.64\sin\left(q_1\right)\\ -5.88\sin(q_1+q_2)\end{bmatrix}.$$

For the development of DTC, each joint is considered as a subsystem of the entire manipulator system interconnected by coupling torque. By separating terms only depending on local variables $(q_i, \dot q_i, \ddot q_i)$ from those terms of other joint variables, each subsystem dynamic model can be formulated in joint space as

$$M_i(q_i)\ddot q_i + C_i(q_i,\dot q_i)\dot q_i + G_i(q_i) + Z_i(q,\dot q,\ddot q) = u_i \quad (44)$$

with

$$\begin{aligned}Z_i(q,\dot q,\ddot q) &= \left\{\sum_{j=1,j\neq i}^{n} M_{ij}(q)\ddot q_j + [M_{ii}(q)-M_i(q_i)]\ddot q_i\right\}\\ &\quad+\left\{\sum_{j=1,j\neq i}^{n} C_{ij}(q,\dot q)\dot q_j + [C_{ii}(q,\dot q)-C_i(q_i,\dot q_i)]\dot q_i\right\}\\ &\quad+\left[\bar G_i(q)-G_i(q_i)\right],\end{aligned}$$

where $q_i, \dot q_i, \ddot q_i, \overline{G}_i(q)$ and $\bar u_i$ are the ith element of the vectors $q, \dot q, \ddot q, G(q)$ and u, $M_{ij}(q)$ and $C_{ij}(q,\dot q)$ are the ijth element of the matrices $M(q)$ and $C(q,\dot q)$, respectively.

Let $x_i=[x_{i1},x_{i2}]^{\mathsf T}=[q_i,\dot q_i]^{\mathsf T}$, (44) can be expressed as

$$\begin{cases}\dot x_{i1}=x_{i2},\\ \dot x_{i2}=f_i(q_i,\dot q_i)+g_i(q_i)u_i+h_i(q,\dot q,\ddot q),\end{cases} \quad (45)$$

where x_i is the state of the ith subsystem, and

$$\begin{aligned}f_i(q_i,\dot q_i)&=M_i^{-1}(q_i)\left[-C_i(q_i,\dot q_i)\dot q_i-G_i(q_i)\right],\\ g_i(q_i)&=M_i^{-1}(q_i),\\ h_i(q,\dot q,\ddot q)&=-M_i^{-1}(q_i)Z_i(q,\dot q,\ddot q).\end{aligned}$$

The desired trajectories of two subsystems are

$$q_d=\begin{bmatrix}q_{1d}\\ q_{2d}\end{bmatrix}=\begin{bmatrix}0.5\cos(t)+0.2\sin(3t)\\ 0.3\cos(3t)-0.5\sin(2t)\end{bmatrix}.$$

The structure of RBFNNs, the initial states of subsystems, as well as the initial states of observers are the same as those of Example 4.1. Let the observer gain matrix be K_{io} = diag[200, 400], the RBFNN weights learning rate of the observer be Γ_{if} = 500 and $\Gamma_{ig}=1$. Let the critic NN be the same structure as Example 4.1 with initial values $\hat W_{1c}=[0.2, 1.5, 1.1]^{\mathsf T}$ and $\hat W_{2c}=[1.2, 0.8, 0.9]^{\mathsf T}$, the weight learning rates of the critic NN be η_{i1} = 0.0001, the gain of the compensator (38) be $\Gamma_{iw}=5$, $Q_i=10I_2$, $R_i=0.001I$.

We can see from Figure 6 that the observation errors of each subsystem are verified to be UUB by using the developed local NN observer. Figures 7 and 8 show that the trajectory tracking and tracking errors satisfy the control performance when using the present DTC (39). The tracking control inputs in Figure 9 demonstrate the closed-loop system of the reconfigurable manipulator asymptotically stable.

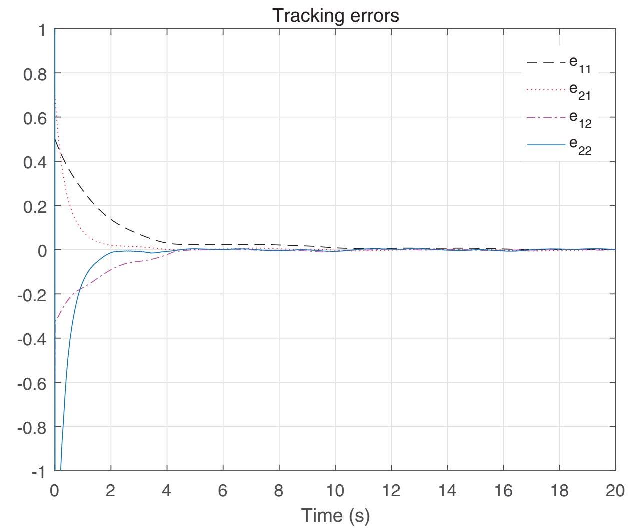

Figure 8. The tracking errors of Example 4.2.

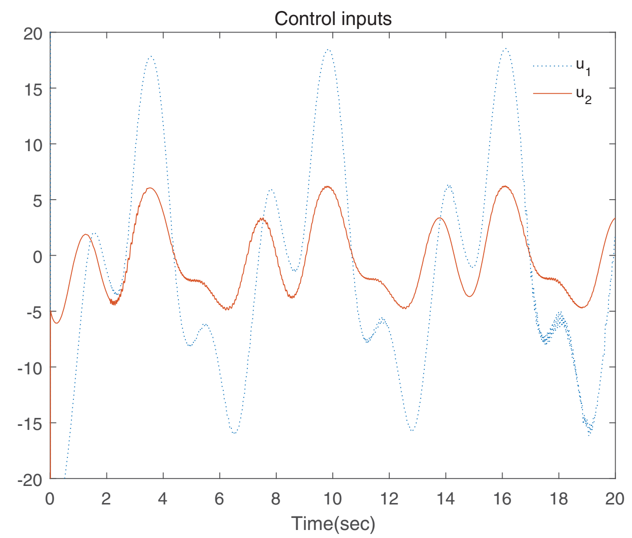

Figure 9. The control inputs of Example 4.2.

In summary, the simulation results of the two examples verify the effectiveness of the proposed scheme.

5. Conclusion

In this paper, we develop a DTC scheme for unknown large-scale nonlinear systems via observer-critic structure-based ADP. A local NN observer is established to identify the unknown subsystem. Hereafter, the local desired control can be derived directly. For error dynamic system, the local value function is approximated by constructing a critic NN, and the local tracking error control can be obtained. Then, the overall error caused by the substitution, observation and approximation of the local tracking error control can be compensated by an adaptive robustifying term. Therefore, the overall DTC can guarantee the closed-loop system to be asymptotically stable by Lyapunov's direct method. Two examples are employed to verify the effectiveness of the proposed DTC scheme.

Acknowledgment

The authors would like to thank the anonymous editors, reviewers for their valuable comments and constructive suggestions to improve the quality of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported in part by the National Natural Science Foundation of China [Grant Numbers U1501251, 61603387, 61533017, 61374051, 61374105 and 61503379]; in part by the Scientific and Technological Development Plan Project in Jilin Province of China [Grant Numbers 20150520112JH and 20160414033GH]; in part by Beijing Municipal Natural Science Foundation [Grant Number 4162065].

Notes on contributors

Bo Zhao received his B.S. degree in Automation, and Ph.D. degree in Control Science and Engineering, all from Jilin University, Changchun, China, in 2009 and 2014, respectively. His is currently a Post-Doctoral Fellow with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China, from 2014. His research interests include adaptive dynamic programming, robot control and fault diagnosis and tolerant control.

Derong Liu received his Ph.D. degree in electrical engineering from the University Of Notre Dame in 1994. He was a Staff Fellow with General Motors Research and Development Centre, from 1993 to 1995. He was an assistant professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, from 1995 to 1999. He joined the University of Illinois at Chicago in 1999, and became a Full Professor of Electrical and Computer Engineering and of Computer Science in 2006. He was selected for the "100 Talents Program" by the Chinese Academy of Sciences in 2008, and he served as the associate director of the State Key Laboratory of Management and Control for Complex Systems at the Institute of Automation, from 2010 to 2015. He is now a full professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His current research interests include adaptive dynamic programming, reinforcement learning and intelligent computing.

Xiong Yang received his B.S. degree in mathematics and applied mathematics from Central China Normal University, Wuhan, China, the M.S. degree in pure mathematics from Shandong University, Jinan, China, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2008, 2011, and 2014, respectively. He served as the assistant professor of the State Key Laboratory of Management and Control for Complex Systems at the Institute of Automation, Chinese Academy of Sciences from 2014 to 2015. He is currently an associate professor with the School of Electrical Engineering and Automation, Tianjin University. His current research interests include adaptive dynamic programming, reinforcement learning, adaptive control, and neural networks.

Yuanchun Li received his Ph.D. degree in General Mechanics from Harbin Institute of Technology, Harbin, China, in 1990. He is currently a Professor in Control Science and Engineering, Changchun University of Technology, Changchun, China. His research interests include complex system modeling and robot control, intelligent control.

ORCID

Bo Zhao http://orcid.org/0000-0002-7684-7342

References

Al-Tamimi, A., Lewis, F. L., & Abu-Khalaf, M. (2008). Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38*(4), 943–949.

Bertsekas, D. P., & Tsitsiklis, J. N. (1995). Neuro-dynamic programming: An overview. In *Proceedings of the 34th IEEE Conference on Decision and Control. 1*,(pp. 560–564).

Bian, T., Jiang, Y., & Jiang, Z. P. (2015). Decentralized adaptive optimal control of large-scale systems with application to power systems. *IEEE Transactions on Industrial Electronics, 62*(4), 2439–2447.

Choi, Y. H., & Yoo, S. J. (2016). Decentralized approximation-free control for uncertain large-scale pure-feedback systems with unknown time-delayed nonlinearities and control directions. *Nonlinear Dynamics.* 85,1053–1066. doi: 10.1007/s11071-016-2743-z

Enns, R., & Si, J. (2003). Helicopter trimming and tracking control using direct neural dynamic programming. *IEEE Transactions on Neural Networks, 14*(4), 929–939.

Fu, Y., Fu, J., & Chai, T. (2015). Robust adaptive dynamic programming of two-player zero-sum games for continuous-time linear systems. *IEEE Transactions on Neural Networks and Learning Systems, 26*(12), 3314–3319.

Gao, W., Jiang, Y., Jiang, Z. P., & Chai, T. (2016). Output-feedback adaptive optimal control of interconnected systems based on robust adaptive dynamic programming. *Automatica, 72*, 37–45.

He, P., & Jagannathan, S. (2005). Reinforcement learning-based output feedback control of nonlinear systems with input constraints. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 35*(1), 150–154.

Henikl, J., Kemmetmller, W., Meurer, T., & Kugi, A. (2016). Infinite-dimensional decentralized damping control of large-scale manipulators with hydraulic actuation. *Automatica, 63*, 101–115.

Hioe, D., Hudon, N., & Bao, J. (2014). Decentralized nonlinear control of process networks based on dissipativity-A Hamilton-Jacobi equation approach. *Journal of Process Control, 24*(3), 172–187.

Hua, C., Li, Y., Wang, H., & Guan, X. (2015). Decentralised fault-tolerant finite-time control for a class of interconnected non-linear systems. *IET Control Theory & Applications, 9*(16), 2331–2339.

Hua, C., Zhang, L., & Guan, X. (2015). Decentralized output feedback adaptive NN tracking control for time-delay stochastic nonlinear systems with prescribed performance. *IEEE Transactions on Neural Networks and Learning Systems, 26*(11), 2749–2759.

Jiang, Y., & Jiang, Z. P. (2012). Robust adaptive dynamic programming for large-scale systems with an application to multimachine power systems. *IEEE Transactions on Circuits and Systems II: Express Briefs, 59*(10), 693–697.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research, 4*, 237–285.

Koo, G. B., Park, J. B., & Joo, Y. H. (2016). Decentralized sampled-data fuzzy observer design for nonlinear interconnected systems. *IEEE Transactions on Fuzzy Systems, 24*(3), 661–674.

Labibi, B., Lohmann, B., Sedigh, A. K., & Maralani, P. J. (2002). Output feedback decentralized control of large-scale systems using weighted sensitivity functions minimization. *Systems & Control Letters, 47*(3), 191–198.

Langarica Crdoba, D., & Ortega, R. (2015). An observer–based scheme for decentralized stabilization of large-scale systems with application to power systems. *Asian Journal of Control, 17*(1), 124–132.

Liu, D., Li, C., Li, H., Wang, D., & Ma, H. (2015). Neural-network-based decentralized control of continuous-time nonlinear interconnected systems with unknown dynamics. *Neurocomputing, 165*, 90–98.

Liu, D., Wang, D., & Li, H. (2014). Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach. *IEEE Transactions on Neural Networks and Learning Systems, 25*(2), 418–428.

Lin, Q., Wei, Q., & Liu, D. (2016). A novel optimal tracking control scheme for a class of discrete-time nonlinear systems using generalised policy iteration adaptive dynamic programming algorithm. *International Journal of Systems Science.* 48(3), 525–534. doi: 10.1080/00207721.2016.1188177

Lu, C., Si, J., & Xie, X. (2008). Direct heuristic dynamic programming for damping oscillations in a large power system. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38*(4), 1008–1013.

Mehraeen, S., & Jagannathan, S. (2011). Decentralized optimal control of a class of interconnected nonlinear discrete-time systems by using online Hamilton-Jacobi-Bellman formulation. *IEEE Transactions on Neural Networks, 22*(11), 1757–1769.

Mu, C., Sun, C., Song, A., & Yu H,. (2016). Iterative GDHP-based approximate optimal tracking control for a class of discrete-time nonlinear systems. *Neurocomputing, 214*, 775–784.

Park, Y. M., Choi, M. S., & Lee, K. Y. (1996). An optimal tracking neuro-controller for nonlinear dynamic systems. *IEEE Transactions on Neural Networks, 7*(5), 1099–1110.

Prokhorov, D. V., & Wunsch, D. C. (1997). Adaptive critic designs. *IEEE Transactions on Neural Networks, 8*(5), 997–1007.

Qin, C., Zhang, H., Luo, Y., & Wang, B. (2014). Finite horizon optimal control of non-linear discrete-time switched systems using adaptive dynamic programming with ϵ-error bound. *International Journal of Systems Science, 45*(8), 1683–1693.

Sadabadi, M. S., Karimi, A., & Karimi, H. (2015). Fixed-order decentralized/distributed control of islanded inverter-interfaced microgrids. *Control Engineering Practice, 45*, 174–193.

Tlili, A. S., & Braiek, N. B. (2014). H_∞ optimization-based decentralized control of linear interconnected systems with nonlinear interconnections. *Journal of the Franklin Institute, 351*(6), 3286–3304.

Tong, S., Zhang, L., & Li, Y. (2016). Observed-based adaptive fuzzy decentralized tracking control for switched uncertain nonlinear large-scale systems with dead zones. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 46*(1), 37–47.

Wang, D., Liu, D., Mu, C., & Ma, H. (2016). Decentralized guaranteed cost control of interconnected systems with uncertainties: A learning-based optimal control strategy. *Neurocomputing*. 214, 297–306. doi: 10.1016/j.neucom.2016.06.020

Wang, D., Liu, D., Zhang, Q., & Zhao, D. (2016). Data-based adaptive critic designs for nonlinear robust optimal control with uncertain dynamics. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 46*(11), 1544–1555.

Wang, D., Mu, C., He, H., & Liu, D. (2016). Event-driven adaptive robust control of nonlinear systems with uncertainties through NDP strategy. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. Advance online publication. doi: 10.1109/TSMC.2016.2592682

Wang, D., Mu, C., Zhang, Q., & Liu, D. (2016). Event-based input-constrained nonlinear H_∞ state feedback with adaptive critic and neural implementation. *Neurocomputing, 214*, 848–856.

Wang, F. Y., Zhang, H., & Liu, D. (2009). Adaptive dynamic programming: An introduction. *IEEE Computational Intelligence Magazine, 4*(2), 39–47.

Wei, Q., & Liu, D. (2014). Adaptive dynamic programming for optimal tracking control of unknown nonlinear systems with application to coal gasification. *IEEE Transactions on Automation Science and Engineering, 11*(4), 1020–1036.

Werbos, P. J. (1992). Approximate dynamic programming for real-time control and neural modeling. *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches, 15*, 493–525.

Yang, A., Naeem, W., Irwin, G. W., & Li, K. (2014). Stability analysis and implementation of a decentralized formation control strategy for unmanned vehicles. *IEEE Transactions on Control Systems Technology, 22*(2), 706–720.

Yang, X., Liu, D., & Wang, D. (2014). Reinforcement learning for adaptive optimal control of unknown continuous-time nonlinear systems with input constraints. *International Journal of Control, 87*(3), 553–566.

Yang, X., Liu, D., Wei, Q., & Wang, D. (2016). Guaranteed cost neural tracking control for a class of uncertain nonlinear systems using adaptive dynamic programming. *Neurocomputing, 198*, 80–90.

Yi, Y., & Zheng, W. X. (2016). Decentralized PI tracking control for non-Gaussian large-scale interconnected distribution systems. *Journal of the Franklin Institute, 353*(1), 235–255.

Zhang, H., Cui, L., & Luo, Y. (2013). Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network ADP. *IEEE Transactions on Cybernetics, 43*(1), 206–216.

Zhang, H., Cui, L., Zhang, X., & Luo, Y. (2011). Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method. *IEEE Transactions on Neural Networks, 22*(12), 2226–2236.

Zhang, H., Luo, Y., & Liu D. (2009). Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraints. *IEEE Transactions on Neural Networks, 20*(9): 1490–1503.

Zhang, H., Song, R., Wei, Q., & Zhang, T. (2011). Optimal tracking control for a class of nonlinear discrete-time systems with time delays based on heuristic dynamic programming. *IEEE Transactions on Neural Networks, 22*(12), 1851–1862.

Zhang, H., Wei, Q., & Luo, Y. (2008). A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear systems via the greedy HDP iteration algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38*(4), 937–942.

Zhang, Z., Zhang, Z., & Zhang, H. (2015). Decentralized robust attitude tracking control for spacecraft networks under unknown inertia matrices. *Neurocomputing, 165*, 202–210.

Zhao, B., & Li, Y. (2014). Local joint information based active fault tolerant control for reconfigurable manipulator. *Nonlinear Dynamics, 77*(3), 859–876.

Zhao, B., Liu, D., & Li, Y. (2016). Online fault compensation control based on policy iteration algorithm for a class of affine non-linear systems with actuator failures. *IET Control Theory & Applications, 10*(15), 1816–1823.

Zhao, B., Liu, D., & Li, Y. (2017). Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems. *Information Sciences, 384*, 21–33.