Decentralized Control for Large-Scale Nonlinear Systems With Unknown Mismatched Interconnections via Policy Iteration

Bo Zhao, *Member, IEEE*, Ding Wang, *Member, IEEE*, Guang Shi, Derong Liu, *Fellow, IEEE*, and Yuanchun Li

***Abstract*—In this paper, the decentralized control problem is solved based on a policy iteration algorithm for large-scale nonlinear systems with unknown mismatched interconnections. The unknown interconnection is approximated by a neural network with local states of isolated subsystem and substituted reference states of coupled subsystems. Then, an adaptive estimation term is utilized to construct the improved local performance index function that reflects the substitution error. Hereafter, the closed-loop large-scale nonlinear system is guaranteed to be ultimately uniformly bounded by the implementation of a set of developed decentralized optimal control policies. Two simulation examples are given to verify the effectiveness of the presented scheme. The significant contribution of this scheme lies in that it removes the common assumptions on satisfying matching condition and upper boundedness of interconnections, when designing the decentralized optimal control for large-scale nonlinear systems.**

***Index Terms*—Adaptive dynamic programming (ADP), decentralized control, large-scale systems, neural networks (NNs), optimal control, policy iteration (PI), reinforcement learning, unknown mismatched interconnections.**

I. Introduction

THE increasing demands of production quality and economic efficiency have led to increasingly large-scale and complex modern systems, such as ecosystems, communication systems, transportation systems, urban traffic systems, and power systems. In general, a large-scale system consists of a set of subsystems coupled by interconnections, which lead to increasing difficulties of analysis and synthesis when utilizing centralized control. To overcome the difficulties in controlling such systems, decentralized control strategy, which utilizes local states of each subsystem, is an efficient and effective approach. In the past few decades, considerable attention has been paid to the design of decentralized controllers for large-scale systems. For example, Wang *et al*. [1] presented an adaptive neural decentralized control approach for stochastic systems with strong interconnected nonlinearities both in drift and diffusion terms. Li *et al*. [2] proposed a decentralized adaptive neural control scheme for a class of interconnected large-scale uncertain systems with input saturation. Zhao *et al.* [3], [4] presented decentralized fault-tolerant control schemes based on self-tuned local feedback gain and local nonlinear velocity observer against actuator failures.

As is well known, the optimal control problem for nonlinear systems can be addressed by the Hamilton–Jacobi–Bellman (HJB) equation, which can be solved by adaptive dynamic programming (ADP) [5] to avoid the difficulty in the "curse of dimensionality" with the help of function approximators, such as neural networks (NNs). Since NNs have a strong approximation capability, Wang *et al.* [6], [7] proposed adaptive neural control schemes for nonlinear systems with dynamic uncertainties and completely unknown dynamics. Li *et al.* [8] tackled the control problems by using NNs for nonlinear systems with unknown dead-zone, time-varying delays [9], [10], and unmodeled dynamics[11]. There are many synonyms used for ADP, such as ADP [12], approximate dynamic programming [13], adaptive critic designs [14], neuro-dynamic programming [15], [16], and reinforcement learning [17]. Recently, ADP algorithms were further employed to solve control problems of continuous-time systems [18]–[20], discrete-time systems [21]–[23], trajectory tracking [24]–[26], input/output constraints [22], [27], external disturbances and uncertainties [28]–[30], zero-sum games [31], fault tolerant [32]–[34], etc. We can see from the literature that ADP algorithms can be categorized into heuristic dynamic programming (HDP) [24], dual HDP (DHP) [35], action-dependent HDP (ADHDP) [36], ADDHP [37], globalized DHP (GDHP) [38], and ADGDHP [14].

As previously mentioned, local controllers should be designed for their corresponding subsystems in decentralized control strategy. Saberi [39] established a decentralized

Manuscript received November 22, 2016; revised February 15, 2017; accepted March 30, 2017. Date of publication April 24, 2017; date of current version September 14, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant U1501251, Grant 61603387, Grant 61533017, Grant 61374051, Grant 61374105, and Grant 61503379, in part by the Scientific and Technological Development Plan Project in Jilin Province of China under Grant 20150520112JH and Grant 20160414033GH, and in part by the Beijing Natural Science Foundation under Grant 4162065. This paper was recommended by Associate Editor Z. Wang. *(Corresponding author: Ding Wang.)*

B. Zhao, D. Wang, and G. Shi are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: zhaobo@ia.ac.cn; ding.wang@ia.ac.cn; shiguang2012@ia.ac.cn).

D. Liu is with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: derong@ustb.edu.cn).

Y. Li is with the Department of Control Science and Engineering, Changchun University of Technology, Changchun 130012, China (e-mail: liyc@ccut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2017.2690665

optimal control for local subsystems of interconnected system. It is shown that the optimal control can be implemented to design a decentralized controller. Several literatures have paid attention to the optimal design in decentralized control. Jiang and Jiang [19] presented a decentralized control design by using robust ADP theory and policy iteration (PI) technique for complex systems with unknown parameters and dynamic uncertainties. And then, Bian *et al.* [40] extended the method to systems with unmatched uncertainties. Lu *et al.* [41] presented a direct HDP method to address nonlinear coordinated control for a large power system with uncertainties. For damping low frequency oscillations in power systems, Molina *et al.* [42] developed an intelligent controller based on local signals by using virtual generators and ADP technique. Bernstein *et al.* [43] presented an optimal PI algorithm to handle the decentralized partially observable Markov decision process. Liu *et al.* [44] developed an NN-based online learning optimal control approach to stabilize nonlinear interconnected large-scale systems, and then by introducing an integral PI algorithm, a model-free optimal control method was extended to unknown interconnected systems [45]. Karimi *et al.* [47] proposed a reinforcement learning-based backsteping decentralized control scheme for electric power systems, where gains of decentralized controllers were tuned by reinforcement learning to adapt to various operating conditions. Mehraeen and Jagannathan [48] solved HJB equation via direct neural dynamic programming for the decentralized near optimal regulation of nonlinear interconnected discrete-time systems. However, most previously mentioned literature focused on the plants in linear or satisfying assumed matching conditions. Actually, interconnections are always unknown and mismatched in many applications. However, ADP-based decentralized control approaches were not presented in previous works for systems in these situations.

Motivated by [44]–[46], this paper addresses the decentralized control problem for large-scale nonlinear systems with unknown mismatched interconnections. By using local states of isolated subsystem and substituted reference states of coupled subsystems, the unknown interconnection is approximated by an NN. Then, the improved local performance index function which reflects the substitution error is constructed with the help of the estimated term. Hereafter, the PI algorithm is developed to solve the HJB equation via the constructed critic NN, and the approximated decentralized control policy can be directly obtained. It is proven that the closed-loop large-scale nonlinear system can be guaranteed to be ultimately uniformly bounded (UUB) based on Lyapunov stability theorem. Two numerical simulation examples are provided to ensure the effectiveness of the proposed scheme.

The main contributions of this paper include the following two aspects.

1) Unlike the literature previously mentioned, this paper extends the ADP algorithm to deal with the decentralized control problem with an improved local performance index function for large-scale nonlinear systems with unknown mismatched interconnections.
2) The unknown mismatched interconnection is estimated by local observer, which utilizes the local states and the substituted reference states of the coupled subsystems. As a result, the proposed scheme avoids the common assumptions on satisfying matching condition and upper boundedness of interconnections of large-scale nonlinear systems in previous ADP-based approaches.

The rest of this paper is organized as follows. In Section II, the problem statement is presented, and approximate the interconnection by employing local states of isolated subsystem and substituted reference states of coupled subsystems. In Section III, the decentralized optimal control policy is derived for the isolated subsystem, and NNs are employed to approximate the interconnection and the critic NN, respectively. Then, the local online PI algorithm is presented. In Section IV, two numerical simulation examples are provided to demonstrate the effectiveness of the developed scheme. In Section V, the conclusion is drawn.

II. Problem Statement

In this paper, we consider a large-scale nonlinear system composed of N subsystems with unknown mismatched interconnections as

$$\begin{cases} \dot{x}_1(t) = f_1(x_1(t)) + g_1(x_1(t))u_1(x_1(t)) + h_1(x(t)) \\ \quad\vdots \\ \dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(x_i(t)) + h_i(x(t)) \\ \quad\vdots \\ \dot{x}_N(t) = f_N(x_N(t)) + g_N(x_N(t))u_N(x_i(t)) + h_N(x(t)). \end{cases} \tag{1}$$

The ith ($i = 1, 2, \ldots, N$) interconnected subsystem is described by

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(x_i(t)) + h_i(x(t)) \tag{2}$$

where $x_i(t) \in \mathbb{R}^{n_i}$ and $u_i(x_i(t)) \in \mathbb{R}^{m_i}$ are the state vector and input vector of the ith subsystem, respectively. $x = [x_1, x_2, \ldots, x_N]^\mathsf{T} \in \mathbb{R}^n$ with $n = \sum_{i=1}^{N} n_i$ denotes the entire system state and $u_1(x_1), u_2(x_2), \ldots, u_N(x_N)$ are local control inputs. For the ith subsystem, $f_i(\cdot)$ and $g_i(\cdot)$ are known, locally Lipschitz and differentiable in their augments with $f_i(0) = 0$. $h_i(x(t))$ is the unknown mismatched interconnection term.

As is well known, fuzzy logic systems, NNs, etc. are excellent approximators for unknown nonlinearities. Furthermore, since the radial basis function NN (RBFNN) has simple structure and excellent approximation capability, RBFNN is employed in a given compact set $\Omega \in \mathbb{R}^n$ to approximate the unknown interconnection $h_i(x(t))$, that is

$$h_i(x(t)) = W_{ih}^\mathsf{T}\sigma_{ih}(x(t)) + \varepsilon_i(x(t)) \tag{3}$$

where $\sigma_{ih}(x(t))$ is called the basis function that is commonly selected as a Gaussian function

$$\sigma_{ih}(x) = \exp\left(\frac{-(x - c_i)^\mathsf{T}(x - c_i)}{b_i^2}\right)$$

where the constant vector c_i is the center of the basis function, and $b_i > 0$ is a real number which is the width of the basis function. The optimal weight vector $W_{ih} = [w_{i1}, w_{i2}, \ldots, w_{ik}]^{\mathsf{T}}$ is defined as

$$W_{ih} = \arg\min_{\hat{W}_{ih}\in\mathbb{R}^k}\left\{\sup_{x\in\Omega}\left|h_i(x) - \hat{W}_{ih}^{\mathsf{T}}\sigma_{ih}(x(t))\right|\right\}$$

and $\varepsilon_i(x)$ is the NN approximation error, which can be decreased by increasing the NN hidden node number k.

Assumption 1: The NN approximation error ε_i is upper bounded, i.e., $|\varepsilon_i(x)| \le \phi_{i1}$, where ϕ_{i1} is an unknown positive constant.

To relax the upper boundedness assumption of interconnections, we approximate the interconnection term in the ith subsystem by RBFNN using the states of local subsystem and the reference states of the coupled subsystems, that is

$$\begin{aligned} h_i(x) &= W_{ih}^{\mathsf{T}}\sigma_{ih}(x_{iD}) + \Delta_i(x, x_{iD}) + \varepsilon_i(x_i)\\ &= h_{id}(x_{iD}) + \Delta_i(x, x_{iD}) + \varepsilon_i(x_i) \end{aligned} \tag{4}$$

where $x_{iD} = [x_{1d}, x_{2d}, \ldots, x_i, \ldots, x_{Nd}]^{\mathsf{T}}$, x_{id} indicates the reference states of the coupled subsystems, $h_{id}(x_{iD}) = W_{ih}^{\mathsf{T}}\sigma_{ih}(x_{iD})$, $\Delta_i(x, x_{iD}) = W_{ih}^{\mathsf{T}}\sigma_{ih}(x) - W_{ih}^{\mathsf{T}}\sigma_{ih}(x_{iD})$ is the substitution error since it arises from the substitution of NN inputs.

Similar to [49], the Gaussian function $\sigma_{ih}(x_i)$ satisfies the global Lipschitz condition, which implies

$$\|\Delta_i\| \le \sum_{j=1, j\neq i}^{N} d_{ij}E_j$$

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

Remark 1: We can observe that (4) can be obtained only by adding and subtracting the term $h_{id}(x_{iD})$, which can be approximated by RBFNN. So it can avoid the common upper boundedness assumption of the interconnection term in the ith subsystem. In other words, the function $h_i(x)$ depends only on the corresponding local states and the reference states, which are shared with each subsystem before the system runs.

For the ith isolated subsystem

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(x_i(t)) + h_{id}(x_{iD}) \tag{5}$$

since $f_i(\cdot)$ and $g_i(\cdot)$ are locally Lipschitz continuous on a set $\Omega_i \in \mathbb{R}^{n_i}$, the subsystem (5) is controllable. Different from the interconnected subsystem (2), the isolated subsystem (5) depends only on its local states.

Remark 2: To eliminate confusions, it is necessary to distinguish the concepts of interconnected subsystems, isolated subsystems, and coupled subsystems. In this paper, we call all the subsystems interconnected with the ith one as coupled subsystems. On the other hand, we call (2) interconnected subsystem, since $h_i(x(t))$ contains the actual states of all the subsystems. Different from it, $h_{id}(x_{iD})$ in the isolated subsystem (5) depends only on the local states and the substituted reference states of the coupled subsystems. That is to say, $h_{id}(x_{iD})$ is independent from interconnection $h_i(x(t))$. So it can be called isolated subsystem.

The main objective of this paper is to find a set of local control policies $u_1(x_1), u_2(x_2), \ldots, u_N(x_N)$ as the decentralized control law to stabilize the system (1). To handle the optimal control problem, we need to obtain the optimal control policy $u_i^*(x_i)$ for the ith subsystem. Thus, it is desired to find the feedback control policy $u_i(x_i)$ to minimize the improved local infinite horizon performance index function as

$$J_i(x_{i0}) = \int_0^\infty \left(\hat{\delta}_i\left\|\nabla J_i^{\mathsf{T}}(x_i(\tau))\right\|E_i + U_i(x_i(\tau), u_i(\tau))\right)\mathrm{d}\tau \tag{6}$$

where $U_i(x_i, u_i) = x_i^{\mathsf{T}}Q_ix_i + u_i^{\mathsf{T}}R_iu_i$ is the utility function, $U_i(0, 0) = 0$, and $U_i(x_i, u_i) \ge 0$ for all x_i and u_i, in which $Q_i \in \mathbb{R}^{n_i\times n_i}$ and $R_i \in \mathbb{R}^{m_i\times m_i}$ are positive definite matrices. $\hat{\delta}_i$ is a positive function which will be defined later. $\nabla J_i(x_i)$ denotes the partial derivative of local performance index function $J_i(x_i)$ with respect to local state x_i, i.e., $\nabla J_i(x_i) = (\partial J_i(x_i)/\partial x_i)$.

III. Decentralized Controller Design and Stability Analysis

In this section, we present the optimal decentralized controller design and stability analysis in detail.

A. Optimal Control

Based on the optimal control theory, the designed feedback control policy must be admissible. Therefore, before the optimal control is presented, the definition of admissible control is introduced.

Definition 1: For the ith isolated subsystem (5), a control policy $u_i(x_i)$ is defined to be admissible with respect to (6) if $u_i(x_i)$ is continuous on a set $\Omega_i \in \mathbb{R}^{n_i}$, $u_i(0) = 0$ and $u_i(x_i)$ stabilizes the isolated subsystem (5), and $J_i(x_{i0})$ in (6), where x_{i0} is the initial state of x_i, is finite for all $x_i \in \Omega_i$.

Consider the ith isolated subsystem (5), for any admissible control policy $u_i(x_i) \in \psi_i(\Omega_i)$, where $\psi_i(\Omega_i)$ denotes the set of admissible control, if the improved local value function

$$V_i(x_i) = \int_0^\infty \left(\hat{\delta}_i\left\|\nabla V_i^{\mathsf{T}}(x_i)\right\|E_i + U_i(x_i(\tau), u_i(\tau))\right)\mathrm{d}\tau \tag{7}$$

is continuously differentiable, then the infinitesimal version of (7) is the so-called local nonlinear Lyapunov equation

$$\begin{aligned} 0 = {}&\hat{\delta}_i\left\|\nabla V_i^{\mathsf{T}}(x_i)\right\|E_i + U_i(x_i, u_i)\\ &+ \nabla V_i^{\mathsf{T}}(x_i)(f_i(x_i) + g_i(x_i)u_i(x_i) + h_{id}(x_{iD})) \end{aligned} \tag{8}$$

with $V_i(0) = 0$.

Define the local Hamiltonian as

$$\begin{aligned} H_i(x_i, u_i, \nabla V_i(x_i)) = {}&\hat{\delta}_i\left\|\nabla V_i^{\mathsf{T}}(x_i)\right\|E_i + U_i(x_i, u_i)\\ &+ \nabla V_i^{\mathsf{T}}(x_i)(f_i(x_i) + g_i(x_i)u_i(x_i)\\ &+ h_{id}(x_{iD})) \end{aligned}$$

and the local value function as

$$V_i^*(x_i) = \min_{u_i\in\Psi_i(\Omega_i)}\int_0^\infty \left(\hat{\delta}_i\left\|\nabla V_i^{*\mathsf{T}}(x_i)\right\|E_i + U_i(x_i(\tau), u_i(\tau))\right)\mathrm{d}\tau. \tag{9}$$

According to optimal control theory, $V_i^*(x_i)$ satisfies HJB equation

$$0 = \min_{u_i \in \Psi_i(\Omega)} H_i\big(x_i, u_i, \nabla V_i^*(x_i)\big) \tag{10}$$

where $\nabla V_i^*(x_i) = (\partial V_i^*(x_i)/\partial x_i)$. Assume the solution $V_i^*(x_i)$ exists and is continuously differentiable, the local optimal control policy can be described as

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

For considered large-scale nonlinear system (1), the local feedback control policies $u_1(x_1), u_2(x_2), \ldots, u_N(x_N)$ should be presented to guarantee the entire closed-loop system stable. To achieve this goal, we will transform the stabilization problem into designing a set of local optimal controllers with proper local value functions.

Theorem 1: For the ith interconnected subsystem (2) with the improved local value function (7), $V_i^*(x_i)$ is the optimal solution of the HJB (10), and $u_i^*(x_i)$ is the optimal control policy by (11). It implies that the control policies $u_1^*(x_1), u_2^*(x_2), \ldots, u_N^*(x_N)$ are the decentralized control law of large-scale nonlinear system (1).

Proof: The theorem can be proved by showing that $V_i^*(x_i)$ is a Lyapunov function. From the definition of each term in (9), we can observe that $V_i^*(x_i) > 0$ for any $x_i \neq 0$ and $V_i^*(x_i) = 0$ for $x_i = 0$, which implies that $V_i^*(x_i)$ is a positive definite function. Therefore, the time derivative of $V_i^*(x_i)$ along the corresponding state of the closed-loop interconnected subsystem is described by

$$\begin{aligned}\dot{V}_i^*(x_i) &= \big(\nabla V_i^*(x_i)\big)^{\mathsf{T}} \dot{x}_i \\ &= \big(\nabla V_i^*(x_i)\big)^{\mathsf{T}} (f_i(x_i) + g_i(x_i) u_i(x_i) \\ &\qquad + h_{iD}(x_{iD}) + \Delta_i + \varepsilon_i).\end{aligned} \tag{12}$$

Denoting $(\nabla V_i^*(x_i))^{\mathsf{T}}$ as $\nabla V_i^{*\mathsf{T}}$ for simplicity, and substituting (8) into (12), we have

$$\begin{aligned}\dot{V}_i^*(x_i) &= -\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) + \nabla V_i^{*\mathsf{T}} (\Delta_i + \varepsilon_i) \\ &\le -\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) + \left\| \nabla V_i^{*\mathsf{T}} \right\| \|\Delta_i\| \\ &\quad + \left\| \nabla V_i^{*\mathsf{T}} \right\| \|\varepsilon_i\|.\end{aligned}$$

Considering Assumption 1, we have

$$\begin{aligned}\dot{V}_i^*(x_i) &\le -\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) \\ &\quad + \left\| \nabla V_i^{*\mathsf{T}} \right\| \sum_{j=1, j\neq i}^{N} d_{ij} E_j + \left\| \nabla V_i^{*\mathsf{T}} \right\| \phi_{i1} \\ &\le -\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) \\ &\quad + \max_{ij}\{d_{ij}\} \left\| \nabla V_i^{*\mathsf{T}} \right\| \sum_{j=1, j\neq i}^{N} E_j + \left\| \nabla V_i^{*\mathsf{T}} \right\| \phi_{i1}.\end{aligned}$$

Therefore

$$\begin{aligned}\dot{V}^* &= \sum_{i=1}^{N} V_i^* \\ &\le \sum_{i=1}^{N} \left(-\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) + \left\| \nabla V_i^{*\mathsf{T}} \right\| \phi_{i1} \right) \\ &\quad + \sum_{i=1}^{N} \max_{ij}\{d_{ij}\} \left\| \nabla V_i^{*\mathsf{T}} \right\| \sum_{j=1, j\neq i}^{N} E_j \\ &\le \sum_{i=1}^{N} \left(-\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) + \left\| \nabla V_i^{*\mathsf{T}} \right\| \phi_{i1} \right) \\ &\quad + \max_{ij}\{d_{ij}\} \sum_{i=1}^{N} \left\| \nabla V_i^{*\mathsf{T}} \right\| \sum_{j=1, j\neq i}^{N} E_j.\end{aligned} \tag{13}$$

Since $E_i = \|x_i - x_{id}\| \ge 0$, (13) becomes

$$\begin{aligned}\dot{V}^* &\le \sum_{i=1}^{N} \left(-\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) + \left\| \nabla V_i^{*\mathsf{T}} \right\| \phi_{i1} \right) \\ &\quad + \max_{ij}\{d_{ij}\} \sum_{i=1}^{N} \left\| \nabla V_i^{*\mathsf{T}} \right\| \sum_{j=1}^{N} E_j \\ &= \sum_{i=1}^{N} \left(-\hat{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) + \left\| \nabla V_i^{*\mathsf{T}} \right\| \phi_{i1} \right) \\ &\quad + N \cdot \max_{ij}\{d_{ij}\} \sum_{i=1}^{N} \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i \\ &\quad - \max_{ij}\{d_{ij}\} \sum_{i=1}^{N} \left\| \nabla V_i^{*\mathsf{T}} \right\| \sum_{j=1}^{N} \big(E_i - E_j\big).\end{aligned}$$

Let $\delta_i = N \cdot \max_{ij}\{d_{ij}\}$. We have

$$\begin{aligned}\dot{V}^* &= \sum_{i=1}^{N} \left(\tilde{\delta}_i \left\| \nabla V_i^{*\mathsf{T}} \right\| E_i - U_i(x_i, u_i) + \left\| \nabla V_i^{*\mathsf{T}} \right\| \phi_{i1} \right) \\ &\quad - \max_{ij}\{d_{ij}\} \sum_{i=1}^{N} \left\| \nabla V_i^{*\mathsf{T}} \right\| \sum_{j=1}^{N} \big(E_i - E_j\big)\end{aligned} \tag{14}$$

where $\tilde{\delta}_i = \delta_i - \hat{\delta}_i$. Denoting $\eta_{i1} = \tilde{\delta}_i \|\nabla V_i^{*\mathsf{T}}\| E_i + \|\nabla V_i^{*\mathsf{T}}\| \phi_{i1} - \max_{ij}\{d_{ij}\} \sum_{i=1}^{N} \|\nabla V_i^{*\mathsf{T}}\| \sum_{j=1}^{N} (E_i - E_j)$, which is assumed to be bounded, i.e., $|\eta_{i1}| \le \Phi_i$, we have

$$\begin{aligned}\dot{V}^* &\le \sum_{i=1}^{N} \left(\Phi_i - x_i^{\mathsf{T}} Q_i x_i - u_i^{\mathsf{T}} R_i u_i \right) \\ &\le \sum_{i=1}^{N} \left(\Phi_i - x_i^{\mathsf{T}} Q_i x_i \right) \\ &\le \sum_{i=1}^{N} \left(\Phi_i - \lambda_{\min}(Q_i) \|x_i\|^2 \right)\end{aligned}$$

where $\lambda_i(\cdot)$ denotes the minimum eigenvalue of the matrix. Hence, we can conclude that $\dot{V}^* < 0$ when x_i lies outside of the compact set

$$\Omega_{x_i} = \left\{ x_i \ : \ \|x_i\| \leq \sqrt{\frac{\Phi_i}{\lambda_{\min}(Q_i)}} \right\}.$$

It implies that $V^*(x)$ is a Lyapunov function. This indicates that $x_i(t)$ will converge to a small neighborhood wherever the initial position is. This completes the proof. ■

Remark 3: In [44], the interconnection term is required to satisfy the assumed matching condition, which plays an important role in guaranteeing the closed-loop isolated subsystem to be stable. Unlike the method in [44], in this paper, we can see from the detailed proof that the strong assumption is relaxed by moving the substituted interconnection into the isolated subsystem, and leaving the bounded term $\|\nabla V_i^*(x_i)\|\phi_{i1}$ to be handled. Furthermore, the UUB stability is guaranteed for the large-scale nonlinear system (1), rather than the isolated subsystem (5). Therefore, the assumption on the matched interconnection can be removed.

B. Neural Network Implementation

In this section, two NNs are employed to approximate the unknown mismatched interconnection and the assumed differentiable local performance index function.

1) Approximation of the Interconnection: In this part, a state observer is employed to estimate the state of interconnected subsystem (2) as

$$\dot{\hat{x}}_i(t) = f_i(\hat{x}_i) + g_i(\hat{x}_i)u_i + \hat{h}_{id}(\hat{x}_{iD}) + l_i e_i \tag{15}$$

where $e_i = x_i - \hat{x}_i$ is the observation error, and $l_i = \text{diag}[l_{i1}, l_{i2}, \ldots, l_{in}] \in \mathbb{R}^{n_i * n_i}$ is the observer gain matrix with all positive elements. Noticing that the approximated unknown mismatched interconnection h_i is shown as (3), it can be approximated by $\hat{h}_{id}$, which is expressed as

$$\hat{h}_{id} = \hat{W}_{ih}^{\mathsf{T}} \sigma_{ih}(\hat{x}_{iD}) \tag{16}$$

whose weight vector is updated by

$$\dot{\hat{W}}_{ih} = \Gamma_{ih} e_i^{\mathsf{T}} \sigma_{ih}(\hat{x}_{iD}) \tag{17}$$

with $\Gamma_{ih} > 0$ a constant.

Combining (2), (4) with (15), we have

$$\begin{aligned}\dot{e}_i = &\left(f_i(x_i) - f_i(\hat{x}_i)\right) + \left(g_i(x_i) - g_i(\hat{x}_i)\right)u_i \\ &+ h_{id}(x_{iD}) + \Delta_i(x, x_{iD}) + \varepsilon_i - \hat{h}_{id}(\hat{x}_{iD}) - l_i e_i.\end{aligned}$$

Since $f_i(\cdot)$ and $g_i(\cdot)$ are locally Lipschitz, we have

$$\begin{aligned}\dot{e}_i &\leq D_{if}\|e_i\| + D_{ig}\|e_i\|u_i + h_{id}(x_{iD}) + \Delta_i(x, x_{iD}) \\ &\quad + \varepsilon_i - \hat{h}_{id}(\hat{x}_{iD}) - l_i e_i \\ &= D_{if}\|e_i\| + D_{ig}\|e_i\|u_i + \tilde{W}_{ih}^{\mathsf{T}}\sigma_{ih}(\hat{x}_{iD}) \\ &\quad + \Delta_i(x, x_{iD}) + \varepsilon_i - l_i e_i\end{aligned}$$

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

Theorem 2: Consider the interconnected subsystem (2), as well as the approximation of the unknown mismatched interconnection (16) and with the updated law as (17), the observation error e_i which is derived by combining (2) with the developed state observer (15) is guaranteed to be UUB.

Proof: Select the Lyapunov function candidate as

$$L_{i1} = \frac{1}{2} e_i^{\mathsf{T}} e_i + \frac{1}{2} \tilde{W}_{ih}^{\mathsf{T}} \Gamma_{ih}^{-1} \tilde{W}_{ih} \tag{18}$$

where $\tilde{W}_{ih} = W_{ih} - \hat{W}_{ih}$ is the weight approximation error.

Denoting $\eta_{i2} = D_{ig}\|e_i\|u_i + \Delta_i(x, x_{iD}) + \varepsilon_i$, the time derivative of (18) is

$$\begin{aligned}\dot{L}_{i1} &= e_i^{\mathsf{T}} \dot{e}_i - \tilde{W}_{ih}^{\mathsf{T}} \Gamma_{ih}^{-1} \dot{\hat{W}}_{ih} \\ &\leq e_i^{\mathsf{T}} \Big(D_{if}\|e_i\| + \tilde{W}_{ih}^{\mathsf{T}} \sigma_{ih}(x_{iD}) + \eta_{i2} - l_i e_i \Big) - \tilde{W}_{ih}^{\mathsf{T}} \Gamma_{ih}^{-1} \dot{\hat{W}}_{ih}.\end{aligned}$$

Suppose that the norm of the entire error is bounded as $\|\eta_{i2}\| \leq \phi_{i2}$ with $\phi_{i2} > 0$ as an unknown constant, we have

$$\begin{aligned}\dot{L}_{i1} \leq &-\left(\lambda_{\min}(l_i) - D_{if}\right)\|e_i\|^2 + \|e_i\|\phi_{i2} \\ &+ \tilde{W}_{ih}^{\mathsf{T}} \Big(e_i^{\mathsf{T}} \sigma_{ih}(x_{iD}) - \Gamma_{ih}^{-1} \dot{\hat{W}}_{ih} \Big).\end{aligned} \tag{19}$$

Substituting (17) into (19), we have

$$\begin{aligned}\dot{L}_{i1} &\leq -\left(\lambda_{\min}(l_i) - D_{if}\right)\|e_i\|^2 + \|e_i\|\phi_{i2} \\ &= \|e_i\|\Big(\left(-\lambda_{\min}(l_i) + D_{if}\right)\|e_i\| + \phi_{i2}\Big).\end{aligned}$$

Therefore, we can conclude that $\dot{L}_{i1} \leq 0$ when e_i lies outside of the compact set

$$\Omega_{e_i} = \left\{ e_i \ : \ \|e_i\| < \frac{\phi_{i2}}{\lambda_{\min}(l_i) - D_{\text{if}}} \right\}$$

where $\lambda_{\min}(l_i) > D_{if}$. According to the Lyapunov's direct method, the observation error is UUB with the approximation and substitution of the unknown mismatched interconnection. This completes the proof. ■

Remark 4: It is reasonable to assume η_{i1} and η_{i2} in Theorems 1 and 2 to be bounded. Take η_{i1} in Theorem 1 as an example, the bounded η_{i1} is necessary to guarantee the closed-loop system to be UUB, since we cannot promise η_{i1} is a positive or negative function. It means that $x_i(t)$ will converge to a small neighborhood, which may be smaller than the given Ω_{x_i}, but never larger than it.

Remark 5: From Theorems 1 and 2, we can see that the summaries are obtained by the boundedness assumptions of η_{i1} and η_{i2}. It indicates that the stability verifications are based on the boundedness of the states, rather than the boundedness on interconnections. Thus, it removes the assumption on available upper boundedness of interconnections in [44].

2) Critic Neural Network: Since the term $\hat{\delta}_i\|\nabla V_i(x_i)\|E_i$ in (7) is not completely known, we need to use parametric structures, such as NNs, to approximate it. We can observe that the unknown part $\nabla V_i(x_i)$ is the gradient along the corresponding state x_i of the critic NN, and it can be indirectly obtained by approximating $V_i(x_i)$ with a single layer NN on the compact set Ω_i as

$$V_i(x_i) = W_{ic}^{\mathsf{T}} \sigma_{ic}(x_i) + \varepsilon_{ic}(x_i) \tag{20}$$

where $W_{ic} \in \mathbb{R}^{l_i}$ is the ideal weight vector, $\sigma_i(x_i) \in \mathbb{R}^{l_i}$ is the activation function, l_i is the number of neurons in the hidden-layer, and $\varepsilon_{ic}(x_i)$ is the approximation error. Then, its gradient along corresponding state x_i is

$$\nabla V_i(x_i) = (\nabla \sigma_{ic}(x_i))^{\mathsf{T}} W_{ic} + \nabla \varepsilon_{ic}^{\mathsf{T}}(x_i) \tag{21}$$

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

From (20), the approximate critic NN can be expressed by

$$\hat{V}_i(x_i) = \hat{W}_{ic}^{\mathsf{T}}\sigma_{ic}(x_i).$$

Then, the gradient of $\hat{V}_i(x_i)$ along the corresponding state is

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

For the isolated subsystem (5), substituting (21) into the nonlinear Lyapunov function (8), we have

$$0 = \hat{\delta}_i\left\|\nabla V_i^{\mathsf{T}}(x_i)\right\|E_i + U_i(x_i,u_i) + \left(W_{ic}^{\mathsf{T}}\nabla\sigma_{ic}(x_i) + \nabla\varepsilon_{ic}(x_i)\right)^{\mathsf{T}} \times (f_i(x_i) + g_i(x_i)u_i(x_i) + h_{id}(x_{iD})).$$

Let $\upsilon_i = \|\nabla V_i^{\mathsf{T}}(x_i)\| - \|\nabla\hat{V}_i^{\mathsf{T}}(x_i)\|$, for the interconnected subsystem (2), the Hamiltonian can be expressed as

$$\begin{aligned} H_i(x_i,u_i,W_{ic}) &= \hat{\delta}_i\left\|\nabla\hat{V}_i^{\mathsf{T}}(x_i)\right\|E_i + U_i(x_i,u_i) \\ &\quad + W_{ic}^{\mathsf{T}}\nabla\sigma_{ic}(x_i)\dot{x}_i \\ &= -\hat{\delta}_i\upsilon_i E_i - \nabla\varepsilon_{ic}^{\mathsf{T}}(x_i)\dot{x}_i \\ &= e_{icH} \end{aligned} \tag{22}$$

where e_{icH} is the approximation error of the critic NN.

Thus, the approximate local Hamiltonian can be obtained by

$$\begin{aligned} H_i\left(x_i,u_i,\hat{W}_{ic}\right) &= \hat{\delta}_i\|\nabla\hat{V}_i^{\mathsf{T}}(x_i)\|E_i + U_i(x_i,u_i) + \hat{W}_{ic}^{\mathsf{T}}\nabla\sigma_{ic}(x_i)\dot{x}_i \\ &= e_{ic}. \end{aligned}$$

Let $\theta_i = \nabla\sigma_{ic}(x_i)\dot{x}_i$. By the steepest descent algorithm, the objective function $E_{ic} = (1/2)e_{ic}^{\mathsf{T}}e_{ic}$ can be minimized in order to adjust the weight vector of the critic NN $\hat{W}_{ic}$, which should be updated by

$$\dot{\hat{W}}_{ic} = -l_{ic}e_{ic}\theta_i \tag{23}$$

where $l_{ic} > 0$ is the learning rate.

Define the weight approximation error as $\tilde{W}_{ic} = W_{ic} - \hat{W}_{ic}$, according to (22) and (23), one has

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

The critic NN weight approximation error can be updated by

$$\dot{\tilde{W}}_{ic} = -\dot{\hat{W}}_{ic} = l_{ic}\left(e_{icH} - \tilde{W}_{ic}^{\mathsf{T}}\theta_i\right)\theta_i. \tag{24}$$

Therefore, according to (11) and (20), the ideal local control policy can be expressed as

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

And it can be approximated as

$$\hat{u}_i(x_i) = -\frac{1}{2}R_i^{-1}g_i^{\mathsf{T}}(x_i)(\nabla\sigma_{ic}(x_i))^{\mathsf{T}}\hat{W}_{ic}. \tag{25}$$

From the above equation, we can observe that the local control policy is derived by the critic NN, and the training of the action NN is no longer required.

Theorem 3: Consider the interconnected subsystem (2), the weight of the critic NN is updated by (24), the dynamics of the weight approximation error vector can be guaranteed to be UUB.

Proof: Select the Lyapunov function candidate as

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

Its time derivative is

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

Hence, $\dot{L}_{i2} < 0$ when $\tilde{W}_{ic}$ lies outside of the compact set

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

where $\|\theta_i\| \le \theta_{iM}$, and θ_{iM} is a positive constant. Based on the Lyapunov stability theorem, the dynamics of the weight approximation error vector is UUB. This completes the proof. ■

Remark 6: Since the convergence rate of RBFNN is higher than that of back propagation NN (BPNN), the RBFNN is employed by the developed local state observer (15). However, the local control policy (25) requires the partial derivative of local critic NN, which has heavy computational burden if RBFNN is employed. To tradeoff between the convergence rate and computational burden, BPNN is selected for local critic NN. Thus, different structures are chosen for these two NNs.

C. Stability Analysis

Theorem 4: Consider the interconnected subsystem (2), together with the improved local value function (7), where $\hat{\delta}_i$ is updated by

$$\dot{\hat{\delta}}_i = \Gamma_{i\delta}\left\|\left(\nabla V_i^*(x_i)\right)^{\mathsf{T}}\right\|E_i \tag{26}$$

and $\Gamma_{i\delta} > 0$ a constant, the N approximated decentralized control policies developed by (25) guarantee the closed-loop large-scale nonlinear system (1) to be UUB. In other words, the control policies $u_1(x_1), u_2(x_2), \ldots, u_N(x_N)$ are the decentralized control law for the large-scale nonlinear system composed of N subsystems as (2).

Proof: Select the Lyapunov function candidate for the ith interconnected subsystem as

$$L_{i3} = \sum_{i=1}^{N}\left(V_i^* + \frac{1}{2}\tilde{\delta}_i^{\mathsf{T}}\Gamma_{i\delta}^{-1}\tilde{\delta}_i\right).$$

Its time derivative is

$$\dot{L}_{i3} = \sum_{i=1}^{N}\left(\nabla V_i^{*\mathsf{T}}\dot{x}_i - \tilde{\delta}_i^{\mathsf{T}}\Gamma_{i\delta}^{-1}\dot{\hat{\delta}}_i\right).$$

According to (8) and (14), we can obtain

$$\dot{L}_{i3} \leq \sum_{i=1}^{N} \Big(\tilde{\delta}_i^{\mathsf{T}} \Big\| \nabla V_i^{*\mathsf{T}} \Big\| E_i - U_i(x_i, u_i) + \Big\| \nabla V_i^{*\mathsf{T}} \Big\| \phi_i - \tilde{\delta}_i^{\mathsf{T}} \Gamma_{i\delta}^{-1} \dot{\hat{\delta}}_i \Big). \quad (27)$$

Substituting (26) into (27), we have

$$\begin{aligned}\dot{L}_{i3} &\leq \sum_{i=1}^{N} \Big(-U_i(x_i, u_i) + \Big\| \nabla V_i^{*\mathsf{T}} \Big\| \phi_i \Big) \\ &\leq \sum_{i=1}^{N} \Big(-\lambda_{\min}(Q_i) \|x_i\|^2 + \Big\| \nabla V_i^{*\mathsf{T}} \Big\| \phi_i \Big).\end{aligned}$$

We can conclude that $\dot{\Sigma} \leq 0$ when x_i lies outside of the compact set

$$\Omega_{x_i} = \left\{ x_i \ : \ \|x_i\| < \sqrt{\frac{\big\| \nabla V_i^{*\mathsf{T}} \big\| \phi_i}{\lambda_{\min}(Q_i)}} \right\}.$$

From Lyapunov stability theorem, the closed-loop large-scale nonlinear system (1) is UUB with the control policies $u_1(x_1), u_2(x_2), \ldots, u_N(x_N)$. This completes the proof. ■

Remark 7: The positive function $\hat{\delta}_i$ defined in (6) can be updated by (26). It cannot guarantee $\hat{\delta}_i$ to be positive at the very beginning of updating. Noticing that the right hand side of (26) is positive when $t > 0$, $\hat{\delta}_i$ will be guaranteed to be positive all the time as long as its initial value $\hat{\delta}_{i0} \geq 0$ for updating. Thus, (7) can be guaranteed to be a Lyapunov equation with a proper initial value.

D. Local Online PI Algorithm

Here, a local online PI algorithm is introduced to solve HJB equations. The local online PI algorithm consists of the local policy evaluation based on (8) and the local policy improvement based on (11), and its iteration process can be described as Algorithm 1.

Algorithm 1 Local Online PI Algorithm

1: For $i = 1, 2, \ldots, N$, select a set of small positive constants ξ_i, let $p = 0$ and $V_i^{(0)}(x_i) = 0$, and begin with admissible control policies $u_i^{(0)}(x_i)$.

2: **(Local policy evaluation)** Let $p > 0$, based on the local control policy $u_i^{(p)}(x_i)$, solve the following local nonlinear Lyapunov equation for $u_i^{(p)}(x_i)$:

$$0 = \hat{\delta}_i \Big\| \nabla V_i^{(p)\mathsf{T}}(x_i) \Big\| E_i + U_i(x_i, u_i^{(p)}) + \nabla V_i^{(p)\mathsf{T}}(x_i)(f_i(x_i) + g_i(x_i)u_i(x_i) + h_{id}(x_{iD})). \quad (28)$$

3: **(Local policy improvement)** Update the local control policy $u_i^{(p)}(x_i)$ by

$$u_i^{(p+1)}(x_i) = -\frac{1}{2} R_i^{-1} g_i^{\mathsf{T}}(x_i) \nabla V_i^{(p)}(x_i). \quad (29)$$

4: If $\Big\| V_i^{(p+1)}(x_i) - V_i^{(p)}(x_i) \Big\| \leq \xi_i$, stop and obtain the approximated optimal control; else, let $p = p + 1$ and return to 2.

From Algorithm 1, we can see that $V_i^{(0)}(x_i) = 0$ is required. It is required to prove the convergence of Algorithm 1, e.g., $V_i^{(p)}(x_i) \to J_i^*(x_i)$ and $u_i^{(p)}(x_i) \to u_i^*(x_i)$ as $p \to \infty$.

Theorem 5: For the ith isolated subsystem (5), given N initial admissible control policies $u_i^{(0)}(x_i)$, where $i = 1, 2, \ldots, N$. Then, using the local PI algorithm described by (28) and (29), the improved local value functions and control policies converge to the optimal ones as $p \to \infty$, i.e., $V_i^{(p)}(x_i) \to J_i^*(x_i)$ and $u_i^{(p)}(x_i) \to u_i^*(x_i)$.

Proof: For the ith subsystem, we have $u_i^{(p)}(x_i) \in \Psi(\Omega_i)$ for any $p \geq 0$ with a given initial admissible control policy $u_i^{(0)}(x_i)$. Furthermore, there exists an integer p_{0i} for any ζ_i such that for any $p \geq p_{0i}$, the following formulas hold simultaneously:

$$\sup_{x_i \in \Omega_i} \Big| V_i^{(p)}(x_i) - J_i^*(x_i) \Big| < \zeta_i \quad (30)$$

$$\sup_{x_i \in \Omega_i} \Big| u_i^{(p)}(x_i) - u_i^*(x_i) \Big| < \zeta_i. \quad (31)$$

Similarly, this conclusion can be extended to the case of N isolated subsystems. Additionally, we denote $p_0 = \max\{p_{0i}\}$. Therefore, there exists any integer p_0 for any ζ, where $\zeta = \max\{\zeta_i\}$, such that for any $p \geq p_0$, (30) and (31) are true for $i = 1, 2, \ldots, N$. That is to say, the algorithm will converge to the improved local optimal value functions and local optimal controls of the N isolated subsystems. This completes the proof. ■

IV. Simulation Study

For large-scale nonlinear systems with unknown mismatched interconnections, two simulation examples are given in order to show the effectiveness of the proposed decentralized control scheme in this section.

Example 1: Consider the following large-scale nonlinear system:

$$\begin{aligned}\dot{x}_1 &= \begin{bmatrix} x_{12} - x_{11} \\ -0.5(x_{11} + x_{12}) - 0.5x_{12}(\cos(2x_{11}) + 2)^2 \end{bmatrix} + \begin{bmatrix} 0 \\ \cos(2x_1) + 2 \end{bmatrix} u_1(x_1) \\ &\quad + \begin{bmatrix} 0 \\ 4(x_{11} + x_{22}) \sin\big(x_{12}^3\big) \cos(0.5x_{21}) \end{bmatrix} \\ \dot{x}_2 &= \begin{bmatrix} x_{22} \\ -x_{21} - 0.5x_{22} + 0.5x_{21}^2 x_{22} \end{bmatrix} + \begin{bmatrix} 0 \\ x_{21} \end{bmatrix} u_2(x_2) \\ &\quad + \begin{bmatrix} 0 \\ 0.5(x_{12} + x_{22}) \cos\big(e^{x_{21}^2}\big) \end{bmatrix}\end{aligned}$$

where $x_i = [x_{i1}, x_{i2}]^{\mathsf{T}} \in \mathbb{R}^2$ and $u_i(x_i) \in \mathbb{R}$ are the state and control input of ith subsystem, respectively. It is assumed that the interconnections are unknown and mismatched.

Let the initial states of the system be $x_{10} = x_{20} = [1, -1]^{\mathsf{T}}$, and the initial states of the observer be $\hat{x}_{10} = [2, -2]^{\mathsf{T}}$ and $\hat{x}_{20} = [1.5, -1.5]^{\mathsf{T}}$, respectively. Because it is a regulation

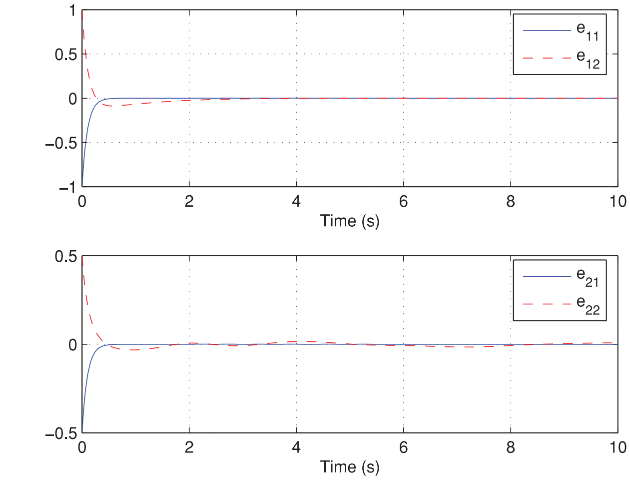

Fig. 1. State estimation errors of Example 1.

problem, the reference states of the coupled subsystems can be chosen as $x_{id} = 0$. In this simulation, the RBFNN in the local observer is chosen as 2–7–1 with 2 input neurons, 7 hidden neurons, and 1 output neuron. Meanwhile, the improved local value function (6) is approximated by a critic NN, whose structure is chosen as 2–3–1 with 2 input neurons, 3 hidden neurons, and 1 output neuron, and the weight vector as $\hat{W}_{ic} = [\hat{W}_{ic1}, \hat{W}_{ic2}, \hat{W}_{ic3}]^{\mathsf{T}}$ with the initial values $W_{1c0} = [1.6, 0.4, 0.6]^{\mathsf{T}}$ and $W_{2c0} = [0.3, 0.4, 1.3]^{\mathsf{T}}$. The activation function of the critic NN is selected as $\sigma_{ic}(x_i) = [x_{i1}^2, \ x_{i1}x_{i2}, \ x_{i2}^2]$. Let $Q_i = 20I_2$, $R_i = 20I$, the weight learning rates of the approximated interconnection and the critic NN be $\Gamma_{ih} = 10$ and $l_{ic} = 0.05$, the updated rate of $\hat{\delta}_i$ in improved local value function (7) be $\Gamma_{i\delta} = 0.0001$, the state observer gain matrix be $l_i = 10I_2$, where I_n denotes the identity matrix with n dimensions, respectively.

The simulation results are shown in Figs. 1–3. Fig. 1 illustrates the state estimation error by using the local state observer (15). It implies that the unknown interconnection can be approximated precisely online. We can see in Fig. 2, the weights of two critic NNs converge to $[2.297408, -0.339473, 1.384527]^{\mathsf{T}}$ and $[2.838536, -1.982101, 3.267496]^{\mathsf{T}}$. From Fig. 3, the system states can converge to zero by using the improved local value function (7) and the developed local PI algorithm. Therefore, the simulation results verify the effectiveness of the proposed decentralized control scheme.

Remark 8: In the considered large-scale nonlinear systems, the known dynamic $f_i(x_i)$ and $g_i(x_i)$ are estimated well when the estimation error is guaranteed to be UUB. Thus, in this case, the unknown mismatched interconnection can be approximated successfully.

Example 2: In order to further show the effectiveness of the proposed decentralized control scheme based on local PI algorithm, a hard spring connected parallel inverted pendulum system [50], [51] is employed in our simulation. The model of the parallel inverted pendulum system shown in Fig. 4 can be expressed by

$$m_1 l_1^2 \ddot{\theta}_1 - m_1 g l_1 \sin\theta_1 + b_1\dot{\theta}_1 - F a_1 \cos(\theta_1 - \beta) = \delta_1 u_1$$
$$m_2 l_2^2 \ddot{\theta}_2 - m_2 g l_2 \sin\theta_2 + b_2\dot{\theta}_2 - F a_2 \cos(\theta_2 - \beta) = \delta_2 u_2 \quad (32)$$

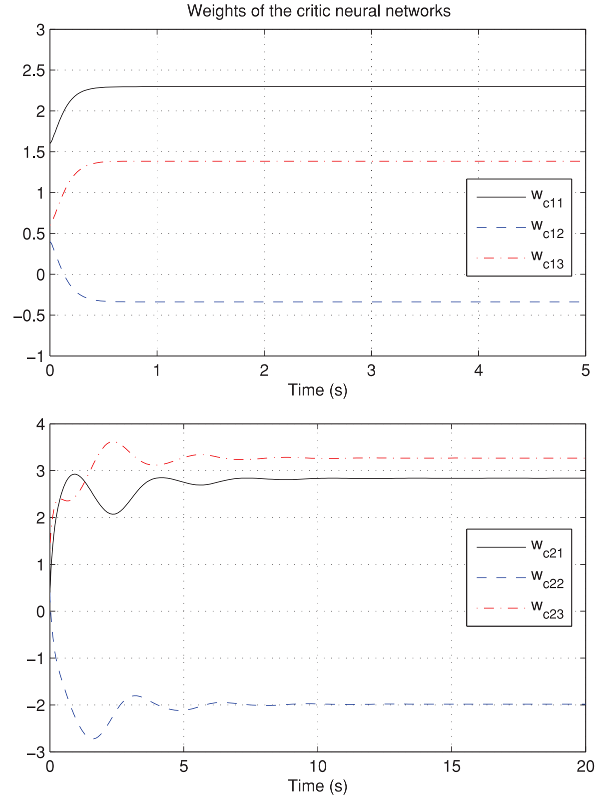

Fig. 2. Weights of critic NNs of Example 1.

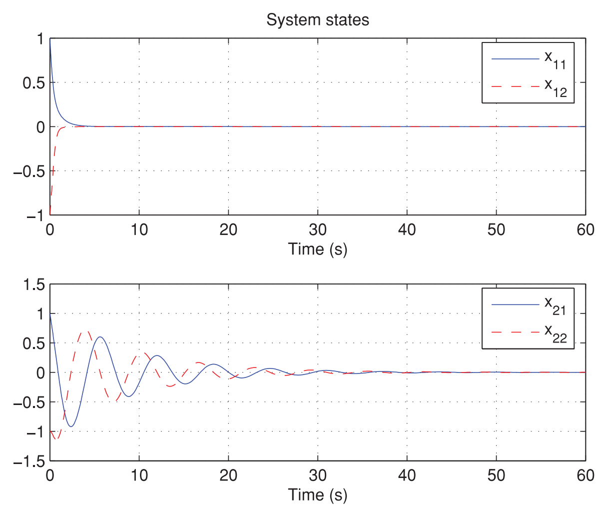

Fig. 3. System states of Example 1.

where b_1 and b_2 are damping coefficients, and

$$F = k\left\{1 + A^2(l_k - l_0)^2\right\}(l_k - l_0), \ |A(l_k - l_0)| < 1$$
$$\beta = \arctan\left(\frac{a_1\cos\theta_1 - a_2\cos\theta_2}{l_0 - a_1\sin\theta_1 + a_2\sin\theta_2}\right)$$
$$l_k = \left\{(l_0 - a_1\sin\theta_1 + a_2\sin\theta_2)^2 + (a_1\cos\theta_1 + a_2\cos\theta_2)^2\right\}^2.$$

In this simulation, 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 $a_1 = a_2 = 0.1$.

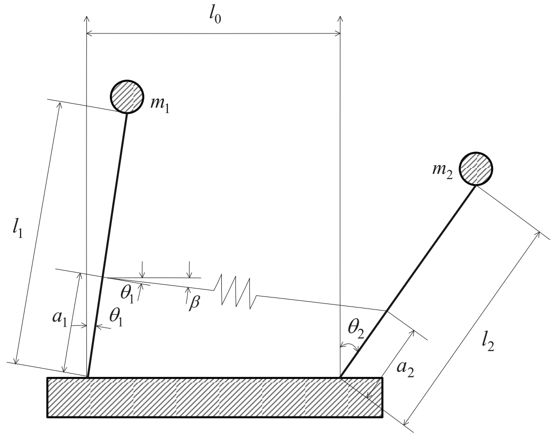

Fig. 4. Parallel inverted pendulum system.

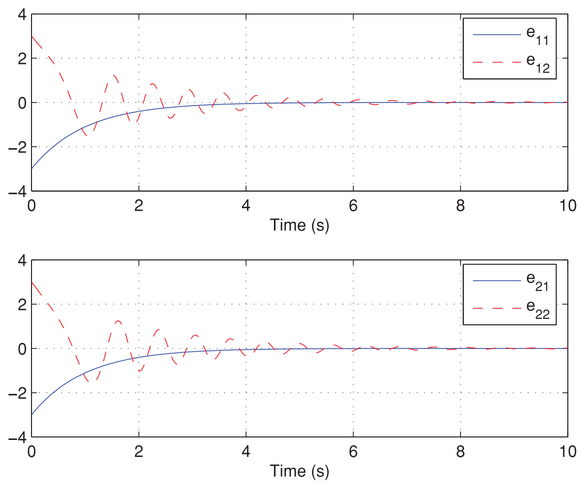

Fig. 5. State estimation errors of Example 2.

Let $x_i = [x_{i1}, x_{i2}]^{\mathsf{T}} = [\theta_i, \dot{\theta}_i]^{\mathsf{T}} \in \mathbb{R}^2$, the model (32) can be expressed as

$$\begin{aligned}\dot{x}_{11} &= x_{12}\\ \dot{x}_{12} &= \delta_1 u_1 + f_1(x_1) + h_1(x)\\ \dot{x}_{21} &= x_{22}\\ \dot{x}_{22} &= \delta_2 u_2 + f_2(x_2) + h_2(x)\end{aligned}$$

where $f_1(x_1) = 5.88 \sin x_{11} - 0.036x_{12}$, $f_2(x_2) = 5.88 \sin x_{21} - 0.036x_{22}$, $h_1(x) = 4Fa_1 \cos(x_{11} - \beta)$, and $h_2(x) = 4Fa_2 \cos(x_{21} - \beta)$.

Let the initial states of the parallel inverted pendulum, the structure of critic NN be the same as those of Example 1. Let initial values of the weight vectors, respectively, be $W_{1c0} = [1,\ 1.8,\ 1.6]^{\mathsf{T}}$ and $W_{2c0} = [1.6,\ 1,\ 1.2]^{\mathsf{T}}$, $Q_i = 0.1I_2$, and $R_i = 0.01I$, the weight learning rates of the approximated interconnection and the critic NN be $\Gamma_{ih} = 100$ and $l_{ic} = 0.2$, the updated rate of $\hat{\delta}_i$ in value function (7) be $\Gamma_{i\delta} = 0.001$, and the state observer gain matrix be $l_i = 10I_2$, respectively.

The simulation results are illustrated in Figs. 5–7. Fig. 5 illustrates the unknown interconnection can be estimated successful online. We can see in Fig. 6, the weights of two critic NNs converge to $[0.896337,\ 1.932218,\ 1.369870]^{\mathsf{T}}$ and $[0.621462,\ 1.340702,\ 1.113269]^{\mathsf{T}}$, respectively. Fig. 7 shows that the system states can converge to zero by using the presented decentralized control policy. The simulation results reveal that the proposed decentralized control scheme can be applied to large-scale nonlinear systems with unknown mismatched interconnections.

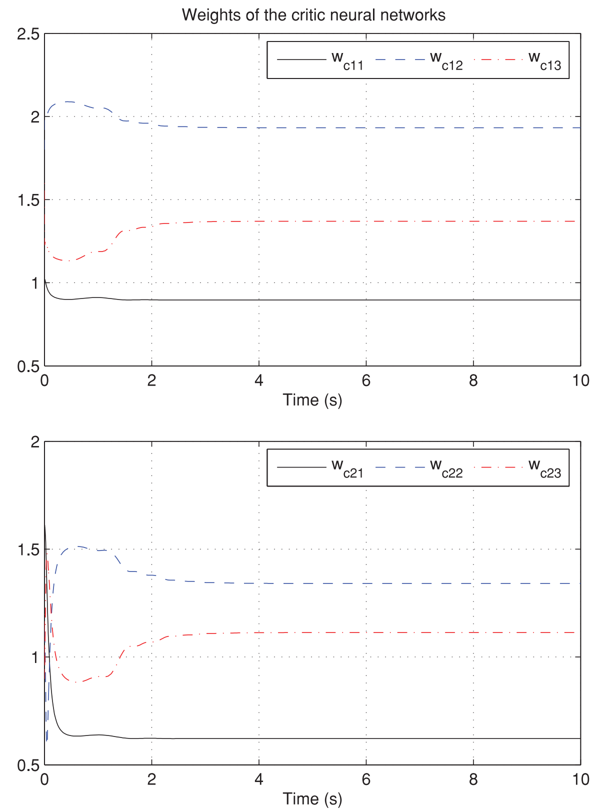

Fig. 6. Weights of critic NNs of Example 2.

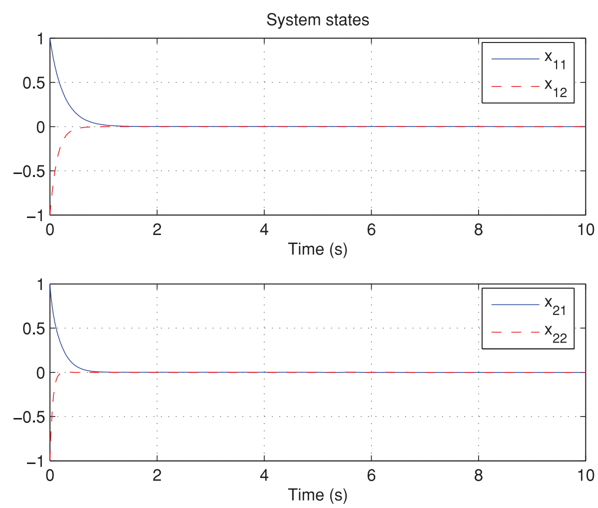

Fig. 7. System states of Example 2.

V. CONCLUSION

In this paper, we proposed a decentralized control scheme based on local PI algorithm for large-scale nonlinear systems with unknown mismatched interconnections. To relax the common boundedness assumption of the interconnection, the local states of isolated subsystem and the substituted reference states of coupled subsystems are employed to approximate interconnection terms. Then, an improved local performance index function is established to reflect the NN substitution error. At last, by the Lyapunov stability theorem, the closed-loop large-scale nonlinear system is guaranteed to be UUB via the developed decentralized control scheme. The simulation results ensure that the proposed decentralized control scheme is effective.

REFERENCES

[1] H. Wang, X. Liu, and K. Liu, "Robust adaptive neural tracking control for a class of stochastic nonlinear interconnected systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 3, pp. 510–523, Mar. 2016.

[2] T. Li, D. Wang, J. Li, and Y. Li, "Adaptive decentralized NN control of nonlinear interconnected time-delay systems with input saturation," *Asian J. Control*, vol. 15, no. 2, pp. 533–542, 2013.

[3] B. Zhao, Y. Li, and D. Liu, "Self-tuned local feedback gain based decentralized fault tolerant control for a class of large-scale nonlinear systems," *Neurocomputing*, vol. 235, pp. 147–156, Apr. 2017.

[4] B. Zhao and Y. Li, "Local joint information based active fault tolerant control for reconfigurable manipulator," *Nonlin. Dyn.*, vol. 77, no. 3, pp. 859–876, 2014.

[5] P. J. Werbos, "A menu of designs for reinforcement learning over time," in *Neural Network Control*. Cambridge, MA, USA: MIT Press, 1990, pp. 67–95.

[6] H. Wang, P. Shi, H. Li, and Q. Zhou, "Adaptive neural tracking control for a class of nonlinear systems with dynamics uncertainties," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2016.2607166.

[7] H. Wang, W. Sun, and P. X. Liu, "Adaptive intelligent control of nonaffine nonlinear time-delay systems with dynamic uncertainties," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2016.2627048.

[8] Z. Li, T. Li, and G. Feng, "Adaptive neural control for a class of stochastic nonlinear time-delay systems with unknown dead zone using dynamic surface technique," *Int. J. Robust Nonlin. Control*, vol. 26, no. 4, pp. 759–781, 2016.

[9] T. Li, Z. Li, D. Wang, and C. L. P. Chen, "Output-feedback adaptive neural control for stochastic nonlinear time-varying delay systems with unknown control directions," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 6, pp. 1188–1201, Jun. 2016.

[10] Y. Li, T. Li, and S. Tong, "Adaptive neural networks output feedback dynamic surface control design for MIMO pure-feedback nonlinear systems with hysteresis," *Neurocomputing*, vol. 198, pp. 58–68, Jul. 2016.

[11] Y. Li, T. Li, B. Miao, and C. L. P. Chen, "Adaptive NN control for a class of stochastic nonlinear systems with unmodeled dynamics using DSC technique," *Neurocomputing*, vol. 149, pp. 142–150, Feb. 2015.

[12] F. Y. Wang, H. Zhang, and D. Liu, "Adaptive dynamic programming: An introduction," *IEEE Comput. Intell. Mag.*, vol. 4, no. 2, pp. 39–47, May 2009.

[13] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, "Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 4, pp. 943–949, Aug. 2008.

[14] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs," *IEEE Trans. Neural Netw.*, vol. 8, no. 5, pp. 997–1007, Sep. 1997.

[15] D. P. Bertsekas and J. N. Tsitsiklis, "Neuro-dynamic programming: An overview," in *Proc. 34th IEEE Conf. Decis. Control*, 1995, pp. 560–564.

[16] D. Wang, C.-X. Mu, and D.-R. Liu, "Data-driven nonlinear near-optimal regulation based on iterative neural dynamic programming," *Acta Automatica Sinica*, vol. 43, no. 3, pp. 366–375, 2017.

[17] L. P. Kaelbling, M. L. Littman, and A. M. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, no. 1, pp. 237–285, 1996.

[18] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, "Adaptive optimal control for continuous-time linear systems based on policy iteration," *Automatica*, vol. 45, no. 2, pp. 477–484, 2009.

[19] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

[20] D. Liu, X. Yang, and H. Li, "Adaptive optimal control for a class of continuous-time affine nonlinear systems with unknown internal dynamics," *Neural Comput. Appl.*, vol. 23, nos. 7–8, pp. 1843–1850, 2013.

[21] D. Wang, D. Liu, Q. Wei, D. Zhao, and N. Jin, "Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming," *Automatica*, vol. 48, no. 8, pp. 1825–1832, 2012.

[22] H. Zhang, Y. Luo, and D. Liu, "Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraints," *IEEE Trans. Neural Netw.*, vol. 20, no. 9, pp. 1490–1503, Sep. 2009.

[23] D. Liu, Q. Wei, and P. Yan, "Generalized policy iteration adaptive dynamic programming for discrete-time nonlinear systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 12, pp. 1577–1591, Dec. 2015.

[24] L. Yang, J. Si, K. S. Tsakalis, and A. A. Rodriguez, "Direct heuristic dynamic programming for nonlinear tracking control with filtered tracking error," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 6, pp. 1617–1622, Dec. 2009.

[25] Q. Wei and D. Liu, "Adaptive dynamic programming for optimal tracking control of unknown nonlinear systems with application to coal gasification," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 4, pp. 1020–1036, Oct. 2014.

[26] B. Zhao, D. Liu, X. Yang, and Y. Li, "Observer-critic structure-based adaptive dynamic programming for decentralised tracking control of unknown large-scale nonlinear systems," *Int. J. Syst. Sci.*, 2017, doi: 10.1080/00207721.2017.1296982.

[27] X. Yang, D. Liu, and D. Wang, "Reinforcement learning for adaptive optimal control of unknown continuous-time nonlinear systems with input constraints," *Int. J. Control*, vol. 87, no. 3, pp. 553–566, 2014.

[28] D. Wang, D. Liu, and H. Li, "Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 627–632, Apr. 2014.

[29] Y. Jiang and Z.-P. Jiang, "Robust adaptive dynamic programming for large-scale systems with an application to multimachine power systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 10, pp. 693–697, Oct. 2012.

[30] D. Wang, D. Liu, H. Li, B. Luo, and H. Ma, "An approximate optimal control approach for robust stabilization of a class of discrete-time nonlinear systems with uncertainties," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 5, pp. 713–717, May 2016.

[31] D. Liu, H. Li, and D. Wang, "Online synchronous approximate optimal learning algorithm for multi-player non-zero-sum games with unknown dynamics," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 8, pp. 1015–1027, Aug. 2014.

[32] Z. Wang, L. Liu, H. Zhang, and G. Xiao, "Fault-tolerant controller design for a class of nonlinear MIMO discrete-time systems via online reinforcement learning algorithm," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 5, pp. 611–622, May 2016.

[33] B. Zhao, D. Liu, and Y. Li, "Online fault compensation control based on policy iteration algorithm for a class of affine non-linear systems with actuator failures," *IET Control Theory Appl.*, vol. 10, no. 15, pp. 1816–1823, Oct. 2016.

[34] B. Zhao, D. Liu, and Y. Li, "Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems," *Inf. Sci.*, vol. 384, pp. 21–33, Apr. 2017.

[35] G. K. Venayagamoorthy, R. G. Harley, and D. C. Wunsch, "Dual heuristic programming excitation neurocontrol for generators in a multimachine power system," *IEEE Trans. Ind. Appl.*, vol. 39, no. 2, pp. 382–394, Mar./Apr. 2003.

[36] D. Fuselli *et al.*, "Action dependent heuristic dynamic programming for home energy resource scheduling," *Int. J. Elect. Power Energy Syst.*, vol. 48, pp. 148–160, Jun. 2013.

[37] S. Song, G. Cai, and X. Lin, "Optimal neuron-controller for fluid triple-tank system via improved ADDHP algorithm," in *Advances in Computational Intelligence*. Berlin, Germany: Springer, 2009, pp. 483–492.

[38] J. Si and Y.-T. Wang, "Online learning control by association and reinforcement," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 264–276, Mar. 2001.

[39] A. Saberi, "On optimality of decentralized control for a class of nonlinear interconnected systems," *Automatica*, vol. 24, no. 1, pp. 101–104, 1988.

[40] T. Bian, Y. Jiang, and Z.-P. Jiang, "Decentralized adaptive optimal control of large-scale systems with application to power systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2439–2447, Apr. 2015.

[41] C. Lu, J. Si, and X. Xie, “Direct heuristic dynamic programming for damping oscillations in a large power system,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 4, pp. 1008–1013, Aug. 2008.
[42] D. Molina, G. K. Venayagamoorthy, J. Liang, and R. G. Harley, “Intelligent local area signals based damping of power system oscillations using virtual generators and approximate dynamic programming,” *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 498–508, Mar. 2013.
[43] D. S. Bernstein, C. Amato, E. A. Hansen, and S. Zilberstein, “Policy iteration for decentralized control of Markov decision processes,” *J. Artif. Intell. Res.*, vol. 34, no. 1, pp. 89–132, 2009.
[44] D. Liu, D. Wang, and H. Li, “Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 418–428, Feb. 2014.
[45] D. Liu, C. Li, H. Li, D. Wang, and H. Ma, “Neural-network-based decentralized control of continuous-time nonlinear interconnected systems with unknown dynamics,” *Neurocomputing*, vol. 165, pp. 90–98, Oct. 2015.
[46] D. Wang, H. He, B. Zhao, and D. Liu, “Adaptive near-optimal controllers for non-linear decentralised feedback stabilisation problems,” *IET Control Theory Appl.*, vol. 11, no. 6, pp. 799–806, Apr. 2017.
[47] A. Karimi, S. Eftekharnejad, and A. Feliachi, “Reinforcement learning based backstepping control of power system oscillations,” *Elect. Power Syst. Res.*, vol. 79, no. 11, pp. 1511–1520, 2009.
[48] S. Mehraeen and S. Jagannathan, “Decentralized optimal control of a class of interconnected nonlinear discrete-time systems by using online Hamilton-Jacobi-Bellman formulation,” *IEEE Trans. Neural Netw.*, vol. 22, no. 11, pp. 1757–1769, Nov. 2011.
[49] W. Chen and J. Li, “Decentralized output-feedback neural control for systems with unknown interconnections,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 1, pp. 258–266, Feb. 2008.
[50] Y. Tang, M. Tomizuka, G. Guerrero, and G. Montemayor, “Decentralized robust control of mechanical systems,” *IEEE Trans. Autom. Control*, vol. 45, no. 4, pp. 771–776, Apr. 2000.
[51] C. Hua, Y. Li, H. Wang, and X. Guan, “Decentralised fault-tolerant finite-time control for a class of interconnected non-linear systems,” *IET Control Theory Appl.*, vol. 9, no. 16, pp. 2331–2339, Oct. 2015.

Bo Zhao (M’16) received the B.S. degree in automation and the Ph.D. degree in control science and engineering from Jilin University, Changchun, China, in 2009 and 2014, respectively.

He 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 current research interests include adaptive dynamic programming, fault diagnosis and tolerant control, neural-network-based control, optimal control, and robot control.

Ding Wang (M’15) received the B.S. degree in mathematics from the Zhengzhou University of Light Industry, Zhengzhou, China, the M.S. degree in operations research and cybernetics from Northeastern University, Shenyang, China, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2007, 2009, and 2012, respectively.

He was a Visiting Scholar with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI, USA, from 2015 to 2017. He is currently an Associate Professor with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His current research interests include adaptive and learning systems, computational intelligence, and intelligent control. He has published over 90 journal and conference papers, and co-authored two monographs.

Dr. Wang was a recipient of the Excellent Doctoral Dissertation Award of Chinese Academy of Sciences in 2013, and a nomination of the Excellent Doctoral Dissertation Award of Chinese Association of Automation (CAA) in 2014. He serves as an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS and *Neurocomputing*. He is a member of Asia–Pacific Neural Network Society and CAA.

Guang Shi received the B.S. degree in automation from Zhejiang University, Hangzhou, China, in 2012. He is currently pursuing the Ph.D. degree with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China.

His current research interests include neural networks, adaptive dynamic programming, and optimal control and energy management in smart grids.

Derong Liu (S’91–M’94–SM’96–F’05) received the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, USA, in 1994.

He was a Staff Fellow with General Motors Research and Development Center, from 1993 to 1995. He was an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, USA, from 1995 to 1999. He joined the University of Illinois at Chicago, Chicago, IL, USA, in 1999, and became a Full Professor of Electrical and Computer Engineering and of Computer Science in 2006. He served as the Associate Director of the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China, from 2010 to 2015. He is currently a Full Professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing. He has published 15 books (six research monographs and nine edited volumes).

Dr. Liu was a recipient of the Faculty Early Career Development Award from the National Science Foundation in 1999, the University Scholar Award from University of Illinois from 2006 to 2009, the Overseas Outstanding Young Scholar Award from the National Natural Science Foundation of China in 2008, and the Outstanding Achievement Award from Asia Pacific Neural Network Assembly in 2014. He was selected for the “100 Talents Program” by the Chinese Academy of Sciences in 2008. He is an Elected AdCom Member of the IEEE Computational Intelligence Society and he is the Editor-in-Chief of the Artificial Intelligence Review. He was the General Chair of 2014 IEEE World Congress on Computational Intelligence and is the General Chair of 2016 World Congress on Intelligent Control and Automation. He is a Fellow of the International Neural Network Society.

Yuanchun Li received the Ph.D. degree in general mechanics from the Harbin Institute of Technology, Harbin, China, in 1990.

He is currently a Professor in Control Science and Engineering with the Changchun University of Technology, Changchun, China. His current research interests include adaptive dynamic programming, complex system modeling, and robot control, intelligent control.