

Decentralized Control for a Class of Nonlinear Interconnected Systems Using Integral Policy Iteration*

Chao Li, Derong Liu, *Fellow, IEEE*, Hongliang Li, and Ding Wang
The State Key Laboratory of Management and Control for Complex Systems
Institute of Automation, Chinese Academy of Sciences
Beijing 100190, China
{lichao2012, derong.liu, hongliang.li, ding.wang}@ia.ac.cn

Abstract—In this paper, we establish a decentralized control law to stabilize a class of nonlinear interconnected large-scale systems using a neural-network-based online model-free integral policy iteration algorithm. The model-free approach can solve the decentralized control problem for the interconnected system which has unknown dynamics. The stabilizing decentralized control strategy is derived based on the optimal control policies of the isolated subsystems. The online model-free integral policy iteration algorithm is developed to solve the optimal control problems for the isolated subsystems with unknown system dynamics. The actor-critic technique and the least squares implementation method are used to obtain the optimal control policies. A simulation example is given to verify the applicability of the decentralized control strategy.

Index Terms—Adaptive dynamic programming, decentralized control, optimal control, policy iteration, neural networks.

I. INTRODUCTION

Decentralized control uses local information of each subsystem in the controller implementation for large-scale interconnected systems. This overcomes the limitations of the traditional control that requires sufficiently information between subsystems. Therefore, the decentralized controllers have simpler architecture, and are more practical than the traditional centralized controllers. Various decentralized controllers have been established for large-scale interconnected systems in the presence of uncertainties and information structure constraints [1]–[7]. Generally speaking, a decentralized control strategy is comprised of some noninteracting local controllers corresponding to the isolated subsystems. In [8], the decentralized control consisted of the optimal control policies of the isolated subsystems was derived for the large-scale system. Therefore, the optimal control method can be applied to facilitate the design process of the decentralized control law.

The optimal control problem of nonlinear systems has been widely studied in the past few decades. The optimal control policy can be obtained by solving Hamilton-Jacobi-Bellman

(HJB) equation which is a partial differential equation. Because of the curse of dimensionality, this is a difficult task even in the case of completely known dynamics. Among the methods of solving the HJB equation, adaptive dynamic programming (ADP) has received increasing attention owing to its learning and optimal capacities [9]–[17]. Reinforcement learning (RL) is another computational method and it can interactively find an optimal policy [18]–[21]. Numerous excellent results have been obtained that greatly promote the development of relevant disciplines. Abu-Khalaf and Lewis [22] established an offline optimal control law for nonlinear systems with saturating actuators. Vamvoudakis and Lewis [23] derived a synchronous policy iteration (PI) algorithm to learn the online continuous-time optimal control with known dynamics. Vrabie and Lewis [24] derived an integral RL method to obtain direct adaptive optimal control for nonlinear input-affine continuous systems with partially unknown dynamics. Jiang and Jiang [25] presented a novel PI approach for continuous-time linear systems with complete unknown dynamics. Liu *et al.* [26] extended the PI algorithm to nonlinear optimal control problem with unknown dynamics and discounted cost function. Lee *et al.* [27], [28] presented an integral Q-learning algorithm for continuous-time systems without the exact knowledge of the system dynamics.

It is difficult to obtain the exact knowledge of the system dynamics for large-scale systems, such as transportation systems and power systems. The novelty of this paper is that we relax the assumptions of exact knowledge of the system dynamics required in the optimal controller design presented in [8]. In this paper, we use an online model-free integral PI to solve the decentralized control of a class of continuous-time nonlinear interconnected systems. We establish the stabilizing decentralized control law by adding feedback gains to the local optimal policies of the isolated subsystems. The optimal control problems for the isolated subsystems with unknown dynamics are related to develop the decentralized control law. To implement this algorithm, a critic neural network (NN) and an action NN are used to approximate the value function and control policy of the isolated subsystem, respectively. The effectiveness of the decentralized control law established in this paper is demonstrated by a simulation example.

* This work was supported in part by the National Natural Science Foundation of China under Grants 61034002, 61233001, 61273140, 61304086, and 61374105.

II. PROBLEM FORMULATION

We consider a continuous-time nonlinear large-scale system Σ composed of N interconnected subsystems described by

$$\begin{aligned}\Sigma : \dot{x}_i(t) &= f_i(x_i(t)) + g_i(x_i(t))(u_i(x_i(t)) + Z_i(x(t))) \\ i &= 1, 2, \dots, N\end{aligned}\quad (1)$$

where $x_i(t) \in \mathbb{R}^{n_i}$ is the state, $u_i(x_i(t)) \in \mathbb{R}^{m_i}$ is the control input vector of the i th subsystem. The overall state of the large-scale system Σ is denoted by $x = [x_1^\top \ x_2^\top \ \dots \ x_N^\top]^\top \in \mathbb{R}^n$, where $n = \sum_{i=1}^N n_i$. The local states are represented by x_1, x_2, \dots, x_N , whereas $u_1(x_1), u_2(x_2), \dots, u_N(x_N)$ are local control laws. For the i th subsystem, f_i is a continuous nonlinear internal dynamics function from \mathbb{R}^{n_i} into \mathbb{R}^{n_i} such that $f_i(0) = 0$. $g_i(x_i)$ is the input gain function from \mathbb{R}^{n_i} into $\mathbb{R}^{n_i \times m_i}$. $Z_i(x(t))$ is the interconnected term for the i th subsystem.

The i th isolated subsystem Σ_i is given by

$$\Sigma_i : \dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(x_i(t)). \quad (2)$$

For the i th isolated subsystem, we assume that the subsystem is controllable, $f_i + g_i u_i$ is Lipschitz continuous on a set Ω_i in \mathbb{R}^{n_i} , and there exists a continuous control policy that asymptotically stabilizes the subsystem. Let $x_i(0) = x_{i0}$ be the initial state of the i th subsystem. Additionally, we let the following assumptions hold through this paper.

Assumption 1: The state vector $x_i = 0$ is the equilibrium of the i th subsystem, $i = 1, 2, \dots, N$.

Assumption 2: The functions $f_i(\cdot)$ and $g_i(\cdot)$ are differentiable in their arguments, where $i = 1, 2, \dots, N$.

Assumption 3: The feedback control vector $u_i(x_i) = 0$ when $x_i = 0$, where $i = 1, 2, \dots, N$.

In this paper, we aim at finding N feedback control policies $u_1(x_1), u_2(x_2), \dots, u_N(x_N)$ as the decentralized control law to stabilize the large-scale system (1). As shown in [8], the decentralized control of the interconnected system is related to the optimal control of the isolated subsystems. We need to find the optimal control policy $u_i^*(x_i)$ of the i th isolated subsystem. The optimal control policy minimizes the following cost function

$$J_i(x_{i0}) = \int_0^\infty \{Q_i(x_i(\tau)) + u_i^\top(x_i(\tau))R_iu_i(x_i(\tau))\}d\tau$$

where $Q_i(x_i)$ is a positive definite function and R_i is a positive definite matrix.

III. COMPUTATIONAL DECENTRALIZED CONTROLLER DESIGN USING ONLINE MODEL-FREE PI ALGORITHM

In this section, we present the decentralized controller design. By adding some local feedback gains to the isolated optimal control policies, the decentralized control law is established. Then we develop a model-free integral PI algorithm to solve the optimal control problem with completely

unknown dynamics. An implementation of the established model-free integral PI algorithm is discussed at last.

A. Decentralized Control Law

According to [2], we modify the local optimal control laws $u_1^*(x_1), u_2^*(x_2), \dots, u_N^*(x_N)$ by proportionally increasing some local feedback gains to obtain a stabilizing decentralized control law for the interconnected system (1). We know that the feedback control laws

$$u_i(x_i) = \pi_i u_i^*(x_i) \quad (3)$$

can ensure that the N closed-loop isolated subsystems are asymptotically stable for all $\pi_i \geq 1/2$. However, only when $\pi_i = 1$, the feedback controls are optimal. In fact, similar results have been given in [29], showing that the optimal controls $u_i^*(x_i)$ are robust in the sense that they have infinite gain margins.

In [8], the stabilizing decentralized control law for the large-scale system (1) is provided: there exist N positive numbers $\pi_i^* > 0$, $i = 1, 2, \dots, N$, such that for any $\pi_i > \pi_i^*$, the feedback controls developed by (3) ensure that the closed-loop interconnected system is asymptotically stable. In other words, the control pair $(u_1(x_1), u_2(x_2), \dots, u_N(x_N))$ is the decentralized control law of interconnected system (1).

To design the decentralized control law, we need to solve the optimal control problems for the N isolated subsystems. We consider the i th isolated subsystem Σ_i in (2). For the feedback control policy $\mu_i(x_i)$, we assume that the associated cost function

$$V_i(x_{i0}) = \int_0^\infty \{Q_i(x_i(\tau)) + \mu_i^\top(x_i(\tau))R_i\mu_i(x_i(\tau))\}d\tau$$

is continuously differentiable. The infinitesimal version of the cost function is the nonlinear Lyapunov equation

$$\begin{aligned}Q_i(x_i) + \mu_i^\top(x_i)R_i\mu_i(x_i) \\ + (\nabla V_i(x_i))^\top(f_i(x_i) + g_i(x_i)\mu_i(x_i)) = 0\end{aligned}\quad (4)$$

with $V_i(0) = 0$. In (4), the term $\nabla V_i(x_i) = \partial V_i(x_i)/\partial x_i$, denotes the partial derivative of the local cost function $V_i(x_i)$ with respect to local state x_i .

The optimal cost function of the i th isolated subsystem can be formulated as

$$J_i^*(x_{i0}) = \min_{\mu_i} \int_0^\infty \{Q_i(x_i(\tau)) + \mu_i^\top(x_i(\tau))R_i\mu_i(x_i(\tau))\}d\tau,$$

and it satisfies the so-called HJB equation

$$0 = \min_{\mu_i} H_i(x_i, \mu_i, \nabla J_i^*(x_i))$$

where $\nabla J_i^*(x_i) = \partial J_i^*(x_i)/\partial x_i$. The Hamiltonian function is defined by

$$\begin{aligned}H_i(x_i, \mu_i, \nabla V_i(x_i)) = Q_i(x_i) + \mu_i^\top(x_i)R_i\mu_i(x_i) \\ + (\nabla V_i(x_i))^\top(f_i(x_i) + g_i(x_i)\mu_i(x_i)).\end{aligned}\quad (5)$$

In view of optimal control theory, by minimizing the Hamiltonian function (5), the optimal control policy for the i th isolated subsystem can be obtained as

$$\begin{aligned} u_i^*(x_i) &= \arg \min_{\mu_i} H_i(x_i, \mu_i, \nabla J_i^*(x_i)) \\ &= -\frac{1}{2} R_i^{-1} g_i^\top(x_i) \nabla J_i^*(x_i). \end{aligned} \quad (6)$$

B. Model-free PI Algorithm

In the following, we make effort to obtain the approximation solutions of the HJB equations related to the optimal control problems. First, we present the online PI algorithm with known dynamics to solve the HJB equations. The online PI algorithm consists of policy evaluation based on (4) and policy improvement based on (6), as shown in [8]. The successive procedure can be demonstrated as follows.

Algorithm 1 Online PI

1: Give a small positive real number ϵ . Let $p = 0$ and start with an initial stabilizing control policy $\mu_i^{(0)}(x_i)$.

2: **Policy Evaluation:**

Based on the control policy $\mu_i^{(p)}(x_i)$, solve the following nonlinear Lyapunov equations for $V_i^{(p)}(x_i)$

$$0 = Q_i(x_i) + (\mu_i^{(p)}(x_i))^\top R_i \mu_i^{(p)}(x_i) + (\nabla V_i^{(p)}(x_i))^\top (f_i(x_i) + g_i(x_i) \mu_i^{(p)}(x_i)). \quad (7)$$

3: **Policy Improvement:**

Update the control policies by

$$\mu_i^{(p+1)}(x_i) = -\frac{1}{2} R_i^{-1} g_i^\top(x_i) \nabla V_i^{(p)}(x_i). \quad (8)$$

4: If $\|V_i^{(p)}(x_i) - V_i^{(p-1)}(x_i)\| \leq \epsilon$, stop and obtain the approximate optimal control law of the i th isolated subsystem; else, set $p = p + 1$ and go to Step 2.

Then, we will develop an online model-free integral PI algorithm for optimal control problems with completely unknown dynamics. To deal with exploration which relaxes the assumptions of exact knowledge on $f_i(x_i)$ and $g_i(x_i)$, consider the following nonlinear subsystem explored by a known bounded piecewise continuous signal $e_i(t)$:

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))[u_i(x_i(t)) + e_i(t)]. \quad (9)$$

The derivative of the cost function with respect to time is calculated as

$$\dot{V}_i(x_i) = \nabla V_i^\top(x_i) (f_i(x_i) + g_i(x_i)[\mu_i(x_i) + e_i]). \quad (10)$$

Integral (10) from t to $t + T$ with any time interval $T > 0$

and considering (7) and (8), we have

$$\begin{aligned} V_i^{(p)}(x_i(t+T)) - V_i^{(p)}(x_i(t)) &= -2 \int_t^{t+T} \left(\mu_i^{(p+1)}(x_i) \right)^\top \\ &\times R_i e_i d\tau - \int_t^{t+T} \left\{ Q_i(x_i) + (\mu_i^{(p)}(x_i))^\top R_i \mu_i^{(p)}(x_i) \right\} d\tau. \end{aligned} \quad (11)$$

Equation (11) plays an important role in relaxing the assumption of the knowledge of system dynamics, since $f_i(x_i)$ and $g_i(x_i)$ do not appear in the equation. It means that the iteration can be done without knowing the system dynamics. Thus, we obtain the online model-free integral PI algorithm.

Algorithm 2 Online Model-free Integral PI

1: Give a small positive real number ϵ . Let $p = 0$ and start with an initial stabilizing control policy $\mu_i^{(0)}(x_i)$.

2: **Policy Evaluation and Improvement:**

Based on the control policy $\mu_i^{(p)}(x_i)$, solve the following nonlinear Lyapunov equations for $V_i^{(p)}(x_i)$ and $\mu_i^{(p+1)}(x_i)$

$$\begin{aligned} V_i^{(p)}(x_i(t)) &= \int_t^{t+T} \left\{ Q_i(x_i) + (\mu_i^{(p)}(x_i))^\top R_i \mu_i^{(p)}(x_i) \right\} d\tau \\ &+ 2 \int_t^{t+T} (\mu_i^{(p+1)}(x_i))^\top R_i e_i d\tau + V_i^{(p)}(x_i(t+T)). \end{aligned} \quad (12)$$

3: If $\|V_i^{(p)}(x_i) - V_i^{(p-1)}(x_i)\| \leq \epsilon$, stop and obtain the approximate optimal control law of the i th isolated subsystem; else, set $p = p + 1$ and go to Step 2.

Remark: In Algorithms 1 and 2, we let $V_i^{(p-1)}(x_i) = 0$, when $p = 0$.

C. Online NN Implementation

In this subsection, an implementation method of the established model-free PI algorithm is discussed, which is based on NN. A critic NN and an action NN are used to approximate the cost function and control policy of the subsystem. We assume that for the i th subsystem, $V_i^{(p)}(x_i)$ and $\mu_i^{(p+1)}(x_i)$ are represented on a compact set Ω_i by single-layer NNs as

$$\begin{aligned} V_i^{(p)}(x_i) &= (w_c^i)^\top \phi_c^i(x_i) + \varepsilon_c^i(x_i) \\ \mu_i^{(p+1)}(x_i) &= (w_a^i)^\top \phi_a^i(x_i) + \varepsilon_a^i(x_i) \end{aligned}$$

where $w_c^i \in \mathbb{R}^{N_c^i}$ and $w_a^i \in \mathbb{R}^{N_a^i}$ are the unknown bounded ideal weight parameters which will be determined by the established model-free PI algorithm, $\phi_c^i(x_i) \in \mathbb{R}^{N_c^i}$ and $\phi_a^i(x_i) \in \mathbb{R}^{N_a^i}$ are the continuously differentiable nonlinear activation functions, and $\varepsilon_c^i(x_i) \in \mathbb{R}$ and $\varepsilon_a^i(x_i) \in \mathbb{R}$ are the bounded NN approximation errors. Here, the subscripts ‘c’

and ‘ a ’ denotes the critic and the action, respectively. Since the ideal weights are unknown, the outputs of critic NN and action NN are

$$\hat{V}_i^{(p)}(x_i) = (\hat{w}_c^i)^\top \phi_c^i(x_i) \quad (13)$$

$$\hat{\mu}_i^{(p+1)}(x_i) = (\hat{w}_a^i)^\top \phi_a^i(x_i) \quad (14)$$

where \hat{w}_c^i and \hat{w}_a^i are the current estimated weights.

Using the expressions (13) and (14), (12) can be rewritten to be a general form

$$[\psi_k^i]^\top \begin{bmatrix} \hat{w}_c^i \\ \hat{w}_a^i \end{bmatrix} = \theta_k^i \quad (15)$$

with

$$\theta_k^i = \int_{t+(k-1)T}^{t+kT} \left\{ Q_i(x_i) + \left(\mu_i^{(p)}(x_i) \right)^\top R_i \mu_i^{(p)}(x_i) \right\} d\tau$$

$$\psi_k^i = \left[\left(\phi_c^i(x_i(t+(k-1)T)) - \phi_c^i(x_i(t+kT)) \right)^\top, \right.$$

$$\left. -2 \int_{t+(k-1)T}^{t+kT} R_i e_i \left(\phi_a^i(x_i) \right)^\top d\tau \right]^\top$$

where the measurement time is from $t+(k-1)T$ to $t+kT$. Since (15) is only a 1-dimensional equation, we cannot guarantee the uniqueness of the solution. Similar to [27], we use the least squares sense method to solve the parameter vector over a compact set Ω_i . For any positive integral K_i , we denote $\Phi_i = [\psi_1^i, \psi_2^i, \dots, \psi_{K_i}^i]$ and $\Theta_i = [\theta_1^i, \theta_2^i, \dots, \theta_{K_i}^i]^\top$. Then, we have the following K_i -dimensional equation

$$\Phi_i^\top \begin{bmatrix} \hat{w}_c^i \\ \hat{w}_a^i \end{bmatrix} = \Theta_i.$$

If Φ_i^\top has full column rank, the parameters can be solved by

$$\begin{bmatrix} \hat{w}_c^i \\ \hat{w}_a^i \end{bmatrix} = (\Phi_i \Phi_i^\top)^{-1} \Phi_i \Theta_i. \quad (16)$$

Therefore, we need to guarantee that the number of collected points K_i satisfies $K_i \geq \text{rank}(\Phi_i) = N_c^i + N_a^i$, which will make $(\Phi_i \Phi_i^\top)^{-1}$ exist. The least squares problem in (16) can be solved in real time by collecting enough data points generated from the system (9).

Clearly, the problem of designing the decentralized control law becomes to derive the optimal controllers for the N isolated subsystems. Based on the online model-free integral PI algorithm and NN techniques, we obtain the approximation solutions of the HJB equations. We can conclude that the approximate optimal control policies $\hat{\mu}_i(x_i)$ can be obtained. As shown in [8], by adding some local feedback gains to the isolated optimal control policies, we have the decentralized control law

$$u_i(x_i) = \pi_i \hat{\mu}_i(x_i). \quad (17)$$

Therefore, the stabilizing decentralized control law of the interconnected large-scale system is derived.

IV. NUMERICAL SIMULATION

A simulation example is provided in this section to demonstrate the effectiveness of the decentralized control law established in this paper.

We consider the following nonlinear interconnected system consisting of two subsystems

$$\begin{aligned} \dot{x}_1 &= \begin{bmatrix} -x_{11} + x_{12} \\ -0.5x_{11} - 0.5x_{12} - 0.5x_{12}(\cos(2x_{11}) + 2)^2 \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \cos(2x_{11}) + 2 \end{bmatrix} (u_1(x_1) + (x_{11} + x_{12}) \sin x_{12}^2 \cos(0.5x_{21})) \\ \dot{x}_2 &= \begin{bmatrix} x_{22} \\ -x_{21} - 0.5x_{22} + 0.5x_{21}^2 x_{22} \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ x_{21} \end{bmatrix} (u_2(x_2) + 0.5(x_{12} + x_{22}) \cos(e^{x_{21}})) \end{aligned} \quad (18)$$

where $x_1 = [x_{11} \ x_{12}]^\top \in \mathbb{R}^2$ and $u_1(x_1) \in \mathbb{R}$ are the state and control variables of subsystem 1, and $x_2 = [x_{21} \ x_{22}]^\top \in \mathbb{R}^2$ and $u_2(x_2) \in \mathbb{R}$ are the state and control variables of subsystem 2. We deal with the optimal control problem of these two isolated subsystems. According to [8], the cost functions of the optimal control problem are

$$\begin{aligned} J_1(x_{10}) &= \int_0^\infty \{x_{11}^2 + x_{12}^2 + u_1^\top u_1\} d\tau \\ J_2(x_{20}) &= \int_0^\infty \{x_{22}^2 + u_2^\top u_2\} d\tau. \end{aligned}$$

Assume that the exact knowledge of the dynamics (18) is fully unknown. We adopt the online model-free PI algorithm to tackle the optimal control problem.

For the isolated subsystem 1

$$\begin{aligned} \dot{x}_1 &= \begin{bmatrix} -x_{11} + x_{12} \\ -0.5x_{11} - 0.5x_{12} - 0.5x_{12}(\cos(2x_{11}) + 2)^2 \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \cos(2x_{11}) + 2 \end{bmatrix} u_1(x_1), \end{aligned}$$

we denote the weight vectors of the critic and action networks as

$$\begin{aligned} \hat{w}_c^1 &= [\hat{w}_{c1}^1 \ \hat{w}_{c2}^1 \ \hat{w}_{c3}^1]^\top \\ \hat{w}_a^1 &= [\hat{w}_{a1}^1 \ \hat{w}_{a2}^1]^\top. \end{aligned}$$

The activation functions are chosen as

$$\begin{aligned} \phi_c^1(x_1) &= [x_{11}^2 \ x_{11}x_{12} \ x_{12}^2]^\top \\ \phi_a^1(x_1) &= [x_{11}(2 + \cos(2x_{11})) \ x_{12}(2 + \cos(2x_{11}))]^\top. \end{aligned}$$

From these parameters, we know $N_c^1 = 3$ and $N_a^1 = 2$, so we conduct the simulation with $K_1 = 10$. We set the initial state and initial weights as $x_{10} = [1 \ -1]^\top$, $\hat{w}_c^1 = [0 \ 0 \ 0]^\top$ and $\hat{w}_a^1 = [-0.3 \ -0.9]^\top$. The period time $T = 0.1s$ and exploration $e_1(t) = 0.5 \sin(2\pi t)$ are used in the learning process. The least squares problem is solved after 10 samples are acquired, and thus the weights of the NNs

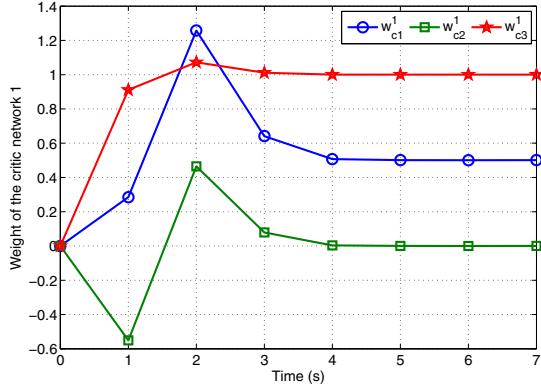


Fig. 1. Evolutions of the weight of the critic network 1.

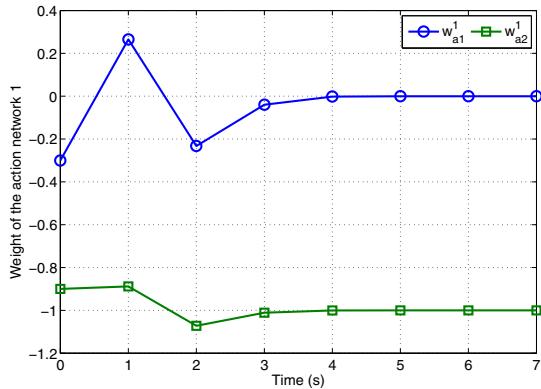


Fig. 2. Evolutions of the weight of the action network 1.

are updated every 1s. According to [23], the optimal cost function and control policy of the isolated subsystem 1 are $J_1^*(x_1) = 0.5x_{11}^2 + x_{12}^2$ and $u_1^*(x_1) = -(\cos(2x_{11}) + 2)x_{12}$. So the optimal weights $w_c^{1*} = [0.5 \ 0 \ 1]^\top$ and $w_a^{1*} = [0 \ -1]^\top$. Figs. 1 and 2 illustrate the evolutions of the critic weights and action weights, respectively. It is clear that the weights approximately converge to the optimal ones. At time $t = 7s$, $\hat{w}_c^1 = [0.5012 \ 0.0003 \ 1.0000]^\top$ and $\hat{w}_a^1 = [-0.0002 \ -1.0000]^\top$.

Similarly, for the isolated subsystem 2, the activation functions are chosen as

$$\begin{aligned}\phi_c^2(x_2) &= [x_{21}^2 \ x_{21}x_{22} \ x_{22}^2]^\top \\ \phi_a^2(x_2) &= [x_{21}^2 \ x_{21}x_{22}]^\top.\end{aligned}$$

As $N_c^2 = 3$ and $N_a^2 = 2$, we conduct the simulation with $K_2 = 10$. We set the initial state and initial weights as $x_{20} = [1 \ -1]^\top$, $\hat{w}_c^2 = [0 \ 0 \ 0]^\top$ and $\hat{w}_a^2 = [0 \ 0]^\top$. The period time $T = 0.1s$ and exploration $e_2(t) = 0.5 \sin(2\pi t)$ are used in the learning process. The optimal cost function and control policy of the isolated subsystem 2 are $J_2^*(x_2) = x_{21}^2 + x_{22}^2$ and $u_2^*(x_2) = -x_{21}x_{22}$. So the optimal

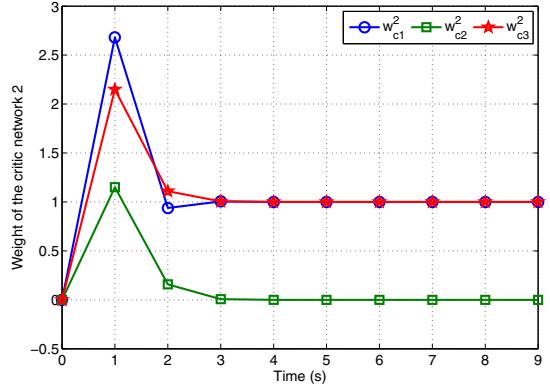


Fig. 3. Evolutions of the weight of the critic network 2.

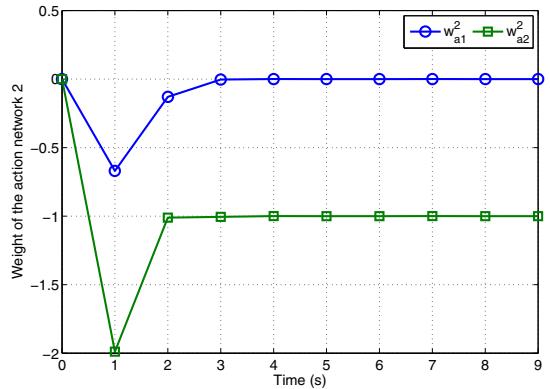


Fig. 4. Evolutions of the weight of the action network 2.

weights are $w_c^{2*} = [1 \ 0 \ 1]^\top$ and $w_a^{2*} = [0 \ -1]^\top$. Figs. 3 and 4 illustrate the evolutions of the critic weights and action weights, respectively. It is clear that the weights approximately converge to the optimal ones. At time $t = 9s$, $\hat{w}_c^2 = [1.0000 \ -0.0000 \ 1.0000]^\top$ and $\hat{w}_a^2 = [-0.0000 \ -1.0000]^\top$.

According to (17), we choose $\pi_1 = \pi_2 = 2$ to obtain the decentralized control law $(\pi_1 \hat{\mu}_1(x_1), \pi_2 \hat{\mu}_2(x_2))$ of the interconnected system (18). By applying the decentralized control law to control the interconnected system for 60s, we obtain the evolution process of the state trajectories shown in Figs. 5 and 6. Obviously, these simulation results testify the applicability of the decentralized control law developed in this paper.

V. CONCLUSION

In this paper, a stabilizing decentralized control law for a class of nonlinear large-scale systems is established. The decentralized control law is derived by the optimal controllers of the isolated subsystems. We use an online model-free integral PI algorithm with an exploration to solve the HJB equations related to the optimal control problem of the

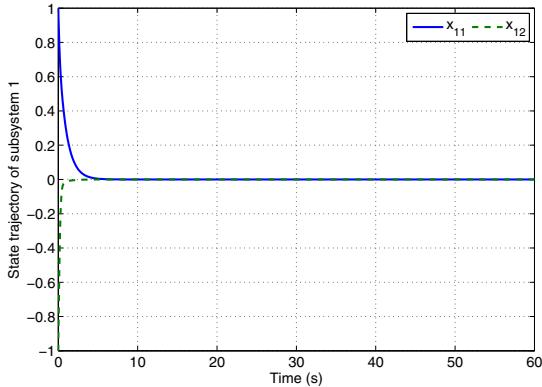


Fig. 5. State trajectory of subsystem 1 under the action of the decentralized control law.

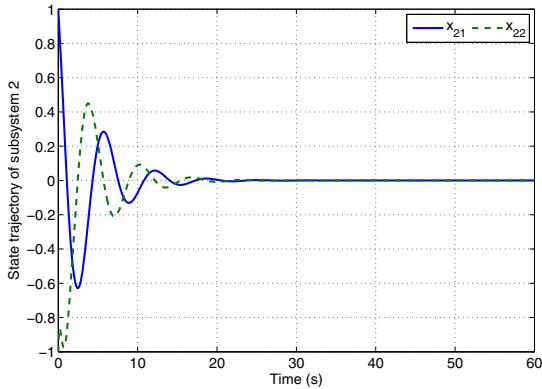


Fig. 6. State trajectory of subsystem 2 under the action of the decentralized control law.

isolated subsystems. To implement the constructed algorithm, we use the actor-critic technique and the least squares implementation method. We demonstrate the effectiveness of the developed decentralized control law by a simulation example.

REFERENCES

- [1] L. Bakule, "Decentralized control: An overview," *Annu. Rev. Control.*, vol. 32, no. 1, pp. 87–98, 2008.
- [2] A. Saberi, "On optimality of decentralized control for a class of nonlinear interconnected systems," *Automatica*, vol. 24, no. 1, pp. 101–104, 1988.
- [3] P. A. Ioannou, "Decentralized adaptive control of interconnected systems," *IEEE Trans. Automat. Contr.*, vol. 31, pp. 291–298, Apr. 1986.
- [4] J. T. Spooner and K. M. Passino, "Decentralized adaptive control of nonlinear systems using radial basis neural networks," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 2050–2057, Nov. 1999.
- [5] K. Kalsi, J. Lian, and S. H. Zak, "Decentralized dynamic output feedback control of nonlinear interconnected systems," *IEEE Trans. Automat. Contr.*, vol. 55, pp. 1964–1970, Aug. 2010.
- [6] J. Lavaei, "Decentralized implementation of centralized controllers for interconnected systems," *IEEE Trans. Automat. Contr.*, vol. 57, pp. 1860–1865, Jul. 2012.
- [7] D. D. Siljak, *Decentralized Control of Complex Systems*. Boston, MA, USA: Academic, 2012.
- [8] 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.*, in press, 2013.
- [9] F. Y. Wang, H. Zhang, and D. Liu, "Adaptive dynamic programming: an introduction," *IEEE Computational Magazine*, vol. 4, pp. 39–47, May 2009.
- [10] H. Zhang, Q. Wei, and D. Liu, "An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games," *Automatica*, vol. 47, no. 1, pp. 207–214, 2011.
- [11] D. Liu and Q. Wei, "Finite-approximation-error-based optimal control approach for discrete-time nonlinear systems," *IEEE Transactions on Cybernetics*, vol. 43, pp. 779–789, Apr. 2013.
- [12] D. Liu, Y. Zhang, and H. Zhang, "A self-learning call admission control scheme for CDMA cellular networks," *IEEE Trans. Neural Netw.*, vol. 16, pp. 1219–1228, Sept. 2005.
- [13] D. Liu, H. Javaherian, O. Kovalenko, and T. Huang, "Adaptive critic learning techniques for engine torque and air-fuel ratio control," *IEEE Trans. Syst., Man, Cybern. B*, vol. 38, pp. 988–993, Aug. 2008.
- [14] 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, pp. 693–697, Oct. 2012.
- [15] T. Dierks and S. Jagannathan, "Online optimal control of affine nonlinear discrete-time systems with unknown internal dynamics by using time-based policy update," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, pp. 1118–1129, July 2012.
- [16] T. Huang and D. Liu, "A self-learning scheme for residential energy system control and management," *Neural Computing and Applications*, vol. 22, no. 2, pp. 259–269, 2013.
- [17] Y. Jiang and Z. P. Jiang, "Robust adaptive dynamic programming with an application to power systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, pp. 1150–1156, July 2013.
- [18] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE Circuits and Systems Magazine*, vol. 9, pp. 32–50, Third Quarter 2009.
- [19] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.
- [20] J. Si, A. G. Barto, W. B. Powell, and D. Wunsch, *Handbook of Learning and Approximate Dynamic Programming*. Hoboken, NJ: Wiley, 2004.
- [21] S. J. Bradtke and B. E. Ydstie, "Adaptive linear quadratic control using policy iteration," in *Proc. American Control Conf.*, Baltimore, Maryland, June 1994, pp. 3475–3479.
- [22] M. Abu-Khalaf and F. L. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779–791, 2005.
- [23] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.
- [24] D. Vrabie and F. L. Lewis, "Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems," *Neural Networks*, vol. 22, no. 3, pp. 237–246, 2009.
- [25] 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.
- [26] D. Liu, X. Yang, and H. Li, "Adaptive optimal control for a class of continuous-time affine nonlinear systems with unknown internal dynamics," *Neural Computing and Applications*, vol. 23, no. 7, pp. 1843–1850, 2012.
- [27] J. Y. Lee, J. B. Park, and Y. H. Choi, "Integral Q-learning and explorized policy iteration for adaptive optimal control of continuous-time linear systems," *Automatica*, vol. 48, no. 11, pp. 2850–2859, 2012.
- [28] J. Y. Lee, J. B. Park, and Y. H. Choi, "Integral reinforcement learning with explorations for continuous-time nonlinear systems," in *Proc. Int. Joint Conf. Neural Networks*, Brisbane, Australia, June 2012, pp. 1042–1047.
- [29] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, 1997.