INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING
Int. J. Adapt. Control Signal Process. 2015; **00**:1–28
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/acs

Optimal Adaptive Control for Weakly Coupled Nonlinear Systems: A Neuro-Inspired Approach†

Luis Rodolfo García Carrillo[1], Kyriakos G. Vamvoudakis[2,*], João Pedro Hespanha[2]

[1] *School of Engineering and Computing Sciences, Texas A&M University - Corpus Christi, TX, 78412-5797 USA*
[2] *Center for Control, Dynamical-systems and Computation (CCDC), University of California, Santa Barbara, CA 93106-9560 USA*

SUMMARY

This paper proposes a new approximate dynamic programming algorithm to solve the infinite-horizon optimal control problem for weakly coupled nonlinear systems. The algorithm is implemented as a three-critics/four-actors approximators structure, where the critic approximators are used to learn the optimal costs, while the actor approximators are used to learn the optimal control policies. Simultaneous continuous-time adaptation of both critic and actor approximators is implemented, a method commonly known as synchronous policy iteration. The adaptive control nature of the algorithm requires a persistence of excitation condition to be a priori guaranteed, but this can be relaxed by using previously stored data concurrently with current data in the update of the critic approximators. Appropriate robustifying terms are added to the controllers to eliminate the effects of the residual errors, leading to asymptotic stability of the equilibrium point of the closed-loop system. Simulation results show the effectiveness of the proposed approach for a sixth-order dynamical example. Copyright © 2015 John Wiley & Sons, Ltd.

Received . . .

KEY WORDS: Weakly coupled systems; large scale systems; adaptive control; approximate dynamic programming; asymptotic stability; optimal control; reinforcement learning.

1. INTRODUCTION

Large-scale systems represent a challenging problem in optimal control [1], because their complexity can make numerical computations infeasible. A common approach for dealing with these kind of systems consists on splitting the large-scale design problem into a set of simpler problems or subsystems. As an example, the subsystems for the regulation of temperature, pressure, and flow, are designed separately in spite of their connection through a chemical plant. Similarly, such connected systems can be found in power systems, aircrafts, cars, communication networks. They are generally characterized by the presence of weak coupling between their subsystems. Practical knowledge may provide some guidance on how to split a large-scale problem into a set of simpler problems. But all these approaches completely neglect the coupling effect and most of the time the obtained results do not have a guaranteed performance level.

Weakly coupled linear systems have been studied extensively since their introduction to the control systems community by Kokotovic et al. [2] (see also for example [3], [4], and the references therein). Those systems have also been studied in mathematics [5], [6], economics [7], power system

*Correspondence to: K. G. Vamvoudakis, Center for Control, Dynamical-systems and Computation (CCDC), University of California, Santa Barbara, CA 93106-9560 USA. E-mail: kyriakos@ece.ucsb.edu
†This material is based upon work supported in part by ARO MURI Grant number W911NF0910553, ARO grant W911NF-09-D-0001 (Inst. for Collaborative Biotechnologies), University of California, Santa Barbara, CA 93106-9560 USA

engineering [8], [9], [10], and in nearly complete decomposable continuous-time and discrete-time Markov chains [11], [12], [13].

Due to the curse of dimensionality and the intractable form of the Hamilton-Jacobi-Bellman (HJB) equations that arise in optimal control, obtaining closed-form optimal control solutions for weakly coupled nonlinear systems is practically impossible. A first attempt for the optimization of coupled nonlinear systems was reported in [14] where the authors proposed a coupling perturbation method for near-optimum design. Approximate solutions of independent reduced-order HJB equations using Successive Galerkin Approximation (SGA) [15], [16], [17] have been proposed as alternative methods for solving the weakly coupled nonlinear optimal control problem. Unfortunately, the SGA method suffers from a computational complexity that increases with the dimension of the system under consideration.

Adaptive dynamic programming (ADP) techniques were proposed by Werbos [18], [19]. ADP brings together the advantages of adaptive and optimal control to obtain approximate and forward in time solutions to difficult optimization problems [20], [21], [22]. But, all the existing algorithms - such as the ones developed in [23], [24], [25], [26], [27], and the references therein - can only guarantee uniform ultimate boundedness of the closed-loop signals, i.e., a *milder* form of stability [28], and require a persistence of excitation condition to be satisfied for all time.

The need for adaptive controllers with the ability to learn optimal solutions for weakly coupled nonlinear systems, while also guaranteeing asymptotic stability of the equilibrium point of the closed-loop system motivates our research. The algorithm proposed, is motivated by a reinforcement learning algorithm called Policy Iteration (PI) [29] which is inspired by behaviorist psychology. To the best of our knowledge, there are not any *asymptotically stable online solutions* to the continuous-time HJB equation for weakly coupled nonlinear systems since couplings add nonlinearities to the HJB and make the problem more difficult.

1.1. Related work

A decoupling transformation that exactly decomposes weakly coupled linear systems composed of two subsystems into independent subsystems was introduced in [30]. These results were extended in [31] and in the book [32] to the general case of linear weakly coupled systems composed of N subsystems, and conditions under which such a transformation is feasible were established. The proposed optimal control algorithm is obtained in the form of a feedback law, where feedback gains are calculated from two independent reduced-order optimal control problems. In a similar way, the optimal control problem for weakly coupled bilinear systems was studied in [35], [33], and [34]. These results, were based on a recursive reduced-order scheme in order to solve the algebraic Riccati equation. Following this reduced-order scheme for solving the algebraic Riccati equation, the authors in [36] proposed a nonlinear optimal control for a weakly coupled nonlinear system based on the solution of two independent reduced-order HJB equations, using successive Galerkin approximation (SGA) [15], [16], [17]. The main drawback of this method is the offline design and that the computational complexity increases with the dimension of the system.

Moreover, in most of the adaptive control algorithms [47], there is a need for guaranteed persistence of excitation (PE) condition which is equivalent to space exploration in reinforcement learning [29], [37], [38]. This condition is restrictive in nonlinear systems and often difficult to guarantee in practice. Hence, convergence cannot be guaranteed. The work of [39] from the adaptive control side, and the works of [40] and [41] from the reinforcement learning side propose some alternatives that rely on concurrently using current and recorded data for adaptation to obviate the difficulty of guaranteeing convergence with PE. Recently the authors in [42] have used concurrent learning in optimal adaptive control but they only prove a *milder* form of stability, namely *uniform ultimate boundedness of the closed-loop signals* by using an approach that is based on integral reinforcement learning.

1.2. Contributions

The contributions of the paper rely on the development of an adaptive learning algorithm to solve the continuous-time optimal control problem with infinite horizon cost for weakly coupled nonlinear

DOI: 10.1002/acs

systems. The online adaptive algorithm is implemented as a three-critic/four-actor approximators structure, which involves continuous-time adaptation of both critic and actor approximators. The proposed algorithm is an appropriate combination of ideas from adaptive control, optimal control and reinforcement learning. Finally, we prove asymptotic stability of the equilibrium point of the closed-loop system.

Structure

The paper is structured as follows. In Section 2 we formulate the optimal control with saturated inputs problem. The approximate solution for the HJB equation is presented in Section 3. The Lyapunov proof that guarantees asymptotic stability of the closed-loop is presented in Section 5. Simulation results demonstrating the performance of the online algorithm acting on a weakly coupled system are given in Section 4. Finally Section 5 concludes and talks about future work.

Notation

The notation used here is standard. $\mathbb{R}^+$ is the set of positive real numbers and $\mathbb{Z}^+$ is the set of positive integer numbers. The superscript $*$ is used to denote the optimal solution, $\lambda_{\min}(A)$ is the minimum eigenvalue of a matrix A, $\lambda_{\max}(A)$ is the maximum eigenvalue of a matrix A and $\mathbf{1}_m$ is the column vector with m ones. The gradient of a scalar-valued function with respect to a vector-valued variable x is denoted as a column vector, and is denoted by $\nabla := \partial/\partial x$. V_x denotes the partial derivative a given function $V(x)$ with respect to x. A function $\alpha : \mathbb{R}^+ \to \mathbb{R}$ is said to belong to class $\mathcal{K}(\alpha \in \mathcal{K})$ functions if it is strictly increasing and $\alpha(0) = 0$.

2. PROBLEM FORMULATION

Consider the following weakly coupled nonlinear continuous-time system,

$$\dot{x} \equiv \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} f_{11}(x_1) + \varepsilon f_{12}(x) \\ \varepsilon f_{21}(x) + f_{22}(x_2) \end{bmatrix} + \begin{bmatrix} g_{11}(x_1) & \varepsilon g_{12}(x) \\ \varepsilon g_{21}(x) & g_{22}(x_2) \end{bmatrix} \begin{bmatrix} u_{11}(t) + \varepsilon u_{12}(t) \\ \varepsilon u_{21}(t) + u_{22}(t) \end{bmatrix}, \tag{1}$$

with an initial condition,

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix},$$

where $x_1 \in \mathbb{R}^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$, with $n_1 + n_2 = n$ are the states that can be measured, $u_{1i} \in \mathbb{R}^{m_1}$, $u_{2i} \in \mathbb{R}^{m_2}$, with $m_1 + m_2 = m$, $i \in \{1, 2\}$ are the control inputs and $\varepsilon \in \mathbb{R}^+$ is a small coupling parameter. Moreover, $x = [x_1^{\mathrm{T}} \;\; x_2^{\mathrm{T}}]^{\mathrm{T}}$ is the full state variable and $u = [u_{11}^{\mathrm{T}} + \varepsilon u_{12}^{\mathrm{T}} \quad \varepsilon u_{21}^{\mathrm{T}} + u_{22}^{\mathrm{T}}]^{\mathrm{T}} \in U \subseteq \mathbb{R}^m$ is the total control input. We assume that $f_{1i}(\cdot) \in \mathbb{R}^{n_1}$, $f_{2i}(\cdot) \in \mathbb{R}^{n_2}$ and $g_{ij}(\cdot) \in \mathbb{R}^{n_i \times m_j}$ are known functions. We also assume that $f_{1i}(0) = 0$ and $f_{2i}(0) = 0$ for $i \in \{1, 2\}$.

It is desired to minimize the following infinite horizon cost functional associated with (1),

$$V = \frac{1}{2} \int_0^\infty r\left(x(\tau), u(\tau)\right) d\tau, \;\forall\, x(0) \tag{2}$$

where,

$$r(x, u) := x^{\mathrm{T}} Q x + u^{\mathrm{T}} R u, \;\; \forall\, x, u, \tag{3}$$

where the matrices $Q \geq 0$ and $R > 0$ have the following weakly coupled structures,

$$Q = \begin{bmatrix} Q_1 & \varepsilon Q_2 \\ \varepsilon Q_2^{\mathrm{T}} & Q_3 \end{bmatrix} \qquad\qquad R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix},$$

Int. J. Adapt. Control Signal Process. (2015)
DOI: 10.1002/acs

with, Q_1, Q_2, Q_3 and R_1, R_2 matrices of appropriate dimensions.

The cost functional for the weakly coupled system can be rewritten as follows,

$$V = V_1(x_1, u_{11}) + \varepsilon V_2(x, u_{11}, u_{12}, u_{21}, u_{22}) + V_3(x_2, u_{22}), \tag{4}$$

where,

$$V_1(x_1, u_{11}) = \frac{1}{2}\int_0^\infty \left\{x_1^{\mathrm{T}} Q_1 x_1 + u_{11}^{\mathrm{T}} R_1 u_{11}\right\} dt, \forall x_1, u_{11}, \tag{5}$$

$$V_2(x, u_{11}, u_{12}, u_{21}, u_{22}) = \int_0^\infty \left\{x_1^{\mathrm{T}} Q_2 x_2 + u_{11}^{\mathrm{T}} R_1 u_{12} + u_{22}^{\mathrm{T}} R_2 u_{21}\right\} dt, \forall x, u_{11}, u_{12}, u_{21}, u_{22} \tag{6}$$

$$V_3(x_2, u_{22}) = \frac{1}{2}\int_0^\infty \left\{x_2^{\mathrm{T}} Q_3 x_2 + u_{22}^{\mathrm{T}} R_2 u_{22}\right\} dt, \forall x_2, u_{22}. \tag{7}$$

The Hamiltonian of the system (1) associated with the cost function (2)-(3) after setting $\varepsilon^2 = 0$ is expressed by the following $\mathcal{O}(\varepsilon^2)$ approximation,

$$H = H_1 + \varepsilon H_2 + H_3, \tag{8}$$

where,

$$H_1 = \frac{1}{2} x_1^{\mathrm{T}} Q_1 x_1 + V_{1x_1}^{\mathrm{T}} f_{11}(x_1) + V_{1x_1}^{\mathrm{T}} g_{11}(x_1) u_{11} + \frac{1}{2} u_{11}^{\mathrm{T}} R_1 u_{11}, \forall x_1, u_{11}, \tag{9}$$

$$\begin{aligned} H_2 = {} & x_1^{\mathrm{T}} Q_2 x_2 + u_{11}^{\mathrm{T}} R_1 u_{12} + u_{22}^{\mathrm{T}} R_2 u_{21} + V_{1x_1}^{\mathrm{T}} f_{12}(x) + V_{3x_2}^{\mathrm{T}} f_{21}(x) \\ & + V_{1x_1}^{\mathrm{T}} g_{11}(x_1) u_{12} + V_{1x_1}^{\mathrm{T}} g_{12}(x) u_{22} \\ & + V_{2x_1}^{\mathrm{T}} g_{11}(x_1) u_{11} + V_{2x_2}^{\mathrm{T}} g_{22}(x_2) u_{22} + V_{2x_1}^{\mathrm{T}} f_{11}(x_1) + V_{2x_2}^{\mathrm{T}} f_{22}(x_2) \\ & + V_{3x_2}^{\mathrm{T}} g_{21}(x) u_{11} + V_{3x_2}^{\mathrm{T}} g_{22}(x_2) u_{21},\ \forall x, u_{11}, u_{12}, u_{21}, u_{22}, \end{aligned} \tag{10}$$

$$H_3 = \frac{1}{2} x_2^{\mathrm{T}} Q_3 x_2 + V_{3x_2}^{\mathrm{T}} f_{22}(x_2) + V_{3x_2}^{\mathrm{T}} g_{22}(x_2) u_{22} + \frac{1}{2} u_{22}^{\mathrm{T}} R_2 u_{22}, \forall x_2, u_{22}. \tag{11}$$

Hence, the ultimate goal is to find the following optimal value function,

$$V^* = \min_{u \in U} \int_t^\infty r(x, u) d\tau, \quad t \geqslant 0, \tag{12}$$

subject to the state dynamics in (1).

The optimal value V^* satisfies the following HJB equation (see [1] for an existence theorem),

$$\frac{1}{2} x^{\mathrm{T}} Q x + \frac{1}{2} u^{*\mathrm{T}} R u^* + V_x^{*\mathrm{T}} f(x) + V_x^{*\mathrm{T}} g(x) u^* = 0, \tag{13}$$

where,

$$f(x) = \left[f_{11} + \varepsilon f_{12} \quad \varepsilon f_{21} + f_{22}\right]^{\mathrm{T}} \quad ; \quad g(x) = \begin{bmatrix} g_{11} & \varepsilon g_{12} \\ \varepsilon g_{21} & g_{22} \end{bmatrix},$$

and u^* is the optimal control that will be found later.

Assumption 1 (Smoothness of solution)
The solution to (13) is smooth, i.e. $V^* \in C^1$, and positive definite with $V^*(0) = 0$. □

Remark 1
Hamilton-Jacobi equations are nonlinear partial differential equations, and it is well-known that in general such equations do not admit global classical solutions and if they do, they may not be smooth. But they may have the so-called viscosity solutions [43]. Under certain local reachability and observability assumptions, they have local smooth solutions [44]. Various other assumptions guarantee existence of smooth solutions, such as that the dynamics not be bilinear and the cost function not contain cross-terms in the state and control input. The latter two assumptions are satisfied for the system (1) and the cost (2) under consideration. □

We shall now, split the solution $V_x^{*\mathrm{T}}$ as follows,

$$V_x^{*\mathrm{T}} = \begin{bmatrix} V_{1x_1}^{*\mathrm{T}} & + & \varepsilon V_{2x_1}^{*\mathrm{T}} \\ \varepsilon V_{2x_2}^{*\mathrm{T}} & + & V_{3x_2}^{*\mathrm{T}} \end{bmatrix}. \tag{14}$$

The optimal control input for the system (1) with the optimal value function (12) can be obtained using the stationarity condition in the Hamiltonian (8),

$$\frac{\partial H}{\partial u} = 0 \Rightarrow u^* = -R^{-1} g^{\mathrm{T}}(x) V_x^{*\mathrm{T}}(x), \tag{15}$$

which can be split into the following control inputs,

$$u_{11}^*(x_1) = -R_1^{-1} g_{11}^{\mathrm{T}}(x_1) V_{1x_1}^{*\mathrm{T}}, \forall x_1, \tag{16}$$
$$u_{12}^*(x) = -R_1^{-1} \left(g_{11}^{\mathrm{T}}(x_1) V_{2x_1}^{*\mathrm{T}} + g_{21}^{\mathrm{T}}(x) V_{3x_2}^{*\mathrm{T}} \right), \forall x, \tag{17}$$
$$u_{21}^*(x) = -R_2^{-1} \left(g_{22}^{\mathrm{T}}(x_2) V_{2x_2}^{*\mathrm{T}} + g_{12}^{\mathrm{T}}(x) V_{1x_1}^{*\mathrm{T}} \right), \forall x, \tag{18}$$
$$u_{22}^*(x_2) = -R_2^{-1} g_{22}^{\mathrm{T}}(x_2) V_{3x_2}^{*\mathrm{T}}, \forall x_2. \tag{19}$$

After substituting the optimal controls (16)-(19) into the Hamiltonians (9)-(11) one has the following 3 HJB equations $H_i^* = 0,\ \forall i \in \{1, 2, 3\}$,

$$0 = \frac{1}{2} x_1^{\mathrm{T}} Q_1 x_1 + V_{1x_1}^{*\mathrm{T}} f_{11}(x_1) + V_{1x_1}^{*\mathrm{T}} g_{11}(x_1) u_{11}^*(x_1) + \frac{1}{2} u_{11}^{*\mathrm{T}}(x_1) R_1 u_{11}^*(x_1), \tag{20}$$

$$\begin{aligned} 0 =& x_1^{\mathrm{T}} Q_2 x_2 + V_{1x_1}^{*\mathrm{T}} f_{12}(x) + V_{3x_2}^{*\mathrm{T}} f_{21}(x) \\ &+ V_{1x_1}^{*\mathrm{T}} g_{12}(x) u_{22}^*(x_2) + V_{2x_1}^{*\mathrm{T}} g_{11}(x_1) u_{11}^*(x_1) + V_{2x_2}^{*\mathrm{T}} g_{22}(x_2) u_{22}^*(x_2) \\ &+ V_{2x_1}^{*\mathrm{T}} f_{11}(x_1) + V_{2x_2}^{*\mathrm{T}} f_{22}(x_2) + V_{3x_2}^{*\mathrm{T}} g_{21}(x) u_{11}^*(x_1), \end{aligned} \tag{21}$$

$$0 = \frac{1}{2} x_2^{\mathrm{T}} Q_3 x_2 + V_{3x_2}^{*\mathrm{T}} f_{22}(x_2) + V_{3x_2}^{*\mathrm{T}} g_{22}(x_2) u_{22}^*(x_2) + \frac{1}{2} u_{22}^{*\mathrm{T}}(x_2) R_2 u_{22}^*(x_2). \tag{22}$$

Due to the nonlinear nature of these three weakly coupled HJB equations, finding their solution is generally difficult or impossible.

The following section shall provide approximate solutions to equations (20), (21), and (22).

3. APPROXIMATE SOLUTION

The next subsections will lay the foundation for updating the optimal value function and the optimal control input simultaneously by using data collected along the closed-loop trajectory.

3.1. Critic approximators and recorded past data

The first step to solve the HJB equations (20), (21), and (22) is to approximate the value function $V^*(x)$ in equation (12) on any given compact set $\Omega \subseteq \mathbb{R}^n$ with a critic approximator as follows,

$$V^*(x) = {W^*}^{\mathrm{T}} \phi(x) + \epsilon(x),\ \forall x, \tag{23}$$

where $W^* \in \mathbb{R}^{N_{tot}}$ are the ideal weights satisfying $\|W^*\| \leqslant W_{\max}$; $\phi(x) : \Omega \to \mathbb{R}^{N_{tot}}$, $\phi(x) = [\varphi_1(x)\ \varphi_2(x)\ \ldots\ \varphi_{N_{tot}}(x)]^{\mathrm{T}}$ are the basis functions, such that $\varphi_i(0) = 0$ and $\nabla \varphi_i(0) = 0$, $\forall i = 1, \ldots, N_{tot}$; N_{tot} is the number of neurons in the hidden layer and $\epsilon(x)$ is the approximation error. It has been shown in [46] that NNs with a single hidden layer and an appropriately smooth hidden layer activation function are capable of arbitrarily accurate approximation to an arbitrary function and its derivatives.

One should pick the basis functions $\varphi_i(x)$, $\forall i \in \{1, 2, \ldots, N_{tot}\}$ as polynomial, radial basis or sigmoidal functions. In this case, V^* and its derivatives,

$$\begin{aligned} V_x^*(x) &= \left[\frac{\partial}{\partial x} \phi(x) \right]^{\mathrm{T}} W^* + \frac{\partial}{\partial x} \epsilon(x), \\ &=: \nabla \phi^{\mathrm{T}} W^* + \nabla \epsilon,\ \forall x \in \Omega, \end{aligned} \tag{24}$$

DOI: 10.1002/acs

can be uniformly approximated on any given compact set Ω. According to Weierstrass higher order approximation Theorem [45], a polynomial suffices to approximate V^* as well as its derivatives when they exist. Moreover, as the number of basis sets N_{tot} increases, the approximation error on a compact set Ω goes to zero, i.e., $\epsilon(x) \to 0$ as $N_{tot} \to \infty$.

We shall require a form of uniformity in this approximation result that is common in neuro-adaptive control and other approximation techniques [46, 47]. We shall now write the *approximate Hamiltonian* as,

$$H^* := \frac{1}{2}x^{\mathrm{T}}Qx + \frac{1}{2}u^{\mathrm{T}}Ru + W^{*\mathrm{T}}\nabla\phi(f(x) + gu),\ \forall x, u, \tag{25}$$

with a residual error given as,

$$\epsilon_H := H^* - H = -\nabla\epsilon^{\mathrm{T}} f(x) - \nabla\epsilon^{\mathrm{T}}(f + gu),\ \forall x, u, \tag{26}$$

with,

$$H^* = H_1^* + \varepsilon H_2^* + H_3^*,$$

and,

$$\epsilon_H = \epsilon_{H_1} + \varepsilon\epsilon_{H_2} + \epsilon_{H_3}.$$

Assumption 2 (Critic Uniform Approximation)
The critic activation functions ϕ, the value function approximation error ϵ, their derivatives, and the Hamiltonian residual error ϵ_H are all uniformly bounded on a set $\Omega \subseteq \mathbb{R}^n$, in the sense that there exist known finite constants $\phi_{\mathrm{m}}, \phi_{\mathrm{dm}}, \epsilon_{\mathrm{m}}, \epsilon_{\mathrm{dm}}, \epsilon_{\mathrm{Hm}} \in \mathbb{R}^+$ such that $|\phi(x)| \leqslant \phi_{\mathrm{m}}$, $|\nabla\phi(x)| \leqslant \phi_{\mathrm{dm}}$, $|\epsilon(x)| \leqslant \epsilon_{\mathrm{m}}$, $|\nabla\epsilon(x)| \leqslant \epsilon_{\mathrm{dm}}$, $|\epsilon_H(x)| \leqslant \epsilon_{\mathrm{Hm}}$, $\forall x \in \Omega$. □

Since the ideal weights for the value function $V^*(x)$ that appear in (23) are unknown, one must consider the actual *critic weight estimates* $\hat{W} \in \mathbb{R}^{N_{tot}}$, associated to,

$$\hat{V}(x) = \hat{W}^{\mathrm{T}}\phi(x),\ \forall x. \tag{27}$$

The approximate solution (27) can be split to obtain the approximate solution of equations (5), (6), and (7), and for that reason we shall use a vector of polynomials $\phi_1 \in \mathbb{R}^{N_1}$, $\phi_2 \in \mathbb{R}^{N_2}$, and $\phi_3 \in \mathbb{R}^{N_3}$ respectively. Hence, these approximations can be expressed as,

$$\hat{V}_1(x_1) = \hat{W}_1^{\mathrm{T}}\phi_1(x_1), \qquad \forall\ x_1, \tag{28}$$

$$\hat{V}_2(x) = \hat{W}_2^{\mathrm{T}}\phi_2(x), \qquad \forall\ x, \tag{29}$$

$$\hat{V}_3(x_2) = \hat{W}_3^{\mathrm{T}}\phi_3(x_2), \qquad \forall\ x_2. \tag{30}$$

Our objective is to find update laws for the weight estimates $\hat{W}_1 \in \mathbb{R}^{N_1}$, $\hat{W}_2 \in \mathbb{R}^{N_2}$, and $\hat{W}_3 \in \mathbb{R}^{N_3}$ where $N_j,\ j = \{1, 2, 3\}$ are the neurons in the hidden layer of each critic approximator. Our objective is for the actual weight estimates to converge to the ideal values in the sense that $\hat{W}_1 \to W_1^*$, $\hat{W}_2 \to W_2^*$, and $\hat{W}_3 \to W_3^*$.

Now, we can write the approximate Hamiltonians (25) with current weight estimates as,

$$\hat{H}_1 \equiv \frac{1}{2}x_1^{\mathrm{T}}Q_1x_1 + \hat{W}_1^{\mathrm{T}}\nabla\phi_1 f_{11}(x_1) + \hat{W}_1^{\mathrm{T}}\nabla\phi_1 g_{11}(x_1)u_{11} + \frac{1}{2}u_{11}^{\mathrm{T}}R_1u_{11}, \forall x_1, u_{11}, \tag{31}$$

$$\begin{aligned}\hat{H}_2 \equiv\ & x_1^{\mathrm{T}}Q_2x_2 + u_{11}^{\mathrm{T}}R_1u_{12} + u_{22}^{\mathrm{T}}R_2u_{21} + \hat{W}_1^{\mathrm{T}}\nabla\phi_1 f_{12}(x) + \hat{W}_3^{\mathrm{T}}\nabla\phi_3 f_{21}(x)\\ & + \hat{W}_1^{\mathrm{T}}\nabla\phi_1 g_{11}(x_1)u_{12} + \hat{W}_1^{\mathrm{T}}\nabla\phi_1 g_{12}(x)u_{22} + \hat{W}_2^{\mathrm{T}}\nabla\phi_{2x_1} g_{11}(x_1)u_{11} + \hat{W}_2^{\mathrm{T}}\nabla\phi_{2x_2} g_{22}(x_2)u_{22}\\ & + \hat{W}_2^{\mathrm{T}}\nabla\phi_{2x_1} f_{11}(x_1) + \hat{W}_2^{\mathrm{T}}\nabla\phi_{2x_2} f_{22}(x_2)\\ & + \hat{W}_3^{\mathrm{T}}\nabla\phi_3 g_{21}(x)u_{11} + \hat{W}_3^{\mathrm{T}}\nabla\phi_3 g_{22}(x_2)u_{21},\ \forall x, u_{11}, u_{12}, u_{21}, u_{22},\end{aligned} \tag{32}$$

$$\hat{H}_3 \equiv \frac{1}{2}x_2^{\mathrm{T}}Q_3x_2 + \hat{W}_3^{\mathrm{T}}\nabla\phi_3 f_{22}(x_2) + \hat{W}_3^{\mathrm{T}}\nabla\phi_3 g_{22}(x_2)u_{22} + \frac{1}{2}u_{22}^{\mathrm{T}}R_2u_{22},\ \forall x_2, u_{22}. \tag{33}$$

DOI: 10.1002/acs

It is obvious that, when we have convergence of the actual weight estimates to the ideal weight values and $u_{11} = u_{11}^*, u_{12} = u_{12}^*, u_{21} = u_{21}^*, u_{22} = u_{22}^*$ then the approximate Hamiltonians also converge to the HJB equations in the sense that $\hat{H}_1 \to H_1^*$, $\hat{H}_2 \to H_2^*$, and $\hat{H}_3 \to H_3^*$, as $t \to \infty$.

Definition 1 ([47])
A vector signal $\Phi(t)$ is *exciting* over the interval $[t, t+T_{\mathrm{PE}}]$, with $T_{\mathrm{PE}} \in \mathbb{R}^+$ if there exists $\beta_1, \beta_2 \in \mathbb{R}^+$ such that $\beta_1 I \leqslant \int_t^{t+T} \Phi(\tau)\Phi^{\mathrm{T}}(\tau) d\tau \leqslant \beta_2 I$ with I an identity matrix of appropriate dimensions. □

To achieve convergence of (31), (32), (33) to the (approximate) Hamiltonian (25) along the closed-loop trajectories, one would typically need persistence of excitation $\forall t \geqslant 0$ (see Definition 1) for the vectors $\omega_1(t), \omega_2(t), \omega_3(t)$ defined by

$$\omega_1 := \nabla\phi_1(f_{11} + g_{11}u_{11}), \tag{34}$$
$$\omega_2 := \nabla\phi_{2x_2}(f_{22} + g_{22}u_{22}) + \nabla\phi_{2x_1}(f_{11} + g_{11}u_{11}), \tag{35}$$
$$\omega_3 := \nabla\phi_3(f_{22} + g_{22}u_{22}). \tag{36}$$

To weaken the need to guarantee a persistence of excitation condition in the sense of Definition 1 for infinite-time, we follow the approach proposed in [48] that uses *past recorded data*, *concurrently with current data*. To this effect, we define the Hamiltonian errors corresponding to the data collected at the *current time* t,

$$e_1(t) := \hat{H}_1 - H_1^* = \hat{H}_1\left(x_1(t), u_{11}(t), \hat{W}_1(t)^{\mathrm{T}}\nabla\phi_1(x_1(t))\right), \tag{37}$$
$$e_2(t) := \hat{H}_2 - H_2^* = \hat{H}_2\Big(x(t), u_{11}(t), u_{22}(t), u_{12}(t), u_{21}(t), \hat{W}_1(t)^{\mathrm{T}}\nabla\phi_1(x_1(t)), \hat{W}_2(t)^{\mathrm{T}}\nabla\phi_2(x(t)), \hat{W}_3(t)^{\mathrm{T}}\nabla\phi_3(x_2(t))\Big), \tag{38}$$
$$e_3(t) := \hat{H}_3 - H_3^* = \hat{H}_3\left(x_2(t), u_{22}(t), \hat{W}_3(t)^{\mathrm{T}}\nabla\phi_3(x_2(t))\right), \tag{39}$$

where the latter equalities in equations (37), (38), and (39) are due to (20), (21), and (22), respectively. Similarly, the errors corresponding to data *previously collected at times* $t_0, t_1, \ldots, t_{k_j} < t, \forall j \in \{1, 2, 3\}$ can be defined as,

$$e_{1\mathrm{buff}_i}(t_i, t) := \hat{H}_1\left(x_1(t_i), u_{11}(t_i), \hat{W}_1(t)^{\mathrm{T}}\nabla\phi_1(x_1(t_i))\right),$$
$$e_{2\mathrm{buff}_i}(t_i, t) := \hat{H}_2\Big(x(t_i), u_{11}(t_i), u_{22}(t_i), u_{12}(t_i), u_{21}(t_i), \hat{W}_1(t)^{\mathrm{T}}\nabla\phi_1(x_1(t_i)), \hat{W}_2(t)^{\mathrm{T}}\nabla\phi_2(x(t_i)), \hat{W}_3(t)^{\mathrm{T}}\nabla\phi_3(x_2(t_i))\Big)$$
$$e_{3\mathrm{buff}_i}(t_i, t) := \hat{H}_3\left(x_2(t_i), u_{22}(t_i), \hat{W}_3(t)^{\mathrm{T}}\nabla\phi_3(x_2(t_i))\right).$$

Note that, while the errors $e_{1\mathrm{buff}_i}(t_i, t)$, $e_{2\mathrm{buff}_i}(t_i, t)$, $e_{3\mathrm{buff}_i}(t_i, t)$ use past *state and input data* $x_1(t_i)$, $x(t_i)$, $x_2(t_i)$ and $u_{11}(t_i)$, $u_{12}(t_i)$, $u_{21}(t_i)$, $u_{22}(t_i)$ respectively, they are defined based on the *current* weight estimates $\hat{W}_1(t)$, $\hat{W}_2(t)$, $\hat{W}_3(t)$.

The current and previous errors defined above can be combined into the following (normalized) quadratic errors,

$$E_1(t) = \frac{1}{2}\left(\frac{e_1^2(t)}{(\omega_1(t)^{\mathrm{T}}\omega_1(t) + 1)^2} + \sum_{i=1}^{k_1} \frac{e_{1\mathrm{buff}_i}^2(t_i, t)}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i) + 1)^2}\right), \forall t$$
$$E_2(t) = \frac{1}{2}\left(\frac{e_2^2(t)}{(\omega_2^{\mathrm{T}}(t)\omega_2(t) + 1)^2} + \sum_{i=1}^{k_2} \frac{e_{2\mathrm{buff}_i}^2(t_i, t)}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i) + 1)^2}\right), \forall t$$
$$E_3(t) = \frac{1}{2}\left(\frac{e_3^2(t)}{(\omega_3^{\mathrm{T}}(t)\omega_3(t) + 1)(t)^2} + \sum_{i=1}^{k_3} \frac{e_{3\mathrm{buff}_i}^2(t_i, t)}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i) + 1)^2}\right), \forall t.$$

The tuning laws for the three critic approximators are obtained by a gradient descent-like rule as follows,

$$\dot{\hat{W}}_1 = -\alpha_1 \frac{\partial E_1}{\partial \hat{W}_1} = -\alpha_1 \frac{\omega_1(t)}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2} e_1(t) - \alpha_1 \sum_{i=1}^{k_1} \frac{\omega_1(t_i)}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2} e_{1\mathrm{buff}_i}(t_i,t), \tag{40}$$

$$\dot{\hat{W}}_2 = -\alpha_2 \frac{\partial E_2}{\partial \hat{W}_2} = -\alpha_2 \frac{\omega_2(t)}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2} e_2(t) - \alpha_2 \sum_{i=1}^{k_2} \frac{\omega_2(t_i)}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2} e_{2\mathrm{buff}_i}(t_i,t), \tag{41}$$

$$\dot{\hat{W}}_3 = -\alpha_3 \frac{\partial E_3}{\partial \hat{W}_3} = -\alpha_3 \frac{\omega_3(t)}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2} e_3(t) - \alpha_3 \sum_{i=1}^{k_3} \frac{\omega_3(t_i)}{(\omega_1(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2} e_{3\mathrm{buff}_i}(t_i,t), \tag{42}$$

$\forall\ t > t_i \geqslant 0$, where $\alpha_1 \in \mathbb{R}^+$, $\alpha_2 \in \mathbb{R}^+$, $\alpha_3 \in \mathbb{R}^+$, are constant gains that determine the speed of convergence.

Now we shall define the weight estimation errors of the critic approximators as,

$$\tilde{W}_1 := W_1^* - \hat{W}_1 \quad \in \mathbb{R}^{N_1} \tag{43}$$

$$\tilde{W}_2 := W_2^* - \hat{W}_2 \quad \in \mathbb{R}^{N_2} \tag{44}$$

$$\tilde{W}_3 := W_3^* - \hat{W}_3 \quad \in \mathbb{R}^{N_3} \tag{45}$$

where $W_i^* \in \mathbb{R}^{N_i},\ i \in \{1,2,3\}$ are the ideal weights, satisfying $\|W_i^*\| \leqslant W_{\mathrm{imax}},\ i \in \{1,2,3\}$ with $W_{\mathrm{imax}} \in \mathbb{R}^+,\ i \in \{1,2,3\}$. Now the weight estimation error dynamics are given as,

$$\begin{aligned}\dot{\tilde{W}}_1 = &-\alpha_1 \left(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2} + \sum_{i=1}^{k_1} \frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2} \right) \tilde{W}_1(t) \\ &+ \alpha_1 \left(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2} \epsilon_{H_1}(t) + \sum_{i=1}^{k_1} \frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2} \epsilon_{H_1}(t_i) \right) \equiv \quad -N_{\mathrm{om1}} + p_{\mathrm{er1}},\end{aligned} \tag{46}$$

$$\begin{aligned}\dot{\tilde{W}}_2 = &-\alpha_2 \left(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2} + \sum_{i=1}^{k_2} \frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2} \right) \tilde{W}_2(t) \\ &+ \alpha_2 \left(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2} \epsilon_{H_2}(t) + \sum_{i=1}^{k_2} \frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2} \epsilon_{H_2}(t_i) \right) \equiv \quad -N_{\mathrm{om2}} + p_{\mathrm{er2}},\end{aligned} \tag{47}$$

$$\begin{aligned}\dot{\tilde{W}}_3 = &-\alpha_3 \left(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2} + \sum_{i=1}^{k_3} \frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2} \right) \tilde{W}_3(t) \\ &+ \alpha_3 \left(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2} \epsilon_{H_3}(t) + \sum_{i=1}^{k_3} \frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2} \epsilon_{H_3}(t_i) \right) \equiv \quad -N_{\mathrm{om3}} + p_{\mathrm{er3}}.\end{aligned} \tag{48}$$

DOI: 10.1002/acs

where $N_{\text{om}i},\ \forall i \in \{1,2,3\}$ are the nominal systems and $p_{\text{er}i},\ \forall i \in \{1,2,3\}$ are the perturbation due to the errors $\epsilon_{H_i},\ \forall i \in \{1,2,3\}$.

Note that, in order to derive the expressions for the components of $\dot{\tilde{W}} = -\dot{\hat{W}}$ we used (40), (41), (42) together with the fact that $\frac{1}{2}x^{\mathrm{T}}Qx + \frac{1}{2}u^{\mathrm{T}}Ru = -W^{*\mathrm{T}}\omega(t) + \epsilon_H(t)$, which is a consequence of (9), (10), (11), and (26).

Theorem 1
Suppose that $\{\omega_j(t_1),\ldots,\omega_j(t_{k_j})\}$ contains $N_j, \forall j \in \{1,2,3\}$ *linearly independent vectors* and that the critic tuning laws are given by (40), (41), (42). Then, for any given control signal $u(t)$ for the nominal systems (i.e. $\epsilon_{H_i} = 0,\ \forall i \in \{1,2,3\}$) we have that,

$$\frac{d}{dt}\left\|\tilde{W}_1(t)\right\|^2 \leqslant -2\alpha_1\lambda_{\min}\left(\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\right)\left\|\tilde{W}_1\right\|^2, \tag{49}$$

$$\frac{d}{dt}\left\|\tilde{W}_2(t)\right\|^2 \leqslant -2\alpha_2\lambda_{\min}\left(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\right)\left\|\tilde{W}_2\right\|^2, \tag{50}$$

$$\frac{d}{dt}\left\|\tilde{W}_3(t)\right\|^2 \leqslant -2\alpha_3\lambda_{\min}\left(\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\right)\left\|\tilde{W}_3\right\|^2, \tag{51}$$

and for bounded $\epsilon_{H_j},\ \forall j \in \{1,2,3\}$, the $\tilde{W}_j,\ \forall j \in \{1,2,3\}$ converge exponentially to the residual sets, $R_{sj} = \left\{\tilde{W}_j \mid \left\|\tilde{W}_j\right\| \leqslant \frac{\left\|\alpha_j\left(\frac{\omega_j(t)\omega_j(t)^{\mathrm{T}}}{(\omega_j(t)^{\mathrm{T}}\omega_j(t)+1)^2}\epsilon_{H_j}(t)+\sum_{i=1}^{k_j}\frac{\omega_j(t_i)\omega_j(t_i)^{\mathrm{T}}}{(\omega_j(t_i)^{\mathrm{T}}\omega_j(t_i)+1)^2}\epsilon_{H_j}(t_i)\right)\right\|}{\lambda_{\min}\left(\sum_{i=1}^{k_j}\frac{\omega_j(t_i)\omega_j(t_i)^{\mathrm{T}}}{(\omega_j(t_i)^{\mathrm{T}}\omega_j(t_i)+1)^2}\right)}\right\}$, $\forall j \in \{1,2,3\}$. □

Remark 2
Ordinary adaptive optimal control algorithms, e.g. [49], do not have the extra past-data term $\sum_{i=1}^{k_j}\frac{\omega_j(t_i)\omega_j(t_i)^{\mathrm{T}}}{(\omega_j(t_i)^{\mathrm{T}}\omega_j(t_i)+1)^2},\ \forall j \in \{1,2,3\}$ in the error dynamics and thus need a persistence of excitation condition on $\frac{\omega_j(t)}{(\omega_j(t)^{\mathrm{T}}\omega_j(t)+1)}$ (typically of the form $\beta_1\mathcal{I} \leqslant \int_t^{t+T}\frac{\omega_j(t)\omega_j(t)^{\mathrm{T}}}{(\omega_j(t)^{\mathrm{T}}\omega_j(t)+1)^2} \leqslant \beta_2\mathcal{I}$ with constants $\beta_1,\ \beta_2,\ T \in \mathbb{R}^+$) that holds for every t from $t=0$ to $t=\infty$. This condition *cannot be verified during learning*. In Theorem 1, the persistence of excitation condition comes through the requirement that at least $N_j, \forall j \in \{1,2,3\}$ of the vectors $\{\omega_j(t_1),\ldots,\omega_j(t_{k_j})\}, \forall j \in \{1,2,3\}$ must be linearly independent, which is equivalent to the matrix $\Lambda_j := \sum_{i=1}^{k_j}\frac{\omega_j(t_i)\omega_j(t_i)^{\mathrm{T}}}{(\omega_j(t_i)^{\mathrm{T}}\omega_j(t_i)+1)^2}, \forall j \in \{1,2,3\}$ being positive definite. In practice, as one collects each additional vector $\omega_j(t_i)$, one adds a new term to the matrix $\sum_{i=1}^{k_j}\frac{\omega_j(t_i)\omega_j(t_i)^{\mathrm{T}}}{(\omega_j(t_i)^{\mathrm{T}}\omega_j(t_i)+1)^2}, \forall j \in \{1,2,3\}$ and one can stop recording points as soon as this matrix becomes full-rank (i.e. $t_{k_j}, \forall j\{1,2,3\}$ time has been reached). From that point forward, one does not need to record new data and the assumption of Theorem 1 holds, *regardless of whether or not future data provides additional excitation*. In spite of the fact that our Theorem, for theoretical purposes requires a very large number of basis sets (i.e. $N_j \to \infty, \forall j \in \{1,2,3\}$) in our numerical simulations it suffices to pick a small number of quadratic or radial basis functions. The selection of the times t_i is somewhat arbitrary, but in our numerical simulations we typically select these values equally spaced in time. □

Remark 3
It is assumed that the maximum number of data points to be stored in the history (i.e., $t_0, t_1, \ldots, t_{k_j} < t, \forall j \in \{1,2,3\}$) is limited due to memory/bandwidth limitations. □

Proof of Theorem 1. See Appendix. ■

3.2. Actor approximators

One could use a single set of weights with a sliding-mode controller as in [51] to approximate both the optimal value functions V_1^*, V_2^*, V_3^* and their gradients ∇V_1^*, ∇V_2^*, ∇V_3^* but instead we independently adjust two sets of weights: the critic weights introduced in (28), (29), (30) to approximate V_1^*, V_2^*, V_3^*, respectively and the actor weights introduced below to approximate u_{11}^*, u_{12}^*, u_{21}^*, u_{22}^* from (16), (17), (18), (19). While this carries additional computational burden, the flexibility introduced by this "over-parameterization" will enable us to establish convergence to the optimal solution and guaranteed Lyapunov-based stability, which seems difficult using only one set of weights.

The optimal control policies (16), (17), (18), (19) can be approximated, respectively, by 4 actor approximators as follows

$$u_{11}^*(x_1) = {W_{u_{11}}^*}^{\mathrm{T}} \phi_{u_{11}}(x_1) + \epsilon_{u_{11}}(x_1), \quad \forall x_1 \tag{52}$$

$$u_{12}^*(x) = {W_{u_{12}}^*}^{\mathrm{T}} \phi_{u_{12}}(x) + \epsilon_{u_{12}}(x), \quad \forall x \tag{53}$$

$$u_{21}^*(x) = {W_{u_{21}}^*}^{\mathrm{T}} \phi_{u_{21}}(x) + \epsilon_{u_{21}}(x), \quad \forall x \tag{54}$$

$$u_{22}^*(x_2) = {W_{u_{22}}^*}^{\mathrm{T}} \phi_{u_{22}}(x_2) + \epsilon_{u_{22}}(x_2), \quad \forall x_2 \tag{55}$$

where $W_{u_{11}}^* \in \mathbb{R}^{N_4 \times m}$, $W_{u_{12}}^* \in \mathbb{R}^{N_5 \times m}$, $W_{u_{21}}^* \in \mathbb{R}^{N_6 \times m}$, and $W_{u_{22}}^* \in \mathbb{R}^{N_7 \times m}$ are the ideal weight matrices, $\phi_{u_{11}(x_1)}$, $\phi_{u_{21}(x)}$, $\phi_{u_{12}(x)}$, and $\phi_{u_{22}(x_2)}$ are the basis functions defined in a similar way than the one used for the critic approximators, N_4, N_5, N_6, N_7 is the number of neurons in the hidden layer of each actor approximator, and $\epsilon_{u_{11}}, \epsilon_{u_{12}}, \epsilon_{u_{21}}, \epsilon_{u_{22}}$ are the four approximation errors. As before, the $u_{11}^*(x_1)$, $u_{12}^*(x)$, $u_{21}^*(x)$, $u_{22}^*(x_2)$ can be uniformly approximated, as expressed by the following assumption. According to Weierstrass higher order approximation theorem [46], a polynomial basis set suffices for proper approximation, and moreover as the number of basis sets N_4, N_5, N_6, N_7 increases, the approximation errors goes to zero, i.e., $\epsilon_{u_{11}} \to 0$, $\epsilon_{u_{12}} \to 0$, $\epsilon_{u_{21}} \to 0$, and $\epsilon_{u_{22}} \to 0$, as $N_4, N_5, N_6, N_7 \to \infty$.

Assumption 3 (Actor Uniform Approximation)
The actor activation functions in $\phi_{u_{11}} \in \mathbb{R}^{N_4}$, $\phi_{u_{12}} \in \mathbb{R}^{N_5}$, $\phi_{u_{21}} \in \mathbb{R}^{N_6}$, $\phi_{u_{22}} \in \mathbb{R}^{N_7}$ and the actor residual errors $\epsilon_{u_{11}}$, $\epsilon_{u_{12}}$, $\epsilon_{u_{21}}$, $\epsilon_{u_{22}}$ are all uniformly bounded on any given compact set Ω, in the sense that there exist known finite constants $\{\phi_{u_{11}\max}, \phi_{u_{12}\max}, \phi_{u_{21}\max}, \phi_{u_{22}\max}\} \in \mathbb{R}^+$ and $\{\epsilon_{u_{11}\max}, \epsilon_{u_{12}\max}, \epsilon_{u_{21}\max}, \epsilon_{u_{22}\max}\} \in \mathbb{R}^+$ such that $|\phi_{u_{11}}(x_1)| \leqslant \phi_{u_{11}\max}$, $|\phi_{u_{12}}(x)| \leqslant \phi_{u_{12}\max}$, $|\phi_{u_{21}}(x)| \leqslant \phi_{u_{21}\max}$, $|\phi_{u_{22}}(x_2)| \leqslant \phi_{u_{22}\max}$, and $|\epsilon_{u_{11}}(x_1)| \leqslant \epsilon_{u_{11}\max}$, $|\epsilon_{u_{12}}(x)| \leqslant \epsilon_{u_{12}\max}$, $|\epsilon_{u_{21}}(x)| \leqslant \epsilon_{u_{21}\max}$, $|\epsilon_{u_{22}}(x_2)| \leqslant \epsilon_{u_{22}\max}$, $\forall x \in \Omega$. □

Since the ideal weights $W_{u_{11}}^* \in \mathbb{R}^{N_4 \times m}$, $W_{u_{12}}^* \in \mathbb{R}^{N_5 \times m}$, $W_{u_{21}}^* \in \mathbb{R}^{N_6 \times m}$, and $W_{u_{22}}^* \in \mathbb{R}^{N_7 \times m}$ are not known, we introduce the *current actor estimate weights* $\hat{W}_{u_{11}}$, $\hat{W}_{u_{12}}$, $\hat{W}_{u_{21}}$, and $\hat{W}_{u_{22}}$ to approximate the optimal controls (52), (53), (54), (55), respectively, by the following actor approximators,

$$\hat{u}_{11}(x_1) = \hat{W}_{u_{11}}^{\mathrm{T}} \phi_{u_{11}}(x_1), \quad \forall x_1, \tag{56}$$

$$\hat{u}_{12}(x) = \hat{W}_{u_{12}}^{\mathrm{T}} \phi_{u_{12}}(x), \quad \forall x, \tag{57}$$

$$\hat{u}_{21}(x) = \hat{W}_{u_{21}}^{\mathrm{T}} \phi_{u_{21}}(x), \quad \forall x, \tag{58}$$

$$\hat{u}_{22}(x_2) = \hat{W}_{u_{22}}^{\mathrm{T}} \phi_{u_{22}}(x_2), \quad \forall x_2. \tag{59}$$

DOI: 10.1002/acs

Our goal is then to appropriately tune $\hat{W}_{u_{11}}$, $\hat{W}_{u_{12}}$, $\hat{W}_{u_{21}}$, $\hat{W}_{u_{22}}$, such that the following quadratic error terms are minimized,

$$E_{u_{11}} = \frac{1}{2}e_{u_{11}}^{\mathrm{T}}(t)e_{u_{11}}(t), \quad \forall t, \tag{60}$$

$$E_{u_{12}} = \frac{1}{2}e_{u_{12}}^{\mathrm{T}}(t)e_{u_{21}}(t), \quad \forall t, \tag{61}$$

$$E_{u_{21}} = \frac{1}{2}e_{u_{21}}^{\mathrm{T}}(t)e_{u_{12}}(t), \quad \forall t, \tag{62}$$

$$E_{u_{22}} = \frac{1}{2}e_{u_{22}}^{\mathrm{T}}(t)e_{u_{22}}(t), \quad \forall t, \tag{63}$$

where,

$$\begin{aligned} e_{u_{11}} &= \hat{W}_{u_{11}}^{\mathrm{T}}\phi_{u_{11}} + R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_1^{\mathrm{T}}\hat{W}_1, \\ e_{u_{12}} &= \hat{W}_{u_{12}}^{\mathrm{T}}\phi_{u_{12}} + R_1^{-1}\left(g_{11}^{\mathrm{T}}(x_1)\nabla\phi_2^{\mathrm{T}}\hat{W}_2 + g_{21}^{\mathrm{T}}(x)\nabla\phi_3^{\mathrm{T}}\hat{W}_3\right), \\ e_{u_{21}} &= \hat{W}_{u_{21}}^{\mathrm{T}}\phi_{u_{21}} + R_2^{-1}\left(g_{22}^{\mathrm{T}}(x_2)\nabla\phi_2^{\mathrm{T}}\hat{W}_2 + g_{12}^{\mathrm{T}}(x)\nabla\phi_1^{\mathrm{T}}\hat{W}_1\right), \\ e_{u_{22}} &= \hat{W}_{u_{22}}^{\mathrm{T}}\phi_{u_{22}} + R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_3^{\mathrm{T}}\hat{W}_3, \end{aligned}$$

are the errors between the estimates (56), (57), (58), (59) and versions of (16), (17), (18), (19), in which V^* is approximated by the estimates of the critic approximators (28), (29), (30).

The tuning laws for the actor approximators are obtained by a gradient descent-like rule as follows

$$\begin{aligned} \dot{\hat{W}}_{u_{11}} &= -\alpha_{u_{11}}\frac{\partial E_{u_{11}}}{\partial \hat{W}_{u_{11}}} = -\alpha_{u_{11}}\phi_{u_{11}}e_{u_{11}} \\ &= -\alpha_{u_{11}}\phi_{u_{11}}\left(\hat{W}_{u_{11}}^{\mathrm{T}}\phi_{u_{11}} + R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_1^{\mathrm{T}}\hat{W}_1\right)^{\mathrm{T}}, \end{aligned} \tag{64}$$

$$\begin{aligned} \dot{\hat{W}}_{u_{12}} &= -\alpha_{u_{12}}\frac{\partial E_{u_{12}}}{\partial \hat{W}_{u_{12}}} = -\alpha_{u_{12}}\phi_{u_{12}}e_{u_{12}} \\ &= -\alpha_{u_{12}}\phi_{u_{12}}\left(\hat{W}_{u_{12}}^{\mathrm{T}}\phi_{u_{12}} + R_1^{-1}\left(g_{11}^{\mathrm{T}}(x_1)\nabla\phi_2^{\mathrm{T}}\hat{W}_2 + g_{21}^{\mathrm{T}}(x)\nabla\phi_3^{\mathrm{T}}\hat{W}_3\right)\right)^{\mathrm{T}}, \end{aligned} \tag{65}$$

$$\begin{aligned} \dot{\hat{W}}_{u_{21}} &= -\alpha_{u_{21}}\frac{\partial E_{u_{21}}}{\partial \hat{W}_{u_{21}}} = -\alpha_{u_{21}}\phi_{u_{21}}e_{u_{21}} \\ &= -\alpha_{u_{21}}\phi_{u_{21}}\left(\hat{W}_{u_{21}}^{\mathrm{T}}\phi_{u_{21}} + R_2^{-1}\left(g_{22}^{\mathrm{T}}(x_2)\nabla\phi_2^{\mathrm{T}}\hat{W}_2 + g_{12}^{\mathrm{T}}(x)\nabla\phi_1^{\mathrm{T}}\hat{W}_1\right)\right)^{\mathrm{T}}, \end{aligned} \tag{66}$$

$$\begin{aligned} \dot{\hat{W}}_{u_{22}} &= -\alpha_{u_{22}}\frac{\partial E_{u_{22}}}{\partial \hat{W}_{u_{22}}} = -\alpha_{u_{22}}\phi_{u_{22}}e_{u_{22}} \\ &= -\alpha_{u_{22}}\phi_{u_{22}}\left(\hat{W}_{u_{22}}^{\mathrm{T}}\phi_{u_{22}} + R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_3^{\mathrm{T}}\hat{W}_3\right)^{\mathrm{T}}, \end{aligned} \tag{67}$$

where $\alpha_{u_{11}} \in \mathbb{R}^+$, $\alpha_{u_{12}} \in \mathbb{R}^+$, $\alpha_{u_{12}} \in \mathbb{R}^+$, and $\alpha_{u_{22}} \in \mathbb{R}^+$ are constant gains that determine the speed of convergence. Defining the weight estimation errors for each one of the actors by

$$\tilde{W}_{u_{11}} = W_{u_{11}}^* - \hat{W}_{u_{11}}, \ \in \mathbb{R}^{N_4 \times m} \tag{68}$$

$$\tilde{W}_{u_{12}} = W_{u_{12}}^* - \hat{W}_{u_{12}}, \ \in \mathbb{R}^{N_5 \times m} \tag{69}$$

$$\tilde{W}_{u_{21}} = W_{u_{21}}^* - \hat{W}_{u_{21}}, \ \in \mathbb{R}^{N_6 \times m} \tag{70}$$

$$\tilde{W}_{u_{22}} = W_{u_{22}}^* - \hat{W}_{u_{22}}. \ \in \mathbb{R}^{N_7 \times m} \tag{71}$$

DOI: 10.1002/acs

and after taking into consideration that (16), (17), (18), (19) with (23) is approximated by (56), (57), (58), (59), respectively, the actor approximators error dynamics can be written as

$$\begin{aligned}\dot{\tilde{W}}_{u_{11}} =& -\alpha_{u_{11}}\phi_{u_{11}}\phi_{u_{11}}^{\mathrm{T}}\tilde{W}_{u_{11}} - \alpha_{u_{11}}\phi_{u_{11}}\left(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_1^{\mathrm{T}}\tilde{W}_1\right)^{\mathrm{T}} \\ &- \alpha_{u_{11}}\phi_{u_{11}}\epsilon_{u_{11}} - \alpha_{u_{11}}\phi_{u_{11}}\left(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\epsilon_1\right)^{\mathrm{T}},\end{aligned} \tag{72}$$

$$\begin{aligned}\dot{\tilde{W}}_{u_{12}} =& -\alpha_{u_{12}}\phi_{u_{12}}\phi_{u_{12}}^{\mathrm{T}}\tilde{W}_{u_{12}} - \alpha_{u_{12}}\phi_{u_{12}}\left(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_2^{\mathrm{T}}\tilde{W}_2 + R_1^{-1}g_{21}^{\mathrm{T}}(x)\nabla\phi_3\tilde{W}_3\right)^{\mathrm{T}} \\ &- \alpha_{u_{12}}\phi_{u_{12}}\epsilon_{u_{12}} - \alpha_{u_{12}}\phi_{u_{12}}\left(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\epsilon_2 + R_1^{-1}g_{21}^{\mathrm{T}}(x)\nabla\epsilon_3\right)^{\mathrm{T}},\end{aligned} \tag{73}$$

$$\begin{aligned}\dot{\tilde{W}}_{u_{21}} =& -\alpha_{u_{21}}\phi_{u_{21}}\phi_{u_{21}}^{\mathrm{T}}\tilde{W}_{u_{21}} - \alpha_{u_{21}}\phi_{u_{21}}\left(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_2^{\mathrm{T}}\tilde{W}_2 + R_2^{-1}g_{12}^{\mathrm{T}}(x)\nabla\phi_1\tilde{W}_1\right)^{\mathrm{T}} \\ &- \alpha_{u_{12}}\phi_{u_{21}}\epsilon_{u_{21}} - \alpha_{u_{21}}\phi_{u_{21}}\left(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\epsilon_2 + R_2^{-1}g_{12}^{\mathrm{T}}(x)\nabla\epsilon_1\right)^{\mathrm{T}},\end{aligned} \tag{74}$$

$$\begin{aligned}\dot{\tilde{W}}_{u_{22}} =& -\alpha_{u_{22}}\phi_{u_{22}}\phi_{u_{22}}^{\mathrm{T}}\tilde{W}_{u_{22}} - \alpha_{u_{22}}\phi_{u_{22}}\left(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_3^{\mathrm{T}}\tilde{W}_3\right)^{\mathrm{T}} \\ &- \alpha_{u_{22}}\phi_{u_{22}}\epsilon_{u_{22}} - \alpha_{u_{22}}\phi_{u_{22}}\left(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\epsilon_3\right)^{\mathrm{T}}.\end{aligned} \tag{75}$$

A pseudocode (with inline comments to provide guidance following after the symbol $\rhd$) that describes the proposed adaptive-optimal control algorithm has the following form,

Algorithm 1: Adaptive-Optimal Control Algorithm for Weakly-Coupled Nonlinear Systems

```
Start with initial state x(0), random initial weights Ŵ_u11(0), Ŵ_u12(0), Ŵ_u21(0), Ŵ_u22(0), Ŵ_1(0), Ŵ_2(0), Ŵ_3(0)
and i = 1
procedure
    Propagate t, x(t) using (1)
    Propagate Ŵ_u11, Ŵ_u12, Ŵ_u21, Ŵ_u22, Ŵ_1, Ŵ_2, Ŵ_3      ▷ {integrate dŴ_u11, dŴ_u12, dŴ_u21, dŴ_u22 as in
(64)-(67) and Ŵ_1, Ŵ_2, Ŵ_3 as in (40)-(42) using any ode solver (e.g. Runge Kutta)}
    Compute V̂_1 = Ŵ_1^T φ_1(x_1), V̂_2 = Ŵ_2^T φ_2(x), V̂_3 = Ŵ_3^T φ_3(x_2)
    Compute û_11 = Ŵ_u11^T φ_u11(x_1), û_12 = Ŵ_u12^T φ_u12(x), û_21 = Ŵ_u21^T φ_u21(x), û_22 = Ŵ_u22^T φ_u22(x_2)
    if i ≠ k then          ▷ {{ω_j(t_1), ω_j(t_2), ..., ω_j(t_i)}, ∀j ∈ {1,2,3} has N_1, N_2 and N_3 linearly
independent elements respectively and t_k is the time instant that this happens}
        Select an arbitrary data point to be included in each history stack (c.f. Remarks 2-3)
        i := i + 1
    end if
end procedure
```

Remark 4
Note that the algorithm runs in real time in a plug-n-play framework and we do not have any iterations within the algorithm. The computational complexity is similar to an adaptive control architecture [47] which increase with the number of the states. □

3.3. *Stability analysis*

The following regularity assumption is needed for the stability analysis presented below.

Assumption 4
The process input functions $g_{11}(.)$, $g_{12}(.)$, $g_{21}(.)$ and $g_{22}(.)$ are uniformly bounded on Ω, i.e., $\sup_{x_1\in\Omega}\|g_{11}(x_1)\| < g_{11\max}$, $\sup_{x\in\Omega}\|g_{12}(x)\| < g_{12\max}$, $\sup_{x\in\Omega}\|g_{21}(x)\| < g_{21\max}$, $\sup_{x_2\in\Omega}\|g_{22}(x_2)\| < g_{22\max}$. □

To remove the effect of the approximation errors ϵ_1, $\epsilon_{u_{11}}$, ϵ_2, $\epsilon_{u_{12}}$, $\epsilon_{u_{21}}$, ϵ_3, $\epsilon_{u_{22}}$, (and their partial derivatives) and obtain a closed-loop system with an asymptotically stable equilibrium point, one

Prepared using acsauth.cls
DOI: 10.1002/acs

needs to add robustifying control terms to (56), (57), (58), and (59), following the work of [52] and write,

$$u_{11}(t) = \hat{W}_{u_{11}}^{\mathrm{T}}\phi_{u_{11}}(x_1) - \frac{x_1^{\mathrm{T}}x_1}{(A + x_1^{\mathrm{T}}x_1)}B_{11}\mathbf{1}_{m_1},\ \forall t, \tag{76}$$

$$u_{12}(t) = \hat{W}_{u_{12}}^{\mathrm{T}}\phi_{u_{12}}(x) - \frac{x^{\mathrm{T}}x}{(A + x^{\mathrm{T}}x)}B_{12}\mathbf{1}_{m_1},\ \forall t, \tag{77}$$

$$u_{21}(t) = \hat{W}_{u_{21}}^{\mathrm{T}}\phi_{u_{21}}(x) - \frac{x^{\mathrm{T}}x}{(A + x^{\mathrm{T}}x)}B_{21}\mathbf{1}_{m_2},\ \forall t, \tag{78}$$

$$u_{22}(t) = \hat{W}_{u_{22}}^{\mathrm{T}}\phi_{u_{22}}(x_2) - \frac{x_2^{\mathrm{T}}x_2}{(A + x_2^{\mathrm{T}}x_2)}B_{22}\mathbf{1}_{m_2},\ \forall t, \tag{79}$$

with $A \in \mathbb{R}^+$, and $B_{11} \in \mathbb{R}^+,\ B_{12} \in \mathbb{R}^+,\ B_{21} \in \mathbb{R}^+,\ B_{22} \in \mathbb{R}^+$ satisfying (80), (81), (82), and (83) respectively,

$$\begin{aligned} B_{11}\|x_1\|^2 \geqslant & \frac{A + \|x_1\|^2}{(W_{1\max}\phi_{1\mathrm{dmax}} + \epsilon_{1\mathrm{dmax}})g_{11\max}}\Bigg\{\frac{1}{4\alpha_1}\left(\left(1 + \lambda_{\max}\left(\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\right)\right)\epsilon_{H1max}\right)^2 \\ & + \frac{\left((\phi_{u_{11}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\epsilon_{1\max} + \phi_{u_{11}\max}\epsilon_{u_{11}\max}) + \phi_{u_{11}\max}\epsilon_{u_{11}\max}\right)^2}{2} \\ & + \frac{g_{11\max}\phi_{u_{11}\max}}{2}(W_{1\max}\phi_{1\mathrm{dmax}} + \epsilon_{1\mathrm{dmax}})^2 + \frac{1}{2}(W_{1\max}\phi_{1\mathrm{dmax}} + \epsilon_{1\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u_{11}\max}^2\Bigg\}, \end{aligned} \tag{80}$$

$$\begin{aligned} B_{12}\|x\|^2 \geqslant & \frac{A + \|x\|^2}{(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}})g_{12\max}}\Bigg\{\frac{1}{4\alpha_2}\left((1 + \lambda_{\max}\left(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\right)\right)\epsilon_{2H_{2\max}}\right)^2 \\ & + \left(\frac{\left(\phi_{u_{12}\max}\left(\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\epsilon_{2\max} + \lambda_{\max}(R_1^{-1})g_{21\max}\nabla\epsilon_{3\max}\right)\right)^2}{2}\right) \\ & + \frac{g_{12\max}\phi_{u_{12}\max}}{2}\left(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}}\right)^2 + \frac{1}{2}(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u_{12}\max}^2\Bigg\}, \end{aligned} \tag{81}$$

$$\begin{aligned} B_{21}\|x\|^2 \geqslant & \frac{A + \|x\|^2}{(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}})g_{21\max}}\Bigg\{\frac{1}{4\alpha_2}\left(\left(1 + \lambda_{\max}\left(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\right)\right)\epsilon_{2H_{2\max}}\right)^2 \\ & + \left(\frac{\left(\phi_{u_{21}\max}\left(\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\epsilon_{2\max} + \lambda_{\max}(R_2^{-1})g_{12\max}\nabla\epsilon_{1\max}\right)\right)^2}{2}\right) \\ & + \frac{g_{21\max}\phi_{u_{21}\max}}{2}\left(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}}\right)^2 + \frac{1}{2}(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u_{21}\max}^2\Bigg\}, \end{aligned} \tag{82}$$

$$\begin{aligned} B_{22}\|x_2\|^2 \geqslant & \frac{A + \|x_2\|^2}{(W_{3\max}\phi_{3\mathrm{dmax}} + \epsilon_{3\mathrm{dmax}})g_{22\max}}\Bigg\{\frac{1}{4\alpha_3}\left(\left(1 + \lambda_{\max}\left(\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\right)\right)\epsilon_{3H_{3\max}}\right)^2 \\ & + \frac{\left((\phi_{u_{22}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\epsilon_{3\max} + \phi_{u_{22}\max}\epsilon_{u_{22}\max}) + \phi_{u_{22}\max}\epsilon_{u_{22}\max}\right)^2}{2} \\ & + \frac{g_{22\max}\phi_{u_{22}\max}}{2}\left(W_{3\max}\phi_{3\mathrm{dmax}} + \epsilon_{3\mathrm{dmax}}\right)^2 + \frac{1}{2}(W_{3\max}\phi_{3\mathrm{dmax}} + \epsilon_{3\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u22\max}^2\Bigg\}. \end{aligned} \tag{83}$$

DOI: 10.1002/acs

The following theorem is the main result of the paper and proves asymptotic stability of the equilibrium point of the closed-loop system dynamics (1), (76)-(79). The closed-loop systems dynamics can be written as,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} f_{11}(x_1) + \varepsilon f_{12}(x) \\ \varepsilon f_{21}(x) + f_{22}(x_2) \end{bmatrix} + \begin{bmatrix} g_{11}(x) & \varepsilon g_{12}(x) \\ \varepsilon g_{21}(x) & g_{22}(x) \end{bmatrix} \left(\begin{bmatrix} \hat{W}_{u_{11}}^{\mathrm{T}} \phi_{u_{11}} + \varepsilon \hat{W}_{u_{12}}^{\mathrm{T}} \phi_{u_{12}} \\ \varepsilon \hat{W}_{u_{21}}^{\mathrm{T}} \phi_{u_{21}} + \hat{W}_{u_{22}}^{\mathrm{T}} \phi_{u_{22}} \end{bmatrix} \right.$$
$$\left. - \begin{bmatrix} \frac{x_1^{\mathrm{T}} x_1}{\left(A + x_1^{\mathrm{T}} x_1\right)} B_{11} \mathbf{1}_{m_1} + \varepsilon \frac{x^{\mathrm{T}} x}{\left(A + x^{\mathrm{T}} x\right)} B_{12} \mathbf{1}_{m_1} \\ \varepsilon \frac{x^{\mathrm{T}} x}{\left(A + x^{\mathrm{T}} x\right)} B_{21} \mathbf{1}_{m_2} + \frac{x_2^{\mathrm{T}} x_2}{\left(A + x_2^{\mathrm{T}} x_2\right)} B_{22} \mathbf{1}_{m_2} \end{bmatrix} \right). \tag{84}$$

Theorem 2

Consider the closed-loop dynamics given by (84) together with the tuning laws for the critic and the actor approximators given by (40)-(42) and (64)-(67), respectively. Suppose that the HJB equations (20)-(22) have a positive definite, smooth solution, the Assumptions 1, 2, 3, and 4 hold, and that $\{\omega_1(t_1), \omega_1(t_2), \ldots, \omega_1(t_{k_1})\}$, $\{\omega_2(t_1), \omega_2(t_2), \ldots, \omega_2(t_{k_2})\}$, and $\{\omega_3(t_1), \omega_3(t_2), \ldots, \omega_3(t_{k_3})\}$ have N_1, N_2, and N_3 *linearly independent elements* respectively. Then, there exists a triple $\Omega_x \times \Omega_W \times \Omega_{W_u} \subset \Omega$, with Ω compact such that the solution $\tilde{Z} := \begin{bmatrix} x(t)^{\mathrm{T}} & \tilde{W}_1(t)^{\mathrm{T}} & \tilde{W}_2(t)^{\mathrm{T}} & \tilde{W}_3(t)^{\mathrm{T}} & \tilde{W}_{u_{11}}(t)^{\mathrm{T}} & \tilde{W}_{u_{12}}(t)^{\mathrm{T}} & \tilde{W}_{u_{21}}(t)^{\mathrm{T}} & \tilde{W}_{u_{22}}(t)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in (\Omega_x \times \Omega_W \times \Omega_{W_u})$ converges *asymptotically* to zero for all initial approximator weights $(\tilde{W}_1(0), \tilde{W}_2(0), \tilde{W}_3(0))$ inside Ω_W, $\big(\tilde{W}_{u_{11}}(0), \tilde{W}_{u_{12}}(0), \tilde{W}_{u_{21}}(0), \tilde{W}_{u_{22}}(0)\big)$ inside Ω_{W_u} and state $x(0)$ inside Ω_x, provided that the following inequalities are satisfied,

$$\frac{1}{\alpha_1}\left(8\alpha_1^2 \lambda_{\min}\left(\sum_{i=1}^{k_1} \frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\right) - 1\right)$$
$$> \phi_{u_{11}\max} \lambda_{\max}(R_1^{-1}) g_{11\max} \nabla\phi_{1\max} + \phi_{u_{21}\max} \lambda_{\max}(R_2^{-1}) g_{12\max} \nabla\phi_{1\max}, \tag{85}$$

$$\frac{1}{\phi_{\mathrm{u}_{11}\max}}\left(2\phi_{\mathrm{u}_{11}\max}^2 - 1\right) > \lambda_{\max}(R_1^{-1}) g_{11\max} \nabla\phi_{1\max} + g_{11\max}, \tag{86}$$

$$\frac{1}{\alpha_2}\left(8\alpha_2^2 \lambda_{\min}\left(\sum_{i=1}^{k_2} \frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\right) - 1\right)$$
$$> \phi_{u_{12}\max} \lambda_{\max}(R_1^{-1}) g_{11\max} \nabla\phi_{2\max} + \phi_{u_{21}\max} \lambda_{\max}(R_2^{-1}) g_{22\max} \nabla\phi_{2\max}, \tag{87}$$

$$\frac{1}{\phi_{\mathrm{u}_{12}\max}}\left(2\phi_{\mathrm{u}_{12}\max}^2 - 1\right) > \lambda_{\max}(R_1^{-1}) g_{11\max} \nabla\phi_{2\max} + \lambda_{\max}(R_1^{-1}) g_{21\max} \nabla\phi_{3\max} + g_{22\max}, \tag{88}$$

$$\frac{1}{\phi_{\mathrm{u}_{21}\max}}\left(2\phi_{\mathrm{u}_{21}\max}^2 - 1\right) > \lambda_{\max}(R_2^{-1}) g_{22\max} \nabla\phi_{2\max} + \lambda_{\max}(R_2^{-1}) g_{12\max} \nabla\phi_{1\max} + g_{21\max}, \tag{89}$$

$$\frac{1}{\alpha_3}\left(8\alpha_3^2 \lambda_{\min}\left(\sum_{i=1}^{k_3} \frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\right) - 1\right)$$
$$> \phi_{u_{22}\max} \lambda_{\max}(R_2^{-1}) g_{22\max} \nabla\phi_{3\max} + \phi_{u_{12}\max} \lambda_{\max}(R_1^{-1}) g_{21\max} \nabla\phi_{3\max}, \tag{90}$$

$$\frac{1}{\phi_{\mathrm{u}_{22}\max}}\left(2\phi_{\mathrm{u}_{22}\max}^2 - 1\right) > \lambda_{\max}(R_2^{-1}) g_{22\max} \nabla\phi_{3\max} + g_{22\max}. \tag{91}$$

When the set Ω that appears in the Assumptions 2, 3, and 4 is the whole $\mathbb{R}^n$, then the triple $\Omega_W \times \Omega_{W_u} \times \Omega_x$ can also be the whole $\mathbb{R}^n$. □

DOI: 10.1002/acs

Proof. See Appendix. ■

Remark 5
For the inequalities (85), (87), and (90) to hold, one needs to pick the tuning gains α_1, α_2, α_3 for the critic approximator sufficiently large since the left hand side of these inequalities are monotonically increasing to $+\infty$ on α_1, α_2 and α_3 respectively. But as noted in adaptive control [47], large adaptive gains can cause high frequency oscillations in the control signal and reduced tolerance to time delays that will destabilize the system. Regarding (86), (88), (89) and (91), since ϕ_{u11max}, ϕ_{u12max}, ϕ_{u21max}, ϕ_{u22max} are simply the upper bounds that appear in Assumption 3, one can select them as large as needed since the left hand side of these inequalities are monotonically increasing to $+\infty$ on ϕ_{u11max}, ϕ_{u12max}, ϕ_{u21max} and ϕ_{u22max} respectively. However, one must keep in mind that large values for these upper bounds, require an appropriate large value for the functions B_{11}, B_{12}, B_{21}, B_{22} in the robustness terms in (76)-(79). It is possible to pick B_{11}, B_{12}, B_{21}, B_{22} high enough to ensure the convergence of the state to an arbitrarily small neighborhood of the equilibrium point. Choosing an increasing or time-varying robustifying term, can lead to asymptotic stability provided that the inequalities (80)-(83) hold $\forall x$. □

Remark 6
From the conclusion of Theorem 2, we shall have that $\left\|\tilde{Z}\right\| \to 0$ which implies $\|x\| \to 0$, it is straightforward that as $t \to \infty$ then from (76)-(79) we have (56)-(59) which are $\epsilon_{u_{11}}, \epsilon_{u_{12}}, \epsilon_{u_{21}}, \epsilon_{u_{22}}$ respectively away from the optimal. □

Remark 7
In order to get ϵ small we assume that we have a large number of basis sets, i.e. $N_1 \to \infty$, $N_2 \to \infty$ and $N_3 \to \infty$. Moreover, in order to get $\epsilon_{u_{11}}$, $\epsilon_{u_{12}}$, $\epsilon_{u_{21}}$, $\epsilon_{u_{22}}$ small we also assume that we have a large number of basis sets, i.e. $N_4 \to \infty$, $N_5 \to \infty$, $N_6 \to \infty$, $N_7 \to \infty$. But note that this is a requirement for theoretical purposes. We have observed in our numerical and simulation examples that picking quadratic basis function can achieve the required result. □

Remark 8
In case the approximation holds over the entire space, i.e. $\Omega \equiv \mathbb{R}^n$, one can conclude global existence of solution provided that the HJB solution V^* is norm coercive (i.e., $V^* \to 0 \quad \Rightarrow \quad x \to 0$), as this suffices to guarantee that the Lyapunov function $\mathcal{V}$ that we use in the proof of Theorem 1 is also norm coercive (see [28]). □

4. NUMERICAL EXAMPLE

This section presents a sixth-order numerical example to illustrate the effectiveness of the proposed optimal adaptive control algorithm for weakly coupled nonlinear systems like the one described by equation (1). The state variables are taken as, $x_1 = [x_{11}\ x_{12}\ x_{13}]^{\mathrm{T}}$ and $x_2 = [x_{21}\ x_{22}\ x_{23}]^{\mathrm{T}}$. The small perturbation parameter is chosen as, $\varepsilon = 0.1$. The matrices of the system under consideration are chosen as

$$f_{11}(x_1) = \begin{bmatrix} -0.001x_{11}^2 \\ -x_{12}x_{11} \\ -x_{13} \end{bmatrix}, \qquad f_{12}(x) = \begin{bmatrix} 0.1x_{23}x_{21}x_{13} \\ -3.26x_{11} \\ -0.25x_{23}^2 \end{bmatrix},$$

$$f_{21}(x) = \begin{bmatrix} -1.3x_{12}^2 \\ 0.95x_{11}x_{21} - 1.03x_{12}x_{22} \\ -2.1x_{13} \end{bmatrix}, \qquad f_{22}(x_2) = \begin{bmatrix} 0 \\ 0.413x_{21} - 0.426x_{22} \\ -0.09x_{23} \end{bmatrix},$$

DOI: 10.1002/acs

$$g_{11}(x_1) = \begin{bmatrix} -4.274x_{11}^2x_{13} \\ -0.01x_{12} \\ 0 \end{bmatrix}, \qquad g_{12}(x) = \begin{bmatrix} 0.1x_{12} \\ -4.5x_{22} \\ x_{13} \end{bmatrix},$$

$$g_{21}(x) = \begin{bmatrix} 0.75x_{11} \\ x_{21}x_{13} \\ 0 \end{bmatrix}, \qquad g_{22}(x_2) = \begin{bmatrix} -0.718x_{21} \\ -0.1x_{23}x_{21} \\ 0.0718x_{23}x_{22}^2 \end{bmatrix}.$$

We consider a cost defined by (2), and with the user-defined matrices Q and R to be identity matrices of appropriate dimensions.

The weights are initialized randomly in the interval $[0,1]$, the critic activation functions were chosen to be quadratic of the form $\phi_1 = [x_{11}^2 \;\; x_{11}x_{12} \;\; x_{11}x_{13} \;\; x_{12}^2 \;\; x_{12}x_{13} \;\; x_{13}^2]$, $\phi_2 = [x_{11}^2 \;\; x_{11}x_{12} \;\; x_{11}x_{13} \;\; x_{11}x_{21} \;\; x_{11}x_{22} \;\; x_{11}x_{23} \;\; x_{12}^2 \;\; x_{12}x_{13} \;\; x_{12}x_{21} \;\; x_{12}x_{22} \;\; x_{12}x_{23} \;\; x_{13}^2 \;\; x_{13}x_{21} \;\; x_{13}x_{22} \;\; x_{13}x_{23} \;\; x_{21}^2 \;\; x_{21}x_{22} \;\; x_{21}x_{23} \;\; x_{22}^2 \;\; x_{22}x_{23} \;\; x_{23}^2]$, $\phi_3 = [x_{21}^2 \; x_{21}x_{22} \; x_{21}x_{23} \; x_{22}^2 \; x_{22}x_{23} \; x_{23}^2]$, and the actor activation functions are picked in a similar way. The initial states are chosen as $x(0) = [3 \;\; -1 \;\; 4.3 \;\; 1.2 \;\; -1.5 \;\; -1]^{\mathrm{T}}$, and the tuning gains were set to $\alpha_1 = \alpha_2 = \alpha_3 = 10$, and $\alpha_{u_{11}} = \alpha_{u_{12}} = \alpha_{u_{21}} = \alpha_{u_{22}} = 2$.

Figure 1 shows the time evolution of the states in the weakly coupled nonlinear system. The convergence of the critic parameters W_c to the optimal cost (12) is shown in Figure 2. The evolution of the actor parameters W_u is shown in Figure 3. The optimal control inputs, i.e. $u_1 = u_{11} + \varepsilon u_{12}$ and $u_2 = \varepsilon u_{21} + u_{22}$, are shown in Figure 4.

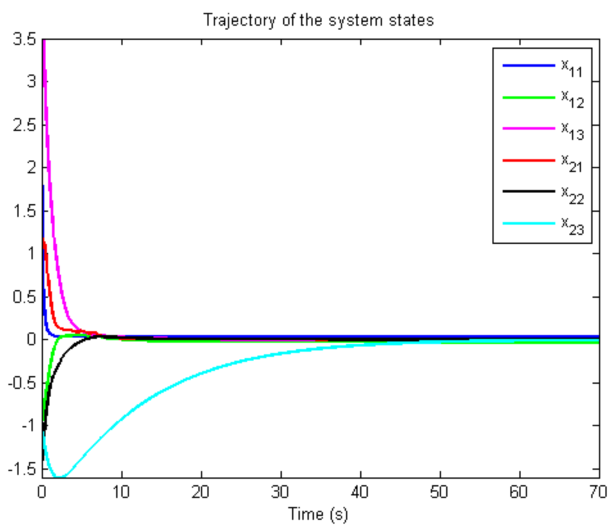

Figure 1. Trajectory of the closed-loop system states.

5. CONCLUSIONS

This paper proposed a new approximate dynamic programming algorithm for controlling weakly coupled nonlinear systems, which also relaxes the persistence of excitation condition by using previously stored data concurrently with current data. The algorithm is implemented as a three-critic/four-actor approximators structure. To suppress the effects of the three critics and four actors approximation errors, robustifying terms have been added to the controllers. We finally prove asymptotic stability of the equilibrium point of the overall closed-loop system. Simulation results illustrate the effectiveness of the proposed approach. Future work will be concentrated on extending the results in completely unknown systems and multiple decision makers.

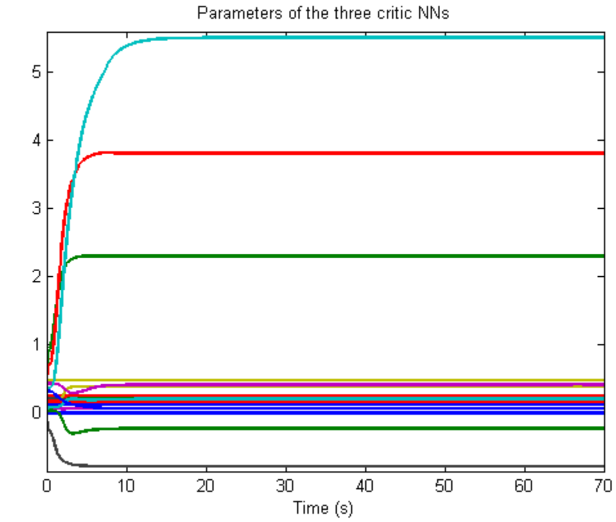

Figure 2. Convergence of the critic parameters.

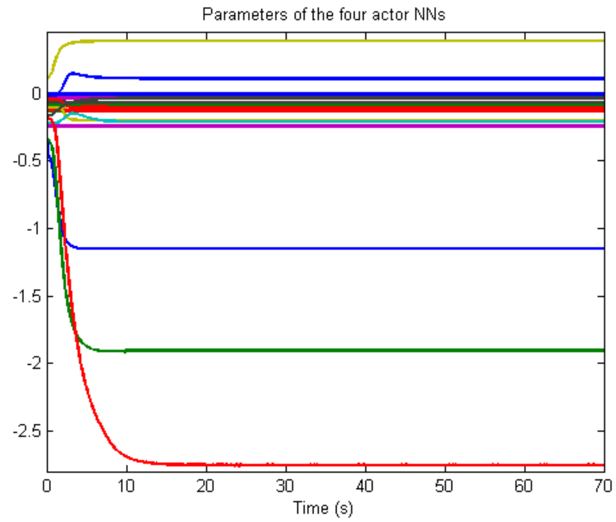

Figure 3. Parameters of the four actors.

APPENDIX

Proof of Theorem 1

Consider the following Lyapunov function, for $t \geqslant 0$

$$\mathcal{L} = \sum_{i=1}^{3} \mathcal{L}_i, \tag{92}$$

DOI: 10.1002/acs

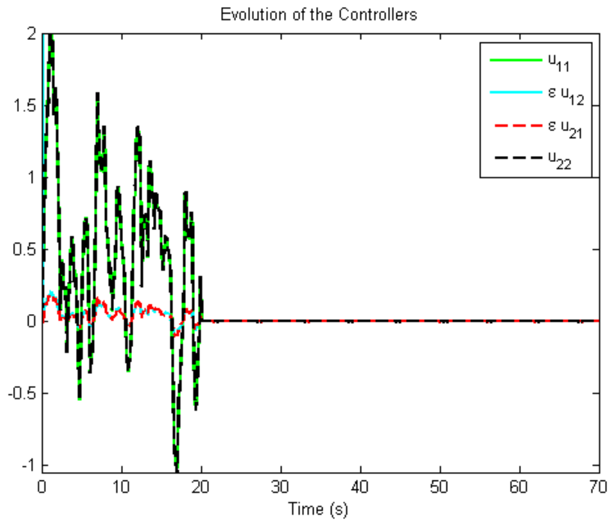

Figure 4. Trajectories for u_{11}, ϵu_{12}, ϵu_{21}, and u_{22}.

where,

$$\mathcal{L}_1 := \frac{1}{2\alpha_1}\tilde{W}_1(t)^{\mathrm{T}}\tilde{W}_1(t), \tag{93}$$

$$\mathcal{L}_2 := \frac{1}{2\alpha_2}\tilde{W}_2(t)^{\mathrm{T}}\tilde{W}_2(t), \tag{94}$$

$$\mathcal{L}_3 := \frac{1}{2\alpha_3}\tilde{W}_3(t)^{\mathrm{T}}\tilde{W}_3(t). \tag{95}$$

By differentiating (92) (i.e. (93), (94), (95)) along the critic error dynamics one has

$$\begin{aligned}\dot{\mathcal{L}}_1 = &-\tilde{W}_1(t)^{\mathrm{T}}\left(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2}+\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\right)\tilde{W}_1(t)\\ &+\tilde{W}_1(t)^{\mathrm{T}}\alpha_1\left(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2}\epsilon_{H_1}(t)+\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\epsilon_{H_1}(t_i)\right),\end{aligned} \tag{96}$$

$$\begin{aligned}\dot{\mathcal{L}}_2 = &-\tilde{W}_2(t)^{\mathrm{T}}\left(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2}+\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\right)\tilde{W}_2(t)\\ &+\tilde{W}_2(t)^{\mathrm{T}}\alpha_2\left(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2}\epsilon_{H_2}(t)+\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\epsilon_{H_2}(t_i)\right),\end{aligned} \tag{97}$$

$$\begin{aligned}\dot{\mathcal{L}}_3 = &-\tilde{W}_3(t)^{\mathrm{T}}\left(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2}+\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\right)\tilde{W}_3(t)\\ &+\tilde{W}_3(t)^{\mathrm{T}}\alpha_3\left(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2}\epsilon_{H_3}(t)+\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\epsilon_{H_3}(t_i)\right).\end{aligned} \tag{98}$$

DOI: 10.1002/acs

Then $\dot{\mathcal{L}}$ is negative definite (see Section 4.9 in [28], where one can prove Input to State Stability (ISS) by treating (46)-(48) as dynamical systems with $\epsilon_{H_j}, j \in \{1,2,3\}$ as input), as long as

$$\|\tilde{W}_j\| > \frac{\left\| \alpha_j \left(\frac{\omega_j(t)\omega_j(t)^{\mathrm{T}}}{(\omega_j(t)^{\mathrm{T}}\omega_j(t)+1)^2}\epsilon_{H_j}(t) + \sum_{i=1}^{k_j} \frac{\omega_j(t_i)\omega_j(t_i)^{\mathrm{T}}}{(\omega_j(t_i)^{\mathrm{T}}\omega_j(t_i)+1)^2}\epsilon_{H_j}(t_i) \right) \right\|}{\lambda_{\min}\left(\sum_{i=1}^{k_j} \frac{\omega_j(t_i)\omega_j(t_i)^{\mathrm{T}}}{(\omega_j(t_i)^{\mathrm{T}}\omega_j(t_i)+1)^2} \right)}, \forall j \in \{1,2,3\}. \tag{99}$$

Equations (49), (50), (51) follow from this and the fact that $\frac{\omega_j(t)\omega_j(t)^{\mathrm{T}}}{(\omega_j(t)^{\mathrm{T}}\omega_j(t)+1)^2} > 0$, $\forall t$ and $\forall j \in \{1,2,3\}$. Since $\{\omega_j(t_1), \ldots, \omega_j(t_{k_j})\}$ has $N_j, \forall j \in \{1,2,3\}$ linearly independent vectors, the matrices Λ_j, $\forall j \in \{1,2,3\}$ are positive definite, from which the exponential stability of the nominal system follows. ■

Proof of Theorem 2

Consider the following Lyapunov function,

$$\mathcal{V} := V^* + V_c + V_u, \tag{100}$$

with

$$\begin{aligned} V^* :=& V_1^* + V_2^* + V_3^*, \\ V_c :=& V_{c1} + V_{c2} + V_{c3} := \tilde{W}_1^{\mathrm{T}}\tilde{W}_1 + \tilde{W}_2^{\mathrm{T}}\tilde{W}_2 + \tilde{W}_3^{\mathrm{T}}\tilde{W}_3, \\ V_u :=& \operatorname{trace}\{\tilde{W}_{u_{11}}^{\mathrm{T}}\tilde{W}_{u_{11}}\} + \operatorname{trace}\{\tilde{W}_{u_{12}}^{\mathrm{T}}\tilde{W}_{u_{12}}\} + \operatorname{trace}\{\tilde{W}_{u_{21}}^{\mathrm{T}}\tilde{W}_{u_{21}}\} + \operatorname{trace}\{\tilde{W}_{u_{22}}^{\mathrm{T}}\tilde{W}_{u_{22}}\}, \end{aligned}$$

where V^*, are the optimal value functions in (12), that is, the positive definite and smooth solution of (20)-(22). Since $\mathcal{V}$ is positive definite, there exist class-$\mathcal{K}$ functions $\gamma_1(.)$ and $\gamma_2(.)$ to write,

$$\gamma_1\left(\left\|\tilde{Z}\right\|\right) \leqslant \mathcal{V} \leqslant \gamma_2\left(\left\|\tilde{Z}\right\|\right),$$

for all $\tilde{Z} \equiv \left[x^{\mathrm{T}}(t) \;\; \tilde{W}_1^{\mathrm{T}}(t) \;\; \tilde{W}_2^{\mathrm{T}}(t) \;\; \tilde{W}_3^{\mathrm{T}}(t) \;\; \tilde{W}_{u_{11}}^{\mathrm{T}}(t) \;\; \tilde{W}_{u_{12}}^{\mathrm{T}}(t) \;\; \tilde{W}_{u_{21}}^{\mathrm{T}}(t) \;\; \tilde{W}_{u_{22}}^{\mathrm{T}}(t) \right]^{\mathrm{T}} \in B_r$ where $B_r \subset \Omega$ is a ball of radius $r \in \mathbb{R}^+$. By taking the time derivative of the first term with respect to the state trajectories with $u(t)$ (see (84)), and the second term with respect to the perturbed critic estimation error dynamics (46), (47), (48), using (49), (50), (51), substituting the update for the actors (64), (65), (66), (67) and grouping terms together, then (100) becomes equation (101).

$$
\begin{aligned}
\dot{\mathcal{V}} =& V_{1x_1}{}^{*\mathrm{T}}\big(f_{11}(x_1) - g_{11}(x_1)\tilde{W}_{u_{11}}^{\mathrm{T}}\phi_{u_{11}} + g_{11}(x_1)(u_{11}^* - \epsilon_{u_{11}}) - g_{11}(x_1)B_{11}\frac{x_1^{\mathrm{T}}x_1\mathbf{1}_{m_1}}{(A + x_1^{\mathrm{T}}x_1)}\big) \\
&+ V_{2x_1}{}^{*\mathrm{T}}\big(f_{12}(x) - g_{12}(x)\tilde{W}_{u_{12}}^{\mathrm{T}}\phi_{u_{12}} + g_{12}(x)(u_{12}^* - \epsilon_{u_{12}}) - g_{12}(x)B_{12}\frac{x^{\mathrm{T}}x\mathbf{1}_{m_1}}{(A + x^{\mathrm{T}}x)}\big) \\
&+ V_{2x_2}{}^{*\mathrm{T}}\big(f_{21}(x) - g_{21}(x)\tilde{W}_{u_{21}}^{\mathrm{T}}\phi_{u_{21}} + g_{21}(x)(u_{21}^* - \epsilon_{u_{21}}) - g_{21}(x)B_{21}\frac{x^{\mathrm{T}}x\mathbf{1}_{m_2}}{(A + x^{\mathrm{T}}x)}\big) \\
&+ V_{3x_2}{}^{*\mathrm{T}}\big(f_{22}(x_2) - g_{22}(x_2)\tilde{W}_{u_{22}}^{\mathrm{T}}\phi_{u_{22}} + g_{22}(x_2)(u_{22}^* - \epsilon_{u_{22}}) - g_{22}(x_2)B_{22}\frac{x_2^{\mathrm{T}}x_2\mathbf{1}_{m_2}}{(A + x_2^{\mathrm{T}}x_2)}\big) \\
&- \frac{\partial V_{c1}}{\partial \tilde{W}_1}^{\mathrm{T}}\left(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2} + \sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\right)\tilde{W}_1 \\
&+ \frac{\partial V_{c1}}{\partial \tilde{W}_1}^{\mathrm{T}}\left(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2}\epsilon_{H_1}(t) + \sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\epsilon_{H_1}(t_i)\right) \\
&- \frac{\partial V_{c2}}{\partial \tilde{W}_2}^{\mathrm{T}}\left(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2} + \sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\right)\tilde{W}_2 \\
&+ \frac{\partial V_{c2}}{\partial \tilde{W}_2}^{\mathrm{T}}\left(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2}\epsilon_{H_2}(t) + \sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\epsilon_{H_2}(t_i)\right) \\
&- \frac{\partial V_{c3}}{\partial \tilde{W}_3}^{\mathrm{T}}\left(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2} + \sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\right)\tilde{W}_3 \\
&+ \frac{\partial V_{c3}}{\partial \tilde{W}_3}^{\mathrm{T}}\left(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2}\epsilon_{H_3}(t) + \sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\epsilon_{H_3}(t_i)\right) \\
&+ \mathrm{trace}\{\tilde{W}_{u_{11}}^{\mathrm{T}}\big(-\phi_{u_{11}}\phi_{u_{11}}^{\mathrm{T}}\tilde{W}_{u_{11}} - \phi_{u_{11}}\big(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_1^{\mathrm{T}}\tilde{W}_1\big)^{\mathrm{T}} - \phi_{u_{11}}\epsilon_{u_{11}} \\
&\qquad - \phi_{u_{11}}\big(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\epsilon_1\big)^{\mathrm{T}}\big)\} \\
&+ \mathrm{trace}\{\tilde{W}_{u_{12}}^{\mathrm{T}}\big(-\alpha_{u_{12}}\phi_{u_{12}}\phi_{u_{12}}^{\mathrm{T}}\tilde{W}_{u_{12}} - \alpha_{u_{12}}\phi_{u_{12}}\big(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_2^{\mathrm{T}}\tilde{W}_2 + R_1^{-1}g_{21}^{\mathrm{T}}(x)\nabla\phi_3\tilde{W}_3\big)^{\mathrm{T}} \\
&\qquad\qquad - \alpha_{u_{12}}\phi_{u_{12}}\epsilon_{u_{12}} - \alpha_{u_{12}}\phi_{u_{12}}\big(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\epsilon_2 + R_1^{-1}g_{21}^{\mathrm{T}}(x)\nabla\epsilon_3\big)^{\mathrm{T}}\big)\} \\
&+ \mathrm{trace}\{\tilde{W}_{u_{21}}^{\mathrm{T}}\big(-\alpha_{u_{21}}\phi_{u_{21}}\phi_{u_{21}}^{\mathrm{T}}\tilde{W}_{u_{21}} - \alpha_{u_{21}}\phi_{u_{21}}\big(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_2^{\mathrm{T}}\tilde{W}_2 + R_2^{-1}g_{12}^{\mathrm{T}}(x)\nabla\phi_1\tilde{W}_1\big)^{\mathrm{T}} \\
&\qquad\qquad - \alpha_{u_{12}}\phi_{u_{21}}\epsilon_{u_{21}} - \alpha_{u_{21}}\phi_{u_{21}}\big(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\epsilon_2 + R_2^{-1}g_{12}^{\mathrm{T}}(x)\nabla\epsilon_1\big)^{\mathrm{T}}\big)\} \\
&+ \mathrm{trace}\{\tilde{W}_{u_{22}}^{\mathrm{T}}\big(-\phi_{u_{22}}\phi_{u_{22}}^{\mathrm{T}}\tilde{W}_{u_{22}} - \phi_{u_{22}}\big(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_3^{\mathrm{T}}\tilde{W}_3\big)^{\mathrm{T}} - \phi_{u_{22}}\epsilon_{u_{22}} \\
&\qquad - \phi_{u_{22}}\big(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\epsilon_3\big)^{\mathrm{T}}\big)\},\ t \geqslant 0, \\
=& T_1 + T_2 + T_3.
\end{aligned} \tag{101}
$$

In equation (101) the three terms T_1, T_2, and T_3 are given by (102), (103), and (104) respectively. Specifically, for the term T_1 one has

$$
\begin{aligned}
T_1 := &-\frac{\partial V_{c1}}{\partial \tilde{W}_1}^{\mathrm{T}}\Big(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2}+\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\Big)\tilde{W}_1 \\
&+\frac{\partial V_{c1}}{\partial \tilde{W}_1}^{\mathrm{T}}\Big(\frac{\omega_1(t)\omega_1(t)^{\mathrm{T}}}{(\omega_1(t)^{\mathrm{T}}\omega_1(t)+1)^2}\epsilon_{H_1}(t)+\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\epsilon_{H_1}(t_i)\Big) \\
&-\frac{\partial V_{c2}}{\partial \tilde{W}_2}^{\mathrm{T}}\Big(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2}+\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\Big)\tilde{W}_2 \\
&+\frac{\partial V_{c2}}{\partial \tilde{W}_2}^{\mathrm{T}}\Big(\frac{\omega_2(t)\omega_2(t)^{\mathrm{T}}}{(\omega_2(t)^{\mathrm{T}}\omega_2(t)+1)^2}\epsilon_{H_2}(t)+\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\epsilon_{H_2}(t_i)\Big) \\
&-\frac{\partial V_{c3}}{\partial \tilde{W}_3}^{\mathrm{T}}\Big(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2}+\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\Big)\tilde{W}_3 \\
&+\frac{\partial V_{c3}}{\partial \tilde{W}_3}^{\mathrm{T}}\Big(\frac{\omega_3(t)\omega_3(t)^{\mathrm{T}}}{(\omega_3(t)^{\mathrm{T}}\omega_3(t)+1)^2}\epsilon_{H_3}(t)+\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\epsilon_{H_3}(t_i)\Big)
\end{aligned}
$$

then,

$$
\begin{aligned}
T_1 \leqslant &-2\alpha_1\lambda_{1\min}\big(\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\big)\left\|\tilde{W}_1\right\|^2 \\
&+\frac{1}{2\alpha_1}\left\|\tilde{W}_1\right\|\Big((1+\lambda_{1\max}\big(\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\big))\epsilon_{H1max}\Big) \\
&-2\alpha_2\lambda_{2\min}\big(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\big)\left\|\tilde{W}_2\right\|^2 \\
&+\frac{1}{2\alpha_2}\left\|\tilde{W}_2\right\|\Big((1+\lambda_{2\max}\big(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\big))\epsilon_{H2max}\Big) \\
&-2\alpha_3\lambda_{3\min}\big(\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\big)\left\|\tilde{W}_3\right\|^2 \\
&+\frac{1}{2\alpha_3}\left\|\tilde{W}_3\right\|\Big((1+\lambda_{3\max}\big(\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\big))\epsilon_{H3max}\Big)
\end{aligned}
$$

DOI: 10.1002/acs

finally one has,

$$\begin{aligned}T_1 \leqslant & -2\alpha_1\lambda_{1\min}\Big(\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\Big)\left\|\tilde{W}_1\right\|^2+\frac{1}{4\alpha_1}\left\|\tilde{W}_1\right\|^2 \\ & +\frac{1}{4\alpha_1}\left((1+\lambda_{1\max}\Big(\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\Big))\epsilon_{H1max}\right)^2 \\ & -2\alpha_2\lambda_{2\min}\Big(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\Big)\left\|\tilde{W}_2\right\|^2+\frac{1}{4\alpha_2}\left\|\tilde{W}_2\right\|^2 \\ & +\frac{1}{4\alpha_2}\left((1+\lambda_{2\max}\Big(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\Big))\epsilon_{H2max}\right)^2 \\ & -2\alpha_3\lambda_{3\min}\Big(\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\Big)\left\|\tilde{W}_3\right\|^2+\frac{1}{4\alpha_3}\left\|\tilde{W}_3\right\|^2 \\ & +\frac{1}{4\alpha_3}\left((1+\lambda_{3\max}\Big(\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\Big))\epsilon_{H3max}\right)^2 \end{aligned} \tag{102}$$

For the term T_2 one has

$$\begin{aligned}T_2 :=\ & \mathrm{trace}\{\tilde{W}_{u_{11}}^{\mathrm{T}}\big(-\phi_{u_{11}}\phi_{u_{11}}^{\mathrm{T}}\tilde{W}_{u_{11}}-\phi_{u_{11}}\left(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_1^{\mathrm{T}}\tilde{W}_1\right)^{\mathrm{T}} \\ & -\phi_{u_{11}}\epsilon_{u_{11}}-\phi_{u_{11}}\left(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\epsilon_1\right)^{\mathrm{T}}\big)\} \\ & +\mathrm{trace}\{\tilde{W}_{u_{12}}^{\mathrm{T}}\big(-\phi_{u_{12}}\phi_{u_{12}}^{\mathrm{T}}\tilde{W}_{u_{12}}-\phi_{u_{12}}\big(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\phi_2^{\mathrm{T}}\tilde{W}_2+R_1^{-1}g_{21}^{\mathrm{T}}(x)\nabla\phi_3\tilde{W}_3\big)^{\mathrm{T}} \\ & \quad -\phi_{u_{12}}\epsilon_{u_{12}}-\phi_{u_{12}}\left(R_1^{-1}g_{11}^{\mathrm{T}}(x_1)\nabla\epsilon_2+R_1^{-1}g_{21}^{\mathrm{T}}(x)\nabla\epsilon_3\right)^{\mathrm{T}}\big)\} \\ & +\mathrm{trace}\{\tilde{W}_{u_{21}}^{\mathrm{T}}\big(-\phi_{u_{21}}\phi_{u_{21}}^{\mathrm{T}}\tilde{W}_{u_{21}}-\phi_{u_{21}}\big(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_2^{\mathrm{T}}\tilde{W}_2+R_2^{-1}g_{12}^{\mathrm{T}}(x)\nabla\phi_1\tilde{W}_1\big)^{\mathrm{T}} \\ & \quad -\phi_{u_{21}}\epsilon_{u_{21}}-\phi_{u_{21}}\left(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\epsilon_2+R_2^{-1}g_{12}^{\mathrm{T}}(x)\nabla\epsilon_1\right)^{\mathrm{T}}\big)\} \\ & +\mathrm{trace}\{\tilde{W}_{u_{22}}^{\mathrm{T}}\big(-\phi_{u_{22}}\phi_{u_{22}}^{\mathrm{T}}\tilde{W}_{u_{22}}-\phi_{u_{22}}\left(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\phi_3^{\mathrm{T}}\tilde{W}_3\right)^{\mathrm{T}} \\ & -\phi_{u_{22}}\epsilon_{u_{22}}-\phi_{u_{22}}\left(R_2^{-1}g_{22}^{\mathrm{T}}(x_2)\nabla\epsilon_3\right)^{\mathrm{T}}\big)\}\end{aligned}$$

then,

$$\begin{aligned}T_2 \leqslant & -\phi_{\mathrm{u}_{11}\max}^2\left\|\tilde{W}_{u_{11}}\right\|^2-\phi_{\mathrm{u}_{11}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{1\max}\left\|\tilde{W}_1\right\|\left\|\tilde{W}_{u_{11}}\right\| \\ & \qquad -\left(\phi_{u_{11}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\epsilon_{1\max}+\phi_{u_{11}\max}\epsilon_{u_{11}\max}\right)\left\|\tilde{W}_{u_{11}}\right\| \\ & -\phi_{\mathrm{u}_{12}\max}^2\left\|\tilde{W}_{u_{12}}\right\|^2-\phi_{\mathrm{u}_{12}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{2\max}\left\|\tilde{W}_2\right\|\left\|\tilde{W}_{u_{12}}\right\| \\ & -\phi_{\mathrm{u}_{12}\max}\lambda(R_1^{-1})g_{21\max}\nabla\phi_{3\max}\left\|\tilde{W}_3\right\|\left\|\tilde{W}_{u_{12}}\right\| \\ & \qquad -\left(\phi_{u_{12}\max}\left(\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\epsilon_{2\max}+\lambda_{\max}(R_1^{-1})g_{21\max}\nabla\epsilon_{3\max}\right)^{\mathrm{T}}\left\|\tilde{W}_{u_{12}}\right\|\right. \\ & -\phi_{\mathrm{u}_{21}\max}^2\left\|\tilde{W}_{u_{21}}\right\|^2-\phi_{\mathrm{u}_{21}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{2\max}\left\|\tilde{W}_2\right\|\left\|\tilde{W}_{u_{21}}\right\| \\ & -\phi_{\mathrm{u}_{21}\max}\lambda_{\max}(R_2^{-1})g_{12\max}\nabla\phi_{1\max}\left\|\tilde{W}_1\right\|\left\|\tilde{W}_{u_{21}}\right\| \\ & \qquad -\left(\phi_{u_{21}\max}\left(\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\epsilon_{2\max}+\lambda_{\max}(R_2^{-1})g_{12\max}\nabla\epsilon_{1\max}\right)^{\mathrm{T}}\left\|\tilde{W}_{u_{21}}\right\|\right. \\ & -\phi_{\mathrm{u}_{22}\max}^2\left\|\tilde{W}_{u_{22}}\right\|^2-\phi_{\mathrm{u}_{22}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{3\max}\left\|\tilde{W}_3\right\|\left\|\tilde{W}_{u_{22}}\right\| \\ & \qquad -\left(\phi_{u_{22}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\epsilon_{3\max}+\phi_{u_{22}\max}\epsilon_{u_{22}\max}\right)\left\|\tilde{W}_{u_{22}}\right\|\end{aligned}$$

finally one has

$$
\begin{aligned}
T_2 \leqslant &-\phi_{u_{11}\max}^2 \left\|\tilde{W}_{u_{11}}\right\|^2 + \phi_{\mathrm{u}_{11}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{1\max}\left(\frac{\left\|\tilde{W}_1\right\|^2}{2}+\frac{\left\|\tilde{W}_{u_{11}}\right\|^2}{2}\right)\\
&+\frac{\left((\phi_{u_{11}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\epsilon_{1\max}+\phi_{u_{11}\max}\epsilon_{u_{11}\max})+\phi_{u_{11}\max}\epsilon_{u_{11}\max}\right)^2}{2}+\left\|\tilde{W}_{u_{11}}\right\|^2\\
&-\phi_{\mathrm{u}_{12}\max}^2 \left\|\tilde{W}_{u_{12}}\right\|^2 + \phi_{\mathrm{u}_{12}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{2\max}\left(\frac{\left\|\tilde{W}_2\right\|^2}{2}+\frac{\left\|\tilde{W}_{u_{12}}\right\|^2}{2}\right)\\
&+\phi_{\mathrm{u}_{12}\max}\lambda(R_1^{-1})g_{21\max}\nabla\phi_{3\max}\left(\frac{\left\|\tilde{W}_3\right\|^2}{2}+\frac{\left\|\tilde{W}_{u_{12}}\right\|^2}{2}\right)\\
&+\left(\frac{\left(\phi_{u_{12}\max}\left(\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\epsilon_{2\max}+\lambda_{\max}(R_1^{-1})g_{21\max}\nabla\epsilon_{3\max}\right)\right)^2}{2}\right)+\frac{1}{2}\left\|\tilde{W}_{u_{12}}\right\|^2\\
&-\phi_{\mathrm{u}_{21}\max}^2 \left\|\tilde{W}_{u_{21}}\right\|^2 + \phi_{\mathrm{u}_{21}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{2\max}\left(\frac{\left\|\tilde{W}_2\right\|^2}{2}+\frac{\left\|\tilde{W}_{u_{21}}\right\|^2}{2}\right)\\
&+\phi_{\mathrm{u}_{21}\max}\lambda_{\max}(R_2^{-1})g_{12\max}\nabla\phi_{1\max}\left(\frac{\left\|\tilde{W}_1\right\|^2}{2}+\frac{\left\|\tilde{W}_{u_{21}}\right\|^2}{2}\right)\\
&+\left(\frac{\left(\phi_{u_{21}\max}\left(\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\epsilon_{2\max}+\lambda_{\max}(R_2^{-1})g_{12\max}\nabla\epsilon_{1\max}\right)\right)^2}{2}\right)+\frac{1}{2}\left\|\tilde{W}_{u_{21}}\right\|^2\\
&-\phi_{u_{22}\max}^2 \left\|\tilde{W}_{u_{22}}\right\|^2 + \phi_{\mathrm{u}_{22}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{3\max}\left(\frac{\left\|\tilde{W}_3\right\|^2}{2}+\frac{\left\|\tilde{W}_{u_{22}}\right\|^2}{2}\right)\\
&+\frac{\left((\phi_{u_{22}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\epsilon_{3\max}+\phi_{u_{22}\max}\epsilon_{u_2\max})+\phi_{u_{22}\max}\epsilon_{u_{22}\max}\right)^2}{2}+\frac{1}{2}\left\|\tilde{W}_{u_{22}}\right\|^2
\end{aligned}
\tag{103}
$$

For the term T_3 one has

$$
\begin{aligned}
T_3 :=&V_{1x_1}{}^{*\mathrm{T}}\big(f_{11}(x_1)-g_{11}(x_1)\tilde{W}_{u_{11}}^{\mathrm{T}}\phi_{u_{11}}+g_{11}(x_1)(u_{11}^*-\epsilon_{u_{11}})-g_{11}(x_1)B_{11}\frac{x_1^{\mathrm{T}}x_1\mathbf{1}_{m_1}}{(A+x_1^{\mathrm{T}}x_1)}\big)\\
&+V_{2x_1}{}^{*\mathrm{T}}\big(f_{12}(x)-g_{12}(x)\tilde{W}_{u_{12}}^{\mathrm{T}}\phi_{u_{12}}+g_{12}(x)(u_{12}^*-\epsilon_{u_{12}})-g_{12}(x)B_{12}\frac{x^{\mathrm{T}}x\mathbf{1}_{m_1}}{(A+x^{\mathrm{T}}x)}\big)\\
&+V_{2x_2}{}^{*\mathrm{T}}\big(f_{21}(x)-g_{21}(x)\tilde{W}_{u_{21}}^{\mathrm{T}}\phi_{u_{21}}+g_{21}(x)(u_{21}^*-\epsilon_{u_{21}})-g_{21}(x)B_{21}\frac{x^{\mathrm{T}}x\mathbf{1}_{m_2}}{(A+x^{\mathrm{T}}x)}\big)\\
&+V_{3x_2}{}^{*\mathrm{T}}\big(f_{22}(x_2)-g_{22}(x_2)\tilde{W}_{u_{22}}^{\mathrm{T}}\phi_{u_{22}}+g_{22}(x_2)(u_{22}^*-\epsilon_{u_{22}})-g_{22}(x_2)B_{22}\frac{x_2^{\mathrm{T}}x_2\mathbf{1}_{m_2}}{(A+x_2^{\mathrm{T}}x_2)}\big)
\end{aligned}
\tag{104}
$$

DOI: 10.1002/acs

Using the HJB equations (20)-(22) in equation (104) yields the following expression for T_3

$$\begin{aligned}
T_3 = &-\frac{1}{2}\big(x_1^{\mathrm{T}} Q_1 x_1 + u_{11}^{*\mathrm{T}} R_1 u_{11}^{*})\big) - V_{1x_1}^{*\mathrm{T}} g_{11}(x_1)\tilde{W}_{u_{11}}^{\mathrm{T}} \phi_{u_{11}} - V_{1x_1}^{*\mathrm{T}} g_{11}(x_1)\epsilon_{u_{11}} - V_{1x_1}^{*\mathrm{T}} g_{11}(x_1) B_{11} \frac{x_1^{\mathrm{T}} x_1 \mathbf{1}_{m_1}}{(A + x_1^{\mathrm{T}} x_1)} \\
&- x_1^{\mathrm{T}} Q_2 x_2 - V_{2x_1}^{*\mathrm{T}} g_{12}(x) \tilde{W}_{u_{12}}^{\mathrm{T}} \phi_{u_{12}} \\
&- V_{2x_1}^{*\mathrm{T}} g_{12}(x)\epsilon_{u_{12}} - V_{2x_1}^{*\mathrm{T}} g_{12}(x) B_{12} \frac{x^{\mathrm{T}} x \mathbf{1}_{m_1}}{(A + x^{\mathrm{T}} x)} \\
&- V_{2x_2}^{*\mathrm{T}} g_{21}(x) \tilde{W}_{u_{21}}^{\mathrm{T}} \phi_{u_{21}} - V_{2x_2}^{*\mathrm{T}} g_{21}(x)\epsilon_{u_{21}} - V_{2x_2}^{*\mathrm{T}} g_{21}(x) B_{21} \frac{x^{\mathrm{T}} x \mathbf{1}_{m_2}}{(A + x^{\mathrm{T}} x)} \\
&- \frac{1}{2}\big(x_2^{\mathrm{T}} Q_3 x_2 + u_{22}^{\mathrm{T}} R_2 u_{22}\big) - V_{3x_2}^{*\mathrm{T}} g_{22}(x_2) \tilde{W}_{u_{22}}^{\mathrm{T}} \phi_{u_{22}} \\
&- V_{3x_2}^{*\mathrm{T}} g_{22}(x_2)\epsilon_{u_{22}} - V_{3x_2}^{*\mathrm{T}} g_{22}(x_2) B_{22} \frac{x_2^{\mathrm{T}} x_2 \mathbf{1}_{m_2}}{(A + x_2^{\mathrm{T}} x_2)}
\end{aligned}$$

then,

$$\begin{aligned}
T_3 \leqslant &-\frac{1}{2}\big(x_1^{\mathrm{T}} Q_1 x_1 + u_{11}^{*\mathrm{T}} R_1 u_{11}^{*}\big) - x_1^{\mathrm{T}} Q_2 x_2 \\
&- (W_{1\max}\phi_{1\mathrm{d}\max} + \epsilon_{1\mathrm{d}\max})\Big(g_{11\max}\phi_{u_{11}\max}\left\|\tilde{W}_{u_{11}}\right\| + (\epsilon_{u_{11}\max} + g_{11\max} B_{11}\frac{x_1^{\mathrm{T}} x_1 \mathbf{1}_{m_1}}{A + x_1^{\mathrm{T}} x_1})\Big) \\
&- (W_{2\max}\phi_{2\mathrm{d}\max} + \epsilon_{2\mathrm{d}\max})\Big(g_{12\max}\phi_{u_{12}\max}\left\|\tilde{W}_{u_{12}}\right\| + (\epsilon_{u_{12}\max} + g_{12\max} B_{12}\frac{x^{\mathrm{T}} x \mathbf{1}_{m_1}}{A + x^{\mathrm{T}} x})\Big) \\
&- (W_{2\max}\phi_{2\mathrm{d}\max} + \epsilon_{2\mathrm{d}\max})\Big(g_{21\max}\phi_{u_{21}\max}\left\|\tilde{W}_{u_{21}}\right\| + (\epsilon_{u_{21}\max} + g_{21\max} B_{21}\frac{x^{\mathrm{T}} x \mathbf{1}_{m_2}}{A + x^{\mathrm{T}} x})\Big) \\
&- \frac{1}{2}\big(x_2^{\mathrm{T}} Q_3 x_2 + u_{22}^{\mathrm{T}} R_2 u_{22}\big) \\
&- (W_{3\max}\phi_{3\mathrm{d}\max} + \epsilon_{3\mathrm{d}\max})\Big(g_{22\max}\phi_{u_{22}\max}\left\|\tilde{W}_{u_{22}}\right\| + (\epsilon_{u_{22}\max} + g_{22\max} B_{22}\frac{x_2^{\mathrm{T}} x_2 \mathbf{1}_{m_2}}{A + x_2^{\mathrm{T}} x_2})\Big).
\end{aligned} \tag{105}$$

DOI: 10.1002/acs

Since $A + x_1^{\mathrm{T}} x_1 > 0$, $A + x^{\mathrm{T}} x > 0$, $A + x_2^{\mathrm{T}} x_2 > 0$, T_3 can be further upper bounded as,

$$
\begin{aligned}
T_3 \leqslant & -\frac{1}{2}\big(x_1^{\mathrm{T}} Q_1 x_1 + u_{11}^{*\mathrm{T}} R_1 u_{11}^*)\big) - x_1^{\mathrm{T}} Q_2 x_2 - \frac{1}{2}\big(x_2^{\mathrm{T}} Q_3 x_2 + u_{22}^{\mathrm{T}} R_2 u_{22}\big) \\
& + \frac{g_{11\max}\phi_{u_{11}\max}}{2}\big((W_{1\max}\phi_{1\mathrm{dmax}} + \epsilon_{1\mathrm{dmax}})\big)^2 \\
& + \frac{g_{11\max}\phi_{u_{11}\max}}{2}\left\|\tilde{W}_{u_{11}}\right\|^2 \\
& + \frac{1}{2}(W_{1\max}\phi_{1\mathrm{dmax}} + \epsilon_{1\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u_{11}\max}^2 \\
& - (W_{1\max}\phi_{1\mathrm{dmax}} + \epsilon_{1\mathrm{dmax}}) g_{11\max} B_{11} \frac{x_1^{\mathrm{T}} x_1 \mathbf{1}_{m_1}}{A + x_1^{\mathrm{T}} x_1} \\
& + \frac{g_{12\max}\phi_{u_{12}\max}}{2}\big((W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}})\big)^2 + \frac{g_{12\max}\phi_{u_{12}\max}}{2}\left\|\tilde{W}_{u_{12}}\right\|^2 \\
& + \frac{1}{2}(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u_{12}\max}^2 \\
& - (W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}}) g_{12\max} B_{12} \frac{x^{\mathrm{T}} x \mathbf{1}_{m_1}}{A + x^{\mathrm{T}} x} \\
& + \frac{g_{21\max}\phi_{u_{21}\max}}{2}\left\|\tilde{W}_{u_{21}}\right\|^2 \\
& + \frac{1}{2}(W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u_{21}\max}^2 \\
& - (W_{2\max}\phi_{2\mathrm{dmax}} + \epsilon_{2\mathrm{dmax}}) g_{21\max} B_{21} \frac{x^{\mathrm{T}} x \mathbf{1}_{m_2}}{A + x^{\mathrm{T}} x} \\
& + \frac{g_{22\max}\phi_{u_{22}\max}}{2}\big((W_{3\max}\phi_{3\mathrm{dmax}} + \epsilon_{3\mathrm{dmax}})\big)^2 \\
& + \frac{g_{22\max}\phi_{u_{22}\max}}{2}\left\|\tilde{W}_{u_{22}}\right\|^2 \\
& + \frac{1}{2}(W_{3\max}\phi_{3\mathrm{dmax}} + \epsilon_{3\mathrm{dmax}})^2 + \frac{1}{2}\epsilon_{u_{22}\max}^2 \\
& - (W_{3\max}\phi_{3\mathrm{dmax}} + \epsilon_{3\mathrm{dmax}}) g_{22\max} B_{22} \frac{x_2^{\mathrm{T}} x_2 \mathbf{1}_{m_2}}{A + x_2^{\mathrm{T}} x_2}.
\end{aligned} \tag{106}
$$

DOI: 10.1002/acs

Finally, after taking into account the bound of $B_{11}\|x_1\|$, $B_{12}\|x\|$, $B_{21}\|x\|$, and $B_{22}\|x_2\|$ from (80), (81), (82), and (83), respectively, we can upper bound (101) as

$$
\begin{aligned}
\dot{\mathcal{V}} \leqslant & -\Bigg(2\alpha_1\lambda_{\min}\Big(\sum_{i=1}^{k_1}\frac{\omega_1(t_i)\omega_1(t_i)^{\mathrm{T}}}{(\omega_1(t_i)^{\mathrm{T}}\omega_1(t_i)+1)^2}\Big)-\frac{1}{4\alpha_1}-\frac{\phi_{u_{11}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{1\max}}{2} \\
& \quad -\frac{\phi_{u_{21}\max}\lambda_{\max}(R_2^{-1})g_{12\max}\nabla\phi_{1\max}}{2}\Bigg)\left\|\tilde{W}_1\right\|^2 \\
& -\Bigg(\phi^2_{u_{11}\max}-\frac{\phi_{u_{11}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{1\max}}{2}-\frac{1}{2}-\frac{g_{11_{\max}}\phi_{u_{11}\max}}{2}\Bigg)\left\|\tilde{W}_{u_{11}}\right\|^2 \\
& -\frac{1}{2}(x_1^{\mathrm{T}}Q_1x_1+u_{11}^{*\mathrm{T}}R_1u_{11}^*)-x_1^{\mathrm{T}}Q_2x_2-\frac{1}{2}(x_2^{\mathrm{T}}Q_3x_2+u_{22}^{\mathrm{T}}R_2u_{22}) \\
& -\Bigg(2\alpha_2\lambda_{\min}\Big(\sum_{i=1}^{k_2}\frac{\omega_2(t_i)\omega_2(t_i)^{\mathrm{T}}}{(\omega_2(t_i)^{\mathrm{T}}\omega_2(t_i)+1)^2}\Big)-\frac{1}{4\alpha_2}-\frac{\phi_{u_{12}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{2\max}}{2} \\
& \quad -\frac{\phi_{u_{21}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{2\max}}{2}\Bigg)\left\|\tilde{W}_2\right\|^2 \\
& -\Bigg(\phi^2_{u_{12}\max}-\frac{\phi_{u_{12}\max}\lambda_{\max}(R_1^{-1})g_{11\max}\nabla\phi_{2\max}}{2}-\frac{\phi_{u_{12}\max}\lambda_{\max}(R_1^{-1})g_{21\max}\nabla\phi_{3\max}}{2} \\
& \quad -\frac{1}{2}-\frac{g_{12_{\max}}\phi_{u_{12}\max}}{2}\Bigg)\left\|\tilde{W}_{u_{12}}\right\|^2 \\
& -\Bigg(\phi^2_{u_{21}\max}-\frac{\phi_{u_{21}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{2\max}}{2}-\frac{\phi_{u_{21}\max}\lambda_{\max}(R_2^{-1})g_{12\max}\nabla\phi_{1\max}}{2} \\
& \quad -\frac{1}{2}-\frac{g_{21_{\max}}\phi_{u_{21}\max}}{2}\Bigg)\left\|\tilde{W}_{u_{21}}\right\|^2 \\
& -\Bigg(2\alpha_3\lambda_{\min}\Big(\sum_{i=1}^{k_3}\frac{\omega_3(t_i)\omega_3(t_i)^{\mathrm{T}}}{(\omega_3(t_i)^{\mathrm{T}}\omega_3(t_i)+1)^2}\Big)-\frac{1}{4\alpha_3}-\frac{\phi_{u_{22}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{3\max}}{2} \\
& \quad -\frac{\phi_{u_{12}\max}\lambda_{\max}(R_1^{-1})g_{21\max}\nabla\phi_{3\max}}{2}\Bigg)\left\|\tilde{W}_3\right\|^2 \\
& -\Bigg(\phi^2_{u_{22}\max}-\frac{\phi_{u_{22}\max}\lambda_{\max}(R_2^{-1})g_{22\max}\nabla\phi_{3\max}}{2}-\frac{1}{2}-\frac{g_{22_{\max}}\phi_{u_{22}\max}}{2}\Bigg)\left\|\tilde{W}_{u_{22}}\right\|^2,\ t\geqslant 0.
\end{aligned}
$$

Then by taking into account the inequalities (85)-(91) (which are the parentheses above) one has $\dot{\mathcal{V}}\leqslant 0,\ \ t\geqslant 0$. From Barbalat's lemma [50] it follows that as $t\to\infty$, then $\|\tilde{Z}\|\to 0$ The result holds as long as we can show that the state $x(t)$ remains in the set $\Omega\subseteq\mathbb{R}^n$ for all times. To this effect, define the following compact set

$$M \coloneqq \left\{x\in\mathbb{R}^n|\mathcal{V}(t)\leqslant m)\right\}\subset\mathbb{R}^n$$

where m is chosen as the largest constant so that $M\subseteq\Omega$. Since by assumption $x_0\in\Omega_x$, and $\Omega_x\subset\Omega$ then we can conclude that $x_0\in\Omega$. While $x(t)$ remains inside Ω, we have seen that $\dot{\mathcal{V}}\leqslant 0$ and therefore $x(t)$ must remain inside $M\subset\Omega$. The fact that $x(t)$ remains inside a compact set also excludes the possibility of finite escape time and therefore one has global existence of solution. ■

REFERENCES

1. Lewis F.L., Vrabie D., Syrmos V.L. *Optimal Control*, John Wiley & Sons., January 2012.
2. Kokotovic P., Perkins W., Cruz J.B., D'Ans G. *ϵ-coupling for near optimum design of large scale linear systems*, Inst. Elect. Eng. Proc. Part D, May 1969; 116(5):889-892.
3. Gajic Z., Petkovski D., ShenX. *Singularly Perturbed and Weakly Coupled Linear Control Systems: A Recursive Approach*, Springer-Verlag, Berlin, Germany, 1990.

4. Gajic Z., Shen X. *Parallel Algorithms for Optimal Control of Large Scale Linear Systems*, Springer, London, U.K., 1992.
5. Kaszkurewics E., Bhaya A., Silja D. *On the convergence of parallel asynchronous block-iterative computations*, Linear Algebra and its Applications, 1990; 131:139-160.
6. Zecevic A., Siljak D. *A block-parallel Newton method via overlapping epsilon decomposition*, SIAM Journal of Matrix Analysis and Applications, 1994; 15:824-844.
7. Okuguchi K. *Matrices with dominant diagonal blocks and economic theory*, Journal of Mathematical Economics, 1978; 5:43-52.
8. Medanic J., Avramovic B. *Solution of load-flow problems in power systems by ϵ-coupling method*, Proceeding of the Institution of Electrical Engineers, August 1975; 122(8):801-805.
9. Ilić-Spong M., Katz N., Dai H., Zaborsky J. *Block diagonal dominance for systems of nonlinear equations with applications to load flow calculations in power systems*, Mathematical Modelling, 1984; 5(5):275-297.
10. Crow M., Ilic M. *The parallel implementation of the waveform relaxation method for transient stability simulations*, IEEE Transactions on Power Systems, August, 1990; 5:922-932.
11. Phillips R., Kokotovic P. *A singular perturbation approach to modeling and control of Markov chains*, IEEE Transactions on Automatic Control, October 1981; 26:1087-1094.
12. Aldhaheri R., Khalil H. *Aggregation method for nearly completely decomposable Markov chains*, IEEE Transactions on Automatic Control, February 1991; 36:178187.
13. Stewart W., *Introduction to Numerical Solution of Markov Chains*, Princeton Univ. Press, Princeton, NJ, 1994.
14. Kokotovic P., Singh G., *Optimization of Coupled Nonlinear Systems*, International Journal of Control, 1971; 14: 51-64.
15. Beard R., McLain T. *Successive Galerkin approximation algorithms for nonlinear optimal and robust control*, International Journal of Control, 1998; 71(5):717-743.
16. Kim Y.J., Kim B.S., Lim M.T. *Composite control for singularly perturbed nonlinear systems via successive Galerkin approximation*, IEE Proceedings on Control Theory and Applications, September 2003; 150(5):483-488.
17. Kim Y.J., Kim B.S., Lim M.T. *Finite-time composite control for a class of singularly perturbed nonlinear systems via successive Galerkin approximation*, IEE Proceedings on Control Theory and Applications, 2005; 152(5):507-512.
18. Werbos P.J. *Brain-like intelligent control: from neural nets to larger-scale systems*, IEEE Conference on Decision and Control, San Diego, CA, USA, December 1997; 3902-3904.
19. Werbos P.J. *Intelligence in the brain: A theory of how it works and how to build it* Neural Networks, April 2009; 22(3):200-212.
20. Murray J.J., Cox C.J., Lendaris G.G., Saeks R. *Adaptive dynamic programming*, IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, May 2002; 32(2):140-153.
21. Bertsekas D.P. *Dynamic Programming and Optimal Control*, Vol. II, Athena Scientific, 3rd edition, 2007.
22. Powell W.B. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, Wiley Series in Probability and Statistics. Wiley, 2007.
23. Bhasin S., Kamalapurkar R., Johnson M., Vamvoudakis K., Lewis F., Dixon W. *A novel actor critic identifier architecture for approximate optimal control of uncertain nonlinear systems*, Automatica, 2013; 49(1):82-92.
24. Lewis F., Liu D. *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. IEEE Press Series on Computational Intelligence, Wiley, 2013.
25. Lewis F., Vrabie D., Vamvoudakis K. *Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers*, IEEE Control Systems Magazine, 2012; 32(6):76-105.
26. Vamvoudakis K.G., Lewis F.L. *Online actor critic algorithm to solve the continuous-time infinite horizon optimal control problem*, Automatica, 2010; 46(5):878-888.
27. Wei Q.L., Zhang H.G., Cui L.L. *Data-based optimal control for discrete-time zero-sum games of 2-d systems using adaptive critic designs*, Acta Automatica Sinica, 2009; 35(6):682-692.
28. Khalil H.K. *Nonlinear systems*, Macmillan Pub. Co., 1992.
29. Sutton R.S., Barto A.G. *Reinforcement learning: An introduction*, MIT Press, Cambridge, MA, 1998.
30. Gajic Z., Shen X. *Decoupling transformation for weakly coupled linear systems*, International Journal of Control, 1989; 50:1515-1521.
31. Gajic Z., Borno I. *General Transformation for Block Diagonalization of Weakly Coupled Linear Systems Composed of N-Subsystems* Transactions on Circuits and Systems, Fundamental Theory and Applications, June 2000; 47(6):909-912.
32. Gajic Z., Lim M.-T., Skataric D., Su W.-C., Kecman V. *Optimal control: weakly coupled systems and applications*, CRC Press, 2008.
33. Aganovic Z., Gajic Z. *Optimal control of weakly coupled bilinear systems*, Automatica, November 1993; 29(6):1591-1593.
34. Aganovic Z., Gajic Z. *Linear optimal control of bilinear systems: With applications to singular perturbations and weak coupling*, Springer, London, U.K., 1995.
35. Cebuhar W., Costanza V. *Approximation procedures for the optimal control fo bilinear and nonlinear systems*, Journal of Optimimal Theory and Applications, 1984; 43(4):615-627.
36. Kim Y.J., Lim M.T. *Parallel Optimal Control for Weakly Coupled Nonlinear Systems Using Successive Galerkin Approximation*, IEEE Transactions on Automatic Control, July 2008; 53(6):1542-1547.
37. Chen Z., Jagannathan S. *Generalized Hamilton-Jacobi-Bellman formulation: Based neural network control of affine nonlinear discrete-time systems*, IEEE Transactions on Neural Networks, January 2008; 19(1):90-106.
38. Lewis F.L., Vrabie D., Vamvoudakis K.G. *Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers*, IEEE Control Systems Magazine, December 2012; 32(6):76-105.
39. Chowdhary G., Johnson E. *Concurrent learning for convergence in adaptive control without persistency of excitation*, In IEEE Conference and Decision and Control, Atlanta, GA, December 2010; 3674-3679.
40. Boyan J.A. *Technical update: Least-squares temporal difference learning*, Machine Learning, 2002; 49(2-3):233-246.

DOI: 10.1002/acs

41. Heydari A., Balakrishnan S.N. *Fixed-final-time optimal control of nonlinear systems with terminal constraints*, Neural Networks, 2013; 48:61-71.
42. Modares H., Lewis F.L., Naghibi-Sistani M.B. *Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems*, Automatica, 2014; 50(1):193-202.
43. M. Bardi, C-D. Italo, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Springer, 2008.
44. A. J. Van der Schaft, *L2-gain analysis of nonlinear systems and nonlinear state feedback H-∞ control*, IEEE Transactions on Automatic Control, 1992; 37(6): 770-784.
45. Abu-Khalaf M., Lewis F.L. *Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network approach*, Automatica, 2005; 41(5):779-791.
46. Hornik K., Stinchcombe M.B., White H. *Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks*, Neural Networks, 1990; 3(5):551-560.
47. Ioannou P., Fidan B. *Adaptive Control Tutorial*, Advances in Design and Control, Society for Industrial and Applied Mathematics, 2006.
48. Chowdhary G., Yucelen T., Mhlegg M., Johnson E.N. *Concurrent learning adaptive control of linear systems with exponentially convergent bounds*, International Journal of Adaptive Control and Signal Processing, 2012.
49. Vamvoudakis K.G., Lewis F.L. *Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem*, Automatica, 2010; 46(5):878-888.
50. Haddad W.M., Chellaboina V.S. *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*, Princeton University Press, 2008.
51. Dierks T., Jagannathan S. *Optimal control of affine nonlinear continuous-time systems*, in American Control Conference, 2010; 1568-1573.
52. Polycarpou M., Farrell J., Sharma M. *On-line approximation control of uncertain nonlinear systems: issues with control input saturation*, in Proc. American Control Conference, 2003; 543-548.