Journal of Computer Science and Cybernetics, V.39, N.2 (2023), 175–189
DOI no 10.15625/1813-9663/18099

OPTIMAL TRACKING CONTROL FOR ROBOT MANIPULATORS WITH INPUT CONSTRAINT BASED REINFORCEMENT LEARNING

NGUYEN DUC DIEN[1], NGUYEN TAN LUY[2,*], LAI KHAC LAI[3], TRAN THANH HAI[4]

[1] *University of Economics – Technology for Industry, Ha Noi, Viet Nam*
[2] *Ho Chi Minh City University of Technology–VNU, Ho Chi Minh City, Viet Nam*
[3] *Thai Nguyen University of Technology, Viet Nam*
[4] *Industrial University of Ho Chi Minh City, Viet Nam*

Abstract. This paper introduces an optimal tracking controller for robot manipulators with saturation torques. The robot model is presented as a strict-feedback nonlinear system. Firstly, the position tracking control problem is transformed into the optimal tracking control problem. Subsequently, the saturated optimal control law is designed. The optimal control law is determined through the solution of the Hamilton-Jacobi-Bellman (HJB) equation. We use a reinforcement learning algorithm with only one neural network (NN) to approximate the solution of the equation HJB. The technique of experience replay is used to relax a persistent citation condition. By Lyapunov analysis, the tracking and the approximation errors are uniformly ultimately bounded (UUB). Finally, the simulation on a robot manipulator with saturation torques is performed to verify the efficiency of the proposed controller.

Keywords. Reinforcement learning; Saturation torques; Saturated optimal tracking control; Robot.

1. INTRODUCTION

Robot manipulators have brought excellent efficiency in manufacturing, healthcare, and services. Controller designs, aiming to improve performance and reduce costs, have continuously received the attention of researchers [8, 24]. Sliding controllers for the robot manipulators from basic to advanced have been implemented, such as adaptive slide controllers [1], terminal slide controllers [5], third-order slide controllers [13], and fixed-time slide controllers [25]. Chwa et al. [2] proposed a sliding controller combined with a system identifier and disturbance observer. A terminal slide controller combined with the backstepping controller was used in [17, 21]. Intelligent controllers combined with advanced controllers have also been implemented, such as an adaptive fuzzy sliding controller [20], an adaptive fuzzy backstepping controller [27], and a terminal slide mode controller combined with a radial basic function (RBF) neural network (NN) controller [4].

*Corresponding author.
E-mail addresses: nddien@uneti.edu.vn (N.D. Dien); ntanluy@hcmut.edu.vn (N.T. Luy); laikhaclai@tnu.edu.vn (L.K. Lai); tranthanhhai@iuh.edu.vn (T.T. Hai)

© 2023 Vietnam Academy of Science & Technology

In practice, the actuators are constrained by the saturation torque, which results in reduced control efficiency or instability. Wei et al. [3] proposed an efficient impedance controller based on the Lyapunov function to solve the input saturation problem. Ling et al. [10] designed an adaptive fuzzy dynamic surface controller in which the smooth function was used with the mean value theorem to solve the difficulties related to input saturation. Yang et al. [22] proposed a bounded barrier Lyapunov function and have designed an auxiliary system to suppress the input saturation effect. In [28], a NN controller combined with the backstepping technique was proposed to apply to robot manipulators with input saturation. The controllers proposed in [3, 10, 22, 28] have effectively solved the input saturation problem for the robot manipulators; however, optimal control has not been presented.

The optimal control problem's solution depends on the HJB equation's solution. Reinforcement learning and NN have been proper methods to approximate online HJB equation solutions [18, 23]. The optimal control for the robot manipulators was proposed in [6, 7, 12]. The algorithm in [6] used 2 NNs, which increased the computational cost. In [14–16], the saturated optimal tracking controller was designed using RL with only one NN. They implemented the proposed algorithm through experiments on a mobile robot [14] and a PMSM system [16]. For robot manipulators with input constraints, Zhao et al. [26] proposed an optimal controller based on RL in discrete time. To the best of our knowledge, the saturated optimal tracking controller has not been resolved for the robot manipulators in continuous time.

From the above analysis, we design an optimal tracking controller for robot manipulators with input constraints based on an RL method using a NN. The contributions of this paper include the following.

1. The dynamics of the robot manipulator are represented as a strict-feedback nonlinear system with saturation inputs. The feedforward control inputs are proposed. Then the tracking control problem is transformed into an optimal control problem.

2. The cost function of the saturation system is proposed. The solution of the HJB equation determines the saturation optimal control law. Therefore, in this paper, we use a reinforcement learning algorithm with only one NN instead of 2 [6, 18] to approximate the solution of the HJB equation. Furthermore, the experience replay is applied when updating the NN parameters to relax a persistent citation condition.

The structure of the paper is organized as follows. The robot manipulators' dynamics and the system transformation problem are presented in Section 2. In Section 3, based on the transformed problem, the saturated optimal control law is designed. In Section 4, the simulation results are shown. Some conclusions are made in section 5.

2. DYNAMICS OF THE ROBOT MANIPULATOR AND SYSTEM TRANSFORMATION

This section presents the robot manipulator dynamics model and proposes a feedforward controller to convert the tracking control problem into the optimal control problem.

2.1. Dynamics of the robot manipulator

Consider a model of robot manipulator Euler – Lagrange systems [8] with the dynamic equation presented as

$$D(\varphi)\ddot{\varphi} + H(\varphi, \dot{\varphi})\dot{\varphi} + G(\varphi) = \tau, \tag{1}$$

where $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_n] \in \mathbb{R}^{n\times 1}$ is the angular position vector, $\dot{\varphi} = [\dot{\varphi}_1, \dot{\varphi}_2, \dots, \dot{\varphi}_n] \in \mathbb{R}^{n\times 1}$ is the angular velocity vector, and $\ddot{\varphi} = [\ddot{\varphi}_1, \ddot{\varphi}_2, \dots, \ddot{\varphi}_n] \in \mathbb{R}^{n\times 1}$ is the angular acceleration vector. The inertia matrix $D(\varphi) \in \mathbb{R}^{n\times n}$ is symmetric positive definite, $H(\varphi, \dot{\varphi}) \in \mathbb{R}^{n\times n}$ is the Coriolis-centripetal matrices, and $G(\varphi) \in \mathbb{R}^{n\times 1}$ is the dynamic friction. $\tau = [\tau_1, \tau_2, \dots, \tau_n] \in \mathbb{R}^{n\times 1}$ is the control input vector, which is bounded by ρ, i.e.,$\|\tau\| \le \rho$.

Property 1. *There exist positive constants b_D, b_H, and b_G such that $\left\|D^{-1}(\varphi)\right\| \le b_D$, $\|H(\varphi, \dot{\varphi})\| \le b_H$, and $\|G(\varphi)\| \le b_G$.*

By defining $\upsilon = \dot{\varphi}$, dynamics (1) is rewritten as a strict-feedback nonlinear system as follows

$$\begin{cases} \dot{\chi}_\varphi = F_\varphi(\chi_\varphi) + g_\varphi(\chi_\varphi)\chi_\upsilon \\ \dot{\chi}_\upsilon = F_\upsilon(\chi_\varphi, \chi_\upsilon) + g_\upsilon(\chi_\varphi, \chi_\upsilon)\tau, \end{cases} \tag{2}$$

where $\chi_\varphi = \varphi$, $\chi_\upsilon = \upsilon$, $F_\varphi(\chi_\varphi) = 0_{n\times 1}$, $g_\varphi(\chi_q) = I_n$, $f_\upsilon(\chi_\varphi, \chi_\upsilon) = -D^{-1}(\varphi)(H(\varphi, \dot{\varphi})\upsilon + G(\varphi)) \in \mathbb{R}^{n\times 1}$, and $g_\upsilon(\chi_\varphi, \chi_\upsilon) = D^{-1}(\varphi) \in \mathbb{R}^{n\times n}$.

Remark 1. *From Properties 1, one has $\|F_\upsilon(\chi_\varphi, \chi_\upsilon)\| \le b_{F_\upsilon}$, $\|g_\upsilon(\chi_\varphi, \chi_\upsilon)\| \le b_{g_\upsilon}$, where b_{f_υ} and b_{g_υ} are known positive constants.*

Define $\chi_{\varphi d}$ as the reference position trajectory, and tracking error $z_\varphi = \chi_\varphi - \chi_{\varphi d}$. The control objective is to find feedback control laws for the system (2), such that $\lim\limits_{t\to\infty} z_{\chi_\varphi}(t) \to 0$. Furthermore, a defined tracking cost function related to tracking errors and control inputs is minimized.

2.2. System transformation

The problem of optimal controller design is easily accomplished with an affine nonlinear system. Therefore, this section presents the steps to convert system (2) to the equivalent affine nonlinear system by applying the backstepping technique [23].

We first define the new variables as follows: $\chi_{\upsilon d} = \chi_{\upsilon d}^* + \chi_{\upsilon d}^a$, $\tau = \tau^* + \tau^a$, where $\chi_{\upsilon d}$ is the augmented pseudo-control input vector, $\chi_{\upsilon d}^* = [\chi_{\upsilon d1}^{*T}, ..., \chi_{\upsilon dn}^{*T}]^T$ will be the saturated virtual optimal control input designed in the next section, $\chi_{\upsilon d}^a$ is the feedforward virtual control input, τ is the real control input vector, $\tau^* = [\tau_1^{*T}, ..., \tau_n^{*T}]^T$ will be the saturated optimal feedback control input designed in the next section, $\tau^a = [\tau_1^{aT}, ..., \tau_n^{aT}]^T$ is the feedforward control input.

Taking the derivative z_φ and z_υ, we get

$$\begin{cases} \dot{z}_\varphi = -\dot{\chi}_{\varphi d} + g_\varphi(\chi_\varphi)\chi_{\upsilon d}^* + g_\varphi(\chi_\varphi)\chi_{\upsilon d}^a + g_\varphi(\chi_\varphi)z_\upsilon \\ \dot{z}_\upsilon = -\dot{\chi}_{\upsilon d} + F_\upsilon(\chi_\varphi, \chi_\upsilon) + g_\upsilon(\chi_\varphi, \chi_\upsilon)\tau^* + g_\upsilon(\chi_\varphi, \chi_\upsilon)\tau^a. \end{cases} \tag{3}$$

We design the feedforward control inputs as follows

$$\begin{cases} \chi_{vd}^{a} = g_{\varphi}^{-1}(\chi_{\varphi})(\Lambda_1 z_v + \dot{\chi}_{\varphi d} + f_{\varphi}(z_{\varphi})) \\ \tau^{a} = g_{v}^{-1}(\chi_{\varphi}, \chi_v)\Big(F_v(z_{\varphi}, z_v) + \dot{\chi}_{vd} - F_v(\chi_{\varphi}, \chi_v) - g_{\varphi}^{T}(\chi_{\varphi}) z_{\varphi} - \Lambda_2 z_v\Big), \end{cases} \tag{4}$$

where $z_v = \dot{\chi}_{\varphi} - \chi_{vd}$, χ_{vd}^{a} and τ^{a} are limited by $(\rho - \sigma\tanh(1))$, i.e.,

$$\begin{cases} \|\chi_{vd}^{a}\| \leq \rho - \sigma\tanh(1) \\ \|\tau^{a}\| \leq \rho - \sigma\tanh(1), \end{cases} \tag{5}$$

where $0 < \sigma < \rho$.

Substituting (4) into (3), the tracking error dynamics become

$$\begin{cases} \dot{z}_{\varphi} = \bar{F}_{\varphi}(z_{\varphi}, z_v) + g_{\varphi}(\chi_{\varphi})\chi_{vd}^{*} + g_{\varphi}(\chi_{\varphi}) z_v \\ \dot{z}_v = \bar{F}_v(z_{\varphi}, z_v) + g_v(\chi_{\varphi}, \chi_v)\tau^{*} - g_{\varphi}^{T}(\chi_{\varphi}) z_{\varphi}, \end{cases} \tag{6}$$

where $\bar{F}_{\varphi}(z_{\varphi}, z_v) = F_{\varphi}(z_{\varphi}) + \Lambda_1 z_v$, $\bar{F}_v(z_{\varphi}, z_v) = F_v(z_{\varphi}, z_v) - \Lambda_2 z_v$.

Lemma 1. *Consider the following error dynamics*

$$\dot{z} = \bar{F}_{\varphi v}(z_{\varphi}, z_v) + g_{\varphi v}(\chi_{\varphi}, \chi_v)\mu^{*}, \tag{7}$$

where $z = [z_{\varphi}^{T}, z_v^{T}]^{T} \in \mathbb{R}^{2n\times 1}$, $\bar{F}_{\varphi v}(z_{\varphi}, z_v) = [\bar{F}_{\varphi}^{T}(z_{\varphi}, z_v), \bar{F}_{v}^{T}(z_{\varphi}, z_v)]^{T} \in \mathbb{R}^{2n\times 1}$,

$g_{\varphi v}(\chi_{\varphi}, \chi_v) = diag[g_{\varphi}(\chi_{\varphi}),\ g_v(\chi_{\varphi}, \chi_v)] \in \mathbb{R}^{2n\times 2n}$, $\mu^{*} = [\chi_{vd}^{*T}, \tau^{*T}]^{\mathrm{T}} \in \mathbb{R}^{2n\times 1}$,

$\mu_a = [\chi_{vd}^{aT}, \tau^{aT}]^{T} \in \mathbb{R}^{2n\times 1}$, $\mu_{\tau} = [\chi_{vd}^{T}, \tau^{T}]^{T} = \mu^{*} + \mu^{a} \in \mathbb{R}^{2n\times 1}$. *Suppose the saturated optimal control law u^{*} is designed to stabilize the system* (7). *In that case, the optimal tracking control problem for system* (2) *and the optimal control problem for system* (7) *are equivalent.*

Proof. Consider a Lyapunov function for system (7) as follows

$$\mathcal{L}_1 = \frac{1}{2} z_{\varphi}^{T} z_{\varphi} + \frac{1}{2} z_v^{T} z_v. \tag{8}$$

Taking derivative (8) along trajectories of (6), one obtains

$$\begin{aligned} \dot{\mathcal{L}}_1 &= z_{\varphi}^{T}\bar{F}_{\varphi}(z_{\varphi}, z_v) + z_{\varphi}^{T} g_{\varphi}(\chi_{\varphi})\chi_{vd}^{*} + z_{\varphi}^{T} g_{\varphi}(\chi_{\varphi}) z_v + z_v^{T}\bar{F}_v(z_{\varphi}, z_v) + z_v^{T} g_v(\chi_{\varphi}, \chi_v)\tau^{*} \\ &\quad - z_v^{T} g_{\varphi}^{T}(\chi_{\varphi}) z_{\varphi} \\ &= z_{\varphi}^{T}\left(\bar{F}_{\varphi}(z_{\varphi}, z_v) + g_{\varphi}(\chi_{\varphi})\chi_{vd}^{*}\right) + z_v^{T}\left(\bar{F}_v(z_{\varphi}, z_v) + g_v(\chi_{\varphi}, \chi_v)\tau^{*}\right) \\ &= z^{T}\left(\bar{F}_{\varphi v}(z_{\varphi}, z_v) + g_{\varphi v}(\chi_{\varphi}, \chi_v)\mu^{*}\right). \end{aligned} \tag{9}$$

On the other hand, for system (7), consider a Lyapunov function as

$$\mathcal{L}_2 = \frac{1}{2} z^{T} z. \tag{10}$$

Taking derivative (10), one obtains

$$\dot{\mathcal{L}}_2 = \frac{1}{2} z^T (\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\mu^*). \tag{11}$$

Comparing (9) with (11), it can be seen that if the control law u^* makes the system (7) stable, i.e., $\dot{\mathcal{L}}_2 < 0$, then $\dot{\mathcal{L}}_1 < 0$, therefore, the closed system (2) is also stable. In other words, the optimal tracking problem for the system (2) is transformed into the optimal control problem for the system (7).

The proof is completed. ∎

Remark 2. *From* (4)*, it can be seen that if the optimal control input vector* τ^* *is designed to satisfy the constraint* $\|\tau^*\| \leq \sigma$*, then the control input vector* τ *satisfies the constraint* $\|\tau\| \leq \rho$*.*

3. OPTIMAL TRACKING CONTROL

In this section, we analyze and design the optimal control law for system (7) with input constraint, such that the tracking error is stable. Furthermore, the cost function is minimized, where the cost function is suggested as

$$\mathcal{J}(z) = \int_t^\infty (z^T \Theta z + \Omega(\mu))dt, \tag{12}$$

where Θ is a positive definite matrix, μ is an approximation of μ^* at each given time, and $\Omega(\mu) \geq 0$ is the energy cost function. Then the problem of optimal tracking control for system (2) is solved by Lemma 1.

Consider system (7) with the cost function defined as (12). For input constraints, one choice for $\Omega(\mu)$ is [19]

$$\Omega(\mu) = 2\varrho \int_0^\mu \tanh^{-T}(s/\varrho) R ds, \tag{13}$$

where $R = \mathrm{diag}(r_1, ..., r_{2n}) > 0$. To facilitate the calculation and implementation of the proposed algorithm, we transform equation (13) as follows:

- Firstly, using integral by parts (13), we get

$$\Omega(\mu) = 2\varrho \tanh^{-T}\left(\frac{s}{\varrho}\right) Rs + \Omega_1(\mu), \tag{14}$$

where $\Omega_1(\mu) = -2\rho\bar{R}\int_0^\mu s\nabla\gamma ds$, $\nabla\gamma = \frac{\partial \tanh^{-T}(s/\varrho)}{\partial s}$.

- Next, defining $\kappa = \tanh^{-T}(s/\varrho) \to s = \rho\tanh(\kappa)$, $\Omega_1(\mu)$ becomes

$$\Omega_1(\mu) = -2\rho^2\bar{R} \int_0^{\tanh^{-T}(\mu/s)} \tanh(\kappa)d\kappa. \tag{15}$$

Defining $\beta = \tanh(\kappa) \rightarrow d\kappa = \frac{d\beta}{\bar{1}-\tanh^2(\kappa)} = \frac{d\beta}{\bar{1}-\beta^2}$, where $\bar{1} = [1, 1, \cdots, 1]^T$, (15) is rewritten as

$$\Omega_1(\mu) = -2\varrho^2 \bar{R} \int\limits_0^{\mu/\varrho} \frac{\beta}{\bar{1}-\beta^2} d\beta = \varrho^2 \bar{R} \ln \left(\bar{1} - \left(\frac{\mu}{\varrho} \right)^2 \right). \tag{16}$$

Substituting (16) into (14), $\Omega(\mu)$ becomes

$$\Omega(\mu) = 2\rho \tanh^{-T} \left(\frac{\mu}{\varrho} \right) R\mu + \varrho^2 \bar{R} \ln \left(\bar{1} - \frac{\mu^2}{\rho^2} \right). \tag{17}$$

Using differential (12) along the system trajectories (7), we get the following Bellman equation

$$z^T \Theta z + \Omega(\mu) + \nabla \mathcal{J}^T(z)(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\mu) = 0, \tag{18}$$

where $\nabla \mathcal{J}(z) = \partial \mathcal{J}(z)/\partial z$. The Hamilton function for the system (7) is defined as

$$H\left(z, \mu, \nabla \mathcal{J}(z)\right) = z^T \Theta z + \Omega(\mu) + \nabla \mathcal{J}^T(z)(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\mu). \tag{19}$$

Let $\mathcal{J}^*(z)$ be the optimal cost function, which is defined as

$$\mathcal{J}^*\left(z\right) = \min_{\mu} \left(\mathcal{J}\left(z\right)\right). \tag{20}$$

Define the optimal Hamilton function as follows

$$H\left(z, \mu, \nabla \mathcal{J}^*(z)\right) = z^T \Theta z + \Omega(\mu) + \nabla \mathcal{J}^{*T}(z)(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\mu). \tag{21}$$

The HJB equation is presented as

$$H\left(z, \mu^*, \nabla \mathcal{J}^*(z)\right) = z^T \Theta z + \Omega(\mu^*) + \nabla \mathcal{J}^{*T}(z)(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\mu^*) = 0, \tag{22}$$

where $\nabla \mathcal{J}^*(z) = \partial \mathcal{J}^*(z)/\partial z$. Applying $\frac{\partial H(z,\mu,\nabla \mathcal{J}^*(z))}{\partial \mu} = 0$ to (19), we derive the saturated optimal control law as follows

$$\mu^* = \arg\min_{\mu} H(z, \mu, \nabla \mathcal{J}^*(z)) = -\rho \tanh \left(\frac{1}{2\rho} R^{-1} g_{\varphi v}^T(\chi_\varphi, \chi_v) \nabla \mathcal{J}^*(z) \right). \tag{23}$$

To solve the optimal control problem, one solves the HJB equation (22), then determines the saturated optimal control law μ^* as (27). In this paper, we solve the HJB equation using RL and NN. $\mathcal{J}^*(z)$ is approximated as follows

$$\mathcal{J}^*(z) = W^T \zeta(z) + \epsilon(z), \tag{24}$$

$$\nabla \mathcal{J}^*(z) = W^T \frac{\partial \zeta(z)}{\partial z} + \frac{\partial \nabla \epsilon(z)}{\partial z} = W^T \nabla \zeta + \nabla \epsilon, \tag{25}$$

where $W \in \mathbb{R}^H$ is the NN weight vector, $\zeta(z) : \mathbb{R}^n \rightarrow \mathbb{R}^H$ is the NN activation function vector with $\zeta(0) = 0$, H is the number of neurons in the hidden layer, $\epsilon(z)$ is the NN approximation error.

Assumption 1. *One can choose $\zeta(z)$ to be a linearly independent basis set with quadratic, tanh, or sigmoid elements to satisfy $\|\zeta(z)\| \leq b_\zeta$, $\|\nabla\zeta\| = \|\partial\zeta(z)/\partial z\| \leq s_{\nabla\zeta}$, $\|\epsilon(z)\| \leq b_\epsilon$, $\|\nabla\epsilon\| = \|\partial\epsilon(z)/\partial z\| \leq s_{\nabla\epsilon}$, where b_ζ, $b_{\nabla\zeta}$, b_ϵ, and $b_{\nabla\epsilon}$ are positive constants [16].*

Substituting (25) into (22), the equation HJB is rewritten as

$$H\left(z, \mu^*, W^T\nabla\zeta\right) = z^T\Theta z + \Omega(\mu^*) + W^T\nabla\zeta(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\mu^*) - \epsilon_H = 0, \quad (26)$$

where $\epsilon_H = -\nabla\epsilon(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)$.

By the weight vector (24) is unknown, $\mathcal{J}^*(z)$ is approximated by

$$\hat{\mathcal{J}}(z) = \hat{W}^T\zeta(z), \quad (27)$$

where $\hat{W} \in \mathbb{R}^H$. The control law (23) is rewritten as

$$\hat{\mu} = -\varrho\tanh\left(\frac{1}{2\varrho}R^{-1}g^T_{\varphi v}(\chi_\varphi, \chi_v)\nabla\zeta^T\hat{W}\right). \quad (28)$$

Substituting (27) into (22), the HJB equation (22) is rewritten as

$$\hat{H}\left(z, \hat{\mu}, \hat{W}^T\nabla\zeta\right) = z^T\Theta z + \Omega(\hat{\mu}) + \hat{W}^T\nabla\zeta(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu}) = e_1, \quad (29)$$

where $e_1 = -\tilde{W}^T\nabla\zeta(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu}) + \epsilon_H$ and $\tilde{W} = W - \hat{W}$.

The weight update law $\hat{W}$ is designed to minimize the squared error of $E = \frac{1}{2}e_1^Te_1$. Apply the algorithm Normalized gradient descent $\dot{\hat{W}} = -\alpha_1\frac{\partial E}{\partial \hat{W}}$ and the technique of experience replay (namely concurrent learning (CL)) [11, 19], the weight update law $\hat{W}$ is proposed as follows:

- If $z^T(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu}) < 0$, then

$$\begin{aligned}\dot{\hat{W}} = &-\alpha_1\frac{\eta}{(\eta^T\eta+1)^2}\left(\eta^T\hat{W} + \mathrm{z}^T\Theta z + \Omega(\hat{\mu})\right)\\ &-\alpha_1\sum\nolimits_{i=1}^{P}\frac{\eta(t_i)}{(\eta(t_i)^T\eta(t_i)+1)^2}\left(\eta(t_i)^T\hat{W} + \mathrm{z}(t_i)^T\Theta z(t_i) + \Omega(\hat{\mu}(t_i))\right).\end{aligned} \quad (30)$$

- If $z^T(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu}) \geq 0$, then

$$\begin{aligned}\dot{\hat{W}} = &-\alpha_1\frac{\eta}{(\eta^T\eta+1)^2}\left(\eta^T\hat{W} + \mathrm{z}^T\Theta z + \Omega(\hat{\mu})\right)\\ &-\alpha_1\sum\nolimits_{i=1}^{P}\frac{\eta(t_i)}{(\eta(t_i)^T\eta(t_i)+1)^2}\left(\eta(t_i)^T\hat{W} + \mathrm{z}(t_i)^T\Theta z(t_i) + \Omega(\hat{\mu}(t_i))\right) + \alpha_2\frac{1}{2}\nabla\zeta Gz,\end{aligned} \quad (31)$$

where $\eta = \nabla\zeta(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu})$, $G = g_{\varphi v}(\chi_\varphi, \chi_v)R^{-1}g^T_{\varphi v}(\chi_\varphi, \chi_v)$, $\alpha_1 > 0$, $\alpha_2 > 0$. The past data is recorded and stored in $\{\eta(t_i)\}_{i=1}^P$, $\{r(t_i)\}_{i=1}^P$, where $r(t_i) = \mathrm{z}^T\Theta z + \Omega(\hat{\mu})$. Note that $\{\eta(t_i)\}_{i=1}^P$ must be linearly independent, i.e., rank $(\eta(t_1), \eta(t_2), ..., \eta(t_P)) = H$, the number of selected is $P \geq H$ [19].

Theorem 1. *Consider the error dynamics to be defined by* (12) *with input constraint, the HJB equation is given by* (22)*, and the saturated optimal control law is given by* (23)*. Let the saturated optimal control law is approximated by NN* (28)*, where the NN weights are tuned online by* (30) *and* (31)*. Then, the tracking errors and the approximation error are UUB stable. Furthermore, the value function* (27) *and the control law* (28) *converge to the near-optimal value, i.e.,* $\left\|\mathcal{J}^*(z) - \hat{\mathcal{J}}(z)\right\| \leq b_{\mathcal{J}}$*,* $\|\mu^* - \hat{\mu}\| \leq b_\mu$*, where* $b_{\mathcal{J}}$ *and* b_μ *are small positive constants.*

Proof. Consider the Lyapunov function for system (7) as follows

$$\mathcal{J}_3 = \frac{1}{2}\alpha_2 z^T z + \frac{1}{2}\text{trace}\left(\tilde{W}^T\tilde{W}\right) = \mathcal{J}_{31} + \mathcal{J}_{32}. \tag{32}$$

Taking the derivative of $\mathcal{J}_3$ with respect to time, one obtains

$$\dot{\mathcal{J}}_3 = \alpha_2 z^T \dot{z} + \tilde{W}^T\dot{\tilde{W}} = \alpha_2 z^T(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu}) + \tilde{W}^T\dot{\tilde{W}}. \tag{33}$$

Firstly, the tuning law (30) is considered. By $\alpha_2 z^T(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu}) < 0$, it is easy to see that $\exists\lambda_0 > 0$ such that

$$\alpha_2 z^T(\bar{F}_{\varphi v}(z_\varphi, z_v) + g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu}) < -\alpha_2\lambda_0 \|z\|. \tag{34}$$

The weight approximation error dynamics can be written as

$$\dot{\tilde{W}} = -\alpha_1\xi\left(\eta^T\tilde{W} - \epsilon_H\right) - \alpha_1\sum_{i=1}^{P}\xi(t_i)\left(\eta^T(t_i)\tilde{W} - \epsilon_H(t_i)\right), \tag{35}$$

where $\xi = \frac{\eta}{(\eta^T\eta+1)^2}$. Taking the derivative $\mathcal{J}_{32}$ along (35), one obtains

$$\dot{\mathcal{J}}_{32} = -\alpha_1\tilde{W}^T\Gamma\tilde{W} + \alpha_1\tilde{W}^T\left(\xi\epsilon_H + \sum_{i=1}^{P}\xi(t_i)\epsilon_H(t_i)\right), \tag{36}$$

where $\Gamma = \xi\eta^T + \sum_{i=1}^{P}\xi(t_i)\eta^T(t_i) > 0$. Using Young's inequality, we have

$$\begin{aligned}\dot{\mathcal{J}}_{32} &\leq -\lambda_{min}(\Gamma)\left\|\tilde{W}\right\|^2 + \left\|\tilde{W}\right\|^2\xi\eta^T + \frac{\alpha_1^2\epsilon_H}{4} + \left\|\tilde{W}\right\|^2\sum_{i=1}^{P}\xi(t_i)\eta^T(t_i) + \frac{\alpha_1^2 P\epsilon_H^2}{4}\\ &\leq -(\alpha_1 - 1)\lambda_{min}(\Gamma)\left\|\tilde{W}\right\|^2 + \frac{{\alpha_1}^2}{4}(P+1)\epsilon_{Hm}^2,\end{aligned} \tag{37}$$

where $\lambda(.)$ is an eigenvalue of the matrix and ϵ_{Hm} is the upper bound of ϵ_H. From (34) and (37), we have

$$\dot{\mathcal{J}}_3 \leq -\alpha_2\lambda_0\|z\| - \psi_1\left\|\tilde{W}\right\|^2 + \psi_2, \tag{38}$$

where $\psi_1 = (\alpha_1 - 1)\lambda_{min}(\Gamma)$, $\psi_2 = \frac{{\alpha_1}^2}{4}(P+1)\epsilon_{Hm}^2$ and $\alpha_1 > 1$. The Lyapunov derivative is negative if

$$\|z\| > \frac{\psi_2}{\alpha_2\lambda_0} = \bar{b}_z, \tag{39}$$

$$\left\|\tilde{W}\right\| > \sqrt{\frac{\psi_2}{\psi_1}} = \bar{b}_{\tilde{W}}. \tag{40}$$

Next, the tuning law (31) is considered. $\dot{\mathcal{J}}_{31}$ is rewritten as

$$\dot{\mathcal{J}}_{31} = \alpha_2 z^T \bar{F}_{\varphi v}(z_\varphi, z_v) + \alpha_2 z^T g_{\varphi v}(\chi_\varphi, \chi_v)\hat{\mu} - \frac{1}{2}\alpha_2 z^T G \nabla\zeta^T \hat{W}. \tag{41}$$

According to Property 1, $\alpha_2 z^T \bar{F}_{\varphi v}(z_\varphi, z_v)$ is bounded by $(-\psi_3 \|z\|^2)$, i.e.,

$$\alpha_2 z^T \bar{F}_{\varphi v}(z_\varphi, z_v) < -\psi_3 \|z\|^2, \tag{42}$$

where $\psi_3 = \alpha_2 b_{\bar{F}}$, $b_{\bar{F}}$ is the upper bound of $\bar{F}_{\varphi v}(z_\varphi, z_v)$. Then, $\dot{\mathcal{J}}_3$ becomes

$$\dot{\mathcal{J}}_3 \le -\psi_3 \left(\|z\| - \frac{\psi_4}{2\psi_3} \right)^2 - \psi_1 \left\| \tilde{W} \right\|^2 + \psi_2 + \frac{\psi_4^2}{4\psi_3^2}, \tag{43}$$

where $\psi_4 = \frac{1}{2}\alpha_2 b_G b_{\nabla\zeta} \|W\|$, $\|G\| \le b_G$. The Lyapunov derivative is negative if

$$\|z\| > \sqrt{\frac{4\psi_3^2 \psi_2 + \psi_4^2}{4\psi_3^3}} + \frac{\psi_4}{2\psi_3} = \bar{\bar{b}}_z, \tag{44}$$

$$\left\| \tilde{W} \right\| > \sqrt{\frac{4\psi_3^2 \psi_2 + \psi_4^2}{4\psi_3^2 \psi_1}} = \bar{\bar{b}}_{\tilde{W}}. \tag{45}$$

From (39), (40), (44), and (45), It can be concluded that if $\|z\|$ or $\left\| \tilde{W} \right\|$ exceeds a compact set b_z or $b_{\tilde{W}}$, then $\dot{\mathcal{J}}_3 < 0$, where $b_z = \max\left(\bar{b}_z, \bar{\bar{b}}_z \right)$ and $b_{\tilde{W}} = \max\left(\bar{b}_{\tilde{W}}, \bar{\bar{b}}_{\tilde{W}} \right)$. Therefore, the tracking and NN errors are UUB stable [9].

From (23), (24), (27), and (28), it can be seen that $\hat{\mathcal{J}}(z)$, $\hat{\mu}$ converge to the near-optimal values, i.e., $\left\| \mathcal{J}^*(z) - \hat{\mathcal{J}}(z) \right\| \le b_{\tilde{W}} b_{\nabla\zeta} + b_\epsilon = b_J$, and $\|\mu^* - \hat{\mu}\| \le \lambda_{min}(R)\, g_{\varphi v \max} \left(b_{\tilde{W}} b_{\nabla\zeta} + b_{\nabla\epsilon} \right) = b_\mu$, with $b_{\mathcal{J}} \ge 0$, $b_\mu \ge 0$. One chooses α_1 and α_2 appropriately to achieve the desired convergence quality.

The proof is completed. ■

4. SIMULATIONS

In this section, the efficiency of the saturated optimal tracking controller (SOTC) is verified by simulation. The updating of the weights of the NN using only PE (persistence of excitation) is performed to show the effect of using CL.

Consider a two-link robot manipulator [12] with saturation torques, the matrices of the dynamics equation are

$$\begin{aligned} D(\varphi) &= \begin{bmatrix} \varrho_1 + 2\varrho_3 \cos(\varphi_2) & \varrho_2 + \varrho_3 \cos(\varphi_2) \\ \varrho_2 + \varrho_3 \cos(\varphi_2) & \varrho_2 \end{bmatrix}, \\ H(\varphi, \dot{\varphi}) &= \begin{bmatrix} -\varrho_3 \sin(\varphi_2)\dot{\varphi}_2 & -\varrho_3 \sin(\varphi_2)(\dot{\varphi}_1 + \dot{\varphi}_2) \\ \varrho_3 \sin(\varphi_2)\dot{\varphi}_1 & 0 \end{bmatrix}, G(\varphi) = \begin{bmatrix} 8.45 \tanh(\dot{\varphi}_1) \\ 2.35 \tanh(\dot{\varphi}_2) \end{bmatrix}, \end{aligned} \tag{46}$$

where $\varrho_1 = 3.473$ kgm^2, $\varrho_2 = 0.196$ kgm^2, $\varrho_3 = 0.242$ kgm^2, the control input vector $\tau = [\tau_1, \tau_2]^T$ is bounded by $\|\tau\| \le 4$ N.m. Let $E = (X, Y)$ be the coordinates of the manipulator's end-effector in the workspace and $E_d = (X_d, Y_d)$ be the desired coordinates.

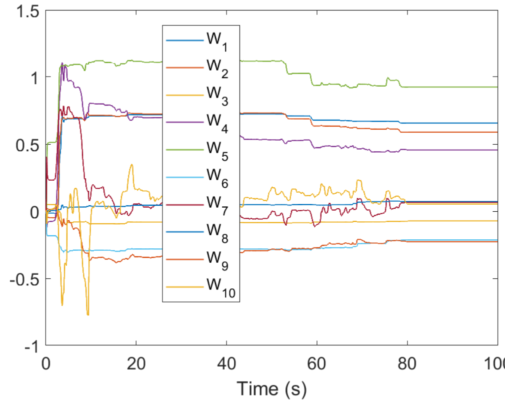

Figure 1: Evolution of NN weights in the case of using only PE.

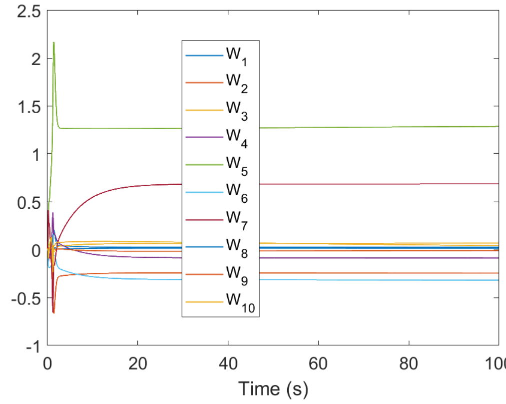

Figure 2: Evolution of NN weights in the case of using CL.

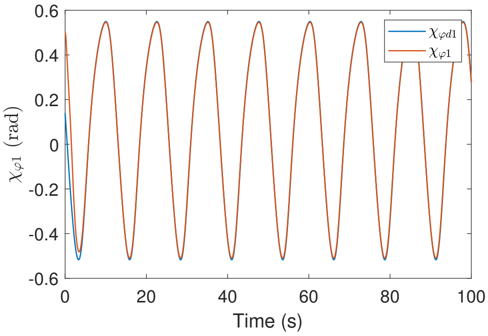

Figure 3: The position of the first joint

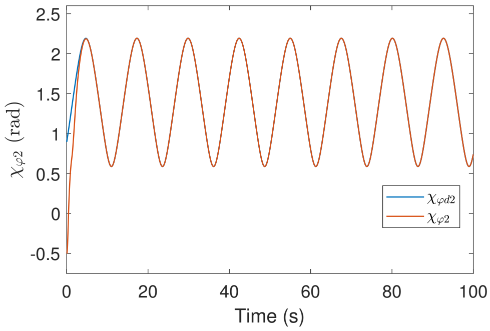

Figure 4: The position of the second joint

The desired path of E_d is given as follows

$$\begin{cases} X_d = 1 + 0.5\sin(0.5t + \frac{\pi}{2}) \\ Y_d = 1 + 0.5\cos(0.5t + \frac{\pi}{2}). \end{cases} \tag{47}$$

The reference position trajectory $\chi_{\varphi d}$ is determined by the inverse kinematics equation, i.e.,

$$\begin{cases} \chi_{\varphi d2} = \arccos \frac{X_d^2 + Y_d^2 - l_1^2 - l_2^2}{2 l_1 l_2} \\ \chi_{\varphi d1} = \arctan \frac{Y_d}{X_d} - \arctan \frac{l_2 \sin(\chi_{\varphi d2})}{l_1 + l_2 \cos(\chi_{\varphi d2})}, \end{cases} \tag{48}$$

where $l_1 = l_2 = 1$ m. The initial position and velocity values are chosen as $\varphi(0) = [0.5, -0.5]^T$, $\dot{\varphi}(0) = [0, 0]^T$.

We choose the SOTC parameters as follows: $\Lambda_1 = \Lambda_2 = \mathrm{diag}[1, 1]$, the activation function is determined as the Kronecker product quadratic polynomial basis vector, i.e., $\zeta(z) = [z_{\varphi 1}^2, z_{\varphi 1} z_{\varphi 2}, z_{\varphi 1} z_{\upsilon 1}, z_{\varphi 1} z_{\upsilon 2}, z_{\varphi 2}^2, z_{\varphi 2} z_{\upsilon 1}, z_{\varphi 2} z_{\upsilon 2}, z_{\upsilon 1}^2, z_{\upsilon 1} z_{\upsilon 2}, z_{\upsilon 2}^2]$, $\Theta = I \in \mathbb{R}^{4\times 4}$, $R = 1$, $\alpha_1 = 50$, $\alpha_2 = 0.01$, $\sigma = 0.6$, $\rho = 4$, initial values for the weights of NN are zeros. The PE condition is applied by adding noise

$$PE = 0.5(\sin^2(t)\cos(t) + \sin^2(2t)\cos(0.1t) + \sin^2(1.2t)\cos(0.5t) + \sin^5(t) + \sin^2(1.12t) + \cos(2.4t)\sin^3(2.4t)),$$

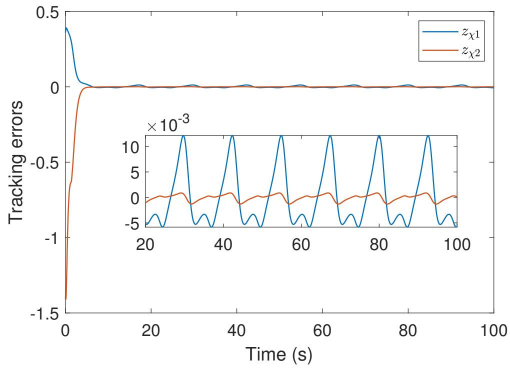

Figure 5: The position tracking errors

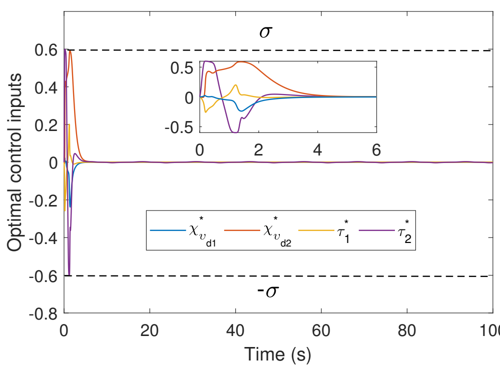

Figure 6: The saturated optimal control inputs

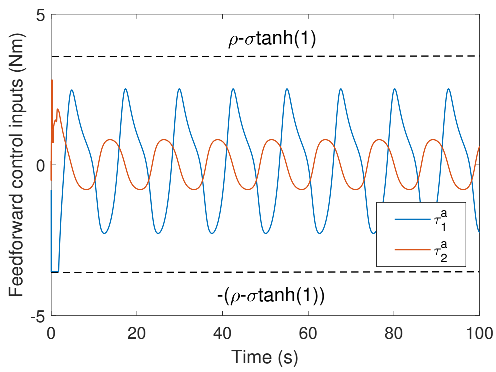

Figure 7: The feedforward control inputs

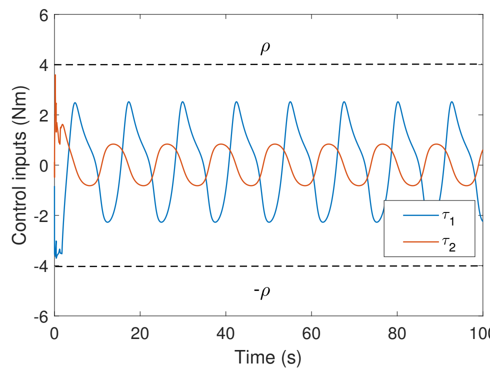

Figure 8: The control torques

to the control inputs. Simulation time is $t = 100$ s with a sampling period of $T = 0.01$ s.

We perform SOTC simulation in case of using only PE condition with $0 \leq t \leq 80$ s. In this case, the results of the weights NN are illustrated in Figure 1, showing that the weights get stuck in a local minimum [19]. It isn't easy to obtain convergent weights using only PE. On the other hand, using PE for a long time affects the actuators and is challenging to implement in practice. SOTC only uses the PE condition for the first 0.4 s, then uses CL with the data stack size as $P = 20$. Figure 2 illustrates the convergence of the NN weights of the SOTC. The weights of NN converge to the optimal values $W = [0.4303, -0.6671, -0.148, 0.0157, 1.14, -0.0216, 0.1913, 0.0074, 0.0284, 0.01572]^T$. Figures 3, 4, and 5 show the position of the joints and the tracking errors. After the algorithm converges, the tracking errors are approximately 10^{-3}. SOTC provides the saturated optimal control inputs value, as shown in Figure 6, the feedforward control inputs value, as shown in Figure 7, and the control inputs value, as shown in Figure 8. It can be seen that the control inputs are within the saturation limit. The result of the actual path of E is shown in Figure 9, where the actual path of E is determined by the forward kinematics equation, i.e.,

$$\begin{cases} X = l_1\cos(\chi_{\varphi 1}) + l_2\cos(\chi_{\varphi 1} + \chi_{\varphi 2}) \\ Y = l_1\sin(\chi_{\varphi 1}) + l_2\sin(\chi_{\varphi 1} + \chi_{\varphi 2}). \end{cases} \tag{49}$$

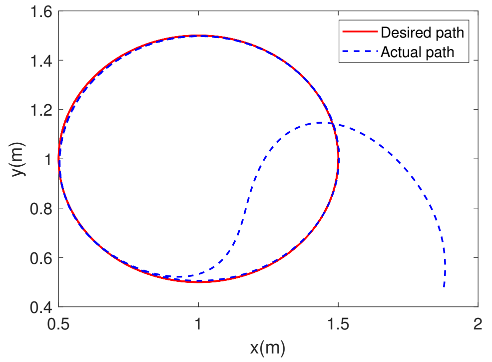

Figure 9: The actual path of E

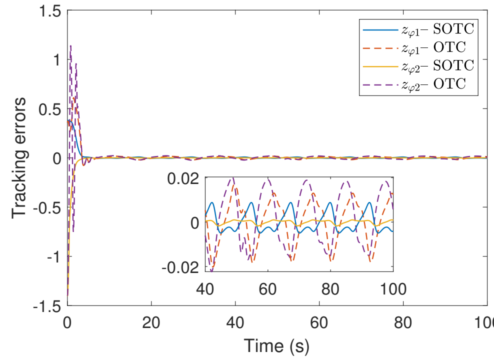

Figure 10: The tracking error of SOTC and OTC

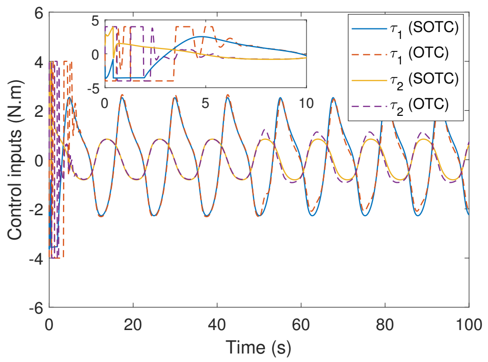

Figure 11: The control torques of SOTC and OTC

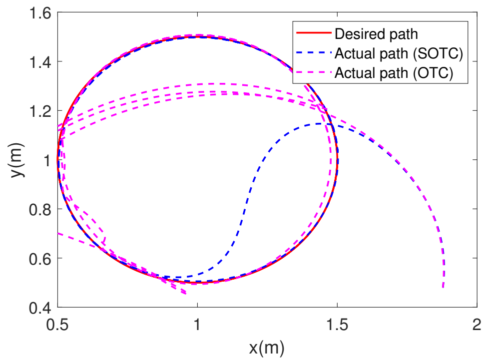

Figure 12: The actual path of E with SOTC and OTC

Next, we simulate a comparison with an optimal tracking controller (OTC) proposed in [7]. Select the parameters for the OTC as follows: $Q = I \in \mathbb{R}^{4\times 4}$, $R = \text{diag}[20, 100]$, $\Lambda = \text{diag}[200, 50]$. Comparative simulation results are shown in Figures 10, 11, and 12. The mean square error (MSE) of the positions of the control algorithms is calculated according to the following formula

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\chi_{\varphi}^{i} - \chi_{\varphi d}^{i}\right)^2, \tag{50}$$

where n is the total number of samples stored. The results of the MSE values are shown in Table 1. It can be seen that the OTC provides significant oscillation control inputs. Furthermore, OTC has a larger tracking error than SOTC. The results show that the saturation inputs problem has been resolved, and the tracking performance of the SOTC has been verified.

Table 1: Comparison of tracking error of SOTC and OTC

MSE	The first joint	The second joint
SOTC	0.0025	0.0084
OTC	0.0059	0.0159

Remark 3. *The proposed algorithm only uses a single NN instead of 2 NNs [6, 18]. The algorithm in [6, 18] uses one NN critic and one NN actor. For system (7), a NN needs ten weights and ten action functions. Therefore, the number of parameters to be stored in the proposed algorithm is 20, and the algorithm in [6, 18] is double that of the proposed algorithm. Thus, the computational cost of the algorithm in [6, 18] increases, which will reduce the convergence speed significantly compared to the proposed algorithm.*

5. CONCLUSION

This paper has solved the problem of designing the optimal tracking controller for robot manipulators with saturation torques. The feedforward control inputs have been proposed to convert the position-tracking control problem into the optimal control problem. The saturated optimal control law was built using a single NN based on the online RL algorithm. When updating the weights of the NN, the CL technique was used to relax a persistent citation condition. The proposed controller ensures that the tracking and approximate errors are UUB stable and the cost function converges to a near-optimal value. In future work, we will apply the proposed controller in an experiment.

ACKNOWLEDGMENT

We acknowledge Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for supporting this study.

REFERENCES

[1] J. Baek, W. Kwon, and C. Kang, "A new widely and stably adaptive sliding-mode control with nonsingular terminal sliding variable for robot manipulators," *IEEE Access*, vol. 8, pp. 43 443–43 454, 2020.

[2] D. Chwa and H. Kwon, "Nonlinear robust control of unknown robot manipulator systems with actuators and disturbances using system identification and integral sliding mode disturbance observer," *IEEE Access*, vol. 10, pp. 35 410–35 421, 2022.

[3] W. He, Y. Dong, and C. Sun, "Adaptive neural impedance control of a robotic manipulator with input saturation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 3, pp. 334–344, 2015.

[4] W. Jie, L. M. Cheol, K. Jaehyung, and K. H. Hee, "Fast fractional-order terminal sliding mode control with RBFNN based sliding perturbation observer for 7-dof robot manipulator," *IEEE Access*, vol. 9, pp. 67 117–67 128, 2021.

[5] W. Jie, Z. Yudong, B. Yulong, H. H. Kim, and M. C. Lee, "Trajectory tracking control using fractional-order terminal sliding mode control with sliding perturbation observer for a 7-dof robot manipulator," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 4, pp. 1886–1893, 2020.

[6] R. Kamalapurkar, H. Dinh, S. Bhasin, and W. E. Dixon, "Approximate optimal trajectory tracking for continuous-time nonlinear systems," *Automatica*, vol. 51, pp. 40–48, 2015.

[7] L. Kong, W. He, C. Yang, and C. Sun, “Robust neurooptimal control for a robot via adaptive dynamic programming,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 6, pp. 2584–2594, 2020.

[8] F. L. Lewis, D. M. Dawson, and C. T. Abdallah, *Robot Manipulator Control: Theory and Practice*. CRC Press, 2003.

[9] F. L. Lewis, K. Liu, and A. Yesildirek, “Neural net robot controller with guaranteed tracking performance,” *IEEE Transactions on Neural Networks*, vol. 6, no. 3, pp. 703–715, 1995.

[10] S. Ling, H. Wang, and P. X. Liu, “Adaptive fuzzy dynamic surface control of flexible-joint robot systems with input saturation,” *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 1, pp. 97–107, 2019.

[11] N. T. Luy, “Reinforecement learning-based optimal tracking control for wheeled mobile robot,” in *2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE, 2012, pp. 371–376.

[12] L. F. Modares H and J. ZP, “$h\infty$ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning,” *IEEE Trans Neural Netw Learn Syst*, vol. 26, no. 10, pp. 2550–62, 2015.

[13] V.-C. Nguyen, A.-T. Vo, and H.-J. Kang, “A non-singular fast terminal sliding mode control based on third-order sliding mode observer for a class of second-order uncertain nonlinear systems and its application to robot manipulators,” *IEEE Access*, vol. 8, pp. 78 109–78 120, 2020.

[14] L. Nguyen Tan, “Distributed optimal control for nonholonomic systems with input constraints and uncertain interconnections,” *Nonlinear Dynamics*, vol. 93, no. 2, pp. 801–817, 2018.

[15] L. N. Tan, “Distributed $h\infty$ optimal tracking control for strict-feedback nonlinear large-scale systems with disturbances and saturating actuators,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 11, pp. 4719–4731, 2018.

[16] L. N. Tan and T. C. Pham, “Optimal tracking control for PMSM with partially unknown dynamics, saturation voltages, torque, and voltage disturbances,” *IEEE Transactions on Industrial Electronics*, vol. 69, no. 4, pp. 3481–3491, 2021.

[17] T. N. Truong, A. T. Vo, and H.-J. Kang, “A backstepping global fast terminal sliding mode control for trajectory tracking control of industrial robotic manipulators,” *IEEE Access*, vol. 9, pp. 31 921–31 931, 2021.

[18] K. G. Vamvoudakis and F. L. Lewis, “Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem,” *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.

[19] K. G. Vamvoudakis, M. F. Miranda, and J. P. Hespanha, “Asymptotically stable adaptive–optimal control algorithm with saturating actuators and relaxed persistence of excitation,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 11, pp. 2386–2398, 2015.

[20] M. Van and S. S. Ge, “Adaptive fuzzy integral sliding-mode control for robust fault-tolerant control of robot manipulators with disturbance observer,” *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 5, pp. 1284–1296, 2020.

[21] M. Van, M. Mavrovouniotis, and S. S. Ge, “An adaptive backstepping nonsingular fast terminal sliding mode control for robust fault tolerant control of robot manipulators,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 7, pp. 1448–1458, 2018.

[22] C. Yang, D. Huang, W. He, and L. Cheng, "Neural control of robot manipulators with trajectory tracking constraints and input saturation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 9, pp. 4231–4242, 2020.

[23] H. Zargarzadeh, T. Dierks, and S. Jagannathan, "Adaptive neural network-based optimal control of nonlinear continuous-time systems in strict-feedback form," *International Journal of Adaptive Control and Signal Processing*, vol. 28, no. 3-5, pp. 305–324, 2014.

[24] S. Zeghloul, M. A. Laribi, and J.-P. Gazeau, "Robotics and mechatronics," in *Proceedings of the 4th IFToMM International Symposium on Robotics and Mechatronics*. Springer, 2015.

[25] L. Zhang, Y. Wang, Y. Hou, and H. Li, "Fixed-time sliding mode control for uncertain robot manipulators," *IEEE Access*, vol. 7, pp. 149 750–149 763, 2019.

[26] X. Zhao, B. Tao, L. Qian, and H. Ding, "Model-based actor-critic learning for optimal tracking control of robots with input saturation," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 6, pp. 5046–5056, 2020.

[27] J. Zhou, E. Liu, X. Tian, and Z. Li, "Adaptive fuzzy backstepping control based on dynamic surface control for uncertain robotic manipulator," *IEEE Access*, vol. 10, pp. 23 333–23 341, 2022.

[28] Q. Zhou, S. Zhao, H. Li, R. Lu, and C. Wu, "Adaptive neural network tracking control for robotic manipulators with dead zone," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3611–3620, 2018.

Received on February 18, 2023
Accepted on May 17, 2023