IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 March 2024, accepted 5 May 2024, date of publication 13 May 2024, date of current version 20 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3400590

RESEARCH ARTICLE

ADP-Based H_∞ Optimal Control of Robot Manipulators With Asymmetric Input Constraints and Disturbances

DIEN NGUYEN DUC[1], LAI LAI KHAC[2], AND LUY NGUYEN TAN[3], (Senior Member, IEEE)

[1]Faculty of Electrical-Automation, University of Economics-Technology for Industry (UNETI), Hanoi 100000, Vietnam
[2]Faculty of Electrical Engineering, Thai Nguyen University of Technology, Thai Nguyen 240000, Vietnam
[3]Faculty of Electric-Electronics Engineering, Ho Chi Minh City University of Technology (HCMUT), Vietnam National University Ho Chi Minh City (VNU-HCM), Ho Chi Minh City 700000, Vietnam

Corresponding author: Luy Nguyen Tan (ntanluy@hcmut.edu.vn)

ABSTRACT Trajectory tracking control for robot manipulators is an attractive topic in the research community. This is a challenging problem because robot manipulators are complex nonlinear systems. Furthermore, the tracking control performance for robot manipulators is greatly affected by input constraints and external disturbances. This paper proposes a novel H_∞ optimal controller for robot manipulators with asymmetric input constraints and external disturbances based on adaptive dynamic programming (ADP). Firstly, a strict feedback nonlinear system is used to represent the robot manipulator dynamics, and then a feedforward controller is designed to construct the tracking error dynamics. Secondly, a value function is introduced, and the Hamilton-Jacobi-Isaacs equation is made and approximated online by the principle of adaptive dynamic programming. Thirdly, the optimal control law and disturbance compensation law are determined. The stability and convergence of the proposed algorithm are analyzed by the Lyapunov technique. Finally, the controller performance is verified through simulation and experimental results with STM32F407 of STMicroelectronics.

INDEX TERMS Robot manipulators, asymmetric input constraints, adaptive dynamic programming, H_∞ optimal control.

I. INTRODUCTION

Tracking control for robot manipulators is a complex problem because robot manipulators are nonlinear systems with unknown and uncertain parameters [1], [2]. Furthermore, the manipulators are often affected by external disturbances and constrained inputs. For many decades, such a problem has attracted much attention from researchers. In [3], a neural adaptive PID control scheme was applied to robot manipulators with uncertain parameters without measuring velocity and acceleration. A self-tuning fuzzy controller was introduced in [4] for robots with uncertain parameters. A robust mode predictive controller was proposed in [5] for robots with disturbances. In [6], an adaptive fuzzy sliding controller and disturbance observer were presented. Additionally, neural adaptive sliding controllers were proposed in [7] and [8], in which radial basis function neural networks (RBFNNs) are applied to estimate unknown nonlinear functions. Overall, using neural networks [7], [9] or fuzzy models [6], the adaptive sliding controller has effectively handled the tracking control problem for robots with disturbances, but they cannot handle input saturation, leading to poor or even unstable control performance [10], [11].

He et al. [12] proposed an adaptive controller based on RBFNN to eliminate the effect of the input saturation. In [13], an adaptive fuzzy dynamic surface controller was introduced for robot manipulators, in which the input saturation was considered using a smooth function along with the mean value theorem. On the other hand, the backend systems were designed in [14], [15], and [16] to solve the input saturation. More generally, in [17], an asymmetric saturation model

The associate editor coordinating the review of this manuscript and approving it for publication was Guilin Yang.

based on a Gaussian error function was applied to handle the influence of asymmetric input saturation. An adaptive neural controller and a hyperbolic tangent function were also proposed in [18] to deal with the phenomenon. Additionally, in [19], a controller for asymmetrically constrained inputs was designed by converting a system with asymmetrically constrained inputs into a system with symmetrically constrained inputs. In [20], a constraint controller was built based on the asymmetric barrier Lyapunov function. However, the proposed algorithms mentioned above do not minimize any cost function, which means they are not optimal. Therefore, the goal of robust optimal control is important and equally challenging.

The robust optimal control problem in an H_∞ control manner is solved if the optimal control law and the external disturbance compensation law are determined. Generally, the laws depend on the solution of the Hamilton-Jacobi-Issac (HJI) equation [21]. Unfortunately, solving the HJI equation is a challenge since it is a higher-order nonlinear differential equation and has no analytical solution. Recently, ADP, a reinforcement learning (RL) version, has become a useful method for approximating online solutions of HJI equations [21], [22]. Robust optimal control algorithms were designed for robots in [23] and [24], but they used function approximations with double NNs. On the contrary, in [25], the H_∞ optimal controller was designed based on ADP combined with the zero-sum game theory and implemented with only one NN, but the input constraint problem was not mentioned. In [26], the saturated optimal tracking controller using RL was offered for a robot manipulator, but the external disturbance rejection was not mentioned. The robust optimal control algorithm was introduced in [27], where rejecting disturbances was offered but dealing with saturated inputs was not considered. In our previous work [28], we designed an H_∞ optimal controller for single-wheel robots with input constraints but did not address the asymmetrically saturated inputs. In addition, an H_∞ optimal control algorithm with asymmetrically saturated inputs was proposed in our previous work [29]; however, the algorithm only considers nonholonomic mobile agents. In [30] and [31], optimal control algorithms with asymmetrically saturated inputs were designed, but the algorithms did not consider external disturbances and were only applied to affine nonlinear systems. We designed a constrained optimal controller for robot manipulators with asymmetrically saturated inputs based on ADP in [32]; however, the algorithm did not consider external disturbances and was only performed on simulations.

To the best of our knowledge, the problem of H_∞ optimal control for robot manipulators with asymmetric input saturation and external disturbance has not been solved. Although, the H_∞ performance indexes of the manipulators can be easily obtained by combining the zero-sum game theory and ADP, the indexes cannot be robust stable under asymmetrically saturated inputs. Therefore, this paper is the first work using the ADP method to design a constrained H_∞ optimal controller (CHOC) for robot manipulators with asymmetric input constraints and external disturbances. The main contributions of this paper are shown as follows:

1) Different from the works in [24], [25], [26], and [27], we design a new feedforward controller to convert the robust optimal tracking control problem with asymmetric input saturation for robot manipulators into a robust optimal control problem for affine nonlinear systems.
2) Unlike the work in [24], we design a robust optimal control method based on ADP and game theory to determine the optimal control and disturbance compensation laws. Furthermore, the RL algorithm uses only one NN instead of two for an online estimate of the solution of the HJI equation to reduce the computational complexity.
3) Instead of using only the persistent excitation (PE) condition, we apply the concurrent learning (CL) technique to avoid affecting the actuator and ensure fast convergence.

The remainder of the article is organized as follows: Section II presents the robot manipulator's dynamics and the control objective. Section III provides the design of the feedforward controller and the CHOC. Section IV shows the simulation and experimental results. Finally, Section V gives the conclusion of the paper.

II. PROBLEM FORMULATION

A. KINEMATIC AND DYNAMIC MODEL OF ROBOT MANIPULATOR

Consider a model of an n-link robot manipulator with asymmetric input constraints and disturbances described by the following kinematic and dynamic equations [33], [34]:

$$P(x_P, y_P, z_P) = F(\delta), \tag{1}$$

$$\delta = I(x_P, y_P, z_P), \tag{2}$$

$$\mathcal{H}(\delta)\ddot{\delta} + \mathcal{C}(\delta, \dot{\delta})\dot{\delta} + \mathcal{G}(\delta) + \mathcal{F}(\dot{\delta}) = \tau - \tau_d, \tag{3}$$

where x_P, y_P, and z_P are the position values of X, Y, and Z, respectively, in the $OXYZ$ coordinate system, $P = [x_P, y_P, z_P]^T$ represents the endpoint position in the workspace, $\delta \in \mathbb{R}^{n\times 1}$ is joint variable, $F(\delta) : \mathbb{R}^n \to \mathbb{R}^3$ is a function of forward kinematics, $I(x_P, y_P, z_P) : \mathbb{R}^3 \to \mathbb{R}^n$ is a function of inverse kinematics, $\mathcal{H}(\delta) \in \mathbb{R}^{n\times n}$ is a symmetric and positive definite inertial matrix, $\mathcal{C}(\delta, \dot{\delta}) \in \mathbb{R}^{n\times n}$ is the Coriolis-centripetal matrix, $\mathcal{G}(\delta) \in \mathbb{R}^{n\times 1}$ is gravity vector, $\mathcal{F}(\dot{\delta}) \in \mathbb{R}^{n\times 1}$ is the static friction vector, $\tau_d \in \mathbb{R}^{n\times 1}$ is the disturbances vector, and $\tau = [\tau_1, \tau_2, \ldots, \tau_n]^T \in \mathbb{R}^{n\times 1}$ is the control inputs vector. τ is bounded by

$$\chi_1 \le \tau_i \le \chi_2, i = 1, 2, \cdots, n, \tag{4}$$

with $|\chi_1| \le |\chi_2|$.

Property 1: The matrices $\mathcal{H}(\delta)$, $\mathcal{C}(\delta, \dot{\delta})$, $\mathcal{G}(\delta)$, τ_d are bounded by $m_{\min} \le \|\mathcal{H}(\delta)\| \le m_{\max}$, $\|\mathcal{C}(\delta, \dot{\delta})\| \le c_{\max}$, $\|\mathcal{G}((\delta))\| \le g_{\max}$, $\|\tau_d\| \leqslant d_{\max}$, where $m_{\min}$, $m_{\max}$, $c_{\max}$, $g_{\max}$, $d_{\max}$ are positive constants.

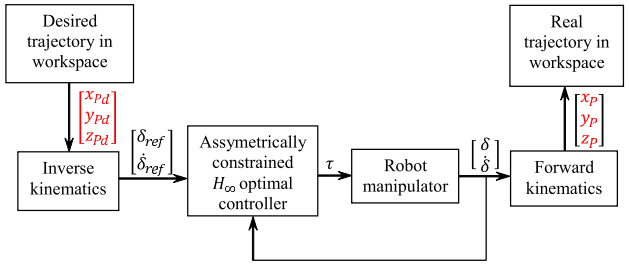

FIGURE 1. Control structure diagram.

Remark 1: Property 1 is practical by δ and $\dot{\delta}$ are bounded. Due to the physical limitations of the electrical drive circuit and motors, the control inputs are constrained. The case of asymmetric input constraints occurs when the actuator is partially reduced in efficiency, changes in the mechanical structure, and the faulty power drive circuit.

In preparation for the CHOC design, the system (3) is represented as a strict-feedback nonlinear system, i.e.,

$$\begin{cases} \dot{\delta} = \mathcal{F}_\delta(\delta) + \mathcal{G}_\delta(\delta)\psi \\ \dot{\psi} = \mathcal{F}_\psi(\delta, \psi) + \mathcal{G}_\psi(\delta, \psi)\tau + \mathcal{K}_\psi(\delta, \psi)\tau_d \end{cases} \tag{5}$$

where $\mathcal{F}_\delta(\delta) = 0_{n\times 1}$, $\mathcal{G}_\delta(\delta) = I_n$, $\psi = \dot{\delta}$, $\mathcal{F}_\psi(\delta, \psi) = -\mathcal{H}^{-1}(\delta)(\mathcal{C}(\delta, \dot{\delta})\psi + \mathcal{G}(\delta) + \mathcal{F}(\dot{\delta})) \in \mathbb{R}^{n\times 1}$, $\mathcal{G}_\psi(\delta, \psi) = \mathcal{H}^{-1}(\delta) \in \mathbb{R}^{n\times n}$, $\mathcal{K}_\psi(\delta, \psi) = -\mathcal{H}^{-1}(\delta) \in \mathbb{R}^{n\times n}$.

Remark 2: $\mathcal{F}_\psi(\delta, \psi)$, $\mathcal{G}_\psi(\delta, \psi)$, and $\mathcal{K}_\psi(\delta, \psi)$ are bounded by $\|\mathcal{F}_\psi(\delta, \psi)\| \leq m_{\min}^{-1}(c_{\max} + g_{\max})\|\psi\|$, $\|\mathcal{G}_\psi(\delta, \psi)\| \leq m_{\min}^{-1}$, $\|\mathcal{K}_\psi(\delta, \psi)\| \leq m_{\min}^{-1}$. The upper bounds of $\mathcal{F}_\psi(\delta, \psi)$, $\mathcal{G}_\psi(\delta, \psi)$, and $\mathcal{K}_\psi(\delta, \psi)$ are only intended to demonstrate stability for closed-loop system and are not used in the control law. Therefore, determining the values of the upper bounds of $\mathcal{F}_\psi(\delta, \psi)$, $\mathcal{G}_\psi(\delta, \psi)$, and $\mathcal{K}_\psi(\delta, \psi)$ is not necessary.

B. CONTROL OBJECTIVE

The controller's objective is to control the tracking of the desired trajectory in the workspace, ensuring that $\lim_{t\to\infty}\|e_P\| = \lim_{t\to\infty}\|P - P_d\| = 0$ when the external disturbances are zeros, with $P_d = [x_{Pd}, y_{Pd}, z_{Pd}]^T$ being the desired position in the workspace. In the first step, we use inverse kinematic equations to convert the desired trajectory in the workspace to joint space. The CHOC is implemented in the second step to ensure tracking the desired trajectory for robot manipulators with asymmetric input constraints in the joint space, i.e., providing $\lim_{t\to\infty}\|\delta(t) - \delta_{ref}(t)\| = 0$ when the external disturbances are zeros, where $\delta_{ref}(t) \in \mathbb{R}^{n\times 1}$ is the reference trajectory in the joint space. However, the external disturbances are non-zeros; thus, the objective is to design the control law to make the tracking errors L_2-bounded [35]. We use forward kinematic equations to convert the trajectory from joint space to workspace. These steps are described in Fig. 1.

Assumption 1: $\delta_{ref}(t)$ is smooth and bounded.

III. ASYMMETRICALLY CONSTRAINED H_∞ OPTIMAL CONTROL

This section designs an CHOC for the robot manipulator. First, the feedforward control law is designed to convert the problem of H_∞ optimal tracking control for the robot manipulator into an issue of H_∞ optimal control for an affine nonlinear system. Then, the asymmetrically constrained H_∞ optimal control law is designed based on ADP method combined with the game theory.

A. FEEDFORWARD CONTROL

In this section, we design feedforward control inputs, then transform the system (5) into an affine nonlinear system, known as an augmented system. The design steps are developed based on the backstepping technique [36].

Step 1: Define the tracking error as $e_\delta = \delta - \delta_{ref}$, $e_\psi = \psi - \psi_{ref}$. Then, the dynamics of the position tracking errors are presented as

$$\dot{e}_\delta = -\dot{\delta}_{ref} + \mathcal{G}_\delta(\delta)\psi_{ref}^* + \mathcal{G}_\delta(\delta)\psi_{ref}^a + \mathcal{G}_\delta(\delta)e_\psi, \tag{6}$$

where $\psi_{ref} = \psi_{ref}^* + \psi_{ref}^a$ are the virtual control inputs, ψ_{ref}^* are the virtual optimal control inputs, ψ_{ref}^a are the feedforward virtual control inputs designed as

$$\psi_{ref}^a = \mathcal{G}_\delta^{-1}(\delta)(\dot{\delta}_{ref} + K_1 e_\psi - K_2 e_\delta). \tag{7}$$

By substituting (7) into (6), the dynamics (6) is rewritten as

$$\begin{aligned} \dot{e}_\delta &= \mathcal{G}_\delta(\delta)\psi_{ref}^* + K_1 e_\psi - K_2 e_\delta + \mathcal{G}_\delta(\delta)e_\psi \\ &= \bar{\mathcal{F}}_\delta(e_\delta, e_\psi) + \psi_{ref}^* + e_\psi, \end{aligned} \tag{8}$$

where $\bar{\mathcal{F}}_\delta(e_\delta, e_\psi) = K_1 e_\psi - K_2 e_\delta$, K_1 and K_2 are the positive definite matrices.

Step 2: The dynamics of the angular velocity tracking errors are calculated as follows:

$$\begin{aligned} \dot{e}_\psi = -\dot{\psi}_{ref} + \mathcal{F}_\psi(\delta, \psi) + \mathcal{G}_\psi(\delta, \psi)\tau^* + \mathcal{G}_\psi(\delta, \psi)\tau^a \\ + \mathcal{K}_\psi(\delta, \psi)\tau_d, \end{aligned} \tag{9}$$

where $\tau = \tau^* + \tau^a$, τ^* are the actual optimal control inputs, τ^a are the feedforward actual control inputs designed as

$$\begin{aligned} \tau^a = \mathcal{G}_\psi^{-1}(\delta, \psi)(\dot{\psi}_{ref} + \mathcal{F}_\psi(e_\delta, e_\psi) - \mathcal{F}_\psi(\delta, \psi) - e_\delta \\ - K_3 e_\delta - K_4 \dot{e}_\delta). \end{aligned} \tag{10}$$

By substituting (10) into (9), the dynamics (9) is rewritten as

$$\dot{e}_\psi = \bar{\mathcal{F}}_\psi(e_\delta, e_\psi) + \mathcal{G}_\psi(\delta, \psi)\tau^* - e_\delta + \mathcal{K}_\psi(\delta, \psi)\tau_d, \tag{11}$$

where $\bar{\mathcal{F}}_\psi(e_\delta, e_\psi) = \mathcal{F}_\psi(e_\delta, e_\psi) - K_3 e_\delta - K_4 \dot{e}_\delta$, K_3 and K_4 are the positive definite matrices.

Lemma 1: Consider the following augmented system:

$$\dot{\mathcal{Z}} = \bar{\mathcal{F}}_{\delta\psi}(e_\delta, e_\psi) + \mathcal{G}_{\delta\psi}(\delta, \psi)\mu^* + \mathcal{K}_{\delta\psi}(\delta, \psi)d, \tag{12}$$

where $\mathcal{Z} = [e_\delta^T, e_\psi^T]^T$, $\bar{\mathcal{F}}_{\delta\psi}(e_\delta, e_\psi) = [\bar{\mathcal{F}}_\delta^T(e_\delta, e_\psi), \bar{\mathcal{F}}_\psi^T(e_\delta, e_\psi)]^T$, $\mathcal{G}_{\delta\psi}(\delta, \psi) = diag\left[\mathcal{G}_\delta(\delta), \mathcal{G}_\psi(\delta, \psi)\right]$, $\mathcal{K}_{\delta\psi}(\delta, \psi) = diag[0_{n\times n}, \mathcal{K}_\psi(\delta, \psi)]$, $\mu^* = [\psi_{ref}^{*T}, \tau^{*T}]^T$, $\mu^a = [\psi_{ref}^{aT}, \tau^{aT}]^T$, $\mu = [\psi_{ref}^T, \tau^T]^T = \mu^* + \mu^a$,

$d = [0_{1\times n}, \tau_d^T]^T$. Suppose the feedforward virtual control inputs are presented as (7), and the feedforward actual control inputs are presented as (10). Then, the H_∞ optimal control issue for the dynamics (5) and dynamics (12) are equivalent.

Proof: For dynamics (5), we select the Lyapunov function form

$$\mathfrak{J}_1 = \frac{1}{2}e_\delta^T e_\delta + \frac{1}{2}e_\psi^T e_\psi. \tag{13}$$

Taking derivative of (13) and combining (8) and (11) yields

$$\begin{aligned}\dot{\mathfrak{J}}_1 = {} & e_\delta^T \bar{\mathcal{F}}_\delta(e_\delta, e_\psi) + e_\psi^T \bar{\mathcal{F}}_\psi(e_\delta, e_\psi) + e_\delta^T \psi_{ref}^* \\ & + e_\psi^T \mathcal{G}_\psi(\delta, \psi)\tau^* + e_\psi^T \mathcal{K}_\psi(\delta, \psi)\tau_d + e_\delta^T e_\psi - e_\psi^T e_\delta.\end{aligned} \tag{14}$$

It can be seen that $e_\delta^T e_\psi = e_\psi^T e_\delta$. Thus, (14) becomes

$$\dot{\mathfrak{J}}_1 = \mathcal{Z}^T(\bar{\mathcal{F}}_{\delta\psi}(e_\delta, e_\psi) + \mathcal{G}_{\delta\psi}(\delta, \psi)\mu^* + \mathcal{K}_{\delta\psi}(\delta, \psi)d). \tag{15}$$

For dynamics (12), we select the Lyapunov function form

$$\mathfrak{J}_2 = \frac{1}{2}\mathcal{Z}^T\mathcal{Z}. \tag{16}$$

Taking derivative (16) and combining (12), one has

$$\dot{\mathfrak{J}}_2 = \mathcal{Z}^T(\bar{\mathcal{F}}_{\delta\psi}(e_\delta, e_\psi) + \mathcal{G}_{\delta\psi}(\delta, \psi)\mu^* + \mathcal{K}_{\delta\psi}(\delta, \psi)d). \tag{17}$$

It can be easily seen that, if μ^* makes the dynamics (12) stable, then $\dot{\mathfrak{J}}_2 < 0$. Through comparison of (17) and (15), $\dot{\mathfrak{J}}_2 < 0$ implies $\dot{\mathfrak{J}}_1 < 0$. It can be concluded that the dynamics (5) is also stable. Therefore, the H_∞ optimal control issue of the dynamics (5) is converted to the H_∞ optimal control issue for dynamics (12).

This completes the proof. ∎

Remark 3: The feedforward virtual control inputs and the feedforward actual control inputs are designed as (7) and (10), and constrained by

$$\begin{cases} \chi_1 - \nu_1 \tanh(1) \leqslant \psi_{refi}^a \leqslant \chi_2 - \nu_2 \tanh(1) \\ \chi_1 - \nu_1 \tanh(1) \leqslant \tau_i^a \leqslant \chi_2 - \nu_2 \tanh(1) \end{cases} \tag{18}$$

where $i = 1, 2, \cdots, n$, $|\nu_1| < |\chi_1|$, $|\nu_2| < |\chi_2|$. We can see that if ψ_{refi}^* and τ_i^* are designed to satisfy the constraint $\nu_1 \leq \psi_{refi}^* \leq \nu_2$, $\nu_1 \leq \tau_i^* \leq \nu_2$, then ν_{refi} and τ_i will be limited by $\chi_1 \leq \psi_{ref} \leq \chi_2$, $\chi_1 \leq \tau_i \leq \chi_2$.

B. H_∞ OPTIMAL CONTROL

This section designs the asymmetrically constrained H_∞ optimal control law for dynamics (12) based on the ADP method combined with the game theory [22]. To derive the controller, we give the following definitions.

Definition 1 [37]: Consider the $\dot{\mathcal{Z}} = \mathcal{F}(\mathcal{Z}, \mu)$, $\mathcal{Z} \in \mathbb{R}^n$ with $\mathcal{Z}_0$ being the equilibrium point. $\mathcal{Z}_0$ is said to be ultimately uniformly bounded (UUB) if there exists a compact set $\mho \in \mathbb{R}^n$ so that for all $\mathcal{Z}_0 \in \mho$, there exists a bound C and time $T(C, \mathcal{Z}_0)$ such that $\|\mathcal{Z} - \mathcal{Z}_0\| \leq C$ for all $t > t_0 + T$.

Definition 2 [21]: The system (12) has L_2-gain less than or equal ρ $\forall d \in L_2[0, \infty]$, $0 \leqslant T \leqslant \infty$ if the following expression is satisfied:

$$\int_0^T (\mathcal{Z}^T\Theta\mathcal{Z} + \digamma(\mu))d\tau \leqslant \rho^2 \int_0^T (d^T d)d\tau, \tag{19}$$

where Θ is the positive definite matrix, $\digamma(\mu)$ is the energy cost function and non-negative, μ is an estimate of μ^* at time t, and $\rho > 0$ is the attenuation level of the disturbance.

Based on the condition (19), the cost function is presented as

$$\mathcal{V}(\mathcal{Z}) = \int_t^\infty \kappa(\tau)d\tau, \tag{20}$$

where $\kappa(t) = \mathcal{Z}^T\Theta\mathcal{Z} + \digamma(\mu) - \rho^2 d^T d$.

By the control inputs are asymmetrically saturated, the energy cost function is chosen as [30] and [40]

$$\digamma(\mu) = (\nu_2 - \nu_1)\int_{\frac{\nu_2+\nu_1}{2}}^{\mu} \tanh^{-T}\left(\frac{2s - \nu_2 - \nu_1}{\nu_2 - \nu_1}\right) Rds, \tag{21}$$

where $R = diag(R_1, \ldots, R_n) > 0$. Defining $\varrho = (\nu_2 + \nu_1)/2$ and $\varsigma = (\nu_2 - \nu_1)/2$, (21) is rewritten as

$$\digamma(\mu) = 2\varsigma\int_{\varrho}^{\mu} \tanh^{-T}\left(\frac{s - \varrho}{\varsigma}\right) Rds. \tag{22}$$

We apply integral by parts and perform some conversion steps. We can derive that

$$\digamma(\mu) = 2\varsigma\tanh^{-T}\left(\frac{\mu - \varrho}{\varsigma}\right)R(\mu - \varsigma) + \varsigma^2\bar{R}\ln\left(\bar{1} - \left(\frac{\mu - \varrho}{\varsigma}\right)^2\right), \tag{23}$$

with $\bar{R} = [R_1, \ldots, R_n]$, $\bar{1} = [1, \ldots, 1]^T \in \mathbb{R}^{n\times 1}$.

Consider the Hamilton function, which has the following form:

$$\begin{aligned}\mathcal{H}(\mathcal{Z}, \mu, d, \mathcal{V}_\mathcal{Z}) = {} & \mathcal{Z}^T\Theta\mathcal{Z} + \digamma(\mu) - \rho^2 d^T d \\ & + \mathcal{V}_\mathcal{Z}^T\left(\bar{\mathcal{F}}_{\delta\psi} + \mathcal{G}_{\delta\psi}\mu + \mathcal{K}_{\delta\psi}d\right),\end{aligned} \tag{24}$$

with $\mathcal{V}_\mathcal{Z} = \partial\mathcal{V}(\mathcal{Z})/\partial\mathcal{Z}$. The optimal value function $\mathcal{V}^*(\mathcal{Z})$ is determined by using zero-sum game theory:

$$\mathcal{V}^*(\mathcal{Z}) = \min_\mu \max_d \int_0^\infty \kappa(\tau)d\tau. \tag{25}$$

ADP and the zero-sum game theory are used to find the optimal value $\mathcal{V}^*(\mathcal{Z})$ that satisfies the Nash condition

$$\mathcal{V}^*(\mathcal{Z}) = \min_\mu \max_d \int_0^\infty \kappa(\tau)d\tau = \max_d \min_\mu \int_0^\infty \kappa(\tau)d\tau. \tag{26}$$

Then, the saddle point (μ^*, d^*) exists, where μ^* is the optimal control law, d^* is the disturbance compensation law.

To ensure the Nash condition (26), the following Isaacs condition must be satisfied:

$$\min_{\mu}\max_{d}\mathcal{H}(\mathcal{Z},\mu,d,\mathcal{V}^*_{\mathcal{Z}})=\max_{d}\min_{\mu}\mathcal{H}(\mathcal{Z},\mu,d,\mathcal{V}^*_{\mathcal{Z}}). \tag{27}$$

The saddle point (μ^*, d^*) is determined through the stationary conditions $\partial\mathcal{H}(\mathcal{Z},\mu,d,\mathcal{V}^*_{\mathcal{Z}})/\partial\mu = 0$, $\partial\mathcal{H}(\mathcal{Z},\mu,d,\mathcal{V}^*_{\mathcal{Z}})/\partial d = 0$. Hence, we obtain μ^* and d^* as follows:

$$\mu^*=-\varsigma\tanh(\frac{1}{2\varsigma}R^{-1}\mathcal{G}^T_{\delta\psi}\mathcal{V}^*_{\mathcal{Z}})+\varrho, \tag{28}$$

$$d^*=\frac{1}{2\rho}\mathcal{K}^T_{\delta\psi}\mathcal{V}^*_{\mathcal{Z}}, \tag{29}$$

where $\mathcal{V}^*_{\mathcal{Z}}=\partial\mathcal{V}^*(\mathcal{Z})/\partial\mathcal{Z}$.

Substituting (28) and (29) into (24), we obtain the HJI equation as follows:

$$\mathcal{H}^*\mathcal{Z},\mu^*,d^*,\mathcal{V}^*_{\mathcal{Z}})=\mathcal{Z}^T\Theta\mathcal{Z}+F\left(\mu^*\right)-\rho^2d^{*T}d^*+\mathcal{V}^{*T}_{\mathcal{Z}}\left(\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\mu^*+\mathcal{K}_{\delta\psi}d^*\right)=0. \tag{30}$$

If $\mathcal{V}^*(\mathcal{Z})$ of the HJI equation (30) is found, one can determine μ^* and d^*. The minimal positive definite smooth solutions $\mathcal{V}^*(\mathcal{Z})$ always exist [21]. However, it cannot be obtained by analytical methods. Therefore, we need to estimate the solution of equation (30). With the requirement of a simple online approximation algorithm, a simple NN is applied and presented as follows:

$$\mathcal{V}^*(\mathcal{Z})=\varpi^T\zeta(\mathcal{Z})+\varepsilon(\mathcal{Z}), \tag{31}$$

$$\mathcal{V}^*_{\mathcal{Z}}=\varpi^T\frac{\partial\zeta(\mathcal{Z})}{\partial\mathcal{Z}}+\frac{\partial\varepsilon(\mathcal{Z})}{\partial\mathcal{Z}}=\varpi^T\zeta_{\mathcal{Z}}+\varepsilon_{\mathcal{Z}}, \tag{32}$$

where NN has a structure with one hidden layer and one output, $\varpi\in\mathbb{R}^{hc}$ is the weight vector, $\zeta(\mathcal{Z}):\mathbb{R}^n\to\mathbb{R}^{hc}$ is a smooth function vector and a completely independent basis set that satisfies Assumption 2, hc is a number of hidden cells, $\varepsilon(\mathcal{Z})$ is the approximation error.

Assumption 2 [39]: $\|\zeta(\mathcal{Z})\|\le s_\zeta$, $\|\zeta_{\mathcal{Z}}\|=\|\partial\zeta(\mathcal{Z})/\partial\mathcal{Z}\|\le s_{\nabla\zeta}$, $\|\varepsilon(\mathcal{Z})\|\le s_\varepsilon$, $\|\varepsilon_{\mathcal{Z}}\|=\|\partial\varepsilon(\mathcal{Z})/\partial\mathcal{Z}\|\le s_{\nabla\varepsilon}$, where s_ζ, $s_{\nabla\zeta}$, s_ε, $s_{\nabla\varepsilon}$ are positive constants.

The Hamilton function (24) becomes

$$\mathcal{H}(\mathcal{Z},\mu^*,d^*,\varpi^T\zeta_{\mathcal{Z}})=\mathcal{Z}^T\Theta\mathcal{Z}+F(\mu^*)-\rho^2d^{*T}d^*+\varpi^T\zeta_{\mathcal{Z}}\left(\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\mu^*+\mathcal{K}_{\delta\psi}d^*\right)=\varepsilon_H, \tag{33}$$

where $\varepsilon_H=-\varepsilon^T_{\mathcal{Z}}(\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\mu^*+\mathcal{K}_{\delta\psi}d^*)$.

As ϖ are unknown, the value function is given as follows:

$$\hat{\mathcal{V}}(\mathcal{Z})=\hat{\varpi}^T\zeta(\mathcal{Z}). \tag{34}$$

Using $\hat{\mathcal{V}}(\mathcal{Z})$ for (28) and (29), the approximate control laws of (28) and (29) are expressed as

$$\hat{\mu}=-\varsigma\tanh(\frac{1}{2\varrho}R^{-1}\mathcal{G}^T_{\delta\psi}\zeta^T_{\mathcal{Z}}\hat{\varpi})+\varrho. \tag{35}$$

$$\hat{d}=\frac{1}{2\rho}\mathcal{K}^T_{\delta\psi}\zeta^T_{\mathcal{Z}}\hat{\varpi}. \tag{36}$$

Similarly, the Hamiltonian function becomes

$$\mathcal{H}(\mathcal{Z},\hat{\mu},\hat{d},\hat{\varpi}^T\zeta_{\mathcal{Z}})=\mathcal{Z}^T\Theta\mathcal{Z}+F(\hat{\mu})-\rho^2\hat{d}^T\hat{d}+\hat{\varpi}^T\zeta_{\mathcal{Z}}(\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\hat{\mu}+\mathcal{K}_{\delta\psi}\hat{d})=\varepsilon_e, \tag{37}$$

where ε_e is the Hamilton function approximation error. From (33) and (37), we can obtain

$$\varepsilon_e=-\tilde{\varpi}^T\zeta_{\mathcal{Z}}(\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\hat{\mu}+\mathcal{K}_{\delta\psi}\hat{d})+\varepsilon_H, \tag{38}$$

where $\tilde{\varpi}=\varpi-\hat{\varpi}$. If we propose a parameter tuning law to minimize $E=\frac{1}{2}\varepsilon^T_e\varepsilon_e$, then $\hat{\varpi}\to\varpi$. Applying the algorithm gradient descent combined with the CL technique, the parameter tuning law designed as

$$\dot{\hat{\varpi}}=-\alpha\frac{\Psi}{(\Psi^T\Psi+1)^2}(\Psi^T\hat{\varpi}+\kappa(t))-\alpha\sum_{i=1}^{P}\frac{\Psi(t_i)}{(\Psi(t_i)^T\Psi(t_i)+1)^2}(\Psi(t_i)^T\hat{\varpi}+\kappa(t_i)), \tag{39}$$

where $\kappa(t)=\mathcal{Z}^T\Theta\mathcal{Z}+F(\hat{\mu})-\rho^2\hat{d}^T\hat{d}$, $\Psi=\zeta_{\mathcal{Z}}(\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\hat{\mu}+\mathcal{K}_{\delta\psi}\hat{d})$, $\alpha>0$ is the learning rate. We use the CL technique in (39) to relax the PE condition and ensure convergence fast. $\Psi(t_i)$, $\kappa(t_i)$ are stored in $\{\Psi(t_i),\kappa(t_i)\}^P_{i=1}$, with $\{\Psi(t_i)\}^P_{i=1}$ must be linearly independent, i.e. $rank(\Psi(t_1),\Psi(t_2),\ldots,\Psi(t_P))=P$.

The approximate error dynamics of the weights can be determined as follows:

$$\dot{\tilde{\varpi}}=-\alpha\bar{\Psi}(\bar{\Psi}^T\tilde{\varpi}-\bar{\varepsilon}_e)-\alpha\sum_{i=1}^{P}\bar{\Psi}(t_i)(\bar{\Psi}^T(t_i)\tilde{\varpi}-\bar{\varepsilon}_e(t_i)), \tag{40}$$

with $\bar{\Psi}=\Psi/(\Psi^T\Psi+1)$, $\|\bar{\varepsilon}_e(s)\|=\|\varepsilon_H/(1+\Psi^T\Psi)\|\le s_{\bar{\varepsilon}}$, where $s_{\bar{\varepsilon}}$ is a positive constant.

C. PROOF OF STABILITY AND CONVERGENCE

The following theorem analyzes the stability and convergence of the CHOC.

Theorem 1: Consider the tracking error dynamics (12) with asymmetrically constrained inputs and disturbances. Let Assumptions 1-2 be satisfied. Assuming the cost function is defined in (20), the optimal control law is determined by (35), and the disturbance compensation law is defined by (36), where the on-line adjustment of the values is determined by (39). Then, the tracking and approximation errors are UUB.

Proof: Select a Lyapunov function as

$$\mathfrak{J}_3=\underbrace{\mathcal{V}^*(\mathcal{Z})}_{\mathfrak{J}_{31}}+\underbrace{\frac{1}{2}trace\left(\tilde{\varpi}^T\tilde{\varpi}\right)}_{\mathfrak{J}_{32}}. \tag{41}$$

Taking derivative $\mathfrak{J}_{31}$ along trajectories of $\dot{\mathcal{Z}}=\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\hat{\mu}+\mathcal{K}_{\delta\psi}\hat{d}$, we can get

$$\dot{\mathfrak{J}}_{31}=\mathcal{V}^{*T}_{\mathcal{Z}}\dot{\mathcal{Z}}=\mathcal{V}^{*T}_{\mathcal{Z}}(\bar{\mathcal{F}}_{\delta\psi}+\mathcal{G}_{\delta\psi}\hat{\mu}+\mathcal{K}_{\delta\psi}\hat{d})$$

$$= \mathcal{V}_{\mathcal{Z}}^{*T}(\bar{\mathcal{F}}_{\delta\psi} + \mathcal{G}_{\delta\psi}\mu^* + \mathcal{K}_{\delta\psi}d^*) + \mathcal{V}_{\mathcal{Z}}^{*T}\mathcal{G}_{\delta\psi}(\hat{\mu} - \mu^*) + \mathcal{V}_{\mathcal{Z}}^{*T}\mathcal{K}_{\delta\psi}(\hat{d} - d^*). \quad (42)$$

Observing (30) and (32), (42) becomes

$$\dot{\mathfrak{J}}_{31} = -\mathcal{Z}^T\Theta\mathcal{Z} - \digamma\left(\mu^*\right) + \rho^2 d^{*T}d^* + \mathcal{G}_{\delta\psi}\left(\zeta_{\mathcal{Z}}^T\varpi + \varepsilon_{\mathcal{Z}}\right)(\hat{\mu} - \mu^*) + \mathcal{K}_{\delta\psi}\left(\zeta_{\mathcal{Z}}^T\varpi + \varepsilon_{\mathcal{Z}}\right)(\hat{d} - d^*). \quad (43)$$

Using Young's inequality to (43), we can derive that

$$\dot{\mathfrak{J}}_{31} \leq -\mathcal{Z}^T\Theta\mathcal{Z} - \digamma\left(\mu^*\right) + \rho^2 d^{*T}d^* + \frac{1}{4}\left\|\mathcal{G}_{\delta\psi}\left(\zeta_{\mathcal{Z}}^T\varpi + \varepsilon_{\mathcal{Z}}\right)\right\|^2 + \left\|\hat{\mu} - \mu^*\right\|^2 + \frac{1}{4}\left\|\mathcal{K}_{\delta\psi}\left(\zeta_{\mathcal{Z}}^T\varpi + \varepsilon_{\mathcal{Z}}\right)\right\|^2 + \left\|\hat{d} - d^*\right\|^2. \quad (44)$$

On the other hand, we can gain $\mathcal{Z}^T\Theta\mathcal{Z} \geq \lambda_{\min}(\Theta)\|\mathcal{Z}\|^2$, with $\lambda(.)$ is an eigenvalue of the matrix. Then, (44) is rewritten as follows:

$$\dot{\mathfrak{J}}_{31} \leqslant -\lambda_{\min}(\Theta)\|\mathcal{Z}\|^2 + \rho^2\left\|\hat{d}\right\|^2 + \frac{1}{4}\left\|\mathcal{G}_{\delta\psi}\left(\zeta_z^T\varpi + \varepsilon_{\mathcal{Z}}\right)\right\|^2 + \left\|\hat{\mu} - \mu^*\right\|^2 + \frac{1}{4}\left\|\mathcal{K}_{\delta\psi}\left(\zeta_{\mathcal{Z}}^T\varpi + \varepsilon_{\mathcal{Z}}\right)\right\|^2 + \left\|\hat{d} - d^*\right\|^2 - \digamma\left(\mu^*\right). \quad (45)$$

Applying inequality $(x+y)^2 \leq 2\left(x^2+y^2\right)$, we can derive that

$$\dot{\mathfrak{J}}_{31} \leq -\lambda_{\min}(\Theta)\|\mathcal{Z}\|^2 + \rho^2\left\|\hat{d}\right\|^2 + \frac{1}{2}(\left\|\mathcal{G}_{\delta\psi}\zeta_{\mathcal{Z}}^T\varpi\right\|^2 + \left\|\mathcal{G}_{\delta\psi}\varepsilon_{\mathcal{Z}}\right\|^2) + \left\|2\varsigma\tanh(\hat{\xi})\right\|^2 + \left\|2\varsigma\tanh(\xi^*)\right\|^2 + \frac{1}{2}\left\|\mathcal{K}_{\delta\psi}\right\|^2\|\zeta_{\mathcal{Z}}\|^2\|\varpi\|^2 + \frac{1}{2}\left\|\mathcal{G}_{\delta\psi}\varepsilon_{\mathcal{Z}}\right\|^2 + \frac{1}{2\rho^2}\left\|\mathcal{K}_{\delta\psi}\right\|^2\|\zeta_{\mathcal{Z}}\|^2\|\tilde{\varpi}\|^2 + \frac{1}{2\rho^2}\left\|\mathcal{K}_{\delta\psi}\right\|^2\|\varepsilon_{\mathcal{Z}}\|^2 - \digamma\left(\mu^*\right), \quad (46)$$

where $\hat{\xi} = \frac{1}{2\varsigma}R^{-1}\mathcal{G}_{\delta\psi}^T\zeta_{\mathcal{Z}}^T\hat{\varpi}$, $\xi^* = \frac{1}{2\varsigma}R^{-1}\mathcal{G}_{\delta\psi}^T(\zeta_{\mathcal{Z}}^T\varpi + \varepsilon_{\mathcal{Z}})$. We can easily deduce that $\digamma(\mu^*) \geq 0$, (46) becomes

$$\dot{\mathfrak{J}}_{31} \leq -\lambda_{\min}(\Theta)\|\mathcal{Z}\|^2 + \eta_2\|\tilde{\varpi}\|^2 + \eta_1, \quad (47)$$

where $\eta_1 = \frac{1}{2}\mathcal{G}_{\delta\psi\max}^2(s_{\nabla\zeta}^2 s_{\varpi}^2 + s_{\varepsilon}^2) + 8\varsigma^2 + \frac{1}{2}\mathcal{K}_{\delta\psi\max}^2(\frac{3}{2}s_{\nabla\zeta}^2 s_{\varpi}^2 + s_{\varepsilon}^2 + \frac{1}{2\rho^2}s_{\varepsilon}^2)$, $\eta_2 = \frac{1}{2\rho^2}\mathcal{K}_{\delta\psi\max}^2 s_{\nabla\varpi}^2$, $\|\varpi\| \leq s_{\varpi}$, $s_{\varpi} > 0$, $\mathcal{G}_{\delta\psi\max} = \max\{1, m_{\min}^{-1}\}$, $\mathcal{K}_{\delta\psi\max} = m_{\min}^{-1}$.

Taking derivative $\mathfrak{J}_{32}$ along (40) we have

$$\dot{\mathfrak{J}}_{32} = -\alpha\tilde{\varpi}^T\Xi\tilde{\varpi} + \alpha\tilde{\varpi}^T(\bar{\Psi}\bar{\varepsilon}_e + \sum_{i=1}^{P}\bar{\Psi}(t_i)\bar{\varepsilon}_e(t_i)), \quad (48)$$

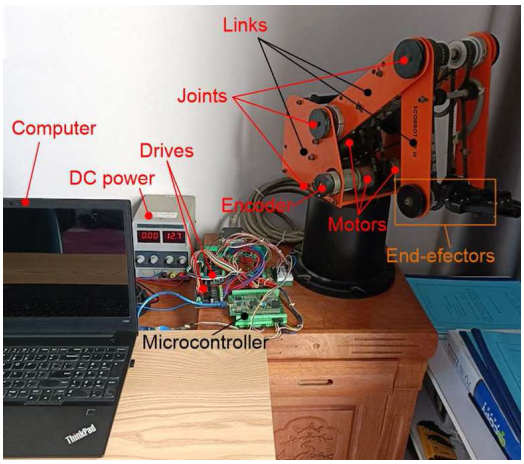

FIGURE 2. Actual Scorbot-ER.

where $\Xi = \bar{\Psi}\bar{\Psi}^T + \sum_{i=1}^{P}\bar{\Psi}(t_i)\bar{\Psi}^T(t_i) > 0$. Applying Young's inequality to (48), we can deduce that

$$\dot{\mathfrak{J}}_{32} \leq -(\alpha - 1)\lambda_{min}(\Xi)\|\tilde{\varpi}\|^2 + \frac{\alpha^2}{4}(P+1)s_{\bar{\varepsilon}}^2. \quad (49)$$

Observing (47) and (49), $\dot{\mathfrak{J}}_3$ becomes

$$\dot{\mathfrak{J}}_3 \leq -\lambda_{\min}(\Theta)\|\mathcal{Z}\|^2 - \eta_3\|\tilde{\varpi}\|^2 + \eta_4, \quad (50)$$

where $\eta_3 = (\alpha - 1)\lambda_{min}(\Xi) - \eta_2$, $\eta_4 = \eta_1 + \frac{\alpha^2}{4}(P+1)s_{\bar{\varepsilon}}^2$, $\alpha > \frac{\eta_2}{\lambda_{min}(\Xi)} + 1$.

$\dot{\mathfrak{J}}_3 < 0$, if and only if

$$\begin{cases} \|\mathcal{Z}\| > \sqrt{\dfrac{\eta_4}{\lambda_{\min}(\Theta)}} = s_{\mathcal{Z}} \\ \|\tilde{\varpi}\| > \sqrt{\dfrac{\eta_4}{\eta_3}} = s_{\tilde{\varpi}} \end{cases} \quad (51)$$

It can be concluded that if the tracking errors $\|\mathcal{Z}\|$ or the approximation errors $\|\tilde{\varpi}\|$ exceed stable regions, then $\dot{\mathfrak{J}}_3 < 0$, i.e., the tracking errors or the approximation errors are pulled inside the stable regions. In other words, the tracking and approximation errors are UUB [37] (see Definition 1).

This completes the proof. ■

IV. SIMULATION AND EXPERIMENT

This section verifies the performance of the CHOC through simulation and experiment. The results are compared with the robust adaptive controller (RAC) [8].

Consider the Scorbot-ER robotic manipulator [38], illustrated in Figs. 2, 3, where $l_1 = 0.35$m, $l_e = 0.025$m, $l_2 = 0.222$m, $l_3 = 0.222$m. The position of P is determined by

$$P = \begin{bmatrix} x_P \\ y_P \\ z_P \end{bmatrix} = \begin{bmatrix} (l_3\cos\delta_3 + l_2\cos\delta_2 + l_e)\cos\delta_1 \\ (l_3\cos\delta_3 + l_2\cos\delta_2 + l_e)\sin\delta_1 \\ l_3\sin\delta_3 + l_2\sin\delta_2 + l_1 \end{bmatrix}. \quad (52)$$

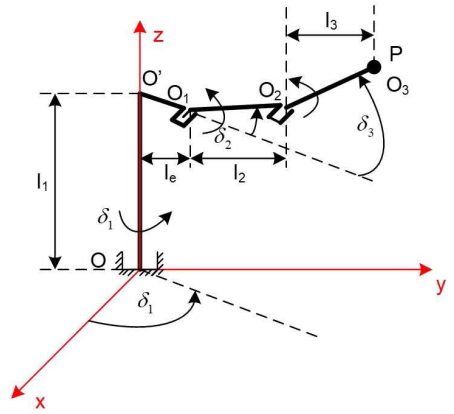

FIGURE 3. The scheme of Scorbot-ER.

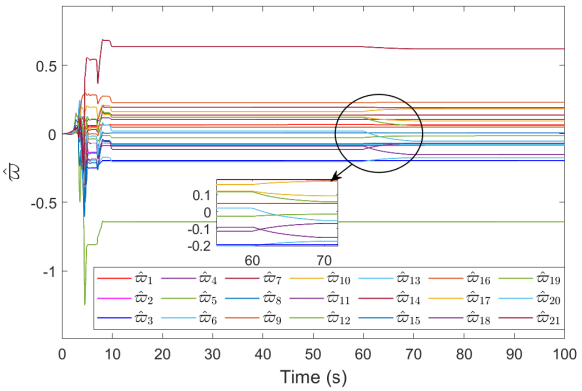

FIGURE 4. Evolution of NN weights.

The matrices of the components in the dynamics equation have the following form:

$$\mathcal{H}(\delta) = \begin{bmatrix} \mathcal{H}_{11} & 0 & 0 \\ 0 & \sigma_6 & l_2\sigma_2\cos(\delta_3-\delta_2) \\ 0 & l_2\sigma_2\cos(\delta_3-\delta_2) & \sigma_7 \end{bmatrix}, \tag{53}$$

$$\mathcal{C}(\delta,\dot{\delta}) = \begin{bmatrix} \Delta_1\dot{\delta}_2+\Delta_2\dot{\delta}_3 & \Delta_1\dot{\delta}_1 & \Delta_2\dot{\delta}_1 \\ -\Delta_1\dot{\delta}_1 & 0 & -\Delta_3\dot{\delta}_3 \\ -\Delta_2\dot{\delta}_1 & \Delta_3\dot{\delta}_2 & 0 \end{bmatrix}, \tag{54}$$

$$\mathcal{F}(\dot{\delta}) = \begin{bmatrix} \sigma_8\dot{\delta}_1+\sigma_{11}\,\mathrm{sgn}(\dot{\delta}_1) \\ \sigma_9\dot{\delta}_2+\sigma_{12}\,\mathrm{sgn}(\dot{\delta}_2) \\ \sigma_{10}\dot{\delta}_3+\sigma_{13}\,\mathrm{sgn}(\dot{\delta}_3) \end{bmatrix}, \tag{55}$$

$$\mathcal{G}(\delta) = \begin{bmatrix} 0 & \sigma_1 g\cos\delta_2 & \sigma_2 g\cos\delta_3 \end{bmatrix}^T, \tag{56}$$

where $\mathcal{H}_{11} = \mathcal{H}_{111}+\mathcal{H}_{112}+\mathcal{H}_{113}+\mathcal{H}_{114}$, $\mathcal{H}_{111} = 2\sigma_1 l_e\cos\delta_2$, $\mathcal{H}_{112} = 2\sigma_2(l_e+l_2\cos\delta_2)\cos\delta_3$, $\mathcal{H}_{113} = 0.5\sigma_3\cos(2\delta_2)$, $\mathcal{H}_{114} = 0.5\sigma_4\cos(2\delta_3)+\sigma_5$, $\Delta_1 = -(\sigma_1 l_e\sin\delta_2+\sigma_2 l_2\sin\delta_2\cos\delta_3+0.5\sigma_3\sin(2\delta_2))$, $\Delta_2 = -(\sigma_2(l_e+l_2\cos\delta_2)\sin\delta_3+0.5\sigma_4\sin(2\delta_3))$, $\Delta_3 = l_2\sigma_2\sin(\delta_3-\delta_2)$, $\sigma_1=\sigma_4=\sigma_7=0.006$, $\sigma_2=0.002$, $\sigma_3=\sigma_5=\sigma_6=0.011$, $\sigma_8=\sigma_9=\sigma_{10}=0.52$, $\sigma_{11}=0.019, \sigma_{12}=\sigma_{13}=0.018, \delta(0)=[0.01,0.01,0.01]^T$, $\dot{\delta}(0)=[0,0,0]^T$, $\chi_1=-0.8$, $\chi_2=0.9$. The desired

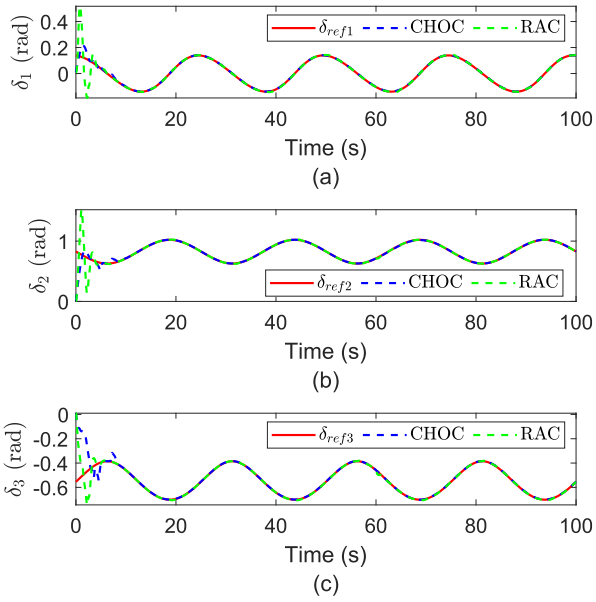

FIGURE 5. Trajectory tracking of CHOC and RAC in joint space: (a) Position δ_1; (b) Position δ_2; (c) Position δ_3.

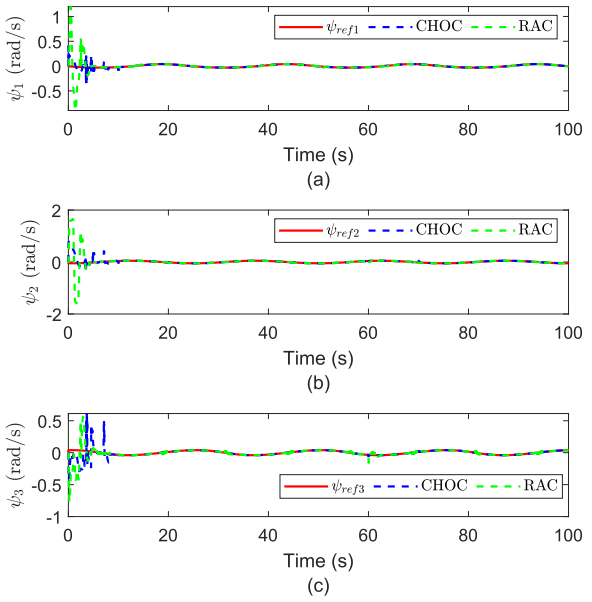

FIGURE 6. Angular velocity of joints of CHOC and RAC: (a) ψ_1; (b) ψ_2; (c) ψ_3.

trajectory of P_d is described as follows:

$$P_d = \begin{bmatrix} x_{P_d} \\ y_{P_d} \\ z_{P_d} \end{bmatrix} = \begin{bmatrix} 0.36+0.05\sin(0.08\pi t) \\ 0.05\cos(0.08\pi t) \\ 0.4 \end{bmatrix}. \tag{57}$$

A. SIMULATION

The parameters of the CHOC are selected as $K_1=K_2=K_3=K_4=diag[10,10,10]$, the initial weights are zeros, $\Theta=I\in\mathbb{R}^{6\times6}$, $R=I\in\mathbb{R}^{6\times6}$, $\alpha=100$, $\nu_2=0.15$, $\nu_1=-0.1$, $\rho=0.5$, $\chi_2=0.9$, $\chi_1=-0.8$, the activation function is

$$\zeta(e)=[e_{\delta1}^2, e_{\delta1}e_{\delta2}, e_{\delta1}e_{\delta3}, e_{\delta1}e_{\psi1}, e_{\delta1}e_{\psi2}, e_{\delta1}e_{\psi3}, e_{\delta2}^2,$$

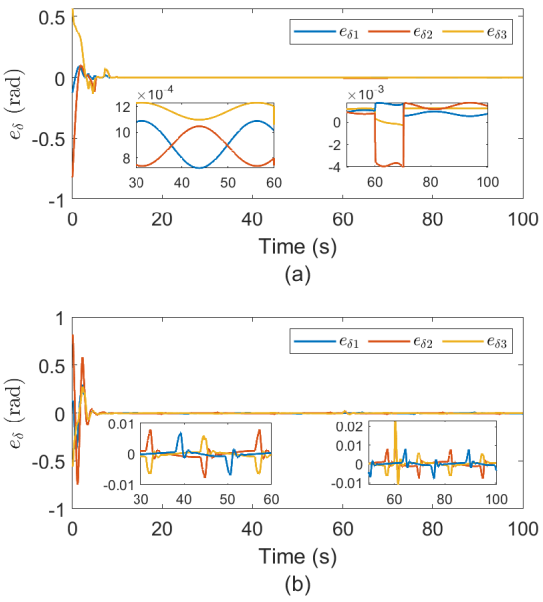

FIGURE 7. Position tracking errors of CHOC and RAC in joint space: (a) CHOC; (b) RAC.

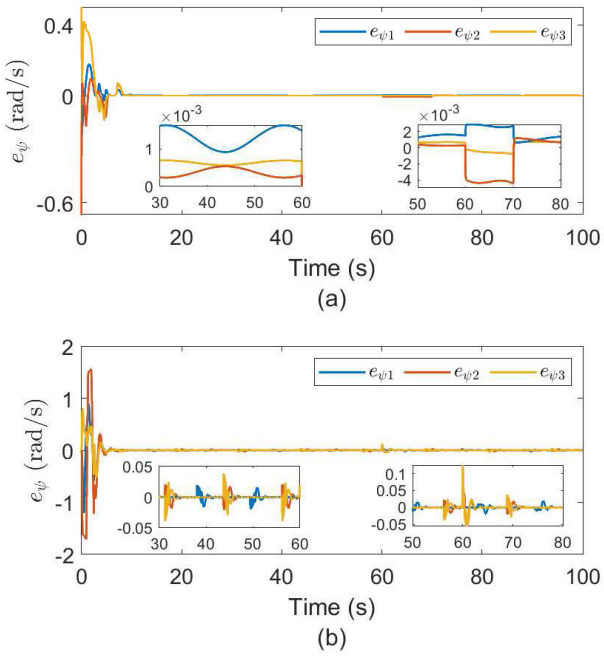

FIGURE 8. Angular velocity tracking errors of CHOC and RAC in joint space: (a) CHOC; (b) RAC.

$$e_{\delta 2}e_{\delta 3}, e_{\delta 2}e_{\psi 1}, e_{\delta 2}e_{\psi 2}, e_{\delta 2}e_{\psi 3}, e_{\delta 3}^2, e_{\delta 3}e_{\psi 1}, e_{\delta 3}e_{\psi 2}, e_{\delta 3}e_{\psi 3}, e_{\psi 1}^2, e_{\psi 1}e_{\psi 2}, e_{\psi 1}e_{\psi 3}, e_{\psi 2}^2, e_{\psi 2}e_{\psi 3}, e_{\psi 3}^2]^T,$$

the sampling period $T = 0.01s$. The control law of RAC is given as $\tau = \hat{f}(z) + K_6 s - \upsilon$, where $\hat{f}(z)$ is RBFNN and is an approximation of $f(z), f(z) = \mathcal{H}\left(\ddot{\delta}_{ref} + K_5\dot{e}_\delta\right) + \mathcal{C}\left(\dot{\delta}_{ref} + K_5 e_\delta\right) + \mathcal{G} + \mathcal{F}$, $K_5 = K_6 = diag[10, 10, 10]$, $s = \dot{e}_\delta + K_5 e_\delta$, $\upsilon = -K_7\,\mathrm{sgn}(s)$, $K_7 = 0.2$. We change the model parameters at the 60th second for both control laws, i.e., the values $\sigma_1, \sigma_2, \ldots, \sigma_{12}$, and σ_{13} are changed to double

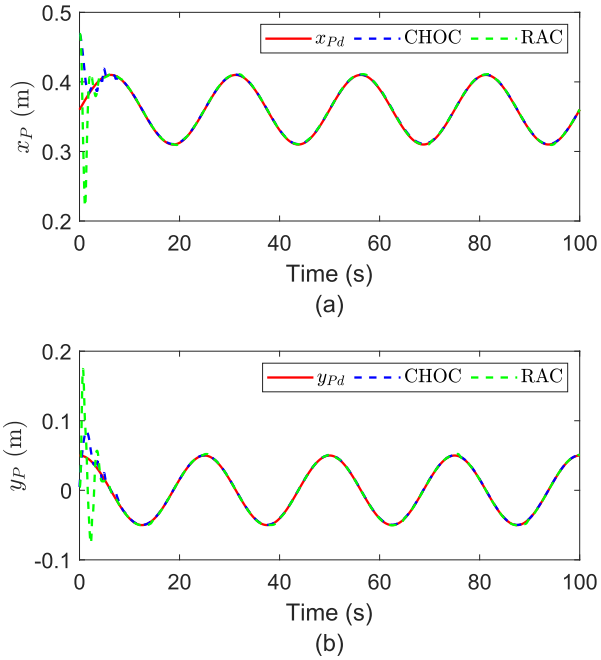

FIGURE 9. Trajectory tracking of CHOC and RAC in the worksapce: (a) Position x_P; (b) Position y_P.

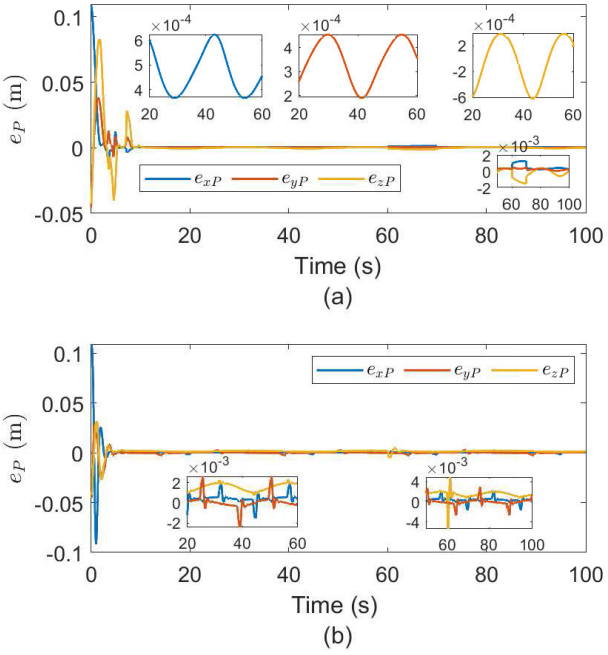

FIGURE 10. Tracking errors of CHOC and RAC in the workspace: (a) CHOC; (b) RAC.

the initial values. Fig. 4 presents the convolution of the NN weights, where the weights converge after 10s. At the time of load change (60th second), the weights change to adapt to the load change and then continue to converge. The results of comparing the trajectories of the positions of CHOC and RAC in joint space are illustrated in Fig. 5. Fig. 6 shows CHOC and RAC's angular velocity tracking trajectory. The position and angular velocity tracking errors of the controllers in joint space are illustrated in Figs. 7, 8. The comparison

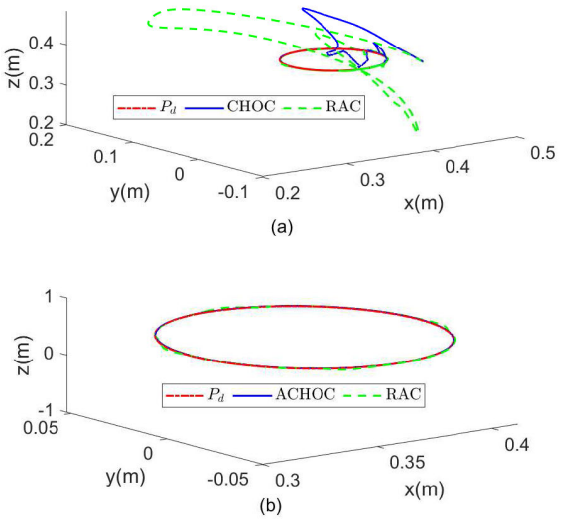

FIGURE 11. Position of point P of CHOC and RAC: (a) During learning; (b) After convergence.

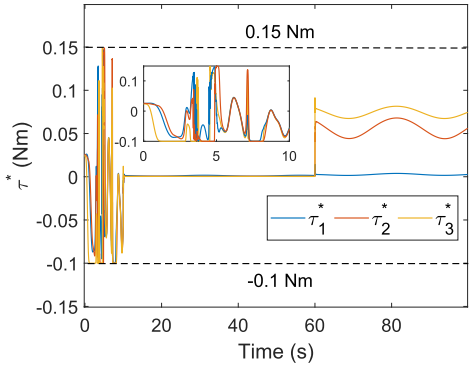

FIGURE 12. The optimal control inputs of CHOC.

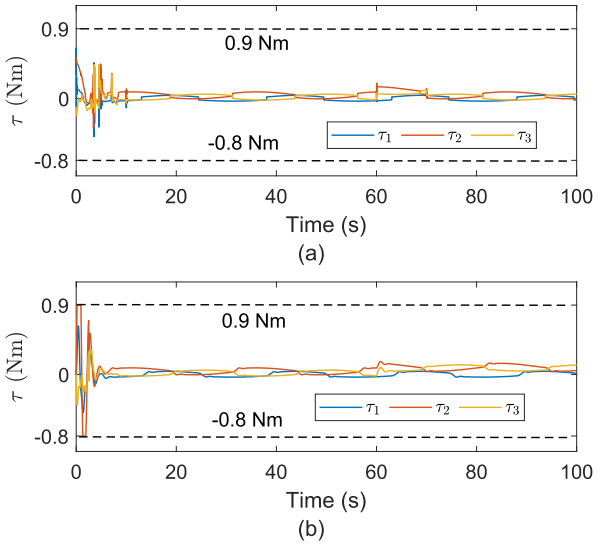

FIGURE 13. Control torques of CHOC and RAC: (a) CHOC; (b) RAC.

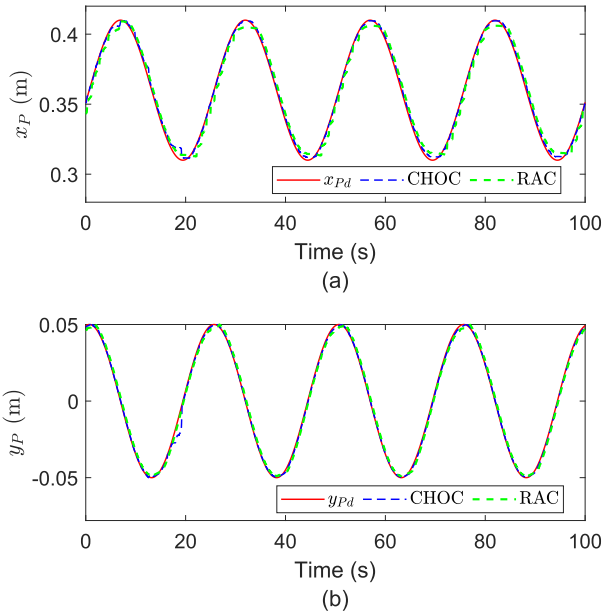

FIGURE 14. Experimental results of the trajectory tracking of CHOC and RAC in the workspace: (a) Position x_P; (b) Position y_P.

results show that CHOC's ability to track trajectories in the joint space is better than RAC's. CHOC's position tracking error does not exceed 12.10^{-4} rad. When the load changes suddenly, the position tracking error increases but does not exceed 4.10^{-3} rad and then returns to the value 12.10^{-4} rad (see Fig. 7). In addition, Fig. 8 shows that CHOC's angular velocity tracking error has a smaller value than RAC.

The position trajectories in the controllers' workspace are shown in Fig. 9. The tracking errors in the controllers' workspace are shown in Fig. 10, which indicates that the tracking errors of CHOC are smaller than RAC. The tracking error of CHOC in the workspace is less than 0.6 mm for the X-position, 0.45 mm for the Y-position, and 0.6 mm for the Z-position (see Fig. 10). The position of point P of CHOC and RAC are illustrated in Fig. 11, where Fig. 11a shows the results during the learning process, and Fig. 11b shows the results after convergence. The control inputs of CHOC are presented in Figs. 12 and 13a, where Fig. 12 shows the optimal feedback control inputs and Fig. 13a illustrates the actual control torque inputs. The control torque results of RAC are presented in Fig. 13b. Fig. 12 shows that the optimal control inputs are within the allowable asymmetry limit and change value when the minimum and maximum thresholds are reached, which leads to the actual control torques not violating the asymmetry constraint (Fig. 13a). Since RAC does not handle the influence of input constraints, the control torques fluctuate enormously in the early stages (Fig. 13b). After the algorithms converge, the actual control torques of CHOC reach the optimal value while the torques of RAC do not (Fig. 13). The simulation results show that CHOC has better control performance than RAC.

B. EXPERIMENT

In this section, we present experimental results on Scorbot-ER. Scorbot's hardware is set up as follows (see Fig. 2): The drive motor of the robot joints is a DC motor with a built-in gearbox with a gear ratio of 127.7:1. In addition,

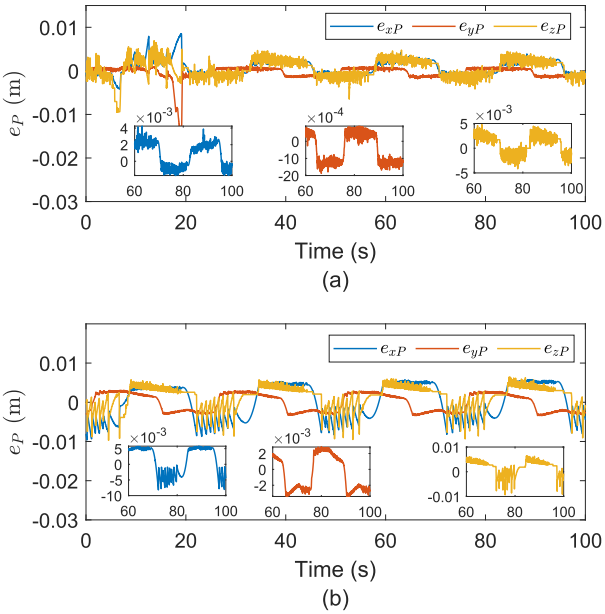

FIGURE 15. Experimental results of the tracking error of CHOC and RAC in the workspace: (a) CHOC; (b) RAC.

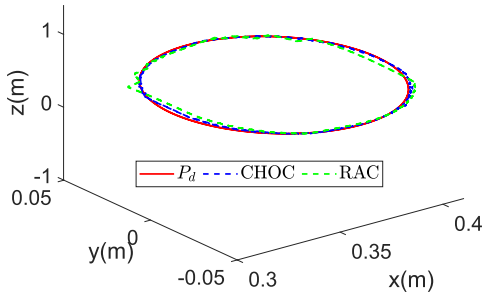

FIGURE 16. Experimental results: Position of point *P* of CHOC and RAC after convergence.

the motors are equipped with optical encoders to measure the angular velocity and angular position of the joints. The control algorithms are implemented on the STM32F407 microcontroller. To interface the STM32F407 with the motors, we use the L298 driver circuit.

For the CHOC and RAC algorithms, we use convergence weights from simulation to make real-time learning faster. The position tracking trajectories in the workspace of CHOC and RAC are presented in Fig. 14. The position tracking errors of CHOC in the workspace are shown in Fig. 15a, indicating that the tracking errors are significant in the early stages and small as the algorithm converges. Fig. 15b shows tracking errors of RAC in the workspace. Comparing the tracking errors of CHOC and RAC, we quickly see that the tracking errors of CHOC are smaller than RAC. Fig. 16 presents the position of point *P* of CHOC and RAC after convergence, showing that CHOC's tracking performance is better than RAC's. Through experimental results, the control effectiveness of CHOC has been verified.

A video illustrating experimental results can be found at the following URL: https://www.youtube.com/watch?v=A-S4Chi5798.

V. CONCLUSION

The H_∞ optimal control algorithm has been proposed for robot manipulators with asymmetric input constraints and disturbances. The feedforward control is proposed to transform the H_∞ optimal control problem for robot manipulators into the H_∞ optimal control problem for the affine systems. The constrained optimal control and disturbance compensation laws have been built based on ADP and the game theory. The proposed control algorithm has been analyzed to show that the closed-loop dynamics is UUB stable and the control parameters converge to the near-optimal values. The performance of the proposed algorithm has been verified through comparative simulation and experimental results. CHOC can be developed to be applied to multi-agent systems, such as decentralized control of robot manipulators or some multiple robot manipulator operations. Future work will focus on decentralized optimal control of robot manipulators and distributed optimal control for multi-robot systems.

ACKNOWLEDGMENT

The authors would like to thank Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, for supporting this study.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

REFERENCES

[1] W. He, Z. Li, and C. L. P. Chen, "A survey of human-centered intelligent robots: Issues and challenges," *IEEE/CAA J. Autom. Sinica*, vol. 4, no. 4, pp. 602–609, Jun. 2017.
[2] F. L. Lewis, D. M. Dawson, and C. T. Abdallah, *Robot Manipulator Control: Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2003.
[3] K. Shojaei, A. Kazemy, and A. Chatraei, "An observer-based neural adaptive PID^2 controller for robot manipulators including motor dynamics with a prescribed performance," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 3, pp. 1689–1699, Jun. 2021.
[4] B. M. Yilmaz, E. Tatlicioglu, A. Savran, and M. Alci, "Self-adjusting fuzzy logic based control of robot manipulators in task space," *IEEE Trans. Ind. Electron.*, vol. 69, no. 2, pp. 1620–1629, Feb. 2022.
[5] L. Dai, Y. Yu, D.-H. Zhai, T. Huang, and Y. Xia, "Robust model predictive tracking control for robot manipulators with disturbances," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4288–4297, May 2021.
[6] M. Van and S. S. Ge, "Adaptive fuzzy integral sliding-mode control for robust fault-tolerant control of robot manipulators with disturbance observer," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 5, pp. 1284–1296, May 2021.
[7] R.-D. Xi, X. Xiao, T.-N. Ma, and Z.-X. Yang, "Adaptive sliding mode disturbance observer based robust control for robot manipulators towards assembly assistance," *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 6139–6146, Jul. 2022.
[8] V. T. Yen, W. Y. Nan, and P. Van Cuong, "Robust adaptive sliding mode neural networks control for industrial robot manipulators," *Int. J. Control, Autom. Syst.*, vol. 17, no. 3, pp. 783–792, Mar. 2019.
[9] M. Boukens, A. Boukabou, and M. Chadli, "Robust adaptive neural network-based trajectory tracking control approach for nonholonomic electrically driven mobile robots," *Robot. Auto. Syst.*, vol. 92, pp. 30–40, Jun. 2017.
[10] W. He, Y. Dong, and C. Sun, "Adaptive neural impedance control of a robotic manipulator with input saturation," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 46, no. 3, pp. 334–344, Mar. 2016.
[11] P. Hippe, *Windup in Control: Its Effects and Their Prevention*. Germany: Springer, 2006.

[12] W. He, Y. Sun, Z. Yan, C. Yang, Z. Li, and O. Kaynak, "Disturbance observer-based neural network control of cooperative multiple manipulators with input saturation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 5, pp. 1735–1746, May 2020.

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

[14] E. Arefinia, H. A. Talebi, and A. Doustmohammadi, "A robust adaptive model reference impedance control of a robotic manipulator with actuator saturation," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 50, no. 2, pp. 409–420, Feb. 2020.

[15] G. Li, J. Yu, and X. Chen, "Adaptive fuzzy neural network command filtered impedance control of constrained robotic manipulators with disturbance observer," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 8, pp. 1–10, Jun. 2021.

[16] C. Yang, D. Huang, W. He, and L. Cheng, "Neural control of robot manipulators with trajectory tracking constraints and input saturation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 9, pp. 4231–4242, Sep. 2021.

[17] J. Ma, S. S. Ge, Z. Zheng, and D. Hu, "Adaptive NN control of a class of nonlinear systems with asymmetric saturation actuators," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 7, pp. 1532–1538, Jul. 2015.

[18] L. Kong, W. He, Y. Dong, L. Cheng, C. Yang, and Z. Li, "Asymmetric bounded neural control for an uncertain robot by state feedback and output feedback," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 3, pp. 1735–1746, Mar. 2021.

[19] L. Kong, W. He, W. Yang, Q. Li, and O. Kaynak, "Fuzzy approximation-based finite-time control for a robot with actuator saturation under time-varying constraints of work space," *IEEE Trans. Cybern.*, vol. 51, no. 10, pp. 4873–4884, Oct. 2021.

[20] X. Yu, H. Luo, S. Shi, Y. Wei, and L. Ou, "Asymmetric constrained control scheme design with discrete output feedback in unknown robot–environment interaction system," *Robotica*, vol. 41, no. 1, pp. 105–125, Jan. 2023.

[21] A. J. Van Der Schaft, "L2-gain analysis of nonlinear systems and nonlinear state feedback H_∞ control," *IEEE Trans. Automat. Control*, vol. 37, no. 6, pp. 770–784, Jul. 1992.

[22] H. Modares, F. L. Lewis, and Z.-P. Jiang, "H_∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2550–2562, 2015.

[23] 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, Jan. 2015.

[24] V. T. Vu, P. N. Dao, P. T. Loc, and T. Q. Huy, "Sliding variable-based online adaptive reinforcement learning of Uncertain/Disturbed nonlinear mechanical systems," *J. Control, Autom. Electr. Syst.*, vol. 32, no. 2, pp. 281–290, Apr. 2021.

[25] B. Dong, T. An, X. Zhu, Y. Li, and K. Liu, "Zero-sum game-based neuro-optimal control of modular robot manipulators with uncertain disturbance using critic only policy iteration," *Neurocomputing*, vol. 450, pp. 183–196, Aug. 2021.

[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 Trans. Ind. Electron.*, vol. 68, no. 6, pp. 5046–5056, Jun. 2021.

[27] L. Kong, W. He, C. Yang, and C. Sun, "Robust neurooptimal control for a robot via adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 6, pp. 2584–2594, Jun. 2021.

[28] L. N. Tan and D. L. Gia, "ADP-based H_∞ optimal decoupled control of single-wheel robots with physically coupling effects, input constraints, and disturbances," *IEEE Trans. Ind. Electron.*, vol. 71, no. 7, pp. 7445–7454, Jul. 2024, doi: 10.1109/TIE.2023.3301537.

[29] N. Le-Dung, P. Huynh-Lam, N. Hoang-Giap, and N. Tan-Luy, "Event-triggered distributed robust optimal control of nonholonomic mobile agents with obstacle avoidance formation, input constraints and external disturbances," *J. Franklin Inst.*, vol. 360, no. 8, pp. 5564–5587, May 2023.

[30] X. Bu, "An improvement of single-network adaptive critic design for nonlinear systems with asymmetry constraints," *J. Franklin Inst.*, vol. 356, no. 16, pp. 9646–9664, Nov. 2019.

[31] L. Xia, Q. Li, R. Song, and H. Modares, "Optimal synchronization control of heterogeneous asymmetric input-constrained unknown nonlinear MASs via reinforcement learning," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 3, pp. 520–532, Mar. 2022.

[32] N. D. Dien, N. T. Luy, and L. K. Lai, "Optimal tracking control for robot manipulators with asymmetric saturation torques based on reinforcement learning," *J. Comput. Sci. Cybern.*, vol. 39, no. 1, pp. 61–77, Mar. 2023.

[33] J. J. Craig, *Introduction to Robotics: Mechanics and Control*. London, U.K.: Pearson Educacion, 2005.

[34] S. Zeghloul, M. A. Laribi, and J.-P. Gazeau, "Robotics and mechatronics," in *Proc. 4th IFToMM Int. Symp. Robot. Mechatronics*, 2015, pp. 1–10.

[35] Q. Jiao, H. Modares, S. Xu, F. L. Lewis, and K. G. Vamvoudakis, "Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control," *Automatica*, vol. 69, pp. 24–34, Jul. 2016.

[36] H. Zargarzadeh, T. Dierks, and S. Jagannathan, "Adaptive neural network-based optimal control of nonlinear continuous-time systems in strict-feedback form," *Int. J. Adapt. Control Signal Process.*, vol. 28, nos. 3–5, pp. 305–324, Mar. 2014.

[37] F. L. Lewis, K. Liu, and A. Yesildirek, "Neural net robot controller with guaranteed tracking performance," *IEEE Trans. Neural Netw.*, vol. 6, no. 3, pp. 703–715, May 1995.

[38] M. Szuster and P. Gierlak, "Approximate dynamic programming in tracking control of a robotic manipulator," *Int. J. Adv. Robotic Syst.*, vol. 13, no. 1, p. 16, Jan. 2016.

[39] L. N. Tan and T. C. Pham, "Optimal tracking control for PMSM with partially unknown dynamics, saturation voltages, torque, and voltage disturbances," *IEEE Trans. Ind. Electron.*, vol. 69, no. 4, pp. 3481–3491, Apr. 2022.

[40] S. Xue, B. Luo, D. Liu, and Y. Gao, "Event-triggered integral reinforcement learning for nonzero-sum games with asymmetric input saturation," *Neural Netw.*, vol. 152, pp. 212–223, Aug. 2022.

DIEN NGUYEN DUC received the degree in electrical engineering technology, majoring in automation from the University of Economics-Technology for Industry, Vietnam, and the master's degree in control and automation engineering from the University of Transport Vietnam, in 2014. Since 2012, he has been a Lecturer with the Department of Control and Automation, Faculty of Electrical-Automation, University of Economics-Technology for Industry. His main research interests include intelligent control, robotics, adaptive optimal control, sustainable optimal control, and ADP.

LAI LAI KHAC received the Ph.D. degree in control and automation engineering from Hanoi University of Science and Technology, Vietnam, in 2003. He received the title of an Associate Professor, in 2007. He is currently with Thai Nguyen University of Technology, Vietnam. His current research interests include adaptive and optimal control, intelligent control of electric drive systems, and renewable energy systems.

LUY NGUYEN TAN (Senior Member, IEEE) received the B.S. and M.Sc. degrees in computer science and automation and control engineering and the Ph.D. degree in automation from Ho Chi Minh City University of Technology (HCMUT), Vietnam, in 1996, 2006, and 2015, respectively.

He is currently with the Faculty of Electric-Electronics Engineering (FEEE), HCMUT, Vietnam National University Ho Chi Minh City (VNU-HCM), Vietnam. His current research interests include adaptive dynamic programming, distributed control, large-scale systems, and deep learning. He served as a Reviewer for many journals, including IEEE Transactions on Neural Network and Learning Systems, IEEE Transactions on Industrial Electronics, IEEE Transactions on Industrial Informatics, IEEE Transactions on Cybernetics, IEEE Transactions on Systems, Man, Cybernetics: Systems, IEEE Access, and Elsevier journals.