Integral Reinforcement Learning Control of an Uncertain 2-DOF Helicopter System With Input Quantization and State Constraints

Zhijia Zhao, *Member, IEEE*, Yan Weng, Zhijie Liu, *Member, IEEE*, Yu Liu, *Senior Member, IEEE*, and Keum-Shik Hong, *Life Fellow, IEEE*

***Abstract*—This study presents an integral reinforcement learning (RL) control strategy for an uncertain two-degree-of-freedom (2-DOF) helicopter system by considering the input quantization and state constraints. Initially, a hysteresis quantizer is utilized to reduce oscillations when transmitting input signals. Within the control framework, the model uncertainty is estimated by an action neural network (NN), while the long-term cost function is approximated by a critic NN. Moreover, integral barrier Lyapunov functions (IBLFs) are employed to directly restrict the states, thereby guaranteeing the convergence of the helicopter system. Lyapunov stability theory demonstrates that all signals within the closed-loop system are semiglobal uniform ultimate bounded. Ultimately, the efficacy of the proposed control strategy is validated via both simulations and experiments.**

***Index Terms*—Input quantization, integral barrier Lyapunov function (IBLF), integral reinforcement learning (RL) control, neural network (NN), two-degree-of-freedom (2-DOF) helicopter system.**

Received 14 May 2024; revised 4 September 2024 and 25 October 2024; accepted 4 December 2024. Date of publication 11 February 2025; date of current version 23 July 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62273112, Grant 62433011, and Grant 62473039; in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2023B1515120018 and Grant 2023B1515120019; in part by the Science and Technology Planning Project of Guangzhou, China, under Grant 2023A03J0120; in part by the Joint Fund of Ministry of Education for Equipment Pre-Research under Grant 8091B03032303; in part by Beijing Nova Program under Grant 20240484561; and in part by the National Research Foundation of Korea funded by the Ministry of Science and ICT, Korea, under Grant IRIS-2023-00207954. *(Corresponding author: Zhijie Liu.)*

Zhijia Zhao and Yan Weng are with the School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China (e-mail: zhaozj@gzhu.edu.cn; 2112207025@e.gzhu.edu.cn).

Zhijie Liu is with the Institute of Artificial Intelligence and the State Key Laboratory of Advanced Metallurgy, University of Science and Technology Beijing, Beijing 100083, China (e-mail: liuzhijie@ustb.edu.cn).

Yu Liu is with the School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China (e-mail: auylau@scut.edu.cn).

Keum-Shik Hong is with the Institute for Future, School of Automation, Qingdao University, Qingdao 266071, China, and also with the School of Mechanical Engineering, Pusan National University, Busan 46241, South Korea (e-mail: kshong@pusan.ac.kr).

Digital Object Identifier 10.1109/TIE.2024.3519581

I. Introduction

UNMANNED aerial vehicle (UAV) technology has experienced remarkable advancements and breakthroughs in recent years [1], [2]. Unmanned helicopters, a prominent category of UAVs, have extensive applications in various scenarios, such as reconnaissance, climate monitoring, and medical transportation, owing to their vertical lifting and hovering capabilities [3], [4], [5], [6]. Nevertheless, helicopter systems are inherently complex multiinput and multioutput (MIMO) systems, characterized by model uncertainty and strong coupling, which pose challenges regarding their control. Hence, designing effective controllers for helicopter systems is crucial [7].

In recent years, significant research efforts have been directed toward addressing the control challenges associated with two-degree-of-freedom (2-DOF) helicopter systems [8], [9], [10]. However, these studies either linearized inherently nonlinear helicopter systems or solely addressed model uncertainty in the context of nonlinear systems without considering potential constraints on inputs and outputs.

In practical engineering applications, the inevitable constraints are of paramount importance. Overlooking these constraints may lead to detrimental effects on the control performance and system stability and even result in system instability. Quantization, which is a technique for converting continuous signals into segmented constant signals using a specific transformation algorithm, reduces the communication rate and consequently alleviates the communication burden. Therefore, the exploration of quantization methods within control systems has seen considerable advancements in recent years [11], [12], [13], [14]. In [15], Liu et al. presented a groundbreaking nonlinear decomposition technique for quantized inputs, successfully overcoming the inherent drawback of conventional quantization methods that require the generation of quantized inputs in discrete segments. Furthermore, Zhang et al. enhance the handling of virtual control signals by incorporating dynamic filtering techniques. They utilize parametric projection for estimator design and introduce an innovative state observer form. This approach effectively achieves integrated control of nonlinear systems with mismatched uncertainties, addressing both input and output quantization conditions [16]. In addition, the barrier Lyapunov function (BLF) has gained

recognition as an effective approach for efficiently addressing constrained problems. Studies have widely employed it when designing control schemes for nonlinear systems considering input or output constraints. Furthermore, the utilization of the BLF has been demonstrated to significantly enhance control performance [17], mitigate output hysteresis [18], expedite the convergence of tracking errors [19], [20], and bolster system robustness [21]. However, traditional logarithmic and tangent BLFs indirectly impose constraints on the state or output by limiting the error, which has certain limitations. To overcome this problem, the integral BLF (IBLF) method has been applied in various studies [22], [23], [24]. In [25], Zhang et al. utilized IBLF to address the impact of output constraints. In a strict-feedback system, the application of IBLF can ensure that all system states are within preset constraints, thus effectively improving the control performance and robustness [26]. Despite successes in studies involving quantization techniques or IBLF, there remains a scarcity of studies that combine both input quantization and state constraints using IBLF, which motivates our research.

In recent times, radial basis function (RBF) neural networks (NNs) have risen to prominence as a significant approach for managing the uncertainties and unknown dynamic parameters in nonlinear systems, attributed to their superior nonlinear approximation capabilities and inherent self-learning attributes [27], [28], [29], [30]. In [31], Zhang et al. proposed an adaptive NN-based variable stiffness control strategy that specifically target nonlinear dynamics and external disturbances in uncertain robotic systems. By combining a high-dimensional integral Lyapunov function with a disturbance observer, the method effectively cope with the complexity of variable stiffness actuators and ensure system stability. Moreover, Zhao et al. proposed an adaptive neural quantization control method for MIMO nonlinear systems and utilized an error-dependent asymmetric time-varying barrier function to constrain the system output [32]. Nevertheless, in complex situations, the performance of adaptive NNs, which relies solely on system errors for weight adjustments, may not fully meet the desired objectives. Reinforcement learning (RL) is a machine learning method in which an intelligent agent interacts with its environment to obtain rewards and penalties that shape its behavioral strategy. It has undergone notable progress and extensive adoption in recent years. Additionally, RL has shown significant potential in the domain of automatic control [33], [34], [35], [36]. In [37], an RL control strategy for a helicopter system was successfully developed by Zhao et al., using the state error and control input to derive an optimal strategy and updating the weights based on current gradients. However, significant time costs or convergence difficulties may be faced by this method. To address the limitations of previous approaches, Moares et al. proposed an integral RL (IRL) algorithm utilizing empirical replay for continuous-time systems with input constraints and partially unknown models, demonstrating through experiments that this method could significantly accelerate convergence speed [38]. Building on this, Zhang et al. developed an adaptive fuzzy fault-tolerant tracking control method using IRL for partially unknown systems with actuator faults [39]. Similarly, Guo et al. introduced an adaptive NN control method based on IRL for continuous-time nonlinear MIMO systems with unknown control directions [40]. All these studies have shown through simulation that IRL-based control methods could effectively enhance control performance and improve system robustness. With the continuous advancement of IRL, researchers have successfully applied this method to tackle the control of increasingly complex nonlinear systems, yielding promising results [41], [42], [43]. However, its application to 2-DOF helicopter systems with input quantization and state constraints is yet to be fully explored. This motivates the present study.

Inspired by the aforementioned studies, we are going to an IRL control (IRLC) strategy for an uncertain 2-DOF nonlinear helicopter system. This strategy aims to achieve high-precision trajectory tracking control of the helicopter system despite the challenges posed by input quantization and unknown time-varying disturbances. The main contributions of this study are as follows.

1) Logarithmic quantizers [44] and sector bounded quantizers [14] are favored by researchers for their simple design and high quantization efficiency. However, frequent quantization switching can result in significant fluctuations in control inputs and a reduction in control accuracy, which may not be suitable for 2-DOF helicopter systems that require high-precision tracking control. Therefore, this article employs a hysteresis quantizer to quantize the system input signals, which not only smooths the impact of quantization errors on control performance but also maintains good quantization efficiency.
2) Different from [32], we propose an IRLC approach based on two NNs to approximate system uncertainties. In this method, a critic NN is used to evaluate the control strategy, while an action NN approximates the system uncertainties. Additionally, known system information is appropriately incorporated into the network training to enhance approximation accuracy. Moreover, unlike [33] and [37], the proposed IRLC strategy introduces an instantaneous cost function with an adjustable error threshold, aiming to reduce training costs, improve network convergence speed, and enhance real-time evaluation capabilities.
3) To ensure control performance and system stability, we utilize an IBLF to directly constrain the system state, and build an adaptive auxiliary system to compensate for the composite term consisting of quantization error and unknown disturbance. Finally, the simulation results of the proposed IRLC strategy under different unknown dynamics of the simulated helicopter system are given, and the effectiveness of the proposed control strategy is further verified by experiments.

II. Problem Formulation and Preliminaries

A. Control Objective and System Model

In this article, we propose an IRLC scheme for an uncertain nonlinear 2-DOF helicopter system to address challenges such as input quantization and time-varying disturbances. The main control objective is to ensure that the pitch and yaw angles of

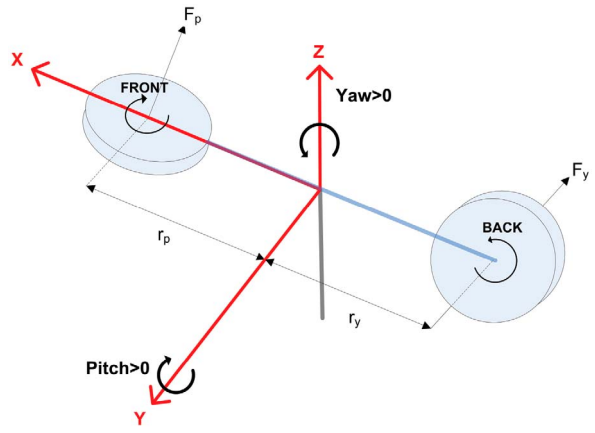

Fig. 1. 2-DOF helicopter model [45].

the helicopter can accurately track a given trajectory despite the high uncertainty, unknown time-varying disturbances, and input quantization of the system. In addition, the introduction of the IBLF ensures that the system state is kept within the defined constraints and the tracking error is kept within the desired range. After rigorous validation, the proposed control scheme ensures uniform boundedness of all signals within the closed-loop system.

Fig. 1 shows the model of a 2-DOF helicopter system. According to the Euler–Lagrange method, we formulated the dynamic of the system as

$$\left(J_p + mL_{cm}^2\right)\ddot{\theta} = K_{pp}V_p + K_{py}V_y - mgL_{cm}\cos\theta - D_p\dot{\theta} - mL_{cm}^2\dot{\phi}^2\sin\theta\cos\theta \quad (1)$$

$$\left(J_y + mL_{cm}^2\cos^2\theta\right)\ddot{\phi} = K_{yp}V_p + K_{yy}V_y - D_y\dot{\phi} + 2mL_{cm}^2\dot{\phi}\dot{\theta}\sin\theta\cos\theta \quad (2)$$

where θ and ϕ refer to the pitch and yaw angles, respectively. J_p and J_y denote the moments of inertia around the pitch and yaw axes, respectively. Similarly, D_p and D_y represent the viscous friction coefficients along the pitch and yaw axes, respectively. The input voltages for the motors are indicated by V_p and V_y, for the pitch and yaw directions, respectively. The parameters g and m stand for the gravity acceleration and the mass of the helicopter, respectively. L_{cm} represents the distance from the mass center to the fixed frame, while K_{pp}, K_{py}, K_{yp}, and K_{yy} are the coefficients representing the thrust torque constants.

To simplify the dynamic equation, we define

$$\begin{aligned} f_1 &= J_p + mL_{cm}^2 \\ f_2 &= -mgL_{cm}\cos\theta - D_p\dot{\theta} - mL_{cm}^2\sin\theta\cos\theta\dot{\phi}^2 \\ g_1 &= J_y + mL_{cm}^2\cos^2\theta \\ g_2 &= -D_y\dot{\phi} + 2mL_{cm}^2\dot{\phi}\dot{\theta}\sin\theta\cos\theta. \end{aligned} \quad (3)$$

Moreover, $x_1 = [\theta, \phi]^T$, $x_2 = [\dot{\theta}, \dot{\phi}]^T$, $u = [V_p, V_y]^T$, and $x = [x_1, x_2]^T$. Subsequently, (1) and (2) can be translated into the following MIMO nonlinear equations:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= L(x) + \Delta L(x) + B(x)u + d \\ y &= x_1 \end{aligned} \quad (4)$$

where $\Delta L(x)$ denotes the uncertainty of the system, d is the unknown time-varying disturbance, and L and B are expressed as follows:

$$L(x) = \begin{bmatrix} \frac{f_2}{f_1} \\ \frac{g_2}{g_1} \end{bmatrix}, B(x) = \begin{bmatrix} \frac{Kpp}{f_1} & \frac{Kpy}{f_1} \\ \frac{Kyp}{g_1} & \frac{Kyy}{g_1} \end{bmatrix}. \quad (5)$$

B. Hysteresis Quantizer

To ensure optimal system performance within a specified bandwidth and to minimize information transmission jitter, a quantizer is implemented to optimize the communication rate effectively. The original input u is represented as the quantized output $H(u_j)$, where u_j is the quantized input signal. Therefore, the hysteresis quantizer is defined as follows:

$$H(u_j) = \begin{cases} u_j^i \text{sign}(u_j), & \frac{u_j^i}{1+\delta} \le |u_j| \le u_j^i, \dot{u}_j < 0 \\ & or\ u_j^i < |u_j| < \frac{u_j^i}{1-\delta} \\ & \dot{u}_j > 0 \\ u_j^i(1+\delta)\text{sign}(u_j), & u_j^i < |u_j| < \frac{u_j^i}{1-\delta}, \dot{u}_j < 0 \\ & or\ \frac{u_j^i}{1-\delta} < |u_j| < \frac{u_j^i(1+\delta)}{1-\delta} \\ & \dot{u}_j > 0 \\ 0, & 0 \le |u_j| < \frac{u_{\min}}{1+\delta}, \dot{u}_j < 0 \\ & or\ \frac{u_{\min}}{1+\delta} \le |u_j| \le u_{\min} \\ & \dot{u}_j > 0 \\ H(u_j(t^-)), & \dot{u}_j = 0, j = 1, 2 \end{cases} \quad (6)$$

where $u_j^i = \rho^{(1-i)}u_{\min}$, $i = 1, 2, \cdots$, $\delta = (1-\rho/1+\rho)$, $\delta \in (0,1)$, $\rho \in (0,1)$ is the quantization density, i is the quantization level, and $u_{\min}$ denotes the quantization dead zone for $H(u_j)$ whereas $H(u_j)$ is in the set $U_h = \{0, \pm u_j^i, \pm u_j^i(1+\delta)\}$. Based on [15], we decompose the hysteresis quantizer into the following equation:

$$H(u) = Mu + N \quad (7)$$

where $M = \text{diag}[m_1, m_2]$ with $1-\delta < m_i < 1+\delta$ and $N = [n_1, n_2]^T$ with $n_i < u_{\min}$. Considering (4) and (7), we obtain

$$\dot{x}_2 = L(x) + \Delta L(x) + GMu + \zeta \quad (8)$$

where $\zeta = GN + d$.

Remark 1: One of the objectives of the action NN is to approximate for the model uncertainty ΔL, helping to reduce z_1 and z_2 to nearly zero. The second objective is to minimize $\hat{Q}(t)$ to $\hat{Q}_d(t) = 0$.

C. Preliminaries

Assumption 1: For any positive constants K_{ci} and Y_{ci}, the desired tracking trajectory x_d and its time derivatives $\dot{x}_d$ satisfy $x_{di} < K_{ci}$ and $|\dot{x}_{di}| < Y_{ci}$.

Lemma 1: For any $Z \in \mathbb{R}$ and $b > 0$, the following inequality holds [46]:

$$0 \le |Z| - Z\tanh\left(\frac{Z}{b}\right) \le 0.2785b. \quad (9)$$

Lemma 2: For any variable $q \in \mathbb{R}$ and a constant $\eth > 0$, the subsequent relationship is established [47]:

$$0 \le |q| - \frac{q^2}{\sqrt{q^2 + \eth^2}} \le \eth. \tag{10}$$

Lemma 3: If a Lyapunov function candidate $V(x)$ satisfies $\Phi_1(||x||) \le V(x) \le \Phi_2(||x||)$, and with a bounded initial value $V(0)$, then $V(x)$ is positive definite and the following inequality can be obtained [48]:

$$\dot{V}(x) \le -\rho V(x) + C \tag{11}$$

where $\rho > 0$ and $C > 0$.

D. Neural Network Approximation

Considering the unknown continuous function F_{nn}, we employed RBFNNs to approximate it

$$F_{nn}(X) = W^{*T}\nu(X) + \varrho(X) \tag{12}$$

where $W^* = [w_1, w_2, \ldots, w_n]^T$ represents the optimal weight of the RBFNN, with $n > 1$ indicating the quantity of nodes in the hidden layer; the input vector X belongs to a subset Ω of $\mathbb{R}^m$; $\varrho(X)$ specifies the approximation error; and $\nu(X) \in [\nu_1(X), \nu_2(X), \ldots, \nu_n(X)]$, where $\nu_i(X)$ is a basis function as

$$\nu_i(X) = \exp\left[\frac{-(X - c_i)^T(X - c_i)}{\varpi^2}\right], \ i = 1, 2, \ldots, n \tag{13}$$

where ϖ denotes the width of the basis function and $c_i \in [c_{i1}, c_{i2}, \ldots, c_{im}]$ is the center of the receptive field. For all $X \in \Omega \subset \mathbb{R}^m$, the optimal weight vector W^* is

$$W^* \triangleq \arg\min_{W \in \mathbb{R}^n} \left\{\sup_{Z \in \Omega} \left|F_{nn}(X) - W^T\nu(X)\right|\right\}. \tag{14}$$

E. Critic Neural Network Design

The long-term cost function is given as

$$Q(t) = \int_t^{\infty} \varsigma^{\frac{-\varepsilon+t}{T}} p(z_1(\varepsilon))d\varepsilon \tag{15}$$

where $\varsigma \in (0, 1)$ is a constant used to discount future costs, and $p(z_1)$ denotes the instantaneous cost function of the following form

$$p(z_{1i}(\varepsilon)) = \begin{cases} 0 & |z_{1i}(\varepsilon)| \le e \\ 1 & |z_{1i}(\varepsilon)| > e \end{cases}, \ \varepsilon \in [t - T, t). \tag{16}$$

Here, $e > 0$ is a customized error threshold and $T > 0$ indicates an integration interval. The evaluation rule based on the instantaneous cost function is as follows: $p = 0$ indicates good tracking performance, whereas $p = 1$ indicates poor tracking performance. The long-term cost function at time $(t - T)$ is as follows:

$$\begin{aligned} Q(t - T) &= \int_{t-T}^{\infty} \varsigma^{\frac{-\varepsilon+t-T}{T}} p(z_1(\varepsilon))d\varepsilon \\ &= \varsigma^{-1}Q(t) + \int_{t-T}^{t} \varsigma^{\frac{-\varepsilon+t-T}{T}} p(z_c(\varepsilon))d\varepsilon \\ &= \varsigma^{-1}(Q(t) + p_e) \end{aligned} \tag{17}$$

where $p_e \in [p_{e1}, p_{e2}]^T$ is the instantaneous cost as follows:

$$\begin{aligned} p_{ei} &= \int_{t-T}^{t} \varsigma^{\frac{-\varepsilon+t}{T}} p(z_{1i}(\varepsilon))d\varepsilon \\ &= \begin{cases} 0 & |z_{1i}(\varepsilon)| \le e \\ \frac{T}{\ln \varsigma}(\varsigma - 1) & |z_{1i}(\varepsilon)| > e. \end{cases} \end{aligned} \tag{18}$$

This implies $\|p_e\| \le b_{p_e}$, where b_{p_e} is a positive constant.

As shown in (17), $Q(t)$ is not directly available. Hence, a critic NN is employed to approximate it

$$Q(t) = W_c^{*T}\nu_c(X_c(t)) + \varrho_c \tag{19}$$

where W_c^* signifies the optimal weight with $\|W_c^*\|_F \le b_{W_c}$. n_c refers to the hidden layer nodes of the critic NN. ν_c represents the basis function. The input vector is denoted by $X_c = [x_1^T, z_1^T]^T$. Finally, ϱ_c denotes the approximation error.

As W_c^* is unknown, we define $\hat{W}_c$ as an estimate of W_c^* and $\widetilde{W}_c = \hat{W}_c - W_c^*$. Then, the estimation of $Q(t)$ is represented as

$$\hat{Q}(t) = \hat{W}_c^T\nu_c(X_c(t)). \tag{20}$$

According to (22), $Q(t - T)$ can be expressed as follows:

$$\hat{Q}(t - T) = \hat{W}_c^T\nu_c(X_c(t - T)). \tag{21}$$

The prediction error is then described as follows:

$$\begin{aligned} e_c &= \hat{Q}(t) - \varsigma\hat{Q}(t - T) + p_e \\ &= \hat{W}_c^T\Delta\nu_c(t) + p_e \end{aligned} \tag{22}$$

where $\Delta\nu_c(t) = \nu_c(X_c(t)) - \varsigma\nu_c(X_c(t - T))$.

We define the error of the critic NN as

$$E_c = \frac{1}{2}e_c^T e_c. \tag{23}$$

Through the application of the gradient descent technique, the adaptive law is derived as follows:

$$\dot{\hat{W}}_c = -\lambda_c \frac{\partial E_c}{\partial \hat{W}_c}. \tag{24}$$

Then, we obtain

$$\dot{\hat{W}}_c = -\lambda_c\left[\Delta\nu_c(t)\left(\hat{W}_c^T\Delta\nu_c(t) + p_e\right)^T + \sigma_c\hat{W}_c\right] \tag{25}$$

where $\lambda_c > 0$ denotes the learning rate of the critic NN and σ_c is a small positive constant.

F. State Constraints

To enforce the state constraints, an IBLF candidate is introduced in the following form:

$$V_1 = \sum_{i=1}^{2} \int_0^{z_{1i}} \frac{\sigma_i K_{ci}^2}{K_{ci}^2 - (\sigma_i + x_{di})^2} d\sigma \tag{26}$$

where K_{ci} denotes the designed boundary.

Lemma 4: If the condition $|x_{di}| < K_{ci}$ is satisfied, then the following equation holds [26]:

$$V_1 \le \sum_{i=1}^{2} \frac{K_{ci}^2 z_{1i}^2}{K_{ci}^2 - x_{1i}^2}. \tag{27}$$

According to Lemma 4, we get the time derivative of V_1 as

$$\dot{V}_1 = \sum_{i=1}^{2} \frac{K_{ci}^2 z_{1i}}{K_{ci}^2 - x_{1i}^2}(z_{2i} + \alpha_{1i} - \dot{x}_{di}) + \sum_{i=1}^{2} \frac{\partial V_1}{\partial x_{di}} \dot{x}_{di}. \quad (28)$$

We define $\sigma_i = \beta z_{1i}$. Using integration by parts, we obtain

$$\frac{\partial V_1}{\partial x_{di}} = z_{1i}\left(\frac{K_{ci}^2}{K_{ci}^2 - x_{1i}^2} - \epsilon(z_{1i}, x_{di})\right) \quad (29)$$

where ϵ is expressed as follows:

$$\begin{aligned}\epsilon(z_{1i}, x_{di}) &= \frac{K_{ci}}{z_{1i}} \int_0^1 \frac{1}{1 - (\frac{\beta z_{1i} + x_{di}}{K_{ci}})^2} d\left(\frac{\beta z_{1i} + x_{di}}{K_{ci}}\right) \\ &= \frac{K_{ci}}{2z_{1i}} \ln \frac{(K_{ci} + z_{1i} + x_{di})(K_{ci} - x_{di})}{(K_{ci} - z_{1i} - x_{di})(K_{ci} + x_{di})}. \end{aligned} \quad (30)$$

Based on L'Hospital's rule, (32) satisfies

$$\begin{aligned}\lim_{z_{1i}\to 0} \epsilon(z_{1i}, x_{di}) &= \lim_{z_{1i}\to 0} \frac{K_{ci}^2}{K_{ci}^2 - (z_{1i} + x_{di})^2} \\ &= \lim_{z_{1i}\to 0} \frac{K_{ci}^2}{K_{ci}^2 - x_{1i}^2}. \end{aligned} \quad (31)$$

Considering (31)–(33), we rewritten (30) as follows:

$$\begin{aligned}\dot{V}_1 &= \sum_{i=1}^{2} \frac{K_{ci}^2 z_{1i}}{K_{ci}^2 - x_{1i}^2}(z_{2i} + \alpha_{1i} - \dot{x}_{di}) \\ &\quad + \sum_{i=1}^{2} \left(\frac{K_{ci}^2}{K_{ci}^2 - x_{1i}^2} - \epsilon(z_{1i}, x_{di})\right) z_{1i}\dot{x}_{di} \\ &= \sum_{i=1}^{2} \frac{K_{ci}^2 z_{1i}(z_{2i} + \alpha_{1i})}{K_{ci}^2 - x_{1i}^2} - \sum_{i=1}^{2} \epsilon(z_{1i}, x_{di})\dot{x}_{di}. \end{aligned} \quad (32)$$

Let the virtual controller $\alpha_{1i} = -K_{1i} z_{1i} + \dot{x}_{di}$, where K_{1i} is the control gain. Thereafter, the time derivative of V_1 is restated as

$$\dot{V}_1 = \sum_{i=1}^{2} \frac{K_{ci}^2}{K_{ci}^2 - x_{1i}^2}(-K_{1i} z_{1i}^2 + z_{1i} z_{2i}). \quad (33)$$

G. Action Neural Network Design

We design the following Lyapunov candidate function V_2 as follows:

$$V_2 = V_1 + \frac{1}{2} z_2^T z_2. \quad (34)$$

Based on (16) and (35), we express the time derivative of V_2 as

$$\begin{aligned}\dot{V}_2 &= \dot{V}_1 + z_2^T \dot{z}_2 \\ &= \sum_{i=1}^{2} \frac{K_{ci}^2}{K_{ci}^2 - x_{1i}^2}(-K_{1i} z_{1i}^2 + z_{1i} z_{2i}) \\ &\quad + z_2^T (L + \Delta L + BMu + \zeta). \end{aligned} \quad (35)$$

For the model uncertainty of the helicopter system ΔL, we employ an action NN to address it

$$\Delta L - \dot{\alpha}_1 = W_a^{*T} \nu_a(X_a) + \varrho_a \quad (36)$$

where W_a^* represents the optimal weight with $\|W_a^*\|_F \le b_{W_a}$, n_a refers to the count of hidden layer nodes, the input vector for the action NN, denoted by X_a, is composed of $[x_1^T, x_2^T, \dot{\alpha}_1^T]^T$, the basis function is symbolized as ν_a with $\|\nu_a\| \le b_{\nu_a}$, and ϱ_a symbolizes the approximation error with $\|\varrho_a\| \le b_{\varrho_a}$.

Since $1 - \delta < m_i < 1 + \delta$, we know that M is bounded. Then, we have $\ell = \inf_{t\ge 0} \lambda_{\min}(M)$ and $\kappa = (1/\ell)$. Moreover, we define $\widetilde{W}_a = \hat{W}_a - W_a^*$ and $\xi = L + \hat{W}_a^T \varphi_a + \tanh((z_2/b_1))\hat{\zeta} + \begin{bmatrix}(K_{c1}^2 z_{11}/K_{c1}^2 - x_{11}^2) \\ (K_{c2}^2 z_{12}/K_{c2}^2 - x_{12}^2)\end{bmatrix} + K_2 z_2$. Subsequently, the IRLC strategy is designed as follows:

$$u = -B^{-1}\left(z_2 \frac{\hat{\kappa}^2 \xi^T \xi}{\sqrt{\hat{\kappa}^2 z_2^T z_2 \xi^T \xi + \eth^2}}\right) \quad (37)$$

where $K_2 = \mathrm{diag}[k_{21}, k_{22}]$ is the control gain and $\eth$ is a small constant.

We define $\widetilde{\kappa} = \kappa - \hat{\kappa}$ and $\widetilde{\zeta} = \zeta - \hat{\zeta}$. Then, the updating laws of $\dot{\hat{\kappa}}$ and $\dot{\hat{\zeta}}$ are given by

$$\dot{\hat{\kappa}} = \lambda_1 \left(z_2^T \xi - \sigma_1 \hat{\kappa}\right) \quad (38)$$

$$\dot{\hat{\zeta}} = \lambda_2 \left(z_2^T \tanh\left(\frac{z_2}{b_1}\right) - \sigma_2 \hat{\zeta}\right) \quad (39)$$

where λ_1, λ_2, σ_1, σ_2, and b_1 are designed positive constants.

Remark 2: One of the objectives of the action NN is to approximate for the model uncertainty ΔL, helping to reduce z_1 and z_2 to nearly zero. The second objective is to minimize $\hat{Q}(t)$ to $\hat{Q}_d(t) = 0$.

We design the prediction error for action NN as

$$e_a = z_2 + \hat{Q}(t) - \hat{Q}_d(t) = z_2 + \hat{W}_c^T \varphi_c(X_c(t)). \quad (40)$$

Let $E_a = (1/2) e_a^T e_a$. The adaptive law of the action NN is then designed as

$$\dot{\hat{W}}_a = \lambda_a [\nu_a (z_2 + \hat{W}_c^T \nu_c)^T - \sigma_a \hat{W}_a] \quad (41)$$

where $\lambda_a > 0$ is the learning rate of the action NN and σ_a is a small positive constant with $\sigma_a > b_{\nu_c} b_{\nu_a}$.

III. Stability Analysis

Theorem 1: Given the helicopter system described in (1)–(5), along with the control strategy in (39), the network weight update rates in (27) and (43), and the adaptive auxiliary system in (40) and (41), the proposed IRLC scheme ensures that all signals in the control system remain semi-global uniform ultimate bounded (SGUUB) for any tracking trajectory x_d satisfying $x_d(0) \in \Omega_0$, provided that appropriate control gain parameters are selected.

Proof: Consider the following Lyapunov candidate equation

$$V_3 = V_n + V_a + V_c \quad (42)$$

where $V_n = V_2 + (\ell/2\lambda_1)\widetilde{\kappa}^2 + (1/2\lambda_2)\widetilde{\zeta}^2$, $V_a = (1/2)\,\mathrm{tr}\left\{\widetilde{W}_a^T \lambda_a^{-1} \widetilde{W}_a\right\}$, and $V_c = (1/2)\mathrm{tr}\left\{\widetilde{W}_c^T \lambda_c^{-1} \widetilde{W}_c\right\}$.

Considering (39) and Lemma 2, we obtain

$$\begin{aligned} z_2^T GMu &= -\frac{z_2^T z_2 GMG^{-1}\hat{\kappa}^2\xi^T\xi}{\sqrt{\hat{\kappa}^2 z_2^T z_2\xi^T\xi+\eth^2}} \\ &\le -\frac{\hat{\kappa}^2 z_2^T z_2\ell\xi^T\xi}{\sqrt{\hat{\kappa}^2 z_2^T z_2\xi^T\xi+\eth^2}} \\ &\le \ell(\eth-\hat{\kappa}z_2^T\xi) \end{aligned} \tag{43}$$

According to (40), (41), (45), and Lemma 1, the time derivative of V_n is given by

$$\begin{aligned} \dot{V}_n \le & \sum_{i=1}^{2}\frac{K_{ci}^2}{K_{ci}^2-x_{1i}^2}(-K_{1i}z_{1i}^2) - z_2^T K_2 z_2 - z_2^T\widetilde{W}_a\nu_a \\ & + \ell(\eth-\hat{\kappa}z_2^T\xi) + z_2^T\tanh\left(\frac{z_2}{b_1}\right)\hat{\zeta} + 0.557b_1\zeta + z_2^T\xi \\ & - \ell\widetilde{\kappa}z_2^T\xi + \ell\sigma_1\widetilde{\kappa}(\kappa-\widetilde{\kappa}) - z_2^T\tanh\left(\frac{z_2}{b_1}\right)\hat{\zeta} \\ & + \sigma_2\widetilde{\zeta}(\zeta-\widetilde{\zeta}) \\ = & \sum_{i=1}^{2}\frac{K_{ci}^2}{K_{ci}^2-x_{1i}^2}(-K_{1i}z_{1i}^2) - z_2^T K_2 z_2 - z_2^T\widetilde{W}_a\nu_a \\ & + \ell\eth + 0.557b_1\zeta + \ell\sigma_1\widetilde{\kappa}\kappa - \ell\sigma_1\widetilde{\kappa}^2 + \sigma_2\widetilde{\zeta}\zeta - \sigma_2\widetilde{\zeta}^2. \end{aligned} \tag{44}$$

Considering (22)–(27), we obtain

$$\begin{aligned} \dot{V}_c \le & -\mathrm{tr}\left\{\widetilde{W}_c^T\Delta\nu_c\left[(\widetilde{W}_c^T+W_c^{*T})\Delta\nu_c+p_e\right]^T\right\} \\ & -\sigma_c\mathrm{tr}\left\{\widetilde{W}_c^T\widetilde{W}_c\right\} - \sigma_c\mathrm{tr}\left\{\widetilde{W}_c^T W_c^*\right\} \\ = & -\mathrm{tr}\left\{\widetilde{W}_c^T\left(\Delta\nu_c\Delta\nu_c^T+\sigma_c\right)\widetilde{W}_c\right\} \\ & + ||\widetilde{W}_c||_F\left\{||\Delta\nu_c||||W_c^{*T}\Delta\nu_c+p_e|| + \sigma_c\left\|W_c^*\right\|_F\right\} \\ = & -\sigma_c\mathrm{tr}\left\{\widetilde{W}_c^T\widetilde{W}_c\right\} + c_1||\widetilde{W}_c||_F \end{aligned} \tag{45}$$

where $c_1 = ||\Delta\nu_c(t)||\,[||\Delta\nu_c(t)||b_{W_c}+b_{p_e}+\sigma_c b_{W_c}]$.

Considering (27) and (43), we derive

$$\begin{aligned} \dot{V}_a = & -\sigma_a\mathrm{tr}\left\{\widetilde{W}_a^T\hat{W}_a\right\} + \mathrm{tr}\left\{\widetilde{W}_a^T\nu_a[\hat{W}_c^T\nu_c]^T\right\} \\ & + \mathrm{tr}\left\{\widetilde{W}_a^T\nu_a z_2^T\right\} \\ \le & -\left(\frac{\sigma_a-b_{\nu_a}b_{\nu_c}}{2}\right)\mathrm{tr}(\widetilde{W}_a^T\widetilde{W}_a) + \frac{b_{\nu_a}b_{\nu_c}}{2}||\widetilde{W}_c||_F^2 \\ & + \frac{\sigma_a b_{W_a}^2}{2} + b_{\nu_a}b_{\nu_c}b_{W_c}||\widetilde{W}_a||_F + \mathrm{tr}\left\{\widetilde{W}_a^T\nu_a z_2^T\right\} \\ \le & -\left(\frac{\sigma_a-b_{\nu_a}b_{\nu_c}}{2}-\Xi_a\right)\mathrm{tr}(\widetilde{W}_a^T\widetilde{W}_a) + c_2 \\ & + \frac{b_{\nu_a}b_{\nu_c}}{2}\mathrm{tr}\left\{\widetilde{W}_c^T\widetilde{W}_c\right\} + \mathrm{tr}\left\{\widetilde{W}_a^T\nu_a z_2^T\right\} \end{aligned} \tag{46}$$

where $c_2 = (\sigma_a b_{W_a}^2)/2 + (b_{\nu_a}^2 b_{\nu_c}^2 b_{W_c}^2)/4\Xi_a$, and $-[(b_{\nu_a}^2 b_{\nu_c}^2 b_{W_c}^2)/4\Xi_a] + b_{\nu_a}b_{\nu_c}b_{W_c}||\widetilde{W}_a||_F \le \Xi_a\mathrm{tr}\left\{\widetilde{W}_a^T\widetilde{W}_a\right\}$ with $0<\Xi_a<(\sigma_a-b_{\nu_a}b_{\nu_c})/2$.

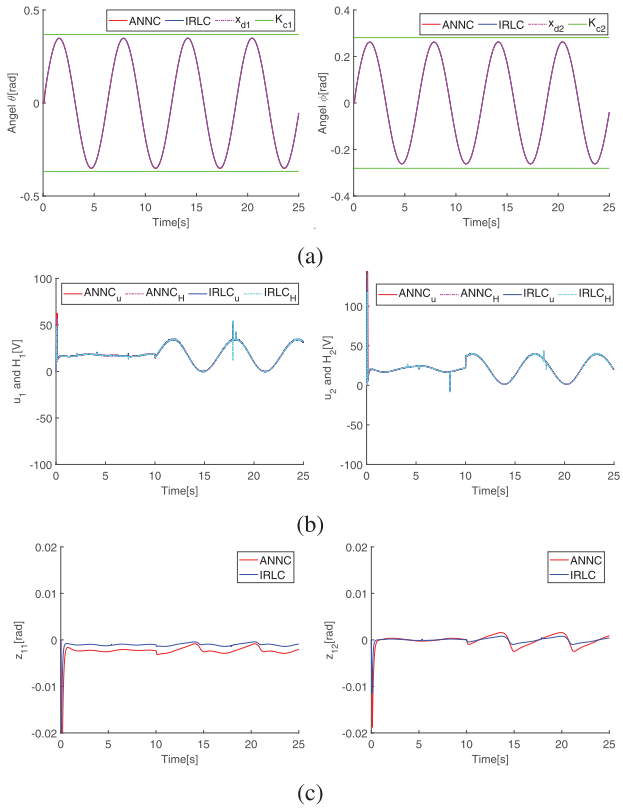

Fig. 2. Control performance under ANNC and IRLC in case 1. (a) Tracking performance. (b) Original signal u and quantized signal H. (c) Tracking errors z_{11} and z_{12}.

Substituting (46)–(48) into (44) and using Young's inequality, we obtain

$$\begin{aligned} \dot{V}_3 \le & \sum_{i=1}^{2}\frac{K_{ci}^2}{K_{ci}^2-x_{1i}^2}(-K_{1i}z_{1i}^2) - z_2^T K_2 z_2 - \frac{1}{2}\ell\sigma_1\widetilde{\kappa}^2 \\ & - \frac{1}{2}\sigma_2\widetilde{\zeta}^2 - \left(\sigma_c - \frac{b_{\nu_a}b_{\nu_c}}{2}\right)\mathrm{tr}\left\{\widetilde{W}_c^T\widetilde{W}_c\right\} + \frac{1}{2}\ell\sigma_1\kappa^2 \\ & + \frac{1}{2}\sigma_2\zeta^2 - \left(\frac{\sigma_a-b_{\nu_a}b_{\nu_c}}{2}-\Xi_a\right)\mathrm{tr}(\widetilde{W}_a^T\widetilde{W}_a) \\ & + c_1||\widetilde{W}_c||_F + c_2 + \ell\eth + 0.557b_1\zeta \\ \le & -\rho V_3 + C \end{aligned} \tag{47}$$

where $\rho = \min(2\lambda_{\min}(K_1), 2\lambda_{\min}(K_2), (\sigma_a - b_{\nu_a}b_{\nu_c} - 2\Xi_a/\lambda_{\max}(\lambda_a^{-1})), (2\sigma_c - b_{\nu_a}b_{\nu_c}/\lambda_{\max}(\lambda_c^{-1})), \lambda_1\sigma_1, \lambda_2\sigma_2)$ and $C = (1/2)\ell\sigma_1\kappa^2 + (1/2)\sigma_2\zeta^2 + c_1||\widetilde{W}_c||_F + c_2 + \ell\eth + 0.557b_1\zeta$. Moreover, the conditions to guarantee $\rho>0$, include $\lambda_{\min}(K_1)>0$, $\lambda_{\min}(K_2)>0$, $(\sigma_a - b_{\nu_a}b_{\nu_c} - 2\Xi_a/\lambda_{\max}(\lambda_a^{-1}))>0$, $(2\sigma_c - b_{\nu_a}b_{\nu_c}/\lambda_{\max}(\lambda_c^{-1}))>0$, $\lambda_1\sigma_1>0$, and $\lambda_2\sigma_2>0$ must hold.

Remark 3: From (44), it is obvious that the Lyapunov candidate V_3 is positive definite. And from the Lemma 3 and the stability theory and analysis in [49], [50], and [51], it can be introduced that $\dot{V}_3$ is negative semidefinite when we choose the appropriate control gain. Therefore, we can conclude that the control system is semi-globally stable and the system signals are bounded.

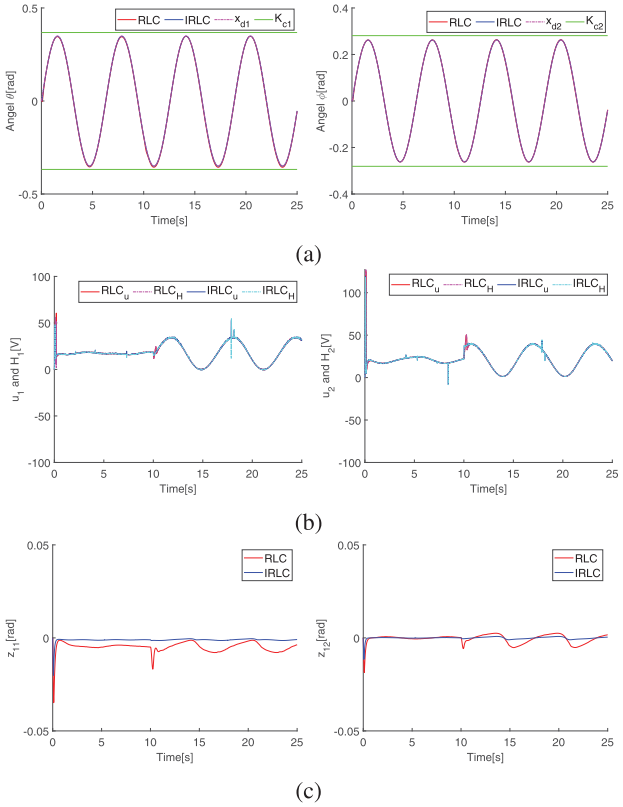

Fig. 3. Control performance under RLC and IRLC in case 1. (a) Tracking performance. (b) Original signal u and quantized signal H. (c) Tracking errors z_{11} and z_{12}.

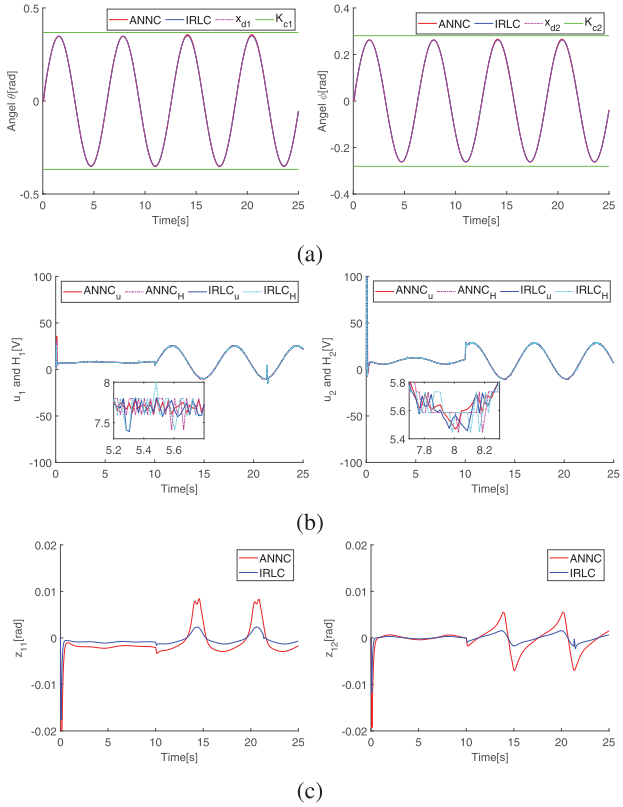

Fig. 4. Control performance under ANNC and IRLC in case 2. (a) Tracking performance. (b) Original signal u and quantized signal H. (c) Tracking errors z_{11} and z_{12}.

Subsequently, multiplying (44) by $e^{\rho t}$ yields, it can be obtain

$$V_3 \leq \left[V_3(0) - \frac{C}{\rho}\right] e^{-\rho t} + \frac{C}{\rho} \leq V_3(0) + \frac{C}{\rho}. \tag{48}$$

Therefore, it is concluded that all signals of the control system are SGUUB.

IV. Simulations

In this section, we outline the development of two control strategies for a 2-DOF helicopter system to validate the efficacy of the proposed IRLC scheme. The model parameters for the 2-DOF helicopter equipment are $K_{pp} = 0.0011\text{N}\cdot\text{m/V}$, $K_{py} = 0.0022\text{N}\cdot\text{m/V}$, $K_{yp} = -0.0027\text{N}\cdot\text{m/V}$, $K_{yy} = 0.0022\text{N}\cdot\text{m/V}$, $J_p = 0.0232\text{kg}\cdot\text{m}^2$, $J_y = 0.0238\text{kg}\cdot\text{m}^2$, $L_{cm} = 0.0071\text{m}$, $D_p = 0.0071\text{N/V}$, $D_y = 0.022\text{N/V}$, and $g = 0.98\text{m/s}^2$. The initial states are set to $x_1(0) = [0, 0]^T$ and $x_2(0) = [0, 0]^T$. The control gains selected are $K_1 = \text{diag}[10, 10]$ and $K_2 = \text{diag}[8, 8]$. The desired trajectory is choosen as $x_d = [x_{d1}, x_{d2}] = [(\pi/9)\sin(t), (\pi/12)\sin(t)]^T$ and the unknown disturbance is set as:

$$d = \begin{cases} 0 & t \leq 10 \\ [2\sin(t), 2\cos(t)]^T & t > 10 \end{cases} \tag{49}$$

Under the proposed IRLC strategy, the action NN generates the control input, and the critic NN evaluates the control performance. The control strategy is adjusted based on the critic error, resulting in input signals with improved control performance through continuous iterative training. The boundaries of the state are set to $K_{c1} = 0.368$ and $K_{c2} = 0.281$. The quantization parameters are selected as $\delta = 0.025$ and $u_{\min} = 0.6$. For the critic NN, the customized error, integration interval, learning rate, NN width, number of nodes, and robustness parameters are selected as $e = 0.004$, $T = 0.002$, $\lambda_c = 4$, $\varpi_c = 1$, $n_c = 64$, and $\sigma_c = 0.1$, respectively. For the action NN, $\lambda_a = 20$, $\varpi_a = 1$, $n_a = 256$, and $\sigma_a = 0.1$ are used. The parameters for the adaptive auxiliary terms in (47) and (48) are designed as $\lambda_1 = \lambda_2 = 10$, $b_1 = 0.25$, and $\sigma_1 = \sigma_2 = 0.1$.

Moreover, an adaptive NN control (ANNC) strategy [32] and a RL control (RLC) strategy [33] are designed for comparison with the IRLC strategy. The simulation results of the three methods are plotted in Figs. 2–5.

A. Case 1

In this case, we set the model uncertainty as $\Delta L = -0.1L$ to simulate the more desirable work environment. The simulation comparison results are presented in Figs. 2 and 3. Fig. 2(a)–2(c) illustrate the comparison of control performance between ANNC and IRLC under case 1. Specifically, Fig. 2(a) shows the tracking of θ and ϕ, Fig. 2(b) displays the control input trajectories and quantized input trajectories, and Fig. 2(c) records the tracking error magnitude. Similarly, Fig. 3(a)–3(c) compare the control performance of RLC and IRLC under the same conditions. From these simulation figures, it is evident that all three control schemes achieve good control performance under

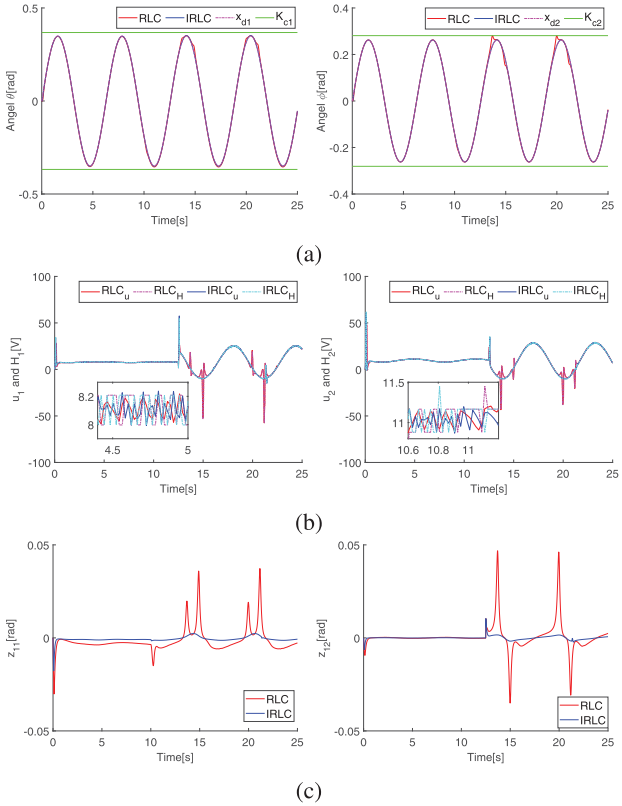

Fig. 5. Control performance under RLC and IRLC in case 2. (a) Tracking performance. (b) Original signal u and quantized signal H. (c) Tracking errors z_{11} and z_{12}.

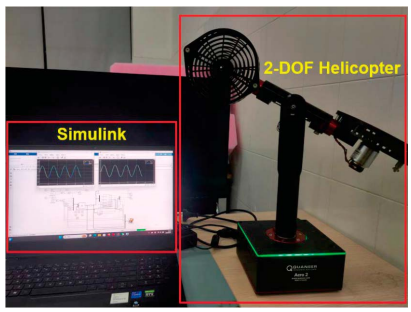

Fig. 6. Quanser Aero 2 experiment platform.

ideal conditions. However, when subjected to strong disturbances, while ANNC and RLC still demonstrate good control capabilities, IRLC outperforms them.

B. Case 2

The uncertainty is set as $\Delta L = -0.6L$ to model the complex effects of high uncertainty with strong perturbations in this case. The comparison of the control performance of ANNC, RLC, and IRLC in case 2 is given in Figs. 4 and 5, respectively. Similarly, the trajectory tracking of the helicopter system in ANNC, RLC, and IRLC is given in Figs. 4(a) and 5(a), respectively. Figs. 4(b) and 5(b) show the input and quantized signals, while Figs. 4(c) and 5(c) show the comparison of tracking error magnitude.

Combining cases 1 and 2, we can see that all control schemes have good control effects under more ideal operating conditions

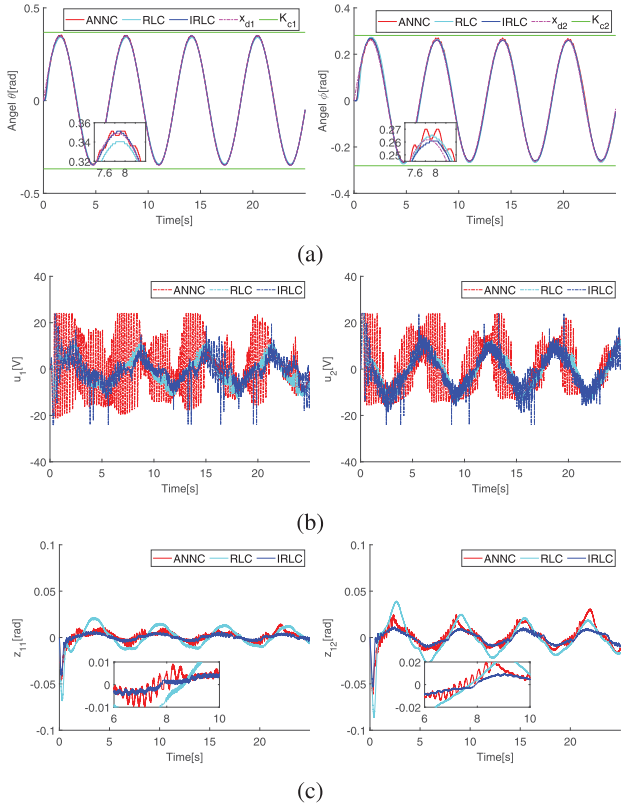

Fig. 7. Control performance under ANNC, RLC, and IRLC in the experiments. (a) Tracking performance. (b) Control input u. (c) Tracking errors z_{11} and z_{12}.

or single nonlinearities, and IRLC is slightly better than ANNC and RLC; however, facing the complex cases with high uncertainties and strong perturbations, ANNC and RLC cannot make the tracking error converge to near 0, and RLC shows multiple input oscillations, which shows poorer performance, while IRLC still has good control performance and stability, and is significantly better than the other two. In summary, the proposed control scheme can be applied to 2-DOF helicopter systems under different operating conditions.

V. Experiments

To further validate the effectiveness of the proposed control strategy, a series of experiments are conducted using the Quanser Aero 2 based on the findings of previous simulations. The Quanser Aero 2 platform is illustrated in Fig. 6. Notably, the input limits are set between -24 and $+24$ V. The strong disturbance in the experiment consists of the disturbance that we simulate with a fan and the air disturbance.

Fig. 7(a) shows that θ and ϕ effectively track the desired trajectories x_{d1} and x_{d2}, with the IBLF ensuring that the system state remains within the specified range. Analyzing Fig. 7(b), it is evident that under ANNC scheme, the input voltage exhibits multiple fluctuations, leading to a heavy communication burden, which may result in system instability. In contrast, RLC and IRLC strategy effectively mitigates quantization oscillations, producing a smoother voltage trajectory. Additionally, Fig. 7(c) reveals that while ANNC achieves a smaller tracking error, its performance is unstable; RLC can effectively address

quantization oscillations but requires a lengthy training period to learn the system and adjust the strategy. In contrast, IRLC efficiently handles input quantization and unknown perturbations, causing the tracking error to rapidly converge to near zero, demonstrating superior control performance and robustness. Based on the above simulation and experimental results, we conclude that the proposed IRLC strategy is suitable for real helicopter systems, offering good communication efficiency and excellent control performance.

VI. Conclusion

This study presented an IRLC strategy for trajectory tracking in a 2-DOF helicopter system. The strategy involved using an action NN to approximate the system uncertainty and adjust the control strategy, whereas a critic NN estimated the cost function and evaluated the current strategy. State constraints were enforced using the IBLF to improve control performance and robustness. Simulation and experimental results validated the effectiveness of the IRLC strategy. Future research directions include applying this strategy to quadcopter and hexacopter UAVs, which require further investigation.

References

[1] R. S. Geronel, R. M. Botez, and D. D. Bueno, "Dynamic responses due to the dryden gust of an autonomous quadrotor UAV carrying a payload," *Aeronaut. J.*, vol. 127, no. 1307, pp. 116–138, Apr. 2022.

[2] Y. Zou, H. Zhang, and W. He, "Adaptive coordinated formation control of heterogeneous vertical takeoff and landing UAVs subject to parametric uncertainties," *IEEE Trans. Cybern.*, vol. 52, no. 5, pp. 3184–3195, May 2022.

[3] Y. Li, M. Chen, P. Shi, and T. Li, "Stochastic anti-disturbance flight control for helicopter systems with switching disturbances under Markovian parameters," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 3, pp. 2933–2946, Jun. 2023.

[4] Y. Ren, Z. Liu, Z. Zhao, and H.-K. Lam, "Adaptive active anti-vibration control for a 3-d helicopter flexible slung-load system with input saturations and backlash," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 1, pp. 320–333, Feb. 2024.

[5] Y.-C. Lai and T.-Q. Le, "Adaptive learning-based observer with dynamic inversion for the autonomous flight of an unmanned helicopter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 3, pp. 1803–1814, Jun. 2021.

[6] H. Xiong and Y. Zhang, "Reinforcement learning-based formation-surrounding control for multiple quadrotor uavs pursuit-evasion games," *ISA Trans.*, vol. 145, pp. 205–224, Feb. 2024.

[7] L. Liu, M. Chen, and T. Li, "Disturbance observer-based LQR tracking control for unmanned autonomous helicopter slung-load system," *Int. J. Control, Automat. Syst.*, vol. 20, no. 4, pp. 1166–1178, Apr. 2022.

[8] R. Ganapathy Subramanian and V. K. Elumalai, "Robust MRAC augmented baseline LQR for tracking control of 2 DOF helicopter," *Robot. Auton. Syst.*, vol. 86, pp. 70–77, Dec. 2016.

[9] A. Fandel, A. Birge, and S. Miah, "Development of reinforcement learning algorithm for 2-DOF helicopter model," in *Proc. IEEE 27th Int. Symp. Ind. Electron. (ISIE)*, 2018, pp. 553–558.

[10] S.-K. Kim, K. S. Kim, and C. K. Ahn, "Order reduction approach to velocity sensorless performance recovery PD-type attitude stabilizer for 2-DOF helicopter applications," *IEEE Trans. Ind. Inform.*, vol. 18, no. 10, pp. 6848–6856, Oct. 2022.

[11] J. Fu and J. Wang, "Adaptive coordinated tracking of multi-agent systems with quantized information," *Syst. Control Lett.*, vol. 74, pp. 115–125, Dec. 2014.

[12] J. Zhou, C. Wen, and G. Yang, "Adaptive backstepping stabilization of nonlinear uncertain systems with quantized input signal," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 460–464, Feb. 2014.

[13] S. M. Schlanbusch, J. Zhou, and R. Schlanbusch, "Adaptive attitude control of a rigid body with input and output quantization," *IEEE Trans. Ind. Electron.*, vol. 69, no. 8, pp. 8296–8305, Aug. 2022.

[14] W. Wang, J. Zhou, C. Wen, and J. Long, "Adaptive backstepping control of uncertain nonlinear systems with input and state quantization," *IEEE Trans. Autom. Control*, vol. 67, no. 12, pp. 6754–6761, Dec. 2022.

[15] Z. Liu, F. Wang, Y. Zhang, and C. L. Philip Chen, "Fuzzy adaptive quantized control for a class of stochastic nonlinear uncertain systems," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 524–534, Feb. 2016.

[16] Z. Zhang, C. Wen, L. Xing, and Y. Song, "Adaptive output feedback control of nonlinear systems with mismatched uncertainties under input/output quantization," *IEEE Trans. Autom. Control*, vol. 67, no. 9, pp. 4801–4808, Sep. 2022.

[17] Y. Mei and Y. Liu, "ILC-RBNNF-based vibration control of a rotatable manipulator with time-varying output constraints," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 53, no. 10, pp. 6416–6425, Oct. 2023.

[18] Y.-J. Liu, W. Zhao, L. Liu, D. Li, S. Tong, and C. L. P. Chen, "Adaptive neural network control for a class of nonlinear systems with function constraints on states," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 6, pp. 2732–2741, Jun. 2023.

[19] X. Yuan, B. Chen, and C. Lin, "Prescribed finite-time adaptive neural tracking control for nonlinear state-constrained systems: Barrier function approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7513–7522, Dec. 2022.

[20] Y. Li, J. Zhang, W. Liu, and S. Tong, "Observer-based adaptive optimized control for stochastic nonlinear systems with input and state constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7791–7805, Dec. 2022.

[21] T. Zhang, M. Xia, and Y. Yi, "Adaptive neural dynamic surface control of strict-feedback nonlinear systems with full state constraints and unmodeled dynamics," *Automatica*, vol. 81, pp. 232–239, Jul. 2017.

[22] Y. Wei, Y. Wang, C. K. Ahn, and D. Duan, "IBLF-based finite-time adaptive fuzzy output-feedback control for uncertain MIMO nonlinear state-constrained systems," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 11, pp. 3389–3400, Nov. 2021.

[23] L. Liu, T. Gao, Y.-J. Liu, S. Tong, C. P. Chen, and L. Ma, "Time-varying iblfs-based adaptive control of uncertain nonlinear systems with full state constraints," *Automatica*, vol. 129, 2021, Art. no. 109595.

[24] W. He and S. S. Ge, "Vibration control of a flexible beam with output constraint," *IEEE Trans. Ind. Electron.*, vol. 62, no. 8, pp. 5023–5030, Aug. 2015.

[25] S. Zhang, Z. Tang, S. S. Ge, and W. He, "Adaptive neural dynamic surface control of output constrained non-linear systems with unknown control direction," *IET Control Theory Appl.*, vol. 11, no. 17, pp. 2994–3003, Sep. 2017.

[26] J. Zhang, W. Jiang, and S. S. Ge, "Adaptive fuzzy control for uncertain strict-feedback nonlinear systems with full-state constraints using disturbance observer," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 53, no. 10, pp. 6145–6156, Oct. 2023.

[27] Z. Liu, Z. Lu, Z. Zhao, M. O. Efe, and K.-S. Hong, "Single parameter adaptive neural network control for multi-agent deployment with prescribed tracking performance," *Automatica*, vol. 156, 2023, Art. no. 111207.

[28] Z. Zhao, Y. Ren, C. Mu, T. Zou, and K.-S. Hong, "Adaptive neural-network-based fault-tolerant control for a flexible string with composite disturbance observer and input constraints," *IEEE Trans. Cybern.*, vol. 52, no. 12, pp. 12843–12853, Dec. 2022.

[29] G. Peng, C. Yang, W. He, and C. L. P. Chen, "Force sensorless admittance control with neural learning for robots with actuator saturation," *IEEE Trans. Ind. Electron.*, vol. 67, no. 4, pp. 3138–3148, Apr. 2020.

[30] B. Hu, Z.-H. Guan, F. L. Lewis, and C. L. P. Chen, "Adaptive tracking control of cooperative robot manipulators with Markovian switched couplings," *IEEE Trans. Ind. Electron.*, vol. 68, no. 3, pp. 2427–2436, Mar. 2021.

[31] L. Zhang, Z. Li, and C. Yang, "Adaptive neural network based variable stiffness control of uncertain robotic systems using disturbance observer," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2236–2245, Mar. 2017.

[32] K. Zhao and J. Chen, "Adaptive neural quantized control of MIMO nonlinear systems under actuation faults and time-varying output constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 9, pp. 3471–3481, Sep. 2020.

[33] H. Shen, X. Yu, H. Yan, J. H. Park, and J. Wang, "Robust fixed-time sliding mode attitude control for a 2-DOF helicopter subject to input saturation and prescribed performance," *IEEE Trans. Transp. Electrific.*, early access, May 17, 2024.

[34] D. Lin, J. Han, K. Li, J. Zhang, and C. Zhang, "Payload transporting with two quadrotors by centralized reinforcement learning method," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 1, pp. 239–251, Feb. 2024.

[35] S. Cao, L. Sun, J. Jiang, and Z. Zuo, "Reinforcement learning-based fixed-time trajectory tracking control for uncertain robotic manipulators with input saturation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 8, pp. 4584–4595, Aug. 2023.
[36] W. Zhao, H. Liu, and F. L. Lewis, "Robust formation control for cooperative underactuated quadrotors via reinforcement learning," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 10, pp. 4577–4587, Oct. 2021.
[37] Z. Zhao, W. He, C. Mu, T. Zou, K.-S. Hong, and H.-X. Li, "Reinforcement learning control for a 2-DOF helicopter with state constraints: Theory and experiments," *IEEE Trans. Automat. Sci. Eng.*, vol. 21, no. 1, pp. 157–167, Jan. 2024.
[38] H. Modares, F. L. Lewis, and M.-B. Naghibi-Sistani, "Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems," *Automatica*, vol. 50, no. 1, pp. 193–202, 2014.
[39] H. Zhang, K. Zhang, Y. Cai, and J. Han, "Adaptive fuzzy fault-tolerant tracking control for partially unknown systems with actuator faults via integral reinforcement learning method," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 10, pp. 1986–1998, Oct. 2019.
[40] X. Guo, W. Yan, and R. Cui, "Integral reinforcement learning-based adaptive NN control for continuous-time nonlinear MIMO systems with unknown control directions," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 50, no. 11, pp. 4068–4077, Nov. 2020.
[41] Y. Lv, H. Chang, and J. Zhao, "Online adaptive integral reinforcement learning for nonlinear multi-input system," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 70, no. 11, pp. 4176–4180, Nov. 2023.
[42] J. Wang, J. Wu, J. Cao, M. Chadli, and H. Shen, "Nonfragile output feedback tracking control for Markov jump fuzzy systems based on integral reinforcement learning scheme," *IEEE Trans. Cybern.*, vol. 53, no. 7, pp. 4521–4530, Jul. 2023.
[43] X. Wang, Q. Wang, and C. Sun, "Prescribed performance fault-tolerant control for uncertain nonlinear MIMO system using actor-critic learning structure," *IEEE Trans. Neural Netwo. Learn. Syst.*, vol. 33, no. 9, pp. 4479–4490, Sep. 2022.
[44] C. Wang, C. Wen, Q. Hu, W. Wang, and X. Zhang, "Distributed adaptive containment control for a class of nonlinear multiagent systems with input quantization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2419–2428, Jun. 2018.
[45] I. Quanser, "Quanser aero 2 laboratory guide," Quanser, Tech. Rep., Dec. 2022. Available: https://quanserinc.box.com/shared/static/93bfbhxihax1v8bm9at44v56gv01vhee.zip
[46] H. Wang, B. Chen, and C. Lin, "Adaptive neural control for strict-feedback stochastic nonlinear systems with time-delay," *Neurocomputing*, vol. 77, no. 1, pp. 267–274, 2012.
[47] S. Song, J. H. Park, B. Zhang, and X. Song, "Event-based adaptive fuzzy fixed-time secure control for nonlinear CPSS against unknown false data injection and backlash-like hysteresis," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 6, pp. 1939–1951, Jun. 2022.
[48] L. Liu, T. Gao, Y.-J. Liu, S. Tong, C. P. Chen, and L. Ma, "Time-varying iblfs-based adaptive control of uncertain nonlinear systems with full state constraints," *Automatica*, vol. 129, 2021, Art. no. 109595.
[49] C. Wen, J. Zhou, Z. Liu, and H. Su, "Robust adaptive control of uncertain nonlinear systems in the presence of input saturation and external disturbance," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1672–1678, Jul. 2011.
[50] P. M. Patre, S. Bhasin, Z. D. Wilcox, and W. E. Dixon, "Composite adaptation for neural network-based controllers," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 944–950, Apr. 2010.
[51] S. Gao, B. Ning, H. Dong, and Y. Chen, "Backstepping-based neural adaptive control for saturated nonlinear systems," in *Proc. 33rd Chin. Control Conf.*, 2014, pp. 3345–3349.

Zhijia Zhao (Member, IEEE) received the B.Eng. degree from the North China University of Water Resources and Electric Power, Zhengzhou, China, in 2010, and the M.Eng. and Ph.D. degrees from the South China University of Technology, Guangzhou, China, in 2013 and 2017, respectively, all in automatic control.

Currently, he is a Professor with the School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou. His research interests include adaptive and learning control, flexible mechanical systems, and robotics.

Yan Weng received the B.Eng. degree in robotics engineering in 2022 from Guangzhou University, Guangzhou, China, where he is currently working toward the M.Eng. degree in control engineering.

His research interests include adaptive and learning control, intelligent control, and robotics.

Zhijie Liu (Member, IEEE) received the B.Sc. degree in electrical engineering and automation from China University of Mining and Technology Beijing, Beijing, China, in 2014, and the Ph.D. degree in control theory and control engineering from Beihang University, Beijing, in 2019.

Currently, he is a Full Professor with the School of Intelligence Science and Technology, University of Science and Technology Beijing, Beijing. In 2017, he was a Research Assistant with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, USA. His research interests include adaptive control, modeling and vibration control for flexible structures, and distributed parameter system.

Yu Liu (Senior Member, IEEE) received the Ph.D. degree in automatic control from the South China University of Technology, Guangzhou, China, in 2009.

Currently, he is a Professor with the School of Automation Science and Engineering, South China University of Technology, also with the R&D Center of Precision Electronic Manufacturing Technology, Guangzhou Institute of Modern Industrial Technology, Guangzhou, and also with the High Performance Motor and Intelligent Control Engineering Technology Research Center of Guangdong Province, Shenzhen, China. His research interests include distributed parameter systems, robot control, intelligent control, intelligent perception and decision making, machine vision.

Prof. Liu is currently an Associate Editor of IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON FUZZY SYSTEMS, and IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS.

Keum-Shik Hong (Life Fellow, IEEE) received the B.S. degree in mechanical design from Seoul National University, Seoul, South Korea, in 1979, the M.S. degree in mechanical engineering from Columbia University, New York, NY, USA, in 1987, and the M.E. degree in applied mathematics and the Ph.D. degree in mechanical engineering from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 1991 and 1997, respectively.

He joined the School of Mechanical Engineering, Pusan National University, Busan, South Korea, in 1993. His research interests include brain–computer interface, nonlinear systems theory, adaptive control, and distributed parameter systems.

Dr. Hong received many awards, including the Best Paper Award from the KFSTS of Korea in 1999 and the Presidential Award of Korea in 2007. He is served as an Associate Editor for *Automatica* from 2000 to 2006, the Editor-in-Chief for JOURNAL OF MECHANICAL SCIENCE AND TECHNOLOGY from 2008 to 2011, and the Editor-in-Chief for *International Journal of Control, Automation, and Systems* from 2018 to 2022. He was the past President of the Institute of Control, Robotics and Systems (ICROS), South Korea, and the Asian Control Association. He is a fellow of the Korean Academy of Science and Technology, an ICROS Fellow, and a member of the National Academy of Engineering of Korea.