Journal of Robotics and Control (JRC)
Volume 5, Issue 6, 2024
ISSN: 2715-5072, DOI: 10.18196/jrc.v5i6.23585

Model-free Optimal Control for Underactuated Quadrotor Aircraft via Reinforcement Learning

Quynh Nga Duong [1], Ngoc Trung Dang [2*]
[1, 2] Faculty of Electrical Engineering, Thai Nguyen University of Technology, Thai Nguyen, Vietnam
Email: [1] duongquynhngaktd@tnut.edu.vn, [2] trungcsktd@tnut.edu.vn
*Corresponding Author

Abstract—**The control of Unmanned Aerial Vehicles (UAVs), especially quadrotor aircraft, has many practical applications such as transporting, mapping, rescue, and agricultural applications. This paper investigates solving the optimal tracking control problem for a quadrotor system. First, an underactuated quadrotor system is considered a highly nonlinear system with six degrees of freedom and four inputs. Second, a hierarchical control structure consisting of position and attitude controller is adopted to address the underactuated problem, the position controller to achieve the desired tracking and generates the references for the attitude controller, and the attitude controller to achieve the reference attitude tracking. Third, to achieve optimal trajectory tracking, two Data- based Reinforcement Learning (RL) algorithms are applied to both position and attitude controllers to find the optimal control input by using the input- output quadrotor system data. Compared with the traditional optimal algorithms which require directly solving the Algebraic Ricatti Equation (ARE) or the Hamilton-Jacobi-Bellman (HJB) equation. It is impossible or difficult to implement due to the high nonlinear dynamic nature of the quadrotor system. By using RL in the proposed method, optimal policies can be learned without the knowledge of quadrotor dynamic information. Applying the learning control policies to the quadrotor system, the vehicle achieves optimal trajectory tracking. Finally, a simulation result is conducted to verify the optimal trajectory tracking for quadrotor with the proposed controller.**

Keywords—Quadrotor System; Reinforcement Learning; Optimal Tracking Control; Data-based Control.

I. INTRODUCTION

In recent years, Unmanned Aerial Vehicles (UAVs) have gained significant attention from the research community and industry. In practice, UAVs are used to assist people in difficult tasks such as rescue missions, fire detection, geological surveys, military operations, agricultural applications, and more [4]-[9]. Among the types of UAVs, the quadrotor aircraft has emerged due to its ability to take off vertically and fly flexibly in space. The quadrotor system is a machine with a rigid body connected to four rotors, so this system is underactuated and strong coupling following to six degrees of freedom and four inputs [1]-[3]. Furthermore, the completed dynamics of the quadrotor are difficult to achieve in practical applications due to the dynamic uncertainties and highly nonlinear nature. Uncertainty dynamics is a common problem of the quadrotor aircraft because the vehicle is usually required to carry unknown payloads in the flight application [29][32], which leads to the model parameter being uncertain. Therefore, the control design problem for the quadrotor system is challenging. To solve this problem, several traditional model-based control approaches are developed for the nonlinear system such as: adaptive control approaches in [4]-[6], robust adaptive control based sliding mode control (SMC) approaches [7]-[9]. Although control methods in [4]-[9] can robust the dynamic uncertainty, they do not guarantee the optimal tracking problem and require the completed dynamic information.

Implementing the optimal control scheme requires the approximate algorithms to solve the Algebraic Ricatti Equation (ARE) or the Hamilton-Jacobi-Bellman (HJB) which are directly difficult to implement for nonlinear quadrotor systems. The development of Reinforcement Learning (RL) has significant implications for addressing the optimal control problem [10][11]. By using the RL method, the optimal control policy can be learned without the knowledge of the system dynamics. In recent years, several RL-based control methods have been developed for the underactuated quadrotor shown in [12]-[15]. In [12], the authors used model-based RL for low-level control of a priori unknown quadrotor dynamics to achieve the stabilize of the vehicle in hover task from the system data. However, this method is designed to model with unchanging dynamics, therefore do not adapt to rapid changes. In [13], the RL method was applied to train a common neural network (NNs) to directly map state to actuators, but this work does not consider physical limits of the quadrotor system because its dynamic model was simplified. In [14], the authors developed an NN-based adaptive optimized controls for the underactuated quadrotor based on backstepping technique, but the dynamic uncertainties do not mention. In [15], the authors developed the integral RL for the position controller and the terminal SMC for the attitude controller to perform the various flying tasks with different loads while guaranteeing optimal performance, but the knowledge of attitude dynamic information is required.

In this paper, the model-free optimal tracking control problem is addressed for the underactuated quadrotor system subject to highly nonlinear and strong coupling. For the traditional optimal approaches in [16] and [17], the optimal control policies are by directly solving the HJB equation or ARE which is difficult to apply to the quadrotor system due to the highly nonlinear nature and the dynamic uncertainties. So, the motivation of this paper is to design an optimal controller to track the quadrotor system following a predefined trajectory without the knowledge of dynamic information. First, the underactuated quadrotor system is divided into two subsystems: the position subsystem and the attitude subsystem. Second, a hierarchical control structure consisting of the position controller and the attitude controller is adopted to address the underactuated problem, the position controller is to achieve the desired position tracking and generates the desired references for the attitude controller, and the attitude controller is to achieve the reference attitude tracking. Then, the data-based RL method is applied for both the position and attitude controller. By using the input-output quadrotor system data, the control policies can be obtained without directly solving the ARE and the HJB equation. Compared with the RL-based control in the previous studies in [14] and [15], the proposed method is to achieve optimal policies without the knowledge of system dynamic information.

The main contributions of this paper are listed following as

- The model-free optimal tracking control for the underactuated quadrotor is solved. By using data-based RL for both the position and the attitude controller, the optimal control policies can be learned by using the input-output system data. The vehicle with the learning policies optimally tracks to the desired reference trajectory which the previous works in [25]-[80] do not mention.
- The proposed controller is to restrain the influence of the high nonlinearity and uncertainties of the quadrotor system which the traditional optimal approaches in [16], [17] are impossible or difficult to implement in practical applications.

The rest of the paper is structured as follows. The quadrotor dynamics model is introduced in Section II. The data-based RL design for the position and attitude controller for the quadrotor illustrated in Section III. A simulation result is conducted to verify the effectiveness of the proposed controller in Section IV. Finally, concluding remarks are contained in Section V.

Notation: I_n is a unit $n \times n$ matrix, $0_{n\times m}$ is a $n \times m$ zeros matrix, $c_{n,m}$ is a column vector with 1 on the m^{th}element and 0 in elsewhere.

II. QUADROTOR MODEL

The quadrotor system is considered as a machine with a rigid body connected to four rotors. Let $E = \left[E_x, E_y, E_z\right]^T$ donates the Earth fixed frame and $B = \left[B_x, B_y, B_z\right]^T$ donates the body fixed frame. Let $p = [x, y, z]^T \in R^3$ donates the position of the mass center of the quadrotor in the Earth fixed frame, $\xi = [\phi,\ \theta,\ \psi]^T \in R^3$ donates the Euler angles following the roll angle ϕ, the pitch angle θ, the yaw angle ψ. The roll and pitch angle satisfy the bounded condition: $|\phi| < \pi/2$ and $|\theta| < \pi/2$ due to the singularity problems. This is an important assumption for quadrotor control design. Moreover, the UAV quadrotor parameters are expressed in the following Table I.

According to [18], the full dynamic model of the quadrotor system consists of the position subsystem and the attitude subsystem is described by (1).

TABLE I. PARAMETER OF QUADROTOR

Mass of the quadrotor	m (kg)
Gravitational acceleration	g (m/s^2)
Inertial matrix	$J = diag(\left[J_x, J_y, \quad J_z\right]) \in R^{3\times 3}$
The arm length	$l\ (m)$
Positive parameter	k_f, k_t

$$\begin{gathered} m\ddot{p} = T_p R_{B2E} c_{3,3} - mg c_{3,3} \\ J\ddot{\xi} = C(\xi, \dot{\xi})\dot{\xi} + \tau \end{gathered} \tag{1}$$

where $R_{B2E} \in SO(3)$ is rotation matrix which transfer the quadrotor coordinates from the body fixed frame to Earth fixed frame.

$$R_{B2E} = \begin{bmatrix} c_\theta c_\psi & c_\psi s_\phi s_\theta - c_\phi s_\psi & s_\phi s_\psi + c_\phi c_\psi s_\theta \\ c_\theta s_\psi & c_\phi c_\psi + s_\phi s_\theta s_\psi & c_\phi s_\theta s_\psi - c_\psi s_\phi \\ -s_\theta & c_\theta s_\phi & c_\phi c_\theta \end{bmatrix} \tag{2}$$

where s_i and c_i with $(i = \phi, \theta, \psi)$ donate the $\sin(i)$ and $\cos(i)$, respectively. The nonlinear Coriolis matrix $C(\xi, \dot{\xi}) = [c_{ij}] \in R^{3\times 3}$ is shown in [19]. The total lift $T_p \in R$ and the control torque $\tau = \left[\tau_\phi, \tau_\theta, \tau_\psi\right] \in R^3$. In practical applications, the total lift and the control torque are completely generated by adjusting the speed of four rotors.

$$\begin{gathered} T_p = k_f(\Omega_1^2 + \Omega_2^2 + \Omega_3^2 + \Omega_4^2) \\ \tau_\phi = l k_f(\Omega_2^2 - \Omega_4^2) \\ \tau_\theta = l k_f(\Omega_1^2 - \Omega_3^2) \\ \tau_\psi = k_t(\Omega_1^2 - \Omega_2^2 + \Omega_3^2 - \Omega_4^2) \end{gathered} \tag{3}$$

where Ω_i $(i = 1,2,3,4)$ donates the speed of rotor i.

Remark 1: *The quadrotor system is underactuated due to six degrees of freedom (DoF) while four independent inputs. Besides, it is easy to observe that the quadrotor dynamics is highly nonlinear and coupling. Therefore, the model-based control approaches are difficult to implement in practical applications.*

Define $u_p = \left[u_{px}, u_{py}, u_{pz}\right]^T \in R^3$ as a virtual position control input satisfies

$$\begin{bmatrix} u_{px} \\ u_{py} \\ y_{pz} \end{bmatrix} = T_p \begin{bmatrix} s_\phi s_\psi + c_\phi c_\psi s_\theta \\ c_\phi s_\theta s_\psi - c_\psi s_\phi \\ c_\phi c_\theta \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ mg \end{bmatrix} \tag{4}$$

The dynamic model of quadrotor can be rewritten as

$$\ddot{p} = m^{-1} u_p \tag{5}$$

$$\ddot{\xi} = J^{-1} C(\xi, \dot{\xi})\dot{\xi} + J^{-1}\tau \tag{6}$$

Remark 2: *The position dynamic equation in Eq. (5) is considered a linear system with the virtual position control input u_p. Therefore, it is possible to apply several linear* *control methods such as Linear Quadratic Regulator (LQR) control as [24] or Linear Quadratic Tracking (LQT) in the position control design. It is noted that the definition of virtual position control input signals does not simplify the nonlinear position dynamic model to linear model.*

The control objective of this paper is to design an optimal controller for the underactuated quadrotor system to achieve optimal trajectory tracking under the influence of dynamic uncertainties. Moreover, the optimal data-based RL controller is developed for both the position and attitude controllers (see Fig. 1). It means that the quadrotor can achieve the reference trajectory tracking with unknown dynamics. In fact, the quadrotor dynamic information is reconstructed by using the quadrotor system data.

III. Control Design

In this section, the optimal position controller and optimal attitude controller are designed based on RL algorithms. To achieve trajectory tracking, the hierarchical control structure of position and attitude controller which is shown in Fig. 1 is adopted, the position controller is to achieve the desired tracking and generates the references for the attitude controller, and the attitude controller is to achieve the reference attitude tracking. By using input-output quadrotor system data, the optimal solutions of the position and attitude controller can be learned without the knowledge of dynamic system. It means that the proposed controller-based RL is to handle the disadvantage of unknown dynamic uncertainty, which has not been studied in conventional methods.

A. Position Controller Design

Define $x_p = [p^T, \dot{p}^T]^T \in R^6$ as the position state of the quadrotor. The position subsystem in Eq. (5) can be rewritten as

$$\dot{x}_p = A_p x_p + B_p u_p \tag{7}$$

where $A_p = \begin{bmatrix} 0_{3\times3} & I_3 \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix} \in R^{6\times6}$ and $B_p = \begin{bmatrix} 0_{3\times3} \\ m^{-1} I_3 \end{bmatrix} \in R^{6\times3}$. Assume that the reference trajectory of the quadrotor can is generated by the system $\dot{x}_{pd} = A_{pd} x_{pd}$, with $x_{pd} = [p_d^T, \dot{p_d}^T] \in R^6$ as the reference position vector. Define $X_p = [x_p^T, x_{pd}^T]^T \in R^{12}$ as the argument position state of the quadrotor following

$$\begin{aligned} \dot{X}_p &= \overline{A}_p X_p + \overline{B}_p u_p \\ E_p &= C_p X_p \end{aligned} \tag{8}$$

where $E_p \in R^3$ as the position tracking error vector, $\overline{A}_p = diag([A_p, A_{pd}]) \in R^{12\times12}$, $B_p = [0_{3\times6} \quad {B_p}^T]^T \in R^{12\times3}$ and $C_p = [I_3 \quad 0_{3\times3} \quad -I_3 \quad 0_{3\times3}] \in R^{3\times12}$. To achieve the optimal position policy, a performance function with a discount factor for the argument position system is given as

$$V_p(X_p) = \int_t^\infty e^{-\gamma_p(\tau-t)} (E_p^T Q_p E_p + u_p^T R_p u_p) d\tau \tag{9}$$

where $\gamma_p > 0$, $Q_p^T = Q_p \in R^{3\times3}$ and $R_p^T = R_p \in R^{3\times3}$ are positive define matrix.

Remark 3: *The positive discount factor γ_p is required in the performance function to ensure that it remains bound when the desired reference does not approach zero as time goes to infinity.*

By using feedback control law $u_p = K_p X_p$ with $K_p \in R^{3\times12}$ is the control gain, the performance function is rewritten as

$$V_p(X_p) = \int_t^\infty e^{-\gamma_p(\tau-t)} X_p \left(\overline{Q}_p + K_p^T R_p K_p\right) X_p d\tau \tag{10}$$

where $Q_p = C_p^T Q_p C_p$, by minimizing the performance function, the optimal position control policy obtains as $u_p^* = K_p^* X_p$ with $K_p^* = -R_p \overline{B}_p \Pi$ and Π is the solution of the ARE equation.

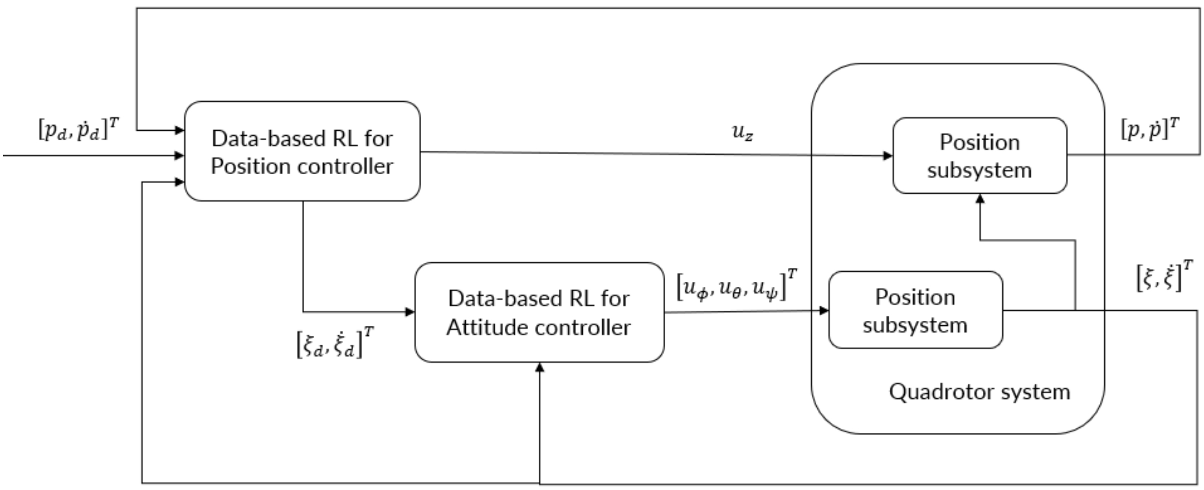

Fig. 1. The quadrotor control structure

$$\overline{A}^T{}_p\Pi + \Pi\overline{A}_p - \gamma_p\Pi + \overline{Q}_p - \Pi\overline{B}_pR_p^{-1}\overline{B}_p\Pi = 0 \tag{11}$$

The direct solution of the ARE equation in Eq. (11) requires the dynamic information of the position subsystem. However, the unknown dynamics resulted of unknown payloads, it is difficult to solve the ARE equation in Eq. (11) to obtain the optimal position control policy. To achieve the model-free optimal control problem for the position subsystem of the quadrotor, a data- based RL is used to learn the optimal position control policy u_p^*. The argument position system is rewritten as

$$\begin{aligned}\dot{X}_p &= \overline{A}_pX_p + \overline{B}_pu_p^i + \overline{B}_p(u_p^0 - u_p^i)\\ &= (\overline{A}_p + \overline{B}_pK_p^i)X_p + \overline{B}_p(u_p^0 - K_p^iX_p)\end{aligned} \tag{12}$$

where u_p^0 is stabilizing and exploring control input and K_p^i is the update control gain in i^{th} iteration. Taking the time derivative of discounted performance function in Eq. (10) along the argument position dynamic system in Eq. (12).

$$\begin{aligned}\frac{dV_p}{dt} &= \left(\Delta V_p^i\right)^T\left(\overline{A}_p + \overline{B}_pK_p^i\right)X_p\\ &\quad + \left(\Delta V_p^i\right)^T\overline{B}_p(u_p^0 - u_p^i)\\ &= -X_p^T\left(\overline{Q}_p + \left[K_p^i\right]^TR_pK_p^i\right)X_p + \gamma_pV_p\\ &\quad - 2\left(u_p^0 - K_p^iX_p\right)^TR_pK_p^{i+1}X_p\end{aligned} \tag{13}$$

By multiplying both sides Eq. (13) with $e^{-\gamma_pt}$ and integrating both side Eq. (13) in $[t, t+T]$, the tracking Bellman equation can be obtained as

$$\begin{aligned}&e^{-\gamma_pT}X_p^T(t+T)\Pi^iX_p(t+T) - X_p^T(t)\Pi^iX_p(t) =\\ &\int_t^{t+T} e^{-\gamma_p(\tau-t)}X_p^T\left(\overline{Q}_p + \left[K_p^i\right]^TR_pK_p^i\right)X_pd\tau\\ &-2\int_t^{t+T} e^{-\gamma_p(\tau-t)}\left(u_p - K_p^iX_p\right)^TR_pK_p^{i+1}X_pd\tau\end{aligned} \tag{14}$$

where T is the interval time. Compared with the ARE in Eq. (11), the Bellman equation in Eq. (14) does not require the knowledge of the dynamic information of the quadrotor. In fact, the quadrotor dynamics is reconstructed by using the input-output system data.

Algorithm 1: Data-based RL for position controller

1. Step 1: Initialization with $u_p^0 = u_p^s + u_p^e$ donate by stabilizing control input u_p^s and the exploring noise input u_p^e. Collect the system data X_p and u_p^0.
2. Solving the Bellman tracking equation by using K_p^i in this previous iteration to obtain Π^i and K_p^{i+1}.

$$\begin{aligned}&e^{-\gamma_pT}X_p^T(t+T)\Pi^iX_p(t+T) - X_p^T(t)\Pi^iX_p(t) =\\ &\int_t^{t+T} e^{-\gamma_p(\tau-t)}X_p^T\left(\overline{Q}_p + \left[K_p^i\right]^TR_pK_p^i\right)X_pd\tau\\ &-2\int_t^{t+T} e^{-\gamma_p(\tau-t)}\left(u_p^0 - K_p^iX_p\right)^TR_pK_p^{i+1}X_pd\tau\end{aligned} \tag{15}$$

3. Set $K_p^i = K_p^{i+1}$ and go to 2 until $\left\|K_p^i - K_p^{i+1}\right\| < \epsilon_p$.

Learning time is considered from the initial time to the convergence of **Algorithm 1.** Furthermore, under the influence of dynamic uncertainties, the proposed RL **Algorithms 1** are implemented with input-ouput quadrotor data collection in the practical system, which are utilized to solve the Eq. (15). Stabilizing and exploring control input is required in the RL methods following [20]. To achieve the initial stability for the quadrotor system, a PD controller is employed, and the exploring control input is a noise input, which is chosen to satisfy the persistency of the excitation (PE) condition. By using the Least-Squares method in [21], the solution of the Bellman tracking equation in Eq. (14) can be obtained and the convergence of **Algorithm 1** can be proven following to [21]. The computational efficiency is absolutely depended on data collection, which are implemented in many time periods. Although the initial stabilizing control makes the quadrotor system perform badly, it is only done to generate and collect input- output quadrotor system data to observe the quadrotor model. After the **Algorithm 1** convergence, the initial stabilizing and exploring control input is replaced by the optimal learning policy. Once the virtual position control policy is determined, following the virtual position control input is defined in Eq. (4), the attitude reference can be obtained as

$$\begin{aligned}T_p &= \frac{u_{pz}}{c_\phi c_\theta}\\ \phi_d &= \arcsin\left(\frac{c_\phi s_\theta s_\psi - u_{py}/T_p}{c_\psi}\right)\\ \theta_d &= \arcsin\left(\frac{-s_\phi s_\psi + u_{px}/T_p}{c_\psi c_\phi}\right)\end{aligned} \tag{16}$$

The desired yaw angle ψ_d usually sets to fixed constant value in practical applications. In this study, the yaw angle is set to zero.

B. Attitude control design

Define the attitude state vector $x_e = \left[\xi^T, \dot{\xi}^T\right]^T \in R^6$ of the quadrotor system. The dynamic model of the attitude subsystem in Eq. (6) is rewritten as

$$\dot{x}_e = f_e(x_e) + B_eu_e \tag{17}$$

where $f_e(x_e) = \begin{bmatrix}\dot{\xi}\\ J^{-1}C(\xi,\dot{\xi})\dot{\xi}\end{bmatrix} \in R^6$ is nonlinear smooth function, $B_p = \begin{bmatrix}0_{3\times3}\\ J^{-1}\end{bmatrix} \in R^{6\times3}$, and $u_e = \tau$ as attitude control input. Following the reference attitude in Eq. (15), let $x_{ed} = \left[\phi_d, \theta_d, \psi_d, \dot{\phi}_d, \dot{\theta}_d, \dot{\psi}_d\right]^T \in R^6$ as attitude reference vector satisfying $\dot{x}_{ed} = f_{ed}(x_{ed})$, with $f_{ed}(x_{ed}) \in R^6$ is a smooth function. Define $X_e = [x_e^T, x_{ed}^T]^T \in R^{12}$ as the argument attitude state vector of the quadrotor following (18).

$$\dot{X}_e = F_e(X_e) + \overline{B}_e u_e$$
$$E_e = C_e X_e \tag{18}$$

where $E_e \in R^3$ as the attitude tracking error vector, $F_e(X_e) = [f_e(x_e)^T \quad f_{ed}(x_{ed})^T]^T \in R^{12}$, $B_p = \begin{bmatrix} 0_{3\times 6} & B_e{}^T \end{bmatrix}^T \in R^{12\times 3}$ and $C_e = [I_3 \quad 0_{3\times 3} \quad -I_3 \quad 0_{3\times 3}] \in R^{3\times 12}$. To achieve the optimal attitude policy, a performance function with a discount factor for the argument position system is given as

$$V_e(X_e) = \int_t^{\infty} e^{-\gamma_e(\tau-t)}(E_e^T Q_e E_e + u_e^T R_e u_e)d\tau$$
$$= \int_t^{\infty} e^{-\gamma_e(\tau-t)}(X_e^T \overline{Q}_e X_e + u_e^T R_e u_e)d\tau \tag{19}$$

where $\overline{Q}_e = C_e^T Q_e C_e$, $Q_e^T = Q_e \in R^{3\times 3}$ and $R_e^T = R_e \in R^{3\times 3}$are positive define matrix. The problem is to find the optimal policy u_e to minimize the performance function in Eq. (19). Differentiating $V_e(X_e)$ in Eq. (19) and using Eq. (18), the Bellman equation as

$$H_e(V_e, u_e) = X_e^T \overline{Q}_e X_e + u_e^T R_e u_e - \gamma_e V_e + \Delta V_e^T (F_e + \overline{B}_e u_e) \tag{20}$$

where $\Delta V_e = \partial V_e / \partial X_e$. Let V_e^* as an optimal performance function. The HJB equation can be obtained as

$$\min_{u_e} H_e(V_e^*, u_e) = 0 \tag{21}$$

By different the HJB equation in Eq. (21) respect to the attitude control input u_e, the optimal control policy u_e^* can be obtained as

$$u_e^* = -\tfrac{1}{2} R_e^{-1} \overline{B}_e \Delta V_e^* \tag{22}$$

Substituting the optimal control policy u_e^* in Eq. (22) to the HJB equation in Eq. (21), the HJB equation yields

$$X_e^T \overline{Q}_e X_e - \gamma_e V_e + [\Delta V_e^*]^T F_e - \tfrac{1}{4}[\Delta V_e^*]^T \overline{B}_e R_e^{-1} \overline{B}_e^T \Delta V_e^* = 0 \tag{23}$$

The HJB equation in Eq. (23) is nonlinear because of the quadrotor system's nonlinear dynamics. As a result, solving the nonlinear HJB equation in Eq. (23) directly through mathematical analysis is challenging. Additionally, the nonlinear HJB equation in Eq. (23) demands an understanding of the quadrotor's dynamics, which is difficult to acquire in practical applications. To achieve the model-free optimal control problem for the attitude subsystem of the quadrotor, a data- based RL is used to learn the optimal attitude control policy. The argument attitude system is rewritten as

$$\dot{X}_e = F_e(X_e) + \overline{B}_e u_e^i + \overline{B}_e(u_e^0 - u_e^i) \tag{24}$$

where u_e^0 is stabilizing and exploring control input and u_e^i is the update control policy in i^{th} iteration. Taking the time derivative of discounted performance function in Eq. (19) along with the argument position dynamic system in Eq. (20).

$$\begin{aligned}\frac{dV_e(t)}{dt} &= (\Delta V_e^i)^T (F_e + \overline{B}_e u_e^i) \\ &\quad + (\Delta V_e^i)^T \overline{B}_e (u_e^0 - u_e^i) \\ &= -X_e^T \overline{Q}_e X_e - [u_e^i]^T R_e u_e + \gamma_e V_e^i \\ &\quad - 2(u_e^0 - u_e^i)^T R_p u_e^{i+1}\end{aligned} \tag{25}$$

It can be easily observed that the quadrotor dynamics F_e and $\overline{B}_e$ are replaced by the input-output system data. Integrating both sides Eq. (21) in $[t, t+T]$, the tracking Bellman equation can be obtained as

$$\begin{aligned}V_e^i(X_e(t+T)) - V_e^i(X_e(t)) &= -\int_t^{t+T} X_e^T(\tau) \overline{Q}_e X_e(\tau) d\tau \\ &\quad - \int_t^{t+T} [u_e^i(\tau)]^T R_e u_e^i(\tau) d\tau \\ &\quad + \gamma_e \int_t^{t+T} V_e(X_e(\tau)) d\tau \\ &\quad - 2\int_t^{t+T} \left(u_e^0 - u_e^i(\tau)\right)^T R_p u_e^{i+1}(\tau) d\tau\end{aligned} \tag{26}$$

where T is the interval time.

Remark 4: *The Bellman equation in Eq. (26) and the Bellman equation in Eq. (23) are equivalent and have the same optimal solution V_e^* and u_e^* following [23]. By using the tracking Bellman equation in Eq. (26), the optimal solution can be found without the knowledge of the quadrotor dynamics.*

Algorithm 2: Data-based RL for attitude controller

---

1. Step 1: Initialization with $u_e^0 = u_e^s + u_e^e$ donate by stabilizing control input u_e^s and the exploring noise input u_e^e satisfying PE condition. Collect the system data X_e and u_e^0.
2. Solving the Bellman tracking equation by using u_e^i in this previous iteration to obtain $V_e^i(X_e)$ and u_e^{i+1}.

$$\begin{aligned}&V_e^i(X_e(t+T)) - V_e^i(X_e(t)) \\ &= -\int_t^{t+T} X_e^T(\tau) \overline{Q}_e X_e(\tau) d\tau \\ &\quad - \int_t^{t+T} [u_e^i(\tau)]^T R_e u_e^i(\tau) d\tau \\ &\quad + \gamma_e \int_t^{t+T} V_e(X_e(\tau)) d\tau \\ &\quad - 2\int_t^{t+T} \left(u_e^0 - u_e^i(\tau)\right)^T R_p u_e^{i+1}(\tau) d\tau\end{aligned} \tag{27}$$

3. Set $u_e^i = u_e^{i+1}$ and go to 2 until $\left\| u_e^i - u_e^{i+1} \right\| < \epsilon_e$.

After the **Algorithm 1** convergence, the position control policy has been applied to the quadrotor system, the collected data from the argument attitude system is implemented, then, the optimal attitude control policy can be learned by **Algorithm 2.** The actor- critic neural networks (NNs) is employed to approximate the optimal solution of **Algorithm 2** following as

$$\begin{aligned} V_e^i(X_e) &= [W_c^i]^T \phi_c(X_e) \\ u_e^i(X_e) &= [W_a^i]^T \phi_a(X_e) \end{aligned} \tag{28}$$

where $W_c^i \in R^{lc\times 1}$ and $W_a^i \in R^{la\times 3}$ donate the weights of the critic and actor NNs, respectively, lc and la are the number of hide layer neurons of the critic and actor NNs. $\phi_c(X_e) \in R^{lc}$ and $\phi_a(X_e) \in R^{la}$ are action functions of critic and actor NNs. By using the Least-squares method in [21], the update optimal attitude policy can be learned under the PE condition. The convergence of **Algorithm 2** is proven as shown in [22].

IV. Simulation Result

This section is a simulation result which is built in MATLAB software to verify the effectiveness of the proposed optimal controller. This quadrotor is used in this simulation with the parameters following as $m = 2$ kg, $g = 9.8\ \mathrm{m/s^2}$, $J = 10^{-3} diag\ ([5, 5, 10.5])$Nm. The parameters of the position and attitude discounted performance functions are chosen as $\gamma_p = \gamma_e = 0.05$, $Q_p = 15I_3$, $Q_e = 100I_3$, and $R_p = R_e = I_3$. The dynamic of reference trajectory is chosen as $\dot{x}_{pd} = A_{pd}x_{pd}$ with the initial state $x_{pd} = [0, 0.5, 1, 0,0, 1]^T$ and the dynamic system matrix satisfy

$$A_{pd} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -0.25 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{29}$$

The initial condition of the quadrotor is $x_p|_{t=0} = [2,0,3,0,1,0]^T$and $x_e|_{t=0} = [0,0,0,0,0,0]^T$. The interval time satisfies that $T = 0.05\ s$. The action function of critic NN is chosen as fourth-order polynomials function and the action function of actual NN is chosen as first-order polynomials function. The exploring noised inputs are chosen as sum of multiple sin function with different frequently $u_p^e = 5 \times \sum_{i=1}^{20} \sin\ (w_{pi}t)$ and $u_e^e = \sum_{i=1}^{20} \sin\ (w_{ei}t)$ where w_{pi}, w_{ei} are random frequently.

To achieve the stability of the quadrotor system, a simple PD controller is chosen to generate the position and attitude system data. Fig. 2 and Fig. 3 show the position and attitude collected data from the quadrotor system. Although the position and attitude responses are low performance, the position and attitude quadrotor data are necessary for learning algorithms. After the collected data process is completed, the quadrotor data will be applied into **Algorithm 1** and **2**, and the training process will begin until the weights of Algorithms converge. Fig. 4 and Fig. 5 show the convergence of two RL Algorithms. After five iterations, the parameters of the control algorithms converge. After both two algorithms converge, the learned control policies will be applied to the quadrotor system. The position and attitude tracking errors with the optimal control law are shown in Fig. 6 and Fig. 7, respectively. Both the position and attitude tracking errors converge to zero within a second. To verify the tracking effectiveness, Fig. 8 shows the trajectory of the quadrotor in 3-dimensional (3D) space. It notes that this paper does not compare the performance of the proposed optimal controller with other controllers. This paper only focuses on studying an optimal control law for the quadrotor with unknown dynamics using reinforcement learning. The simulation results show that the proposed control method achieves the optimal quadrotor trajectory tracking with unknown dynamic information by using RL approach. By using the input-output system data, the control policies can be learned without the knowledge of system dynamics. However, the limitation of the method and the simulation is to not consider the influence of external disturbance, and it is also future work.

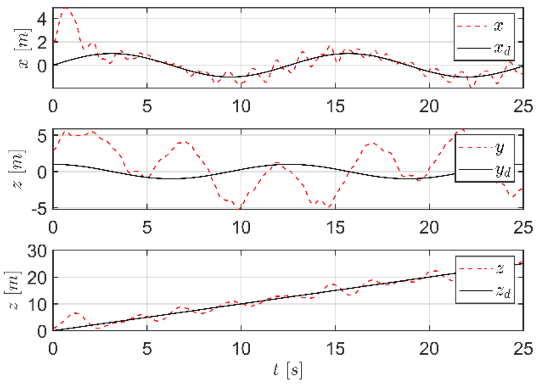

Fig. 2. The collected position data with the initial controller

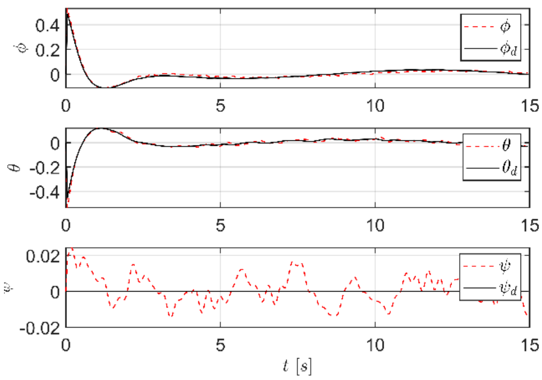

Fig. 3. The collected attitude data with the initial controller

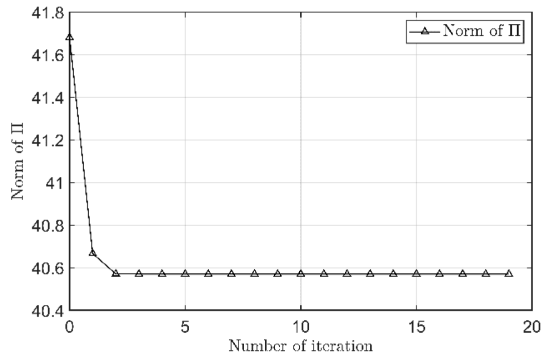

Fig. 4. The convergence of the Algorithm 1

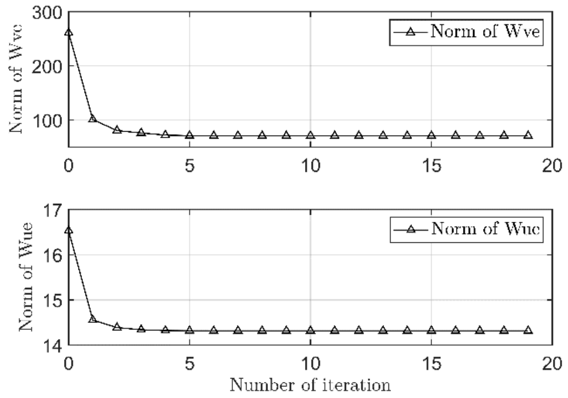

Fig. 5. The convergence of the Algorithm 2

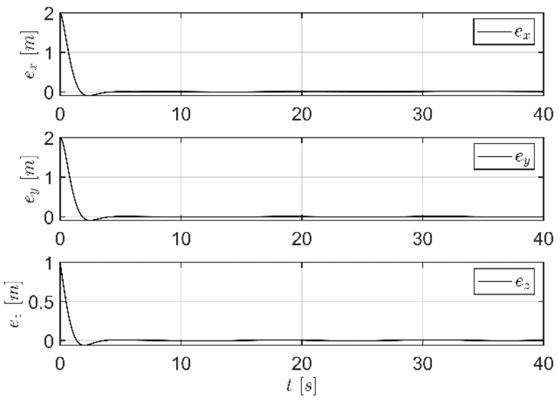

Fig. 6. The position tracking error with the proposed controller

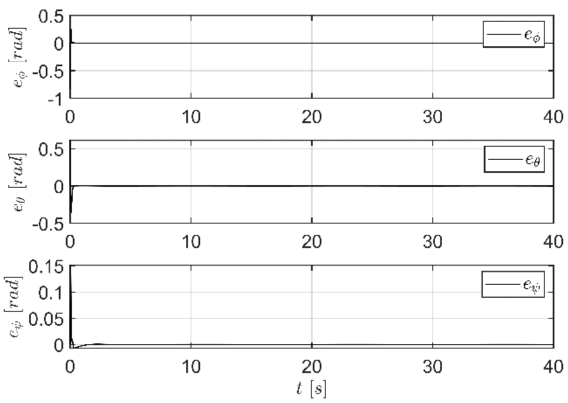

Fig. 7. The attitude tracking error with the proposed controller

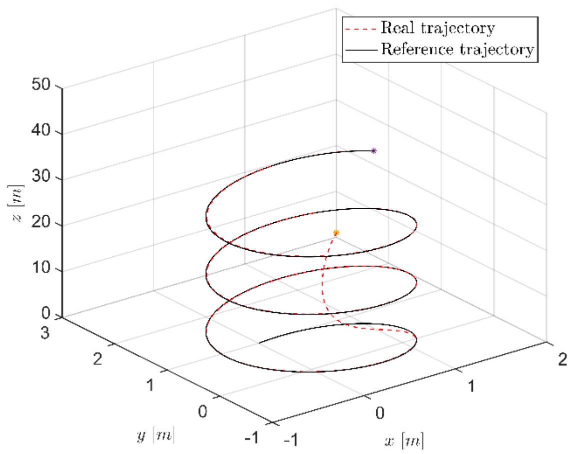

Fig. 8. The trajectory of quadrotor in 3D

V. Conclusion

This paper focuses on solving the optimal tracking control problem for the underactuated quadrotor system with unknown dynamics. A hierarchical control structure consists of the position and attitude control used to address this underactuated of the quadrotor system. By using data-based RL for both positions and attitude controllers, optimal control policies can be learned by using the input-output data system. Compared with traditional optimal control method, the proposed control method achieves the optimal objective without the knowledge of dynamic information. A simulation result is conducted to verify the optimal trajectory tracking for quadrotor with the proposed controller. However, this paper does not consider the influence of external disturbance in the quadrotor dynamics. To achieve the optimal tracking control for the underactuated quadrotor system under the influence of external disturbance, it requires to solve the Hamilton–Jacobi–Isaacs (HJI) equation. It is a future work and motivation of this paper.

Acknowledgment

This research was supported by the Research Foundation funded by the Thai Nguyen University of Technology.

References

[1] G. Hoffmann, S. Waslander, and C. Tomlin, "Quadrotor helicopter trajectory tracking control," *In AIAA guidance, navigation and control conference and exhibit*, p. 7410, 2008, doi: 10.2514/6.2008-7410.

[2] A. Das, F. Lewis, and K. Subbarao, "Backstepping approach for controlling a quadrotor using lagrange form dynamics," *Journal of Intelligent and Robotic Systems*, vol. 56, pp. 127-151, 2009, doi: 10.1007/s10846-009-9331-0.

[3] Z. Zuo, "Trajectory tracking control design with command-filtered compensation for a quadrotor," *IET control theory & applications*, vol. 4, no. 11, pp. 2343-2355, 2010, doi: 10.1049/iet-cta.2009.0336.

[4] M. Huang, B. Xian, C. Diao, K. Yang, and Y. Feng, "Adaptive tracking control of underactuated quadrotor unmanned aerial vehicles via backstepping," *Proceedings of the 2010 American Control Conference*, pp. 2076-2081, 2010, doi: 10.1109/ACC.2010.5r531424.

[5] Y. C. Liu and T. W. Ou, "Non-linear adaptive tracking control for quadrotor aerial robots under uncertain dynamics," *IET Control Theory & Applications*, vol. 15, no. 8, pp. 1126-1139, 2021, doi: 10.1049/cth2.12112.

[6] O. Mofid and S. Mobayen, "Adaptive sliding mode control for finite-time stability of quad-rotor UAVs with parametric uncertainties," *ISA transactions*, vol. 72, pp. 1-14, 2018, doi: 10.1016/j.isatra.2017.11.010.

[7] M. Labbadi and M. Cherkaoui, "Robust adaptive nonsingular fast terminal sliding-mode tracking control for an uncertain quadrotor UAV subjected to disturbances," *ISA transactions*, vol. 99, pp. 290-304, 2020, doi: 10.1016/j.isatra.2019.10.012.

[8] M. Labbadi and M. Cherkaoui, "Robust adaptive backstepping fast terminal sliding mode controller for uncertain quadrotor UAV," *Aerospace Science and Technology*, vol. 93, p. 105306, 2019, doi: 10.1016/j.ast.2019.105306.

[9] T. Huang *et al*., "Robust tracking control of a quadrotor UAV based on adaptive sliding mode controller," *Complexity*, vol. 2019, no. 1, p. 7931632, 2019, doi: 10.1155/2019/7931632.

[10] F. L. Lewis, D. Vrabie, and V. L. Syrmos. *Optimal control*. John Wiley & Sons, 2012, doi: 10.1002/9781118122631.

[11] B. Kiumarsi *et al*., "Optimal and autonomous control using reinforcement learning: A survey," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 6, pp. 2042-2062, 2017.

[12] N. O. Lambert, D. S. Drew, J. Yaconelli, S. Levine, R. Calandra, and K. S. J. Pister, "Low-Level Control of a Quadrotor With Deep Model-Based Reinforcement Learning," in *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4224-4230, 2019, doi: 10.1109/LRA.2019.2930489.

[13] J. Hwangbo, I. Sa, R. Siegwart, and M. Hutter, "Control of a Quadrotor With Reinforcement Learning," in *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 2096-2103, 2017, doi: 10.1109/LRA.2017.2720851.

[14] G. Wen, W. Hao, W. Feng and K. Gao, "Optimized Backstepping Tracking Control Using R,einforcement Learning for Quadrotor Unmanned Aerial Vehicle System," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 8, pp. 5004-5015, Aug. 2022, doi: 10.1109/TSMC.2021.3112688.

[15] S. Li, P. Durdevic, and Z. Yang, "Optimal tracking control based on integral reinforcement learning for an underactuated drone," *IFAC-PapersOnLine*, vol. 52, no. 8, pp. 55-60. 2019, doi: 10.1016/j.ifacol.2019.08.048.

[16] H. Zhang, F. L. Lewis, and A. Das, "Optimal Design for Synchronization of Cooperative Systems: State Feedback, Observer and Output Feedback," in *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1948-1952, 2011, doi: 10.1109/TAC.2011.2139510.

[17] Nusawardhana, S. H. Zak, and W. A. Crossley, "Nonlinear synergetic optimal controllers," *Journal of guidance, control, and dynamics*, vol. 30, no. 4, pp. 1134-1147, 2007, doi: 10.2514/1.27829.

[18] H. Liu, D. Li, Z. Zuo, and Y. Zhong, "Robust Three-Loop Trajectory Tracking Control for Quadrotors With Multiple Uncertainties," in *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2263-2274, 2016, doi: 10.1109/TIE.2016.2514339.

[19] H. Liu, Y. Wang, and J. Xi, "Completely distributed formation control for networked quadrotors under switching communication topologies," *Systems & Control Letters*, vol. 147, p. 104841, 2021, doi: 10.1016/j.sysconle.2020.104841.

[20] F. A. Yaghmaie and D. J. Braun, "Reinforcement learning for a class of continuous-time input constrained optimal control problems," *Automatica*, vol. 99, pp. 221-227, 2019, doi: 10.1016/j.automatica.2018.10.038.

[21] Y. Jiang and Z. P. Jiang "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699-2704, 2012, doi: 10.1016/j.automatica.2012.06.096.

[22] H. -N. Wu and B. Luo, "Neural Network Based Online Simultaneous Policy Update Algorithm for Solving the HJI Equation in Nonlinear H∞ Control," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 12, pp. 1884-1895, 2012, doi: 10.1109/TNNLS.2012.2217349.

[23] H. Modares and F. L. Lewis, "Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning," *Automatica* vol. 50, no. 7, pp. 1780-1792, 2014, doi: 10.1016/j.automatica.2014.05.011.

[24] H. Du, W. Zhu, G. Wen, Z. Duan, and J. Lü, "Distributed Formation Control of Multiple Quadrotor Aircraft Based on Nonsmooth Consensus Algorithms," in *IEEE Transactions on Cybernetics*, vol. 49, no. 1, pp. 342-353, 2019, doi: 10.1109/TCYB.2017.2777463.

[25] H. S. Nguyen, D. T. Le, V. T. Dang, D. H. Nguyen, A. T. Le, and T. L. Nguyen, "Advanced Motion Control of a Quadrotor Unmanned Aerial Vehicle based on Extended State Observer," *2023 International Conference on System Science and Engineering (ICSSE)*, pp. 381-386, 2023, doi: 10.1109/ICSSE58758.2023.10227212.

[26] X. Shao, G. Sun, W. Yao, J. Liu, and L. Wu, "Adaptive Sliding Mode Control for Quadrotor UAVs With Input Saturation," in *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 3, pp. 1498-1509, 2022, doi: 10.1109/TMECH.2021.3094575.

[27] S. Lian *et al*., "Adaptive Attitude Control of a Quadrotor Using Fast Nonsingular Terminal Sliding Mode," in *IEEE Transactions on Industrial Electronics*, vol. 69, no. 2, pp. 1597-1607, 2022, doi: 10.1109/TIE.2021.3057015.

[28] Z. Liu, X. Liu, J. Chen, and C. Fang, "Altitude Control for Variable Load Quadrotor via Learning Rate Based Robust Sliding Mode Controller," in *IEEE Access*, vol. 7, pp. 9736-9744, 2019, doi: 10.1109/ACCESS.2018.2890450.

[29] L. Yu, G. He, X. Wang, and L. Shen, "A Novel Fixed-Time Sliding Mode Control of Quadrotor With Experiments and Comparisons," in *IEEE Control Systems Letters*, vol. 6, pp. 770-775, 2022, doi: 10.1109/LCSYS.2021.3086389.

[30] J. Xiong, J. Pan, G. Chen, X. Zhang, and F. Ding, "Sliding Mode Dual-Channel Disturbance Rejection Attitude Control for a Quadrotor," in *IEEE Transactions on Industrial Electronics*, vol. 69, no. 10, pp. 10489-10499, 2022, doi: 10.1109/TIE.2021.3137600.

[31] B. Li, W. Gong, Y. Yang, B. Xiao, and D. Ran, "Appointed Fixed Time Observer-Based Sliding Mode Control for a Quadrotor UAV Under External Disturbances," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 1, pp. 290-303, 2022, doi: 10.1109/TAES.2021.3101562.

[32] A. Eltayeb, M. F. Rahmat, M. A. M. Basri, M. A. M. Eltoum, and M. S. Mahmoud, "Integral Adaptive Sliding Mode Control for Quadcopter UAV Under Variable Payload and Disturbance," in *IEEE Access*, vol. 10, pp. 94754-94764, 2022, doi: 10.1109/ACCESS.2022.3203058.

[33] N. P. Nguyen, H. Oh, and J. Moon, "Continuous Nonsingular Terminal Sliding-Mode Control With Integral-Type Sliding Surface for Disturbed Systems: Application to Attitude Control for Quadrotor UAVs Under External Disturbances," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 6, pp. 5635-5660, 2022, doi: 10.1109/TAES.2022.3177580.

[34] O. Mofid and S. Mobayen, "Adaptive Finite-Time Backstepping Global Sliding Mode Tracker of Quad-Rotor UAVs Under Model Uncertainty, Wind Perturbation, and Input Saturation," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 1, pp. 140-151, 2022, doi: 10.1109/TAES.2021.3098168.

[35] B. Liu, Y. Wang, O. Mofid, S. Mobayen, and M. H. Khooban, "Barrier Function-Based Backstepping Fractional-Order Sliding Mode Control for Quad-Rotor Unmanned Aerial Vehicle Under External Disturbances," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 1, pp. 716-728, 2024, doi: 10.1109/TAES.2023.3328801.

[36] W. Liu, M. Chen, and P. Shi, "Fixed-Time Disturbance Observer-Based Control for Quadcopter Suspension Transportation System," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 11, pp. 4632-4642, 2022, doi: 10.1109/TCSI.2022.3193878.

[37] V. K. Tripathi, A. K. Kamath, L. Behera, N. K. Verma, and S. Nahavandi, "An Adaptive Fast Terminal Sliding-Mode Controller With Power Rate Proportional Reaching Law for Quadrotor Position and Altitude Tracking," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 6, pp. 3612-3625, 2022, doi: 10.1109/TSMC.2021.3072099.

[38] Q. Han, Z. Liu, H. Su, and X. Liu, "Filter-Based Disturbance Observer and Adaptive Control for Euler–Lagrange Systems With Application to a Quadrotor UAV," in *IEEE Transactions on Industrial Electronics*, vol. 70, no. 8, pp. 8437-8445, 2023, doi: 10.1109/TIE.2022.3224167.

[39] H. Maqsood and Y. Qu. "Nonlinear disturbance observer based sliding mode control of quadrotor helicopter," *Journal of Electrical Engineering & Technology*, vol. 15, pp. 1453-1461, 2020, doi: 10.1007/s42835-020-00421-w.

[40] M. Moussa and M. Cherkaoui, "Robust integral terminal sliding mode control for quadrotor UAV with external disturbances," *International Journal of Aerospace Engineering*, vol. 2019, no.1, p. 2016416, 2019, doi: 10.1155/2019/2016416.

[41] H. Razmi and S. Afshinfar, "Neural network-based adaptive sliding mode control design for position and attitude control of a quadrotor UAV," *Aerospace Science and technology*, vol. 91, pp. 12-27, 2019, doi: 10.1016/j.ast.2019.04.055.

[42] X. Lin, Y. Wang, and Y. Liu, "Neural-network-based robust terminal sliding-mode control of quadrotor," *Asian Journal of Control*, vol. 24, no. 1, pp. 427-438, 2022, doi: 10.1002/asjc.2478.

[43] R. Falcón, H. Ríos, and A. Dzul, "Comparative analysis of continuous sliding-modes control strategies for quad-rotor robust tracking," *Control Engineering Practice*, vol. 90, pp. 241-256, 2019, doi: 10.1016/j.conengprac.2019.06.013.

[44] M. Vahdanipour and M. Khodabandeh, "Adaptive fractional order sliding mode control for a quadrotor with a varying load," *Aerospace Science and Technology*, vol. 86, pp. 737-747, 2019, doi: 10.1016/j.ast.2019.105306.

[45] X. Wang *et al*., "Quadrotor fault tolerant incremental sliding mode control driven by sliding mode disturbance observers," *Aerospace Science and Technology*, vol. 87, pp. 417-430, 2019, doi: 10.1016/j.ast.2019.03.001.

[46] Z. Hou, Zhiwei, P. Lu, and Z. Tu, "Nonsingular terminal sliding mode control for a quadrotor UAV with a total rotor failure," *Aerospace*

Science and Technology, vol. 98, p. 105716, 2020, doi: 10.1016/j.ast.2020.105716.

[47] G. Guoqiang *et al*., "Output Feedback Adaptive Dynamic Surface Sliding-Mode Control for Quadrotor UAVs with Tracking Error Constraints," *Complexity*, vol. 2020, no. 1, p. 8537198, 2020, doi: 10.1155/2020/8537198.

[48] L. Chen *et al*., "Robust trajectory tracking control for a quadrotor using recursive sliding mode control and nonlinear extended state observer," *Aerospace Science and Technology*, vol. 128, p. 107749, 2022, doi: 10.1016/j.ast.2022.107749.

[49] A. Noordin *et al*., "Adaptive PID controller using sliding mode control approaches for quadrotor UAV attitude and position stabilization," *Arabian Journal for Science and Engineering*, vol. 46, pp. 963-981, 2021, doi: 10.1007/s13369-020-04742-w.

[50] X. Ai and J. Yu. "Fixed-time trajectory tracking for a quadrotor with external disturbances: A flatness-based sliding mode control approach," *Aerospace Science and Technology*, vol. 89, pp. 58-76, 2019, doi: 10.1016/j.ast.2019.03.059.

[51] H. Hassani, A. Mansouri, and A. Ahaitouf, "Robust autonomous flight for quadrotor UAV based on adaptive nonsingular fast terminal sliding mode control," *International Journal of Dynamics and Control*, vol. 9, pp. 619-635, 2021, doi: 10.1007/s40435-020-00666-3.

[52] S. Ullah *et al*., "Robust integral sliding mode control design for stability enhancement of under-actuated quadcopter," *International Journal of Control, Automation and Systems*, vol. 18, pp. 1671-1678, 2020, doi: 10.1007/s12555-019-0302-3.

[53] M. Labbadi and M. Cherkaoui, "Novel robust super twisting integral sliding mode controller for a quadrotor under external disturbances," *International Journal of Dynamics and Control*, vol. 8, no. 3, pp. 805-815, 2020, doi: 10.1007/s40435-019-00599-6.

[54] H. Ghadiri, M. Emami, and H. Khodadadi, "Adaptive super-twisting non-singular terminal sliding mode control for tracking of quadrotor with bounded disturbances," *Aerospace Science and Technology*, vol. 112, p. 106616, 2021, doi: 10.1016/j.ast.2021.106616.

[55] S. C. Yogi, V. K. Tripathi, and L. Behera, "Adaptive Integral Sliding Mode Control Using Fully Connected Recurrent Neural Network for Position and Attitude Control of Quadrotor," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 12, pp. 5595-5609, 2021, doi: 10.1109/TNNLS.2021.3071020.

[56] J. Pan *et al*., "Attitude control of quadrotor UAVs based on adaptive sliding mode," *International Journal of Control, Automation and Systems*, vol. 21, no. 8, pp. 2698-2707, 2023, doi: 10.1007/s12555-022-0189-2.

[57] L. -X. Xu, H. -J. Ma, D. Guo, A. -H. Xie, and D. -L. Song, "Backstepping Sliding-Mode and Cascade Active Disturbance Rejection Control for a Quadrotor UAV," in *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 6, pp. 2743-2753, 2020, doi: 10.1109/TMECH.2020.2990582.

[58] M. Pouzesh and S. Mobayen, "Event-triggered fractional-order sliding mode control technique for stabilization of disturbed quadrotor unmanned aerial vehicles," *Aerospace Science and Technology*, vol. 121, p. 107337, 2022, doi: 10.1016/j.ast.2022.107337.

[59] Z. Zhu and S. Cao. "Back-stepping sliding mode control method for quadrotor UAV with actuator failure," *The Journal of Engineering*, vol. 2019, no. 22, pp. 8374-8377, 2019, doi: 10.1049/joe.2019.1084.

[60] X. Wu, B. Xiao, and Y. Qu, "Modeling and sliding mode-based attitude tracking control of a quadrotor UAV with time-varying mass," *ISA transactions*, vol. 124, pp. 436-443, 2022, doi: 10.1016/j.isatra.2019.08.017.

[61] O. Mofid *et al*., "Desired tracking of delayed quadrotor UAV under model uncertainty and wind disturbance using adaptive super-twisting terminal sliding mode control," *ISA transactions*, vol. 123, pp. 455-471, 2022, doi: 10.1016/j.isatra.2021.06.002.

[62] K. Alattas *et al*., "Barrier function adaptive nonsingular terminal sliding mode control approach for quad-rotor unmanned aerial vehicles," *Sensors* vol. 22, no. 3, p. 909, 2022, doi: 10.3390/s22030909.

[63] W. Gong *et al*., "Fixed-time integral-type sliding mode control for the quadrotor UAV attitude stabilization under actuator failures," *Aerospace Science and Technology* vol. 95, p. 105444, 2019, doi: 10.1016/j.ast.2019.105444.

[64] L. Chen *et al*., "Robust adaptive recursive sliding mode attitude control for a quadrotor with unknown disturbances," *ISA transactions*, vol. 122, pp. 114-125, 2022, doi: 10.1016/j.isatra.2021.04.046.

[65] V. K. Tripathi *et al*., "Finite-time super twisting sliding mode controller based on higher-order sliding mode observer for real-time trajectory tracking of a quadrotor," *IET Control Theory & Applications*, vol. 14, no. 16, pp. 2359-2371, 2020, doi: 10.1049/iet-cta.2020.0348.

[66] Z. Zhao *et al*., "High-order sliding mode observer-based trajectory tracking control for a quadrotor UAV with uncertain dynamics," *Nonlinear Dynamics*, vol. 102, pp. 2583-2596, 2020, doi: 10.1007/s11071-020-06050-2.

[67] T. Jiang, T. Song, and D. Lin, "Integral sliding mode based control for quadrotors with disturbances: Simulations and experiments," *International Journal of Control, Automation and Systems*, vol. 17, pp. 1987-1998, 2019, doi: 10.1007/s12555-018-0500-4.

[68] A. Najafi *et al*., "Adaptive barrier fast terminal sliding mode actuator fault tolerant control approach for quadrotor UAVs," *Mathematics*, vol. 10, no. 16, p. 3009, 2022, doi: 10.3390/math10163009.

[69] W. Alqaisi *et al*., "Three-loop uncertainties compensator and sliding mode quadrotor control," *Computers & Electrical Engineering*, vol. 81, p. 106507, 2020, doi: 10.1016/j.compeleceng.2019.106507.

[70] Z. Hou, X. Yu, and P. Lu, "Terminal sliding mode control for quadrotors with chattering reduction and disturbances estimator: Theory and application," *Journal of Intelligent & Robotic Systems*, vol. 105, no. 4, p. 71, 2022, doi: 10.1007/s10846-022-01679-0.

[71] G. Xu *et al*., "Adaptive prescribed performance terminal sliding mode attitude control for quadrotor under input saturation," *IET Control Theory & Applications*, vol. 14, no. 17, pp. 2473-2480, 2020, doi: 10.1049/iet-cta.2019.0488.

[72] Y. Nettari, M. Labbadi, and S. Kurt, "Adaptive backstepping integral sliding mode control combined with super-twisting algorithm for nonlinear UAV quadrotor system," *IFAC-PapersOnLine*, vol. 55, no. 12, pp. 264-269, 2022, doi: 10.1016/j.ifacol.2022.07.322.

[73] S. Ullah *et al*., "Neuro-adaptive fast integral terminal sliding mode control design with variable gain robust exact differentiator for under-actuated quadcopter UAV," *ISA transactions*, vol. 120, pp. 293-304, 2022, doi: 10.1016/j.isatra.2021.02.045.

[74] T. Kusznir and J. Smoczek, "Sliding mode-based control of a UAV quadrotor for suppressing the cable-suspended payload vibration," *Journal of Control Science and Engineering*, vol. 2020, no. 1, p. 5058039, 2020, doi: 10.1155/2020/5058039.

[75] G. Xu *et al*., "Adaptive sliding mode disturbance observer–based funnel trajectory tracking control of quadrotor with external disturbances," *IET Control Theory & Applications*, vol. 15, no. 13, pp. 1778-1788, 2021, doi: 10.1049/cth2.12159.

[76] S. C. Yogi, L. Behera, and S. Nahavandi, "Adaptive Intelligent Minimum Parameter Singularity Free Sliding Mode Controller Design for Quadrotor," in *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 2, pp. 1805-1823, 2024, doi: 10.1109/TASE.2023.3243660.

[77] Z. Li, X. Ma, and Y. Li, "Robust tracking control strategy for a quadrotor using RPD-SMC and RISE," *Neurocomputing*, vol. 331, pp. 312-322, 2019, doi: 10.1016/j.neucom.2018.11.070.

[78] B. Gao, Y. J. Liu, and L. Liu, "Adaptive neural fault-tolerant control of a quadrotor UAV via fast terminal sliding mode," *Aerospace Science and Technology*, vol. 129, p. 107818, 2022, doi: 10.1016/j.ast.2022.107818.

[79] Ö. Bingöl and H. M. Güzey, "Finite-time neuro-sliding-mode controller design for quadrotor uavs carrying suspended payload," *Drones* vol. 6, no. 10, p. 311, 2022, doi: 10.3390/drones6100311.

[80] B. Wang, *et al*., "An adaptive sliding mode fault-tolerant control of a quadrotor unmanned aerial vehicle with actuator faults and model uncertainties," *International Journal of Robust and Nonlinear Control*, vol. 33, no. 17, pp. 10182-10198, 2023, doi: 10.1002/rnc.6631.