IJRCS
International Journal of Robotics and Control Systems
Vol. 5, No. 6, 2025, pp. 3267-3283
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs
ASCEE

Robust Optimal Tracking Control for Wheel Mobile Robot: An Actor–Critic Approach

Ngoc Trung Dang [a,1,*], Van Dai Pham [a,2]

[a] Faculty of Electrical Engineering, Thai Nguyen University of Technology, Thai Nguyen, Vietnam
[1] trungcsktd@tnut.edu.vn; [2] phamvandai2772@gmail.com
[*] Corresponding Author

ARTICLE INFO

Article History
Received October 03, 2025
Revised November 13, 2025
Accepted December 25, 2025

Keywords
Non-Holonomic Wheel Mobile Robot;
Optimal Tracking Control;
Disturbance Observer;
Actor-Critic Based Control;
Reinforcement Learning;
Adaptive Dynamic Programming

ABSTRACT

The robust optimal trajectory tracking control problem for a non-holonomic wheel mobile robot usually remains a difficult problem due to the nonlinearity and non-holonomic constraint of this robot. Moreover, the presence of external disturbances acting on the system may lead to degraded control performance or even instability. To overcome these challenges, a hierarchical control structure, including kinematic and dynamic control loops, is developed to address the non-holonomic constraints of the robot. In the dynamic control loop, a nonlinear disturbance observer is proposed to estimate and compensate for disturbances. Furthermore, based on the actor-critic control strategy, an optimal controller is developed, which achieves the optimal trajectory tracking. redBy this way, the non-holonomic wheel mobile robot equipped with the proposed controller not only achieve the optimal trajectory tracking but also maintains robustness against disturbances. A simulation result built in the MATLAB software is conducted to confirm the superior performance of the proposed method. Compared to PD control approach, the proposed control approach achieves superior performance indices.

1. Introduction

Wheeled mobile robots (WMRs) have emerged as one of the most widely adopted robotic platforms owing to their superior maneuverability, ease of operation, fast and stable locomotion, high energy efficiency, and strong potential for system integration. Consequently, WMRs have been extensively utilized in a wide range of applications, including autonomous navigation, surveillance, smart warehousing, industrial inspection, environmental monitoring, autonomous delivery, and healthcare services [1]–[5].

The trajectory-tracking control of WMRs has long been recognized as a fundamental and challenging problem in robotics and control engineering. Designing an effective control strategy for WMRs requires addressing several inherent difficulties, such as nonlinear coupling effects and non-holonomic motion constraints [5], [40]–[43]. Besides, WMRs operating in complex and dynamic environments are inevitably subject to external disturbances, which further exacerbate the control challenge. To address these issues, a wide range of control methodologies have been proposed in the literature, including classical proportional–integral–derivative (PID) control [6], [44]–[46], sliding mode control (SMC) [7], [8], [47]–[49], fuzzy logic control [9], [50]–[53], and neural network–based control [10], [54]–[57]. Although these approaches have demonstrated satisfactory tracking performance, they still face several limitations. Although SMC [7], [8], [47]–[49] exhibits remarkable robustness and insensitivity to matched uncertainties and external disturbances, the discontinuous control action induces a chattering effect, which leads to high-frequency oscillations in the control input. This effect deteriorates tracking accuracy, reduces actuator lifespan, and may even destabilize the system when unmodeled dynamics are present. Some adaptive control strategies, such as fuzzy logic control [9], [50]–[53] or neural network-based control [10], [54]–[57], have been developed to achieve smooth control actions and adaptive compensation for modeling uncertainties and external disturbances. Another approach to handling disturbances is to employ a disturbance observer or an extended state observer, which aims to accurately estimate the disturbances and compensate for them in the control input. However, these methods rarely consider the optimal trajectory tracking problem, which remains an open issue in the control of nonholonomic wheel mobile robots (NWMRs).

To address the optimal control problem, it usually requires solving the Riccati equation for the linear system and the Hamilton-Jacobi-Bellman (HJB) equation for the nonlinear system. It should be noted that the HJB equation is a fundamental nonlinear partial differential equation used in optimal control theory to determine the value function of a control problem. However, solving the HJB to obtain the optimal solution usually remains a difficult problem, which is difficult or impossible to address through only mathematical analysis. This difficulty arises because the HJB is a highly nonlinear partial differential equation (PDE) that explicitly depends on both the system dynamics and the cost function. Therefore, traditional optimal control methods often provide only a theoretical foundation rather than being widely applied in practical implementations. In a control engineering context, Reinforcement Learning (RL) and Adaptive/Approximate Dynamic Programming (ADP) bridge the gap between traditional optimal control and adaptive control algorithms [11]–[13], [58]–[60] . The main goal is to design a learning algorithm to learn the optimal solution and cost function for a nonlinear system. Different from the traditional optimal control, the optimal control based on RL or ADP can find the solution of Riccati or HJB equations online in real time by minimizing the cost function. One of the representative approaches of optimal control methods based on RL or ADP is the critic-only control method [14], [15], [33], in which a single neural network (NN) is employed to approximate the cost function. It should be noted that in the critic-only learning scheme, the critic neural network learns to optimally minimize the cost function, but this does not guarantee the stability of the closed-loop system. Therefore, an additional stabilizing control term is usually incorporated to ensure system stability during the learning process. Consequently, this approach can only achieve near-optimal control performance rather than the exact optimal solution. To address this problem, a second neural network, called the actor NN, is designed to learn the optimal control policy from the critic NN while ensuring the stability of the system [16], [17]. This approach, which employs two neural networks, one critic NN to approximate the cost function and one actor NN to approximate the optimal control policy under the stability framework of the system, is referred to as the actor–critic method. However, these optimal control approaches address only the optimal regulation problem for the nominal system, which is assumed to be unaffected by external disturbances and model uncertainties. In practical scenarios, disturbances frequently affect the system; therefore, it is essential to develop an optimal control law that simultaneously ensures robustness against such disturbances.

Motivated by these considerations, this paper aims to develop a robust and optimal trajectory-tracking control strategy for NWMR, considering their inherent nonlinearities, nonholonomic constraints, and the influence of input disturbance. To cope with these issues, this work adopts a hierarchical control framework consisting of kinematic and dynamic control loops. The kinematic loop handles nonholonomic constraints, while the dynamic loop focuses on robustness and optimality. In the dynamic layer, a nonlinear disturbance observer is designed to estimate and compensate for disturbances in real time. Furthermore, an actor–critic learning-based optimal control strategy is developed to achieve optimal trajectory tracking performance. The main contribution of this paper is listed as follows

- In this work, a hierarchical control architecture is adopted, consisting of kinematic and dynamic control loops. It is important to highlight that previous research works [18], [19] mainly focused on controlling the kinematic model of nonholonomic wheeled mobile robots. Such simplifications may result in degraded tracking accuracy and reduced control performance in real-world applications.
- This paper proposes a nonlinear disturbance observer (DO) to estimate and compensate for input disturbances, thereby enhancing the robustness of the control law and enabling accurate trajectory tracking for the NWMR.
- Different from the nonlinear control methods [6]–[10], this paper considers the optimal trajectory tracking for NWMR. By using the actor-critic control strategy, the optimal control policy can be learned and applied to the robot system online in real-time.

This paper is organized as follows: In Section 2, the system model of NWMR and the control problem are established. In Section 3, the detailed control design is presented. In Section 4, a simulation result is provided. Finally, Section 5 concludes this paper.

2. Problem Formulation

2.1. Mathematical Model of NWMR

Considering a non-holonomic wheel mobile robot (NWMR) system as shown in Fig. 1, and its model parameters are listed in the Table 1. First of all, the position and orientation state vector is defined as shown in Table 2.

$$q = [x, y, \theta]^T \in \mathbb{R}^3 \tag{1}$$

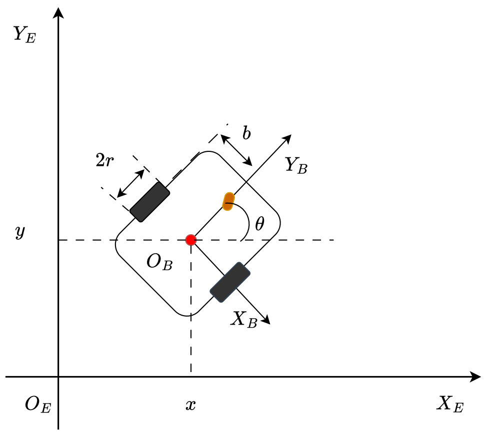

Fig. 1. The dynamics of NWMR platform

Table 1. The model parameters of the WMR

Symbols	Description	Value	Unit
m	Mass of NWMR	4.5	kg
I	Inertia of NWMR	2.7	m
b	Distance from mass center to each wheel	0.1	m
r	Wheel radius	0.05	kg.m^2

Table 2. Nomination

Symbols	Description	Unit
x	Position in the x-direction	m
y	Position in the y-direction	m
θ	Heading angle	rad
v	Linear velocity	m/s
w	Angular velocity	m/s
τ_r	Control torque input of right wheel	Nm
τ_l	Control torque input of left right wheel	Nm

It is well-known that the NWMR system is subjected to the non-holonomic constraint, which is expressed by

$$\dot{x}\sin(\theta) - \dot{y}\cos(\theta) = 0 \tag{2}$$

This non-holonomic constraint implies that the robot's motion is restricted to the direction along the rolling direction of the wheels and lateral motion. The kinematics of the NMWR system is given by

$$\dot{q} = \begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{\theta}\end{bmatrix} = \begin{bmatrix}\cos\theta & 0\\ \sin\theta & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}v\\ w\end{bmatrix} = S(q)\eta \tag{3}$$

According to [20]–[22], based on the Euler–Lagrange formulation, the dynamic model of the NWMR system is given by

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) + B(q)\tau_d = B(q)\tau - A^T(q)\lambda \tag{4}$$

Where $\tau_d \in \mathbb{R}^2$ is the vector of disturbances and $\tau = [\tau_r, \tau_l]^T \in \mathbb{R}^2$ is the control input vector. The matrices $M(q)^{3\times3}$ is the inertia matrix, $C(q,\dot{q})^{3\times3}$ is the Coriolis matrix, $G(q) \in \mathbb{R}^{3\times1}$ is the gravitational vector, $B(q) \in \mathbb{R}^{3\times2}$ is the input transformation matrix, $A(q) \in \mathbb{R}^{1\times3}$ is the constraint matrix, and λ is a Lagrange multiplier. According to [21], [22], these terms $M(q)$, $C(q,\dot{q})$, $G(q)$, $B(q)$, $A(q)$ and λ can be described as

$$\begin{gathered} M(q) = \begin{bmatrix}m & 0 & 0\\ 0 & m & 0\\ 0 & 0 & I\end{bmatrix};\quad C(q,\dot{q}) = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix};\quad G(q) = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix} \\ B(q) = \frac{1}{b}\begin{bmatrix}\cos(\theta) & \cos(\theta)\\ \sin(\theta) & \sin(\theta)\\ r & -r\end{bmatrix};\quad A^T(q) = \begin{bmatrix}-\sin(\theta)\\ \cos(\theta)\\ 0\end{bmatrix} \\ \lambda = -m(\dot{y}\sin(\theta) + \dot{x}\cos(\theta))\dot{\theta} \end{gathered} \tag{5}$$

2.2. Problem Statement

The non-holonomic constraint (2) can be rewritten by

$$A(q)\dot{q} = 0 \tag{6}$$

Meanwhile, we have the fact $S(q)^T A(q) = 0$. Taking the time derivative of (2), we obtain

$$\ddot{q} = [A^T(q), 0_{3\times1}]\eta + S(q)\dot{\eta} \tag{7}$$

Multiplying both sides of (4) with $S(q)^T$, the system model of NWMR is rewritten as follows

$$\begin{aligned} \dot{q} &= S(q)\eta \\ \dot{\eta} &= [\bar{M}]^{-1}\bar{B}\tau + [\bar{M}]^{-1}\bar{\tau}_d \end{aligned} \tag{8}$$

Where

$$\bar{M} = S^T(q)M(q)S^T(q) = \begin{bmatrix} m & 0 \\ 0 & I \end{bmatrix}$$
$$\bar{B} = S^T(q)B(q) = \frac{1}{b}\begin{bmatrix} 1 & 1 \\ r & -r \end{bmatrix} \quad (9)$$
$$\bar{\tau}_d = -S^T(q)B(q)\tau_d$$

It is worth emphasizing that the matrices $\bar{M}$ and $\bar{B}$ are positive definite matrices.

Definition 1 Let's introduce a smooth trajectory reference for the NMWR

$$q_d = [x_d, y_d, \theta_d]^T \quad (10)$$

Which must satisfy the non-holonomic constraint (2). In this case, we call q_d feasible; otherwise, it is non-feasible [23].

In this paper, it is assumed that the trajectory reference is generated by a trajectory generator, and its dynamics are described by

$$\begin{aligned} \dot{x}_d &= v_r \cos(\theta_d) \\ \dot{y}_d &= v_r \sin(\theta_d) \\ \dot{\theta}_d &= w_r \end{aligned} \quad (11)$$

Where $\eta_r = [v_r, w_r]^T \in \mathbb{R}^2$ is the vector of desired reference velocities.
Control problem: The main objective of this paper is to design an optimal controller to achieve both accurate trajectory tracking with respect to the reference trajectory (11) and robustness against the influence of disturbances. Consequently, the position and orientation errors remain within predefined bounds, expressed as

$$\begin{aligned} \lim_{t\to\infty} ||[x(t) - x_d(t), y(t) - y_d(t)]^T|| &\le \epsilon_{xy} \\ \lim_{t\to\infty} ||\theta(t) - \theta_d(t)|| &\le \epsilon_\theta \end{aligned} \quad (12)$$

3. Control Design

For the WMR system modeled by (8), a hierarchical control structure is proposed, including kinematic and dynamic control loops, as shown in Fig. 2. The kinematic controller is designed to generate the desired velocity reference for the dynamic control loop. By designing the dynamic controller to optimally track the desired velocity reference, the accurate trajectory tracking objective can be achieved. Furthermore, a disturbance observer is designed to compensate for the disturbances. The NWMR can achieve both the optimal trajectory tracking and the robustness against unknown disturbances.

3.1. Kinematic Control Loop

Inspired by [22], the tracking error of the WMR $e_q = [e_x, e_y, e_\theta]^T \in \mathbb{R}^3$ is formulated by

$$\begin{aligned} e_x &= (x_r - x)\cos\theta + (y_r - y)\sin\theta \\ e_y &= -(x_r - x)\sin\theta + (y_r - y)\cos\theta \\ e_\theta &= \theta_r - \theta \end{aligned} \quad (13)$$

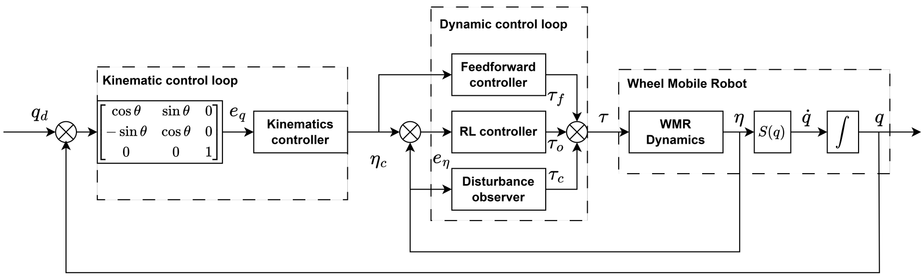

Fig. 2. The hierarchical control structure of the NWMR

For the kinematic dynamics (3) and the reference trajectory (11), the kinematic tracking error dynamics of the NWMR is formulated by

$$\begin{aligned} \dot{e}_x &= e_y w + v_r \cos e_\theta - v \\ \dot{e}_y &= -e_x w + v_r \sin e_\theta \\ \dot{e}_\theta &= w_r - w \end{aligned} \tag{14}$$

Hence, the desired linear and angular velocities $\eta_d = [v_d, w_d]^T \in \mathbb{R}^2$ are designed as follows

$$\begin{aligned} v_d &= c_1 e_x + v_r \cos(e_\theta) \\ w_d &= w_d + c_2 v_r e_y + c_3 v_r \sin(e_\theta) \end{aligned} \tag{15}$$

Where c_1, c_2 and c_3 are positive constants.

Theorem 1 For the kinematic tracking error dynamics (14), in the term of the virtual velocities (15), the tracking error e_q converges asymptotically to zero as $t \to \infty$.

Proof. Choosing the Lyapunov candidate function as follows

$$V_1 = \frac{1}{2}e_x^2 + \frac{1}{2}e_y^2 + \frac{1}{c_2}\big(1 - \cos(e_\theta)\big) \tag{16}$$

Taking the time derivative of (16) along the tracking error trajectory (14) using (17), it yileds

$$\begin{aligned} \dot{V}_1 =& e_x \dot{e}_x + e_y \dot{e}_y + \frac{1}{c_2}\sin(e_\theta)\dot{e}_\theta \\ =& e_x(e_y w + v_r \cos e_\theta - v) + e_y(-e_x w + v_r \sin e_\theta) \\ &+ \frac{1}{c_2}\sin(e_\theta)(w_r - w) \\ =& -c_1 e_x^2 - \frac{c_3}{c_2} v_r \sin^2(e_\theta) \end{aligned} \tag{17}$$

It is noted that the desired linear velocity v_r is positive, thus $V_1 \leq 0$. Based on the Lyapunov stability, the kinematic tracking error dynamics (14) applied to the virtual velocities (17) is asymptotically stable.

3.2. Dynamic Control Loop

The main control objective of this section is to design an optimal controller to achieve the optimal tracking with respect to the desired velocities (15). However, due to the influence of the input disturbance, a nonlinear disturbance observer (DO) is proposed to estimate and compensate for these disturbances. Furthermore, in the optimal control framework, it is noted that the actor-critic approach does not guarantee trajectory tracking with respect to a time-varying reference trajectory. To tackle the challenge, a feedforward control term is added to the control input to ensure accurate trajectory tracking.

3.2.1. Disturbance Observer Design

In this section, a nonlinear disturbance observer (DO) is utilized to estimate disturbances $d(t) = \bar{M}^{-1}\bar{\tau}_d$, and its formulation is given by

$$\begin{aligned} \hat{d} &= \zeta + r(\eta) \\ \dot{\zeta} &= -\frac{\partial r(\eta)}{\partial \eta}\Big(\zeta + r(\eta) + [\bar{M}]^{-1}\bar{B}\tau\Big) \end{aligned} \tag{18}$$

Where $\zeta \in \mathbb{R}^3$ is the internal states, $\hat{d}$ is the estimated value of d, and $r(\eta)$ is a nonlinear function.

Assumption 1 The disturbance is assumed to be slowly time-varying, such that its time derivative $\dot{d}$ is small.

Remark 1 It is worth emphasizing that the **Assumption**. 1 is well-known in disturbance observer design [28]–[32], as it facilitates the stability analysis and estimated disturbance convergence.

Theorem 2 Under the **Assumption**. 1, the disturbance estimation error $\tilde{d} = d - \hat{d}$ is ultimately uniformly bounded (UUB) stable.

Proof. Select the Lyapunov candidate function as follows:

$$V_2 = \frac{1}{2}\tilde{d}^T\tilde{d} \tag{19}$$

Taking the time derivative of V_2 using (18), it yields

$$\begin{aligned} \dot{V}_2 =& \tilde{d}^T\dot{\tilde{d}} \\ =& -\tilde{d}^T\Big(\dot{\zeta} + \dot{r}(\eta)\Big) + \tilde{d}^T\dot{d} \\ \le& -\tilde{d}^T\bigg(-\frac{\partial r(\eta)}{\partial \eta}\Big(\zeta + r(\eta) + [\bar{M}]^{-1}\bar{B}\tau\Big) + \frac{\partial r(\eta)}{\partial \eta}\dot{\eta}\bigg) + \frac{1}{2}\tilde{d}^T\tilde{d} + \frac{1}{2}\dot{d}^T\dot{d} \\ =& -\tilde{d}^T\bigg(-\frac{\partial r(\eta)}{\partial \eta}\Big(\hat{d} + [\bar{M}]^{-1}\bar{B}\tau\Big) + \frac{1}{2}\tilde{d}^T\tilde{d} + \frac{\partial r(\eta)}{\partial \eta}\Big([\bar{M}]^{-1}\bar{B} + d\Big)\bigg) + \frac{1}{2}\dot{d}^T\dot{d} \\ \le& -\tilde{d}^T\bigg(\frac{\partial r(\eta)}{\partial \eta} - \frac{1}{2}\bigg)\tilde{d} + \frac{1}{2}\dot{d}^T\dot{d} \end{aligned} \tag{20}$$

Let $\Xi = \partial r(\eta)/\partial \eta$ and $\epsilon = 0.5\dot{d}^T\dot{d}$. Hence, we obtain

$$\dot{V}_2 \le -2\Big(\lambda_{\min}(\Xi) - 0.5\Big)V_2 + \epsilon \tag{21}$$

Where $\lambda_{\min}(\Xi)$ is the minimum eigenvalue of Ξ. By choosing Ξ as a positive definite matrix satisfying $\lambda_{\min}(\Xi) > 0.5$, based on the Lyapunov stability theory, the disturbance estimation is UBB stable.

Remark 2 It is worth emphasizing that the boundary of ϵ is small under the **Assumption**. 1, which implies that the disturbance estimation error converges exponentially to a small neighborhood around the origin.

3.2.2. Optimal Control Design

Let $e_\eta = [e_v, e_w]^T = \eta - \eta_d$ as the velocity tracking error vector, and its dynamics can be expressed by

$$\dot{e}_\eta = [\bar{M}]^{-1}\bar{B}\tau + d - \dot{\eta}_d \tag{22}$$

As mentioned earlier, the control input is decomposed into three components: the disturbance compensation term τ_c, the feedforward term τ_f, and the optimal term τ_o, as follows

$$\tau = \tau_c + \tau_f + \tau_o \tag{23}$$

with

$$\begin{aligned} \tau_c &= -\big([\bar{M}]^{-1}\bar{B}\big)^{-1}\hat{d} \\ \tau_f &= \big([\bar{M}]^{-1}\bar{B}\big)^{-1}\dot{\eta}_d \end{aligned} \tag{24}$$

Subsituting (23) to (22), it yields

$$\dot{e}_\eta = [\bar{M}]^{-1}\bar{B}\tau_o + \tilde{d} \tag{25}$$

It should be noted that the disturbance estimation error $\tilde{d}$ will converge to a small neighborhood of zero as $t \to \infty$. Therefore, by eliminating the term $\tilde{d}$, the system dynamics of the NWMR is rewritten as a nominal system, which is described by

$$\dot{e}_\eta = f + g\tau_o \tag{26}$$

Where $f = [0, 0, 0]^T$ and $g = [\bar{M}]^{-1}\bar{B}$. By defining the control input (23) with the disturbance compensation term τ_c and the feedforward control term τ_f in (24), the control objective of the dynamic control loop is to develop an optimal controller for the nominal system (26) to achieve the optimal tracking with respect to the desired velocities (15). Firstly, let's define an infinite-horizon cost function, which is given by

$$J(e_\eta, \tau_o) = \int_0^\infty \big(e_\eta^T Q e_\eta + \tau_o^T R \tau_o\big) dt \tag{27}$$

Where $Q \in \mathbb{R}^{3\times3}$ and $R \in \mathbb{R}^{3\times3}$ are positive-definite symmetric matrices. The goal is to design an optimal controller to ensure the tracking error e_η converges to the origin by minimizing the cost function (28). Hence, the value function can be defined as

$$V(e_\eta, \tau_o) = \int_t^\infty \big(e_\eta^T Q e_\eta + \tau^T R \tau\big) ds \tag{28}$$

To derive the necessary conditions for optimality, the Hamiltonian function is introduced as follows

$$H(e_\eta, \tau, \frac{\partial V^T}{\partial e_\eta}) = e_\eta^T Q e_\eta + \tau_o^T R \tau_o + \left[\frac{\partial V}{\partial e_\eta}\right]^T (f + g\tau_o) \tag{29}$$

Then, the optimal value function $V^*(e_\eta)$ satisfies the HJB equation

$$\min_{\tau_o} \Big[H(e_\eta, \tau_o, \nabla V^*_{e_\eta})\Big] = 0 \tag{30}$$

Where $\nabla V_{e_\eta} = \partial V / \partial e_\eta$. According to the Bellman optimality principle, the optimal control policy can be given by

$$\tau_o^* = -\frac{1}{2} R^{-1} g^T \nabla V^* \tag{31}$$

Vol. 5, No. 6, 2025, pp. 3267-3283

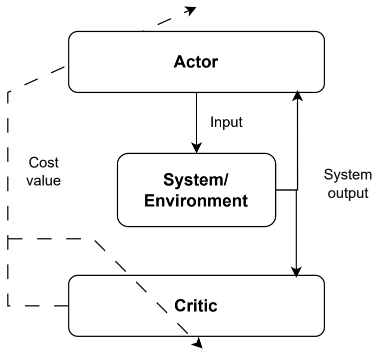

Fig. 3. Actor-critic control structure

To obtain the optimal control policy, it requires solving the HJB for the value function (28), which remains a difficult problem due to the nonlinearity of the HJB equation and the quadratic in the cost function gradient [24] shown in Fig. 3. To overcome this challenge, an actor-critic strategy can be utilized to find the optimal control policy as follows

$$\begin{aligned} \hat{V}(e_\eta) &= \hat{W}_c^T \Phi(e_\eta) \\ \hat{\tau}_o(e_\eta) &= -\frac{1}{2} R^{-1} g^T \left[\frac{\partial \Phi(e_\eta)}{\partial e_\eta}\right]^T \hat{W}_a \end{aligned} \tag{32}$$

Where $\hat{W}_c = [\hat{W}_{c1}, \hat{W}_{c2}, ..., \hat{W}_{cn}]^T \in \mathbb{R}^{n\times 1}$ and $\hat{W}_a = [\hat{W}_{a1}, \hat{W}_{a2}, ..., \hat{W}_{an}]^T \in \mathbb{R}^{n\times 1}$ are the critic and actor weight matrices, respectively; $\Phi(e_\eta) = [\phi_1(e_\eta), \phi_2(e_\eta), ..., \phi_n(e_\eta)]^T \in \mathbb{R}^{n\times 1}$ is the action function vector. The updated laws for both weight matrices are given as

$$\begin{aligned} \dot{\hat{W}}_c &= -k_c \Gamma \frac{w w^T}{1 + \sigma w^T \Gamma w} \delta_{hjb} \\ \dot{\hat{W}}_a &= -k_{a1} \frac{1}{\sqrt{1 + w^T w}} \frac{\partial \Phi}{\partial e_\eta} g R^{-1} g^T \left[\frac{\partial \Phi}{\partial e_\eta}\right]^T (\hat{W}_a - \hat{W}_c)\delta_{hjb} - k_{a2}(\hat{W}_a - \hat{W}_c) \end{aligned} \tag{33}$$

With

$$\begin{aligned} w &= \frac{\partial \Phi}{\partial e_\eta}(f + g\hat{\tau}_o) \\ \delta_{hjb} &= e_\eta^T Q e_\eta + \hat{\tau}_o^T R \hat{\tau}_o + \hat{W}_c^T w \end{aligned} \tag{34}$$

Where $k_c, k_{a1}, k_{a2}, \sigma$ are positive constants and $\Gamma \in \mathbb{R}^{n\times n}$ is an adaptive gain, which is updated by an adaptive law as follows

$$\begin{aligned} \dot{\Gamma} &= -k_c \Gamma \frac{w w^T}{1 + \sigma w^T \Gamma w} \Gamma \\ \Gamma(0) &= \kappa I, \quad \kappa > 0 \end{aligned} \tag{35}$$

According to [37]–[39], under the Persistent Excitation (PE) condition, both critic and actor weight matrices converge, and the optimal control policy can be found.

Remark 3 According to [16], [36], by adding the PE noise input to the control input, the actor-critic algorithm converges. In general RL-based control problem, the exploration noise must be required to satisfy the PE condition to learn the optimal control policy. Several PE noise input have been adopted, such as random noise [25], sum of sinusoidal functions with different frequencies [26] or exponentially decreasing probing noise [27],.

4. Simulation Result

redIn this section, a brief simulation result is presented to verify the effectiveness of the proposed controller in both trajectory tracking and disturbance robustness objectives. Furthermore, a performance comparison between the proposed method and the conventional PD controller is carried out to highlight the superior tracking performance of the proposed controller.

The system parameter of the NWMR is described in the Table. 1. The initial condition of the NWMR is set up as $q(0) = [1, 1, \pi/6]^T$ and $\eta(0) = [0, 0]^T$. It is assumed that the input disturbances $\tau_d = [\tau_{d1}, \tau_{d2}]^T$ acting on the NWMR system can be described by

$$\begin{aligned} \tau_{d1} &= 0.3sin(t + 0.5)sin(0.2t) \\ \tau_{d2} &= 0.2cos(t + 0.4)sin(0.2t) \end{aligned} \tag{36}$$

The reference trajectory is generated by a trajectory generator (11) with $v_r = 0.5$ and $w_r = 0.4$. The control parameters of both kinematic and dynamic controllers are described in the Table. 3. The critic and actor weight matrices are initialized as $\hat{W}_c = [10, 1, 20]^T$ and $\hat{W}_a = [0, 0, 0]^T$, respectively.

Table 3. The control parameters of both kinematic and dynamic controllers

Parameter	Value	Parameter	Value	Parameter	Value
c_1	2	c_2	4	c_3	5
Ξ	$10I_2$	Q	$5I_2$	R	I_2
k_c	10	k_{a1}	50	k_{a2}	20
σ	0.005	κ	2000		

Similar to [34], [35], the action function is chosen as $\Phi(e_\eta) = [e_v^2, e_v e_w, e_w^2]^T$. During the first 2 seconds, an exploring disturbance $\tau_{pe} = [\tau_{pe1}, \tau_{pe2}]^T$ satisfying the PE condition is added to the control input, which is described as follows

$$\begin{aligned} \tau_{pe1} &= 0.1e^{-0.5t}\big(cos(t)sin(t)^2 + cos(0.1t)sin(2t)^2\big) \\ \tau_{pe2} &= 0.1e^{-0.5t}\big(cos(0.5t)sin(-1.2t)^2 + sin(t)^5\big) \end{aligned} \tag{37}$$

The result in Fig. 8 shows the disturbance estimation. After 5 seconds, the estimated values of disturbances converge to the real values. The convergence of both critic and actor weight matrices is shown in Fig. 9. After 2 seconds, both critic and actor weight matrices converge to the value of $\hat{W}_c = \hat{W}_a = [2.28, -0.03, 20.84]^T$. The position and orientation responses are shown in Fig. 4, and the responses of the linear and angular velocities are illustrated in Fig. 5. In addition, The position and orientation tracking errors are shown in Fig. 6, and thethe linear and angular velocity tracking errors are illustrated in Fig. 7. During the first 2 seconds, the tracking performance of the proposed controller is relatively poor due to the learning process. However, after this initial period, it can be clearly seen that the proposed controller achieves significantly improved tracking performance. After 5 seconds, the state responses track to the desired reference, and the NWMR achieves accurate trajectory tracking. As shown in Fig. 6 and Fig.7, the proposed controller achieves higher tracking performance with the tracking error converging to zero when $t \to \infty$, while the PD controller does not ensure the convergence of tracking error due to the influence of disturbances.

To analyze the tracking performance of the quadrotor under the different controllers, four performance indices in are employed as follows:

- Integral of Squared Error (ISE), this index penalizes large errors more heavily and emphasizes overall error energy.

$$\mu_{ISE}^i = \int_0^T [e_i(t)]^2 dt, \qquad i = x, y, \theta, u, v \tag{38}$$

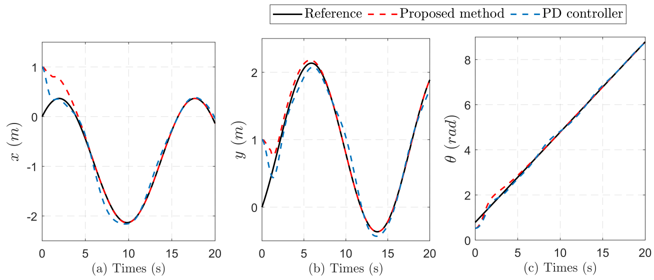

Fig. 4. The position and orientation responses

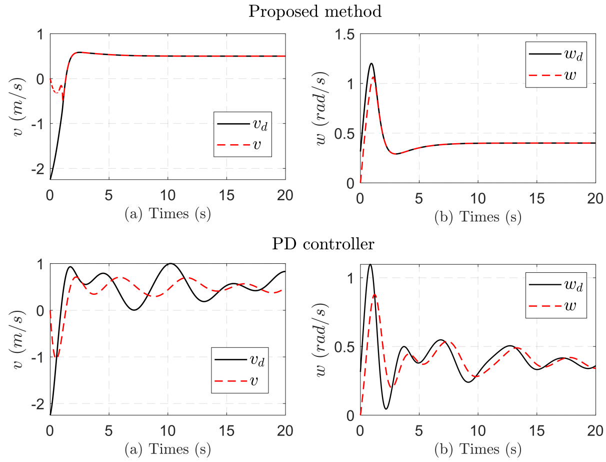

Fig. 5. The responses of the linear and angular velocities

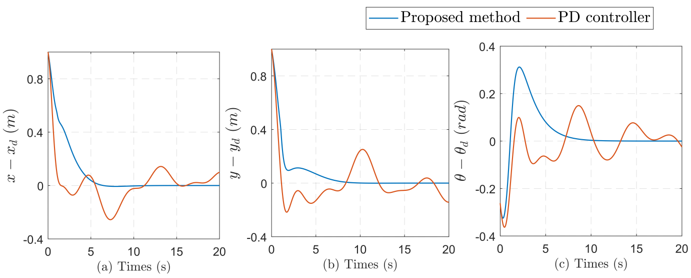

Fig. 6. The position and orientation tracking errors

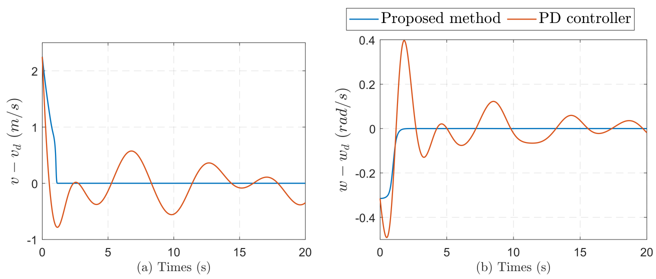

Fig. 7. The responses of the linear and angular velocity tracking errors

- Integral of Time-weighted Squared Error (ITSE):

$$\mu^{i}_{ITSE} = \int_{0}^{T} t[e_i(t)]^2 dt, \qquad i = x, y, \theta, u, v \tag{39}$$

It increases the weight of errors occurring at later times, encouraging faster settling.

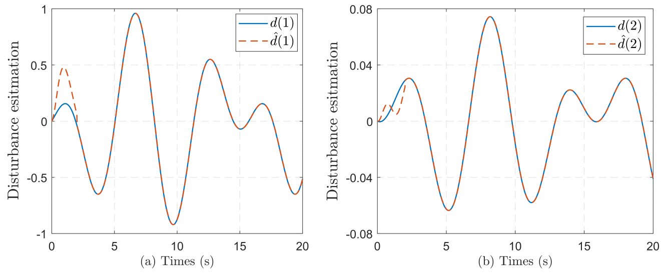

Fig. 8. The disturbance estimation

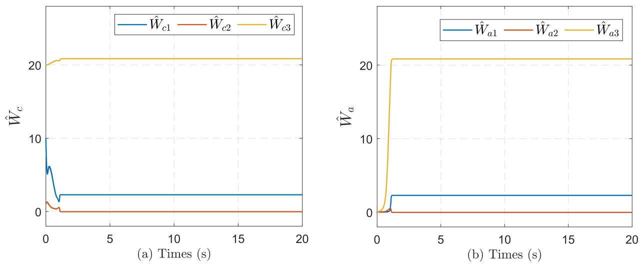

Fig. 9. The convergence of both critic and actor weight matrices

- Integral of Squared Error (IAE), this criterion reduces long-term steady-state errors and overshoot, improving transient behavior.

$$\mu_{IAE}^{p} = \int_{0}^{T} |e_i(t)| dt, \qquad i = x, y, \theta, u, v \tag{40}$$

- Integral of Time-weighted Absolute Error (ITAE), it measures the total accumulated error magnitude, providing a balanced assessment of tracking accuracy.

$$\mu_{IAE}^{p} = \int_{0}^{T} t|e_i(t)| dt, \qquad i = x, y, \theta, u, v \tag{41}$$

The tracking error performance indices of both the proposed controller and the conventional PD controller are shown in Fig. 12. It can be clearly observed that the proposed controller achieves superior tracking performance in ITSE and ITAE indicating a faster error convergence and enhanced transient response compared to the conventional PD controller. Once the system reaches steady state, the tracking error converges to zero and remains negligibly small, demonstrating the controller's effectiveness in ensuring stable and accurate tracking. For the IAE and ITAE indices, the tracking performance of both controllers appears nearly comparable since, during the first 2 seconds, the proposed controller requires a learning phase in which the trajectory exhibits relatively poor motion quality. This initial degradation is attributed to the adaptation and parameter learning process; however, once the learning stabilizes, the proposed controller achieves smoother motion and more accurate trajectory tracking. To further the trajectory tracking performance, a 2D trajectory of the NWMR is shown in Fig. 11. The control torque inputs are shown in Fig. 10. It should be noted that the chattering observed in the control signal during the first 2 seconds arises from the injected input disturbance that satisfies the PE condition. It is emphasized that the exploration noise is essential for the learning process.

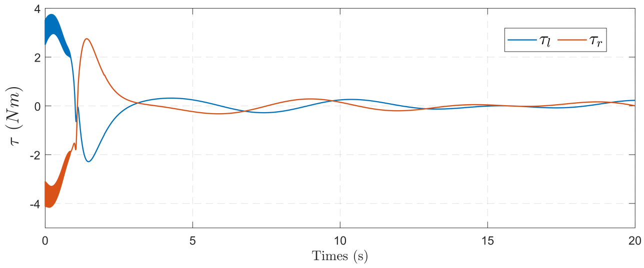

Fig. 10. The control input

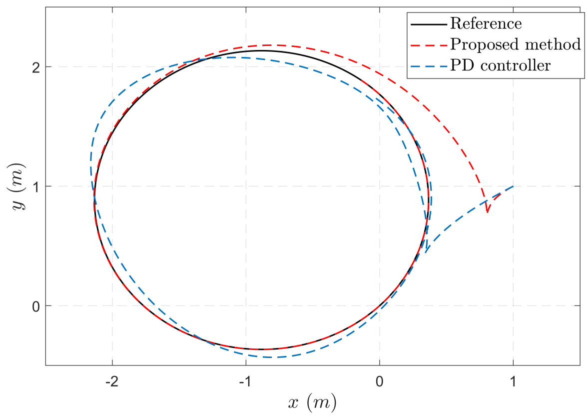

Fig. 11. The 2D trajectory of the NWMR

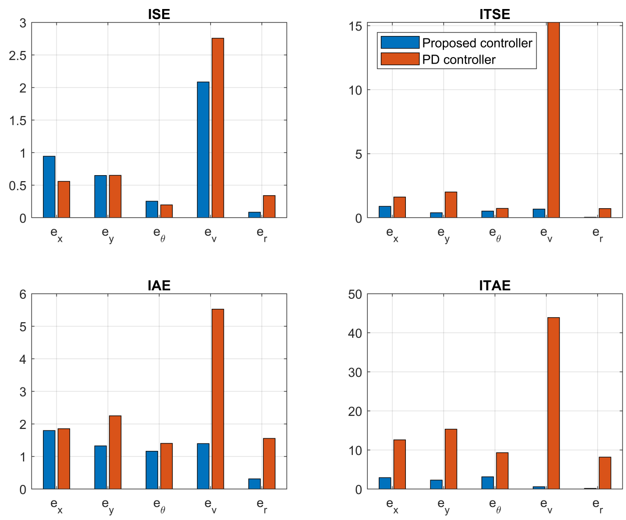

Fig. 12. The performance indices

5. Conclusion

A robust optimal trajectory tracking controller for a nonholonomic wheeled mobile robot has been presented. By employing a hierarchical control structure with a nonlinear disturbance observer and an actor–critic optimal controller, the proposed approach ensures accurate tracking and strong robustness against external disturbances. red Under the PE condition, the actor-critic algorithm converges and the optimal solution of the HJB equation can be learned. Simulation results validated the superior performance of the proposed method, which will be further investigated in real-time experiments. redIn particular, comparative studies with a conventional PD controller demonstrated that the proposed controller achieves significantly improved trajectory tracking performance under external disturbances, with the tracking errors converging rapidly and maintaining near-zero steady-state values. To discussion more, the proposed method requires full dynamic information of the system. Therefore, the future work is to develop an robust optimal controller, which learns the optimal control policy without the knowlegde of system dynamcis.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Aknowledgment: This research was supported by the Research Foundation, which was funded by the Thai Nguyen University of Technology.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] L. Zhao, G. Wang, X. Fan, and Y. Li, "The Analysis of trajectory control of non-holonomic mobile robots based on internet of things target image enhancement technology and backpropagation neural network," *Frontiers in Neurorobotics*, vol. 15, 2021, https://doi.org/10.3389/fnbot.2021.634340.

[2] L. Ding *et al*., "Definition and Application of Variable Resistance Coefficient for Wheeled Mobile Robots on Deformable Terrain," in *IEEE Transactions on Robotics*, vol. 36, no. 3, pp. 894-909, 2020, https://doi.org/10.1109/TRO.2020.2981822.

[3] A. Amirkhani, M. Shirzadeh, M. H. Shojaeefard, and A. Abraham, "Controlling wheeled mobile robot considering the effects of uncertainty with neuro-fuzzy cognitive map," *ISA transactions*, vol. 100 pp. 454-468, 2020, https://doi.org/10.1016/j.isatra.2019.12.011.

[4] L. Zhao, J. Li, H. Li, and B. Liu, "Double-loop tracking control for a wheeled mobile robot with unmodeled dynamics along right angle roads," *ISA transactions* vol. 136, pp. 525-534, 2023, https://doi.org/10.1016/j.isatra.2022.10.045.

[5] L. Li, W. Cao, H. Yang, and Q. Geng, "Trajectory tracking control for a wheel mobile robot on rough and uneven ground," *Mechatronics*, vol. 83, p. 102741, 2022, https://doi.org/10.1016/j.mechatronics.2022.102741.

[6] G. Cao, X. Zhao, C. Ye, S. Yu, B. Li, and C. Jiang, "Fuzzy adaptive PID control method for multi-mecanum-wheeled mobile robot," *Journal of Mechanical Science and Technology*, vol. 36, pp. 2019-2029, 2022, https://doi.org/10.1007/s12206-022-0337-x.

[7] S. T. Dang, X. M. Dinh, T. D. Kim, H. L. Xuan, and M. H. Ha, "Adaptive backstepping hierarchical sliding mode control for 3-wheeled mobile robots based on RBF neural networks," *Electronics*, vol. 12, no. 11, p. 2345, 2023, https://doi.org/10.3390/electronics12112345.

[8] M. Qin, S. Dian, B. Guo, X. Tao, and T. Zhao, "Fractional-order SMC controller for mobile robot trajectory tracking under actuator fault," *Systems Science & Control Engineering*, vol. 10, no. 1, pp. 312-324, 2022, https://doi.org/10.1080/21642583.2021.2023683.

[9] A. A. Tilahun, T. W. Desta, A. O. Salau, and L. Negash, "Design of an Adaptive Fuzzy Sliding Mode Control with Neuro-Fuzzy system for control of a differential drive wheeled mobile robot," *Cogent Engineering*, vol. 10, no. 2, p. 2276517, 2023, https://doi.org/10.1080/23311916.2023.2276517.

[10] N. Hassan and A. Saleem, "Neural Network-Based Adaptive Controller for Trajectory Tracking of Wheeled Mobile Robots," in *IEEE Access*, vol. 10, pp. 13582-13597, 2022, https://doi.org/10.1109/ACCESS.2022.3146970.

[11] D. Wang, N. Gao, D. Liu, J. Li and F. L. Lewis, "Recent Progress in Reinforcement Learning and Adaptive Dynamic Programming for Advanced Control Applications," in *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 1, pp. 18-36, 2024, https://doi.org/10.1109/JAS.2023.123843.

[12] B. Lian, W. Xue, F. L. Lewis, H. Modares, and B. Kiumarsi, *Integral and Inverse Reinforcement Learning for Optimal Control Systems and Games*, Springer Cham, 2024, https://doi.org/10.1007/978-3-031-45252-9.

[13] B. Kiumarsi, K. G. Vamvoudakis, H. Modares and F. L. Lewis, "Optimal and Autonomous Control Using Reinforcement Learning: A Survey," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2042-2062, 2018, https://doi.org/10.1109/TNNLS.2017.2773458.

[14] N. T. Dang, Q. N. Duong and G. K. Dinh, "Disturbance Observer-Based On-Policy Adaptive Dynamic Programming for a Quadrotor UAV with External Disturbance," *Journal of Electrical Engineering & Technology*, pp. 1-14, 2025, https://doi.org/10.1007/s42835-025-02404-1.

[15] Y. Ouyang, L. Dong and C. Sun, "Critic Learning-Based Control for Robotic Manipulators With Prescribed Constraints," in *IEEE Transactions on Cybernetics*, vol. 52, no. 4, pp. 2274-2283, 2022, https://doi.org/10.1109/TCYB.2020.3003550.

[16] F. L. Lewis and D. Liu, *Reinforcement learning and approximate dynamic programming for feedback control*, John Wiley & Sons, 2013, https://books.google.co.id/books?id=wH4Uy-ZFluwC&hl=id&source=gbs_navlinks_s.

[17] P. N. Dao, Q. P. Nguyen, and M. H. Vu, "Adaptive optimal coordination control of perturbed Bilateral Teleoperators with variable time delays using Actor–Critic Reinforcement Learning algorithm," *Mathematics and Computers in Simulation*, vol. 229, pp. 151-175, 2025, https://doi.org/10.1016/j.matcom.2024.09.007.

[18] L. Zhao, J. Jin, and J. Gong, "Robust zeroing neural network for fixed-time kinematic control of wheeled mobile robot in noise-polluted environment," *Mathematics and Computers in Simulation*, vol. 185, pp. 289-307, 2021, https://doi.org/10.1016/j.matcom.2020.12.030.

[19] T. L. Nguyen, X. S. Mai, and P. N. Dao, "A robust distributed model predictive control strategy for Leader–Follower formation control of multiple perturbed wheeled mobile robotics," *European Journal of Control*, vol. 81, p. 101160, 2025, https://doi.org/10.1016/j.ejcon.2024.101160.

[20] L. Li, T. Wang, Y. Xia, and N. Zhou, "Trajectory tracking control for wheeled mobile robots based on nonlinear disturbance observer with extended Kalman filter," *Journal of the Franklin Institute*, vol. 357, no. 13, pp. 8491-8507, 2020, https://doi.org/10.1016/j.jfranklin.2020.04.043.

[21] W. Liu, X. Wang and S. Liang, "Trajectory Tracking Control for Wheeled Mobile Robots Based on a Cascaded System Control Method," *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, pp. 396-401, 2020, https://doi.org/10.1109/IECON43393.2020.9255003.

[22] N. Hassan and A. Saleem, "Neural Network-Based Adaptive Controller for Trajectory Tracking of Wheeled Mobile Robots," in *IEEE Access*, vol. 10, pp. 13582-13597, 2022, https://doi.org/10.1109/ACCESS.2022.3146970.

[23] L. Yan, and B. Ma, "Adaptive practical leader-following formation control of multiple nonholonomic wheeled mobile robots," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 17, pp. 7216-7237, 2020, https://doi.org/10.1002/rnc.5165.

[24] D. Liu, X. Yang, and H. Li, "Adaptive optimal control for a class of continuous-time affine nonlinear systems with unknown internal dynamics," *Neural Computing and Applications*, vol. 23, no. 7, pp. 1843-1850, 2013, https://doi.org/10.1007/s00521-012-1249-y.

[25] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, "Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control," *Automatica*, vol. 43, no. 3, pp. 473-481, 2007, https://doi.org/10.1016/j.automatica.2006.09.019.

[26] 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, https://doi.org/10.1016/j.automatica.2012.06.096.

[27] K. G. Vamvoudakis and F. L. Lewis, "Multi-player non-zero-sum games: Online adaptive learning solution of coupled Hamilton–Jacobi equations," *Automatica*, vol. 47, no. 8, pp. 1556-1569, 2011, https://doi.org/10.1016/j.automatica.2011.03.005.

[28] W. -H. Chen, J. Yang, L. Guo and S. Li, "Disturbance-Observer-Based Control and Related Methods—An Overview," in *IEEE Transactions on Industrial Electronics*, vol. 63, no. 2, pp. 1083-1095, 2016, https://doi.org/10.1109/TIE.2015.2478397.

[29] S. Ding, Q. Hou and H. Wang, "Disturbance-Observer-Based Second-Order Sliding Mode Controller for Speed Control of PMSM Drives," in *IEEE Transactions on Energy Conversion*, vol. 38, no. 1, pp. 100-110, 2023, https://doi.org/10.1109/TEC.2022.3188630.

[30] D. Zhang, J. Hu, J. Cheng, Z. -G. Wu and H. Yan, "A Novel Disturbance Observer Based Fixed-Time Sliding Mode Control for Robotic Manipulators with Global Fast Convergence," in *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 3, pp. 661-672, 2024, https://doi.org/10.1109/JAS.2023.123948.

[31] Y. Chen, J. Liang, Y. Wu, Z. Miao, H. Zhang and Y. Wang, "Adaptive Sliding-Mode Disturbance Observer-Based Finite-Time Control for Unmanned Aerial Manipulator With Prescribed Performance," in *IEEE Transactions on Cybernetics*, vol. 53, no. 5, pp. 3263-3276, 2023, https://doi.org/10.1109/TCYB.2022.3168030.

[32] G. K. Dinh, H. V. Nguyen, V. T. Dang, D. T. Le, D. H. Nguyen and T. L. Nguyen, "Extended State Observer Based Robust Control for Active Suspensions with Electro-Hydraulic Actuators," *2023 12th International Conference on Control, Automation and Information Sciences (ICCAIS)*, pp. 471-476, 2023, https://doi.org/10.1109/ICCAIS59597.2023.10382244.

[33] N. T. Dang, Q. N. Duong, and G. K. Dinh, "Disturbance Observer-Based On-Policy Adaptive Dynamic Programming for a Quadrotor UAV with External Disturbance," *Journal of Electrical Engineering & Technology*, pp. 1-14, 2025, https://doi.org/10.1007/s42835-025-02404-1.

[34] G. K. Dinh, V. H. N. Thi, and P. N. Dao, "Model-free Optimal Formation Control Strategy for Multiple Quadrotors," *Journal of Control, Automation and Electrical Systems*, pp. 1-11, 2025, https://doi.org/10.1007/s40313-025-01193-2.

[35] W. Zhao, H. Liu and F. L. Lewis, "Robust Formation Control for Cooperative Underactuated Quadrotors via Reinforcement Learning," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 10, pp. 4577-4587, 2021, https://doi.org/10.1109/TNNLS.2020.3023711.

[36] G. K. Dinh, T. T. Bui, V. N. Tran, and P. N. Dao, "Data-Based Optimal Tracking Control for Underactuated Quadrotor System with Unknown Dynamic and External Disturbance," *Advances in Data Science and Optimization of Complex Systems*, pp. 443–451, 2025, https://doi.org/10.1007/978-3-031-90606-0_37.

[37] B. Lian, W. Xue, F. L. Lewis, H. Modares, and B. Kiumarsi, *Integral and Inverse Reinforcement Learning for Optimal Control Systems and Games*, Spring Cham, 2024, https://doi.org/10.1007/978-3-031-45252-9.

[38] B. Kiumarsi, K. G. Vamvoudakis, H. Modares and F. L. Lewis, "Optimal and Autonomous Control Using Reinforcement Learning: A Survey," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2042-2062, 2018, https://doi.org/10.1109/TNNLS.2017.2773458.

[39] F. L. Lewis and D. Liu, *Reinforcement learning and approximate dynamic programming for feedback control*, Jhon Wiley & Sons, 2013, https://books.google.co.id/books?id=wH4Uy-ZFluwC&hl=id&source=gbs_navlinks_s.

[40] Y. Zheng, J. Zheng, K. Shao, H. Zhao, H. Xie and H. Wang, "Adaptive Trajectory Tracking Control for Nonholonomic Wheeled Mobile Robots: A Barrier Function Sliding Mode Approach," in *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 4, pp. 1007-1021, 2024, https://doi.org/10.1109/JAS.2023.124002.

[41] J. J. Zhang, Z. L. Fang, Z. Q. Zhang, R. Z. Gao, and S. B. Zhang, "Trajectory tracking control of nonholonomic wheeled mobile robots using model predictive control subjected to lyapunov-based input constraints," *International Journal of Control, Automation and Systems*, vol. 20, no. 5, pp. 1640-1651, 2022, https://doi.org/10.1007/s12555-019-0814-x.

[42] L. Yan, B. Ma, Y. Jia and Y. Jia, "Observer-Based Trajectory Tracking Control of Nonholonomic Wheeled Mobile Robots," in *IEEE Transactions on Control Systems Technology*, vol. 32, no. 3, pp. 1114-1121, 2024, https://doi.org/10.1109/TCST.2024.3351073.

[43] Y. Wu, Y. Wang, and H. Fang, "Full-state constrained neural control and learning for the nonholonomic wheeled mobile robot with unknown dynamics," *ISA transactions*, vol. 125, pp. 22-30, 2022, https://doi.org/10.1016/j.isatra.2021.06.012.

[44] B. M. Yousuf, A. S. Khan, and S. M. Khan, "Dynamic modeling and tracking for nonholonomic mobile robot using PID and back-stepping," *Advanced Control for Applications: Engineering and Industrial Systems*, vol. 3, no. 3, 2021, https://doi.org/10.1002/adc2.71.

[45] C. B. Ben and H. Seddik, "Design of a PID optimized neural networks and PD fuzzy logic controllers for a two-wheeled mobile robot," *Asian Journal of Control*, vol. 23, no. 1, pp. 23-41, 2021, https://doi.org/10.1002/asjc.2356.

[46] J. -H. Park, "Optimal Controller Design for a Mobile Robot Using Genetic Algorithm and Adaptive PID Controller," in *IEEE Access*, vol. 13, pp. 86167-86184, 2025, https://doi.org/10.1109/ACCESS.2025.3570472.

[47] H. Xie, J. Zheng, Z. Sun, H. Wang and R. Chai, "Finite-time tracking control for nonholonomic wheeled mobile robot using adaptive fast nonsingular terminal sliding mode," *Nonlinear Dynamics* vol. 110, no. 2, pp. 1437-1453, 2022, https://doi.org/10.1007/s11071-022-07682-2.

[48] Y. Zheng, J. Zheng, K. Shao, H. Zhao, H. Xie and H. Wang, "Adaptive Trajectory Tracking Control for Nonholonomic Wheeled Mobile Robots: A Barrier Function Sliding Mode Approach," in *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 4, pp. 1007-1021, 2024, hhttps://doi.org/10.1109/JAS.2023.124002.

[49] H. Xie, J. Zheng, R. Chai, and H. T. Nguyen, "Robust tracking control of a differential drive wheeled mobile robot using fast nonsingular terminal sliding mode," *Computers & Electrical Engineering*, vol. 96, p. 107488, 2021, https://doi.org/10.1016/j.compeleceng.2021.107488.

[50] A. Y. Damani, Z. A. Benselama, and R. Hedjar, "Formation control of nonholonomic wheeled mobile robots via fuzzy fractional-order integral sliding mode control," *International Journal of Dynamics and Control*, vol. 11, no. 5, pp. 2273-2284, 2023, https://doi.org/10.1007/s40435-022-01109-x.

[51] P. T. Phuc *et al*., "Design of adaptive fuzzy sliding mode controller for mobile robot," *International Journal of Mechanical Engineering and Robotics Research*, vol. 10, no. 2, pp. 54-59, 2021, https://www.ijmerr.com/uploadfile/2021/0125/20210125022222502.pdf.

[52] Y. Zheng *et al*., "Adaptive fuzzy sliding mode control of uncertain nonholonomic wheeled mobile robot with external disturbance and actuator saturation," *Information Sciences*, vol. 663, p. 120303, 2024, https://doi.org/10.1016/j.ins.2024.120303

[53] Y. A. Wendemagegn *et al*., "Enhancing trajectory tracking accuracy in three-wheeled mobile robots using backstepping fuzzy sliding mode control," *Engineering Research Express*, vol. 6, no. 4, p. 045204, 2024, https://doi.org/10.1088/2631-8695/ad79b9.

[54] N. Hassan and A. Saleem, "Neural Network-Based Adaptive Controller for Trajectory Tracking of Wheeled Mobile Robots," in *IEEE Access*, vol. 10, pp. 13582-13597, 2022, https://doi.org/10.1109/ACCESS.2022.3146970.

[55] T. Beharu Arega, Y. Mersha Tesfa and C. M. Abdissa, "Three-Wheeled Mobile Robot Trajectory Tracking Control Using Nonlinear PID Controller Based Neural Network Combined With Backstepping Controller," in *IEEE Access*, vol. 13, pp. 100167-100182, 2025, https://doi.org/10.1109/ACCESS.2025.3577269.

[56] Y. Wu, Y. Wang, and H. Fang, "Full-state constrained neural control and learning for the nonholonomic wheeled mobile robot with unknown dynamics," *ISA transactions*, vol. 125, pp. 22-30, 2022, https://doi.org/10.1016/j.isatra.2021.06.012.

[57] C. G. Yun, Y. C. Sin, H. R. Ri, and K. N. Jo, "Trajectory tracking control of a three-wheeled omnidirectional mobile robot using disturbance estimation compensator by RBF neural network," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 45, no. 432, 2023, https://doi.org/10.1007/s40430-023-04340-5.

[58] K. Nguyen, V. T. Dang, D. D. Pham, and P. N. Dao, "Formation control scheme with reinforcement learning strategy for a group of multiple surface vehicles," *International Journal of Robust and Nonlinear Control*, vol. 34, no. 3, pp. 2252-2279, 2024, https://doi.org/10.1002/rnc.7083.

[59] G. Wang, B. Luo, and S. Xue, "Integral reinforcement learning-based optimal output feedback control for linear continuous-time systems with input delay," *Neurocomputing*, vol. 460, pp. 31-38, 2021, https://doi.org/10.1016/j.neucom.2021.06.073.

[60] D. Wang, N. Gao, D. Liu, J. Li and F. L. Lewis, "Recent Progress in Reinforcement Learning and Adaptive Dynamic Programming for Advanced Control Applications," in *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 1, pp. 18-36, 2024, https://doi.org/10.1109/JAS.2023.123843.