Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Robust Optimal Tracking Control for Center-of-Mass-Offset Mecanum Mobile Robots Using Off-Policy Integral Reinforcement Learning

Minh Dong Nguyen
Electric Power University
Duy Phong Pham
phongphd@epu.edu.vn

Electric Power University

Research Article

Keywords: Four-wheeled Mecanum mobile robot, Time-Varying RISE, optimal tracking control, Actor-critic, reinforcement learning

Posted Date: October 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10922716/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Robust Optimal Tracking Control for Center-of-Mass-Offset Mecanum Mobile Robots Using Off-Policy Integral Reinforcement Learning

Minh Dong Nguyen[1] and Phong Pham Duy[1*]

[1]Faculty of Electronics and Telecommunications, Electric Power University, 235 Hoang Quoc Viet, Nghia Do, 11355, Hanoi, Vietnam.

*Corresponding author(s). E-mail(s): phongphd@epu.edu.vn;
Contributing authors: dongnm@epu.edu.vn;

Abstract

This paper proposes an off-policy integral reinforcement learning-based robust optimal tracking control method for four-wheeled Mecanum mobile robots subject to COM-induced and slip-related uncertainties. Based on the kinematic and dynamic model of the Mecanum platform, the trajectory tracking problem is transformed into an autonomous affine nonlinear system. A robust optimal control problem is then formulated using the Hamilton–Jacobi–Isaacs framework to simultaneously improve tracking accuracy, reduce control effort, and attenuate the effect of worst-case disturbances. Since the resulting HJI equation is difficult to solve analytically, an off-policy integral reinforcement learning algorithm is developed to approximate its solution from collected input–state data. Three neural networks are employed to estimate the optimal value function, the robust optimal control policy, and the worst-case disturbance policy. Lyapunov-based stability analysis proves that all closed-loop signals are uniformly ultimately bounded. Simulation results under time-varying load-distribution and wheel–ground interaction uncertainties demonstrate that the proposed controller achieves better tracking performance, smoother control torques, and stronger robustness than conventional actor–critic reinforcement learning methods.

Keywords: Four-wheeled Mecanum mobile robot, Time-Varying RISE, optimal tracking control, Actor-critic, reinforcement learning

1 Introduction

Four-wheeled Mecanum mobile robots have been widely investigated and deployed in robotic applications because of their omnidirectional mobility, compact mechanical structure, and high maneuverability in confined environments. Since the original Mecanum wheel concept was introduced in [1], omnidirectional wheeled mobile robots have become an important class of mobile robotic systems, with their structural and motion properties systematically studied in [2]. Owing to the special roller arrangement of Mecanum wheels, the robot can move longitudinally, laterally, and rotationally without changing its body orientation. This capability makes Mecanum mobile robots particularly suitable for logistics automation, intelligent warehousing, service robotics, and industrial transportation.

In recent years, Mecanum mobile robots have also been increasingly used as mobile bases for robotic manipulators, forming mobile manipulation systems capable of performing both locomotion and manipulation tasks [3, 4]. However, the integration of a payload or robotic arm introduces additional challenges to the control design. In practical operation, the center of mass of the overall system may deviate from the geometric center of the mobile base due to payload variation or manipulator motion [5]. Such center-of-mass deviation changes the inertia distribution, causes nonuniform wheel loading, and strengthens the coupling between translational and rotational dynamics. Furthermore, the wheel–ground interaction of Mecanum wheels is inherently more complicated than that of conventional wheels because the passive rollers introduce additional slip components. These effects may result in COM-induced and slip-related uncertainties, which can significantly degrade trajectory tracking performance and may even cause oscillatory motion or instability [6–8].

A considerable number of studies have addressed the modeling and trajectory tracking control of Mecanum-wheeled mobile robots. Kinematic and dynamic models of Mecanum platforms have been developed in [9–12], while adaptive and model-based control approaches have been investigated in [13–17]. For trajectory tracking, several nonlinear control techniques have been proposed, including kinematic control, trajectory linearization, robust tracking control, adaptive integral terminal sliding mode control, and sliding mode control under uncertainties [18–23]. These methods can achieve acceptable tracking performance under certain assumptions. Nevertheless, kinematic controllers usually neglect dynamic effects caused by load redistribution and wheel slippage, whereas robust and adaptive controllers often require conservative gain tuning or prior knowledge of uncertainty bounds. Moreover, most of these approaches mainly focus on stability and tracking accuracy, while the optimal trade-off between tracking performance and control effort is not explicitly considered.

Reinforcement learning and adaptive dynamic programming have recently emerged as effective tools for optimal control of nonlinear systems. In general, optimal control of nonlinear systems is associated with the Hamilton–Jacobi–Bellman equation, while robust optimal control with disturbance attenuation is related to the Hamilton–Jacobi–Isaacs equation [24–27]. Since these equations are nonlinear partial differential equations and are difficult to solve analytically, adaptive dynamic programming and reinforcement learning have been developed to approximate their solutions using neural networks [24–27]. For trajectory tracking problems, ADP-based approaches have

been extended to continuous-time systems, where the value function and control policy depend on tracking error dynamics [28, 29]. However, the time-varying nature of tracking problems makes the stability analysis and optimal policy learning more challenging.

Actor–Critic reinforcement learning structures have been widely used to approximate the optimal value function and control policy without directly solving the Hamilton–Jacobi–Bellman equation [30–36]. These methods provide a promising framework for learning-based optimal tracking control of mobile robots. In particular, reinforcement learning-based tracking control has been applied to Mecanum mobile robots in [17, 37]. However, most existing Actor–Critic-based methods are formulated mainly under the optimal control framework and do not explicitly consider the worst-case disturbance attenuation problem. When the robot is subject to time-varying load distribution, wheel–ground interaction uncertainty, and external disturbances, a robust optimal control formulation is more appropriate because it can simultaneously optimize tracking performance and attenuate disturbance effects.

From the robust optimal control perspective, the Hamilton–Jacobi–Isaacs framework formulates the control input and the disturbance input as two opposite players in a zero-sum differential game. In this formulation, the controller seeks to minimize the performance index, whereas the disturbance seeks to maximize it. The corresponding optimal solution provides both the robust optimal control policy and the worst-case disturbance policy. Nevertheless, solving the HJI equation requires system dynamic information and is generally intractable for uncertain nonlinear systems. To overcome this difficulty, off-policy integral reinforcement learning can be used to learn the solution of the HJI equation from collected input–state data. The off-policy IRL formulation has been shown to provide an equivalent Bellman equation and to converge to the robust optimal control and disturbance policies under suitable conditions [38, 39].

Motivated by the above observations, this paper proposes an off-policy integral reinforcement learning-based robust optimal tracking controller for four-wheeled Mecanum mobile robots subject to COM-induced and slip-related uncertainties. Different from conventional Actor–Critic optimal tracking controllers, the proposed method approximates not only the optimal value function and control policy, but also the worst-case disturbance policy. Three neural networks are employed, including a critic network for the value function, a control actor network for the robust optimal control policy, and a disturbance actor network for the worst-case disturbance policy.

The main contributions of this paper are summarized as follows:

1. A robust optimal tracking control framework is developed for four-wheeled Mecanum mobile robots subject to COM-induced and slip-related uncertainties. The trajectory tracking problem is transformed into an autonomous affine nonlinear system and formulated within the Hamilton–Jacobi–Isaacs framework.
2. An off-policy integral reinforcement learning algorithm is proposed to approximate the solution of the Hamilton–Jacobi–Isaacs equation using three neural networks, corresponding to the optimal value function, the robust optimal control policy, and the worst-case disturbance policy. The UUB stability of the closed-loop system is proven using Lyapunov theory.

2 Kinematic and Dynamic Modeling

Consider a four-wheeled Mecanum mobile robot moving on a horizontal plane, as shown in Fig. 1. An inertial frame $\{O_I, X_I, Y_I\}$ is fixed to the ground, while a body-fixed frame $\{O_R, X_R, Y_R\}$ is attached to the geometric center of the robot platform.

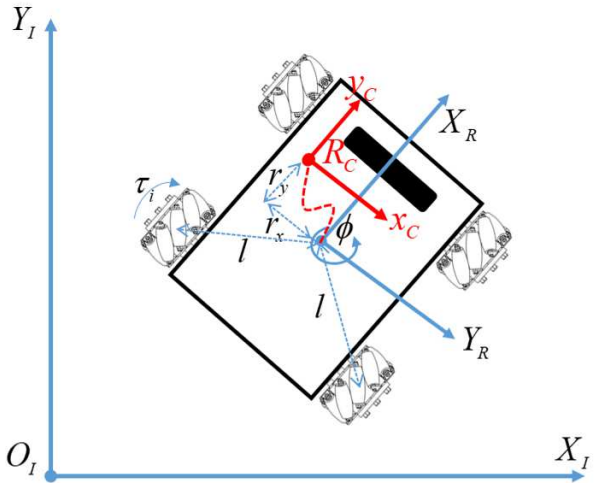


Fig. 1 Dynamic model of the four-wheeled Mecanum mobile robot with time-varying COM offset.

The generalized coordinate vector is defined as

$$q = \begin{bmatrix} x & y & \phi \end{bmatrix}^T, \tag{1}$$

where x and y denote the position of the robot in the inertial frame, and ϕ is the heading angle.

Due to payload variation or manipulator motion, the actual center of mass may deviate from the geometric center of the mobile base, as illustrated in Fig. 1. The center-of-mass offset vector is defined in the body-fixed frame as

$$R_c = \begin{bmatrix} x_c & y_c \end{bmatrix}^T, \tag{2}$$

where x_c and y_c are the longitudinal and lateral COM offsets, respectively. The presence of R_c changes the inertia distribution, produces nonuniform wheel loading, and introduces additional coupling between translational and rotational motions.

The relationship between the inertial-frame velocity $\dot{q}$ and the body-fixed velocity v_q is expressed as

$$v_q = H(\phi)\dot{q}, \tag{3}$$

where

$$v_q = \begin{bmatrix} v_x & v_y & \omega \end{bmatrix}^T, \quad H(\phi) = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

Equivalently, the robot kinematics can be written as

$$\dot{q} = H^{-1}(\phi) v_q. \tag{5}$$

The dynamic model of the Mecanum mobile robot with COM-induced and slip-related uncertainties is described by

$$M(R_c)\ddot{q} + C(q, \dot{q}, R_c)\dot{q} + D(R_c, \rho) = B\tau, \tag{6}$$

where $M(R_c) \in \mathbb{R}^{3\times3}$ is the inertia matrix affected by the COM offset, $C(q, \dot{q}, R_c) \in \mathbb{R}^{3\times3}$ denotes the Coriolis and centrifugal matrix, $D(R_c, \rho)$ represents lumped uncertainties caused by COM-induced coupling, wheel–ground interaction, and wheel slippage, $B \in \mathbb{R}^{3\times4}$ is the input transformation matrix, and

$$\tau = \begin{bmatrix} \tau_1 & \tau_2 & \tau_3 & \tau_4 \end{bmatrix}^T \tag{7}$$

is the vector of wheel torques.

The inertia matrix is given by

$$M(R_c) = \begin{bmatrix} m_{11} & 0 & m_{13} \\ 0 & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}, \tag{8}$$

where

$$\begin{aligned}
m_{11} &= m_b + 4\left(m_w + \frac{I_w}{r^2}\right), \\
m_{22} &= m_b + 4\left(m_w + \frac{I_w}{r^2}\right), \\
m_{13} &= m_{31} = m_b\left(x_c \sin\phi + y_c \cos\phi\right), \\
m_{23} &= m_{32} = m_b\left(-x_c \cos\phi + y_c \sin\phi\right), \\
m_{33} &= m_b\left(x_c^2 + y_c^2\right) + I_b + 8\left(m_w + \frac{I_w}{r^2}\right) l^2 \sin^2\left(\frac{\pi}{4} - \alpha\right).
\end{aligned} \tag{9}$$

Here, m_b and I_b are the mass and moment of inertia of the robot body, m_w and I_w are the mass and moment of inertia of each wheel, r is the wheel radius, l is the distance from the robot center to each wheel, and α is the roller angle of the Mecanum wheel.

The Coriolis and centrifugal matrix induced by the eccentric mass distribution is expressed as

$$C(q, \dot{q}, R_c) = \begin{bmatrix} 0 & 0 & m_b\dot{\phi}(x_c \cos\phi - y_c \sin\phi) \\ 0 & 0 & m_b\dot{\phi}(x_c \sin\phi + y_c \cos\phi) \\ 0 & 0 & 0 \end{bmatrix}. \tag{10}$$

The input transformation matrix is defined as

$$B = \frac{1}{r}\begin{bmatrix} -(C_\phi - S_\phi) & -(C_\phi + S_\phi) & C_\phi - S_\phi & C_\phi + S_\phi \\ -(S_\phi + C_\phi) & -(S_\phi - C_\phi) & S_\phi + C_\phi & S_\phi - C_\phi \\ -\sqrt{2}lS_\alpha & -\sqrt{2}lS_\alpha & -\sqrt{2}lS_\alpha & -\sqrt{2}lS_\alpha \end{bmatrix}, \tag{11}$$

where

$$C_\phi = \cos\phi, \quad S_\phi = \sin\phi, \quad S_\alpha = \sin\left(\frac{\pi}{4} - \alpha\right). \tag{12}$$

The lumped uncertainty term $D(R_c, \rho)$ in (6) includes the effects of COM-induced dynamic coupling, nonuniform wheel loading, wheel slippage, friction variation, and unmodeled disturbances. In this paper, these effects are considered bounded but unknown.

Taking the time derivative of (3) gives

$$\dot{v}_q = H(\phi)\ddot{q} + \dot{H}(\phi)\dot{q}. \tag{13}$$

From (6), the generalized acceleration is obtained as

$$\ddot{q} = M^{-1}(R_c)B\tau - M^{-1}(R_c)C(q, \dot{q}, R_c)\dot{q} - M^{-1}(R_c)D(R_c, \rho). \tag{14}$$

Substituting (14) into (13), and using $\dot{q} = H^{-1}(\phi)v_q$, the body-frame velocity dynamics can be rewritten as

$$\dot{v}_q = \bar{f}(q, v_q)v_q + \bar{g}(q)\tau + \bar{g}(q)d(q, v_q), \tag{15}$$

where

$$\bar{f}(q, v_q) = \dot{H}(\phi)H^{-1}(\phi) - H(\phi)M^{-1}(R_c)C(q, \dot{q}, R_c)H^{-1}(\phi), \tag{16}$$

$$\bar{g}(q) = H(\phi)M^{-1}(R_c)B, \tag{17}$$

$$\bar{g}(q)d(q, v_q) = -H(\phi)M^{-1}(R_c)D(R_c, \rho). \tag{18}$$

Define the system state vector as

$$x = \begin{bmatrix} q^T & v_q^T \end{bmatrix}^T \in \mathbb{R}^6. \tag{19}$$

Then, the control-oriented affine nonlinear model of the Mecanum mobile robot is obtained as

$$\dot{x} = f(x) + g(x)\tau + g(x)d(x), \tag{20}$$

where

$$f(x) = \begin{bmatrix} H^{-1}(\phi)v_q \\ \bar{f}(q, v_q)v_q \end{bmatrix}, \quad g(x) = \begin{bmatrix} 0_{3\times 4} \\ \bar{g}(q) \end{bmatrix}. \tag{21}$$

The term $d(x)$ denotes the matched lumped uncertainty associated with COM-induced and slip-related effects. The affine form (20) will be used in the following sections to formulate the robust optimal tracking control problem and develop the off-policy integral reinforcement learning-based controller.

3 Controller Design

This section develops the proposed robust optimal tracking controller. The tracking error dynamics is first transformed into an autonomous affine nonlinear system. Then, the HJI-based robust optimal control problem is formulated and solved approximately using an off-policy integral reinforcement learning algorithm.

3.1 Robust Optimal Tracking Problem Formulation

Based on the affine nonlinear model obtained in (20), the control-oriented dynamics of the four-wheeled Mecanum mobile robot can be rewritten as

$$\dot{x} = f(x) + g(x)\tau + g(x)d(x), \tag{22}$$

where $x \in \mathbb{R}^6$ is the system state, $\tau \in \mathbb{R}^4$ is the wheel torque input, and $d(x) \in \mathbb{R}^4$ denotes the matched lumped uncertainty caused by COM-induced and slip-related effects.

Let the desired reference state be described by

$$x_d = \begin{bmatrix} q_d^T & v_{qd}^T \end{bmatrix}^T \in \mathbb{R}^6, \tag{23}$$

where

$$q_d = \begin{bmatrix} x_d & y_d & \phi_d \end{bmatrix}^T \tag{24}$$

is the desired position and heading vector, and v_{qd} is the desired body-fixed velocity. The reference system is assumed to satisfy

$$\dot{x}_d = h_d(x_d), \tag{25}$$

where $h_d(x_d)$ is a bounded smooth function.

The tracking error is defined as

$$e = x - x_d. \tag{26}$$

Taking the time derivative of (26) and using (22) and (25), one obtains

$$\dot{e} = f(x) + g(x)\tau + g(x)d(x) - h_d(x_d). \tag{27}$$

To separate the feedforward reference compensation from the feedback control input, the wheel torque input is decomposed as

$$\tau = \tau_d + u, \tag{28}$$

where $u \in \mathbb{R}^4$ is the feedback control input to be designed, and τ_d is the feedforward reference input defined by

$$\tau_d = g^{+}(x_d)\left[h_d(x_d) - f(x_d)\right]. \tag{29}$$

Here, $g^{+}(x_d)$ denotes the generalized inverse of $g(x_d)$.

Substituting (28) into (27) gives

$$\dot{e} = f(x) - h_d(x_d) + g(x)\tau_d + g(x)u + g(x)d(x). \tag{30}$$

To transform the time-varying tracking problem into an autonomous regulation problem, the augmented state vector is introduced as

$$z = \begin{bmatrix} e^T & x_d^T \end{bmatrix}^T \in \mathbb{R}^{12}. \tag{31}$$

Then, the augmented error dynamics can be expressed in the affine form

$$\dot{z} = F(z) + G(z)u + K(z)d, \tag{32}$$

where

$$\begin{aligned} F(z) &= \begin{bmatrix} f(x) - h_d(x_d) + g(x)\tau_d \\ h_d(x_d) \end{bmatrix}, \\ G(z) &= \begin{bmatrix} g(x) \\ 0_{6\times 4} \end{bmatrix}, \\ K(z) &= \begin{bmatrix} g(x) \\ 0_{6\times 4} \end{bmatrix}. \end{aligned} \tag{33}$$

In (32), $d \in \mathbb{R}^4$ represents the matched disturbance input associated with the lumped uncertainty.

The objective is to design a robust optimal tracking controller $u^*(z)$ such that the tracking error is driven to a small neighborhood of the origin while the control effort is minimized and the influence of the worst-case disturbance is attenuated. To this end, the infinite-horizon cost function is defined as

$$J(z, u, d) = \int_0^\infty \left[z^T Q z + u^T R u - \gamma^2 d^T d \right] dt, \tag{34}$$

where $Q = Q^T > 0$ and $R = R^T > 0$ are weighting matrices, and $\gamma > 0$ is the disturbance attenuation level.

The robust optimal tracking problem is formulated as a zero-sum differential game between the control input u and the disturbance input d. The optimal value function is defined as

$$V^*(z) = \min_u \max_d J(z, u, d). \tag{35}$$

In this formulation, the controller seeks to minimize the performance index, whereas the disturbance seeks to maximize it. The solution of (35) provides the robust optimal control policy and the worst-case disturbance policy.

For the subsequent controller design, the following assumptions are introduced.

Assumption 1: The functions $F(z)$, $G(z)$, and $K(z)$ are locally Lipschitz continuous on a compact set $\Omega_z \subset \mathbb{R}^{12}$ containing the origin.

Assumption 2: The lumped disturbance d is bounded, i.e.,

$$\|d\| \leq \bar{d}, \tag{36}$$

where $\bar{d} > 0$ is an unknown constant.

Assumption 3: There exists an admissible control policy $u(z)$ that stabilizes the augmented system (32) and makes the cost function (34) finite.

Under these assumptions, the robust optimal tracking control problem can be solved through the Hamilton–Jacobi–Isaacs equation, which will be derived in the next subsection.

3.2 HJI-Based Robust Optimal Control

For the augmented system (32), the robust optimal tracking problem can be solved by formulating the Hamilton–Jacobi–Isaacs equation. Define the Hamiltonian function associated with the cost function (34) as

$$\begin{aligned} H(z, u, d, \nabla V) = z^T Q z + u^T R u - \gamma^2 d^T d \\ + \nabla V^T \left[F(z) + G(z)u + K(z)d \right], \end{aligned} \tag{37}$$

where $V(z)$ is the value function and

$$\nabla V = \frac{\partial V}{\partial z}. \tag{38}$$

The robust optimal tracking control problem is to find a control policy $u^*(z)$ that minimizes the Hamiltonian, while the disturbance policy $d^*(z)$ maximizes it. Therefore, the saddle-point condition is given by

$$H(z, u^*, d, \nabla V^*) \leq H(z, u^*, d^*, \nabla V^*) \leq H(z, u, d^*, \nabla V^*). \tag{39}$$

The stationary conditions of the Hamiltonian are

$$\frac{\partial H}{\partial u} = 0, \qquad \frac{\partial H}{\partial d} = 0. \tag{40}$$

From (37) and (40), the robust optimal control policy is obtained as

$$u^*(z) = -\frac{1}{2} R^{-1} G^T(z) \nabla V^*(z), \tag{41}$$

and the worst-case disturbance policy is given by

$$d^*(z) = \frac{1}{2\gamma^2} K^T(z) \nabla V^*(z). \tag{42}$$

Substituting (41) and (42) into (37), the Hamilton–Jacobi–Isaacs equation can be written in the compact form as

$$H^*(z, \nabla V^*) = 0. \tag{43}$$

More explicitly, the HJI equation is given by

$$\begin{aligned} H^*(z, \nabla V^*) = {} & z^T Q z + \nabla V^{*T}(z) F(z) \\ & - \frac{1}{4} \nabla V^{*T}(z) G(z) R^{-1} G^T(z) \nabla V^*(z) \\ & + \frac{1}{4\gamma^2} \nabla V^{*T}(z) K(z) K^T(z) \nabla V^*(z) = 0. \end{aligned} \tag{44}$$

Equation (44) is a nonlinear partial differential equation. For the uncertain nonlinear dynamics of the Mecanum mobile robot, it is generally difficult to obtain an analytical solution. Therefore, the robust optimal value function $V^*(z)$, the control policy $u^*(z)$, and the worst-case disturbance policy $d^*(z)$ will be approximated using an off-policy integral reinforcement learning algorithm in the next subsection.

Once the approximate robust optimal control policy $\hat{u}(z)$ is obtained, the actual wheel torque input applied to the Mecanum mobile robot is given by

$$\tau = \tau_d + \hat{u}(z), \tag{45}$$

where τ_d is the feedforward reference input defined in (29).

3.3 Off-Policy Integral Reinforcement Learning-Based Controller

Since the HJI equation (44) cannot be solved analytically for the uncertain nonlinear Mecanum mobile robot, an off-policy integral reinforcement learning method is developed to approximate its solution. The main idea is to learn the robust optimal value function, the robust optimal control policy, and the worst-case disturbance policy from collected input–state data.

For any admissible control policy $u_i(z)$ and disturbance policy $d_i(z)$, the corresponding value function $V_i(z)$ satisfies the Bellman equation

$$\begin{aligned} & z^T Q z + u_i^T R u_i - \gamma^2 d_i^T d_i \\ & \quad + \nabla V_i^T(z) \left[F(z) + G(z) u_i(z) + K(z) d_i(z) \right] = 0. \end{aligned} \tag{46}$$

To construct an off-policy learning formulation, the augmented system (32) is rewritten as

$$\begin{aligned} \dot{z} = {} & F(z) + G(z) u_i + K(z) d_i \\ & + G(z) \left(u - u_i \right) + K(z) \left(d - d_i \right), \end{aligned} \tag{47}$$

where u is the behavior control input applied to the system for data collection, while u_i and d_i are the target policies to be evaluated and improved.

Taking the derivative of $V_i(z)$ along (47) gives

$$\begin{aligned}\dot{V}_i(z) = &\nabla V_i^T(z)\left[F(z) + G(z)u_i + K(z)d_i\right] \\ &+ \nabla V_i^T(z)G(z)(u - u_i) \\ &+ \nabla V_i^T(z)K(z)(d - d_i).\end{aligned} \tag{48}$$

Using (46), (48) can be rewritten as

$$\begin{aligned}\dot{V}_i(z) = &-z^T Q z - u_i^T R u_i + \gamma^2 d_i^T d_i \\ &+ \nabla V_i^T(z)G(z)(u - u_i) \\ &+ \nabla V_i^T(z)K(z)(d - d_i).\end{aligned} \tag{49}$$

Integrating (49) over the time interval $[t, t+T]$ yields the off-policy integral Bellman equation

$$\begin{aligned}&V_i(z(t+T)) - V_i(z(t)) \\ &= -\int_t^{t+T}\left[z^T Q z + u_i^T R u_i - \gamma^2 d_i^T d_i\right] d\zeta \\ &\quad + \int_t^{t+T} \nabla V_i^T(z)G(z)(u - u_i)d\zeta \\ &\quad + \int_t^{t+T} \nabla V_i^T(z)K(z)(d - d_i)d\zeta.\end{aligned} \tag{50}$$

Equation (50) allows the value function and the policies to be learned from collected data without explicitly requiring the complete drift dynamics $F(z)$.

According to the universal approximation property of neural networks, the optimal value function can be approximated over a compact set Ω_z as

$$V^*(z) = W_c^T \sigma(z) + \varepsilon_c(z), \tag{51}$$

where $W_c \in \mathbb{R}^N$ is the ideal critic weight vector, $\sigma(z) \in \mathbb{R}^N$ is the activation function vector, and $\varepsilon_c(z)$ is the approximation error.

The gradient of the optimal value function is then given by

$$\nabla V^*(z) = \nabla \sigma^T(z) W_c + \nabla \varepsilon_c(z). \tag{52}$$

Based on (41) and (42), the robust optimal control policy and the worst-case disturbance policy can be approximated as

$$\hat{u}(z) = -\frac{1}{2} R^{-1} G^T(z) \nabla \sigma^T(z) \hat{W}_a, \tag{53}$$

and

$$\hat{d}(z) = \frac{1}{2\gamma^2} K^T(z) \nabla \sigma^T(z) \hat{W}_d, \tag{54}$$

where $\hat{W}_a$ and $\hat{W}_d$ are the estimated weight vectors of the control actor network and disturbance actor network, respectively.

The critic network is defined as

$$\hat{V}(z) = \hat{W}_c^T \sigma(z), \tag{55}$$

where $\hat{W}_c$ is the estimated critic weight vector.

Substituting (55), (53), and (54) into the Hamiltonian gives the Bellman error

$$\begin{aligned} \delta = {} & z^T Q z + \hat{u}^T R \hat{u} - \gamma^2 \hat{d}^T \hat{d} \\ & + \nabla \hat{V}^T(z) \left[F(z) + G(z)\hat{u} + K(z)\hat{d} \right]. \end{aligned} \tag{56}$$

For implementation, the critic regression vector is defined as

$$\omega = \nabla \sigma(z) \left[F(z) + G(z)\hat{u} + K(z)\hat{d} \right]. \tag{57}$$

Then, the Bellman error can be written as

$$\delta = z^T Q z + \hat{u}^T R \hat{u} - \gamma^2 \hat{d}^T \hat{d} + \hat{W}_c^T \omega. \tag{58}$$

To minimize the Bellman error, the critic weight update law is chosen as

$$\dot{\hat{W}}_c = -\eta_c \frac{\omega}{1 + \omega^T \omega} \delta, \tag{59}$$

where $\eta_c > 0$ is the critic learning rate.

The control actor weight update law is designed as

$$\dot{\hat{W}}_a = -\eta_a \left(\hat{W}_a - \hat{W}_c \right), \tag{60}$$

where $\eta_a > 0$ is the learning rate of the control actor network.

Similarly, the disturbance actor weight update law is given by

$$\dot{\hat{W}}_d = -\eta_d \left(\hat{W}_d - \hat{W}_c \right), \tag{61}$$

where $\eta_d > 0$ is the learning rate of the disturbance actor network.

After the learning process, the approximate robust optimal control policy is obtained from (53). Therefore, the actual wheel torque input applied to the Mecanum mobile robot is

$$\tau = \tau_d + \hat{u}(z), \tag{62}$$

where τ_d is the feedforward reference input defined in (29).

The proposed off-policy integral reinforcement learning-based controller can be summarized as follows:

1. Collect input–state data from the Mecanum mobile robot under a behavior control input.
2. Construct the augmented state z and compute the Bellman error.
3. Update the critic network using (59).
4. Update the control actor network and disturbance actor network using (60) and (61).
5. Apply the control input $\tau = \tau_d + \hat{u}(z)$ to the robot.

3.4 Stability Analysis

This subsection analyzes the stability of the closed-loop system under the proposed off-policy integral reinforcement learning-based controller. The proof follows the uniformly ultimately bounded stability criterion in [40]. The off-policy IRL formulation for solving the HJI equation is based on the equivalent integral Bellman equation discussed in [38, 39].

Let W_c, W_a, and W_d denote the ideal weights of the critic network, the control actor network, and the disturbance actor network, respectively. The corresponding weight estimation errors are defined as

$$\begin{aligned} \tilde{W}_c &= W_c - \hat{W}_c, \\ \tilde{W}_a &= W_a - \hat{W}_a, \\ \tilde{W}_d &= W_d - \hat{W}_d. \end{aligned} \tag{63}$$

The following assumptions are used for stability analysis.

Assumption 4: The ideal neural network weights are bounded, i.e.,

$$\|W_c\| \le \bar{W}_c, \quad \|W_a\| \le \bar{W}_a, \quad \|W_d\| \le \bar{W}_d, \tag{64}$$

where $\bar{W}_c$, $\bar{W}_a$, and $\bar{W}_d$ are positive constants.

Assumption 5: The activation function $\sigma(z)$ and its gradient are bounded on the compact set Ω_z, i.e.,

$$\|\sigma(z)\| \le \bar{\sigma}, \quad \|\nabla\sigma(z)\| \le \bar{\sigma}_z. \tag{65}$$

Assumption 6: The neural network approximation error and its gradient are bounded on Ω_z, i.e.,

$$\|\varepsilon_c(z)\| \le \bar{\varepsilon}_c, \quad \|\nabla\varepsilon_c(z)\| \le \bar{\varepsilon}_z. \tag{66}$$

This assumption is consistent with the universal approximation property of neural networks over compact sets [41].

Theorem 1 *Consider the augmented tracking error system (32) under Assumptions 1–6. Let the robust optimal control policy, the worst-case disturbance policy, and the value function be approximated by the three neural networks in (53), (54), and (55), respectively. If the learning gains η_c, η_a, and η_d are chosen as positive constants and the update laws (59), (60), and (61) are applied, then the augmented tracking error z, the critic weight estimation error $\tilde{W}_c$, the control actor weight estimation error $\tilde{W}_a$, and the disturbance actor weight estimation error $\tilde{W}_d$ are uniformly ultimately bounded. Consequently, the original tracking error $e = x - x_d$ is also uniformly ultimately bounded.*

Proof Consider the Lyapunov candidate function

$$\begin{aligned} L = V^*(z) + \frac{1}{2}\tilde{W}_c^T \Gamma_c^{-1} \tilde{W}_c \\ + \frac{1}{2}\tilde{W}_a^T \Gamma_a^{-1} \tilde{W}_a + \frac{1}{2}\tilde{W}_d^T \Gamma_d^{-1} \tilde{W}_d, \end{aligned} \tag{67}$$

where $\Gamma_c = \Gamma_c^T > 0$, $\Gamma_a = \Gamma_a^T > 0$, and $\Gamma_d = \Gamma_d^T > 0$ are constant positive definite learning gain matrices. Since $V^*(z)$ is positive definite and the weight error terms are quadratic, L is positive definite with respect to z, $\tilde{W}_c$, $\tilde{W}_a$, and $\tilde{W}_d$.

Taking the time derivative of (67) along the closed-loop trajectory gives

$$\begin{aligned} \dot{L} = \nabla V^{*T}(z)\dot{z} + \tilde{W}_c^T \Gamma_c^{-1} \dot{\tilde{W}}_c \\ + \tilde{W}_a^T \Gamma_a^{-1} \dot{\tilde{W}}_a + \tilde{W}_d^T \Gamma_d^{-1} \dot{\tilde{W}}_d. \end{aligned} \tag{68}$$

Using the augmented system (32), the derivative of the optimal value function can be written as

$$\dot{V}^*(z) = \nabla V^{*T}(z) \left[F(z) + G(z)\hat{u} + K(z)d\right]. \tag{69}$$

Adding and subtracting the optimal policies u^* and d^* yields

$$\begin{aligned} \dot{V}^*(z) = \nabla V^{*T}(z) \left[F(z) + G(z)u^* + K(z)d^*\right] \\ + \nabla V^{*T}(z)G(z)(\hat{u} - u^*) \\ + \nabla V^{*T}(z)K(z)(d - d^*). \end{aligned} \tag{70}$$

From the HJI equation (44), the first term in (70) satisfies

$$\begin{aligned} \nabla V^{*T}(z) \left[F(z) + G(z)u^* + K(z)d^*\right] \\ = -z^T Q z - u^{*T} R u^* + \gamma^2 d^{*T} d^*. \end{aligned} \tag{71}$$

Substituting (71) into (70), one obtains

$$\begin{aligned} \dot{V}^*(z) = -z^T Q z - u^{*T} R u^* + \gamma^2 d^{*T} d^* \\ + \nabla V^{*T}(z)G(z)(\hat{u} - u^*) \\ + \nabla V^{*T}(z)K(z)(d - d^*). \end{aligned} \tag{72}$$

From (41), (42), (53), and (54), and using Assumptions 4–6, there exist positive constants c_u, c_d, $\bar{\varepsilon}_u$, and $\bar{\varepsilon}_d$ such that

$$\begin{aligned} \|\hat{u} - u^*\| \leq c_u \|\tilde{W}_a\| + \bar{\varepsilon}_u, \\ \|\hat{d} - d^*\| \leq c_d \|\tilde{W}_d\| + \bar{\varepsilon}_d. \end{aligned} \tag{73}$$

Using the update laws (59), (60), and (61), together with Young's inequality and the boundedness assumptions, the derivative of the Lyapunov function can be upper bounded by

$$\begin{aligned} \dot{L} \leq -\lambda_z \|z\|^2 - \lambda_c \|\tilde{W}_c\|^2 \\ - \lambda_a \|\tilde{W}_a\|^2 - \lambda_d \|\tilde{W}_d\|^2 + \Delta, \end{aligned} \tag{74}$$

where λ_z, λ_c, λ_a, and λ_d are positive constants, and $\Delta > 0$ is a bounded residual term caused by neural network approximation errors and bounded disturbances.

Define

$$\xi = \left[z^T \;\; \tilde{W}_c^T \;\; \tilde{W}_a^T \;\; \tilde{W}_d^T\right]^T. \tag{75}$$

Then, (74) can be rewritten as

$$\dot{L} \leq -\lambda_{\min} \|\xi\|^2 + \Delta, \tag{76}$$

where

$$\lambda_{\min} = \min\{\lambda_z, \lambda_c, \lambda_a, \lambda_d\}. \tag{77}$$

It follows from (76) that $\dot{L} < 0$ whenever

$$\|\xi\| > \sqrt{\frac{\Delta}{\lambda_{\min}}}. \tag{78}$$

Therefore, ξ is uniformly ultimately bounded according to the UUB stability criterion in [40]. Hence, z, $\tilde{W}_c$, $\tilde{W}_a$, and $\tilde{W}_d$ are uniformly ultimately bounded. Since e is a component of z, the original tracking error $e = x - x_d$ is also uniformly ultimately bounded. This completes the proof. □

4 Simulation Results

This section presents numerical simulations to verify the tracking performance and robustness of the proposed control scheme for the Mecanum mobile robot. Two simulation cases are considered. In Case 1, the center-of-mass (COM) offset is assumed to be constant. In Case 2, the robot is affected by time-varying COM-induced and slip-related uncertainties. The performance of Actor-Critic RL and Off-policy IRL is compared in terms of trajectory tracking, tracking errors, control inputs, and neural network weight responses.

4.1 Simulation Setup

The simulations are conducted in MATLAB/Simulink to evaluate the tracking performance and robustness of the proposed control framework for the four-wheeled Mecanum mobile robot. Two learning-based controllers are implemented for comparison: the Actor-Critic RL controller and the Off-policy IRL controller. Both controllers are tested under the same robot parameters, desired trajectory, initial conditions, and uncertainty scenarios to ensure a fair comparison.

The physical parameters of the Mecanum mobile robot used in the simulations are listed in Table 1.

Table 1 Physical parameters of the Mecanum mobile robot

Robot Parameter	Description	Value
m_b	Body mass	25 kg
m_w	Wheel mass	0.35 kg
r	Wheel radius	0.05 m
a	Half-length of robot base	0.2 m
b	Half-width of robot base	0.3 m
I_b	Body moment of inertia	0.5 kg m^2
I_w	Wheel moment of inertia	4.0378×10^{-4} kg m^2
α	Roller angle	$\pi/4$ rad

The main tracking, learning, and compensation parameters of the two controllers are summarized in Table 2.

Table 2 Controller parameter settings

Parameter	Actor-Critic RL	Off-policy IRL
K_q	$\mathrm{diag}(4,4,4)$	$\mathrm{diag}(4,4,4)$
K_v	$\mathrm{diag}(3,3,5)$	$\mathrm{diag}(3,3,5)$
R	I_3	I_3
Q	$10I_6$	$10I_6$
γ	10^{-3}	0.5
η_c	2.3	0.08
η_{a1}	8	0.2
η_{a2}	10^{-3}	50
β	10^{-3}	10^{-3}
ν	10^{-3}	10^{-3}
η_a	10^{-3}	–
λ_1	–	$4I_3$
λ_2	–	50
β_1	–	6
k_s	–	100

where R and Q are the weighting matrices in the cost function, which penalize the control effort and tracking/state errors, respectively. The parameter γ denotes the discount factor or disturbance attenuation-related coefficient used in the learning formulation. The parameters η_c, η_{a1}, η_{a2}, and η_a are learning rates used to update the critic and actor neural network weights. The constants β and ν are auxiliary adaptation parameters used to regulate the online learning process and improve numerical robustness.

For the Off-policy IRL controller, λ_1 and λ_2 are RISE feedback gains associated with the filtered tracking error dynamics. The parameter β_1 is the gain of the signum term in the RISE compensation law, while k_s is an additional RISE gain used to enhance disturbance rejection. The symbol "–" indicates that the corresponding parameter is not used in that controller.

4.2 Case 1: Constant Center-of-Mass Offset

In the first case, the center-of-mass (COM) offset is assumed to be constant during the entire simulation. This condition represents a fixed payload displacement or mechanical assembly error. The COM offset is expressed as

$$x_c = \bar{x}_c, \qquad y_c = \bar{y}_c, \tag{79}$$

where $\bar{x}_c$ and $\bar{y}_c$ denote the constant COM offset values along the longitudinal and lateral directions, respectively.

Fig. 2 shows the trajectory tracking performance under constant COM offset. Both Actor-Critic RL and Off-policy IRL can drive the robot from the initial position to the desired circular trajectory. After a short transient phase, the two trajectories converge closely to the reference path. This indicates that the fixed COM-induced model mismatch can be effectively compensated by both learning-based controllers.

The reference and actual trajectory components are presented in Fig. 3. The x- and y-position responses of Off-policy IRL follow the desired sinusoidal profiles with small

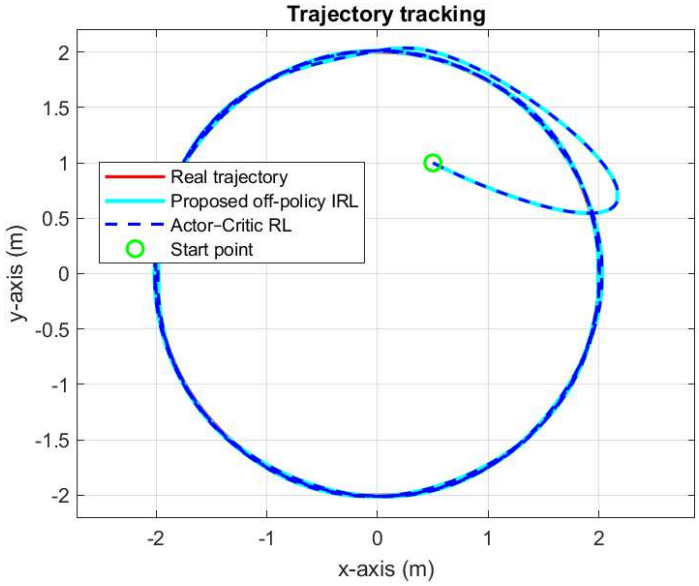


Fig. 2 Trajectory tracking in Case 1.

deviations after the initial transient. The heading response also converges rapidly and remains bounded, showing that Off-policy IRL can preserve tracking stability even in the presence of COM-induced coupling.

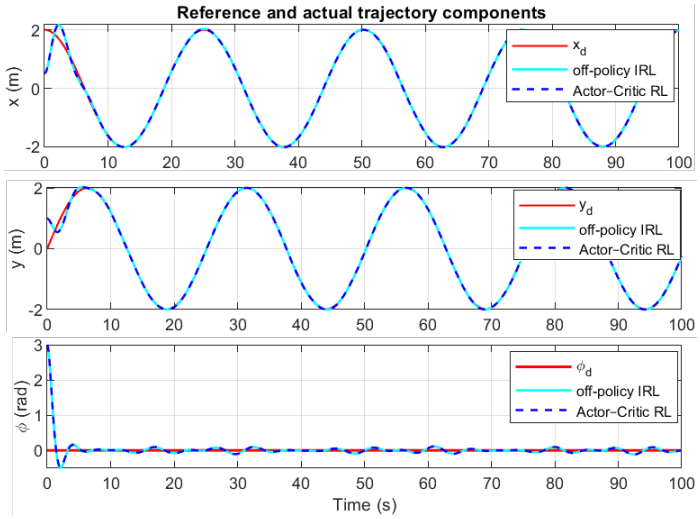


Fig. 3 Reference and actual trajectory components in Case 1.

Fig. 4 shows the tracking error comparison. The errors in the x, y, and ϕ channels decrease rapidly and then remain in a small neighborhood around zero. The error

responses of Off-policy IRL are very close to those of Actor-Critic RL, which is consistent with the RMSE comparison in Fig. 7. This result shows that, for a fixed COM offset, Off-policy IRL achieves tracking accuracy comparable to Actor-Critic RL while maintaining robust learning behavior.

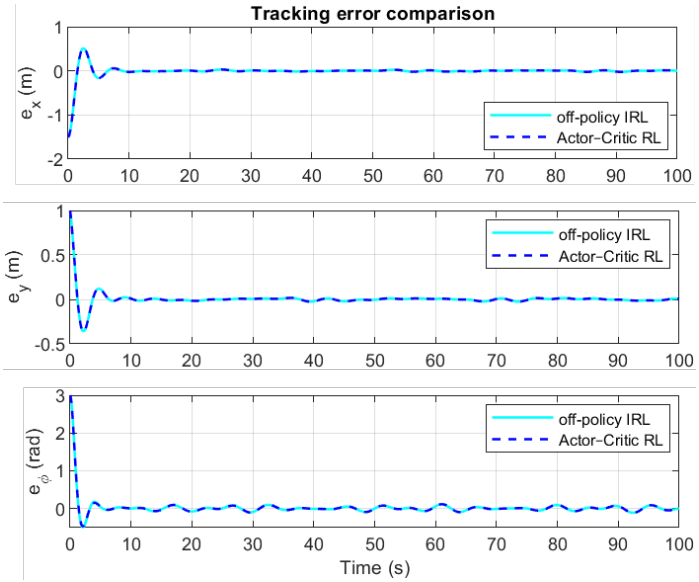


Fig. 4 Tracking error comparison in Case 1.

The wheel torque inputs of Off-policy IRL are shown in Fig. 5. The torque responses are relatively large only during the initial transient phase and then become bounded and smooth periodic signals. This demonstrates that Off-policy IRL compensates for the constant COM-induced uncertainty without producing excessive or chattering control inputs.

Fig. 6 presents the critic weights W_c, control actor weights W_a, and disturbance actor weights W_d of Off-policy IRL. The neural network weights adapt during the transient phase and remain bounded throughout the simulation. This confirms the stable online learning behavior of Off-policy IRL under constant COM-induced uncertainty.

Fig. 7 provides the RMSE comparison between the two controllers. The RMSE values are nearly identical for the x, y, and ϕ channels, indicating that both controllers achieve similar tracking accuracy in this constant-offset scenario. Therefore, Case 1 verifies that Off-policy IRL can maintain stable tracking, bounded control effort, and bounded neural network weights under fixed COM uncertainty.

4.3 Case 2: Tracking Under Time-Varying COM-Induced and Slip-Related Uncertainties

In the second case, the robot is subjected to time-varying COM-induced uncertainty, slip-related uncertainty, and payload mass variation. Compared with Case 1, this case

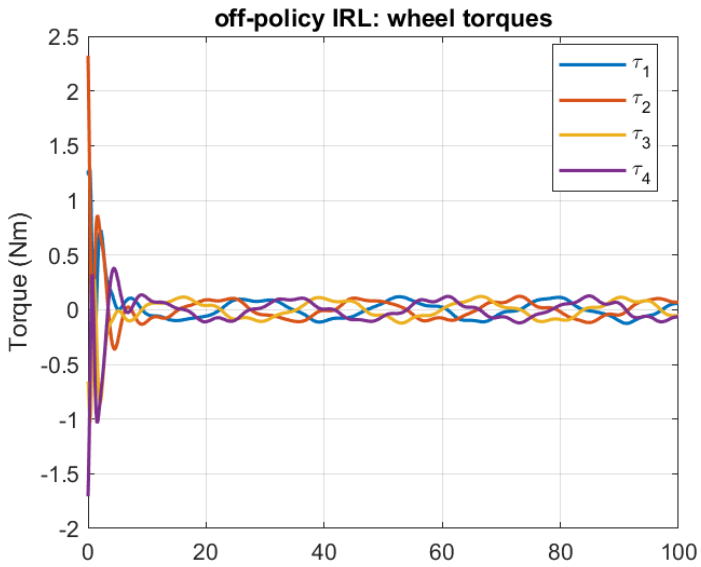


Fig. 5 Wheel torque responses of Off-policy IRL under constant COM offset.

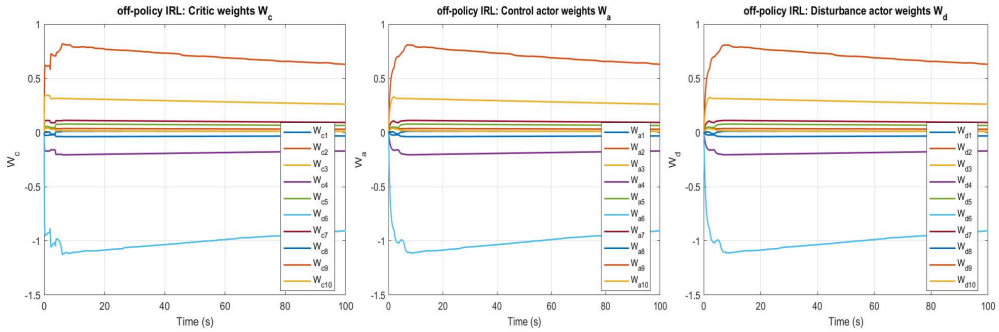


Fig. 6 Critic, control actor, and disturbance actor weight responses of Off-policy IRL in Case 1.

is more challenging because the inertial coupling, wheel–ground interaction, and total mass of the robot vary during motion. The time-varying COM offsets are selected as

$$x_c(t) = 0.035\sin(0.8t) + 0.020\cos(1.5t) + 0.020\sin(2.5t), \tag{80}$$

$$y_c(t) = 0.025\sin(0.3t) + 0.040\cos(1.2t) + 0.025\sin(2.8t) + 0.010\cos(3.5t). \tag{81}$$

In addition, the payload mass variation is set as $\Delta m = 3$ kg.

Fig. 8 shows the trajectory tracking performance under time-varying COM-induced uncertainty, slip-related uncertainty, and payload mass variation. Both controllers can still track the desired circular trajectory despite the stronger uncertainty effects. Compared with the constant COM-offset case, larger trajectory deviations appear during motion because the robot dynamics are affected simultaneously by COM variation,

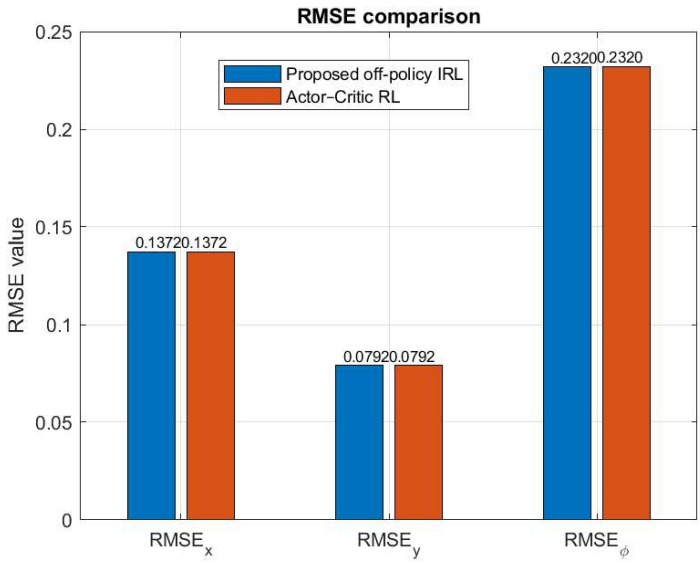


Fig. 7 RMSE comparison in Case 1.

wheel slip, and mass variation. Nevertheless, the Off-policy IRL controller remains close to the reference trajectory, showing its ability to preserve stable tracking under compounded uncertainty.

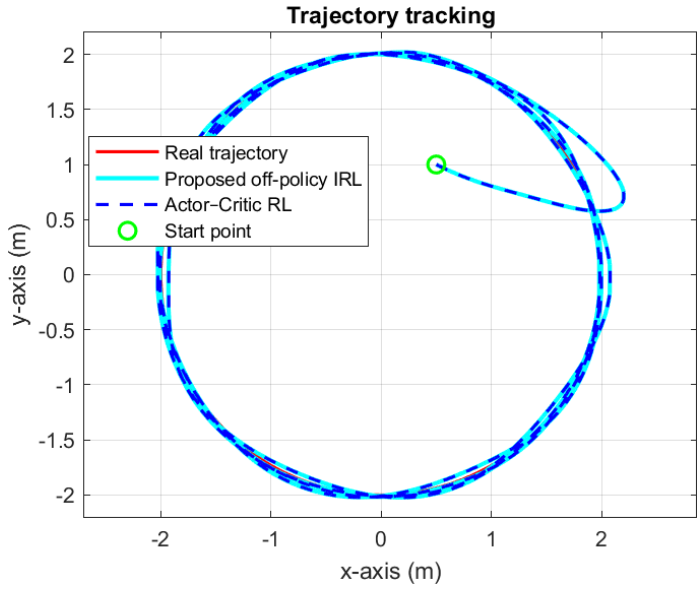


Fig. 8 Trajectory tracking in Case 2.

The reference and actual trajectory components are shown in Fig. 9. The x- and y-position responses of Off-policy IRL follow the desired trajectories with small deviations after the initial transient. The heading response exhibits more visible oscillations than in Case 1, which is mainly caused by the combined effects of time-varying COM coupling, wheel slip, and payload mass variation. However, the response remains bounded and follows the desired heading trend.

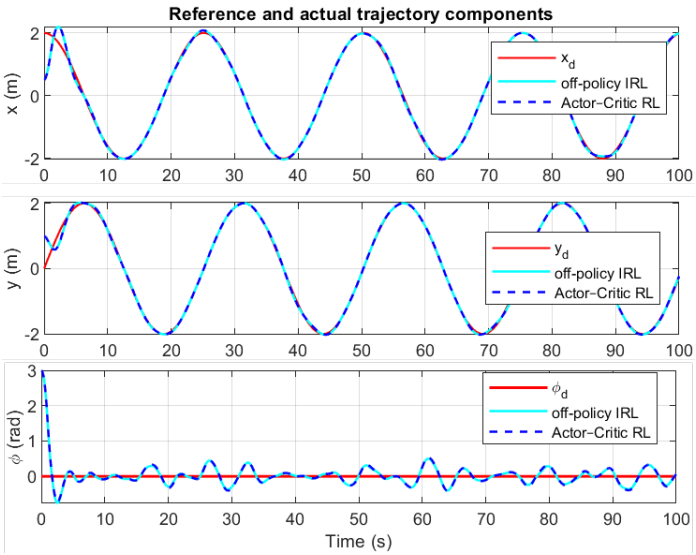


Fig. 9 Reference and actual trajectory components in Case 2.

Fig. 10 presents the tracking error comparison. The errors in the x, y, and ϕ channels converge rapidly after the initial transient and remain bounded during the whole simulation. Compared with Case 1, the residual errors become more oscillatory due to the time-varying uncertainty. Even under this condition, Off-policy IRL maintains small tracking errors and shows slightly better heading-error attenuation than Actor-Critic RL.

The wheel torque responses of Off-policy IRL are shown in Fig. 11. A large torque variation appears only at the beginning of the simulation, corresponding to the initial tracking correction. Afterward, the torque inputs remain bounded and evolve smoothly without chattering. This confirms that Off-policy IRL can compensate for the combined COM-induced, slip-related, and payload-mass uncertainties without requiring excessive control effort.

The RMSE comparison is presented in Fig. 12. The RMSE values of Off-policy IRL and Actor-Critic RL are close in the x- and y-tracking channels. For the heading angle, Off-policy IRL achieves a smaller RMSE_ϕ value, namely 0.3269 compared with 0.3316 for Actor-Critic RL. This indicates that Off-policy IRL provides slightly better robustness in the orientation channel, where the combined effects of time-varying COM coupling, wheel slip, and payload mass variation are more significant.

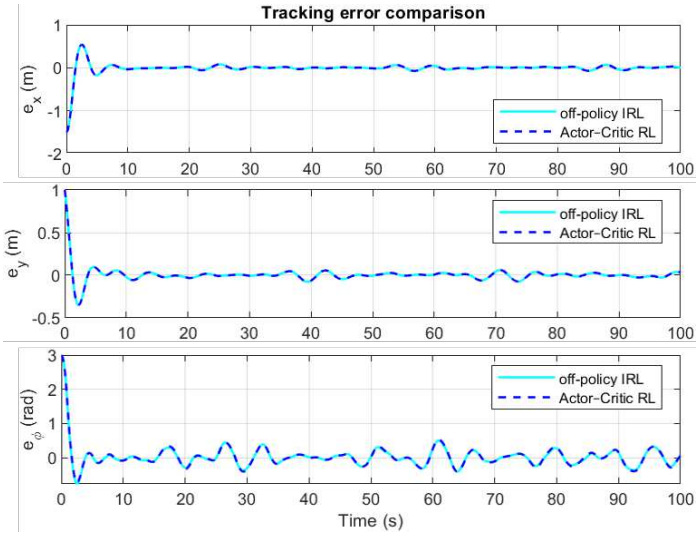


Fig. 10 Tracking error comparison in Case 2.

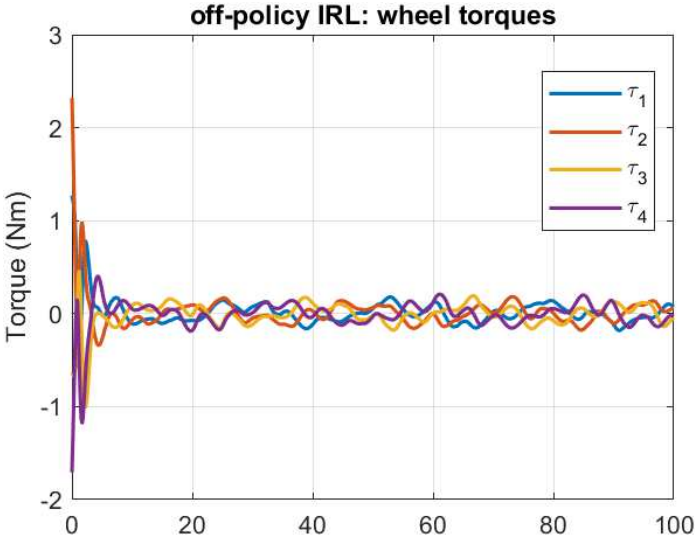


Fig. 11 Wheel torque responses of Off-policy IRL in Case 2.

Overall, Case 2 demonstrates that Off-policy IRL can maintain stable trajectory tracking under more severe time-varying uncertainty. Although the tracking errors become more oscillatory than in the constant-offset case, the controller keeps the errors and torque inputs bounded. The improved heading RMSE also suggests that Off-policy IRL is more suitable for handling uncertainty effects that strongly influence the rotational dynamics of the Mecanum mobile robot.

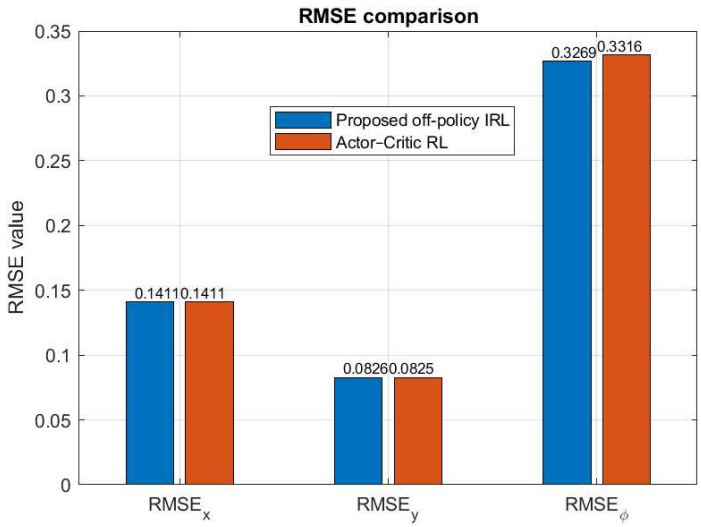


Fig. 12 RMSE comparison in Case 2.

5 Conclusion

This paper proposed an Off-policy IRL-based robust optimal tracking controller for a four-wheeled Mecanum mobile robot subject to COM-induced and slip-related uncertainties. The trajectory tracking problem was transformed into an autonomous affine nonlinear system and formulated within the Hamilton–Jacobi–Isaacs framework. To approximate the solution of the HJI equation, three neural networks were employed to estimate the optimal value function, the robust optimal control policy, and the worst-case disturbance policy. Lyapunov-based stability analysis proved that the closed-loop system is uniformly ultimately bounded.

Simulation results under constant COM offset and time-varying COM-induced, slip-related, and payload-mass uncertainties verified the effectiveness of the proposed controller. In particular, Off-policy IRL maintained bounded tracking errors and smooth control inputs, while achieving slightly better heading tracking performance than Actor-Critic RL under the combined uncertainty scenario. Future work will focus on experimental validation on a real Mecanum mobile robot platform.

- **Funding** This study received no external funding. The article processing charge (APC) will be jointly covered by the authors.

- **Data availability** No datasets were generated or analyzed during the current study.
- **Author contribution** Minh Dong Nguyen: Conceptualization, methodology, mathematical modeling, controller design, software, simulation, data analysis, visualization, and writing—original draft. Phong Pham Duy: Conceptualization, methodology, theoretical analysis, validation, supervision, writing—review and editing, and project administration. All authors reviewed and approved the final manuscript.

References

[1] Ilon, B.E.: Wheels for a course stable self-propelling vehicle movable in any desired direction on the ground or some other base. U.S. Patent 3,876,255, Apr. 8, 1975 (1975)

[2] Campion, G., Bastin, G., Dandrea-Novel, B.: Structural properties and classification of kinematic and dynamic models of wheeled mobile robots. IEEE Trans. Robot. Autom. **12**(1), 47–62 (1996) https://doi.org/10.1109/70.481750

[3] Xing, H., Xu, Y., Ding, L., Chen, J., Gao, H., Tavakoli, M.: Trajectory tracking control of wheeled mobile manipulators with joint flexibility via virtual decomposition approach. IEEE Trans. Autom. Sci. Eng. **22**, 11808–11825 (2025) https://doi.org/10.1109/TASE.2025.3540860

[4] Yamamoto, Y., Yun, X.: Modeling and compensation of the dynamic interaction of a mobile manipulator. In: Proc. IEEE Int. Conf. Robot. Autom., San Diego, CA, USA, pp. 2187–2192 (1994). https://doi.org/10.1109/ROBOT.1994.350960

[5] Lynch, K.M., Park, F.C.: Modern Robotics: Mechanics, Planning, and Control. Cambridge Univ. Press, Cambridge, U.K. (2017). https://doi.org/10.1017/9781316661239

[6] Zeidis, I., Zimmermann, K.: Dynamics of a four-wheeled mobile robot with Mecanum wheels. ZAMM-J. Appl. Math. Mech. **99**(12) (2019) https://doi.org/10.1002/zamm.201900173 . Art. no. e201900173

[7] Taheri, H., Qiao, B., Ghaeminezhad, N.: Kinematic model of a four Mecanum wheeled mobile robot. Int. J. Comput. Appl. **113**(3), 6–9 (2015) https://doi.org/10.5120/19804-1586

[8] Abdelrahman, M., Zeidis, I., Bondarev, O., Adamov, B., Becker, F., Zimmermann, K.: A description of the dynamics of a four-wheel Mecanum mobile system as a basis for a platform concept for special purpose vehicles for disabled persons. In: Proc. 58th Ilmenau Scientific Colloquium, Ilmenau, Germany (2014)

[9] Hernandez, J.C.O., Almeida, D.I.R.: Experimental control approach of a Mecanum-wheeled mobile robot for slippage error and energy consumption reduction on different surfaces. J. Mech. Sci. Technol. **38**, 6309–6318 (2024) https://doi.org/10.1007/s12206-024-1042-8

[10] Sun, Z., et al.: Trajectory-tracking control of Mecanum-wheeled omnidirectional mobile robots using adaptive integral terminal sliding mode. Comput. Electr. Eng. **96** (2021) https://doi.org/10.1016/j.compeleceng.2021.107500 . Art. no. 107500

[11] Liu, Y., *et al.*: Omni-directional mobile robot controller based on trajectory linearization. Robot. Auton. Syst. **56**(5), 461–479 (2008) https://doi.org/10.1016/

j.robot.2007.08.007

[12] Hendzel, Z., Kolodziej, M.: Robust tracking control of omni-Mecanum wheeled robot. In: Proc. Conf. Autom. Springer, ??? (2021). https://doi.org/10.1007/978-3-030-74893-7_21

[13] Solea, R., Filipescu, A., Nunes, U.: Sliding-mode control for trajectory-tracking of a wheeled mobile robot in presence of uncertainties. In: Proc. 7th Asian Control Conf., Hong Kong, China, pp. 1701–1706 (2009)

[14] Slotine, J.-J.E., Li, W.: Applied Nonlinear Control. Prentice-Hall, Englewood Cliffs, NJ, USA (1991)

[15] Vamvoudakis, K.G.: An online actor/critic algorithm for event-triggered optimal control of continuous-time nonlinear systems. In: Proc. Amer. Control Conf., Portland, OR, USA, pp. 1–6 (2014). https://doi.org/10.1109/ACC.2014.6859198

[16] Kiumarsi, B., Vamvoudakis, K.G., Modares, H., Lewis, F.L.: Optimal and autonomous control using reinforcement learning: A survey. IEEE Trans. Neural Netw. Learn. Syst. **29**(6), 2042–2062 (2018) https://doi.org/10.1109/TNNLS.2017.2773458

[17] Lewis, F.L., Liu, D. (eds.): Reinforcement Learning and Approximate Dynamic Programming for Feedback Control. Wiley, Hoboken, NJ, USA (2013). https://doi.org/10.1002/9781118453988

[18] Tlale, N., Villiers, M.: Kinematics and dynamics modelling of a Mecanum wheeled mobile platform. In: Proc. 15th Int. Conf. Mechatronics Mach. Vis. Pract., Auckland, New Zealand, pp. 657–662 (2008). https://doi.org/10.1109/MMVIP.2008.4749608

[19] Lin, L.-C., Shih, H.-Y.: Modeling and adaptive control of an omni-Mecanum-wheeled robot. Intell. Control Autom. **4**(2), 166–179 (2013) https://doi.org/10.4236/ica.2013.42021

[20] Becker, F., Bondarev, O., Zeidis, I., Zimmermann, K., Abdelrahman, M., Adamov, B.: An approach to the kinematics and dynamics of a four-wheel Mecanum vehicle. Sci. J. IFToMM Probl. Mech. **2**(55), 27–37 (2014)

[21] Padois, V., Fourquet, J.-Y., Chiron, P.: Kinematic and dynamic model-based control of wheeled mobile manipulators: A unified framework for reactive approaches. Robotica **25**(2), 157–173 (2007) https://doi.org/10.1017/S0263574707003360

[22] Yamamoto, Y., Yun, X.: Modeling and compensation of the dynamic interaction of a mobile manipulator. In: Proc. IEEE Int. Conf. Robot. Autom., San Diego, CA, USA, pp. 2187–2192 (1994). https://doi.org/10.1109/ROBOT.1994.350960

[23] Lee, D., Byun, J., Kim, H.J.: RISE-based trajectory tracking control of an aerial manipulator under uncertainty. IEEE Control Syst. Lett. **6**, 3379–3384 (2022) https://doi.org/10.1109/LCSYS.2022.3184820

[24] Lewis, F.L., Vrabie, D.L., Syrmos, V.L.: Optimal Control. Wiley, Hoboken, NJ, USA (2012). https://doi.org/10.1002/9781118122631

[25] Abu-Khalaf, M., Lewis, F.L.: Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. Automatica **41**(5), 779–791 (2005) https://doi.org/10.1016/j.automatica.2004.11.034

[26] Vamvoudakis, K.G., Lewis, F.L.: Online solution of nonlinear two-player zero-sum games using synchronous policy iteration. Int. J. Robust Nonlinear Control **22**(13), 1460–1483 (2012) https://doi.org/10.1002/rnc.1760

[27] Bhasin, S., Kamalapurkar, R., Johnson, M., Vamvoudakis, K.G., Lewis, F.L., Dixon, W.E.: A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems. Automatica **49**(1), 82–92 (2013) https://doi.org/10.1016/j.automatica.2012.09.019

[28] Zhang, H., Cui, L., Zhang, X., Luo, Y.: Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method. IEEE Trans. Neural Netw. **22**(12), 2226–2236 (2011) https://doi.org/10.1109/TNN.2011.2168538

[29] Dierks, T., Jagannathan, S.: Optimal control of affine nonlinear continuous-time systems using an online hamilton–jacobi–isaacs formulation. In: Proc. IEEE Conf. Decision Control, Atlanta, GA, USA, pp. 3048–3053 (2010). https://doi.org/10.1109/CDC.2010.5717676

[30] Zeidis, I., Zimmermann, K.: Dynamics of a four-wheeled mobile robot with Mecanum wheels. ZAMM-J. Appl. Math. Mech. **99**(12) (2019) https://doi.org/10.1002/zamm.201900173 . Art. no. e201900173

[31] Taheri, H., Qiao, B., Ghaeminezhad, N.: Kinematic model of a four Mecanum wheeled mobile robot. Int. J. Comput. Appl. **113**(3), 6–9 (2015) https://doi.org/10.5120/19804-1586

[32] Abdelrahman, M., Zeidis, I., Bondarev, O., Adamov, B., Becker, F., Zimmermann, K.: A description of the dynamics of a four-wheel Mecanum mobile system as a basis for a platform concept for special purpose vehicles for disabled persons. In: Proc. 58th Ilmenau Scientific Colloquium, Ilmenau, Germany (2014)

[33] Shao, S., Zhang, K.: RISE-adaptive neural control for robotic manipulators with unknown disturbances. IEEE Access **8**, 97729–97736 (2020) https://doi.org/10.1109/ACCESS.2020.2997383

[34] Xian, B., Queiroz, M.S., Dawson, D.M., McIntyre, M.L.: A continuous asymptotic tracking control strategy for uncertain nonlinear systems. IEEE Trans. Autom. Control **49**(7), 1206–1211 (2004) https://doi.org/10.1109/TAC.2004.831148

[35] Fischer, N., Kan, Z., Kamalapurkar, R., Dixon, W.E.: Saturated RISE feedback control for a class of second-order nonlinear systems. IEEE Trans. Autom. Control **59**(4), 1094–1099 (2014) https://doi.org/10.1109/TAC.2013.2286913

[36] Asl, H.J., Narikiyo, T., Kawanishi, M.: RISE-based prescribed performance control of euler–lagrange systems. J. Franklin Inst. **356**(13), 7144–7163 (2019) https://doi.org/10.1016/j.jfranklin.2019.06.033

[37] Shijin, C.S., Udayakumar, K.: Speed control of wheeled mobile robots using PID with dynamic and kinematic modelling. In: Proc. Int. Conf. Innov. Inf., Embedded Commun. Syst., Coimbatore, India, pp. 1–7 (2017). https://doi.org/10.1109/ICIIECS.2017.8275962

[38] Wu, H.-N., Luo, B.: Neural network based online simultaneous policy update algorithm for solving the HJI equation in nonlinear H_∞ control. IEEE Trans. Neural Netw. Learn. Syst. **23**(12), 1884–1895 (2012) https://doi.org/10.1109/TNNLS.2012.2217349

[39] Kiumarsi, B., Modares, H., Lewis, F.L.: Optimal tracking control of uncertain systems: On-policy and off-policy reinforcement learning approaches. In: Control of Complex Systems: Theory and Applications, pp. 165–186. Elsevier, Amsterdam, The Netherlands (2016). https://doi.org/10.1016/B978-0-12-805246-4.00005-7

[40] Khalil, H.K.: Nonlinear Systems, 3rd edn. Prentice Hall, Upper Saddle River, NJ, USA (2002)

[41] Hornik, K.: Approximation capabilities of multilayer feedforward networks. Neural Networks **4**(2), 251–257 (1991) https://doi.org/10.1016/0893-6080(91)90009-T