Sound & Vibration 2025, 59(6), 3820.
doi: https://doi.org/10.59400/sv3820

Article

Integral reinforcement learning based adaptive control of a RTG crane in twisting motion

Jialu Lv[1], Yongli Zhang[1,*], Bingdong Jiang[2], Changli Zhang[2], Aihua Jiang[2]

[1] Tianjin Key Laboratory of Information Sensing and Intelligent Control, School of Automation and Electrical Engineering, Tianjin University of Technology and Education, Tianjin 300222, China

[2] Guangzhou Academy of Special Equipment Inspection and Testing, Guangzhou 510510, China

*** Corresponding author:** Yongli Zhang, zhangyl@tute.edu.cn

CITATION

Lv J, Zhang Y, Jiang B, et al. Integral reinforcement learning based adaptive control of a RTG crane in twisting motion. Sound & Vibration. 2025; 59(6): 3820. https://doi.org/10.59400/sv3820

ARTICLE INFO

Received: 20 October 2025
Revised: 14 November 2025
Accepted: 21 November 2025
Available online: 11 December 2025

Abstract: The rubber-tyred gantry (RTG) crane is employed as an essential piece of equipment for container handling in port operations. The RTG crane owns time-varying characteristics and parametric uncertainties. Meanwhile, the twisting of the container during its operation has a detrimental effect on the operation efficiency. In view of this, an improved adaptive control method based on integral reinforcement learning (IRL) is proposed in this paper. The mechanism model of the RTG system is developed, and the dynamic characteristics are analysed. Then, an IRL-based adaptive controller is designed and the involved positive definite Lyapunov matrix P is optimised to improve the control performance. In contrast to classical adaptive control methods, the proposed method calculates P based on real-time state variables, thereby eliminating model reliance and guaranteeing adaptive capacity. Finally, the effectiveness of the proposed method in enhancing anti-twisting performance is verified by digital and physical experiments. In the digital experiments, compared with the classical adaptive method, the load twisting settling time is reduced by 1 s, and the maximum twisting angle is decreased by approximately 0.7 degrees. In the physical experiments, despite the influence of practical friction and disturbances, the settling time is still reduced by about 1 s. These results show that the proposed scheme possesses both theoretical effectiveness and engineering practicality.

Keywords: RTG crane; time-varying; adaptive control; integral reinforcement learning; dynamic parameter updating

1. Introduction

In the crane industry, there are two types of control problems. One is anti-swing control and the other is anti-twisting control. For the crane anti-swing problem, researchers have proposed various control strategies. Scholars have reduced the crane load swing and improved the anti-interference ability of the system through adaptive control [1–3], fuzzy control [4], PID control [5] and other methods. Zhou et al. [6] proposed a fuzzy PID control algorithm integrated with frequency converter speed regulation. The method effectively reduced the load swing angle and shortened adjustment time.

Kim et al. [7] proposed an output feedback-based tracking control strategy using an adaptive traceless Kalman filter to estimate the unmodelled friction and state variables. The swing phenomenon of the crane was effectively suppressed. Lu et al. [8] developed an online trajectory planning strategy to suppress swing in rubber-tyred

gantry cranes. This method significantly reduced container swing and twisting by optimising trajectory generation and feedback control. Zhang et al. [9] proposed a double-input double-output inversion-based feedforward control (DIDOIBFC) scheme. This scheme realized fast point-to-point (PTP) transportation for portal cranes by combining Bessel curve trajectory planning, gain scheduling feedback control (GSFC) and a PSO-optimized LQR controller. Bandong et al. [10] proposes a sliding mode method for rubber-tyred gantry (RTG) crane container positioning and anti-swing control. This approach is designed specifically for such nonlinear multivariable system problems. The controller design integrates Particle Swarm Optimization to adjust key parameters. Experimental results demonstrate the effectiveness of the container motion control scheme. The proposed control strategy significantly reduces the load swing during operation. It simultaneously ensures the final precise positioning of the container at the target location. To mitigate boom oscillations in slewing cranes induced by wind disturbances and operator inputs, Xu et al. [11] developed a hybrid anti-swing strategy. This strategy merges a cable-driven mechanism powered by four motors with a fuzzy PID controller. By formulating a dynamic model for the lifting apparatus and implementing a dedicated control platform, the method achieves coordinated motor actuation governed by intelligent algorithms. Verification tests confirm that the in-plane swing angle θ_1 is reduced by 88% and the out-of-plane angle θ_2 by 75%, improving safety and productivity. Ren et al. [12] introduce an innovative payload positioning mechanism employing a parallel cable-driven configuration, integrated with an adaptive tension-regulation strategy. Through real-time feedback, the system dynamically retracts or releases the anti-swing cables to regulate their length and tension, spatially confining the payload and suppressing its swing. Experimental and simulation validations confirm that this method reduces swing by more than 89.86%, without compromising the crane's designated load-handling capacity.

To enhance performance under parameter uncertainties and changing rope lengths, scholars have applied adaptive control [13] and sliding mode control [14, 15]. These methods effectively achieve precise positioning and efficient swing suppression for crane systems. The results show improved robustness and reduced load swing in complex operating conditions.

Beyond the aforementioned anti-swing control strategies that directly suppress payload oscillations, the coordinated performance of the crane's actuation system itself forms the foundational basis for sway mitigation. Zhang et al. [16] focused on enhancing synchronization control for hydraulic cylinders in heavy prefabricated component lifting systems. To address low precision and partial load challenges, a novel control method that combines IGWO-BP-PID with state-derivative feedback was developed. Experimental validation through co-simulation and field tests confirmed the method's effectiveness. Results showed no overshoot, faster stabilization, and higher synchronization accuracy compared to conventional PID control.

In terms of anti-twisting control, Ye et al. [17] proposed a new dynamic model and control method and the swing and twisting of the payload were effectively suppressed. The robustness and control performance were verified through numerous simulations and experiments. However, the method is structured as a feedforward

open-loop mechanism grounded in a precise dynamic model. Its operational efficacy is fundamentally contingent upon the accuracy of offline system parameter identification. The design is explicitly tailored to fixed linearized operational frequencies. This approach inherently restricts its applicability in scenarios involving real-time parameter fluctuations, such as varying suspension cable length or heterogeneous payload mass distribution. Furthermore, the robustness characteristic is principally afforded by predefined static frequency-insensitivity bands. This design does not incorporate capabilities for real-time dynamic state observation or adaptive control adjustment. Ngo et al. [18] proposed a PID-based hydraulic control method to suppress skew by dynamically adjusting the tensions of four cables, with simulations verified its efficacy. However, this research relies on linearized models and fixed-parameter assumptions, making it unable to handle the time-varying dynamics present in actual operations, including abrupt wind shifts and load variations. Its PID controller does not incorporate an online adaptive mechanism, leading to reduced robustness under changing environmental conditions. Tho et al. [19] addressed the parameter uncertainties during the anti-twisting process of spreaders by proposing a robust nonlinear integral sliding mode control method. Compared with H_∞ control, this method demonstrated superior anti-twisting performance in both simulations and experiments, making it more suitable for practical applications.

The research on anti-twisting control has rarely been investigated, whereas anti-swing control has made significant progress and the related methods have been widely applied.

In the anti-twisting control of cranes, the length of the rope and the load weight are time-varying, along with the uncertainties of certain factors, such as rope elasticity, wind interference, friction and so on. Therefore, an improved adaptive controller is designed for this time-varying system to stabilize the twisting angle of the container in this paper.

Adaptive control is a control method that can effectively cope with system uncertainties and external disturbances. This method can effectively compensate for unknown dynamics and external disturbances of the system by adjusting the controller parameters in real time. Currently, adaptive control is often combined with techniques such as sliding mode control [20–22] and robust control [23] to solve nonlinear and under-driven problems. With the development of artificial intelligence technology, the combination of adaptive control with reinforcement learning, neural networks [24] and other techniques further expands its application scope.

However, in traditional adaptive controllers, the Lyapunov matrix P is constant matrix, which may affect the control effectiveness of time-varying uncertainty systems. Therefore, how to adjust P online during the control process becomes the key of our current research.

Integral reinforcement learning (IRL) is a data-driven control method based on a value function in integral form [25, 26]. It can solve the control problem for systems with unknown dynamics, and achieve the optimal control through online learning. Liu et al. [27] proposed an online data-driven IRL algorithm for nonlinear systems with constrained inputs and achieved efficient solution of the optimal control

problem through policy iteration and neural network approximation. In addition, the combination of IRL with backstepping techniques [28], adaptive control [29, 30], and robust control [31] further enhances the performance of the control system and expands the scope of IRL applications in control systems.

Lv et al. [32] proposed an adaptive integral reinforcement learning (AIRL) scheme to solve optimal control for partially unknown multi-input nonlinear systems without directly solving the Hamilton-Jacobi-Bellman equation.

Salamat et al. [33] proposed a data-driven IRL algorithm to achieve potential energy shaping and asymptotically stable control for underactuated mechanical systems via a novel value function and actor-control policy.

Focusing on the container twisting problem, this paper achieves stable angle control by combining adaptive control with IRL.

The structure of this paper is outlined below.

Section 2 describes the structure of the RTG crane and the mechanism of the anti-twisting device. The mathematical model of the system is developed using the Lagrangian method.

Section 3 designs the IRL-based adaptive controller and analyses the stability of the system.

Section 4 presents and analyses the experimental results.

Section 5 summarizes the work conducted in this paper, analyses the limitations of the work, and provides an outlook on future research directions.

2. Anti-twisting control system for RTG cranes

2.1. Mathematical model of control system

A typical RTG crane, as shown in **Figure 1**, consists of a metal structure comprising the girder, portal legs, saddle beam, and bracing members. This structure serves as the crane's framework, supporting the entire weight and load while ensuring overall stability and shape retention. The structure of RTG crane is mainly composed of gantry bracket, power transmission system, hoisting mechanism, trolley mechanism, skew trolley mechanism and telescopic spreader. **Figure 2** shows the structural diagram of the skew trolley and spreader.

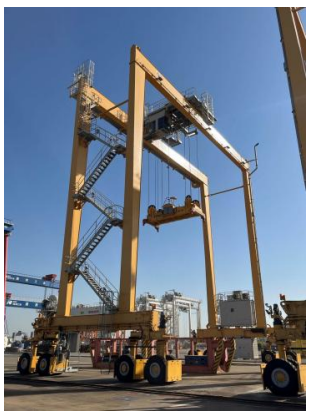

Figure 1. A RTG crane.

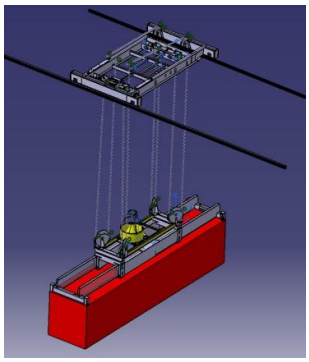

Figure 2. The structural diagram of the trolley and spreader.

The simplified structure of the skew trolley and spreader is shown in **Figure 3**. Neglecting the elastic damping characteristics of the rope and assuming that the weight of the container is uniformly distributed. The container is twisted around the z-axis, with φ serving as the twisting angle.

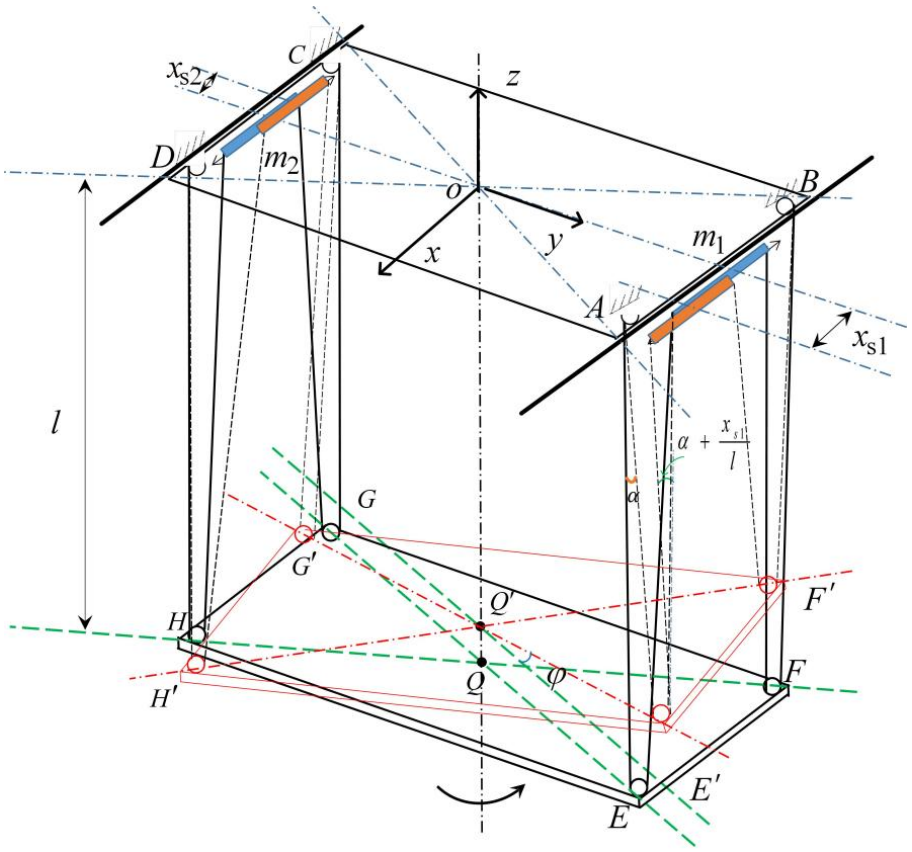

Figure 3. The simplified structure of the skew trolley and spreader.

The length of the rope is *l*, the diagonal length of the container is 2*r*, the length of the container is *a*, the width is *b*, the mass of the skew trolley and the container is *m* and *M*, respectively.

The displacement of the skew trolley is x_{s1}, x_{s2}. The anti-twisting control utilizes dual torque, taking $x_{s1} = -x_{s2}$. The tension forces in the ropes on the four pulleys are assumed to be equal.

As shown in **Figure 3**, the angle between the ropes is 2α, $\sin\alpha = x_{s1}/2l$. Then

the tension force on the rope is

$$f \approx \frac{Mg}{8}\cos\alpha \tag{1}$$

The horizontal tension force on the rope is

$$f_x \approx \frac{Mg}{8}\sin\alpha \tag{2}$$

For the small-angle oscillations pertinent to this study, the air damping can be modeled as a linear moment [34, 35]. The damping moment of the air against the container is

$$\tau = -\frac{k_c a\dot{\varphi}}{3}$$

where k_c is the damping coefficient.

According to the **Figure 3**, when the skew trolley moves x_{s1}, the vertical distance of the pulley from the upper rail is

$$\mathrm{h}' = \sqrt{\left(\frac{2l}{2}\right)^2 - \left(\frac{x_{s1}}{2}\right)^2} = \frac{1}{2}\sqrt{4l^2 - {x_{s1}}^2} \tag{4}$$

According to geometric relationships, the projection length of h' on the horizontal plane of the container is

$$\mathrm{h}'' = \sqrt{\left(a\sin\frac{\varphi}{2}\right)^2 - \left(\frac{x_{s1}}{2}\right)^2} \tag{5}$$

Therefore, the coordinates of the centre of gravity of the container are $(0,0,-\mathrm{h}_\perp)$

$$\begin{aligned} \mathrm{h}_\perp &= \sqrt{\left(\tfrac{1}{2}\sqrt{4l^2 - {x_{s1}}^2}\right)^2 - \left(\sqrt{\left(a\sin\tfrac{\varphi}{2}\right)^2 - \left(\tfrac{x_{s1}}{2}\right)^2}\right)^2} \\ &= \sqrt{l^2 - \left(2a\sin\tfrac{\varphi}{2}\right)^2} \end{aligned} \tag{6}$$

The container's velocity along the z-axis is therefore expressed as

$$\begin{aligned} v_z &= \tfrac{d\mathrm{h}_\perp}{dt} = \tfrac{1}{2}\left(l^2 - \left(2a\sin\tfrac{\varphi}{2}\right)^2\right)^{-\frac{1}{2}} \tfrac{d}{dt}\left(l^2 - \left(2a\sin\tfrac{\varphi}{2}\right)^2\right) \\ &= -\left(l^2 - \left(2a\sin\tfrac{\varphi}{2}\right)^2\right)^{-\frac{1}{2}} \left(2a\sin\tfrac{\varphi}{2}\right) \tfrac{d}{dt}\left(2a\sin\tfrac{\varphi}{2}\right) \\ &= -\left(l^2 - \left(2a\sin\tfrac{\varphi}{2}\right)^2\right)^{-\frac{1}{2}} 2a^2\sin\tfrac{\varphi}{2}\cos\tfrac{\varphi}{2}\dot{\varphi} \end{aligned} \tag{7}$$

$$v_z^2 = 4a^4\left(l^2 - \left(2a\sin\frac{\varphi}{2}\right)^2\right)^{-1}\left(\sin\frac{\varphi}{2}\cos\frac{\varphi}{2}\right)^2\dot{\varphi}^2 \tag{8}$$

The potential energy is taken to be zero on the Oxy plane. Then the potential energy of the container is

$$V_c = -Mg\mathrm{h}_\perp \tag{9}$$

The kinetic energy of the container is

$$T_c = \frac{1}{2}Mv_z^2 + \frac{1}{2}J_z\dot{\varphi}^2 \tag{10}$$

where $J_z = M(a^2+b^2)/12$.

When the container twists an angle, there is a tendency to turn back. The magnitude of the torque is related to the length of the rope. Assuming that the twist stiffness of the rope is k_l. When the length of the rope is fixed, the magnitude of the elastic rotary force is proportional to the angle of twisting, then, the elastic potential energy of twisting is

$$V_\varphi = \frac{1}{2}k_l\varphi^2 \tag{11}$$

The Lagrangian function takes the following form

$$\begin{aligned} L &= T_c - V_c - V_\varphi \\ &= \tfrac{1}{2}Mv_z^2 + \tfrac{1}{2}J_z\dot{\varphi}^2 + Mg\mathrm{h}_\perp - \tfrac{1}{2}k_l\varphi^2 \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial L}{\partial \dot{\varphi}} &= \frac{\partial}{\partial \dot{\varphi}}\left(2a^4M\left(l^2-\left(2a\sin\tfrac{\varphi}{2}\right)^2\right)^{-1}\left(\sin\tfrac{\varphi}{2}\cos\tfrac{\varphi}{2}\right)^2\dot{\varphi}^2 + \tfrac{1}{2}J_z\dot{\varphi}^2\right) \\ &\quad + \frac{\partial}{\partial \dot{\varphi}}\left(Mg\sqrt{l^2-\left(2a\sin\tfrac{\varphi}{2}\right)^2} - \tfrac{1}{2}k_l\varphi^2\right) \\ &= 4a^4M\left(l^2-\left(2a\sin\tfrac{\varphi}{2}\right)^2\right)^{-1}\left(\sin\tfrac{\varphi}{2}\cos\tfrac{\varphi}{2}\right)^2\dot{\varphi} + J_z\dot{\varphi} \end{aligned} \tag{13}$$

Neglecting higher order infinitesimals and obtain the following results

$$\frac{\partial L}{\partial \varphi} = -k_l\varphi \tag{14}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\varphi}}\right) = \frac{d}{dt}\left(4a^4M\left(l^2-\left(2a\sin\frac{\varphi}{2}\right)^2\right)^{-1}\left(\sin\frac{\varphi}{2}\cos\frac{\varphi}{2}\right)^2\dot{\varphi} + J_z\dot{\varphi}\right) \approx J_z\ddot{\varphi} \tag{15}$$

Assuming the horizontal force exerted by the servomotor on one end of the rope is F_l, with its direction aligned with the positive x-axis.

Then

$$J_z\ddot{\varphi} + \frac{k_c a}{2}\dot{\varphi} + k_l\varphi = (F_l + f)\,a = \left(ma\ddot{x}_{s1} + \frac{Mga}{2l}x_{s1}\right)a \tag{16}$$

Considering that the rope cannot transmit the force directly to the container, $m\ddot{x}_{s1}$ can be ignored. The above equation becomes

$$J_z\ddot{\varphi} + \frac{k_c a}{2}\dot{\varphi} + k_l\varphi = \frac{Mga}{2l}x_{s1} \tag{17}$$

2.2. State-space model and stability analysis of the system

The transfer function is

$$G(s) = \frac{\Phi(s)}{X_{s1}(s)} = \frac{Mga}{l(2J_zs^2 + k_cas + 2k_l)} \tag{18}$$

Therefore, the characteristic equation is given by

$$2J_zs^2 + k_c\alpha s + 2k_l = 0 \tag{19}$$

Since J_z is sufficiently large, $k_c^2a^2 - 16J_zk_l < 0$. Thus, the eigenvalues of the

system are obtained as

$$s_{1,2}=\frac{-k_c a\pm j\sqrt{16J_z k_l-k_c^2 a^2}}{4J_z} \tag{20}$$

This pole configuration causes sustained oscillations with slow decay in system response. Systems operating near stability boundaries show high sensitivity to parameter and disturbance changes, and small modeling errors or external perturbations can alter pole locations. Such changes may worsen system overshoot and oscillations in normal operation conditions, and in severe cases, poles can cross into the right half-plane, causing instability and divergence. Therefore, this system lacks adequate stability margin for reliable practical application, so compensation through controller correction is needed to increase damping and improve robustness. This shifts poles deeper into the left half-plane, enhancing both stability and dynamic performance.

The state vector is taken as $\boldsymbol{x}=\begin{bmatrix}\varphi & \dot{\varphi}\end{bmatrix}^{\mathrm{T}}$, and the control input is taken as $u=x_{s1}$. The state space equation of the system can be obtained as

$$\dot{\boldsymbol{x}}=A\boldsymbol{x}+Bu \tag{21}$$

where $A=\begin{bmatrix}0 & 1\\ -\frac{k_l}{J_z} & -\frac{k_c a}{2J_z}\end{bmatrix}$, $B=\begin{bmatrix}0\\ \frac{Mga}{2lJ_z}\end{bmatrix}$.

During crane operation, l is time-varying, k_c and k_l have uncertainties. Time-varying or uncertain parameters in the system's parameter matrix cause varying dynamic behavior under different operating conditions. This significantly increases the difficulty of controlling the container's twisting angle.

3. IRL-based adaptive control

In order to enhance the system's adaptive performance, an IRL-based adaptive controller is designed.

3.1. Fundamentals of classical adaptive control

Adaptive control systems update controller parameters based on information about the input, state or output of the system. The updates follow predefined laws to cope with the dynamic changes and uncertainties of the system [36]. There are many methods that have been used to design adaptive controllers, for example, local parameter optimisation, Lyapunov stability theory and hyperstability theory, etc.

In this paper, Lyapunov stability theory is used to design the adaptive controller, and the specific design process is as follows.

The state space equation for the reference model is

$$\dot{\boldsymbol{x}}_m=A_m\boldsymbol{x}_m+B_m r \tag{22}$$

where r is the input vector and $\boldsymbol{x}_m$ is the state vector of the reference model, A_m and B_m are the state and input matrices, respectively.

For the system, the controller is designed so that the state $\boldsymbol{x}$ of the actual system can be consistent with the state $\boldsymbol{x}_m$ of the reference model. The output signal of the

controller is

$$u = Gr - F\boldsymbol{x} \tag{23}$$

where F denotes the feedback compensation matrix, G denotes the feedforward gain matrix.

Substituting (23) into (21) yields

$$\dot{\boldsymbol{x}} = (A - BF)\,\boldsymbol{x} + BGr \tag{24}$$

The generalized error state vector is

$$\boldsymbol{e} = \boldsymbol{x}_m - \boldsymbol{x} \tag{25}$$

While its derivative can be expressed as

$$\dot{\boldsymbol{e}} = \dot{\boldsymbol{x}}_m - \dot{\boldsymbol{x}} = A_m\boldsymbol{e} + (A_m - A + BF)\,\boldsymbol{x} + (B_m - BG)\,r \tag{26}$$

When the system model is the same as the reference model, the controllers are F_m and G_m. So

$$\begin{cases} A_m = A - BF_m \\ B_m = BG_m \end{cases} \tag{27}$$

Substituting (27) into (26) yields

$$\dot{\boldsymbol{e}} = A_m\boldsymbol{e} - B_m{G_m}^{-1}\widetilde{F}\boldsymbol{x} + B_m{G_m}^{-1}\widetilde{G}r \tag{28}$$

with $\widetilde{F} = F_m - F$, $\widetilde{G} = G_m - G$.

In order to prove the convergence of the system, the Lyapunov function is chosen as

$$V_a = \boldsymbol{e}^{\mathrm{T}}P\boldsymbol{e} + \mathrm{tr}\left(\widetilde{F}^{\mathrm{T}}{P_F}^{-1}\widetilde{F}\right) + \mathrm{tr}\left(\widetilde{G}^{\mathrm{T}}{P_F}^{-1}\widetilde{G}\right) \tag{29}$$

where P is the solution to the Lyapunov equation $A^{\mathrm{T}}P + PA = -Q$, P_F and P_G are positive definite (PD) symmetric matrices.

The derivative of the Lyapunov function is then given by

$$\begin{aligned} \dot{V}_a &= \dot{\boldsymbol{e}}^{\mathrm{T}}P\boldsymbol{e} + \boldsymbol{e}^{\mathrm{T}}P\dot{\boldsymbol{e}} + \mathrm{tr}\left(\dot{\widetilde{F}}^{\mathrm{T}}P_F^{-1}\widetilde{F} + \widetilde{F}^{\mathrm{T}}{P_F}^{-1}\dot{\widetilde{F}}\right) \\ &\quad + \mathrm{tr}\left(\dot{\widetilde{G}}^{\mathrm{T}}{P_G}^{-1}\widetilde{G} + \widetilde{G}^{\mathrm{T}}{P_G}^{-1}\dot{\widetilde{G}}\right) \\ &= -\boldsymbol{e}^{\mathrm{T}}Q\boldsymbol{e} + 2\mathrm{tr}\left(B^{\mathrm{T}}P\boldsymbol{e}Z^{\mathrm{T}}\varphi^{\mathrm{T}}\right) + 2\mathrm{tr}\left(\dot{\Phi}\Gamma^{-1}\Phi^{\mathrm{T}}\right) \end{aligned} \tag{30}$$

where Q is taken as a symmetric PD matrix, $Z = [\boldsymbol{x} \quad r]^{\mathrm{T}}$, $\Phi = [-\widetilde{F} \quad \widetilde{G}]$, $\Gamma = \begin{bmatrix} P_F & 0 \\ 0 & P_G \end{bmatrix}$.

Let

$$2\mathrm{tr}\left(B^{\mathrm{T}}P\boldsymbol{e}Z^{\mathrm{T}}\Phi\right) + 2\mathrm{tr}\left(\dot{\Phi}\Gamma^{-1}\Phi\right) = 0 \tag{31}$$

Gives

$$\Phi = -B^{\mathrm{T}} P \boldsymbol{e} Z^{\mathrm{T}} \Gamma \tag{32}$$

Thus, the adaptive law is designed as

$$\begin{cases} \dot{F} = -P_F B^{\mathrm{T}} P \boldsymbol{e} \boldsymbol{x}^{\mathrm{T}} \\ \dot{G} = P_G B^{\mathrm{T}} P \boldsymbol{e} r^{\mathrm{T}} \end{cases} \tag{33}$$

3.2. IRL-based improvements

In classical adaptive control, P is fixed and cannot be dynamically adapted to the parameter changes of the time-varying system. This may lead to degradation of control performance. Therefore, in this paper, a dynamic adjustment strategy is used to enhance the adaptability to the time-varying and uncertainty of the system. The IRL algorithm is utilized to calculate P.

In modern intelligent control theory, reinforcement learning is a method centered on environmental interaction. It uses feedback rewards to gradually approximate the optimal decision-making mechanism. This capability grants reinforcement learning significant importance in complex system control. Its core principle relies on modeling with Markov decision processes and balancing the exploration of new actions with the exploitation of known ones. This balance enables autonomous learning even when system models are unknown or uncertain. Consequently, reinforcement learning exhibits strong adaptability and generalization capabilities. IRL forms a critical branch of reinforcement learning by employing integral Bellman equations to directly design and analyse optimal controllers for continuous-time systems.

In recent years, scholars have extensively applied IRL to robotic control [37], autonomous system formation [38] and power system optimization. This method addresses key challenges like unknown models, underactuation, and high energy consumption. Mahmoud et al. [39] propose a model-free static output-feedback optimal control method based on IRL. It solely utilizes input-output data and eliminates the need for any system model or state observer. The approach achieves a balanced performance, stability, and practicality within an off-policy framework. Its overall effectiveness outperforms existing model-free methods. Lv et al. [40] employed online adaptive IRL for finite-time optimal control of multi-input nonlinear systems, incorporating Nash equilibrium. This method uses dual neural networks to approximate the performance index, avoiding reliance on internal system dynamics. It has been successfully applied to models like the F-16 aircraft, achieving optimized performance within a finite time.

Compared to other reinforcement learning methods, such as Q-learning or deep deterministic policy gradient (DDPG), its advantages are mainly reflected in the following aspects.

(i) In contrast to discrete-time methods like Q-learning, IRL operates within a rigorous continuous-time framework. This approach fundamentally eliminates model errors resulting from time discretization. Its theoretical structure is inherently more rigorous and systematic.

(ii) IRL executes integration in the continuous domain, differing from discrete-time algorithms like DDPG. It eliminates the need for a known or precisely identified system dynamics model. This method achieves adaptive optimal control solely via online input-output data, minimizing prior knowledge dependence.

(iii) The learning process incorporates Lyapunov stability theory. This provides enhanced stability guarantees for closed-loop control systems and addresses the assurance limitations of traditional reinforcement learning.

Therefore, this paper employs IRL-based adaptive control to regulate the torsional angle of the crane load. This algorithm evaluates control performance via an online system, utilizing IRL to dynamically optimize key adaptive control law parameters. For time-varying systems, conventional fixed parameter matrices rely heavily on costly prior experience. Consequently, they struggle to balance convergence speed and control smoothness under dynamic variations. The proposed method overcomes this by dynamically optimizing these matrices online using real-time system states. This enables real-time tuning of the update law itself, thereby enhancing the controller's adaptability to time-varying uncertainties. The system control block diagram is shown in **Figure 4**.

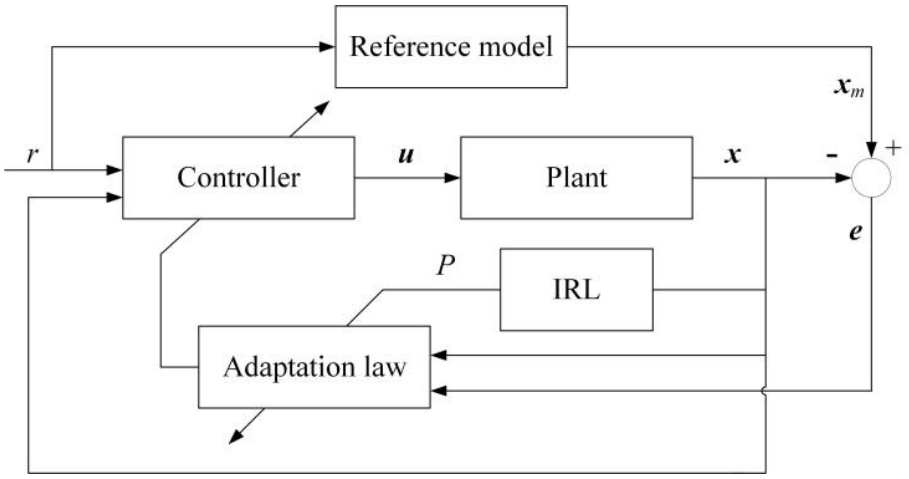

Figure 4. The system control block diagram.

When $r=0$, the control law is $u=-F\boldsymbol{x}$ and the system $\dot{\boldsymbol{x}}=(A-BF)\boldsymbol{x}$ is a stable closed-loop system. Then the value function is

$$V(\boldsymbol{x}(t))=\int_t^{\infty}\boldsymbol{x}^{\mathrm{T}}(\tau)Q\boldsymbol{x}(\tau)d\tau=\boldsymbol{x}^{\mathrm{T}}(t)P\boldsymbol{x}(t) \tag{34}$$

where P is a real symmetric PD solution of the equation $(A-BF)^{\mathrm{T}}P+P(A-BF)=-Q$.

The integral strengthening form of the value function is

$$V(\boldsymbol{x}(t))=\int_t^{t+T}\boldsymbol{x}^{\mathrm{T}}(\tau)Q\boldsymbol{x}(\tau)d\tau+V(\boldsymbol{x}(t+T)) \tag{35}$$

IRL strategy iterative algorithm

$$\boldsymbol{x}^{\mathrm{T}}(t)P_i\boldsymbol{x}(t)=\int_t^{t+T}\boldsymbol{x}(\tau)^{\mathrm{T}}Q\boldsymbol{x}(\tau)d\tau+\boldsymbol{x}^{\mathrm{T}}(t+T)P_i\boldsymbol{x}(t+T) \tag{36}$$

$$\dot{F}_i = -P_F B^{\mathrm{T}} P_i E \tag{37}$$

with $E = \boldsymbol{e}\boldsymbol{x}^{\mathrm{T}}$. $i \in \mathbb{Z}^+$ denotes the iteration count in policy iteration. $\dot{F}_i$ is the derivative of the feedback matrix.

Equation (36) corresponds to the policy evaluation step of the policy iteration algorithm. The left-hand side of this formula represents the estimated total future cost at the current moment. The first term on the equation's right side quantifies the cumulative immediate stage cost over a future finite time window. This term measures the system performance loss during the specific operating period described above. The second right-hand term provides the cost estimate at the end of the stated time window. This equation establishes the theoretical foundation for policy performance evaluation in continuous-time systems.

Equation (37) corresponds to the policy improvement step in the policy iteration algorithm. This step calculates new feedback control gain matrices based on the previously evaluated cost information. It transforms performance insights from cost evaluation into executable controller parameters. This mechanism completes a full iteration from policy evaluation to policy generation. The new gain updates the system's actual control law.

Lemma 1. *Assume the system $\dot{\boldsymbol{x}} = A_i \boldsymbol{x}$ with $A_i = A_i - BF_i$ is stable, P_i being a solution to (36) is equivalent to finding a solution to the Lyapunov equation.*

$$A_i^{\mathrm{T}} P_i + P_i A_i = -Q \tag{38}$$

Proof. Since A_i is a stable matrix and $Q > 0$, there exists a unique solution to (38). Moreover, since $V(\boldsymbol{x}(t)) = \boldsymbol{x}^{\mathrm{T}}(t) P \boldsymbol{x}(t)$, $\forall \boldsymbol{x}(t)$ is a Lyapunov function of the system $\dot{\boldsymbol{x}} = A_i \boldsymbol{x}$ and

$$\frac{d(\boldsymbol{x}^{\mathrm{T}}(t) P_i \boldsymbol{x}(t))}{dt} = \boldsymbol{x}^{\mathrm{T}}(t)(A_i^{\mathrm{T}} P_i + P_i A_i)\boldsymbol{x}(t) = -\boldsymbol{x}^{\mathrm{T}}(t) Q \boldsymbol{x}(t) \tag{39}$$

Then, $\forall T > 0$, the unique solution of the Lyapunov equation satisfies

$$\begin{aligned} \int_t^{t+T} \boldsymbol{x}(\tau)^{\mathrm{T}} Q \mathrm{x}(\tau) d\tau &= \int_t^{t+T} \frac{d(\boldsymbol{x}^{\mathrm{T}}(\tau) P_i \boldsymbol{x}(\tau))}{d\tau} d\tau \\ &= \boldsymbol{x}^{\mathrm{T}}(t) P_i \boldsymbol{x}(t) - \boldsymbol{x}^{\mathrm{T}}(t+T) P_i \boldsymbol{x}(t+T) \end{aligned} \tag{40}$$

The above equation is equivalent to (36). Thus, assume the system $\dot{\boldsymbol{x}} = A_i \boldsymbol{x}$ is asymptotically stable, with the solution of (36) being the unique solution of (38). □

To calculate P_i, the method described in Vrabie et al.'s study [41] was employed. $\boldsymbol{x}^{\mathrm{T}} P_i \boldsymbol{x}$ is expressed as

$$\boldsymbol{x}^{\mathrm{T}} P_i \boldsymbol{x} = \overline{\boldsymbol{p}}_i^{\mathrm{T}} \overline{\boldsymbol{x}} \tag{41}$$

where $\overline{\boldsymbol{x}}$ denotes the Kronecker product quadratic polynomial basis vector having elements $\{x_i(t)x_j(t)\}_{i=1,n;j=1,n}$. The parameter vector $\overline{\boldsymbol{p}}_i$ contains the elements of the matrix P_i ordered by columns and with the redundant elements removed.

By utilizing (41), (36) can be expressed as

$$\overline{\boldsymbol{p}}_i^{\mathrm{T}}(\overline{\boldsymbol{x}}(t)-\overline{\boldsymbol{x}}(t+T)) = \int_t^{t+T} \boldsymbol{x}^{\mathrm{T}}(\tau)Q\boldsymbol{x}(\tau)d\tau \tag{42}$$

where $\overline{\boldsymbol{p}}_i$ is the vector of unknown parameters and $\overline{\boldsymbol{x}}(t)-\overline{\boldsymbol{x}}(t+T)$ acts as a regression vector. The right-hand side

$$d(\overline{\boldsymbol{x}}(t)) = \int_t^{t+T} \boldsymbol{x}^{\mathrm{T}}(\tau)Q\boldsymbol{x}(\tau)d\tau \tag{43}$$

is a desired value or target function to which $\overline{\boldsymbol{p}}_i^{\mathrm{T}}(\overline{\boldsymbol{x}}(t)-\overline{\boldsymbol{x}}(t+T))$ is equal when the parameter vector $\overline{\boldsymbol{p}}_i$ contains the correct parameters.

To calculate it efficiently, define a controller state $V_1(t)$ and add the state equation $\dot{V}_1(t)$ to the controller dynamics.

$$\dot{V}_1(t) = \boldsymbol{x}^{\mathrm{T}}(t)\mathrm{Q}\boldsymbol{x}(t) \tag{44}$$

By measuring $V_1(t)$ along the system trajectory, the value of $d(\overline{\boldsymbol{x}}(t))$ can be computed by using $d(\overline{\boldsymbol{x}}(t)) = V_1(t+T)-V_1(t)$. The right-hand side of (42) must be computed at $N \geq n(n+1)/2$ (where *n* is the dimension of the matrix *P*) points $\overline{\boldsymbol{x}}^i$ in the state space, over time intervals *T*. Then, the batch least-squares solution is obtained as

$$\overline{\boldsymbol{p}}_i = (XX^{\mathrm{T}})^{-1}XY \tag{45}$$

with $X = \begin{bmatrix} \overline{\boldsymbol{x}}_\Delta^1 & \overline{\boldsymbol{x}}_\Delta^2 & \cdots & \overline{\boldsymbol{x}}_\Delta^N \end{bmatrix}^{\mathrm{T}}$, $\overline{\boldsymbol{x}}_\Delta^i = \overline{\boldsymbol{x}}^i(t)-\overline{\boldsymbol{x}}^i(t+T)$, $Y = \begin{bmatrix} d(\overline{\boldsymbol{x}}^1) & d(\overline{\boldsymbol{x}}^2) & \cdots & d(\overline{\boldsymbol{x}}^N) \end{bmatrix}^{\mathrm{T}}$.

The design procedure of the IRL-based adaptive control algorithm can be summarized as follows:

(i) Initialize the controller parameters.
(ii) Calculate the value of $\overline{\boldsymbol{p}}_i$ according to (45).
(iii) Verify whether the calculated $\overline{\boldsymbol{p}}_i$ satisfies the specified conditions.
(iv) Determine the matrix P_i based on the relationship between $\overline{\boldsymbol{p}}_i$ and P_i given in (41).
(v) Update the adaptive controller parameters.

The proposed optimization process for the self-adaptive controller effectively reduces errors caused by system model uncertainties. By dynamically adjusting its parameters, the controller's effectiveness and practicality are improved, thereby enhancing system adaptability and control accuracy in complex conditions.

3.3. Stability analysis

The stability of the controller is analysed below. When $x_m = 0$, (25) can be written as $\boldsymbol{e} = -\boldsymbol{x}$, so $E = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}\begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} -x_1^2 & -x_1x_2 \\ -x_1x_2 & -x_2^2 \end{bmatrix}$, and $-E = \begin{bmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{bmatrix}$. Since all the principal minors of $-E$ are nonnegative, $-E$ is positive semidefinite.

Therefore, E is negative semidefinite.

Theorem 1. *Assume that the control strategy F_i is stabilising at iteration i and that V_i=$\boldsymbol{x}^T P_i \boldsymbol{x}$ is the relevant value. To update the control strategy with (37) yields a stabilized strategy F_{i+1}, which secures the stability of the system.*

Proof. The PD cost function V_i is used as a candidate for the Lyapunov function for state trajectories generated when using the controller F_{i+1}. Taking the derivative of V_i along the trajectory generated by F_{i+1}, one obtains

$$\begin{aligned}\dot{V}_i &= \boldsymbol{x}^{\mathrm{T}}[P_i(A-BF_{i+1}) + (A-BF_{i+1})^{\mathrm{T}}P_i]\boldsymbol{x} \\ &= \boldsymbol{x}^{\mathrm{T}}[P_i(A-BF_i) + (A-BF_i)^{\mathrm{T}}P_i]\boldsymbol{x} \\ &+ \boldsymbol{x}^{\mathrm{T}}[P_iB(F_i-F_{i+1}) + (F_i-F_{i+1})^{\mathrm{T}}B^{\mathrm{T}}P_i]\boldsymbol{x}\end{aligned} \tag{46}$$

When $\Delta t \to 0$, $\dot{F}_i = \lim\limits_{\Delta t\to 0} \frac{F_{i+1}-F_i}{\Delta t}$. The second term above can be written as

$$\begin{aligned}&\Delta t \boldsymbol{x}^{\mathrm{T}}[P_iB\tfrac{F_i-F_{i+1}}{\Delta t} + (\tfrac{F_i-F_{i+1}}{\Delta t})^{\mathrm{T}}B^{\mathrm{T}}P_i]\boldsymbol{x} \\ &= \Delta t \boldsymbol{x}^{\mathrm{T}}[\lim_{\Delta t\to 0} E^{+}(\tfrac{F_i-F_{i+1}}{\Delta t})^{\mathrm{T}}P_F^{-1}\tfrac{F_i-F_{i+1}}{\Delta t}] \\ &+ \lim_{\Delta t\to 0}(\tfrac{F_i-F_{i+1}}{\Delta t})^{\mathrm{T}}P_F^{-1}\tfrac{F_i-F_{i+1}}{\Delta t}\quad E^{+}]\,\boldsymbol{x}\end{aligned} \tag{47}$$

where E^+ is computed from E through eigenvalue decomposition.

Using (38), the first term in (46) can be written as $-\boldsymbol{x}^{\mathrm{T}}Q\boldsymbol{x}$, and adding the resulting two terms

$$\begin{aligned}\dot{V}_i &= -\boldsymbol{x}^{\mathrm{T}}Q\boldsymbol{x} \\ &+\Delta t\boldsymbol{x}^{\mathrm{T}}\Big[\lim_{\Delta t\to 0}\mathrm{E}^{+}(\tfrac{F_i-F_{i+1}}{\Delta t})^{\mathrm{T}}P_F^{-1}\tfrac{F_i-F_{i+1}}{\Delta t} \\ &+ \lim_{\Delta t\to 0}(\tfrac{F_i-F_{i+1}}{\Delta t})^{\mathrm{T}}{P_F}^{-1}\tfrac{F_i-F_{i+1}}{\Delta t}\quad E^{+}]\,\boldsymbol{x}\end{aligned} \tag{48}$$

Thus, when $Q > 0$, $V_i < 0$, so the updated control strategy $u = -F_{i+1}\,\boldsymbol{x}$ is stable. Therefore, the matrix P_i can be obtained by solving (36) and the system is always stable during the iteration. □

4. Experiments

This part of the paper assesses the effectiveness and practicality of the control method through digital and physical experiments. In the digital experiments, take M = 40,000 kg, g = 9.81 m/s^2, α = 12 m, k_l = 0.01 and k_c = 0.01. The value of l varies within the range of 5 m to 20 m.

The control requirements of the system are: when the initial angle is 2.5 degrees, the twisting angle of the container reaches basic stability within 6 s. Meanwhile, during the control process, the control input is in the interval range of [−0.4, 0.4].

4.1. Digital experiments

The digital experiments verify the performance of the IRL-based adaptive control method for anti-twisting of RTG cranes. The variation of the rope length of the crane spreader is shown in **Figure 5**. The figure simulates the process of lifting a container

at one time. Take the controller parameters as $P_F = 10$ and $P_G = 1$ respectively.

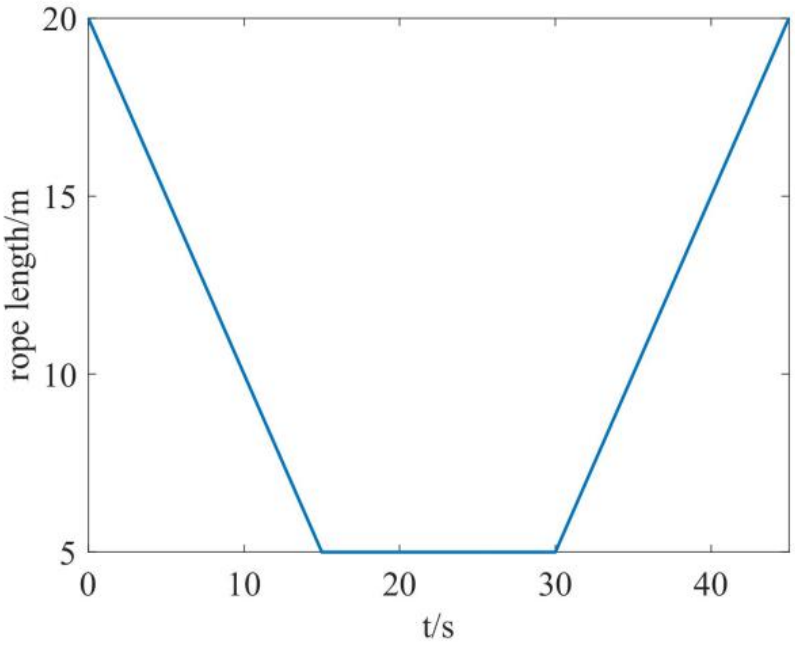

Figure 5. The variation curve of rope length.

Note: The first 15 s is the hoisting phase. The rope length decreases from 20 m to 5 m at 1 m per second. This simulates lifting the cargo from the ground. The 15 to 30 s is the horizontal transport phase with a constant rope length. The final 15 s is the lowering phase. The rope length increases at the same rate. This simulates placing the cargo at the designated location.

In the simulation, the container's twisting angle is controlled using three distinct approaches: classical adaptive controller, IRL-based adaptive controller and LQR controller. The simulation curves of container twisting angle and skew trolley displacement are shown in **Figures 6** and **7**.

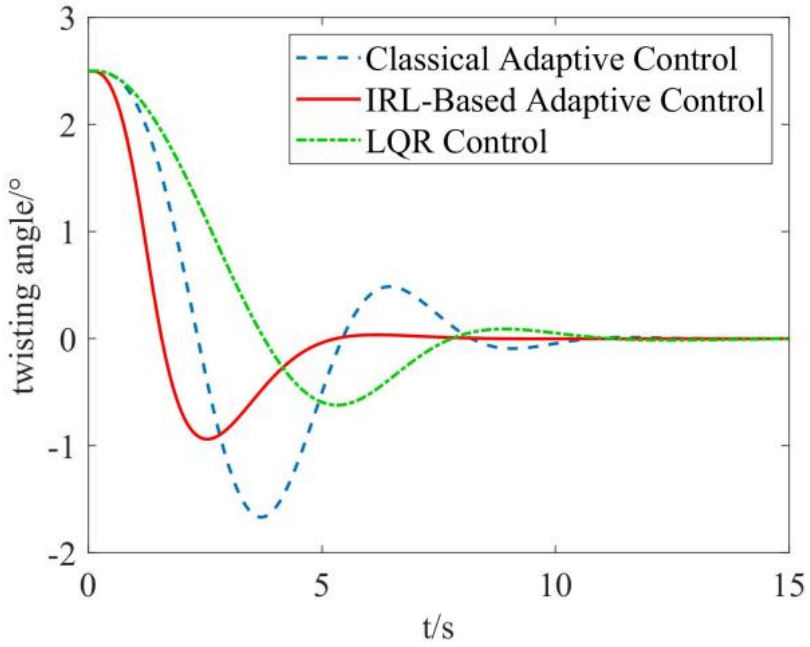

Figure 6. Twisting angle variation curves.

Note: The figure compares three control strategies: traditional adaptive control, LQR control and the IRL-based adaptive control proposed in this paper. All cases share the same initial condition with a swing angle of 10 degrees.

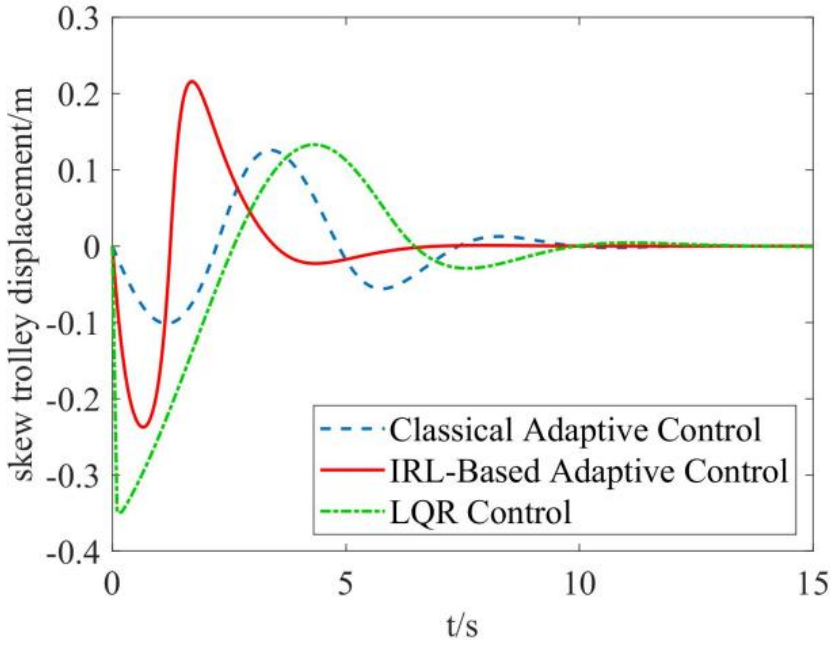

Figure 7. Displacement variation curves of the skew trolley.
Note: The figure compares the skew trolley displacement under three control strategies: traditional adaptive control, LQR control and the IRL-based adaptive control. The operational displacement is bounded within ±0.4 m.

The analysis of digital experiment results shows that the optimisation of the P in the IRL-based adaptive controller results in an effective improvement of the system.

The IRL-based adaptive control achieves faster system response than classical adaptive control. The container's twisting angle decays rapidly to within ±1° and stabilizes within 6 s. At the same time, the time for the system to reach stabilisation is reduced by 3 s and the maximum twisting angle of the container during control is reduced by 0.75 degrees. In addition, the displacement of the skew trolley is always in the interval [−0.4, 0.4]. This satisfies the control requirements. Moreover, compared with LQR control, the adaptive control can achieve rapid angle stabilization with smaller control effort.

To facilitate a more intuitive comparison of the overall performance across different control methods, **Table 1** presents key performance indicators for three distinct approaches.

Table 1. Performance comparison of different control methods.

Control method	Maximum twisting angle/°	Settling time/s	Algorithm Execution Time/ms	Maximum Displacement/m
LQR	0.62	11	0.06	0.35
Classical adaptive control	1.67	10.4	0.35	0.11
IRL-based adaptive control	0.92	5.54	3.54	0.24

Regarding algorithm real-time performance, the adaptive controller based on inverse reinforcement learning proposed in this paper requires longer computation time per iteration compared to the other two methods. This increase mainly stems from the iterative computational demands of online policy evaluation and parameter updates. However, the execution time of this controller remains shorter than the system's defined

sampling period. It can complete all computations and output control commands within one sampling interval, thereby avoiding system instability or performance degradation due to computational delays. Therefore, despite its relatively higher computational complexity, this scheme still satisfies the real-time constraints of control systems.

Therefore, IRL-based adaptive control is able to achieve fast control of the angle while satisfying the displacement constraints. This control strategy exhibits enhanced practical applicability.

4.2. Physical experiments

In a verification of the engineering practicability of the proposed method, this study conducts physical experiments based on an electronic anti-twisting and anti-swing crane experimental system.

As shown in **Figure 8**, a small-scale RTG model is designed.

Figure 8. A small-scale RTG model, consisting of an angle detection system, a motion control system, and a monitoring system.

The system consists of a servo drive system, an angle detection system, a motion control system, and a monitoring system, forming a complete closed-loop control system. The servo drive system includes two servo motors, each with a matching driver, to provide a horizontal force couple for the skew trolley. The angle detection system uses a photoelectric encoder to measure the load angle in real time. A motion control card is employed for collecting system status and outputting servo motor drive commands. The monitoring system, built with an industrial computer, performs data acquisition and human-machine interaction.

The most difficult control situation is adopted during experiments. The maximum rope length is 1.5 m and the total load mass is 5 kg during experiments. The responses of the anti-twisting control are shown in **Figure 9**. The proposed method can achieve the rapid stabilisation of the container twisting angle compared with other method.

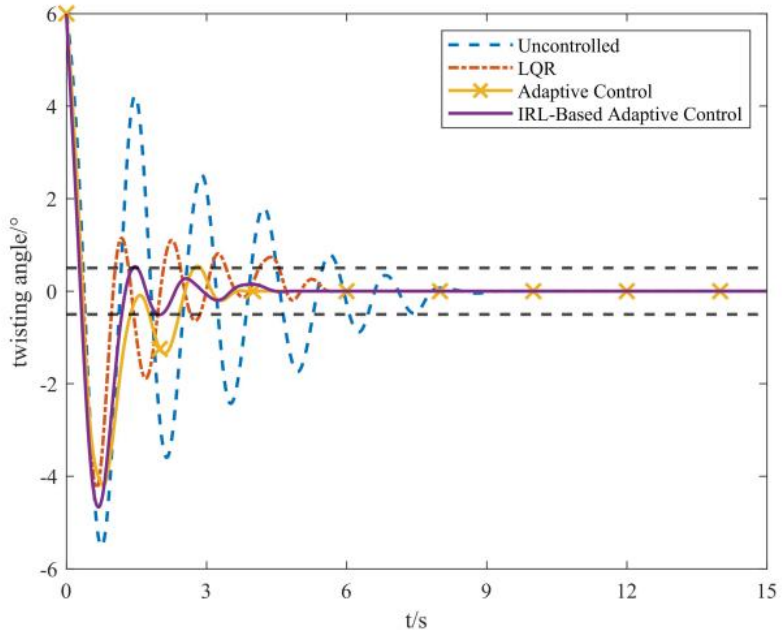

Figure 9. The twisting angle curves of the load.

Note: This figure presents a comparison among four cases: the uncontrolled case, adaptive control, LQR control, and the proposed IRL-based control. A common initial twisting angle of 6 degrees and a load mass of 5 kg are set for all cases.

As shown in **Figure 9**, under uncontrolled conditions, the load's twisting angle exhibits a decaying oscillation due to air resistance and inherent damping. However, the attenuation is slow, taking about 6 cycles for the system to stabilize. This indicates insufficient natural damping, which fails to meet the practical requirement for rapid stabilization in engineering. After implementing the LQR control strategy, the system's dynamic response has been improved. The oscillation amplitude of the load twisting angle is reduced, while the angle decays faster, entering the ±0.5° error band in 4.5 s. The system achieves a stable state within 5 oscillation cycles. This indicates that the LQR control enhances the system's damping performance and suppresses load twisting.

After implementing adaptive control, the dynamic performance of the system shows some improvement. The previous oscillatory behavior is reduced, with the response entering the error band around 2.8 s and essentially stabilizing at approximately 4 s.

The proposed control scheme further accelerates angle decay. It achieves rapid attenuation within one cycle, enters the error band around 2 s, and thus reduces overshoot and oscillation duration. The system reaches essential stability after just 3 cycles. Compared with the LQR and adaptive control methods, the control scheme proposed in this paper reduces the settling time of the load's twisting angle by 3.03 s and 1.36 s, respectively.

Therefore, the IRL-based adaptive controller designed in this paper can effectively achieve the expected control performance while maintaining control inputs within predetermined limits. In addition, for the time-varying characteristic and uncertainty of the system, the method is able to adjust the controller parameters online according to the state of the system to improve the control performance.

5. Conclusion

5.1. Summary

During crane operations, external disturbances and parameter uncertainties cause twisting of the spreader. This paper adopts the Lagrange method to establish a nonlinear mathematical model of the system. The model describes the dynamic behaviors of trolley motion, hoisting mechanisms, and spreader twisting. Subsequently, the model is linearized appropriately and the stability of the system is analysed.

Secondly, an enhanced adaptive controller design is presented, integrating an iterative reinforcement learning algorithm for real-time optimization of control strategies. This method employs IRL, transforming the adaptation process into an online decision-making problem aimed at long-term optimal performance. It enables the adaptive strategy to perform forward-looking online adjustments based on real-time system performance feedback.

The proposed control method's performance is verified through digital and physical experiments.

The experimental results show effective suppression of load torsion phenomena. This approach exhibits better performance compared to traditional adaptive and LQR controllers. Therefore, the proposed control method exhibits practical value and broad application potential.

5.2. Limitations and future work

This study has verified the fundamental anti-swing performance of the proposed algorithm on an experimental system, yet several limitations persist. First, the physical experiment was conducted on a dedicated RTG crane anti-swing simulation testbed. The dynamic characteristics of this platform differ from those of a full-scale crane due to scaling effects. Therefore, the algorithm's performance on a prototype system requires further validation. Second, the experimental setup represented an ideal environment. Practical wind disturbances, such as sustained winds or sudden gusts, were not systematically introduced. Consequently, a more comprehensive evaluation of the system's dynamic response and robustness under extreme conditions is necessary in future work.

To advance technology towards practical engineering applications, subsequent research should address two principal aspects. First, large-scale prototype experiments or high-fidelity simulations must be conducted, with emphasis on scale effects such as control delays and structural coupling. This phase is required to validate the algorithm's scalability and extended adaptability. Second, the system's performance must be examined under realistic and complex operational environments. Accordingly, systematic experiments should be designed and executed under varied and extreme working conditions, including wind-tunnel and real-wind-field settings with multiple operational scenarios. These tests will provide a comprehensive evaluation of the algorithm's performance under complex wind loads, encompassing both sustained and sudden wind conditions, thereby enabling a complete assessment of its dynamic response and robustness against actual disturbances. Finally, future work will also

explore integrating the proposed method with H_∞ or L2-gain robust control theory. The performance metrics for suppressing non-stationary disturbances, such as the L2 gain γ, will be incorporated into the IRL cost function learning process. This integration aims to ensure the controller maintains effectiveness under strong disturbances.

Author contributions: All authors contributed equally to the conception, design, data collection, analysis, and writing of this study. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (62541324), the China Academic Theme Case Project (ZT-2410066005), and the Science and Technology Project of Guangzhou Municipal Administration for Market Regulation (2025KJ14).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No datasets were generated or analysed during the current study.

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

References

1. Yang L, Ouyang H. Precision-positioning adaptive controller for swing elimination in three-dimensional overhead cranes with distributed mass beams. ISA Transactions. 2021; 127: 449–460.
2. Xi H, Wu Q, Ouyang H. Nonlinear control of ship-mounted rotary crane based on adaptive dynamic programming. IEEE Access. 2024; 12: 104869–104877.
3. Zhao B, Ouyang H, Iwasaki M. Motion trajectory tracking and sway reduction for double-pendulum overhead cranes using improved adaptive control without velocity feedback. IEEE/ASME Transactions on Mechatronics. 2022; 27(5): 3648–3659.
4. Miao X, Zhao B, Wang L, et al. Trolley regulation and swing reduction of underactuated double-pendulum overhead cranes using fuzzy adaptive nonlinear control. Nonlinear Dynamics. 2022; 109: 837–847.
5. Yan Y, Qin Y, Zhang L, et al. Swing suppression control in quayside crane by using fuzzy logic and improved particle swarm optimization algorithm. Iranian Journal of Science and Technology, Transactions of Mechanical Engineering. 2022; 47: 1131–1144.
6. Zhou Q, Wang K, Xiong X, et al. Optimization of bridge crane control system using fuzzy PID control and speed control of frequency converter. Journal of Physics: Conference Series. 2021; 1802: 032007.
7. Kim J, Kiss B, Kim D, et al. Tracking control of overhead crane using output feedback with adaptive unscented Kalman filter and condition-based selective scaling. IEEE Access. 2021; 9: 108628–108639.
8. Lu B, Cao H, Hao Y, et al. Online anti-swing trajectory planning for a practical rubber tire container gantry crane. IEEE Transactions on Industrial Electronics. 2022; 69(6): 6193–6203.
9. Zhang Y, Li M. Fast point-to-point transportation of the portal crane via double-input inversion-based feedforward and feedback control. Optimal Control Applications and Methods. 2025; 46(4): 1402–1416.
10. Bandong S, Napitupulu CM, Nazaruddin YY. Optimal RTGC non-linear control system based on sliding mode controller. In: Proceedings of the 2024 SICE International Symposium on Control Systems; 25–31 March 2024; Higashi-Hiroshima, Japan.
11. Xu M, Liu L, Wang J, et al. Research on anti-swing control system of slewing crane based on fuzzy PID. PLOS One. 2024; 19(10): e0311701.
12. Ren Z, Huang Z, Zhao T, et al. Dynamic modelling and experimental analysis of an offshore crane payload positioning system with a parallel cable-driven method. Polish Maritime Research. 2024; 31(2): 29–45.

13. Li G, Ma X, Li Y. Adaptive anti-swing control for 7-DOF overhead crane with double spherical pendulum and varying cable length. IEEE Transactions on Automation Science and Engineering. 2024; 21(4): 5240–5251.
14. Zhang M, Jing X, Zhou Z, et al. Transportation for 4-DOF tower cranes: A periodic sliding mode control approach. IEEE Transactions on Intelligent Transportation Systems. 2024; 25(11): 15909–15921.
15. Li G, Ma X, Li Y. Adaptive sliding mode control based on time-delay estimation for underactuated 7-DOF tower crane. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2025; 55: 2277–2288.
16. Zhang C, Wang G, Yang J. Research on a hydraulic cylinder's synchronous control of lifting equipment for large prefabricated components based on IGWO-BP-PID. In: Proceedings of the 2nd International Electronic Conference on Machines and Applications Session Automation and Control Systems; 18–20 June 2024; Basel, Switzerland.
17. Ye J, Huang J. Analytical analysis and oscillation control of payload twisting dynamics in a tower crane carrying a slender payload. Mechanical Systems and Signal Processing. 2021; 158: 107763.
18. Ngo HQ, Hong K, Kim HK, et al. Skew control of a container crane. In: Proceedings of the 2008 International Conference on Control, Automation and Systems; 14–17 October 2008; Seoul, Korea.
19. Tho DH, Terashima K. Robust control designs of payload's skew rotation in a boom crane system. IEEE Transactions on Control Systems Technology. 2019; 27(4): 1608–1621.
20. Zhang M, Xu W, Gu X, et al. Model reference adaptive sliding mode control of overhead crane with uncertainties. In: Proceedings of the 2019 Chinese Control Conference; 27–30 July 2019; Guangzhou, China. pp. 405–410.
21. Cuong HM, Dong HQ, Trieu PV, et al. Adaptive fractional-order terminal sliding mode control of rubber-tired gantry cranes with uncertainties and unknown disturbances. Mechanical Systems and Signal Processing. 2021; 154: 107601.
22. Khadija D, Said AN. A discrete repetitive adaptive sliding mode control for DC-DC buck converter. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering. 2021; 235: 1698–1708.
23. Anh LT, Quang H, Pham TV. Observer-based nonlinear robust control of floating container cranes subject to output hysteresis. Journal of Dynamic Systems, Measurement, and Control. 2019; 141(11): 111002.
24. Xu C, Hu J. Adaptive robust control of a class of motor servo system with dead zone based on neural network and extended state observer. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering. 2022; 236: 1724–1737.
25. Yang S, Meng D, Wang H, et al. A novel learning function for adaptive surrogate-model-based reliability evaluation. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2024; 382(2264): 20220395.
26. Meng D, Zhu SP. Multidisciplinary Design Optimization of Complex Structures Under Uncertainty. CRC Press; 2024.
27. Liu C, Chu Z, Li Y. Online adaptive data-driven control for unknown nonlinear systems with constrained-input. In: Proceedings of the 2022 First International Conference on Cyber-Energy Systems and Intelligent Energy; 14–15 January 2023; Shenyang, China. pp. 1–6.
28. Yan L, Liu Z, Chen PC, et al. Adaptive optimal consensus of nonlinear multi-agent systems with unknown dynamics using off-policy integral reinforcement learning. Neurocomputing. 2025; 621: 129185.
29. Chen C, Modares H, Xie K, et al. Reinforcement learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics. IEEE Transactions on Automatic Control. 2019; 64(11): 4423–4438.
30. Guo L, Xiong W, Zhao H, et al. A nearly optimal adaptive saturation function tuning method for quasi-sliding mode control based on integral reinforcement learning. Neurocomputing. 2025; 623: 129363.
31. Wang X, Ye X. Optimal robust control of nonlinear uncertain system via off-policy integral reinforcement learning. In: Proceedings of the 2020 Chinese Control Conference (CCC); 27–29 July 2020; Shenyang, China. pp. 1928–1933.
32. Lv Y, Chang H, Zhao J. Online adaptive integral reinforcement learning for nonlinear multi-input system. IEEE Transactions on Circuits and Systems II: Express Briefs. 2023; 70(11): 4176–4180.
33. Salamat B, Bencic D, Elsbacher G, et al. Investigating integral reinforcement learning to achieve asymptotic stability in underactuated mechanical systems. IEEE Robotics and Automation Letters. 2024; 9(1): 191–198.
34. Colin M, Thomas O, Grondel S, et al. Very large amplitude vibrations of flexible structures: Experimental identification and validation of a quadratic drag damping model. Journal of Fluids and Structures. 2020; 97: 103056.
35. Salamon R, Kamiński H, Fritzkowski P. Estimation of parameters of various damping models in planar motion of a pendulum. Meccanica. 2020; 55: 1655–1677.
36. Dong N. Adaptive Control. Beijing Institute of Technology Press; 2009. (in Chinese)
37. Jiang H, An T, Zhang Z, et al. Adaptive fuzzy optimal control of modular robot manipulators systems via integral reinforcement learning-based value iteration algorithm. In: Proceedings of the 2024 IEEE 14th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems; 16–19 July 2024; Copenhagen,

Denmark. pp. 388–393.
38. Zhang T, Yan J, Yang X, et al. Digital twin-driven formation control of ROVs: An integral reinforcement learning-based solution. IEEE Transactions on Industrial Informatics. 2024; 20(12): 14277–14286.
39. Mahmoud E, Mammar S, Smaïli M. Model-free optimal static output feedback control using integral reinforcement learning. In: Proceedings of the 2025 33rd Mediterranean Conference on Control and Automation; 10–13 June 2025; Tangier, Morocco; pp. 25–30.
40. Lv Y, Zhang W, Zhao J, et al. Finite-horizon optimal control for nonlinear multi-input systems with online adaptive integral reinforcement learning. IEEE Transactions on Automation Science and Engineering. 2025; 22: 802–812.
41. Vrabie D, Vamvoudakis KG, Lewis FL. Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles. The Institution of Engineering and Technology; 2012.